TOWARDS A QUANTITATIVE EVALUATION
OF LAYOUT USING GRAPHIC DESIGN PRINCIPLES

A thesis submitted to the
Kent State University Honors College
in partial fulfillment of the requirements
for Departmental Honors

by

Daniel J. Mosora

May, 2012
Table of Contents

LIST OF FIGURES ........................................................................................................................ iv
LIST OF TABLES .......................................................................................................................... v
ACKNOWLEDGMENTS ................................................................................................................ vi

CHAPTER

I. INTRODUCTION .............................................................................................................. 1
   a. Contributions ......................................................................................... 2
   b. Terminology .......................................................................................... 3
   c. Organization .......................................................................................... 4

II. RELATED WORK .......................................................................................................... 5
   i. Computer Science .................................................................................. 5
   ii. Graphic Design ..................................................................................... 9
   iii. Psychology .......................................................................................... 14
   iv. Summary .............................................................................................. 16

III. APPROACH TO QUANTIFYING LAYOUT ............................................................. 17
   i. Understanding Layout .......................................................................... 17
   ii. Characteristics of Layout ...................................................................... 20
      a. Element Attributes ........................................................................... 21
      b. Positioning Attributes ....................................................................... 23
      c. Inferred Modifications ..................................................................... 26
   iii. Algorithmic Considerations .............................................................. 32

IV. CASE STUDY .............................................................................................................. 35
   i. Identifying Elements ............................................................................. 35
ii. Element Attributes ................................................................. 37
iii. Positioning Attributes .......................................................... 38
iv. Inferred Modifications .......................................................... 40

V. POSSIBLE APPLICATIONS ....................................................... 45
VI. CONCLUSIONS & FUTURE WORK .......................................... 47

REFERENCES .............................................................................. 50
LIST OF FIGURES

Figure 1. Square Composition Focal Points................................................................. 19
Figure 2. Two Disks Arranged in a Square.................................................................. 19
Figure 3. Chromostereopsis ....................................................................................... 23
Figure 4. Visual Angle................................................................................................. 34
Figure 5. Case Study Poster Example........................................................................ 36
LIST OF TABLES

Table 1. Characteristics for Automating Layout .................................................. 20
ACKNOWLEDGMENTS

First of all, I would like to express my thanks to my advisor, Dr. Jonathan I. Maletic, for not only allowing me to pursue my crazy ideas with this thesis, but encouraging me all along the way. I would also like to thank the members of my defense committee, Prof. AnnMarie LeBlanc, Dr. Ye Zhao, and Dr. Sara Newman, for their willingness to read over my work, their comments, and patience as I finished formatting everything. Also, I am very grateful for my professors in the School of Visual Communication Design, Shawn Simmons and Chris Ransom, as well as Aaron Rosenberg and Dr. Jill Folk, for their willingness to listen and discuss this idea throughout the process and their recommendations of literature to review. I also would like to thank my sponsor, Colleen Van Lent, for taking interest in my research topic and, through her generous donation, enabling me to pursue this achievement without worrying about printing costs, something I’m not quite used to from my experience in the School of Visual Communication Design.

In my personal life, I cannot be more grateful for those that are close to me. To my mom, dad, brother, sister, my best friends, and my girlfriend, thank you a thousand times. I would have probably lost my mind multiple times over if not for your continuous support. For being there to help me through the final push, I’d like to thank the members of To Write Love On Her Arms UChapter at Kent State for providing me with a huge amount of support without me even asking.

Finally, to anyone who I have spoken with over the past two years, thank you. You have no idea, but simply through our interaction, however so brief, you have contributed to helping me get through some of the hardest days, weeks, and months of my undergraduate career.
CHAPTER I
INTRODUCTION

The process of human perception is a complex system refined over millions of years of evolution and perfected over the course of our early development. We use our senses daily for survival as well as to learn about the world in which we live. Among these, our sense of sight is a main component of our daily interactions with the world around us. From advertisements, to road signs, to perceiving the faces of people we meet, there are subliminal heuristics that deal with the massive flood of information. With technology becoming an integral part of daily routine through the wide usage of smart phones and tablet PCs, the amount of information to which we are introduced each day is at an all time high.

However, the creation of visual information (e.g., advertisements, web sites, and mobile applications) is still done via a purely manual process. The technology has been developed to deliver content quickly to devices of varying screen sizes, but its presentation is left up to contrived methods of presentation that must be manually drafted. Outside of the profession of graphic design, there is a shortage of practical knowledge of how human visual perception operates. There must be a way to leverage the refined processes of our own perception in a way that technology can take advantage of and, subsequently, increase general comprehension and usability of visual presentations by some degree.

This thesis proposes a framework for investigating this problem and applying it to objective layout analysis. Generally in the context of this thesis, objective layout analysis refers to a process of evaluating a visual composition, digital or print-based, through quantitative methods that may be algorithmically defined and automated. Although there
have been pursuits towards layout analysis in many contexts such as interface design, this work is concerned exclusively with traditional poster design to establish a basic framework. However, it is not inconceivable for this research to be applied more generally in the future.

A few problems arise in considering the possibility of creating a framework for objective layout analysis. Firstly, graphic designers approach analysis of layout from both a syntactic and semantic standpoint, using not only the visual appearance of the information being presented, but also the symbolic meaning of the content. This is unsuitable for a basic system of automation for the reason that semantic information is so culturally variable that a human familiar with the information would be most capable to evaluate it. For this reason, the framework described herein focuses only on the syntactic properties of design.

With this in mind, another issue arises in both translating human perception into an objective method of analysis and algorithmically combining the facets of the visual system into a usable metric. Much experimentation will need to be completed to discern the numeric effects of different aspects of graphic design on the overall measurement. For that reason, this thesis focuses upon identifying techniques manipulated by graphic designers and how they would affect a single metric to be used in layout analysis. Also, the objective nature of the study ensures that the process is free from stylistic and semantic influences.

**Contributions**

The result and main contribution of this thesis is a problem analysis and literature review, and a basis with which to conduct further experimentation in the pursuit of an objective layout evaluation metric and algorithmic method. The proposed metric will be referred to as the salience of an element in a composition, and is a measurement of its
visual importance in the layout. The salience of all elements in a composition construct a visual hierarchy that influences a viewer to scan the page in a way determined by the designer. The development of this conceptual metric consists of identifying relevant characteristics of layouts based on general techniques applied in design, grouping them into similar categories, and exploring each one’s effect on the overall perception of a composition.

**Terminology**

This area of research involves many different disciplines, each with their own terminology. For this reason, I will first define some terms that will be commonly used throughout the discussion in order to limit possible confusion.

The terms *element* and *object* are used to refer to a visual “node” in the composition. These are points of interest and could be holistic (e.g., a headline) or partial (e.g., a specific portion of an image or texture). Within the context of these terms, the term *salience* is used to refer to the amount of attention that an element draws to itself. *Weight* and *gravity* may also be used synonymously with the salience of an element.

In the context of a grouping of elements, the term *composition* or *layout* refer to any form of visual presentation, regardless of media (screen, print, interface, etc.). The region that bounds a composition is referred to as the *page* or *compositional area*. *Visual hierarchy* describes the visual “flow” that a viewer’s eyes follow around the page, guided by the composition present. This describes the ranking of elements in the order they appear in this flow and is directly influenced by their salience.

In reference to process, the terms *design* or *layout*, when used as a verb, are used to indicate the act of manipulating visual hierarchy in a composition by sizing and positioning elements. Within this, the term *characteristic* is a generic term used to refer to a property of elements and layouts that is modified by graphic designers in the design
process to achieve hierarchy. Finally, the term *technique* refers to the way that a graphic designer manipulates and applies characteristics to elements in a layout.

Throughout the discussion of this research, I consistently use the term *objective* or *objectivity*. Within the context of my research, the use of this term is within a specific scope. It refers to the process being used to study these topics as a product of the scientific method. This distinction is necessary as the method of study involves human subjects, and objectivity as it relates to humans can never be completely objective. The range of unique human experience does not allow this, and I only attempt to be objective in the context of following a methodical process of study and consistent method of measurement.

**Organization**

Initially, in Chapter 2, I begin with an overview of related work drawn from many diverse fields that contribute to the topic at hand. In the first section of Chapter 3, I describe the background that initially inspired and remained a driving force for my work. In the next section, I present a list of characteristics grouped into three categories that must be investigated further and, finally, an initial high-level algorithmic structure in the last section. An analysis of some of these characteristics is applied to an actual design layout to illustrate how they interoperate in Chapter 4. Possible applications of this research in varying fields are proposed in Chapter 5, with a look towards the future and conclusion in Chapter 6.
CHAPTER II
RELATED WORK

The following work relates directly to the research presented within this thesis or served as an initial inspiration [1-28] to the investigation. These references are categorized into their main subject area, though many are multi-disciplinary in nature.

Section 1: Psychology

From millions of years of development of perception come an equal number of questions that ask why we see things the way that we do. A primary inspiration to the method of which further research on the topic presented in this thesis will be conducted comes from the concepts and methods of the psychological research field called psychophysics. Research in this area focused on not only what physical systems cause us to see, but also what psychological processes tell us what we see. The first bits of research were completed by E. H. Weber in the 19th century and were concerned with measuring a difference threshold of perception. As described early on in Psychophysics: The Fundamentals by Gescheider [10] (p. 1-10), this is defined as the minimal amount of change in stimulus that could be detected by an individual. The psychologist G. T. Fechner expanded upon this research and further investigated the relationship between physical stimulus and perception.

Fechner’s work was originally published in the book Elements of Psychophysics in 1860, [10, 29] which heralded a new age of psychological research as a scientific, as opposed to a philosophical, discipline. He derived an initial general formula for Weber’s work that describes the sensation magnitude with respect to the stimulus intensity on a
logarithmic scale. The notion of a non-linear relationship with regard to perception, as is well known to apply to perception of sound intensity, is very important because many statistical models of human behavior are also non-linear. It is not out of the question to wonder if perception of form and layout also follow a similar non-linear relationship.

The work of Fechner spurred an increase of study in the field with regard to the physical perception of stimuli among one dimension, such as brightness of a visual stimulus. However, within the later developments in this field, the most relevant methods to the problem of perception of visual hierarchy in layout are those applied to studying multidimensional stimuli. Rao and Wilcox discuss [20] the application of multidimensional scaling to psychophysical research. They present the problem as a need to determine the specific transformation functions used by subjects to give responses to a set of multidimensional stimuli. A similar transformation function is exactly what I am in search of with the research presented here.

Another early examination of perception appears in the field of Gestalt psychology, developed by Kurt Koffka in the early twentieth century [12]. Many theories proposed by Gestalt psychology have since been adopted by graphic designers and applied by these people in practice. Those that have application to this work are related to the laws of perceptual organization that are characteristic of Gestalt psychology as well as the idea of figure-ground segregation. These include the laws of prägnanz (simplicity), good continuation (or continuity), and proximity (or grouping).

The laws of good continuation and proximity as well as figure-ground segregation will be discussed in detail later, as they appear as a portion of my proposed metrics. However, the Gestalt law of prägnanz is an important part of perception. This law proposes that every stimulus pattern is seen in such a way that the simplest resulting structure is perceived. This is one of the most important of the Gestalt laws, as it describes the idea of a top-down processing method where features of a visual stimulus are
combined to form a cohesive figure, then analyzed individually if desired by the viewer. Photo-mosaics are a good example of this, where the viewer sees the intended image first, but then may focus in on each individual image to perceive it separately. This idea, while not applicable as a metric directly, is a very important part of analysis because it describes perception in a top-down manner. Gestalt psychology is usually summed up in this way with the phrase: “The whole is greater than the sum of its parts.”

An interesting analysis of visual perception from the mid-twentieth century that is motivated by both psychology and art theory is present in the book *Art and Visual Perception* by Rudolf Arnheim [1]. He discusses a number of visual properties of artwork, but many are related to the semantic qualities of the visual presentation in different art, from children’s art to the great masters. Most interesting in the context of the problem under discussion are the concepts of balance and tension.

These two concepts describe visual “forces” that push and pull on the perceptual center of a composition as its elements are moved about the canvas. This idea of forces is a main component of my analysis as the visual weight and gravity of an element in a composition is the root of my proposed metric. Balance describes the visual forces acting upon an element that is homogeneous, or non-directional, such as a circle. Tension on the other hand, refers to elements that are heterogeneous in nature and thereby contain forces within themselves that affect their overall interaction within the composition.

From visual forces between objects, the concept of visual forces within objects is discussed by Treisman and Gelade in the Feature Integration Theory of Visual Attention (FIT) [26]. In this work, a theory is presented to describe the perception of an element that is distinguishable from its surroundings as a combination of different features within the element. They propose that focused foveal attention is required in order to integrate the features of an object into a whole, whereas unattended objects are just seen at the feature level. Foveal is a word used to describe the center of the field of vision where it is
possible to perceive the most finite detail. This process combines bottom-up process and top-down processing, differing slightly from the ideas proposed by Gestalt psychology. They test this theory and a few hypotheses through the experiments relying upon visual search, texture segregation, identification and localization, and using multiple dimensions of features to be combined into complex whole elements.

Interestingly, they found that unattended objects in activities such as visual search acted as a kind of interference to the process of integrating features in the foveal area. This insight adds to my current and future research, as the weight of an element may need to be dynamically adjusted as the viewer’s eyes move along the composition of the hierarchy. The effect of foveal versus peripheral attention on the actual importance of an object in hierarchy is a very intriguing subject that must be investigated further and accounted for in the initial research phases.

An interesting experiment that tests some of the effects of interference on visual search comes from Jan Theeuwes [30] in 1992. The focus of this study was on the pre-attentive parallel stage of perception, which is the stage that I consider when discussing layout perception. Theeuwes conducted an experiment in which subjects were to locate a salient element of a specific color in a field of elements of a single color. During this experiment, there was also an irrelevant salient element that served as interference. Salience is determined by a particular characteristic of a visual element that causes it to stand out more than others, in this case, color. The result of this particular experiment was that subjects could not train themselves to ignore the irrelevant stimulus no matter how much practice they had with the task. This is interesting in that it is a concrete example of the salience of an element in a composition overriding top-down processing when viewing a visual stimulus.

When discussing visual hierarchy and the importance of elements in a composition, the concept of visual salience is of direct relevance. Parkhurst et al published
a study [17] in 2002 where a saliency map was created of multiple images, and its accuracy was tested against the results of an eye-tracking study. This research is directly related to my proposed metric, as the development of a saliency map is basically what I am concerned with when considering an element’s importance. Their methods of verification are useful as tools with which to conduct my own verification in future work. However, the saliency maps constructed for this study only take three attributes into consideration: color, intensity and orientation. With an expanded repertoire of characteristics to take into account when evaluating salience, it is possible that the flow of fixations in an eye-scanning pattern may be better predicted.

Section 2: Computer Science

**Human-Computer Interaction**

In the field of Human-Computer Interaction (HCI), layout evaluation and automatic layout are discussed within the problem domain of user-interface (UI) design. Lok and Feiner [14] provide a relatively comprehensive examination of the methods for automatic layout creation and layout evaluation for user-interfaces that existed in 2001. Though some methods presented within this survey dealt with the problem of developing metrics to evaluate automatically generated UI layouts, they relied primarily on the usability of the interface as an indicative measurement. My approach relies on more fundamental characteristics developed in the field of graphic design for evaluation, and as a result I also consider more complex layout properties besides element size and position.

Krzysztof Gajos implemented another automatic user-interface layout technique that considers a few other necessary portions of the problem as it is used in practice. The SUPPLE automatic layout tool [9] provides personalized user interfaces that are tuned to a single user’s preferences in a short amount of time based on things such as their preference, task, and abilities. The most interesting part of this work is how it reduces
the problem of actually generating a layout into an optimization problem that prunes a solution space of up to $10^{17}$ possible solutions in near real-time. This lightweight approach to the problem of automatic layout is an important consideration to make in both future research and in building the basis for future research on this topic.

More recently, Yeonsoo Yang and Scott Klemmer [28] created another method for automatically generating layouts. Their approach utilizes design heuristics and Gestalt principles for generation of handheld device UI layout. The design principles of simplicity, structuring, and proportion used are important design considerations, however, these are more high-level design constructs. In my methods, the amount of elements (simplicity), contextual grouping (structure) and proportions of the space are left to either the person creating the layout or arise from application of proper placement characteristics, as with structuring using Gestalt grouping, for example.

Duchowski’s *Eye Tracking Theory and Methodology* [7] is a very thorough account of the history and research methods of eye tracking, which will be a primary tool in evaluation of the concepts presented in this thesis. Not only that, it begins with a look at prior research in psychophysics and other related areas of psychology and neuroscience that affect visual attention. This work played a critical role by introducing me to related research in the area of visual attention, a field directly related to my research. Many of the psychological fields mentioned in the first chapter are represented in the Psychology section of related work presented here. It also provided some insights when initially investigating how to measure the size of an element on the page and gave inspiration for a possible application of this work.

**Computer Vision:**

The topic of the human perceptual system is a frequent area of research in the field of Computer Vision. Finding algorithmic methods of distilling the perceptual system of human vision is directly related to my work. One such algorithm arises out
of research completed by Eric Saund. In this paper [23], Saund presents an algorithmic method of leveraging the Gestalt law of good continuation in order to detect implied closure of contours in an image. As stated previously, Gestalt psychology is an important part of graphic design, and the law of continuity is a main component of an important characteristic of visual design, which will be elaborated upon later. This algorithm relies upon bidirectional graph processing to detect closed paths via a few metrics related to the closed structure of a graph constructed from a contour drawing of the image. However, as their cited processing time of a complex image is 8 seconds, application of this sort of algorithm may present a problem with overall speed of my methods in the future. Nevertheless, it is an excellent example of an initial method of visually processing path closure objectively.

Saund again utilizes design principles in the tool called ScanScribe [24], which leverages Gestalt principles to increase comprehensibility of diagrams as a diagram editing application. This application of Gestalt principles of continuity and grouping to recognize perceptually related elements and figure-ground segregation is a very low level perceptual process. It is directly related to the goals that this work aspires to achieve, but on a much more complete scale. However, it serves as a proof of concept that in a limited scope of application and use of a few principles of design, automatic analysis of layout is possible.

Gestalt principles appear again in a study by Sheel Dhande and Andrew Fiore [6], this time with the Gestalt law of proximity. They completed a thorough study of identifying Gestalt groupings of web page components by automatic methods. This kind of work is extremely useful as it provides a means for identifying groupings based on how they will be perceived. They used three different approaches to accomplish this: K-means clustering, hierarchical clustering, and mixture modeling. As a first step towards this kind of analysis, their study evaluated the effectiveness of each approach with human subjects. They found that the most accurate method used was hierarchical clustering with similarity
based on distance of objects from one another. In future work, these methods will be examined as a portion of salience evaluation.

Generally speaking, visual hierarchy consists of regions of interest and the sequence in which they are viewed. Privitera and Stark conducted a comparison study [19] of six different context-free layout algorithms. They judged the relationship between these by pairing them and analyzing the performance of each based on three similarity metrics: region of interest (ROI) loci, sequences, and transitions. The most important portion of this study as it relates to the goals of this thesis is the similarity of sequences metric, as this is the defining characteristic of visual hierarchy in layout if the actual elements, or ROIs, are known \textit{a priori}. Interestingly, the similarity of sequences between any two given algorithms was surprisingly low, with a maximum ratio of 0.38 and a majority below 0.2. This implies that the sequences generated by these algorithms is not agreed upon between them, and must be examined further.

Reisfeld describes [22] active foveated visual systems as requiring both context-dependent and context-free attentional mechanisms. In this context, Foveated is used to describe systems of computer vision in which there exists a fovea, or central area for analysis much like the fovea in a human eye. In the context of my work, the context-dependent mechanisms are related to the task of the viewer and the meaning of the visual elements being portrayed, which are analogous to the semantic properties not under consideration in this thesis. The context-free mechanisms are particularly of interest to this work because they are applied purely as an objective view of the visual presentation. However, the methods of image analysis presented in this work are driven by a generalization drawn from local symmetry of the forms present. This is not as applicable for graphic design layouts, because although it may be able to locate areas of interest, it has no way of ranking them in order of visual hierarchy.
Data Visualization:

The field of Data Visualization has much to gain from automatic visual analysis. Currently, there are not many ways to objectively decide whether or not a particular visual presentation is more effective over another. It is because of this open-ended question that it is surprising that not much research on this problem exists within this field. One example that is related to layout analysis especially is in research completed by Buja et al. In the paper, “Interactive Data Visualization Using Focusing and Linking” [3], he discusses two techniques for presenting data in a way to help maximize comprehensibility. These techniques, focusing and linking, are relevant to my research as they describe some of the same principles in the context of data visualization and inspire further applications of this research in the form of evaluating the effect of focusing and linking using the saliency metric proposed here.

Focusing is a simplification of views that allows a person to center in on important information with little distraction. This concept is related to Treisman’s work in Feature Integration Theory [26] by removing the interference from unattended stimuli by simply not including it. In graphic design, this technique is related to the characteristics that contribute to visual hierarchy to reduce this same sort of interference in a single view.

Linking is a complex technique that may be applied in a few different ways. Over time, linking involves smooth movement of data points within a view. This sort of movement is very helpful in dynamic data visualization, but is not relevant in the context of the current research under consideration. In parallel, objects may be linked in a single view through either Gestalt continuity or similarity of color and style. These two techniques are used frequently in graphic design through alignments and management of color to enhance visual grouping of semantically related elements that may not be near each other in the layout.
Section 3: Graphic Design

Empirical studies, such as those in the aforementioned works in Computer Science and Psychology, are lacking in the field of graphic design. However, there is a rich past in design that interacts with intuitive explanations that come from art theory as well as the principles described in Gestalt psychology and psychophysics. As such, the following works describe the marriage of these subjective principles that rely upon the viewer’s individual experience for explanation and the quantitative, logical methods of the psychological fields that deal with perception research.

One of the most basic and comprehensive introductions to the principles of graphic design is *Graphic Design: The New Basics* by Ellen Lupton [16]. In this book, Lupton describes the process of design within a framework that is similar to that which is present in a design curriculum’s introductory courses. She begins with the classic example of point, line and plane, and how they interact within a composition, but quickly moves on to more complex topics such as texture, color, figure-ground segregation and hierarchy. These chapters are very important in defining many of the characteristics that make up my framework for analysis of layout.

Lupton again presents many key concepts of graphic design in her book, *Thinking With Type* [15]. Typographic design and typeface design are two areas that form the heart of application of graphic design principles. Hierarchy is present both within a typeface and externally in the way the typeface is used. This book focuses primarily on the problem of how typography is used. In this context, she describes hierarchy [15] (p. 94) in a way that I will later define as literary hierarchy. This description is that hierarchy is a way of organizing information in a text block using things such as font (bold or italic), visual cues (such as bullets and numbers), and spacing (such as tabs).

Discussion of the importance of design principles as they apply within a typeface
is left up to typographers. A thorough representation of this is present in the book [4] *Typographic Design: Form and Communication* by Rob Carter, Ben Day, and Philip Meggs. In this work, there is much discussion of using typographic syntax to create a cohesive form of communication. As far as syntax goes, they discuss nuances within typefaces [4] (p. 39), such as x-height and thick/thin contrast, that improve or reduce a typeface’s legibility at certain point sizes. For example, a typeface with a large x-height, the height of a lowercase letter, is more readable at lower point sizes in blocks of text as it emphasizes the counterform of each letter and allows the reader to more easily distinguish each letterform.

Within the context of creating communication, the authors discuss organizational methods that are used to achieve proper informational hierarchy through visual hierarchy. This is directly related to my work as it both relies upon graphic design principles as well as creating a proper hierarchy for communication. Examples are given that take advantage of things such as an ordered grid structure and ABA repetition to provide both organization and segment off areas of a composition to allow for most effective Gestalt proximity relationships. One of these examples will be expanded upon later in Chapter 4 as a case study.

An intriguing example of automatic layout generation comes from the work of the Institute of Artificial Art Amsterdam (IAAA) [2]. The IAAA is an independent organization that pursues algorithmic methods of creating art in many forms including dance, architecture and graphic design. Their research in artificial design spans graphic design, information visualization and architecture. With respect to graphic design layouts, their work focuses on generalizing a corporate identity for use in generating new documents that fit into the existing visual language. This is accomplished by defining a style grammar in the form of multiple XML schemas and using that as input to a two-cycle generator to create the layout with the desired elements. While this is a very
interesting look at the creation of layouts and a great step forward for the problem of automatic layout, a human designer must still specify the design schemas for this system.

**Section 4: Summary**

Throughout Psychology, Computer Science and Graphic Design, there are crosscutting interests that provoke a multidisciplinary approach to the problem of visual perception. As a result of its long history in psychological research, work in other fields related to perception draws from a solid and consistent theoretical base. With the focus on pre-attentive processing and Gestalt psychology, my work continues much in the same way. In the next chapter, I will begin describing the concepts that this work draws from with the background of what has been already discussed. Furthermore, I will formally define and clarify the key concepts relating to the problem of applying graphic design techniques to quantitative layout evaluation.
CHAPTER III
APPROACH TO QUANTIFYING LAYOUT

Throughout past psychological research, there exists a pre-attentive processing stage to perception that occurs before the viewer is consciously aware of what he or she is looking at. We see this with the feature integration of Treisman's work [26], the pre-attentive parallel stage examined by Jan Theeuwes [30], and in the law of prägnanz described by Gestalt psychology [12]. It is because of this, I propose the following work as a take on the stages of pre-attentive processing from the standpoint of a practitioner of graphic design.

This chapter begins with a background of the ideas and theories that stimulate the conceptual development of a metric, called salience, with which to evaluate layout and a formal description of hierarchy as it relates to graphic design. Then the characteristics of salience are described in detail.

**Section 1: Understanding Layout**

Gestalt perceptual psychology tells us that the whole is greater than the sum of its parts [12]. Simply stated, this refers to the notion that a perceptual scene being evaluated by the viewer is seen as a whole unit, instead of just a collection of unrelated elements. The implication of this is that there are unseen aspects of visual analysis beyond that which is physically encoded in the data. This additional information is drawn from interactions between multiple elements or between an element and the page itself.

These unseen effects of visual layout have been the study of graphic design for centuries, whether explicitly or implicitly under the guise of artistic intuition. Designers
have perfected methods of presenting information in a visual manner over hundreds of years of trial and error in the form of print-based advertising. They have been forced to contend with competition for attention in the ever-growing world, which led to the necessity of developing techniques which appeal to a viewer’s pre-attentive processing and grabs their attention effectively. This is accomplished in part by creating visual tension in a layout that guides a viewer to discover the information in a visual flow pattern dictated by the designer. This flow pattern is derived from the semantic hierarchy of the information presented and is analogous to the concept of a visual hierarchy.

The visual hierarchy, or flow, of a composition is arguably the most important characteristic of a design. It is the culmination of every technique applied by a graphic designer in order to present information in an intuitive manner that is comprehensible, legible, and pleasing to the eye. In order to develop such a flow, a designer must balance the application of visual techniques and styles to adjust the amount of visual weight that a certain element possesses within the composition.

Each element, when present in a layout, is implicitly assigned a hierarchical rank by the viewer’s visual processes. For the purposes of simplicity and generality, the aforementioned metric will describe this hierarchical rank, and will be referred to as the salience of an element, from the work of Parkhurst [17]. The actual process of evaluation is unclear, however, the techniques employed by a designer approximate this enough to be able to dictate a visual hierarchy.

To clarify the concept of an element’s hierarchical rank, it can be compared to that of gravity, as described in Einstein’s theory of relativity. With gravity, the theory states that matter displaces space-time in a way much like a bowling ball on a trampoline. In this situation, the bowling ball pushes the trampoline down in the center, creating a funnel-like shape around it. If one were to put a tennis ball or baseball on the trampoline, it would roll towards the bowling ball. Visually, the weight of an element can be thought
of as the element displacing the white space around it, upsetting the blank flow of the composition area. Now, depending upon the physical properties of the element, white space around the element, and its proximity to other elements, its overall weight and pull for the user’s gaze is affected.

Arnheim discusses [1] this concept of visual forces and weight displacement and interaction in his discussion of balance in art pieces. He presents the hypothetical situation of moving a disk around a square compositional space. As the disk is moved nearer the edge, he notes certain hidden structural locations within the compositional space of the square where the disk appears to be in balance with the space around it. This concept is illustrated by Arnheim using a rough diagram [1] (p. 3) as seen in Figure 1. He further extends this idea nearer to a typical composition by introducing a second disk in the arrangement [1] (p. 7), shown in Figure 2. He notes that in addition to the individual forces that draw an element to the compositional border, there are forces that interact

![Figure 1: Square Composition Focal Points](image1)

![Figure 2: Two Disks Arranged in a Square](image2)
between the two objects as well.

From the single characteristic of element positioning through the aforementioned works, an element’s importance appears to be directly related to the amount of pull it has on the gaze of the observer. However, the method of layout analysis proposed here is a multi-dimensional approach consisting of much more than an element’s placement. What must be determined through further empirical studies is how the unseen forces described in Gestalt Psychology and by researchers like Arnheim are affected by the characteristics under review. My analysis begins with an inventory of the various techniques, attributes and characteristics of design used by designers to achieve proper hierarchy of information.

**Section 2: Characteristics of Layout**

As stated previously, there are many different techniques and characteristics identified by graphic designers and employed in the creation of a layout. These are outlined in Table 1. Though Gestalt psychology describes a purely top-down processing model of perception, these properties are presented in the form of a partial bottom-up model, similar to Treisman’s work [26]. What this means in the context of my analysis

<table>
<thead>
<tr>
<th>Element Attributes</th>
<th>Positioning Attributes</th>
<th>Inferred Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>Location</td>
<td>Gestalt Grouping</td>
</tr>
<tr>
<td>Weight</td>
<td>Page-edge Relationship</td>
<td>Gestalt Continuity</td>
</tr>
<tr>
<td>Color</td>
<td>Rule of Thirds</td>
<td>Texture</td>
</tr>
<tr>
<td></td>
<td>Cropping</td>
<td>Anomaly</td>
</tr>
<tr>
<td></td>
<td>Intervals</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alignment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leading Lines</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contrast</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Figure-Ground Segregation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Literary Hierarchy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Visual Hierarchy</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Characteristics for Automating Layout
is the evaluation begins with the physical features of each element, then moves to the location-based properties, and ends with inferred modifications of salience based on interactions between elements. It is not completely bottom-up, however, because each level of characteristics requires a grouping and balancing with every element below it. As we have seen that the addition of another element causes interactions between the elements that have already been in place. Therefore, each level of processing will show a more refined coherent view, until the ideal whole image of the Gestalt model is constructed.

**Sub-section 1: Element Attributes**

The effect of these attributes will be determined by psychophysical interpretation of them. They’re reliant upon the biology of the visual system itself as well as more low-processing perception. However, there is certainly some psychological connotation to the overall weight that each of these attributes adds to an element. Since these connotations deal with semantic meaning of an element, the element attributes are limited to only the physical characteristics of the element itself and the human visual system's interpretation of them. We now detail each of the characteristics as presented in Table 1.

**Size**

The size describes the actual physical size of the object on the page. Though it is one of the most fundamental of the characteristics being described, care must be taken when using it in the context of perception. Size is very important to this phase of perception because it affects the amount of salience each of the other characteristics generated. A large object with a certain weight or color will have a proportionally greater effect than a smaller object of the same weight or color.

At its most basic level, the perception of size is based on relativity. For example, we are able to judge the height of an object in space based on the familiar objects it is
near and the distance at which we perceive it. Therefore, a calculation of an element’s salience based on size will be directly affected by its size relative to other elements in the composition. However, this sort of calculation will be handled in a later stage (See Contrast in Inferred Modifications). It is for this reason that the calculation of this definition of size will be limited to the element’s absolute size on the page.

Weight

An element’s weight describes its overall gray-scale “heaviness” and is drawn from a typographic background. In the context of typography, this is related to the weight of the typeface, as assigned by the typographer who created it. For example, there are many typefaces, such as Univers that come in many different weights, ranging from very light to heavy or extra bold. Additionally, this could be used to describe images and other elements by their relative heaviness.

When evaluating an object’s weight, it is important to consider what scale the overall weight analysis maps to the object’s salience. This is directly related to the work of Parkhurst in constructing a saliency map of an image, where the weight is analogous to his intensity measurement. Objective analysis of this characteristic will require the use of eye-tracking research much like the methods utilized in the previous saliency research.

Color

Color describes the color of an element based on the visual spectrum. Organizations such as Pantone have focused research efforts into the psychological effect of color choice. Leatrice Eiseman thoroughly describes this research in her book, Color: Messages & Meanings [8]. However, there has not been as much research that examines the overall visual pull that a color contributes to an element. Furthermore, color becomes more complex when relating to other elements (See Contrast in Inferred Modifications).

Much like weight, color contributes to the intensity of an element and must
be evaluated using experimental means in a similar way. However, a roadblock to the measurement of color’s contribution to salience occurs when one considers the effect of color blindness. Those with color blindness perceive a different gamut of colors than normal, and so would perceive different salience values for these. Colin Ware [27] (p. 54) mentions this with an example of red text and blue text on a black background, as shown in Figure 3. Most people perceive the red text in front of the blue text, however those with a condition called chromostereopsis perceive the opposite, with the blue text in the foreground.

*Sub-section 2: Positioning Attributes*

These attributes are also determined by the physical qualities of a specific element, however they are not related to the object itself, but rather its placement on the page. Thus, it is a higher-order processing task that a viewer uses to analyze the effect on an element’s importance as the subconscious evaluation is between an element and the edges of the presentation area, page or screen. The characteristics within this categorization are analogous to Arnheim’s studies of compositional balance [1], especially the rule of thirds.

![Figure 3: Chromostereopsis](image)
and page-edge relationship, shown in Figure 1. Many of these attributes are inter-related and build off of one another, as indicated in their descriptions.

**Location**

Location is the basic attribute related to positioning and describes the element's absolute coordinates within the layout. This attributes serves as a basis for evaluation of the following positioning attributes. As a result, evaluation of the salience as it relates to location is purely contingent on the psychological evaluations discussed below.

**Page-edge Relationship**

Page-edge relationship is a function of an element’s positioning as it relates to the edge of the compositional area. It is a direct measurement of the distance between an element and the page edge. As a physical property of the positioning of the element, it is affected purely by the element's location within the frame.

The overall effect of this characteristic on the salience of an element is unclear. However, the subjective view of it is that an element creates more tension the smaller it is and the closer it is to the edge of the page. Arnheim [1] documents this effect in his analysis of a single element in the square compositional space. There will need to be a great deal of testing to determine the amount of salience gained by this characteristic as it relates to the element’s base salience derived from evaluation of the element attributes.

**Rule of Thirds**

The rule of thirds is a perceptual heuristic used widely in photography that describes active focal points in a layout with the intersection of lines that divide the space into thirds. The result is four points where focus is used to its greatest degree. This concept is illustrated in Arnheim's diagram of the active locations in a square compositional space [1] [Figure 1].
The effect of the rule of thirds is with respect to the natural focal points in the compositional space. Elements that are placed nearer to them will need to be modified with a higher saliency value. Two problems that must be examined occur when considering the effect of this characteristic. The first is that it is unclear how much saliency is affected based on the relative position of an element to these focal points. The second problem was identified in Arnheim's analysis when adding another element to the composition (p. 7). The interactions between elements may have such an effect that they overpower the salience gained through the absolute positioning of a single element.

**Cropping**

Cropping is similar to the page-edge relationship attribute, in that it is a special case where the relationship is zero. If an element is cropped off the edge of the compositional area, it must be handled differently than one with even the slightest white space between it and the area's edge. This is because of the buildup in tension as an element approaches the edge with a subsequent lessening of tension as it is cropped. Again, Arnheim subjectively describes this phenomenon when examining the effect of positioning a circular element within a square layout space [1] (p. 3).

Cropping poses a lot of unique questions with respect to its overall effect of an element. Semantically, the issue of whether or not a crop is “awkward” is a complicated problem to solve automatically, as it relates to the content of the element instead of its visual properties. Visually, the problem posed is the overall effect that the cropping has on the element's normal weight if it had not been cropped. It may even have a relationship to the counterforms generated by the newly cropped element (See Figure-Ground Segregation in Inferred Modifications). Because of these complicated problems, cropping will require a great deal of precise research to identify its effect on the layout as a whole.
Sub-section 3: Inferred Modifications

Inferred modifications are those attributes that arise from an element’s interaction with other elements around it. These are called inferred modifications since they are rooted in the stages of pre-attentive processing discussed earlier that are applied by our subconscious when viewing a composition. It is these attributes that constitute much of the unseen effect of specific properties of elements on the overall layout evaluation as implied by Gestalt psychology. These are also related to the tension and forces between two elements described previously.

Also, they are applied as the last portion of analysis as they depend on both the prior analysis of the individual element attributes as well as the element’s positioning attributes. As a result, these are the most complex graphic design principles that will require the largest amount of investigation to be able to be used to more fully describe the amount of salience an element possesses.

Gestalt Grouping

The law of proximity describes that placing elements near each other causes them to be perceived as being related in some way. This has application in layout design by visually segregating informational elements that have a semantic relationship among them. For example, a set of three points will be segregated from the rest of a composition to show that they are different from the body text present and are semantically related to each other.

There has been much work in perceptually identifying Gestalt groupings by automatic means, as seen through the work of Dhande and Fiore. Evaluation of how this information relates to the salience of elements is a tricky task. The concept of a group inherently adds visual importance to all elements within the group based on the salience given to the group itself as if it were an element in the design. For this reason, it seems best
to analyze Gestalt groupings as if they were elements, and adjust their weight based on the structure within them in a recursive fashion.

*Gestalt Continuity*

The Gestalt law of good continuation refers to an implied continuity that is present between two linear elements that align with each other in any orientation. This is applied through implied graphic alignment of typographic chunks as well as literal continuity between linear elements as they appear on their own or as a part of an image. As it relates to the example above, there should be some form of continuity between the three elements that make up the grouping. Whether straight or curved, the edges of each element will create an implied leading line that continues and may be used to direct the viewer to another portion of the composition.

Evaluation of the effect of Gestalt continuity on object salience may draw from the work of Saund [23] on identifying perceptually closed paths. This evaluation presents another problem that must be taken into account. Namely, each element in a composition may participate in not only a visual alignment as was investigated by Saund, but also an implied alignment based on its interaction of other elements. Identifying individual alignments between elements can get rather complex as the shape of each element becomes more complex and the number of elements under consideration increases.

*Texture*

Texture is an attribute that relates to the repeated occurrence of similar motifs in a certain area of the layout. This attribute is very complicated as it can result from any combination of element and positioning attributes, as long as the textural space is bounded in some way. Once a texture has been established, the analysis shifts from individual elements to the overall effect of the area as described by an extreme application of gestalt grouping.
The result of texture in a layout is that it is seen as a large element in itself, with a weight that is directly influenced by the overall gray value of the texture created. A side effect of some textures are other gestalt relationships such as continuity created by lines within the pattern or resulting from the spread of color in a linear fashion. Because of this, a sophisticated technique must be developed to deal with the interaction of lower-order attributes within a texture.

Anomaly

Anomaly is a focal technique used to bring attention to a single element within a repetitive sequence by changing it enough to make it stand out. This is accomplished by adjusting the element attributes of a specific element in the sequence. There is often a direct relationship between anomaly and texture, as anomaly requires the repetition of a motif in order to differentiate one element from the rest.

As a result of its relationship with texture, anomaly must be handled in much the same way as texture, with the weight adjustment beginning with a general weight assigned to the elements in the texture or pattern itself.

Intervals

The concept of intervals is a very important topic in graphic design. It describes the distribution of white space between elements in a composition. Interval adjustment may be used to set off Gestalt groupings or simply provide a kinetic visual rhythm to the composition. This will be seen in the case study in Chapter 5.

Evaluation of intervals effect on element salience is directly related to its application in graphic design. When placing objects, designers attempt to create dynamic intervals, meaning intervals that are not very similar among each other. As such, it is possible to measure all intervals present in a composition and calculate an overall similarity between the members of this set. From that, groups of intervals that
are too similar can be identified and salience can be adjusted accordingly. An important consideration when calculating this similarity is the location of the intervals since similar intervals may exist in a composition if they are separated by enough variation. A sufficient measurement of this locality should be determined experimentally.

Alignment

Alignment is directly related to gestalt continuity as a special case of it. When creating alignments in a composition, the edges of an element form linear sub-components that imply a line connecting them with other elements in the composition. If an element is off alignment by a small amount from what is intended, the visual flow of the layout is greatly reduced as elements appear disjoint or connected illogically with unrelated elements.

The effect of alignment on the overall salience of an element is proportional to the effect of the elements it is connected with. Also, if a gestalt grouping is formed by multiple elements being in alignment, the structure of the entire group must be considered. As a result of this connection with the Gestalt principles mentioned above, evaluation of the effect of alignment on salience will involve the methods utilized for the principles of grouping and continuity.

Leading Lines

Leading lines are elements that are linear in nature and affect the flow of a composition much like alignment. The difference between leading lines and alignment are that leading lines are physically present in the composition and, as such, contain element attributes.

The effect of this attribute on salience is a combination of the specific weight of the linear element as well as the effect of the gestalt continuity on the general flow. It lends weight to the element being pointed to by it based on a function of its own weight as
determined by length (size), color, and thickness. From these considerations, this may be evaluated as a special case of Gestalt continuity.

**Contrast**

Contrast is another important element used in developing a visual hierarchy. This is the characteristic of an element that differentiates itself in importance from its neighboring elements or from the background itself. Any of the previously mentioned attributes may have contrast applied to them in a layout.

As a result of its generality, contrast can also be very complex, and it may have a different method of analysis for its effect on each attribute. For example, the difference between the color of two objects involves different perceptual connotations than the difference in sizes of two objects. Therefore, contrast must be evaluated within each category of characteristics. It is contained within inferred modifications, as opposed to it being identified within each prior characteristic, because it involves the interaction between two elements and may be applied in the last phase of calculation.

**Figure-Ground Segregation**

Figure-ground segregation is the final Gestalt principle being discussed here. It describes an element as being perceptually separate from the background. This characteristic is a function of contrast. Also called foreground-background separation, figure-ground segregation is a property that was proposed by Koffka [12] in Gestalt psychology.

The problems that arise with this characteristic as it relates to element salience come from the need to decide when a figure becomes ground or vice-versa. This problem is open-ended and may be a result of a combination of previous characteristics as they relate to contrast.
**Literary Hierarchy**

The concept of a literary hierarchy is a new proposition based on examination of eye-tracking research. Much like the rule of thirds, this characteristic is hidden in the structure of the page and is related to the natural page-scanning patterns a viewer goes through when presented with textual information. Eye-tracking research has described this common pattern as the “F-pattern”, named as such because it creates the shape of the letterform “F” [7].

As it relates to the overall salience of an element, this page-scanning pattern will be investigated much like rule of thirds, as it is a perceived connection and not a factual connection. The assumption to be tested is that the elements in the upper left-hand corner and those connected along the F-pattern hold a more prominent position by default. It is because of this connection between elements that this characteristic is referred to as an Inferred Modification.

**Visual Hierarchy**

Visual hierarchy, as the final goal of objective analysis, is indeed a characteristic that arises from inferred modifications. An important point to take note of when considering the final stage of visual hierarchy is that it is ambiguous. Unlike the literary hierarchy that is described above, visual hierarchy does not necessarily conform to any eye-scanning pattern. While affected by the F-pattern described in literary hierarchy, a designer may choose to include the most important elements at the bottom of the page while simultaneously lowering the importance of elements that appear higher in the composition to maintain balance and preserve the intended hierarchy of information.

This is listed as a characteristic because it is a property of visual layout. However, as it relates to an objective metric for analysis, it is a combination of previously considered characteristics and results from the algorithmic method used to analyze a layout.
Section 3: Algorithmic Considerations

We have seen many possibilities for constructing an algorithmic method to evaluate layout in the related work presented previously. As far as an algorithm goes for applying these characteristics to evaluation, one must take care not to reverse the effect of lower-level characteristics too much in application of higher-level characteristics such as those identified as inferred modifications. It is for this reason that it appears a balancing stage of evaluation must be completed, where calculated values are normalized and higher-order attributes applied with minimal error introduction.

The high-level, conceptual process that I have created through the course of this study proceeds as follows:

1. To begin with, the elements under consideration must be identified. This can be accomplished either by user-input, an XML-based representation of the file as is present in Adobe InDesign documents, or through algorithmic methods utilized in image processing. Ideally, the first two methods are most desirable.

2. These components must then be evaluated for their element attributes, followed by a combination step that aggregates the effects of each of these characteristics into a base salience value for each individual element.

3. Positioning attributes will then be evaluated, aggregated in a similar fashion and combined with the values gained from the previous step.

4. Inferred modification evaluation will proceed in much the same way, being evaluated separately where possible, combined with each other, and then combined to determine the overall salience value of each element.

This method, while very abstract, provides a basic structure for proceeding while allowing for expansion in the future. However, some important considerations must be
taken into account when implementing this structure.

First of all, all of the inferred modifications must draw from element attributes for their individual evaluation. It is for this reason that once each step is completed, the values are not discarded, but attached to each element as a value that may be utilized in later stages.

The method of aggregation of values is another concern that must be taken into account. It seems most apparent to determine a normalization function based on perceptual research to transfer each characteristic’s value to a scale on a set range, from 0 to 100 for example. These values can then be combined using linear means to create an overall view of the hierarchy. It is important to remember that the goal of this pursuit is to identify a high-level hierarchy of information present, a very coarse set of rankings. Therefore, it is possible to make some judgment calls on the aggregation methods that reduce accuracy slightly while improving algorithmic complexity. Further testing of the accuracy of different methods must be completed in the future.

When describing human perception, a consideration must be made to the actual area in which the user can see fine enough detail to evaluate these characteristics. Our eyes have a certain area where resolution and perceptual propensity is at its greatest. Outside of this area is where our foveal field of vision transitions to our peripheral field of vision. This area is described in terms of a visual angle. Visual angle is a function of the distance a person is from the stimulus relative to the size of the stimulus, shown in Figure 4. In this figure, A refers to the visual angle, D is the distance from the stimulus, and S is the size of the stimulus. Duchowski [7] (p. 31) calls the edge of foveal vision the “useful” visual field and states that it extends to a visual angle of about 30°. The area outside of this field has poor visual acuity and is mostly used for perception of motion. Ideally, when perceiving a layout, the useful visual field will cover enough of it to be able to lead the viewer to the next logical focal point in the hierarchy. However, when viewing a poster
up close to read text, for example, the entire composition is not available to the viewer's perceptual process. For this reason, one must consider the distance from which a layout is to be viewed when aggregating values.

Finally, granularity of evaluation must be taken into account when implementing the final algorithmic process. Doing pixel-by-pixel analysis of an entire composition is an extremely computationally intensive task especially when considering the number of characteristics presented above. Care will have to be taken to abstract out to higher levels of granularity, such as element-by-element, wherever applicable.
CHAPTER IV
CASE STUDY

For this case study, I will be examining a poster created by a designer named Todd Timney in 2006 and cited from *Typographic Design: Form and Communication* [4] (page 124). This poster contains some very complex elements, such as the guitar, but there are also some basic interactions that may serve as a sufficient example of the characteristics described previously. The goal of this case study is to analyze this design from the standpoint of the framework that has been presented and further illustrate the characteristics with a visual example. Exact measurements and mathematical methods of metric aggregation cannot be included, since these portions will require further research. The analysis follows the structure presented at the end of the previous chapter.

**Identifying Elements**

This design, perceptually, gives many options for possible salient elements. Through manual visual analysis, we can identify at least 23 different possibilities. In numeric order as labeled in Figure 4, with descriptive names assigned, they are: Deer Image\(^1\), Deer 12 Cloud\(^2\), Headline\(^3\), Name Subhead\(^4\), Left Gun\(^5\), Right Gun\(^6\), Guitar\(^7\), White MK Logo\(^8\), Guitar Text Orange\(^9\), Guitar Text White\(^10\), Man Icon\(^11\), Man 12 Cloud\(^12\), Top Left Text 1\(^{13}\), Top-Left Text 2\(^{14}\), Top-Right Text\(^15\), Body Headline\(^16\), Body Text P1\(^{17}\), Body Text P2\(^{18}\), Body Text P3\(^{19}\), 12 Star\(^20\), Bottom Subhead\(^21\), Bottom Body Copy\(^22\), and Background Texture\(^23\). These may even be able to be subdivided into different sub-elements as will be discussed below.

Determining the elements automatically is possible using methods in image
Figure 5: Case Study Poster Example
processing by adjusting the sensitivity of the algorithm used to identify them to both
discover the foreground elements as well as the subtle background variations. These
variations in the texture image are important because they dictate much of the structure of
the composition in this case. The next step is evaluation of the characteristics, and only the
most prominent examples of each will be considered for brevity. Also, I will present more
specific possibilities for further study of the salience value for each that have been drawn
from these examples.

**Element Attributes**

*Size*

The size characteristic must be measured on some scale. Within this framework,
it will be measured with respect to the total size of the compositional area to assign an
initial salience value to an element and normalize values from the start. Since the size of
each element is a very simple measurement, I will discuss the special case of the Guitar element. Depending on the final desired print size of this poster, the large size of the guitar
could cause a strange phenomenon. When a viewer is examining this poster, he or she will
not be able to view the entire guitar element within the usable field of vision previously
described.

This causes a certain problem when assigning an initial weight value to the guitar element itself, where it can be subdivided into smaller elements because of its size and shape. For example, the body of the guitar holds more weight by size than the neck or headstock. Dealing with this sort of segmentation is an image processing concern and can be solved during the previous phase.

*Weight*

There is a large variation in weight among the elements in this composition. For
the body copy in the top right corner, the spacing between each line causes the
overall weight to appear lighter than if there were less space, overall it is a light grey value and will be assigned a lower salience value based on weight. On the other hand, the body of the guitar is entirely black and draws a much greater amount of attention by weight because of this and its salience value must be modified accordingly.

**Color**

As far as color goes, the entire composition follows a rather simple color scheme. Every element is one of four possibilities: orange, white, black or some shade of grey. However, it is unclear which of these four colors are the most salient. Further research needs to be conducted to determine which colors are more likely to attract attention as they relate to others in a composition. The goal in attaining a proper evaluation is to have a ranking system of colors in order of inherent salience. Possible hurdles to this concept are that the relationship may not be linear, and also it will be affected by perceptual disorders that alter a person's perception of color.

**Aggregation Phase**

These values interact at a very basic level. For example, the amount of attention drawn to an element based on weight is also related to how large it is. A small element with a high weight value will attract more attention than a larger element with a low weight value. Because of this, color and weight can be thought of as scalars that modify the base weight granted by the element's size. Using the guitar as an example, the body of the guitar has high values for weight and size, so it draws more attention to itself than the deer or the man icon. Care must be taken when combining the values for these characteristics to preserve this sort of behavior.

**Positioning Attributes**

**Location**

As a valued measurement, location is an absolute metric consisting of the X and Y
coordinates of an element. It does not necessarily have a basic salience value assigned to it, however it is important in determining the effect of the other two characteristics.

One problem that must be considered when measuring location is determining from which point within the element to measure the location. In his discussion of balance, Arnheim [1] (page 8) mentions the concept of the perceptual center of a shape and gives examples of some basic shapes and their perceptual centers. This is referred to frequently in his work, when describing the effect of tension within an element as well. For example, much of the weight for the man icon is located on the man’s torso. This location is a visual approximation and may be related to the size of the element. If the image of the man were larger, perhaps a more salient area for the perceptual center would be his face, however this would be the result of semantic data. Basically, we want to look at the face more because it's an image of a person.

One possible method of calculating this location within an element based on syntactic data can draw from methods used in calculus for finding the center of gravity in a two-dimensional shape. Another possibility would be to use the element itself in a sub-processing task as input to the same evaluation process described here, with the highest ranked portion of the element within its sub-hierarchy being used as its perceptual center.

Page-edge Relationship

The effect of this characteristic is one that is based upon how it makes the viewer feel. For example, the right gun is a bit more salient than the left gun because its perceptual center is closer to the page edge. The perceptual effect is that an element becomes more unstable, or salient, as it approaches the page edge. Further study of the rate at which this occurs must be conducted through future research.

Rule of Thirds

The rule of thirds is so prevalent in all visual forms such as art, design and photography that its effect is undisputed in these fields. Within the composition
being analyzed, notice that the headline, the semantically most important element for informational purposes, is placed most near a point of interest dictated by the rule of thirds, namely the bottom left intersection point. Measurement of the effect of this characteristic is similar to the page-edge relationship and also must be studied further.

**Aggregation Phase**

The positioning attributes are a special case in evaluation. Each one of these characteristics interacts with the others to determine the base positioning salience of an element. The result of this interaction is a basic salience map of the compositional area, consisting of hot spots much like those described by Arnheim [1] in Figure 1. Discovering how the shape, size and proportions of the page affects these hot spots is a subject for further research.

Aggregating the effect of the evaluation of the salience values based on positioning with those calculated in the last step is slightly unclear. What must be decided is if these values are linear modifications of the previous values, or scaling multipliers of the element attribute values. It seems appropriate that the method of combination would be linear. If a multiplicative calculation is used, the effect could possibly cause previously calculated values to become skewed. However, a properly balanced measurement and aggregation method based on further research could be viable for both strategies.

**Inferred Modifications**

Because these characteristics are the most complex of all those mentioned, from interactions between multiple elements, each will require its own specialized method of study to determine exact values for salience. For this reason, I will only describe examples of each that are present in the composition being studied as an illustrative example and will omit the Aggregation Phase from this section. Possible research methods for some have already been described in the previous chapters, within the context of the related
work.

**Gestalt Grouping**

Perceptual grouping of elements is evident in a few places in this composition. From visual analysis, we can identify at least two major groups. The most prominent consisting of the headline, deer, deer 12 cloud, and name subhead. The second most apparent is in the top right corner, containing the body copy and related information to it. A third possible grouping could include the elements between these two groups that are gathered within and near the guitar element.

**Gestalt Continuity**

There are a few wonderfully evident examples of Gestalt continuity, in both subtle and explicit situations, present in the composition. Explicitly, there is a continuity created between the text of the name subhead and the left gun. This forms a relationship between these two elements that adds salience to each of them, with a push and pull between the reading direction of the text and the opposite directionality of the gun image. A subtler example of continuity can be seen in the top boundary of the background texture. It is split by both the neck of the guitar and the body copy element, though it is still perceived as a continuous, horizontal line across the page. Moreover, it is formed by the boundaries of a texture made up of repetition of dots.

**Texture**

The presence of texture in this composition is a very subtle one, being seen in the dot matrix pattern of the background texture. The sub elements that make up this texture are inherent in the style of dot matrix representation and create a few important implications to weight. Firstly, the overall effect of the texture is that the bottom two thirds contains a bit greater weight than the top portion. Also, there is an intricate set of sub-elements present in the variations within the texture, which influence perceived alignments and groupings in a small way and lend a better sense of balance overall.
Anomaly

There isn’t a prime example of anomaly within this composition as the motifs contained within it are not repeated enough to form a pattern which can be broken by an anomaly.

Intervals

The intervals present in this composition are just variable enough to not appear disjoint, but are not too redundant. This is mainly caused by the interaction between the diagonal element, the guitar\(^7\), and the vertical elements, such as the grouping of body copy in the top-right.\(^{17, 18, 19}\) In contrast to the changing distances of the guitar neck to the vertical elements, the vertical intervals between the top-left text \(^1\), the top of the background texture\(^2\), and the deer\(^1\) as well as between the bottom of the headline\(^3\) and the bottom text\(^{21, 22}\) are similar enough to help maintain balance. The main concern for study here is the relative similarity between the set of interval lengths.

Alignment

Implied alignments between elements are what gives structure to the layout and connects elements in a subtle way. There are both perfect and imperfect alignments present, meaning those elements that align exactly with one another and those that are off by a small amount. There is a perfect alignment between the baselines of the top-left text \(^1\) and the top-right text\(^15\), giving them an implied connection and also bounding the space at the top. An example of an imperfect alignment is between the left edge of the body copy\(^{47, 18, 19}\) and the right edge of the headline\(^3\). Though they are not exactly in alignment, there is still a flow that travels down from the body copy back to the headline, lending more salience to each.

Leading Lines

A very good example of leading lines in this composition is the graphic form created by the neck of the guitar\(^7\). The function of this leading line is to add emphasis to
the top-left text\textsuperscript{13, 14} as the viewer’s gaze travels up after reading the headline\textsuperscript{3}. It also serves as a method to discover the text that is both inscribed within it and lying along its edge.\textsuperscript{9, 10}

\textit{Contrast}

Being a very diverse characteristic, I will discuss the contrast as it is present here in two forms: size and weight. The most extreme example of contrast in size within this composition is between the man icon\textsuperscript{11} and the guitar\textsuperscript{7}. Being so close together emphasizes this relationship as well, indicating that the effect of contrast may have a positional component. As for weight, the ideal contrast in weight exists between the two extreme values of weight, white and black. This is used here to set off both the headline\textsuperscript{3} and MK logo\textsuperscript{8} from the black of the deer\textsuperscript{1} and guitar.\textsuperscript{7}

\textit{Figure-Ground Segregation}

In an overall context, the form of the elements that are contained within the orange of the background for a figure-ground relationship, where the figure becomes the element present, such as the man icon,\textsuperscript{11} and the background color/texture. This relationship is inverted within the neck of the guitar,\textsuperscript{7} where the orange text\textsuperscript{9} becomes the figure and the guitar\textsuperscript{7} becomes the ground through inscription of the text. It is important to note that some contributing factors to this characteristic are the contrast of weight, the page-edge relationship, and intervals.

\textit{Literary Hierarchy}

The basic structure of the literary “F-pattern” is leveraged in this composition by the placement of the text relating to the location of the event.\textsuperscript{13, 14} This text, as it is small, light in weight and contains relatively low contrast with the background, gains salience by being in the primary position for a culture that reads from left to right.

\textit{Visual Hierarchy}

Overall, the main flow of this composition is circular in nature. Perceptually estimated, the ordered ranking of elements begins, naturally, at the headline.\textsuperscript{3} It then
follows to the guitar\textsuperscript{7} and guns\textsuperscript{5, 6}, up the guitar neck to the top-left text\textsuperscript{14, 15}, to the body copy,\textsuperscript{17, 18, 19} then back down to the headline\textsuperscript{3} to repeat again. That being said, the elements on the bottom of the page\textsuperscript{21, 22} are not entirely left out of the hierarchy as they do benefit from some alignment with the headline\textsuperscript{3} letterforms and the deer image.\textsuperscript{1} These elements are simply deemphasized within the hierarchy by excluding them from the constructed flow.
CHAPTER V
POSSIBLE APPLICATIONS

The applications of this research are as varied as the fields involved in it. Graphic design may find use in education or practical layout assisting tools. Computer science and document engineering fields may find use in automatic layout of a number of visual elements, including visualization, web sites and paper layouts. The field of cognitive psychology may use it experimentally in a fashion that further explores the nature of human perception and refines the algorithmic weights for a specific demographic or individual.

In graphic design, layout evaluation could enable a computer game to be developed to aid in design education. In this game, situations could be presented where a student is given an amount of text and images to be laid out on a page. Through analysis as the student works, the application can provide automatic feedback about the student's progress that can guide him or her in understanding these abstract design principles and help build compositional awareness.

This same sort of evaluation as applied in a limited scope such as the game mentioned above could be used in an Adobe InDesign plugin. Adobe InDesign is one of the prominent layout creation applications being used in graphic design. It contains sophisticated means of manipulating elements on a page and has automatic means of aiding designers in aligning elements properly. A plugin that leverages evaluation of layout could provide similar recommendations to improve a design as it is being created.

Another application may be found within the scope of digital cross-platform representations, as in web layout and open sourced software. There is heavy research in
these areas, with respect to information design, and ways to handle multiple screen sizes in practice with web development. However, web developers are forced to create two versions of a site, one for mobile and one for desktop viewing. Perhaps a set of general aesthetic principles such as those proposed here may be employed to adjust a layout based upon extreme changes in screen size or other usability information that is gathered.

Another possible application that takes advantage of computer technology could be in the form of a system that can be integrated into an iPad application. The purpose of this system would be to evaluate a user’s distance from the device using the front-facing camera and adjusts layout of whatever is being viewed based on their gaze and calculated foveal field. This could be accomplished by using eye-detection techniques and measuring the size of the user’s eyes to determine distance. The approximate direction of the user’s gaze can also be determined by the position of the iris area on the image his or her eyes. One example of how this could be useful is in a visualization application that determines where a user is looking and arranges the information in a way that facilitates faster scanning of the data sets.

Finally, this research could, much like in the educational example, be employed to automatically generate a rough hierarchy utilizing semantic input from the user in the form of a visual presentation. This could be in any form of application from prototyping a web page to generating quick ideas for starting points for poster and print design. It would be utilized as a low-cost brainstorming activity that may weed out bad designs in the early stages of development, instead of allowing the possibility of spending much time and money polishing a bad design that doesn't accomplish the goals of the project.
CHAPTER VI
CONCLUSIONS & FUTURE WORK

With the many techniques being used by graphic designers in practice, it is clear that the creation of a weight value to determine a saliency map of a composition is a large undertaking. By identifying the methods employed by graphic designers to achieve hierarchy, I am constructing a survey of techniques for future analysis. Much research will need to be completed into the basic nature of the aforementioned characteristics in order to build a system that takes advantage of the extensive research completed in the past through graphic design research. Through an open-ended system of evaluation beginning with the simplest element attributes and building up the effect of the higher-order attributes on the weight of an element, the evaluation will be refined over the course of future research. Towards this evaluation, this work contributes a problem analysis and literature review, resulting in a basis with which to conduct further experimentation in the pursuit of an objective layout evaluation metric and algorithmic method.

In the future, I intend to further explore these specific topics and pursue an implementation of my proposed algorithmic methods, resulting ultimately in a usable tool for an application described previously. This investigation will include user studies, eye-tracking surveys and cross-platform application development with the intent of experimental studies on perception. Resulting from this research will be a master’s thesis and the possibility of a continuing research program if initial investigations indeed prove fruitful.

To conclude and inspire further research in identifying the actual effect of these characteristics on element salience as it relates to human visual perception, I would like to
include a quote from Rory Sutherland in a discussion article posted on the website for the TED: Ideas Worth Spreading conferences:

Let me continue a point: Actually quite a lot of things are counterintuitive, in many ways. Metrics are also dangerous because, in particular, with relation to humans, very few metrics are linear. If you have anything involving either groups of people or networks of people, or just human stuff in general, you may, for example, notice that there's a huge difference in personal satisfaction, in changing a train service from hourly to half hourly. You may also notice there's a big improvement in satisfaction between getting it down from half hourly to, let's say, every 20 minutes. However, what you will probably find, then, is that beyond that point, not necessarily that point, but beyond some point, it becomes pretty much irrelevant. See, quite a lot of metrics in terms of human psychology and satisfaction and value, are actually non-linear.

This quote encapsulates the issues that must be dealt with when measuring the effect of each of the physical property changes as they relate to a person's perception of a composition as a whole. His take on metrics as they relate to humans is something that has been examined many times in nature through statistical analysis. Many things that occur naturally are indeed counterintuitive and more still are non-linear. It is for this reason that future research into the topics described within this thesis must be conducted without any preconceived notions.

It is now clear that, even with the sophisticated technical advancements that have been seen over the past decade, the problem of automatic analysis of layout remains open and incompletely solved. Many attempts have been made to distill perception into a means
available to automation, but there are extreme intricacies present in professional design that still elude analysis. This field is in its infancy, and future research has the potential to bring great things to the way that we perceive visual design itself.
REFERENCES


