2011-04-22

Outdoor Scenes for Data Visualization

Benjamin A. Hillery
Brigham Young University - Provo

Outdoor Scenes for Data Visualization

Benjamin A. Hillery

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Robert P. Burton, Chair
Michael D. Jones
Sean C. Warnick

Department of Computer Science

Brigham Young University

June 2011

ABSTRACT

Outdoor Scenes for Data Visualization

Benjamin A. Hillery
Department of Computer Science, BYU
Master of Science

Recent cognitive research indicates that the human brain possesses special abilities for processing information relating to outdoor scenes. Simulated outdoor scenes are presented as an effective method of displaying higher dimensional data in an efficient and comprehensible manner. We demonstrate novel methods of using outdoor objects and scenes to display multidimensional content in a way that that is intuitive for humans to understand, and to exploit various cues commonly found in scenes from the natural world to communicate the values of multiple variables.

Keywords: data visualization, natural scenes, graph comprehension, visual cognition

ACKNOWLEDGMENTS

I express my gratitude to Dr. Burton for his tireless support, insightful advice, and unwavering optimism. I am indebted to Dr. Stacey and my other colleagues at ITT corporation for advocating my pursuit of higher education. E-on software provided essential tools for constructing the concept figures, and example images provided by Doug McCune were invaluable. I owe thanks to my colleagues Luther, Daniel, Mike, Lee, and Jacob for helping refine the associated presentations. I dedicate this thesis to my wife Krista, for without her faith and encouragement this work would not have been possible.

Contents

List of Figures

Chapter 1

Introduction

Multivariate data visualization presents a significant challenge, in part due to the limits of human perception. There is a finite number of attributes that our visual systems are capable of extracting from a given image[1]. We encounter difficulty when we want to show more variables than we have attributes to display them with. Since three variables represents somewhat of a barrier to what is known to be unambiguously representable in a distance-preserving presentation, throughout this work we use the term multivariate to mean "having more than three variables."

One might be tempted simply to avoid using multivariate data by suppressing variables or by avoiding sets that do not map well to the natural dimensions of human visual perception. Unfortunately, the collection of problems involving multivariate data is vast and growing in a way that makes finding an elegant solution both interesting and desirable (Nowell, 2011). An example of a data set that includes many variables of simultaneous interest is found in the context of geographic information systems, since there are many geographic traits that map to the same two-dimensional space. Population density, earthquake risk, average rainfall, temperature, altitude, barometric pressure, bedrock type, native vegetation cover, and pollution content all represent variables that could be associated with a single location on a map. It would be desirable to be able to show all of the information in a single presentation and in a way that allows us to interpret, relate, and analyze it efficiently.

There are several proposed solutions to the problem of higher dimensional information presentation, some more effective than others. Common techniques include having the computer abstract away certain details that might confuse a human viewer, unfolding dimensions so that they

[1] We entertain a more detailed discussion of these attributes in chapter four

all appear simultaneously, or even re-encoding the data as a shape or object more recognizable to a human observer. A summary of the most important of these techniques is included in the related works section.

The varied success of the different proposed methods invites the question, “What makes a good visualization method?” Since humans are the end consumers of information, the human visual system is a critical part of the visualization process. Despite this, it is unusual for authors introducing new visualization techniques to attempt to explain how a technique functions from a cognitive standpoint. Indeed, not one of the visualization technique papers cited in the related works section contains a reference to the cognitive science literature. The traditional approach defers to the reader’s tacit understanding of human vision by showing demonstration images and having the reader subjectively evaluate their qualities.

Perceptual benchmarking has been proposed as a more rigorous alternative to direct subjective evaluation of visualization methods (Ward and Theroux, 1997). However, benchmarking with human test subjects is tedious and time consuming, making it difficult to use in the the development of new techniques and completely impractical when creating new visualization instances. While it is conceivable that a visualization system with objective human-based feedback could converge on a solution to a new problem through trial-and-error, the time and expense of this “1000 monkies” design approach are prohibitive. In the case of either direct subjective or perceptual benchmarking methodologies, we are left without the ability to predict the quality of a visualization system’s output prior to showing it to an actual human.

To avoid these pitfalls, we diverge from tradition and attempt to explain the cognitive basis for our visualization technique. Our aim is to provide insight and direction in addition to the traditional visual confirmation. Our approach is to combine established principles of graph comprehension with the latest research in human cognition.In this thesis, we investigate the use of outdoor scenes to display multivariate data. We begin with a review of relevant prior research, then move on to discuss the fundamentals of data visualization. We then present our justification for using outdoor scenes to visualize multidimensional data, and discuss recent cognitive and neurological research into scene perception. After presenting possible methods for constructing scene visualizations, we compare our proof-of-concept examples to examples produced by similar existing methods. We demonstrate that given certain constraints, outdoor scenes can be an effective visualization technique.

Chapter 2

Related Work

2.1 Multidimensional Visualization Techniques

Higher dimensional data visualization is a long-standing problem investigated by researchers from a multitude of disciplines. Many techniques have been proposed but all have deficiencies, especially for datasets where there are many independent variables that need to be analyzed simultaneously. In this chapter, we detail the most important of these approaches and the tradeoffs involved in their use.

Higher-dimensional visualization techniques can be classified into two groups: methods that reduce the displayed dimensionality to less than what is in the original data set, and methods that attempt to preserve all of the original variables. Since many visualizations use a combination of methods, any given presentations may include elements from both categories.

2.1.1 Dimension-Reducing Methods

A common dimension-restricting technique is to project the data from a higher dimensional space onto a lower one. Examples of projection-style techniques include volume rendering (Drebin et al., 1988; Levoy, 1988; Sabella, 1988; Upson and Keeler, 1988; Schneider and Westermann, 2003), draftsman's plots (Tukey and Tukey, 1981), and star coordinates (Kandogan, 2001). In volume rendering, data with three or more dimensions is projected onto a two-dimensional plane. In draftsman's plots[1], pairs of variables are plotted against one another in an array of two-dimensional graphs. Star coordinates assigns a two-dimensional vector to each of the variables that are contained

[1] Draftsman's plots are sometimes called scatterplot matrices

in a data point, then scales each vector by the data values before summing them to acheive a resulting point in two-dimensional space.

The biggest advantage of projection-style techniques is that they preserve at least some spatial configuration of the data, allowing the user to draw conclusions about relationships within the data set. For datasets orignating from a spatial domain such as geographical or medical imaging, this allows the information to be viewed within the original context. For non-spatial data, similar points appear close together in the resulting projections, facilitating cluster analysis.

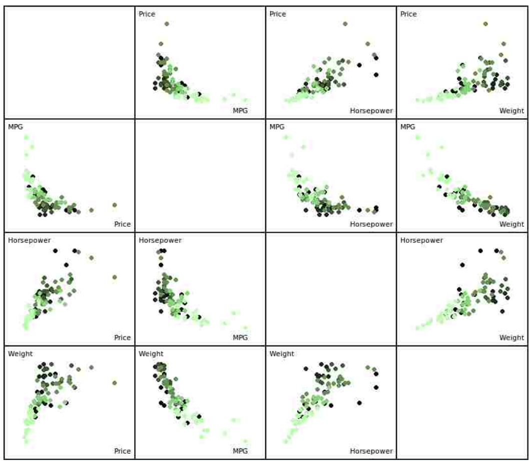
A significant problem in general with dimension-reducing techniques is that explicit information from the original dataset often is lost in the transformation. The many-to-one nature of dimension-reducing methods means that each point in the output space could have come from any of several points in the input space. The most common negative consequence is that dissimilar points in data space appear to be close together in the visualization. Occlusion is an extreme case of this, where multiple distinct data points are rendered indistinguishable in the output. Such ambiguitiy is inevitable if the output space is smaller than the input space.

2.1.2 Dimension-Preserving Methods

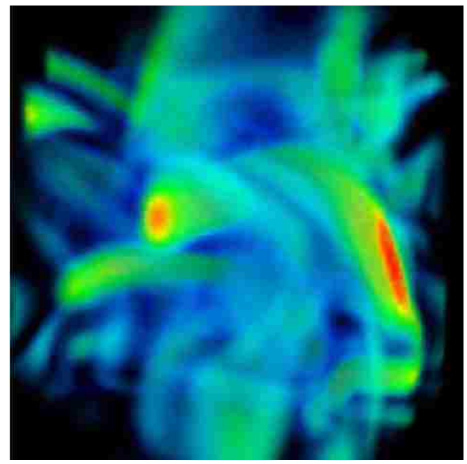
Dimension-preserving techniques typically work by “unfolding” or “unstacking” extra dimensions so that they all can be viewed simultaneously. In three dimensions, these techniques are analogous to unfolding a box and flattening it to provide an unambiguous view of all six sides. Flat data presentation techniques include tables, glyphs, parallel axes, and bar graphs. Figure 2.2 shows examples of several of these methods.

A typical table might present data points as rows of numerical values with the variables organized in columns. While tables preserve the individual values as well as any technique, they do not lend themselves well to the detection of patterns, relationships, or exceptions.

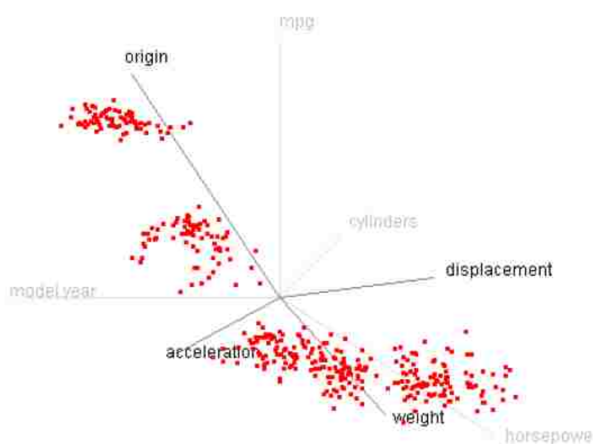
Introduced by Anderson (1960), glyphs are graphical data display objects designed to represent points in n-dimensional space. We define glyphs as graphical objects, representing single points in a k-dimensional sample space, which possess k visible attributes varying independently



(a) Draftsman's Plot

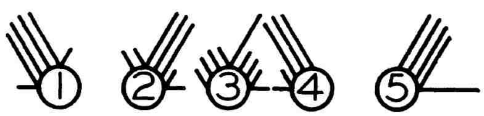


(b) Volume Rendering

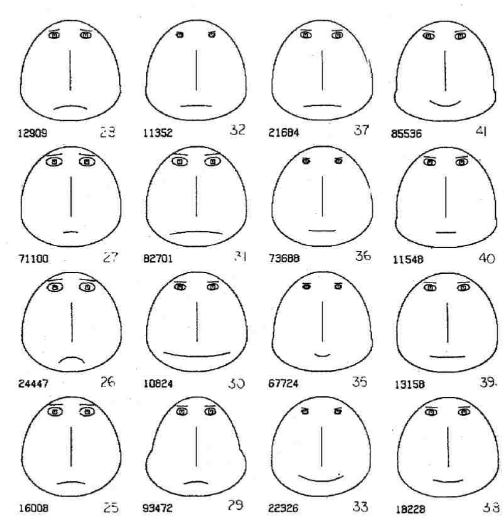


(c) Star Coordinates

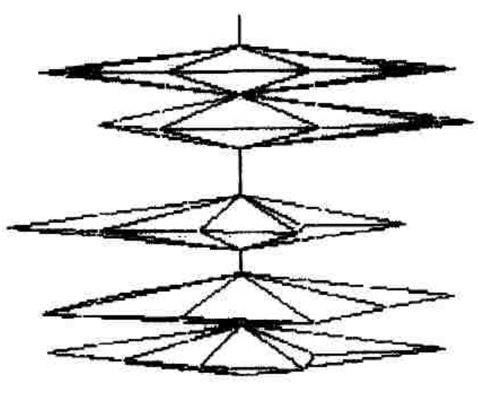
Figure 2.1: Dimension-Reducing Methods



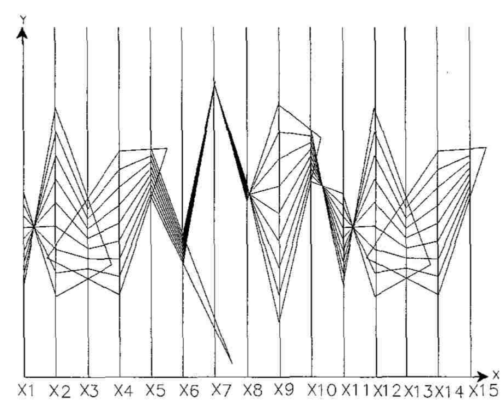
(a) Anderson Glyphs



(b) Chernoff Faces



(c) Pile



(d) Parallel Axes

Figure 2.2: Dimension-Preserving Methods

with the k data parameters (Ward, 2002). While some authors use the term "glyph" to refer specifically to Anderson's work, it has come be used more generally for any technique that meets this definition. Glyphs thus encompass a variety of multidimensional display techniques such as star plots (Chambers et al., 1983), piles (Melville and Burton, 1997), and Chernoff faces (Chernoff, 1973).

Star plots consist of a number of lines radiating outward from of a central focal point, with a polygon encircling it whose vertexes coincide with the line endpoints (Chambers et al., 1983). The distance of each vertex down the line indicates the value of a given variable along that axis. Star plots are used to qualitatively compare various data points and to identify groups of similar and dissimilar points.

Piles extend star plots into three dimensions so that the stars appear closer together in the presentation, thereby making relationships between individual point more apparent (Melville and Burton, 1997). However, since star plots have now been brought into the three-dimensional domain, occlusion problems may require interactive visualization to resolve.

Chernoff faces are glyphs that uses human facial features as the visual elements (Chernoff, 1973). The faces are intended to exploit the capability of the human visual system to intuitively and rapidly detect even small variations in facial features. However, Chernoff faces have several significant deficiencies. They tend to perform badly in perceptual benchmarking studies (Lee et al., 2003; Morris et al., 2000) and have a reputation for a relatively unprofessional appearance (Cluff et al., 1991), perhaps due to a problem common to many computer-generated faces known as the "uncanny valley". The uncanny valley refers to a level of unnaturalness that some synthetic faces and bodies have that makes them uncomfortable for humans to view (Mori, 1970).

The parallel axes technique is similar to star plots, except that instead of using polygons to represent data points, polylines are used instead (Inselberg, 1985). The individual variables of a data point are represented as vertical, parallel lines, and the values of the variables are points along those lines. A data point is then represented as a polyline that connects the points on the vertical lines. To visualize a given dataset, the data points in the set are all plotted onto the same vertical lines (parallel axes) so that they can be compared and analyzed. Parallel axes' drawbacks include low information density and lack of intuitive feel. The technique has been extended in the form of parallel planes (Xiaoxi and Burton, 1992) and parallel volumes (Xun and Burton, 1997).

A general disadvantage of dimension-unfolding is that the configuration of the original space is no longer preserved. Like a cubist painting, seeing all sides of everything laid out flat provides a distorted perspective. Data points that are close together in data space may appear far apart in the presentation, making interrelationships between points more difficult to determine. These inter-spatial relationships form a set of implicit information that may be important to understanding the data set within the context of its target domain. Of course, the implicit information could be recovered by mapping the data back to the original space, but to do so usually requires significant mental effort. This effort limits flat presentations to smaller datasets, typically consisting of a few hundred points or less. Additionally, since the structure of the source space is not preserved, the structure of the output depends entirely on how the unfolding takes place. A single point, "unfolded" in two different ways using the same technique, can produce two entirely different representations. The star plot for a given point may appear alternately spiky or lopsided, depending on how the variables are assigned. Variable assignment is therefore a critical consideration when using any of these techniques.

Despite this, not all domains are spatial, or can be mapped to a spatial domain in any meaningful way. An example of such a domain would be the various parts of speech in the English language. Since there is no standard measure of the distance between parts of speech, it is difficult to form a non-arbitrary spatial relation between them. Forming such a relation is akin to asking how far a verb is from a noun. For information from non-measurable domains, which we call categorical information, a significant disadvantage of dimension-unfolding essentially vanishes.

2.2 Other Important Related Research

2.2.1 Statistical Dimension-Reducing Methods

Another way of reducing data dimensionality is by analyzing the statistical properties of the data set in an attempt to find an alternative set of variables that compresses the data set into fewer dimensions while retaining most of the relevant information. Techniques of this style include principal component analysis, multidimensional scaling, and correlation methods. The drawback of these methods is that any statistical analysis must make assumptions about the structure of the data set, assumptions which will not hold for arbitrary data sets (Jacoby, 1998).

2.2.2 Statistical Cluster Analysis

Another statistical method for reducing data complexity is cluster analysis, which classifies groups of datapoints into clusters according to statistical similarities. The clusters then represent a new meta-dataset with fewer datapoints than the original. Cluster analysis can be thought of as the dual of statistical dimension reduction, where the number of variables remains the same but the number of datapoints is reduced. The drawbacks to cluster analysis are similar to those with dimension reduction. Most clustering methods rely on distance metrics that make assumptions about the data structure that may not be correct for all situations. Further information on cluster analysis can be found in Kaufman (2005) and Tan et al. (2005).

Chapter 3

Mental and Visual Representations

Connecting external visual representations with our internal mental models is the fundamental purpose of data visualization. By better understanding how this happens, we can better understand how to improve it. In this chapter, we discuss foundational aspects of data visualization in preparation for the more involved discussions surrounding scene visualization. We start by introducing the concept of mental representation, our internal mental model of the world around us. We then talk about how humans have extended their representational abilities with external tools, before moving on to the more specific topics of visual representation and graphical comprehension.

3.1 Mental Representation

The mental ability to focus on certain stimuli while ignoring others is known as attention, and is recognized as an important cognitive faculty. In order to make rational decisions in a complex environment with limited mental resources, it is necessary to be able to discern the important information from the unimportant. Attended stimili are integrated into a mental model of the surrounding environment.

This model, a mental abstraction of an external system, is known as mental representation (McNamara, 1999; O'Brien and Opie, 2004). Representation, in a general sense, consists of three parts: a represented object, the vehicle representing that object, and an interpretation connecting the two (Eckardt, 1993). In the case of mental representation, the object can be anything imaginable[1], the interpretation is provided by the conscious mind, and the vehicle is our mental model.

[1] We mean this in the most literal sense. The target domain of mental representation is defined only by the limits of human comprehension.

Our mental model is constantly updated based on the information we observe, and retained for use when we are not observing it. This is akin to taking a photograph in a dark room using a large-format film camera and a bright, narrow-beam flashlight. Although the beam is too narrow to expose the entire photographic plate at once, a complete exposure can be formed by directing the beam of the flashlight around the room. Different parts of the film become exposed until eventually the plate contains an image of the whole environment. This “spotlight” analogy to directing attention was introduced by LaBerge (1983), while describing attributes of human visual attention.

3.1.1 Mental Representation Versus Indication

While a detailed discussion of the issues surrounding mental representation is outside the scope of this work, the distinction between representation and indication is a fine point we must address[2]. In the mind, representation and indication are related, yet distinct phenomena. Both representations and indications are internal mental constructs which can be created in response to external stimuli. However, representations are source-independent, transformable, and share properties with represented objects, whereas indications are source-dependent, non-transformable, and portable (Cummins and Poirier, 2004).

We say that representation is source-independent because any number of sources can result in the same model. For example, you might know the route from home to work because you drove a car along it, but you may just as easily have consulted a map or have listened as another person described it to you. However you received it, the meaning and content of the representation is the same.

Representations are transformable, allowing you to manipulate them in your mind to make inferences and gain new insight from past observations. You are able to reconstruct a response to a question to which you did not immediately know the answer because of your ability to mentally manipulate your perspective of the model and thereby draw conclusions from it.Mental

[2]The terms *representation* and *indication* are used in this work following the convention established by Cummins and Poirier Cummins and Poirier (2004), but terminology in the cognitive science literature varies on this topic. For example, O’Brien and Opie O’Brien and Opie (2004) describe what we term indication as a form of representation grounded in causation.

representations share common attributes with their target objects because they structurally resemble them at an abstract level.

Indicators are not transformable because they do not resemble the condition that caused them, but merely indicate the fact that the condition occurred. Indicators convey state information rather than structure. It is impossible to draw inferences from an indicator without an awareness of the context surrounding it. This is also why we say indication is portable, because the same stimulus can be made to signify different things depending on the situation.

3.2 Visual Representations

Visual descriptions are external representations that preserve information in a visual format. They can serve many purposes, but in the context of data visualization we focus on just two: exploration and communication. In the case of exploration, there is some class of objects we wish to understand. We don't have a specific enumeration of the objects in question, but we do know at least some of their properties. To conduct an exploration, we choose a description style that we hope will enable the human visual system to detect and identify those types of objects we want to find. In communication, there is a known set of objects that we hope to convey to others. The objects are explicitly enumerated and their properties are well understood. We construct visual descriptions of these objects to more efficiently communicate them to others.

The reason we create visual descriptions is because vision is a primary way that humans gather information about the world, and we know that humans are very good at solving visual problems. By creating visual representations of abstract informational content, we hope to apply human visual search and recognition abilities to arbitrary domains. In this sense, data visualization is a tool for exploring and expressing information from arbitrary domains in a human-compatible format.

The objective of data visualization is to enable humans to swiftly construct inner representations of the objects represented by the data. More precisely, we define data visualization as any process that uses a visual description to facilitate human mental representation of a system resembled by a data vehicle. To understand data visualization, it is useful to know why we have visualizations, how they are made, and how they are understood.

3.2.1 Construction of Visual Descriptions

Visualizations are constructed by transforming data into visual forms that are intended to be processed by the human visual system into a mental representation of some desired target system. This is typically done by encoding the relevant information and corresponding relationships into symbols and symbolic relationships. Symbols are visual objects with properties that scale in accordance with data parameters. To facilitate interpretation of symbol properties, many visualizations include non-data scale objects such as coordinate axes and legends.

3.2.2 Comprehension of Visual Descriptions

The process by which humans use visual descriptions to construct internal representations is known as graph comprehension. While details in the literature vary, most accounts of graph comprehension include the common themes of encoding, detection, identification, and interpretation (Winn, 1994; Pinker, 1990; Cleveland, 1993). The comprehension process starts when the visual system encodes the visual array into discrete elements, then detects the presence of objects and discriminates among them. The objects are then identified and their meaning is interpreted, eventually resulting in a detailed mental representation.

Graph comprehension is often described as having two phases. The first phase, called the preattentive phase, is where the visual system recognizes the existence of, establishes relationships among, and differentiates between the various symbols in a visualization. These processes are called preattentive because they can operate in the absence of any attentive activity. They occur rapidly and in parallel, independent of the conscious mind. The preattentive phase can be viewed as a sort of preprocesser for conscious thought, detecting important graphical objects and establishing relationships between those objects. Preattentively established relationships presuppose subsequent processes to certain interpretations, and can have a strong influence on attentive performance.

The attentive phase is where graphical objects are integrated into the into the system of mental representation. Attentive processes all require some level of focus or concentration, and occur one at a time in a serial manner. This demand placed on the conscious mind makes the attentive phase a cognitive bottleneck in the visual comprehension pathway.

The first step in the preattentive phase is encoding. Encoding happens early in the visual system, where the retinal image is deconstructed into constituent parts by bundles of neurons that respond to specific visual features. Neurons have been found that respond selectively to visual-field attributes like edges, size, orientation, and movement. These neurons can be thought of as indicators for low-level image features, and the types of stimulus that elicit their response effectively define the limits of visual perception.

Recognizing the presence of an object, or detection, is the next step in the preattentive phase. A visual object is any collection of visual elements for which attention is not divided. Detection of objects occurs very quickly, on the order of 100ms or less, and is a prerequisite for any attentional processing. Undetected features are essentially invisible to the conscious mind.

Accompanying detection are the steps of discrimination, establishing precedence, and configuration. Discrimination is the ability to discern differences between visual elements, establishing precedence is recognizing hierarchies within the visualization, and configuration is where groups and patterns emerge. The visual system discriminates objects by comparing their detected sub-elements. Objects that have similar subcomponents will be perceived to be identical unless the differences can be detected. Precedence and configuration together establish relationships between objects. Objects are detected in a progressive hierarchy starting with the most prominent object first, the object's most important features second, the important sub-features third, and so on. Between two objects, the one that gets detected first is said to have the highest precedence. Relationships are also established by configuring objects into patterns or groups. Objects are seen as a group when they are visually close together, are located within a visual boundary, or share symmetry. The rules by which object relationships are established are known as the Gestalt laws of grouping.

The outputs of the preattentive processes feed the later attentive operations. Parts of the attentive phase include identification, interpretation, and comprehension. Identification is the process of connecting an object in the visual field to a corresponding stored representation. Objects are thought to be identified when the subcomponents identified in the encoding phase correlate strongly with an existing representation. Unrecognized objects are iteratively re-identified and reconfigured until identification is achieved. After identification, the meaning of the object is interpreted by comparing it to known representations. Through interpretation, the conscious mind can either integrate the information into a known schema or adapt an existing schema to accommodate it. Comprehension is the end state of visualization, and occurs when all of the relevant information in the graphic has been integrated as an appropriate mental representation.

3.3 Discussion

The fundamental purpose of data visualization is to foster meaningful mental representation of a target group or class of objects. With reference to the principles of graph comprehension, an effective visualization is one that facilitates swift, accurate formation of robust, transformable representations. Greater speed allows us to process more information in a shorter amount of time, higher accuracy permits greater confidence in our conclusions, and more detailed representations enable more complicated transformations and deeper insights into the target domain. Speed, accuracy, and detail are the primary qualities that we use to define visualization performance.

Chapter 4

Scenes and Data Visualization

In this chapter, we discuss the motivations and cognitive basis for using naturalistic scenes as representational vehicles for digital information. We draw a connection to Chernoff faces, a closely related technique, and examine the recent research into face and scene perception.

4.1 Motivation for Scene Visualizations

Humans are especially capable at certain visual tasks such as size, motion, and orientation detection recognition (Snowden et al., 2006). These abilities all correspond to task-specific cerebral structures which help reduce the load on the conscious mind. For example, in a natural environment, objects that take up the largest proportion of the visual field are likely to be the most important, being either close to the observer or very large and distant. Rather than to have to attentively assess the size of every object in the field before concentrating on the most important, it would make sense for the visual system to quickly identify those areas that are most important and direct attention towards them. In a series of Nobel-prize winning experiments, Hubel and Wiesel (1959) identified cells in the visual cortex that respond selectively to the orientation of lines in the visual field . Similar studies have found structures that respond to color (Hadjikhani et al., 1998), motion (Movshon and Newsome, 1992), and distance (Pettigrew, 1972). These hard-wired groups of cells all aid in the rapid, parallel processing of visual information, operating preattentively and independent of conscious mental faculties.

Many visualization methods correlate well with these known strengths of the human visual system. Interpreting a bar graph requires judging the relative sizes of adjacent rectangles. Line graphs are useful for showing trends because line orientation is related to slope, and measuring slope is equivalent to differentiation.

Looking towards multidimensional data, elementary attributes like size and orientation have the capacity to encode only a single variable. To encode multivariate data, some combination of elementary attributes has to be used. However, mechanisms for doing so in an effective manner can be elusive. The use of many elementary attributes simultaneously tends to lead to a crowded presentation that is difficult to interpret. Many overlapping elements lead to ambiguities.

What we need is a coherent way of integrating many elementary attributes into a single multidimensional presentation. There are statistical methods that compress multidimensional data into the fewest relevant dimensions, easing the transition to lower-dimensional visualizations, but using these methods may be just forestalling the inevitable. There will invariably be a dataset with too much relevant information to compress.

Another possibility is finding a better way of presenting data to the human visual system. We would like to enable the visual system to simultaneously integrate properties of multiple low-level attributes to provide a single coherent representation of multiple variables. Such a mechanism would contradict traditional feature-integration theory, which asserts that processing combinations of elementary attributes must always require attentive processing (Treisman and Gelade, 1980). Despite this, if feature-integrational mechanisms did in fact exist, we could significantly improve the rate of visualization comprehension by bypassing serial attentive processing.

Encouragingly, certain areas of the preattentive visual system have been identified that may possess the qualities we desire. These domain-specific regions in the visual cortex respond selectively to certain configurations of multiple elementary attributes. In 2006, Downing et al. published the results of a study empirically confirming the existence of these domain-specific areas. To conduct their experiment, they flashed images on a monitor in front of a human test subject and observed the resulting neural response using functional magnetic resonance imaging (fMRI). In analyzing their results, they found three distinct regions of the primary visual cortex, each of which corresponds to a specific non-elementary class of visual stimulus. These regions, corresponding to faces, bodies, and outdoor scenes, are termed the fusiform face area (FFA), extrastriate body area (EBA), and parahippocampal place area (PPA), respectively. The identified regions all lie in the occipitotemporal pathway, the structure of the visual cortex most commonly associated with categorizing objects in the visual field into predefined schema.

Since faces, bodies, and outdoor scenes all contain multiple visual attributes, these types of objects present tempting candidates for visualization techniques. If the mind is able to preattentively form robust representations from these types of visual content, they could potentially revolutionize data visualization. Indeed, the facial processing aspect was investigated for visualization purposes by Chernoff (1973). While his work predates the major research into face perception, Chernoff cited anecdotal evidence that humans have a remarkable ability to detect and discern small differences in facial features. He reasoned that it could be useful to use this ability to aid comprehension of multidimensional information. Bodies have likewise been investigated for use in visualization (Perry and Donath, 2004), but we are unaware of any attempts to explicitly use cognitive scene processing abilities for the purpose of data visualization.

Being relatively unexplored, scenes may be a fruitful area in which to investigate new visualization techniques. In addition to the anticipated feature-integrational benefits, there are a number of other reasons why outdoor scenes merit consideration as a visualization technique. Outdoor scenes offer the potential for a high density of information due to the large number of possible symbols that can be represented. They may be able to avoid some of the aesthetic and implementation issues issues with Chernoff faces, while retaining the same intended benefits. These predictions are the primary motivations for this study, and balance of this work focuses on evaluating them.

4.2 Cognitive Research into Domain-Specific Areas

Before we entertain further discussion of scene visualizations, we need to more firmly establish our cognitive basis for doing so. By understanding the mental processes involved with the PPA, EBA, and FFA, we can gain better insight into any potential benefits and pitfalls of using them for data visualization.

We can start by defining more precisely the capabilities of the domain-specific brain regions we hope to exploit. Of the abilities in question, face processing has been the most studied topic, possibly because science has been aware of its special nature for the longest time. For this reason, we begin our discussion with a review of the issues surrounding faces and their use as visual representational vehicles. This will allow us to frame our later discussion on the less-studied nature of scene perception.

4.2.1 A Closer Look at Chernoff Faces

Cognitive tasks involving faces are traditionally grouped into three categories, namely detection, recognition, and feature-state identification[1]. Detection is the process of recognizing that an object within the visual field is a human face, recognition is identifying whose face it is, and feature-state identification is extracting meaning from the state of the various facial features.

The nature of human visual face detection has been a subject of intensive study. One of the central questions that scientists have tried to answer is whether or not face detection it is a preattentive process. An important method for identifying preattentive processes has been the visual-search paradigm. As established by Treisman and Gelade (1980), preattentively processed visual elements will tend to "pop-out" of a scene full of dissimilar elements, leading to rapid parallel visual search. Thus, we can detect whether an element is preattentively processed by measuring the time it takes a human to search for it. Initial attempts to detect pop-out for faces by Nothdurft (1993) and Brown et al. (1997) resulted with negative outcomes. However, since these early experiments used pieces of faces as distractors, there was some debate as to whether the visual-search paradigm was correctly implemented. More recent experiments by Lewis and Edmonds (2003) and Hershler and Hochstein (2005; 2006) have succeeded in producing a pop-out effect, but their results have not been uncontested (VanRullen, 2006).

[1]The literature has many different terms for these stages. What we call detection, recognition, and feature-state, some authors call categorization, identification, and emotion, respectively

Harder to dispute is new evidence being provided by the neuroscience community on the preattentive nature of face detection. By combing high spatial-resolution imaging techniques like fMRI with high time-resolution techniques like magnetoencephalography (MEG), researchers can trace the path of a visual stimulus as it propagates through the brain. Using these tools, several face-specific temporal signatures have been identified within the FFA (Bentin et al., 1996; Liu et al., 2002; Bayle and Taylor, 2010). The earliest signature, designated M100, has been correlated with face detection (Liu et al., 2002) but not attentive activity (Okazaki et al., 2008), indicating that face detection is preattentive. The results of these imaging studies bolster the case presented by the psychological experiments in support of preattentive face detection.

Concerning identifying the state of facial features, evidence is mounting that certain facial expressions are processed holistically in the preattentive phase. Cognitive scientists have long studied the “face-in-crowd” affect, where a strongly emotional face stands out from a large group. While psychologists have argued both for (Hansen and Hansen, 1988; Calvo and Nummenmaa, 2008; Williams and Mattingley, 2006) and against (Purcell et al., 1996; Coelho et al., 2009) the preattentive hypothesis, the more recent neurological evidence tends to support the preattentive conclusion (Bayle and Taylor, 2010; Liu and Ioannides, 2010).

In contrast to face detection and expression, face identification has not been shown to be a preattentive process. It is intuitively obvious that face recognition is an attentive endeavor, since we all know that finding a familiar face in a large crowd is a time-consuming task. Supporting our intuition, visual-search experiments into face identification indicate serial attentive processing (Tong and Nakayama, 1999). In corroborating neurological studies, the temporal response signatures correlated with face recognition (Liu et al., 2002) are modulated by attentional affects (Okazaki et al., 2008).

Examining these results, a pattern emerges that bears weight for our discussion. While there are some aspects of face processing that appear to be preattentive, namely detection of faces and facial emotion, the outputs of the associated tasks are mainly indicational. It is as if the face-specific regions control a light that turns on whenever a face is present, or an emotion-strength gauge with a needle that peaks for angry faces in peripheral visual field. The properties of individual components of the face are not considered by the on/off nature of the light or the scale of the meter, merely the presence of a face or some abstract measure of emotion distributed across many features.The one task that most directly involves representational content, face identification, is not preattentive. That the preattentive outputs of the face-specific regions of the visual cortex appear to be indicator signals and not true representations is a major concern for Chernoff faces. Since the primary purpose of data visualization is data representation, and preattentive processing has been touted as a primary motivation for face representations of data, the inability to preattentively process representational face content would seem to be a critical deficiency.

The results of experiments conducted directly on Chernoff faces support this conclusion. Morris et al. (2000) conducted an experiment to test whether individual facial features could be discriminated in a preattentive fashion. Chernoff faces were strobed in front of a number of test subjects a rate known to preclude (most) attentive processing. Although the data represented by the faces was strictly binary in nature, human judgment of the state of the facial features was measured to be no better than random chance. The researchers concluded that individual facial features are not processed preattentively.

Barring preattentive feature processing, Chernoff faces are seemingly left in a position no better than any other glyph. As ordinary glyphs, faces would be expected to perform sub-optimally due to the low element scale range and difficulty comparing magnitudes across features. In a study conducted by Lee, Reilly, and Butavicius, faces were shown to perform poorly when compared to star glyphs and spatial visualization methods (Lee et al., 2003).

4.2.2 Research on Scene Perception

Having reviewed some of the research on the brain's face-processing abilities, we must question whether the same principles apply the other known domain-specific areas of the visual cortex, especially outdoor scenes. The purpose of these areas may be non-representational as well, possibly invalidating our pretense for their use in preattentively processing scene representations. Based on what we have learned about faces, we can make two main predictions about scenes: first, that identifying the configuration of individual scene elements is a slow attentive process, and second, that the preattentive components of scene processing produce indicational rather than representational content.

While the research on scene perception is less developed than that of face processing, there is sufficient data to improve confidence in our assumptions. Early experiments into scene perception showed that humans can perform complicated analysis tasks with a surprisingly brief stimulus. Most modern accounts of scene perception generally support the standard two-phase paradigm for visual processing, with well-defined preattentive and attentive tasks.

Aspects of scenes that are processed preattentively are scene vs. non-scene discrimination (Marois et al., 2004), basic scene-type categorization (Li et al., 2002; Grill-Spector and Kanwisher, 2005), and certain global scene attributes (Greene and Oliva, 2009; Joubert et al., 2007). Global scene attributes such as naturalness, temperature, navigability, and depth have been shown through stroboscopic psychological experiments to be identifiable in visual exposures as short as 19ms, and with a mean minimum exposure time of 34ms (Greene and Oliva, 2009). Other global attributes such as mean element size, mean object orientation, and scene pleasantness can also be identified rapidly. Basic scene-type categorization, such as identifying scenes as containing mountains, oceans, forests, and deserts occurs with exposures well under 100ms (Greene and Oliva, 2009; Evans et al., 2005).

Elements of the early visual system, in particular the PPA, play an important role in accomplishing these tasks. The PPA has been shown been shown to respond to scenes in the absence of attention (Marois et al., 2004), and is known to be involved in processing scene spatial layout (Epstein and Kanwisher, 1998).

Among attentively processed attributes are scene object position (Tatler et al., 2003; Evans et al., 2005), object spatial relationships (Tatler et al., 2003), detailed categorizations (Evans et al., 2005) and full-fledged scene representations (Tatler et al., 2003). Representational detail improves with increased exposure time as more scene elements are processed serially (Tatler et al., 2003).

In keeping with our predictions, all of the preattentive scene abilities are non-representational. Knowing that a scene appears to be hot, contains large objects, and is natural does not imply knowledge of spatial configurations or specific inter-element relations. While it may be possible to form a representation given enough such details, doing so would be a strictly attentive endeavor.

4.2.3 Discussion

Looking back at this body of research, we can critically evaluate our motivations for creating scene visualizations. The fast-processing regions of the brain that we had hoped to use for robust data representations have turned out to be producing indicator signals instead. Indicator signals, while useful in a more limited fashion, lack any shared structure with their sources. Indicator outputs are unable to undergo the transformations necessary to make strong inferences about the events that initiated them. This seems to cast a disappointing outlook for scene visualizations, especially in light of the poor performance of Chernoff faces in perceptual benchmarking studies.

While the benchmarking evidence against Chernoff faces may seem particularly troubling, there is at least one aspect where the studies may be criticized. The experiments in question did not consider directly the properties of faces that we know to be processed preattentively. Instead, the studies focused on tasks that required judgment of the state of individual facial features, an apparently attentive process. In at least one case the authors were aware of this deficiency (see Morris et al. (2000)), but did not attempt to account for it in their experimental design. We are therefore left to question whether preattentive abilities like face detection or emotion detection can be effectively used for data visualization.

To be fair, the researches involved with these studies had good reason to be reluctant to use detection or emotion in their experiments, as there are significant challenges in using either effectively. Consider the case of face detection, where the domain-specific hardware in the brain acts as a detector of an object's "faceness". To use the face-detector to encode information, we would need to modulate the "faceness" of objects, possibly by morphing between a face and some equivalent non-face glyph. There are a number of obvious implementation issues with this approach, from aesthetically dubious results to issues in visually relating face and non-face elements. Application of facial emotions to visualization is also troublesome. Emotions are correlated with individual facial features in a highly specific manner. To use emotions effectively, we would need to somehow detect correlations between variables in a dataset and map them to facial features in a plausible manner. Although there are statistical methods for detecting such correlations, they cannot guarantee that the proper configurations exist, and there are drawbacks even when they do. Morris et al. (2000) cited the difficulty in detecting and applying these correlations as justification for not doing so.

In light of these issues, scenes have at least ones superficial advantage over faces. The diversity of valid scene content would seem to provide a broader palette of global features than the range of valid human facial expressions. For datasets with known correlations, the availability of more preattentive properties should simplify implementation. In datasets with unknown correlations, the more diverse nature of scenes should make detecting unknown correlations more likely. These conclusions could be verified with user studies and perceptual benchmarking, respectively.

While non-representational, it is encouraging that preattentive scene attributes appear to reflect top-down rather than bottom-up processing. Properties of elements across the entire scene are integrated to provide a holistic overview, rather than the whole scene being constructed from the sum of small parts. This multi-element integrational approach means that the properties of many scene elements can be evaluated in parallel, hopefully facilitating rapid visual search. Using global scene attributes to integrate element features is essentially a variable-reduction technique, in a vein similar to statistical methods principal component analysis. The primary difference is that all of the variables are presented in view for direct examination, where inherent assumptions are less apt to be obscured by black-box statistical calculations.

In light of the fundamental principles presented in the previous chapter, scene visualizations can be viewed as a method of enhancing the speed of the graph comprehension process. While scene visualization may not have the best accuracy or level of detail, its fast-search nature could serve to complement a slower, more precise format, for those applications that require it.

Chapter 5

Scene Visualization Construction

Having elaborated the purpose and motivation for scene visualizations, we now confront the topic of scene construction. Most visualizations are constructed by first selecting specific visual elements, transforming those elements in relation to the data, and then assembling them to form a complete presentation. Our concept of scene construction follows a similar pattern. We begin by presenting a number of visual elements, typical to outdoor scenes, which may be useful for encoding information. We then discuss methods of integrating elements to form discrete visual representations. The properties of the resulting scene depend on the properties of the constituent elements and inter-element relationships. Since we will identify elements with both dimension-preserving and dimension-reducing properties, scene visualizations have the potential to accommodate either paradigm. We provide examples from both categories. Style choice will be dictated by the nature of the application.

5.1 Scene Elements

There are many aspects of outdoor scenes can be used as visual elements for data representation. Common scene objects, such as mountains, forests, lakes, rivers, and even buildings, all can used to encode information. Our present concern is how to use these objects to display information.

From our model of data, the types of information we hope to display are quantitative, categorical, and relational. We now discuss a number of scene elements, and the types of information they are suitable for representing[1].

[1]It is important to note that computer graphics techniques already exist for modeling and rendering all of the elements described here. Some data visualization packages even include the capability to integrate many of the elements

One important consideration is whether we want to display relationships between variables, the way one variable in our dataset varies in some interesting manner in relation to another variable. Research has shown that inter-variable relations are unlikely to be noticed in graphical representations unless at least one of the variables is represented spatially (Pinker, 1990). That is to say that the relationship between two variables is more likely to get noticed if they are plotted as position and color rather than orientation and color. This recalls our previous discussion on dimension-preserving techniques and preserving the "shape" of the variable space. If we want to see the relationship between two variables, at least one of them needs to be visualized with a relationship-preserving method.

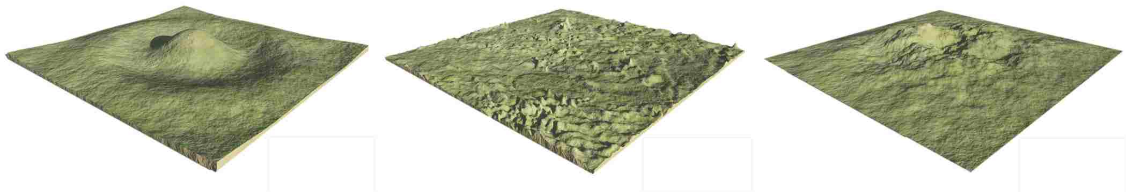
5.1.1 Element Classes

5.1.1.1 Topography

Geographic elements are useful in scene visualizations in at least three ways. First, they can be used to display dependent variables in a spatial context. Second, individual geographic elements can be used as glyphs for representing independent variables. Finally, geography is useful as a neutral backdrop, providing visual coherence for other scene elements.

In a spatial context, topography can be used to display at least three variables, two independent and one dependent. Many datasets resemble mountains, hills, and valleys when one of the variables is displayed as height when plotted against two other variables. Figure 5.1 contains examples of data plotted as topographic elements. 5.1a displays the function $z = \frac{\sin\left(\sqrt{x^2+y^2}\right)}{\sqrt{x^2+y^2}}$, and 5.1b shows the magnitude of the earth's magnetic field over a portion of North America. For comparison to standard computer-generated terrain, 5.1c shows a procedural mountain created using Vue. Since displaying data in this manner involves creating a three-dimensional surface (and therefore a four-variable construct), it may suffer some of the drawbacks of a dimension-reducing technique once the surface is projected into the scene.

we present. We make no claim to be the first to use any of the following elements to display data. We have, in fact borrowed elements from previous work whenever possible. The novel aspect of our approach is the investigation into the explicit usage of the recently uncovered domain-specific mental capabilities relating to outdoor scenes.



(a) sombrero function (b) geomagnetic field (c) procedural terrain

Figure 5.1: Topographic Examples

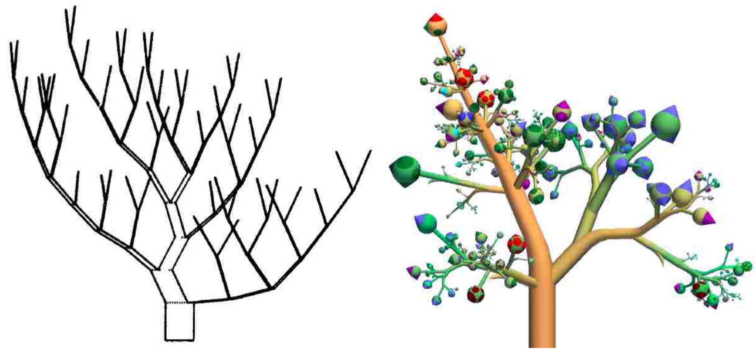
Scene topography can be used to display discrete variables by selecting individual topographic elements and manipulating their properties. geographic features can be made into glyph elements.

As a neutral backdrop, a terrain element can provide a setting in which to place other, data-representing scene elements, without representing any data itself. As will be discussed in a later section, plausibility is an important consideration during scene construction, and an outdoor scene with the ground missing hardly seems plausible.

5.1.1.2 Vegetation

Plants have many attributes that can be used to produce glyph data representations. Plant attributes can be manipulated at a low-level "micro" level or a high-level "macro" level, both methods having advantages and drawbacks.

Low-level methods construct plant visualizations from the bottom-up by manipulating basic plant structures, such as branches, leaves, twigs, and fruit to represent properties of the object dataset. By combing a number of these low-level components, a complete plant-style representation can be formed. When constructing plant from basic features, it is important to recognize that in nature the attributes of these plant structures are not uncorrelated. For example, the size of a branch farther up a tree depends on the size of the branch it connects to. If these correlations are not considered, unnatural-appearing plant structures will result. Most methods that use low-level



(a) Kleiner-Hartigan tree (b) botanical treemap

Figure 5.2: Plant Visualizations

attributes for producing botanical data representations from have a mechanism for preserving these relationships.

Kleiner-Hartigan trees (Kleiner and Hartigan, 1981) and botanical treemaps (Kleiberg et al., 2001) are examples of bottom-up plant-style data representations. Kleiner-Hartigan trees use only the smallest branches in the tree, the “leaf” nodes, to indicate variable values. All of the sub-branches are used to show the relationships between variables. The tree structure is inferred from unstructured data by means of a distance measure. The variables with the closest values will appear closest together in the tree structure, and the most unrelated branches will meet only at the bottom. The thickness of the lower branches is calculated from the number of variables “above” them, helping create a botanically plausible appearance. Figure 5.2a, taken from Kleiner’s paper, shows a tree representation of voting statistics from 48 states in the 1972 United States presidential election. Each small “twig” at the end of a “branch” is a state, and the distance metric was calculated from the proportion of Republication votes in each state.

Kleiberg’s botanical treemaps were created to represent an inherently tree-like source, the directory structure of a computer file system. Figure 5.2b, from Kleiberg’s paper, shows a representation of the root directory tree from a Unix machine.

High-level plant attributes like size, species, and color are all derived from interrelationships between various low-level attributes. Using these macro-attributes can be viewed as a top-down approach to creating plant glyphs. Starting with a nearly-complete digital description of a plant, we can alter higher-level properties without interfering with the basic structure of the plant's primitive elements. For example, given a polygon mesh describing the structure of a tree, it is possible to scale the vertices of the mesh to alter the apparent size and proportions of the plant without considering details such as branch topology or trunk length. Likewise, color-space transformations can be applied to polygon surface maps to change the color of the plant without altering the texture.

At a yet higher level, properties of groups of plants can convey useful information. Within a terrain context, variables could be expressed as the spectrum between a few isolated saplings or a dense thicket. In scene plots, one could use the seasonal range between a leafless winter tree and a full summer crown of foliage.

5.1.1.3 Weather

Weather elements can be added to an outdoor scene. Discrete weather events, such as tornadoes and lightning, are potential glyph candidates. The iconic weather map used in local television newscasts is a familiar example of weather-style glyphs. In the case of scene representations, weather-type icons could indicate variables not related to meteorology. From a dimension, reducing perspective, clouds or fog could conceivably be used for volume rendering.

5.1.1.4 Bodies of Water

Bodies of water often occupy the lowest elevation levels in a topographic area. In spatial terrain, we can use this principle as a sort of thresholding mechanism to obscure elevations that are so low as to not be of interest, or to provide a division between high and low elevation. In scene plots, water level can be used to add an extra independent variable to a scene. For example, the water level of a small stream running through the scene could be used to indicate the value of a variable.

5.2 Scene Construction

Having reviewed candidate scene elements, we can now discuss creating complete scenes. One basic scene visualization method is scene plots. Scene plots are a means of showing discrete observations taken from a multidimensional sample space. Intended as improved glyphs, they directly correspond with the motivation and purpose of Chernoff Faces. As a glyph, they fit under the category of dimension-preserving visualization methods, and assume the corresponding benefits and drawbacks. The comparative advantage of scene plots over conventional glyphs is intended to be in the exploitation of human visual system's ability to decode outdoor scenes. Scene plots may be used by themselves as stand-alone visualizations, or in collections of multiple scene instances.

By itself, a single scene can be used to show an overall situational outlook for a large system. The state of a complex system can be thought of as a point in a multivariate state space. We can map this point onto a single outdoor scene in order to achieve an integrated representation of all variables in question.

To realize a scene visualization of a single multidimensional data point, we start by creating a nominal scene description. To create the nominal scene, we pick a theme for the scene, select appropriate scene elements, then create a base element arrangement. Element selection, while constrained by the overall theme, is also guided by the dictates of global element integration and convention. Global integrational relationships are important for maximizing the potential for preattentive scene processing. Convention can influence the perception process by biasing humans towards certain interpretations. The elements we have selected for the nominal scene are then altered in proportion to the value of the data variables, forming a data-representative scene description. Using the modified scene description, a visual representation is then rendered and displayed in the display medium.

5.2.1 Scene Theme and Element Coherence

The first step in creating a nominal scene is choosing an appropriate theme, such as "tropical island," "alpine meadow," or "forest thicket". Selecting a theme is important because we want the individual scene elements to be perceptually congruent. Inter-element congruence helps avoid inconsistencies that could attract disproportionate attention to certain features. An element that draws a disproportionate level of attention is likely to influence scene interpretation, since attention

Figure 5.3: Nominal Scene

has a strong influence on comprehension of visual representations. Intuitively, it seems likely that a cactus depicted in the middle of a frozen arctic tundra would draw more attention than other, more expected objects, potentially affecting performance. Confirming our intuition, experiments have shown that objects incongruent with a scene take longer to identify (Joubert et al., 2008) and attract more attention (Underwood et al., 2008) than congruent objects. A well-chosen theme guides element selection to maintain uniform attentive consistency across the entire scene, diminishing the influence of object congruence over interpretation.

In addition to element selection, the theme also guides element placement. Elements in incongruent locations and configurations can slow detection and processing. Inverted scenes are more difficult to detect (Walther et al., 2009), as well as objects at unexpected locations within a scene (Lewis and Edmonds, 2003).

Figure 5.3 shows an example nominal scene. The theme was modeled after the terminal lakes of the North American great basin region, such as Mono Lake or the Great Salt Lake. The scene contains three plant elements, two terrain elements, and a water element.

5.2.2 Global Integration Properties

It is critical to consider how each of the chosen scene elements affects the global properties of the scene. As amply demonstrated by feature-specific experiments conducted on Chernoff faces, any preattentive advantage of scene visualizations will be lost if the individual elements do not contribute to the global perspective in a coherent manner. We can increase the chance that element properties will integrate globally by examining statistical properties of the target data set and selecting elements accordingly. If certain correlations between variables are expected or known, scene elements and their associated attribute manipulations can be assigned such that those correlations manifest themselves as a given global attribute. For example, if a number of variables are encoded as scene element sizes, a correlation between those variables is likely to manifest itself in the mean size of the elements displayed in the scene.

Global properties that are manipulable at the individual-element level include mean element size, mean element orientation, scene temperature, scene object center of mass, scene openness, scene naturalness, and many others. Center of mass could be a particularly interesting property to exploit. By placing scene elements at fixed locations where proximity is related to level of correlation, then scaling the objects according to their value, we can create a scene where the visual center of mass indicates where the datapoint sits, within a sort of abstract state-space.

5.2.3 Scene Convention

Graphical conventional is the notion that certain visual configurations, having already been connected to specific interpretations in the past, will presuppose a viewer to those interpretations when viewing those configurations in the future. While not a preattentive process, conventional interpretation can have a powerful influence over the speed and accuracy with which a visualization is comprehended (Winn, 1994). As such, conventional interpretations should be an important consideration during element selection.

One way convention can improve visualization comprehension is by aiding the application of graphical rules. Many types of visualizations have an accompanying list of rules that must be applied in order to properly interpret the graphic, often displayed in the form of a legend. Applying complicated graphical rules requires time and effort on behalf of the viewer, especially in the case where the rules are unfamiliar. Use of graphical convention has been shown to reduce or eliminate the need to consciously apply these rules, improving comprehension efficiency (Winn et al., 1991).

In the context of scenes, if we know the conventional meanings of the scene elements we have selected, we can use them to aid in comprehension and guide interpretations.

Many of the discussed scene elements are known to have conventional interpretations. It has been shown that humans have a generally positive response to plants, and further research indicates human preference for certain plant colors (Kaufman and Lohr, 2004) and types (Lohr and Pearson-Mims, 2006). Weather is known to affect human mood (Chotai et al., 2004) and could be used to add positive or negative connotations to a scene.

Interpretation of graphics has also been shown to be influenced by linguistic patterns. The direction we read text influences what parts of a graphic we associate with past and future. Most English speakers will recognize many sayings that attach emotional, quantitative, or qualitative meaning to outdoor scene elements. Examples include phrases such as “don’t rain on my parade,” “she has a mountain of work to do,” or “he’s underwater in his mortgage.” We expect that scene elements are associated with conventional interpretations in many other ways not yet identified in the formal scientific literature.

5.2.4 Altering Scene Elements to Represent Data

There are many elements that can appear in outdoors, many ways of combining those elements to create a scene, and many ways of altering those elements to represent data. However, not all things that could potentially appear in an outdoor scene are equally effective for the purpose of data visualization. Many issues come in to play when altering elements in proportion to data values, and some of them are less obvious than others. Compounding factors include perceptual linearity, dynamic range, and cross-modulation. Before altering the properties of nominal scene elements, it is important to plan ahead to minimize adverse effects on scene data representations. How the element is altered is often just as important as what the element is.

In the absence of a formal theory of data visualization it is difficult to describe exactly what we mean when we say “perceptual linearity,” but in this work we will use the following definition: A scene element is perceptually self-linear if our interpretation of the element’s value varies in proportion to the variable it represents. For example, in a perceptually linear visualization, if scene A uses the height of a tree to indicate value x, the perceived height of that same tree in scene B indicating the value 2x will be exactly twice that of the tree in scene A. While linearity is not a requirement for every visualization, if we can linearize a given element, we can guarantee that a non-linear representation will be perceived in the way we expect it to.

Perceptual non-linearity can occur in a variety of ways. If an element is scaled so that its borders do not fit within the window of the presentation, the viewer will be left to guess the size of the element using less information than from objects that fit completely in view. Linearly scaling all of an object's dimensions will produce a quadratic increase in the visible area occupied by the object. Some non-linear effects occur within the visual system itself. Many of these properties are commonly used to create optical illusions. To understand and avoid such problems, we refer the reader to any good book on human vision, such as (Snowden et al., 2006).

Scene element cross-modulation is another linearity issue. Cross-modulation occurs when changing the value of one scene element affects the perceived value of another element. A simple example of cross-modulation is occlusion, where one element overlaps the area of another element, affecting the apparent size of that element. Element cross-modulation can manifest itself in unexpected ways. For example, suppose we are using a bush to represent two variables, where the size of the bush represents variable A and the color of the bush represents variable B. If a computer monitor is used to display the visualization, we may find that if we diminish the size of the bush, the color appears to darken as well. Surprisingly, object scale and perceived color are not always independent in computer imaging (Cook et al., 2007). This happens because as the size of the object decreases, the relative level of detail per unit of screen area increases. The highlights in the image become averaged with the shadows on a per-pixel basis, resulting in decreased contrast levels and a dull appearance. In this situation, it may be necessary to linearize the perceived color by boosting contrast while diminishing size.

Encompassing these linearity issues is relational dynamic range. In signal processing, system dynamic range is often defined as the ratio of the largest undistorted signal to the smallest signal distinguishable from noise. We are concerned about a very similar quantity in our visualizations. Since not all scene elements are definable in terms of big and small (is red bigger than blue?), we can also define dynamic range in terms of the ratio of the full linear scale to the smallest distinguishable gradation. If we want to preserve all of the important information in a variable, we must ensure that the dynamic range of the associated vehicle element is sufficient for the task.

We can maximize the available dynamic range by carefully choosing how we scale the selected scene element. In some situations, this may be as simple as scanning the dataset to find the maximum value, determining the maximum linear extent of the element, and linearly scaling all datapoints to fit in between. However, if a variable has significant information spread across very high and very low values, the dynamic range may exceed what is representable with the chosen element. In this case, it may be necessary to sacrifice linearity by scaling the element with a formula that stretches dynamic range while preserving some important aspect of the data. Logarithmic scaling, which preserves monotonicity, is a common scaling method.

Since there is an innumerable number of ways that altering scene elements can have unintended consequences, their presence will have to be considered on a per-application basis. It may prove to be impossible to eliminate all non-linear effects in scene perception, as many are contingent on the properties of the human visual system, and every individual is different. The best we can do is to be aware of them and adapt as appropriate.

Figure 5.4 shows scenes based on the nominal scene in figure 5.3, with all elements altered in proportion to a common scale factor. The sizes of three plant elements and the background terrain element are scaled down from their proportions in the nominal scene, and the hue of the plant elements is altered in HSV color space. The foreground terrain and water elements are retained as scene backdrops. Figure 5.4a shows all elements at their respective maximums, 5.4b shows them at mid-scale, and in 5.4c they are near their practical minima.

(a) Maximum (b) Mid-Scale (c) Minimum

Figure 5.4: Scene Element Scaling

5.2.4.1 Rendering the Scene

The final step in creating a scene visualization is rendering. Rendering can be done schematically or photorealistically, using raster of vector methods, and output can be via monitor or paper. There is a large amount of literature on rendering, and we will not revisit it here. For this study, we have chosen a moderately realistic rendering with three-dimensional scenes displayed mainly on computer monitors. We have implemented the majority of the original figures for this project using python scripts within the Vue software environment. While not marketed as a data visualization package, Vue is a powerful outdoor scene modeling and rendering tool that incorporates all of the visual elements discussed here.

5.2.5 Multiple Scene Plots

While we know it is possible to encode any machine-representable dataset as a single point (Godel, 1992; Homer and Selman, 2001), there are situations where representing all points within a single scene is a disadvantage. Consider the case where the dataset is primarily categorical, such as a collection of responses from a multiple-choice survey. We can represent categorical data with categorical scene elements, but we face a problem when placing many such elements within a single scene. By placing a large amount of categorical information within a spatial context, we risk implying spatial relationships where none exist.

For situations where a single scene is inappropriate, scene arrays can be used to represent discrete points from a dataset. Each point can be mapped to a different scene, with multiple points forming a collection of scenes, allowing comparison of discrete datapoints. The visual boundaries within the array serve to break up spatial continuity between the discrete observations. Figure 5.5 shows a collection of scene plots representing various attributes of automobiles.

5.3 Landscape Plots

While glyphs are useful for representing individual datapoints, a drawback is that they are ineffective at displaying relationships between datapoints across multiple variables. For datasets where such relationships are important, it is necessary to look beyond glyphs for an effective representation. To be effective, we need to select scene components that have attributes better suited for displaying inter-variable relationships. As discussed in Chapter 3, relationships between variables are unlikely to be noticed visually, unless at least one of the variables is represented spatially. To represent variables spatially in a scene visualization, we need to use scene elements with a spatial component. The scene elements that come closest to meeting this criteria are topography and weather. Of the two, topography seems to offer the simplest path to implementation.

While not displaying any dependent variables by itself terrain, can still provide a way of indicating relationships between other datapoints by providing spatial context for additional scene elements. For example, the apparent position of a non-terrain scene element within the terrain could be used to indicate additional element variables.

The most basic topographic element possible is the plain. While a plain may seem to be too simple to display useful information, it can be a valuable construct for our purposes. A plain can be used as the reference surface for allowing glyph-style elements to be compared on a spatial basis. That is to say, while not containing any dependent variables by itself, a plain can indicate two independent variables against which to plot another dependent scene element, by including spatial location among the element's attributes. The plain allows a viewer to estimate the represented values by judging the element's relative position. By applying datapoints to many spatially-mapped vegetation elements, a "forest" can be created, a sort of scatter plot made of scene elements.

While a plain is useful by itself, more varied terrain offers opportunities for greater variable density. One way to do this is described in the section on terrain elements, where one of the data

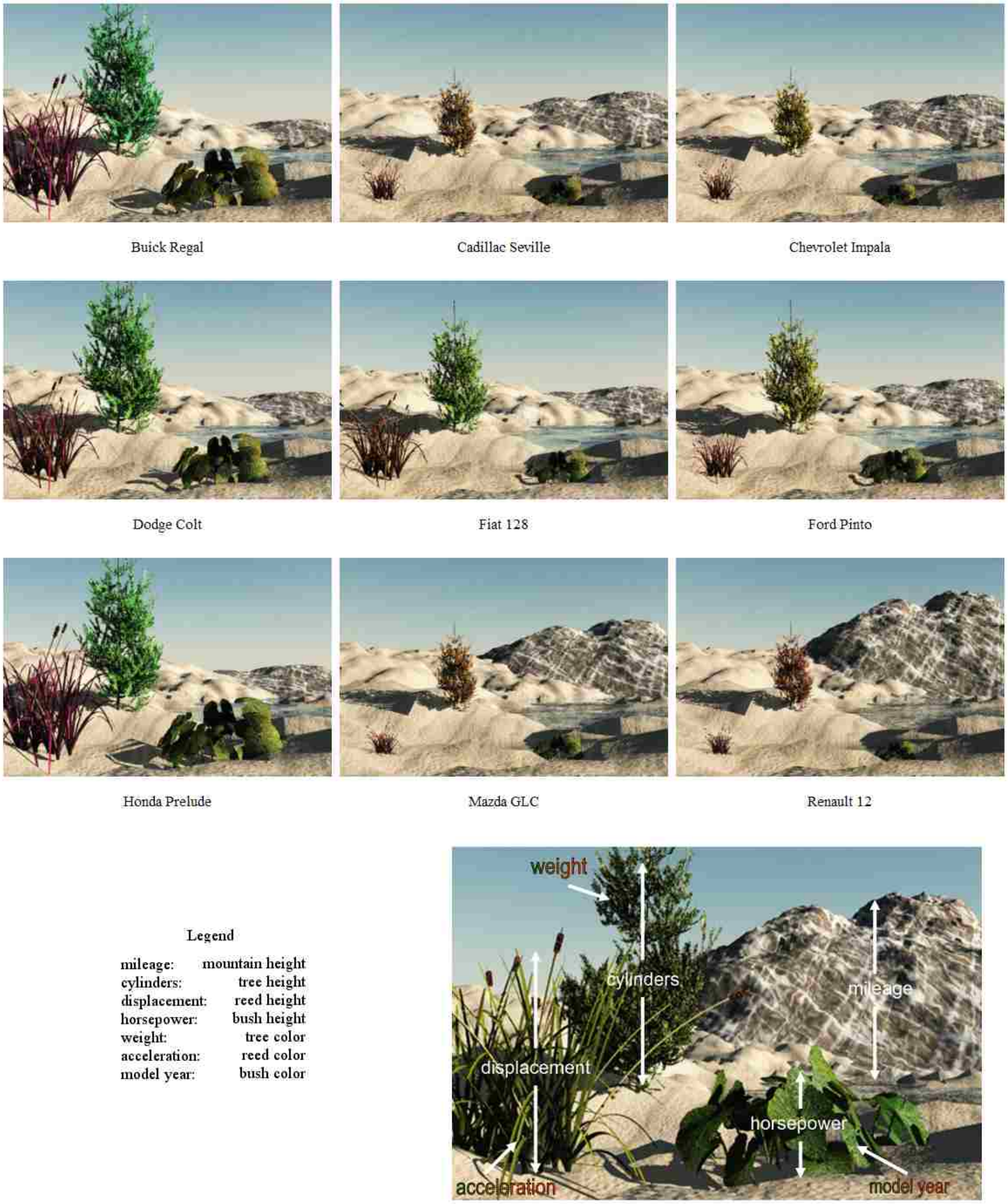


Figure 5.5: Scene Plots: Cars

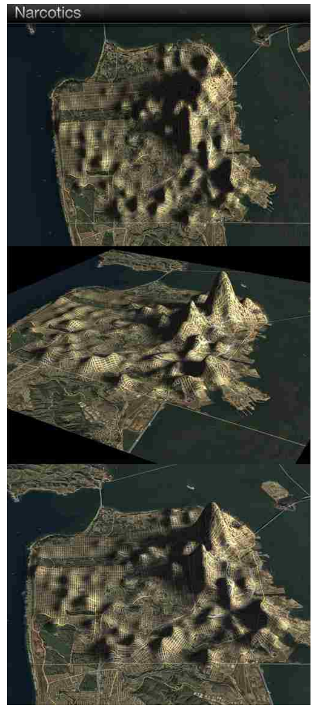


Figure 5.6: Data Landscape: San Francisco Narcotics

variables is applied to the terrain's surface elevation. To make a plausible-looking terrain, heights will have to be scaled to an appropriate level and some form of interpolation may be required to fill in sparsely populated regions.

Another addition is to map an image onto to the topography element. Figure 5.6[2] shows a terrain plot of narcotics crime in San Francisco overlaid with a map of the city. The plot is remarkable in that it allows the viewer to rapidly correlate high crime areas with street-level geography within the same presentation. Application of the information in the map is similarly intuitive. For example, to avoid crime-ridden areas, one needs simply to head downhill.

In this application the image overlay is used primarily for position referencing, allowing the viewer to better estimate the value of position variables over uneven terrain. Its purpose is therefore most accurately stated as one of disambiguation, as opposed to representing truly unique information. To represent another variable, a common method is to create an image through tone mapping, where the variable's values are mapped to an appropriate intensity scale, then referenced

[2] Image courtesy Doug McCune. Used with permission.

against the two corresponding spatially-represented variables. More sophisticated methods for representing multiple variables are also possible (van Wijk, 1991).

5.4 Discussion

Our initial investigation suggests great diversity in the ways that scenes can be used as representational vehicles. We have shown single-point, multipoint, and spatially coherent scene construction paradigms. It can be expected that there are numerous other useful combinations of scene elements.

The greatest concern with scene construction appears to be representational linearity. The multitude of self-linearity and cross-modulation issues will require careful consideration during implementation. Unskilled treatment of the method can produce misleading results. Until an automated system for resolving these problems is devised, the required per-application effort may limit the general use of scene visualization.

Chapter 6

Application and Comparison

With an understanding of the basic concepts that guide scene construction, the discussion now turns to more application-specific issues. We start by describing the types of situations where scene visualizations are likely to have the most impact. We then examine one of the finest examples of Chernoff faces, and compare it to an equivalent scene representation. Finally, we compare scene visualization to surface plots, a method which bears some interesting similarities.

6.1 Applications

The strong points of scene visualizations are preattentive processing of global scene attributes and conventional interpretation of common scene elements. Applications that can benefit from these strengths will benefit most from the technique. A dataset ideally suited for scene plots will have pointwise inter-variable correlations, variable categories that map well to conventional scene interpretations, and variables with limited or compressible dynamic range. In the case of data landscapes, it is also useful for the dataset to have two independent variables against which to plot the others. For scene plots, the size of the dataset should be limited to less than a few hundred points. Candidate datasets include geographical studies, environmental reports, and demographics. Data landscapes are useful when the dataset has two independent variables and one or more dependent variables.

6.2 Comparison to Other Techniques

In the absence of formal theory, visualization performance measurement is as much art as science. The current gold standard seems to be perceptual benchmarking, and while such tests may seem comfortingly objective, the lack of a theoretical standard means that results can be influenced though experimental design. Human performance testing is also expensive and time-consuming, making the cost difficult to justify for a preliminary investigation.

Many authors resort to direct subjective comparison with other techniques, using side-by side examples in the hopes that the results will be obvious to a reasonable human observer. This work follows a similar pattern. To judge the qualities of our hypothetical visualization method, we compare two contrasting examples of outdoor scene visualizations side-by-side with respectively similar techniques. The comparative visualizations were chosen for their alternate similarity in both intended purpose and in substantive practice in relation to scene visualizations. We start by looking at Chernoff faces, which were created with similar goals to scene visualizations. We then examine surface plots, a technique that tacitly exploit principles similar to scene visualizations. Using the knowledge gleaned from the comparisons, we discuss our success relative to our goals.

6.2.1 Chernoff Faces

Among existing visualization techniques, Chernoff faces is the most similar to outdoor scenes, and have been a focus of scrutiny throughout this work. For comparative purposes, let us examine one of the most noteworthy applications of Chernoff faces, Eugine Turner's "Life in Los Angeles" map (figure 6.1a). Turner uses Chernoff faces to denote the demographic makeup of the various regions in the city. As such, the map incorporates a variety of strategies that maximize the potential of face visualizations. For example, quantitative variables like unemployment rate have been reduced to categorical ones by sorting them into low, medium, and high categories, alleviating dynamic range issues. In addition, the number of faces has been kept low by the coarse nature of the city subdivisions. Most importantly, the map uses correlations in the data to create facial expressions that influence human interpretation of the data. Theses factors combine for a highly effective presentation. The meaning of the map is so intuitively clear that the provided legend is almost unnecessary.

Despite the efficiency of Turner's presentation, it can be difficult to apply the same techniques in the general case. Not all variables can be categorized so easily, nor can all datasets be made to have so few data points. A greater issue is the strong data correlations needed to create coherent facial expressions. Not all data can be assigned easily to an emotion, nor is it always desirable to do so.

To evaluate whether outdoor scenes provide a viable alternative to Chernoff faces, we have re-created the map as an outdoor scene (figure 6.1b). The map of Los Angeles is shown covered in a forest of trees instead of a crowd of face glyphs. Affluence is represented by tree height, unemployment rate by vegetation density, urban stress by tree color, and ethnic makeup by tree species[1].

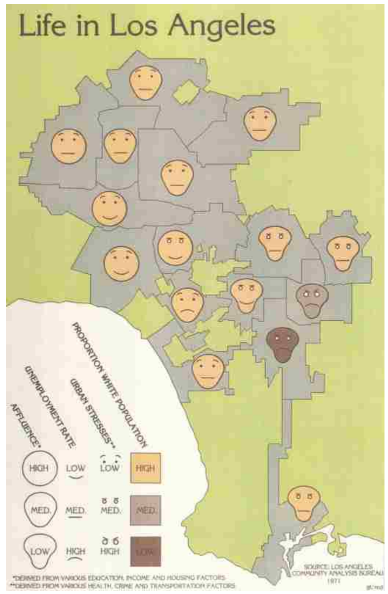
Like the Turner map, global integration and convention has been used to influence interpretation. A tall, dense, green forest signifies the most desirable locations in the city, whereas a sparse, stunted, dry area signals disadvantaged regions. The intuitive nature of the map is retained without the use of faces, and variable assignment is greatly simplified.

6.2.2 Surface Plots

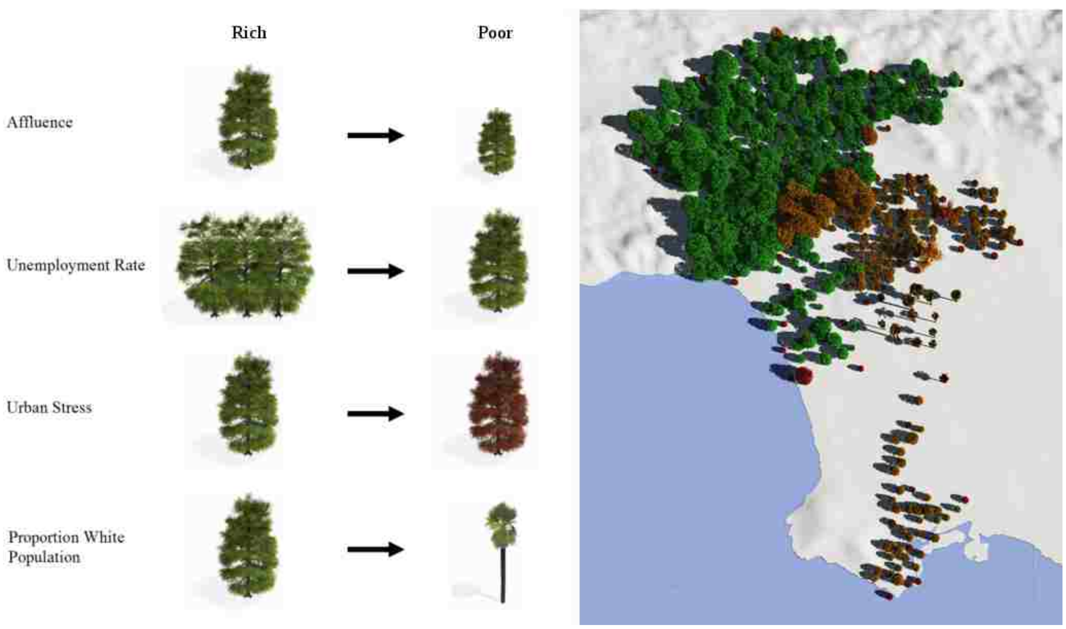
Many types of graphs were in use prior to the modern era of cognitive science. Line graphs existed for centuries before the orientation-responsive areas of the visual cortex were discovered. The creators of these graphs did not know the precise details about why we can interpret such graphics, they just knew that they worked. They applied their tacit knowledge of human perception to make to make effective simple, effective visualizations.

If graph-makers were using orientation-detecting cells centuries ago, perhaps an analogous situation exists for scene visualizations. It is possible that graphics utilizing the same principles as

[1] Species identification can be difficult for small trees. Shadows provide a valuable cue in this situation.



(a) Eugene Turner's "Life in Los Angeles"



(b) "Life in Los Angeles" as an Outdoor Scene

Figure 6.1: Los Angeles Map Comparison

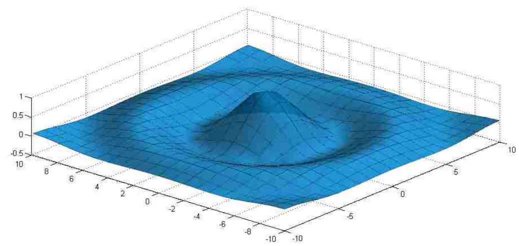


Figure 6.2: Surface Plot

scene visualizations have existed for some time, having been created under a tacit understanding of the human visual system.

One possible tacit analog to outdoor scene visualization is surface plots. Surface plots are a common way of displaying functions of two variables. Figure 6.2 shows a surface plot of the sombrero function, used previously as a terrain example.

Like the retina, the functions represented by surface plots have two independent variables and one dependent variable. These functions therefore could be represented easily and unambiguously as a direct intensity-level map onto the retina. Paradoxically, the process of creating a surface plot is considerably more complicated. Creating a surface plot involves representing the dependent variable as the height of a three-dimensional surface when plotted against the two independent variables. To do this, the data typically is converted into a polygon mesh whose vertices are then rotated, scaled, translated, and projected onto the two-dimensional viewing surface. To form a wire frame visual, the vertices are then connected by drawing lines between them. To form a solid surface, the polygons must then be shaded by calculating the polygon surface normals and applying a lighting model.

The elaborate nature of surface plot construction raises some interesting questions. Why expend so much effort to create a three-dimensional surface, and risk information loss through a dimension-reducing projection? The answer may be that surface plots have been created in a way that exploits the human visual system's ability to process the kind spatial relationships commonly found in outdoor scenes. It is at least plausible that the same areas of the brain that help process images of mountains and valleys also work to interpret the peaks and grooves in a surface plot. Such a conclusion could be verified readily with functional magnetic resonance imaging (fMRI).

If this is true, it raises questions about the motivation for using naturalistic scenes in a visualization setting at all. It may be possible to replace all of the outdoor-themed symbols we have introduced with analogs that are more efficient and have more predictable interpretations. If we knew exactly how the brain processes outdoor scenes, we might eliminate extraneous naturalistic elements and produce a cleaner presentation. Our synthetic scene plot might be more precise and more comprehensible than a natural scene, and may not look much like an outdoor scene at all. Outdoor scenes might be an intermediate hurdle in the way of some other, more effective visualization method.

That said, our choice of the outdoor scene context was not made arbitrarily. Outdoor scenes elicited the strongest response in the fMRI domain-specificity study (Downing et al., 2006), and the naturalness of a visual scene is the fastest-detectable global scene property (Greene and Oliva, 2009). There is something fundamental to outdoor scenes that triggers a response very early in the visual system. While it is presently unclear exactly what about outdoor scenes elicits this response, it is possible that removing naturalistic elements could somehow affect cognitive performance. At the very least, the outdoor scene context provides a convenient mechanism for maintaining scene congruence.

6.3 Discussion

Having experimented with outdoor scenes for data visualization, we can assess the degree of accuracy in our initial predictions. We anticipated that scene visualizations would have increased information density, simpler implementation, intuitive interpretation, and decreased attentive processing when compared to other techniques.

Relative to our prediction of increased information density, results are mixed. The feature-integrational nature of scene perception makes it possible to use very large numbers of objects to display data variables, each object contributing to the overall interpretation. On the other hand, an individual tree glyph will likely have less dynamic range than an equivalent schematic glyph.It is essentially a quantity-accuracy tradeoff, meaning that the provided level of information is contingent on the metric most relevant to the application.

Implementation difficulty is likewise a mixed bag. Scene element correlations are certainly easier to manage than facial features, but irregular element geometry complicates scene construction, and naturalistic elements can be prone to cross-modulation issues. The effect is readily visible in the Los Angeles map, where all of the palm trees are short and therefore more difficult to identify than than other tree species. In this case, the value of the affluence variable is altering detection of the ethnicity variable. It seems likely that a scene composed of schematic glyphs could be created with fewer linearity problems.

On the count of intuitive interpretation, our approach has significant promise. Many elements of outdoor scenes can be shown to have conventional interpretations useful for reducing conscious effort. An untrained viewer of the Los Angeles forest map would likely to be able identify specific scene traits, such as the most affluent regions, without difficulty. In addition, conventional interpretations of outdoor scenes appear to be more adaptable to the general case than Chernoff faces, where individual elements have strong interdependencies, and edge cases create visually incongruent results.

On the prediction of using the brain's innate capabilities to speed comprehension, we have another mixed result. The scene-specific regions of the brain do not appear to produce robust representations with preattentive efficiency. Despite this, the feature-integrational aspects should facilitate rapid visual search within multiple-scene presentations.

Chapter 7

Conclusion and Future Work

7.1 Summary

We have investigated the use of outdoor scenes for data visualization. We considered the nature of mental representation and the principles of effective data visualization. We elaborated our motivation and cognitive basis for creating scene visualizations. We introduced scene construction techniques, including scene elements, scene plots for discrete points and data landscapes for spatial information. We showed concept images demonstrating the technique, and compared them to contemporary methods.

7.2 Conclusion

Faster comprehension, higher information density, intuitive interpretation, and simpler implementation were the predictions that motivated this study. Implementation is simpler for scenes than for Chernoff faces, but more difficult than for schematic glyphs. Increasing the element count introduces a quantity/quality tradeoff that limits the maximum information density. Highly intuitive presentations are possible, and preattentive feature integration enables rapid visual search.

The most significant drawback to scene visualization is linearity. Without a specific formula for resolving linearity problems, scenes must be crafted on an application-to-application basis. The effort required to balance linearity issues may tend to limit the application of scene representations.

Despite the varied accuracy of our initial predictions, we find that the prospects for scene visualization are positive. Intuitive interpretation and preattentive analysis of scene properties are benefits that cannot be discounted. The "Life in Los Angeles" map demonstrates that the results can justify the time invested in scene construction. The key is recognizing those applications where the strengths of the technique complement the purpose of the visualization.

7.3 Future Work

7.3.1 Practical Implementation

The system we used to create the proofs-of-concepts used scenes hard-coded into Python scripts running on top of a proprietary software package using data read from a non-standard file format. This implementation is inadequate for anything but the most technical purposes. More intensive exploration of the scene plot concept will require a more sophisticated implementation. A more efficient scene creation and manipulation interface will be especially critical for producing visualizations of more varied data types.

7.3.2 Perceptual Benchmarking of Feature-Integration and Visual Search

While supported by neurological and cognitive evidence, our claims of enhanced visual search through preattentive scene feature integration remain untested. An objective evaluation through perceptual benchmarking is warranted. To avoid the pitfalls encountered with similar studies, careful consideration will be required to ensure that elements used in test scenes contribute to global scene attributes known to be preattentively processed.

Bibliography

ANDERSON, E. 1960. A semigraphical method for the analysis of complex problems. *Technometrics 2,* 3, 387–391.

BAYLE, D. J. AND TAYLOR, M. J. 2010. Attention inhibition of early cortical activation to fearful faces. *Brain Research 1313*, 113–123.

BENTIN, S., ALLISON, T., PUCE, A., PEREZ, E., AND MCCARTHY, G. 1996. Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience 8,* 6, 551–565.

BROWN, V., HUEY, D., AND FINDLAY, J. M. 1997. Face detection in peripheral vision: Do faces pop out? *Perception 26,* 12, 1555–1570.

CALVO, M. G. AND NUMMENMAA, L. 2008. Detection of emotional faces: Salient physical features guide effective visual search. *Journal of Experimental Psychology: General 137,* 3, 471–494.

CHAMBERS, J. M., CLEVELAND, W. S., KLEINER, B., AND TUKEY, P. A. 1983. *Graphical Methods for Data Analysis*. Chapman and Hall, New York.

CHERNOFF, H. 1973. Using faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association 68*, 361–368.

CHOTAI, J., SMEDH, K., JOHANSSON, C., NILSSON, L., AND ADOLFSSON, R. 2004. An epidemiological study on gender differences in self-reported seasonal changes in mood and behaviour in a general population of northern sweden. *Nordic Journal of Psychiatry 58,* 6, 429–437.

CLEVELAND, W. S. 1993. A model for studying display methods of statistical graphics. *Journal of Computational and Graphical Statistics 2*, 323–343.

CLUFF, E., BURTON, R. P., AND BARRET, W. A. 1991. A survey and characterization of multidimensional presentation techniques. *Journal of Imaging Technology 17,* 4, 142–153.

COELHO, C. M., CLOETE, S., AND WALLIS, G. 2009. The face-in-the-crowd effect: When angry faces are just cross(es). *Journal of Vision 10,* 1, 1–14.

COOK, R. L., HALSTEAD, J., PLANCK, M., AND RYU, D. 2007. Stochastic simplification of aggregate detail. *ACM Transactions on Graphics 26,* 3, 79.

CUMMINS, R. E. AND POIRIER, P. 2004. Representation and indication. In *Representation in Mind*, H. Clapin, Ed. Elsevier, 21–40.

DOWNING, P. E., CHAN, A. W.-Y., PEELEN, M. V., DODDS, C. M., AND KANWISHER, N. 2006. Domain specificity in visual cortex. *Cerebral Cortex Advance Access 16*, 1453–1461.

DREBIN, R. A., CARPENTER, L., AND HANRAHAN, P. 1988. Volume rendering. *SIGGRAPH Computer Graphics 22*, 65–74.

ECKARDT, B. V. 1993. *What is Cognitive Science?* MIT Press.

EPSTEIN, R. AND KANWISHER, N. 1998. A cortical representation of the local visual environment. *Nature 392,* 6676, 598–601.

EVANS, K. K., TREISMAN, A., LI, F., VANRULLEN, R., KOCH, C., AND PERONA, P. 2005. Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance 31*, 1476–1492.

GODEL, K. 1992. *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. Dover Publications. Translation B. Meltzer.

GREENE, M. R. AND OLIVA, A. 2009. Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology 58,* 2, 137–176.

GRILL-SPECTOR, K. AND KANWISHER, N. 2005. Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science 16,* 2, 152–160.

HADJIKHANI, N., LIU, A. K., DALE, A. M., CAVANAGH, P., AND TOOTELL, R. B. 1998. Retinotopy and color sensitivity in human visual cortical area V8. *Nature Neuroscience 1,* 3, 235–241.

HANSEN, C. H. AND HANSEN, R. D. 1988. Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology 54,* 6, 917–924.

HERSHLER, O. AND HOCHSTEIN, S. 2005. At first sight: A high-level pop out effect for faces. *Vision Research 45,* 13, 1707–1724.

HERSHLER, O. AND HOCHSTEIN, S. 2006. With a careful look: Still no low-level confound to face pop-out. *Vision Research 46,* 18, 3028–3035.

HOMER, S. AND SELMAN, A. L. 2001. *Computability and Complexity Theory*. Texts in Computer Science. Springer.

HUBEL, D. H. AND WIESEL, T. N. 1959. Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology 148*, 574–591.

INSELBERG, A. 1985. The plane with parallel coordinates. *The Visual Computer 1,* 2, 69–91.

JACOBY, W. G. 1998. Statistical graphics for visualizing multivariate data. In *Sage University Papers Series on Quantitative Applications in the Social Sciences 07-120*. Sage Publications.

JOUBERT, O. R., FIZE, D., ROUSSELET, G. A., AND FABRE-THORPE, M. 2008. Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *Journal of Vision 8,* 13, 11.1–11.18.

JOUBERT, O. R., ROUSSELET, G. A., FIZE, D., AND FABRE-THORPE, M. 2007. Processing scene context: Fast categorization and object interference. *Vision Research 47,* 26, 3286–3297.

KANDOGAN, E. 2001. Visualizing multi-dimensional clusters, trends, and outliers using star coordinates. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'01)*. ACM, New York, NY, USA, 107–116.

KAUFMAN, A. AND LOHR, V. 2004. Does plant color affect emotional and physiological responses to landscapes? In *XXVI International Horticultural Congress*. Vol. 639. Acta Horticulturae, 229–233.

KAUFMAN, L. AND ROUSSEEUW, P. J. 2005. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series on Probability and Statistics. Wiley-Interscience.

KLEIBERG, E., VAN DE WETERING, H., AND VAN WIJK, J. J. 2001. Botanical visualization of huge hierarchies. In *Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)*. IEEE Computer Society, Washington, DC, USA, 87–94.

KLEINER, B. AND HARTIGAN, J. A. 1981. Representing points in many dimensions by trees and castles. *Journal of the American Statistical Association 76*, 260–276.

LABERGE, D. 1983. Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance 9,* 3, 371–379.

LEE, M. D., REILLY, R. E., AND BUTAVICIUS, M. A. 2003. An empirical evaluation of Chernoff faces, star glyphs, and spatial visualisations for binary data. In *Australian Symposium on*

Information Visualisation, (invis.au'03), T. Pattison and B. Thomas, Eds. CRPIT Series, vol. 24. ACS, Adelaide, Australia, 1–10.

LEVOY, M. 1988. Display of surfaces from volume data. *IEEE Computer Graphics and Applications 8*, 29–37.

LEWIS, M. B. AND EDMONDS, A. J. 2003. Face detection: Mapping human performance. *Perception 32,* 8, 903–920.

LI, F. F., VANRULLEN, R., KOCH, C., AND PERONA, P. 2002. Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America 99,* 14, 9596–9601.

LIU, J., HARRIS, A., AND KANWISHER, N. 2002. Stages of processing in face perception: An MEG study. *Nature Neuroscience 5,* 9, 910–916.

LIU, L. AND IOANNIDES, A. A. 2010. Emotion separation is completed early and it depends on visual field presentation. *PLoS ONE 5,* 3, e9790.

LOHR, V. AND PEARSON-MIMS, C. 2006. Responses to scenes with spreading, rounded, and conical tree forms. *Environment and Behavior 38,* 5, 667–688.

MAROIS, R., YI, D.-J., AND CHUN, M. M. 2004. The neural fate of consciously perceived and missed events in the attentional blink. *Neuron 41,* 3, 465–472.

MCNAMARA, T. P. 1999. Single-code versus multiple-code theories in cognition. In *The Nature of Cognition*, R. J. Sternberg, Ed. The MIT Press, 114–135.

MELVILLE, J. D. AND BURTON, R. P. 1997. Piles for hyperdimensional graphics. *Computers and Graphics 21,* 1, 51–60.

MORI, M. 1970. The uncanny valley. *Energy 7,* 4, 33–35.

MORRIS, C. J., EBERT, D. S., AND RHEINGANS, P. 2000. An experimental analysis of the effectiveness of features in Chernoff faces. In *28th AIPR Workshop: 3D Visualization for Data Exploration and Decision Making, Proceedings of SPIE*. 12–17.

MOVSHON, J. AND NEWSOME, W. 1992. Neural foundations of visual motion perception. *Current Directions in Psychological Science 1,* 1, 35–39.

NOTHDURFT, H. C. 1993. Faces and facial expressions do not pop out. *Perception 22,* 11, 1287–1298.

NOWELL, L. T. 2011. The science of visual analysis at extreme scale. P. C. Wong, J. Park, M. C. Hao, C. Chen, K. Borner, D. L. Kao, and J. C. Roberts, Eds. *Visualization and Data Analysis (VDA'11) 7868,* 1, 786802.

O'BRIEN, G. AND OPIE, J. 2004. Notes toward a structuralist theory of mental representation. In *Representation in Mind*, H. Clapin, Ed. Elsevier.

OKAZAKI, Y., ABRAHAMYAN, A., STEVENS, C., AND IOANNIDES, A. 2008. The timing of face selectivity and attentional modulation in visual processing. *Neuroscience 152,* 4, 1130–1144.

PERRY, E. AND DONATH, J. S. 2004. Anthropomorphic visualization: A new approach for depicting participants in online spaces. In *ACM Conference on Human Factors in Computing Systems Extended Abstracts*, E. Dykstra-Erickson and M. Tscheligi, Eds. ACM, 1115–1118.

PETTIGREW, J. D. 1972. The neurophysiology of binocular vision. *Scientific American 227,* 2, 84–95.

PINKER, S. 1990. A theory of graph comprehension. In *Artificial Intelligence and the Future of Testing*, R. Feedle, Ed. Erlbaum, 73–126.

PURCELL, D. G., STEWART, A. L., AND SKOV, R. B. 1996. It takes a confounded face to pop out of a crowd. *Perception 25,* 9, 1091–1108.

SABELLA, P. 1988. A rendering algorithm for visualizing 3d scalar fields. *SIGGRAPH Computer Graphics 22*, 51–58.

SCHNEIDER, J. AND WESTERMANN, R. 2003. Compression domain volume rendering. In *Proceedings of the IEEE Visualization 2003 Conference (VIS'03)*. IEEE Computer Society, Washington, DC, USA, 293–300.

SNOWDEN, R., THOMPSON, P., AND TROSCIANKO, T. 2006. *Basic Vision*. Oxford University Press, New York.

TAN, P.-N., STEINBACH, M., AND KUMAR, V. 2005. *Introduction to Data Mining* US Ed. Addison Wesley.

TATLER, B. W., GILCHRIST, I. D., AND RUSTED, J. 2003. The time course of abstract visual representation. *Perception 32,* 5, 579–592.

TONG, F. AND NAKAYAMA, K. 1999. Robust representations for faces: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance 25,* 4, 1016–1035.

TREISMAN, A. M. AND GELADE, G. 1980. A feature-integration theory of attention. *Cognitive Psychology 12,* 1, 97–136.

TUKEY, J. W. AND TUKEY, P. A. 1981. Graphical display of data sets in 3 or more dimensions. In *Interpreting Multivariate Data*. John Wiley & Sons, Chichester, U.K., 189–275.

UNDERWOOD, G., TEMPLEMAN, E., LAMMING, L., AND FOULSHAM, T. 2008. Is attention necessary for object identification? Evidence from eye movements during the inspection of real-world scenes. *Consciousness and Cognition 17,* 1, 159–170.

UPSON, C. AND KEELER, M. 1988. V-buffer: Visible volume rendering. *SIGGRAPH Computer Graphics 22*, 59–64.

VAN WIJK, J. J. 1991. Spot noise texture synthesis for data visualization. In *18th Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 309–318.

VANRULLEN, R. 2006. On second glance: Still no high-level pop-out effect for faces. *Vision Research 46*, 3027.

WALTHER, D. B., CADDIGAN, E., FEI-FEI, L., AND BECK, D. M. 2009. Natural scene categories revealed in distributed patterns of activity in the human brain. *Journal of Neuroscience 29,* 34, 10573–10581.

WARD, M. 2002. A taxonomy of glyph placement strategies for multidimensional data visualization. *Information Visualization 1,* 3, 194–210.

WARD, M. AND THEROUX, K. 1997. Perceptual benchmarking for multivariate data visualization. *Scientific Visualization Conference*, 314–321.

WILLIAMS, M. A. AND MATTINGLEY, J. B. 2006. Do angry men get noticed? *Current Biology 16,* 11, R402–R404.

WINN, W. 1994. Contributions of perceptual and cognitive processes to the comprehension of graphics. In *Comprehension of Graphics*, W. Schnotz and R. Kulhavy, Eds. Elsevier, 3–27.

WINN, W., LI, T.-Z., AND SCHILL, D. 1991. Diagrams as aids to problem solving: Their role in facilitating search and computation. *Educational Technology Research and Development 39*, 17–29.

XIAOXI, T. AND BURTON, R. P. 1992. Parallel planes coordinates for hyperdimensional computer graphics. *Journal of Imaging Science and Technology 36,* 6, 591–595.

XUN, P. J. AND BURTON, R. P. 1997. Parallel volumes coordinates for hyperdimensional computer graphics. *Journal of Imaging Science and Technology 41,* 1, 59–70.