Building a perceptual visualisation architecture

Christopher G. Healey
Computer Science Department, North Carolina State University

Abstract. Scientific datasets are often difficult to analyse or visualise, due to their large size and high dimensionality. We propose a multistep approach to address this problem. We begin by using data management techniques to identify areas of interest within the dataset. This allows us to reduce a dataset's size and dimensionality, and to estimate missing values or correct erroneous entries. We display the results using visualisation techniques based on perceptual rules. Our visualisation tools are designed to exploit the power of the low-level human visual system. The result is a set of displays that allow users to perform rapid and accurate exploratory data analysis.

In order to demonstrate our techniques, we visualised an environmental dataset being used to model salmon growth and migration patterns. Data mining was used to identify significant attributes and to provide accurate estimates of plankton density. We used colour and texture to visualise the significant attributes and estimated plankton densities for each month for the years 1956 to 1964. Experiments run in our laboratory showed that the colours and textures we chose support rapid and accurate element identification, boundary detection, region tracking, and estimation. The result is a visualisation tool that allows users to quickly locate specific plankton densities and the boundaries they form. Users can compare plankton densities to other environmental conditions like sea surface temperature and current strength. Finally, users can track changes in any of the dataset's attributes on a monthly or yearly basis.

CR Categories: H.5.2 [Information Interfaces and Presentation]: User Interfaces—ergonomics, screen design, theory and methods; I.3.6 [Computer Graphics]: Methodology and Techniques—ergonomics, interaction techniques; J.2 [Physical Sciences and Engineering]: Earth and Atmospheric Sciences

Keywords: colour, computer graphics, data mining, human vision, knowledge discovery, multidimensional dataset, perception, preattentive processing, scientific visualisation, texture

1 Introduction

This paper describes our investigation of methods for visualising certain types of large, multidimensional datasets. These datasets are becoming more and more common; examples include scientific simulation results, geographic information systems, satellite images, and biomedical scans. The overwhelming amount of information contained in these datasets makes them difficult to analyse using traditional mathematical or statistical techniques. It also makes them difficult to visualise in an efficient or useful manner.

The size of a dataset can be divided into three separate characteristics: the number of elements in the dataset, the number of attributes or dimensions embedded in each element, and the range of values possible for each attribute. All three characteristics may need to be considered during visualisation.

We are building a visualisation architecture specifically designed to address our problem environment. This system combines an initial data filtering step with a perceptual visualisation step. Numerous papers in the visualisation (Brown et al., 1988; Cambell et al., 1989; Smith and Van Rosendale, 1998; Rosenblum, 1994; Treinish et al., 1989; Treinish and Goettsche, 1991; Wolfe and Franzel, 1988) and databases communities (Silbershatz et al., 1990) have stressed the need for advanced data management techniques during visualisation. We use data mining classification algorithms to identify dependencies, to estimate missing or correct erroneous values, and to compress a dataset's size and dimensionality. Results are displayed to the user in a manner that takes advantage of the low-level human visual system. Offloading the majority of the analysis task on the low-level visual system allows users to very rapidly and accurately perform exploratory visualisation on large multielement displays. Trends and relationships, unexpected patterns or results, and other areas of interest can be quickly identified within the dataset. These data subsets can then be further visualised or analysed as required.

Because our perceptual rules are built on the basic workings of the low-level visual system, they can be applied to a wide range of visualisation environments. This flexibility can be exploited through the use of a "visualisation assistant", an AI-based

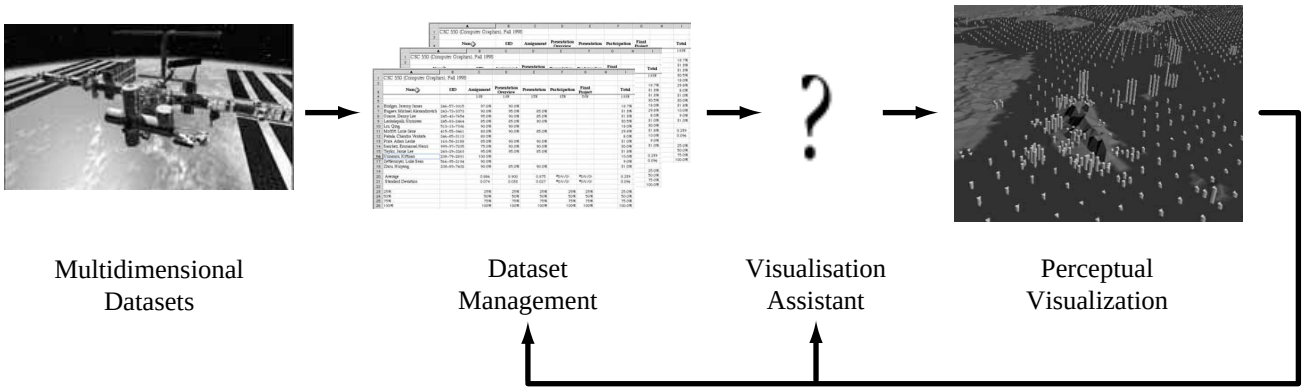


Figure 1: An overview of our visualisation architecture: large, complex, multidimensional datasets are prefiltered using advanced database techniques; the results are sent to a visualisation assistant that combines dataset characteristics and perceptual rules to build a set of appropriate data-feature mappings; some or all of these mappings are applied to visualise the original dataset

planning system that uses our perceptual rules and information about the underlying dataset (e.g. the number, type, importance ordering, and spatial frequency of the attributes, as well as the tasks the user wants to perform) to propose an appropriate collection of data-feature mappings. Some work in this area has been conducted (Bergman et al., 1995; Rogowitz and Treinish, 1993), although current systems are restricted to simple table-lookup schemes. We believe new perceptual rules and a more sophisticated planning system will yield a richer set of visualisation choices (Healey et al., 1999).

This paper focuses on perceptual rules we built for choosing colour and texture during multidimensional data visualisation. The design, execution, and analysis of results from three psychophysical experiments are described in detail. We conclude with a brief overview of our data management techniques, then apply our findings to a real-world dataset, the visualisation of environmental conditions used to conduct salmon migration simulations in the Northwest Pacific Ocean.

2 Oceanography simulations

Our current visualisation testbed for this work is a set of simulations being run in the Westwater Research Centre at the University of British Columbia and the Department of Zoology at North Carolina State University. Researchers in oceanography are studying the growth and movement patterns of different species of salmon in the northern Pacific Ocean. Underlying environmental conditions like plankton density, sea surface temperature (SST), current direction, and current strength affect where the salmon live and how they move and grow (Thomson et al., 1992; Thomson et al., 1994). For example, salmon like cooler water and tend to avoid ocean locations above a certain temperature. Since the salmon feed on plankton blooms, they will try to move to areas where plankton density is highest. Currents will "push" the salmon as they swim. Finally, SST, current direction, and current strength affect the size and location of plankton blooms as they form.

The oceanographers are designing models of how they believe salmon feed and move in the open ocean. These simulated salmon will be placed in a set of known environmental conditions, then tracked to see if their behaviour mirrors that of the real fish. For example, salmon that migrate back to the Fraser River to spawn chose one of two routes. When the Gulf of Alaska is warm, salmon make landfall at the north end of Vancouver Island and approach the Fraser River primarily via a northern route through the Johnstone Strait (the upper arrow in figure 2). When the Gulf of Alaska is cold, salmon are distributed further south, make landfall on the west coast of Vancouver Island, and approach the Fraser River primarily via a southern route through the Juan de Fuca Strait (the lower arrow in figure 2). The ability to predict salmon distributions from prevailing environmental conditions would allow the commercial fishing fleet to estimate how many fish will pass through the Johnstone and Juan de Fuca straits. It would also allow more accurate predictions of the size of the salmon run, helping to ensure that an adequate number of salmon arrive at their spawning grounds.

In order to test their hypotheses, the oceanographers have created a database of SSTs and ocean currents for the region 35° north latitude, 180° west longitude to 62° north latitude, 120° west longitude (figure 2). Measurements within this region are available at $1^\circ \times 1^\circ$ grid spacings. This array of values exists for each month for the years 1956 to 1964, and 1980 to 1989.

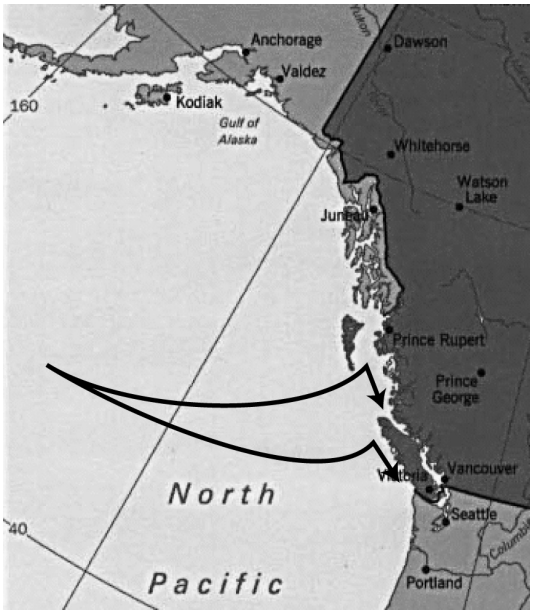


Figure 2: Map of the North Pacific; arrows represent possible salmon migration paths as they pass through the either Johnstone Strait (upper arrow) or the Strait of Juan de Fuca (lower arrow)

Plankton densities have also been collected and tabulated; these are obtained by ships that take readings at various positions in the ocean. Unfortunately, these measurements are much more sparse than the SST and current values. For the years 1956 to 1964, only 1,542 plankton densities are available. This leaves the oceanographers with a number of problems that need to be addressed before their salmon growth and movement models can be tested.

1. A method of estimating plankton densities is required. Currently, spatial interpolation is used to provide the missing values, but this does not work well for months where few (or no) actual densities are available.

2. The oceanographers would like to know how plankton density is related to other environmental conditions like SST, current direction, and current strength. A similar problem will follow: the determination of how salmon growth and open ocean migration patterns are related to underlying environmental conditions.

3. A method is needed for visualising the dataset. This method will be used to display both static (e.g. environmental conditions for a particular month and year) and dynamic results (e.g. a real-time display of environmental conditions as they change over time, possibly with the overlay of salmon locations and movement).

Although the first two problems might be thought to lie outside the scope of visualisation, we feel that management of the underlying data is an inherent part of the visualisation process, particularly for large and complex datasets. The need for data management has been addressed in numerous papers on visualisation (Stonebraker et al., 1993; Treinish, 1993; Treinish et al., 1989). Moreover, this problem was cited as an important area of future research in NSF reports from both the database (Silbershatz et al., 1990) and visualisation communities (Rosenblum, 1994; Smith and Van Rosendale, 1998). To this end, we have implemented extended versions of four data mining classification algorithms that are designed to address the types of problems present in the oceanography datasets.

After using data mining to process the dataset, we must display it on-screen. We have approached the problems of dataset size and dimensionality by trying to exploit the power of the low-level human visual system. Research in computer vision and cognitive psychology provides insight on how the visual system analyses images. A careful mapping of data attributes to visual features (e.g. colour, intensity, and texture) will allow users to perform rapid visual analysis on their data. We must also avoid visual interference effects that can occur when different visual features are combined at the same spatial location. We are currently conducting experiments on the use of colour and texture for multidimensional data visualisation. Results from these experiments are used to visualise the oceanography datasets.

3 Perceptual visualisation

Researchers in computer vision and cognitive psychology are studying how the low-level visual system analyses images. One very interesting result has been the discovery of a limited set of visual features that are processed preattentively, without the need for focused attention (Duncan and Humphreys, 1989; Julész, 1984; Triesman, 1985; Wolfe, 1994). These features can be used to perform certain visual tasks very rapidly and accurately. Examples include searching for elements with a unique visual feature, identifying the boundaries between groups of elements with common features, tracking groups of elements as they move in time and space, and estimating the number of elements with a specific feature. These tasks are preattentive because they can be performed on large multielement displays in less than 200 msec. Moreover, the time required to complete the tasks is independent of the number of data elements being displayed. Eye movements take at least 200 msec to initiate, and random locations of the elements in the display ensure that attention cannot be prefocused on any particular location, yet subjects report that these tasks can be completed with very little effort. This suggests that certain information in the display is processed in parallel by the low-level visual system.

Our interest is focused on identifying relevant results in the vision and psychology literature, then extending these results and integrating them into a visualisation environment. We are currently studying perceptual aspects of colour, orientation, and texture. Results from our experiments allow us to build visualisation tools that use these visual features to effectively represent multidimensional datasets. Because our tools take advantage of the low-level visual system, they offer a number of important advantages:

- Visual analysis is rapid and accurate, since preattentive tasks need an exposure duration of 200 msec or less. (Healey et al., 1995) showed that tasks performed on static frames extended to a dynamic environment, where frames are shown one after another in a movie-like fashion (i.e. tasks that can be performed on a single frame in 200 msec can also be performed on a sequence of frames shown at five frames per second).

- Our tasks are insensitive to display size (to the limits of the display device); increasing the number of elements in the display results in little or no increase in the amount of time required to visually analyse the display. Again, this is a direct result of the fact that preattentive tasks are independent of display size.

- Certain combinations of visual features cause interference patterns in the low-level visual system that can mask information in a display. Our experiments were designed to identify these situations. This means our visualisation tools can be built to avoid data-feature mappings that might interfere with the analysis task.

We chose to use two well-known features to visualise the oceanography datasets: texture and colour. Both features are commonly used in existing visualisation systems. Moreover, both can be decomposed into a number of fundamental "dimensions" (e.g. size, contrast, and regularity for texture, or chromaticity and luminance for colour). This suggests it may be possible to visualise multiple attributes using only these two visual features.

4 Texture selection

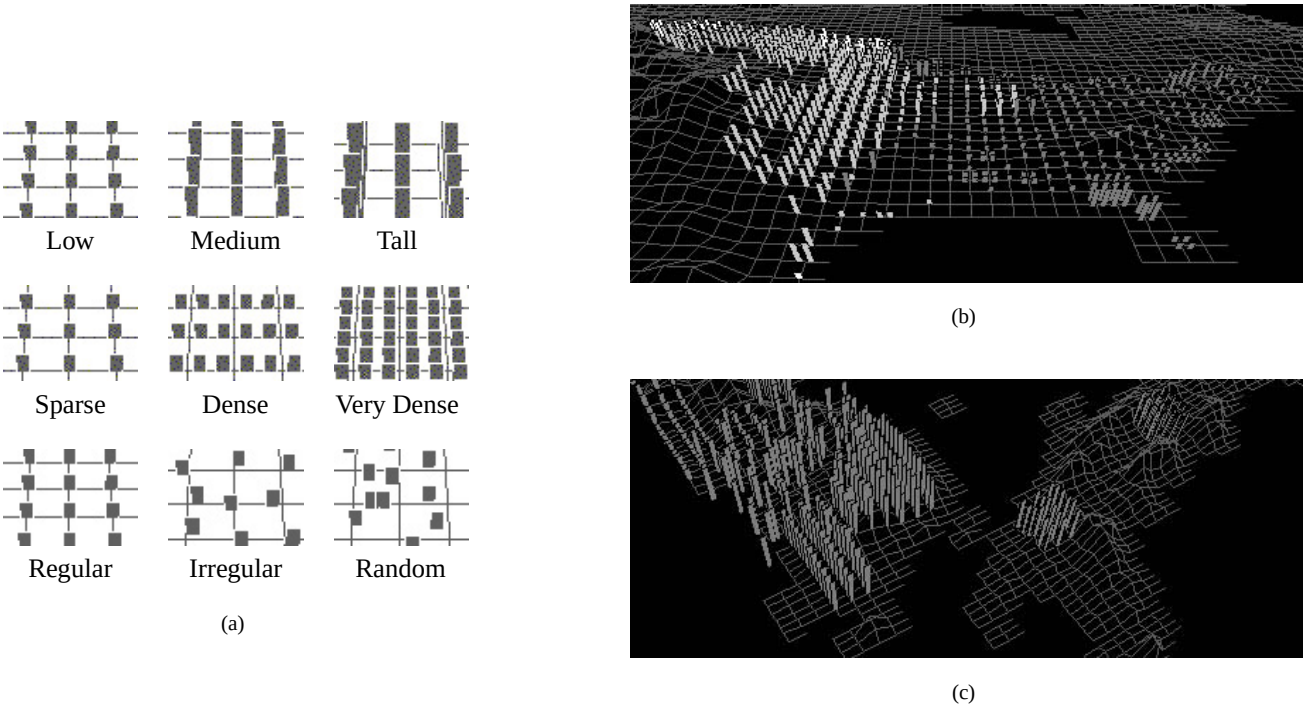
Experiments are being conducted to study the use of perceptual texture elements (or pexels) for multidimensional data visualisation. Texture has been studied extensively in the computer vision and psychology communities (Hallett and Hofmann, 1991; Haralick et al., 1973; Julész et al., 1973; Julész, 1975; Julész et al., 1978; Rao and Lohse, 1993b; Rao and Lohse, 1993a; Reed and Hans Du Buf, 1993; Tamura et al., 1978). One important result was the identification of a collection of "perceptual texture dimensions". Rather than computing statistical measurements, perceptual properties like orientation, contrast, regularity, and granularity can be used to segment and classify regions within an image. (Julész et al., 1973; Julész, 1975; Julész et al., 1978) conducted numerous experiments that verified a viewer's ability to distinguish among texture patches with different contrast or regularity (corresponding to a difference in their first or second-order statistic, respectively). Both (Tamura et al., 1978) and (Rao and Lohse, 1993b; Rao and Lohse, 1993a) asked viewers to divide a collection of texture patterns (Brodatz images) into similar groups. Tamura et al. identified coarseness, contrast, directionality, line-likeness, regularity, and roughness as the perceptual properties their subjects used to classify the textures. Rao and Lohse applied multidimensional scaling to construct three texture dimension axes that corresponded to regularity, directionality, and complexity.

We decided to vary regularity, density, and height during our experiments. Both regularity and density have been proposed by numerous researchers as fundamental perceptual texture dimensions. Although height might not be considered an "intrinsic textural cue", we note that height is one aspect of element size, and that size is an important property of a texture pattern. Results from psychophysical experiments have shown that differences in height are detected preattentively (Triesman, 1985),

moreover, viewers properly correct for perspective foreshortening when they perceive that elements are being displayed in 3D (Aks and Enns, 1996).

We plan to use pexels to visualise multidimensional datasets. As opposed to "texture maps" (patterns that are mapped onto regions of a graphical object), pexels are arrays of elements with visual and spatial characteristics that are controlled by the underlying data being displayed. A number of visualisation systems represent information via texture through the use of glyphs (Pickett and Grinstein, 1988), Wold features (Liu and Picard, 1994), Markov random fields (Li and Robertson, 1995), and Gabor filters (Ware and Knight, 1995). Although we also chose to use individual texture elements to display multidimensional data, our elements are built on results from psychophysical experiments that study how the low-level human visual system "sees" different perceptual texture dimensions. This allows us to harness the visual system's strengths while at the same time avoiding its limitations. Our pexels were designed to support the variation of:

- *regularity:* whether elements are spatially arranged as a regular grid, or with a random distribution,
- *density:* how closely elements are packed together, and
- *height:* how tall or short an element is.



(a)

(b)

(c)

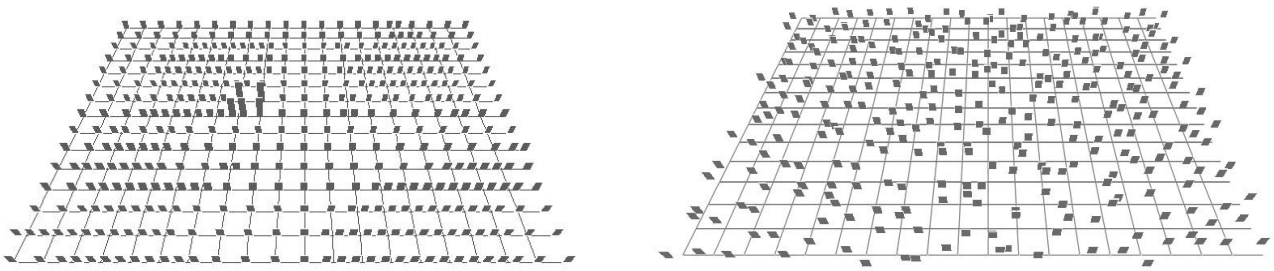
Figure 3: Pexel representation and use: (a) Variation of perceptual texture dimensions height (top row), density (middle row), and regularity (bottom row) across three discrete values; (b) a map of North America, pexels represent areas of high cultivation, height mapped to level of cultivation, density mapped to ground type, greyscale mapped to vegetation type; (c) a map of Japan and the Korean peninsula, height and greyscale mapped to cultivation and vegetation, regularity mapped to ground type

We designed a pexel that looks like a collection of paper strips; at each data element position, a pexel is displayed. The user maps attributes in the dataset to the regularity, density (which controls the number of strips in each pexel), and height of each pexel. A data element controls a pexel's visual appearance by modifying the appropriate texture dimensions based on the element's attribute values. Examples of each of these perceptual dimensions are shown in figure 3a. Figure 3b shows an environmental dataset visualised with texture and greyscale (we used greyscale for printing purposes only; colour is used to display on-screen images). Locations on the map that contain pexels represent areas in North America with high levels of cultivation. Height shows the level of cultivation (75-99% for short pexels, 100% for tall pexels), density shows the ground type

(sparse for alluvial, dense for wetlands), and greyscale shows the vegetation type (dark grey for plains, light grey for forest, and white for woods). Users can easily identify lower levels of cultivation in the central and eastern plains. Areas containing wetlands can be seen as dense pexels in Florida, along the eastern coast, and in the southern parts of the Canadian prairies. Figure 3c shows a map of central Japan and the Korean peninsula. As in figure 3b, height is mapped to cultivation level and greyscale is mapped to vegetation type. In this image, however, regularity is mapped to ground type: regular for alluvial, and irregular for wetlands. Wetlands (i.e. pexels with random placement) can be seen in the northwestern regions of the peninsula.

In order to test our perceptual dimensions and the interactions that occur between them during visualisation, we ran a set of psychophysical experiments designed to investigate a user's ability to rapidly and accurately identify target pexels defined by a particular height, density, or regularity. Observers were asked to determine whether a small group of pexels with a particular type of texture (e.g. a group of taller pexels, as in figure 4a) was present or absent in a 20×15 array. Conditions like target pexel type, exposure duration, target group size, and background texture dimensions differed for each display. This allowed us to test for preattentive task performance, visual interference, and a user preference for a particular target type. In all cases, user accuracy was used to measure performance.

4.1 Experiment 1: Pexels



(a) (b)

Figure 4: Two display types from the taller and regular pexel experiments: (a) a target of medium pexels in a sea of short pexels with a background density pattern (2×2 target group located left of center); (b) a target of regular pexels in a sea of irregular pexels with no background texture pattern (2×2 target group located 3 grid steps right and 7 grid steps up from the lower-left corner of the array)

Each experimental display contained a regularly-spaced 20×15 array of pexels rotated $45°$ about the x-axis (figure 4). All displays were monochrome (i.e. grey and white), to avoid variations in colour or intensity that might mask the underlying texture pattern. Grid lines were drawn at each row and column, to ensure users perceived the pexels as lying on a tilted 3D plane. The array was displayed for a short exposure duration, then removed. Users were then asked whether a group of pexels with a particular target value was present or absent. In order to avoid confusion, each user searched for only one type of target pexel: taller, shorter, sparser, denser, more regular, or more irregular. The appearance of the pexels in each display was varied to test for preattentive performance, visual interference, and feature preference. For example, the following experimental conditions were used to investigate a user's ability to identify taller pexels:

- two target-background pairings: a target of medium pexels in a sea of short pexels, and a target of tall pexels in a sea of medium pexels; different target-background pairings allowed us to test for a subject preference for a particular target type,
- three display durations: 50 msec, 150 msec, and 450 msec; we varied exposure duration to test for preattentive performance, specifically, does the task become more difficult during shorter exposures,
- three secondary texture dimensions: none (every pexel is sparse and regular), density (half the pexels are randomly chosen to be sparse, half to be dense), and regularity (half the pexels are regular, half are random); we added a "background"

texture feature to test for visual interference, that is, does the task become more difficult when a secondary texture dimension appears at random spatial locations in the display, and

- two target group sizes: 2×2 pexels and 4×4 pexels; we used different target group sizes to see how large a group of pexels was needed before the target could be detected by a viewer.

Our experimental conditions produced 36 different display types (two target-background pairings by three display durations by three secondary features by two target group sizes). Users were asked to view 16 variations of each display type, for a total of 576 trials. For each display type, half the trials were randomly chosen to contain a group of target pexels; the remaining half did not.

Examples of two display types are shown in figure 4. Both displays include target pexels. Figure 4a contains a 2×2 target group of medium pexels in a sea of short pexels. The density of each pexel varies across the array, producing an underlying density pattern that is clearly visible. This display type simulates two dimensional data elements being visualised with height as the primary texture dimension and density as the secondary texture dimension. Figure 4b contains a 2×2 target group of regular pexels in a sea of random pexels, with a no background texture pattern. The taller target in figure 4a is very easy to find, while the regular target in figure 4b is almost invisible.

The heights, densities, and regularities we used were chosen through a set of pilot studies. Two patches were placed side-by-side, each displaying a pair of heights, densities, or regularities. Viewers were asked whether the patches were easily discriminable from one another. We tested a range of values for each dimension, although the spatial area available for an individual pexel during our experiments limited the maximum amount of density and irregularity we were able to display. The final values we chose could be rapidly and accurately identified in this limited setting.

The experiments used to test the other five target types (shorter, sparser, denser, more regular, and more irregular) were designed in a similar fashion, with one exception. Experiments that tested regularity had only one target-background pairing: a target of regular pexels in a sea of random pexels (for regular targets), or random pexels in a sea of regular pexels (for irregular targets). Our pilot studies showed that users had significant difficulty discriminating an irregular patch from a random patch. As mentioned above, this was due in part to the small spatial area available to each pexel. Although restricting our regularity conditions to a single target-background pairing meant there were only 18 different display types, users were still asked to complete 576 trials. Thirty-two variations of each display type were shown, 16 of which contained the target pexels, 16 of which did not.

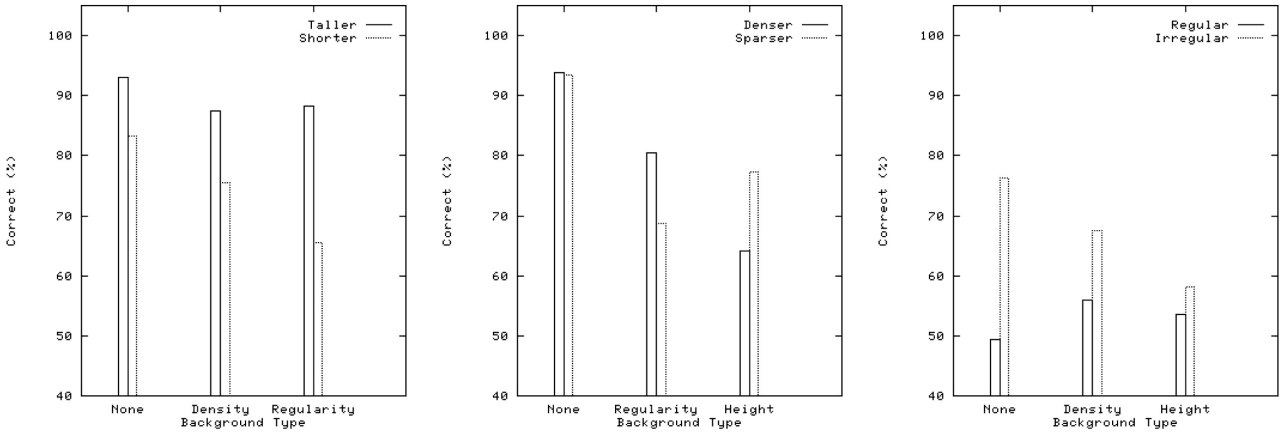
Thirty-eight users (10 males and 28 females) ranging in age from 18 to 26 with normal or corrected acuity participated as observers during our studies. Twenty-four subjects (six per condition) completed the taller, shorter, denser, and regular conditions. Fourteen subjects (seven per condition) completed the sparser and irregular conditions. Subjects were told before the experiment that half the trials would contain a target, and half would not. We used a Macintosh computer with an 8-bit colour display to run our studies. Responses (either "target present" or "target absent") for each trial an observer completed were recorded for later analysis.

4.2 Experiment 1: Results

Each user response collected during our experiments was classified by condition: target type, target-background pairing, exposure duration, secondary texture dimension, target group size, and target present or absent. Trials with the same conditions were combined, and the results were tested for significance using a multi-factor analysis of variance (ANOVA). We used a standard 95% confidence interval to denote significant variation in mean values. In summary, our results showed:

- taller pexels can be identified at preattentive exposure durations (i.e. 150 msec or less) with very high accuracy (approximately 93%); background density and regularity patterns produce no significant interference,
- shorter, denser, and sparser pexels are more difficult to identify than taller pexels, although good results are possible at both 150 and 450 msec; height, regularity, and density background texture patterns cause interference for all three target types,
- irregular pexels are difficult to identify; accuracy was barely acceptable (approximately 76%) even at 150 and 450 msec with no background texture pattern, and
- regular pexels cannot be accurately identified; the percentage of correct results approached chance (i.e. 50%) for every condition.

Taller targets were identified preattentively with very high accuracy (Figure 5a). Background density and regularity patterns caused no significant interference ($F(2,10) = 4.165$, $p = 0.292$). Although accuracy for shorter targets was somewhat lower, it was still acceptable when there was either no background texture pattern or a density texture pattern (83% and 75%, respectively). Both background regularity and density caused a statistically significant reduction in performance ($F(2,10) = 25.965$, $p = 0.0001$). Results showing taller targets were "more salient" than shorter targets was not unexpected; similar asymmetries have been documented by both (Triesman, 1985) and (Aks and Enns, 1996).



(a) (b) (c)

Figure 5: Graphs showing the percentage of correct target detection responses for the six target types, horizontal axis represents background texture pattern, vertical axis represent percentage of correct responses (percentage correct averaged over all trials for the given target type and background pattern): (a) results for taller and shorter targets; (b) results for denser and sparser targets; (c) results for regular and irregular targets

As with height, dense in sparse targets were easier to identify than sparse in dense, particularly with a background regularity pattern. Accuracy with no background texture pattern was as high as for taller targets (Figure 5b). In both cases, a significant interference effect occurred when a background texture was present ($F(2,10) = 77.007$, $p = 0.0001$ and $F(2,10) = 43.343$, $p = 0.0001$ for denser and sparser targets, respectively). Height reduced accuracy dramatically for denser targets, while both height and regularity interfered with the identification of sparser targets.

Performance was poorest for regular and irregular targets. Accuracy for irregular targets was poor (approximately 76%) when there was no background texture pattern. Results were significantly lower for displays that contained a variation in either density or height ($F(2,12) = 7.147$, $p = 0.0118$, with correct responses of 68% and 58%, respectively). Observers were completely unable to detect regular targets in a sea of irregular pexels (see also figure 4b). Even with no background texture pattern, correct responses were only 49%. Similar near-chance results (i.e. correct responses of 50%) occurred when height and regularity texture patterns were displayed. We concluded that subjects resorted to guessing whether the target was present or absent.

For target group sizes, results showed that 4×4 targets are significantly easier to find than 2×2 targets for four target types: taller, shorter, denser and sparser ($F(1,4) = 20.067$, $p = 0.0009$, $F(1,4) = 93.607$, $p = 0.0001$, $F(1,4) = 26.506$, $p = 0.0003$, and $F(1,4) = 8.041$, $p = 0.014$, respectively). There were no significant within-condition F-values, suggesting the effect of target group size (larger easier than smaller) was consistent for each display type. Finally, only shorter and sparser targets showed any significant effect of display duration ($F(2,10) = 25.965$, $p = 0.0001$ and $F(2,10) = 43.343$, $p = 0.0001$, respectively). Again, there were no within-condition F-values; increasing the display duration for shorter or sparser targets resulted in a consistent increase in performance, regardless of the display type being shown.

5 Colour selection

We are also interested in using colour to represent attributes in our multidimensional datasets. Colour is one of the most commonly-used visual features in scientific and data visualisation. Because of this, a large body of past work has studied the use of colour for a variety of visualisation tasks. For example, (Ware and Beatty, 1988) used colour to detect coherence in five-dimensional datasets. (Ware, 1988) also studied the use of luminance and hue for the display of metric and form data on a continuous two-dimensional surface. (Levkowitz and Herman, 1992) designed a linear-optimal continuous colour scale. Finally, researchers at IBM have developed a rule-based perceptual colour selection technique called PRAVDAColor (Bergman et al., 1995).

Our requirements were somewhat different from those of the papers listed above, in part because we wanted to pick a set of discrete colours rather than building a continuous colour scale. We studied the following questions during our investigation:

- How can we choose colours that are easy to identify?
- How can we guarantee that each colour is equally easy to identify from all the other colours being displayed?
- What is the maximum number of colours we can display simultaneously, while still providing these properties?

Previous work has identified a number of issues that may need to be considered during colour selection. We decided to test three promising criteria:

- *colour distance:* the perceived colour distance from each colour to its nearest neighbour(s) is equal and above a minimum threshold (Wyszecki and Stiles, 1982); distance is measured in a perceptually balanced colour model, in our case, using CIE LUV,
- *linear separation:* each colour must be linearly separable from all the other colours (Bauer et al., 1996; D'Zmura, 1991), again by a minimum threshold measured in a perceptually balanced colour model, and
- *colour category:* each colour must occupy a uniquely named colour region (Kawai et al., 1995).

Perceived colour difference between pairs of colours is often measured using a perceptually balanced colour model that roughly equates Euclidean distance to perceived difference. To understand linear separation and colour category, consider a display containing n colours $c_1, \ldots, c_n$ displayed simultaneously, where an observer searchers for a target colour t, $t \neq c_i \; \forall i = 1, \ldots, n$, that is randomly present or absent in the display. If t can be separated from every c_i using a single straight line in colour space, we say t is linearly separable from its background colours. If t occupies a named colour region different from the named regions occupied by every c_i, we say t has a unique colour category. Experimental results have shown that linear separation (Bauer et al., 1996; D'Zmura, 1991) and unique colour category (Kawai et al., 1995) significantly increase the perceptual distinguishability of t, relative to displays where one or both of these criteria are not satisfied.

We began our investigation by controlling colour distance and linear separation, but not colour category. This choice was made for two reasons. First, there is no clear consensus in the cognitive psychology community on whether the way we name objects (in our case, colours) can affect our perceived difference between them. Second, traditional methods for subdividing a colour slice into named colour regions are time consuming and difficult to control. We proceeded under the assumption that the user might choose to search for any one of the available data elements at any given time. This is typical during exploratory data analysis; users will often change the focus of their investigation based on the data they see as the visualisation unfolds. Our requirement meant that the colour selection technique had to allow for rapid and accurate identification of any of the elements being displayed.

5.1 CIE LUV colour model

Before describing our experiment design, we provide a brief overview of the CIE LUV colour model. This model was used to measure colour distance during our studies.

The CIE LUV colour model was proposed by the Commission Internationale de L'Èclairge (CIE) in 1976 (CIE, 1976). Colours are specified using the three dimensions (L^*, u^*, v^*). L^* encodes the luminance or intensity of a given colour, while u^* and v^* encode its chromaticity (u^* and v^* roughly correspond to the red-green and blue-yellow opponent colour channels).

CIE LUV is a perceptually balanced colour model. This provides two useful properties for controlling perceived colour difference. First, colours with the same L^* are isoluminant (i.e. they are perceived as having the same intensity). Second,

Euclidean distance and perceived colour difference (specified in ΔE^* units) can be interchanged, since the colour difference between two colour stimuli x and y is roughly:

$$\Delta E^*_{xy} = \sqrt{(\Delta L^*_{xy})^2 + (\Delta u^*_{xy})^2 + (\Delta v^*_{xy})^2} \quad (1)$$

Our techniques do not depend on CIE LUV; we could have chosen to use any perceptually balanced colour model (e.g. CIE Lab, Munsell, or the Optical Society of America uniform colour system). We picked CIE LUV because it is reasonably well known, because it was designed for emissive displays like a colour monitor, and because its (L^*, u^*, v^*) dimensions are continuous (as opposed to being discrete like the dimensions used in the Munsell colour model).

5.2 Experiment 2: Colour distance and linear separation

We ran four studies to investigate the tradeoff between the number of colours displayed and the time required to determine the presence or absence of a target element. Each study displayed a different number of unique colours, which we identify using the names red, yellow-red, yellow, green-yellow, green, blue-green, blue, purple-blue, purple, and red-purple (or R, YR, Y, GY, G, BG, B, PB, P, and RP):

- *three-colour study:* each display contained three different colours (i.e. one colour for the target and two for the non-targets) that we named R, GY, and PB
- *five-colour study:* each display contained five different colours: R, Y, GY, B, and P
- *seven-colour study:* each display contained seven different colours: R, Y, GY, G, BG, P, and RP
- *nine-colour study:* each display contained nine different colours: R, YR, Y, GY, G, BG, PB, P, and RP

Every colour in a given study was tested as a target. For example, the three-colour study was run three times, first with an R element acting as a target (and GY and PB elements acting as non-targets), next with a GY target (and R and PB non-targets), and finally with a PB target (and R and GY non-targets). Faster search times for certain targets would have to be explained in terms of colour distance, linear separation, or colour category. Each of the four studies were themselves divided in the following manner:

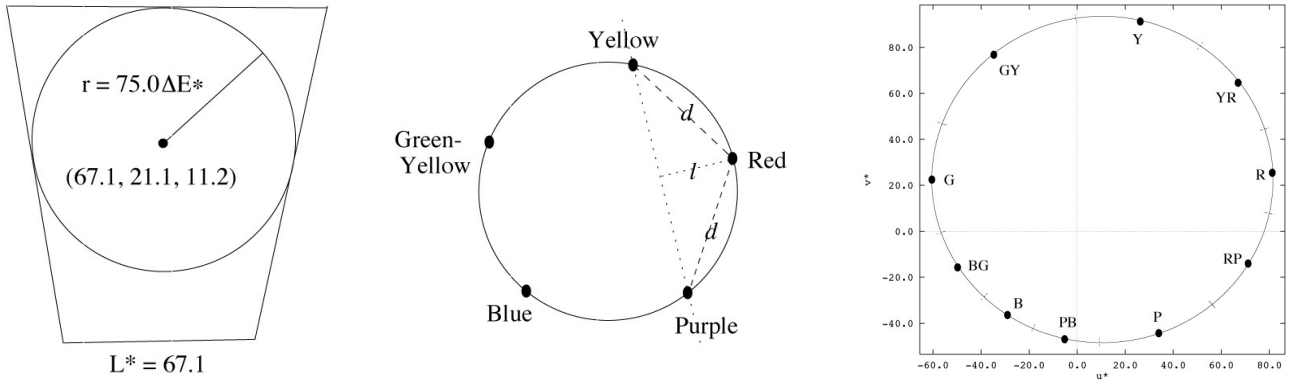
- elements in each display were drawn as coloured squares, and were randomly located on an underlying 9×9 grid that covered the entire viewing area of the monitor
- half of the displays were randomly chosen to contain an element that used the target colour; the other half did not
- one-third of the displays contained a total of 17 elements (one target and 16 non-targets if the target was present, or 17 non-targets if the target was absent); one-third of the displays contained a total of 33 elements; one-third of the displays contained a total of 49 elements

We began the colour selection process by obtaining the chromaticities of our monitor's triads. We also measured the luminance of the monitor's maximum intensity red, green, and blue with a spot photometer. These values were needed to convert colours from CIE LUV into the monitor's RGB gamut.

We wanted to ensure that the colours we chose had the same perceived intensity. Previous research has suggested that random variation in intensity can interfere with an observer's ability to perform visualisation tasks based on colour (Callaghan, 1984). In order to guarantee isoluminance, all the colours were chosen from a single u^*, v^*-slice through the CIE LUV colour space at $L^* = 67.1$. We wanted to maximize the number of available colours, while still maintaining control over colour distance and linear separability. To do this, we computed the boundary of the monitor's gamut in the $L^* = 67.1$ slice. We then found the largest circle inscribed within the gamut (figure 6a).

Given the maximum inscribed circle, we chose colours that were equally spaced around its circumference. For example, during the five-colour study we chose colours at positions 14°, 86°, 158°, 230°, and 302° counterclockwise rotation from the x-axis (figure 6b). This method ensured that neighbouring colours had a constant colour distance d between each other. It also ensured that any colour acting as a target had a constant linear separation l to every other (non-target) colour. A similar technique was used to select colours for the three-colour, seven-colour, and nine-colour studies.

Thirty-eight users with normal or corrected acuity participated as observers during our studies. After ensuring they were not colour blind, each observer was asked to complete one or more target blocks. Each target block consisted of 360 displays



(a) (b) (c)

Figure 6: Choosing colours from the monitor's gamut: (a) the boundary of the gamut at $L^* = 67.1$, along with the maximum inscribed circle centered at $(L^*, u^*, v^*) = (67.1, 21.1, 11.6)$, radius $75.0\,\Delta E^*$; (b) five colours chosen around the circle's circumference; each element has a constant colour distance d with its two neighbours, and a constant linear separation l from the remaining (non-target) elements; (c) the same circle automatically subdivided into ten named colour regions

testing a single colour target from one of the four studies. A total of 72 target blocks were completed, three for each of the 24 different target colours. We used a Macintosh computer with an 8-bit colour display to run our studies. Responses (either "target present" or "target absent") and the time to respond for each display an observer completed were recorded for later analysis.

5.3 Experiment 2: Results

Observers had very little difficulty identifying targets during the three-colour and five-colour studies. Graphs of mean response time across display size were relatively flat for every colour (figures 7a and 7b show results for the three-colour study). Mean response times ranged from 459 msec to 549 msec during the three-colour study, and from 508 msec to 661 msec during the five-colour study. Mean response error during both studies was approximately 2.5%. We concluded that users could accurately identify the target in all cases, and that the time required to do so was relatively independent of display size. This suggests that, even when using five different colours, the visual system can search for any one of the colours in parallel.

Target identification became significantly more difficult for certain colours during the seven-colour and nine-colour studies (figures 7c and 7d shows results for the seven-colour study). Mean response error during the seven-colour study was still low, at approximately 3.3%. The P, Y, R, BG, and RP targets all exhibited relatively flat response time graphs. Mean response time for these elements ranged from 611 msec to 870 msec. The G and GY targets, however, gave response times typical of serial search. An increase in the number of elements being displayed brought on a corresponding increase in response time. The increase for target-absent displays (approximately 19 msec per additional element for GY, and 17 msec per element for G) was roughly twice that for target-present displays (7 msec and 8 msec per element for GY and G, respectively). Observers had to search through, on average, half the elements before they found the target in target-present displays. In target-absent displays, however, they had to search through all the elements to confirm that no target existed. This explains why per item search time increased roughly twice as fast for target-absent displays. A similar set of results was obtained during the nine-colour study. Overall mean response error increased to 8.1%; it was lowest for the P, Y, and PB targets (approximately 3.4%), and highest for the G, GY, and RP targets (approximately 14%).

5.4 Experiment 3: Colour category integration

Results from our four studies showed that controlling colour distance and linear separation alone is not enough to guarantee consistently good (or consistently bad) target identification for every colour. Results from (Kawai et al., 1995) suggest that colour category can also have a strong effect on the amount of time required to identify a colour target. We decided to see whether colour category results could explain the asymmetric response times we observed during the seven-colour and nine-colour studies.

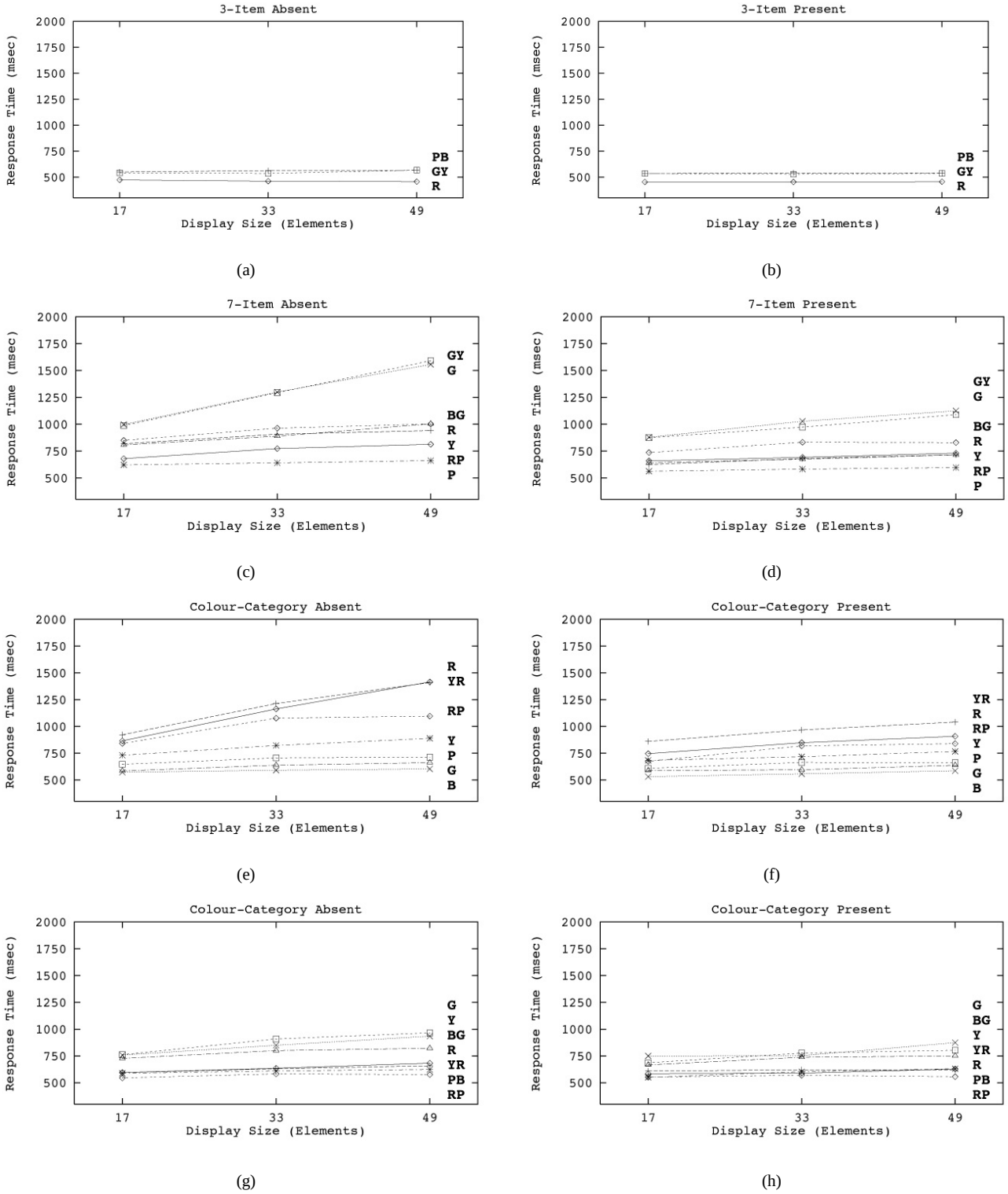


Figure 7: Response time graphs for four of the six studies, the graph on the left represents displays where the target was absent, while the graph on the right represents displays where the target was present: (a,b) response time as a function of display size (i.e. total number of elements shown in the display) for each target from the three-colour study; (c,d) response times for each target from the seven-colour study; (e,f) response times for each target from the first colour category study; (g,h) response times for each target from the combined distance-separation-colour category study

We constructed a simple technique to quickly divide any part of a colour model (in this case, the circumference of a circle embedded in CIE LUV) into named regions. Our technique uses an automatic step to initially divide the model into ten named regions. This was done by translating colours into the Munsell colour space, then using the hue dimension of this space to name the colour. Next we performed a short experimental step, where individual users named representative colours from each of the ten regions (figure 6c, the black dots at the center of each named region were used as a representative colour for that region). This allows us to measure the amount of name similarity that occurs between different colour regions. This similarity can be used to estimate the perceived colour category overlap between these regions.

During the seven-colour study the P, Y, R, BG, and RP targets gave good performance. The G and GY targets gave poor performance. An examination of our category experiment suggests that poor target colours had a high category overlap with at least one of the background distractor colours. For example, during our colour naming experiments almost every observer named both the G and GY representative colours "green". This suggests colours from the G and GY regions are perceived to be lying in a common colour category called green. Similar results were found for representative colours from the B and PB regions (observers uniformly named these colours "blue"). A Spearman correlation between a ranking of user performance and total category overlap was $r_{rank} = 0.821$, confirming our hypothesis that higher mean response times for a given target correspond to a higher category overlap with background colours. A similar correlation value ($r_{rank} = 0.762$) was found during the nine-colour experiment.

Results from investigating colour category integration might imply that effective colours can be selected by controlling colour category alone. In our experiments most observers named representative colours from both the G and GY regions as "green". Similarly, representative colours from B and PB were uniformly named "blue". Both these regions were compressed, resulting in a total of eight categories with low perceptual overlap. We chose seven new colours from these low-overlap colour regions, then reran our experiments. Although response times are somewhat better than those from the original seven-colour study, several colours still exhibited poor search performance (figures 7e and 7f). This was explained by examining the distance between neighbouring pairs of colours, and the linear separation for each colour when it acted as a target. Colours that gave the worst search performance had the smallest neighbour distances and linear separation; in fact, our new R, YR, and RP targets had a linear separation that was smaller than the one used during the nine-colour study. Colours that gave the best search performance had the largest neighbour distances and linear separation.

From these results we concluded that colour category alone cannot be used to ensure consistently good identification based on colour. Colour distance and linear separation need to be considered, since they can affect search performance. We used a more systematic colour selection technique to choose another set of seven colours that guaranteed low category overlap while maximizing colour distance and linear separation. This was done by selecting a single colour at the center of the G-GY region, then choosing another six colours spaced equally about the remainder of the colour circle. Results from displays using these colours as targets are shown in figure 7g and 7h. Mean response error was 5.6%. Response time graphs for all seven colours are much flatter than in the original seven-colour study, although G and Y are still given mixed results during target-absent displays. It appears that seven isoluminant colours is the maximum we can display at one time, while still allowing rapid and accurate identification of any one of the colours.

6 Data mining classification

Data mining or knowledge discovery, as it is sometimes referred to, is a relatively new area of database research. Data mining is defined as "the nontrivial extraction of implicit, previously unknown, and potentially useful information from data" (Frawley et al., 1991). This is done by combining a variety of database, statistical, and machine learning techniques. We are interested in data mining algorithms that perform classification. We believe that these algorithms can be used to improve the efficiency of visualising large, multidimensional datasets. Their advantages are twofold. First, they can be used to reduce the amount of data that needs to be displayed. Second, they can be used to "discover" previously unknown and potentially useful information.

We have implemented extended versions of four existing data mining techniques. These techniques are being applied to a number of real-world datasets, to see if they offer improved efficiency or usefulness compared to visualisation without any form of data management. Two of the algorithms are based on decision trees (Agrawal et al., 1992; Quinlan, 1986), one algorithm is based on statistical tables (Chan and Wong, 1991), and one algorithm is based on rough sets (Ziarko, 1991). All four algorithms build their classification rules from a user-supplied training set. The decision tree algorithms use chi-squared tests to maximize information gain at each level in the tree; leaves in the tree hold a single classification value. The statistical algorithm constructs a table containing a probability for every possible *(attribute value, classification value)* pair; the classification value with the highest total probability is then assigned to an unclassified element. The rough set algorithm uses equivalence relations to compute importance weightings for each data attribute, and rules that map combinations of attribute values to a particular

classification value; the rule whose matching attributes sum to the highest importance weighting is used to classify unknown elements.

The data mining algorithms are designed to process a training set, then provide classification values for one or more unclassified elements. During visualisation, however, users often require more than a simple classification value. We modified and extended the algorithms to provide additional results, in particular, the ability to identify attributes that are significant to a given classification task, confidence weights for each classification performed, and the ability to compare the viability of different classification values for an unclassified element.

Our initial concern for the oceanography datasets was accurate estimation of plankton densities. We created a training set that contained all available density measurements (a total of 1,542 elements) for the years 1956 to 1964. Each of these readings included a latitude, longitude, and the month and year the reading was taken. We added to each element the corresponding SST, current direction, and current strength (these were taken directly from the environment database for the given month, year, latitude, and longitude). Continuous values (SST, current direction and strength, and plankton density) were divided into five equal-width ranges; each value's range was used during classification. Although the data mining algorithms will automatically range continuous data, we found that more accurate results are obtained when a user familiar with the dataset chooses the bounds for each range. The ranges used for SST, current U and V direction, current strength, and plankton density are shown in table 1.

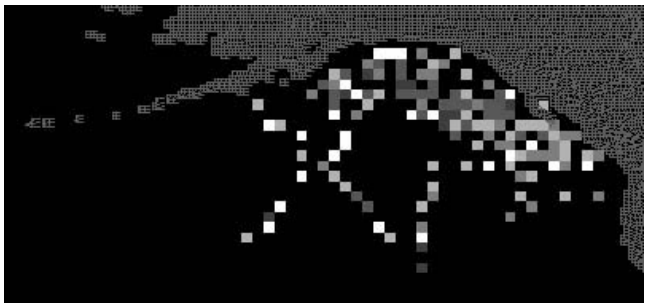
SST (degrees C)	*very cold* $\text{SST} < 6.34$	*cold* $6.34 \leq \text{SST} < 8.98$	*warm* $8.98 \leq \text{SST} < 11.82$	*hot* $11.82 \leq \text{SST} < 14.80$	*very hot* $\text{SST} \geq 14.80$
Current	*N-NE* $(U,V) < 72^\circ$	*NE-SE* $72^\circ \leq (U,V) < 144^\circ$	*SE-SW* $144^\circ \leq (U,V) < 216^\circ$	*SW-NW* $216^\circ \leq (U,V) < 288^\circ$	*NE-N* $(U,V) \geq 288^\circ$
Strength (cm/s)	*very weak* $\text{Str} < 6.09$	*weak* $6.09 \leq \text{Str} < 9.02$	*steady* $9.02 \leq \text{Str} < 11.57$	*strong* $11.57 \leq \text{Str} < 14.54$	*very strong* $\text{Str} \geq 14.54$
Plankton (g/m^3)	*very sparse* $\text{Plk} < 10$	*sparse* $10 \leq \text{Plk} < 28$	*normal* $28 \leq \text{Plk} < 53$	*dense* $53 \leq \text{Plk} < 114$	*very dense* $\text{Plk} \geq 114$

Table 1: Boundaries used to divide SST (measured in degrees Celsius), current (U, V) direction (directions are divided into five 72° quadrants, where 0° represents North), current strength (measured in centimeters per second), and plankton density (measured in grams per meter cubed) into five equal-width ranges.

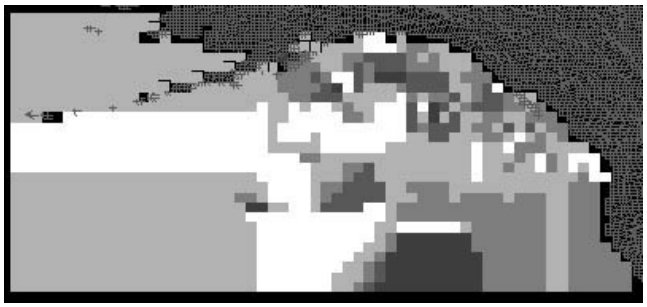
We started by reading the training set with each of our four data mining algorithms, then using significance weights to identify which attributes were being used to classify (i.e. estimate) plankton density. All four algorithms reported similar results: month was the most important attribute to use during classification, followed by current strength and SST. Other attributes (current direction and year) had a significance weight of zero. The oceanographers concurred with these results; plankton densities display a seasonal variability, large current upwellings will produce larger plankton blooms, and higher ocean temperatures cause faster plankton production and higher overall densities. These results allowed us to restrict our visualisations to month, SST, strength, and plankton density. The oceanographers searched these displays for temperature and current patterns, and their relationship to the corresponding plankton densities.

Once rules are built from the training set, the data mining mining algorithms can assign an estimated plankton density to unknown ocean positions based on SST, current strength, and month. This was done for all missing plankton densities for the years 1956 to 1964. We used the interval classification algorithm (Agrawal et al., 1992), since it showed the smallest sensitivity to errors in its training set during prior testing. Approximately 11% of the estimated plankton densities exhibited low confidence weights. Although these elements are included during visualisation, we are studying them in isolation, to try to determine why the data mining algorithm had difficulty assigning them a density value. Initial investigation suggests that elements with certain combinations of month, SST, and current strength were not present in our training set. As a result, the data mining algorithms were uncertain about how to analyse these kinds of elements during classification.

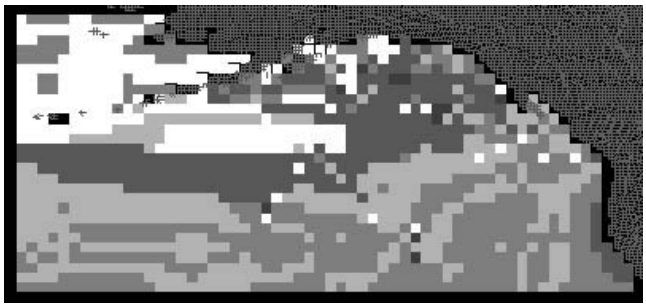
An example of our results is shown in figure 8. The plankton densities that were actually available are shown in figure 8a. Figure 8b shows missing values that have been estimated using spatial interpolation. As expected, this technique performs poorly for locations in the ocean where no initial values are present. Most of the northwest and southwest quadrants have been classified to have moderate density; there is almost certainly more variation in this region. Data mining, on the other hand, uses the month, along with the underlying SSTs and current strengths, to estimate plankton density. In figure 8c, the northwest and southwest quadrants have variability similar to that which exists across the known densities (figure 8a). It is impossible to conclude that the values provided by the data mining algorithm are "more correct" than the interpolated values. However, since



(a)



(b)



(c)

Figure 8: Known and estimated plankton densities for August 1956, greyscale used to represent density (dark grey for low to white for high): (a) known densities; (b) missing densities estimated using interpolation (note the banding that occurs at the boundaries of the array); (c) missing densities estimated using data mining (patterns within the array correspond to the prevailing SSTs and current strengths)

the data mining algorithms classify each element independent of its spatial or temporal neighbours, our algorithms are not at a significantly higher disadvantage when no real data values exist near the value we want to estimate.

7 Oceanography visualisation

We chose to visualise SST and current strength with plankton density, since these attributes (along with month) were significant during data mining. Displaying the three attributes together allows the oceanographers to search for relationships between plankton density, current strength, and SST. Plankton is displayed using colour; SST and current strength are displayed using texture. Colours for the five plankton ranges were chosen using our colour selection technique. Although other colour scales were available (for example, by (Ware, 1988)), our colours are specifically designed to highlight outliers, and to show clearly the boundaries between groups of elements with a common plankton density. We display the five plankton density ranges from low to high using blue (monitor RGB=36, 103, 151), green (monitor RGB=18, 127, 45), brown (monitor RGB=134, 96, 1), red (monitor RGB=243, 51, 55), and purple (monitor RGB=206, 45, 162),

For the underlying texture, we mapped current strength to height and SST to density. Our choices were guided by results we observed from texture experiments, specifically:

- differences in height are easier to detect, compared to differences in density or regularity,
- variation in height may mask differences among sparse, regular, or irregular elements; this appears to be due to the occlusion that occurs when tall pexels in the foreground hide short pexels in the background; this will be less important when users can control their viewpoint into the dataset (our visualisation tool allows the user to interactively manipulate the viewpoint), and

- tightly spaced grids can support up to three easily distinguishable density patterns; placing more strips in a single pexel (e.g. arrays of three by three or four by four strips) will either cause the strips to overlap with their neighbours, or make each strip too thin to easily identify.

Because there may be a feature preference for height over density, and because current strength was deemed "more important" than SST during data mining, we used height to represent currents and density to represent SSTs. The five ranges of current strength are mapped to five different heights. We do not use a linear mapping, rather the lower two ranges (corresponding to the weakest currents) are displayed using two types of short pexels, and the upper three ranges (corresponding to the strongest currents) are displayed using three types of tall pexels. This allows a user to rapidly locate boundaries between weak and strong currents, while still being able to identify each of the five ranges. Although it was not necessary for this particular visualisation, we could have use continuous ranges of height with discrete gaps between them to represent the currents (e.g. heights of 0.1–0.3 and 0.5–0.7 for the lower two ranges of weak currents, and heights of 1.4–1.6, 1.8–2.0, and 2.2–2.4 for the upper three ranges of strong currents). We have used exactly this kind of technique for a weather visualisation and typhoon tracking application (Healey and Enns, 1999). Although it is difficult to determine an absolute attribute value from a particular height (e.g. this pexel's height means windspeed is 12.5 metres per second) viewers had no difficulty comparing relative heights (e.g. since this pexel is about twice as tall as that one, its windspeed must be about twice as large). Our visualisation technique is designed for the high-level exploration and analysis of trends, relationships, and values in the dataset, rather than the determination of the exact attribute values that produced a particular type of pexel.

For SSTs, the lower three ranges (corresponding to the coldest SSTs) are displayed with a pexel containing a single strip, while the upper two ranges (corresponding to the warmest SSTs) are displayed with pexels containing arrays of two and four strips, respectively. The densities we chose allow a user to see clearly the boundaries between cold and warm temperature regions. If necessary, users can change the range boundaries to focus on different SST gradients.

The oceanographers want to traverse their datasets in monthly and yearly steps. Experiments run in our laboratory have shown that preattentive tasks performed on static frames can be extended to a dynamic environment, where displays are shown one after another in a movie-like fashion (Healey et al., 1995). Our visualisation tool was designed to allow users to scan rapidly forwards and backwards through the dataset. This makes it easy to compare changes in the value and location of any of the environmental variables being displayed. The oceanographers can track seasonal changes in current strength, SST, and plankton density as they move month by month through a particular year. They can also see how interannual variability affects the environmental conditions and corresponding plankton densities for a particular month across a range of years.

Figure 9 shows three frames from the oceanography dataset: February and June, 1956. Colour shows the seasonal variation in plankton densities. Height and density allow the oceanographers to track current strengths and SSTs. In February (figure 9a), most plankton densities are less than 28 g/m^3 (i.e. blue and green strips). Currents are low in the north-central Pacific; a region of weak currents also sits off the south coast of Alaska. Most of the ocean is cold (sparse pexels), although a region of higher temperatures can easily be seen as dense pexels in the south. In June (figure 9b) dense plankton blooms (red and purple strips) are present across most of the northern Pacific. The positions of the strong currents have shifted (viewing the entire dataset shows this current pattern is relatively stable for the months March to August). Warmer SSTs have pushed north, although the ocean around Alaska and northern British Columbia is still relatively cold. By October the plankton densities have started to decrease (green, brown, and red strips); few high or low density patches are visible. Current strengths have also decreased in the eastern regions. Overall a much larger percentage of the ocean is warm (i.e. dense pexels). This is common, since summer temperatures will sometimes last in parts of the ocean until October or November.

8 Conclusions

This paper described our two-step approach to visualising complex scientific datasets. We begin by using data mining algorithms to identify significant trends, classify elements, and focus the dataset. The results are then visualised using perceptual features. We demonstrated our techniques by analysing and visualising an environmental dataset being used to run salmon growth and migration simulations. We used data mining to estimate missing plankton densities, and to identify the attributes significant to this estimation. The resulting sea surface temperatures, ocean current strengths, and plankton densities were visualised using colour and texture. The colours and textures (built as arrays of paper strips with varying height and density) were chosen based on results from perceptual experiments run in our laboratory. We exploit the low-level human visual system with our visualisation tools. This makes a large part of the visual analysis automatic; little effort or focused attention is required by the user to perform exploratory tasks like target identification, boundary detection, region tracking, and estimation. These tasks can be carried out on sequences of displays shown one after another at relatively high frame rates (e.g. 100 to 200 msec

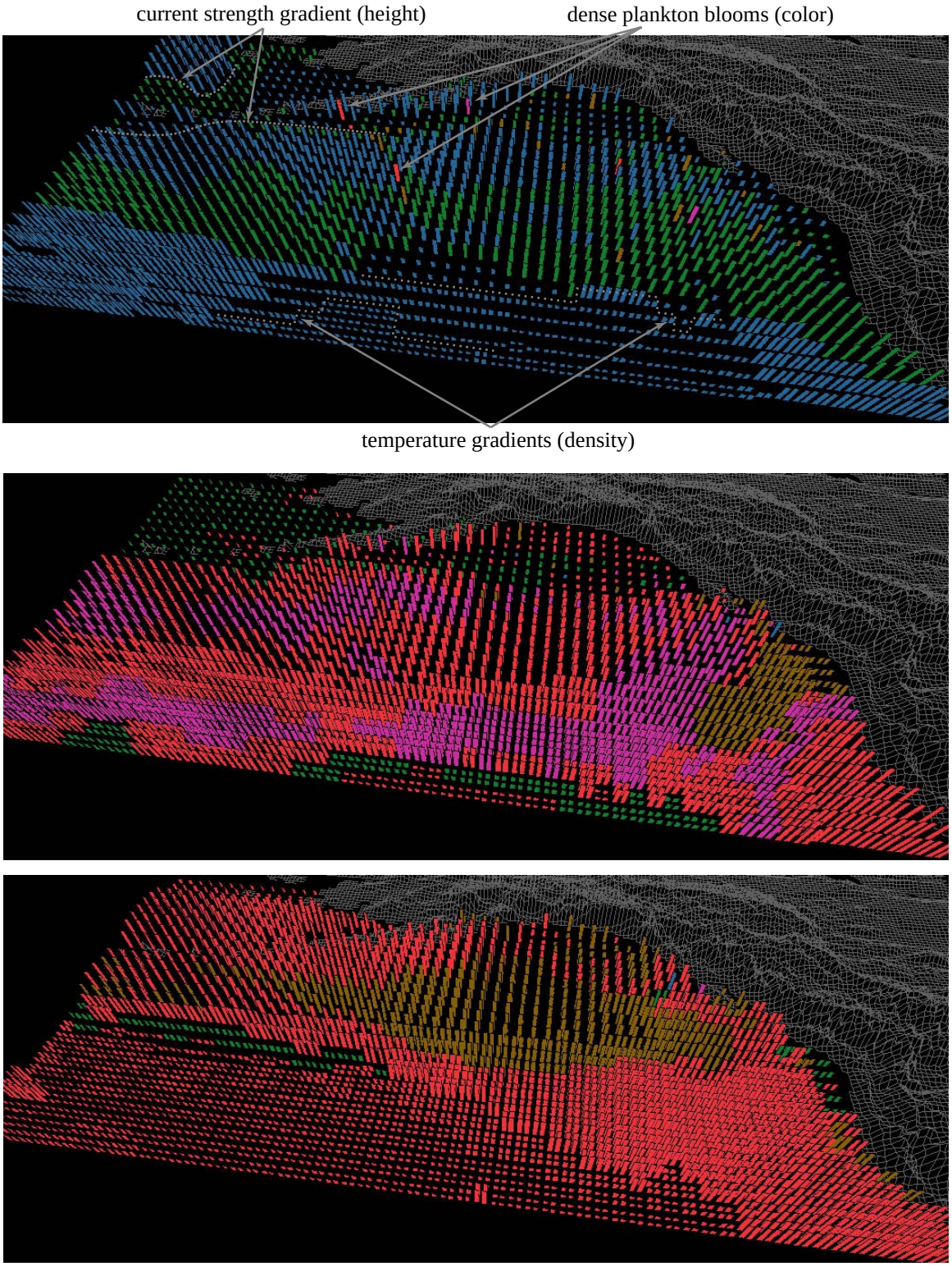


Figure 9: Visualisation of the oceanography datasets, colour used to represent plankton density (blue, green, brown, red, and purple represent lowest to highest densities), height used to represent current strength, texture density used to represent SST: (top) February, 1956; (middle) June, 1956; (bottom) October, 1956

per frame). This technique allows the oceanographers to scan through their datasets month by month or year by year to view seasonal or interannual changes in environmental conditions.

Although our practical example in this paper was an oceanographic dataset, data mining and perceptual visualisation can be applied to a wide range of visualisation environments. We have used perceptual colour selection to highlight regions of interest in reconstructed medical volumes (Tam et al., 1997). We have also used data mining to estimate sea surface temperatures in an environmental dataset from NASA (Healey, 1996); we showed that our results were more accurate than estimates produced by bilinear interpolation. We will continue to test the flexibility of our techniques with new visualisation problems and datasets.

9 Acknowledgments

I would like to thank Dr. Peter Rand and Dr. Michael Healey for access to their data and their expertise from the salmon growth simulations. Dr. James Enns, Dr. Vince Di Lollo, and Dr. Kellogg Booth provided valuable technical advice during my research. I would also like to thank Jeanette Lum for coordinating and running our experiment sessions. This research was funded in part by the National Science and Engineering Research Council of Canada, and by the Office of Naval Research (Grant N00014-96-1120) and the Ballistic Missile Defense Organization through the Multiuniversity Research Initiative.

References

Agrawal, R., Ghosh, S., Imielinski, T., Iyer, B., and Swami, A. (1992). An interval classifier for database mining applications. In *Proceedings 18th Very Large Database (VLDB) Conference*, pages 560–573.

Aks, D. J. and Enns, J. T. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Perception and Performance*, 22(6):1467–1481.

Bauer, B., Jolicoeur, P., and Cowan, W. B. (1996). Visual search for colour targets that are or are not linearly-separable from distractors. *Vision Research*, 36:1439–1446.

Bergman, L. D., Rogowitz, B. E., and Treinish, L. A. (1995). A rule-based tool for assisting colormap selection. In *Proceedings Visualization '95*, pages 118–125, Atlanta, Georgia.

Brown, M. D., Greenberg, D., Keeler, M., Smith, A. R., and Yaeger, L. (1988). The visualization roundtable. *Computers in Physics*, 2(3):16–26.

Callaghan, T. C. (1984). Dimensional interaction of hue and brightness in preattentive field segregation. *Perception & Psychophysics*, 36(1):25–34.

Cambell, W. J., Short, Jr., N. M., and Treinish, L. A. (1989). Adding intelligence to scientific data management. *Computers in Physics*, 3(3):26–32.

Chan, K. C. C. and Wong, A. K. C. (1991). A statistical technique for extracting classificatory knowledge from databases. In Piatetsky-Shapiro, G. and Frawley, W. J., editors, *Knowledge Discovery in Databases*, pages 107–123. AAAI Press/MIT Press, Menlo Park, California.

CIE (1976). *CIE Publication No. 15, Supplement Number 2 (E-1.3.1): Official Recommendations on Uniform Color Spaces, Color-Difference Equations, and Metric Color Terms*. Commission Internationale de L'Èclairge.

Duncan, J. and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96(3):433–458.

D'Zmura, M. (1991). Color in visual search. *Vision Research*, 31(6):951–966.

Frawley, W. J., Piatetsky-Shapiro, G., and Matheus, C. J. (1991). Knowledge discovery in database: An overview. In Piatetsky-Shapiro, G. and Frawley, W. J., editors, *Knowledge Discovery in Databases*, pages 1–27. AAAI Press/MIT Press, Menlo Park, California.

Hallett, P. E. and Hofmann, M. I. (1991). Segregation of some mesh-derived textures evaluted by free viewing. *Vision Research*, 31:1701–1716.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on System, Man, and Cybernetics*, SMC-3(6):610–621.

Healey, C. G. (1996). *Effective Visualization of Large, Multidimensional Datasets.* Ph.D. thesis, The University of British Columbia, Canada.

Healey, C. G., Booth, K. S., and Enns, J. T. (1995). Real-time multivariate data visualization using preattentive processing. *ACM Transactions on Modeling and Computer Simulation*, 5(3):190–221.

Healey, C. G. and Enns, J. T. (1999). Large datasets at a glance: Combining textures and colors in scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, 5(2):145–167.

Healey, C. G., St. Amant, R., and Elhaddad, M. (1999). ViA: A perceptual visualization assistant. In *28th Workshop on Advanced Imagery Pattern Recognition (AIPR-99)*, Washington, DC.

Julész, B. (1984). A brief outline of the texton theory of human vision. *Trends in Neuroscience*, 7(2):41–45.

Julész, B. (April, 1975). Experiments in the visual perception of texture. *Scientific American*, pages 34–43.

Julész, B., Gilbert, E. N., and Shepp, L. A. (1973). Inability of humans to discriminate between visual textures that agree in second-order statistics—revisited. *Perception*, 2:391–405.

Julész, B., Gilbert, E. N., and Victor, J. D. (1978). Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, 31:137–140.

Kawai, M., Uchikawa, K., and Ujike, H. (1995). Influence of color category on visual search. In *Annual Meeting of the Association for Research in Vision and Ophthalmology*, page #2991, Fort Lauderdale, Florida.

Levkowitz, H. and Herman, G. T. (1992). Color scales for image data. *IEEE Computer Graphics & Applications*, 12(1):72–80.

Li, R. and Robertson, P. K. (1995). Towards perceptual control of Markov random field textures. In Grinstein, G. and Levkowitz, H., editors, *Perceptual Issues in Visualization*, pages 83–94. Springer-Verlag, New York, New York.

Liu, F. and Picard, R. W. (1994). Periodicity, directionality, and randomness: Wold features for perceptual pattern recognition. In *Proceedings 12th International Conference on Pattern Recognition*, pages 1–5, Jerusalem, Israel.

Pickett, R. and Grinstein, G. (1988). Iconographic displays for visualizing multidimensional data. In *Proceedings of the 1988 IEEE Conference on Systems, Man, and Cybernetics*, pages 514–519, Beijing and Shenyang, China.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1):81–106.

Rao, A. R. and Lohse, G. L. (1993a). Identifying high level features of texture perception. *CVGIP: Graphical Models and Image Processing*, 55(3):218–233.

Rao, A. R. and Lohse, G. L. (1993b). Towards a texture naming system: Identifying relevant dimensions of texture. In *Proceedings Visualization '93*, pages 220–227, San Jose, California.

Reed, T. R. and Hans Du Buf, J. M. (1993). A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, 57(3):359–372.

Rogowitz, B. E. and Treinish, L. A. (1993). An architecture for rule-based visualization. In *Proceedings Visualization '93*, pages 236–243, San Jose, California.

Rosenblum, L. J. (1994). Research issues in scientific visualization. *IEEE Computer Graphics & Applications*, 14(2):61–85.

Silbershatz, A., Stonebraker, M., and Ullman, J. D. (1990). The "Lagunita" report of the NSF invitational workshop on the future of database systems research. Technical Report TR-90-22, Department of Computer Science, University of Texas at Austin.

Smith, P. H. and Van Rosendale, J. (1998). Data and visualization corridors report on the 1998 CVD workshop series (sponsored by DOE and NSF). Technical Report CACR-164, Center for Advanced Computing Research, California Institute of Technology.

Stonebraker, M., Chen, J., Nathan, N., Paxson, C., Su, A., and Wu, J. (1993). Tioga: A database-oriented visualization tool. In *Proceedings Visualization '93*, pages 86–93, San Jose, California.

Tam, R., Healey, C. G., and Flak, B. (1997). Volume visualization of abdominal aortic aneurysms. In *Proceedings Visualization '97*, pages 43–50, Phoenix, Arizona.

Tamura, H., Mori, S., and Yamawaki, T. (1978). Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-8(6):460–473.

Thomson, K. A., Ingraham, W. J., Healey, M. C., LeBlond, P. H., Groot, C., and Healey, C. G. (1992). The influence of ocean currents on the latitude of landfall and migration speed of sockeye salmon returning to the Fraser River. *Fisheries Oceanography*, 1(2):163–179.

Thomson, K. A., Ingraham, W. J., Healey, M. C., LeBlond, P. H., Groot, C., and Healey, C. G. (1994). Computer simulations of the influence of ocean currents on Fraser River sockeye salmon (*oncorhynchus nerka*) return times. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(2):441–449.

Treinish, L. A. (1993). Unifying principles of data management for scientific visualization. In Earnshaw, R. and Watson, D., editors, *Animation and Scientific Visualization*, pages 141–170. Academic Press, New York, New York.

Treinish, L. A., Foley, J. D., Campbell, W. J., Haber, R. B., and Gurwitz, R. F. (1989). Effective software systems for scientific data visualization. *Computer Graphics*, 23(5):111–136.

Treinish, L. A. and Goettsche, T. (1991). Correlative visualization techniques for multidimensional data. *IBM Journal of Research and Development*, 35(1/2):184–204.

Triesman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics and Image Processing*, 31:156–177.

Ware, C. (1988). Color sequences for univariate maps: Theory, experiments, and principles. *IEEE Computer Graphics & Applications*, 8(5):41–49.

Ware, C. and Beatty, J. C. (1988). Using colour dimensions to display data dimensions. *Human Factors*, 30(2):127–142.

Ware, C. and Knight, W. (1995). Using visual texture for information display. *ACM Transactions on Graphics*, 14(1):3–20.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1(2):202–238.

Wolfe, J. M. and Franzel, S. L. (1988). Binocularity and visual search. *Perception & Psychophysics*, 44:81–93.

Wyszecki, G. and Stiles, W. S. (1982). *Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Edition.* John Wiley & Sons, Inc., New York, New York.

Ziarko, W. (1991). The discovery, analysis, and representation of data dependencies in databases. In Piatetsky-Shapiro, G. and Frawley, W. J., editors, *Knowledge Discovery in Databases*, pages 195–209. AAAI Press/MIT Press, Menlo Park, California.