Choosing Effective Colours for Data Visualization

Christopher G. Healey[1]

Department of Computer Science, University of British Columbia

ABSTRACT

In this paper we describe a technique for choosing multiple colours for use during data visualization. Our goal is a systematic method for maximizing the total number of colours available for use, while still allowing an observer to rapidly and accurately search a display for any one of the given colours. Previous research suggests that we need to consider three separate effects during colour selection: colour distance, linear separation, and colour category. We describe a simple method for measuring and controlling all of these effects. Our method was tested by performing a set of target identification studies; we analysed the ability of thirty-eight observers to find a colour target in displays that contained differently coloured background elements. Results showed our method can be used to select a group of colours that will provide good differentiation between data elements during data visualization.

CR Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces - *ergonomics, screen design (graphics, colour)*; I.3.6 [**Computer Graphics**]: Methodology and Techniques - *ergonomics, interaction techniques*

1 INTRODUCTION

Scientific visualization in computer graphics is a rapidly expanding area of research. This is due in large part to the dramatic increase in both the size and the number of datasets that need to be visualized [5, 18]. To date, many application-specific tools have been built to help analyse individual datasets. Much less work has focused on developing guidelines for the design of visualization techniques [17]. Our work is intended to address one aspect of this more general question.

A typical method of visualizing a dataset involves mapping data attributes to visual features (*e.g.,* shape, size, spatial location, and orientation). Colour is an important and frequently-used feature. Examples include colour temperature gradients on maps and charts, colour-coded vector fields in flow visualization, or colour icons displayed by real-time simulation systems.

If we use colour to represent our data, an important question to ask is: "How can we choose effective colours that provide good differentiation between data elements during the visualization task?" We address this problem by trying to answer three related questions:

- How can we allow rapid and accurate identification of individual data elements through the use of colour?
- What factors determine whether a "target" element's colour will make it easy to find, relative to differently coloured "non-target" elements?
- How many colours can we display at once, while still allowing for rapid and accurate target identification?

Previous work has addressed the issue of choosing colours for certain types of data visualization. For example, Ware and Beatty describe a simple colour visualization technique for displaying correlation in a five-dimensional dataset [22]. Research on the design of military systems sometimes quotes anecdotal evidence that suggests ". . .the general guideline for computer-generated images is no more than five to seven colours at a time. . ." [11], although they offer no explanation for why this might be the case. Robertson, Ware, Rheingans and Tebbs, and Levkowitz and Herman discuss various methods for building effective colour gamuts and colourmaps [15, 21, 14, 13]. Recent work at the IBM Thomas J. Watson Research Center has focused on a rule-based visualization tool which considers how a user perceives visual features like hue, luminance, height, and so on [16, 2].

None of these techniques were intended to investigate the rapid and accurate identification of individual data elements based on colour. Also, since the colour gamut and colourmap work uses continuous colour scales to encode information, they do not address the question of how many colours we can effectively display at once, while still providing good differentiation between individual data elements.

An intuitive first step to gaining more control over colour would be to use a perceptual colour model like CIE LUV, CIE Lab, or Munsell [23]. Unfortunately, fixing the colour distance in a perceptual colour model to a constant value does not guarantee that each colour will be equally easy to detect. Other factors can affect how groups of coloured elements interact with one another.

Our technique does, in fact, use the CIE LUV colour model to provide control over colour distance and isoluminance. We also exploit two specific results related to colour target detection: linear separation [7, 1] and colour category [12]. These effects are controlled to allow for the rapid and accurate identification of colour targets. Target identification is a necessary first step towards performing other types of exploratory data analysis. If we can rapidly and accurately differentiate elements based on their colour, we can apply our results to other important visualization techniques like detection of data boundaries, the tracking of data regions in real-time, and enumeration tasks like counting and estimation [20, 19, 9].

2 BACKGROUND

Before we describe our technique in detail, we provide a brief overview of the CIE LUV colour model, as well as a description of the linear separation and colour category effects.

[1] Department of Computer Science, Unviersity of British Columbia, 2366 Main Mall, Vancouver, British Columbia, V6T 1Z2, Canada

2.1 CIE LUV

The CIE LUV colour model was proposed by the Commission Internationale de L'Èclairge (CIE) in 1976 [6]. Colours are specified using the three dimensions L^* (which encodes luminance), u^*, and v^* (which together encode chromaticity). CIE LUV provides two useful properties for controlling perceived colour difference. First, colours with the same L^* are isoluminant. 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:

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

2.2 Linear Separation

The linear separation effect was originally described by D'Zmura [1991] . He was investigating how the human visual system finds a target colour in a sea of background non-target colours. D'Zmura ran experiments which asked observers to determine the presence or absence of an orange target. Two groups of differently coloured non-target elements were also present in each display (*e.g.,* in one experiment half the non-targets in each display were coloured green and half were coloured red). Results showed that when the target was linearly separable from its non-targets in colour space (Figure 1, target T and non-targets A and C), the time required to determine the target's presence or absence was constant, and independent of the total number of elements being displayed. This suggests detection occurs in parallel in the low-level visual system. When the target was collinear with its non-targets (Figure 1, target T and non-targets A and B), the time required to identify the target was linearly proportional to the number of elements being displayed. Observers had to search serially through each display to determine whether the target was present or absent.

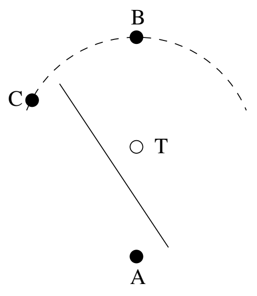


Figure 1: Example of a target with both collinear and separable non-targets, the colours are shown in a u^*, v^*-slice from the CIE LUV colour model, notice that target T is equidistant from all three non-targets A, B, and C; in Case 1, target colour T is collinear with the non-target colours A and B; in Case 2, target T is linearly separable from its non-targets A and C

Work by Bauer [1996] provides a number of results that strengthen D'Zmura's hypothesis. First, he showed that perceptual colour models cannot be used to overcome linear separation. In his experiments the target–non-target distance was fixed to a constant value in CIE LUV (Figure 1). In spite of this, the target required serial searching when it was collinear with its non-targets, but was significantly easier to detect when it was separable. Bauer replicated his initial findings in three additional colour regions: green, blue, and green-yellow. This suggests linear separation applies to colours from different parts of the visible colour domain.

2.3 Colour Category

Kawai [1995] reported results which suggest that the time required to identify a colour target depends in part on the named colour regions occupied by the target and its non-targets. If target identification depended only on the colour distance between a target and its non-targets, the time taken to detect presence or absence of the target should decrease uniformly as target–non-target distance increases. Kawai ran a set of experiments which showed this was not true.

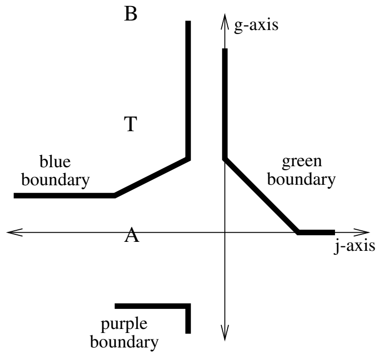


Figure 2: A target T and two non-targets A and B shown in a (j, g) slice from the OSA colour model; the boundaries of the blue, green, and purple colour regions are shown as thick lines; T and B occupy the same named colour region "blue", but A does not

Kawai tested displays which contained a uniquely colour target and a constant number of uniformly coloured non-targets. He divided an isoluminant colour slice used during his experiments into named colour regions. His results showed that search times decreased dramatically whenever the non-target was moved outside the target's colour region. For example, finding a target coloured T in a set of non-targets coloured B was significantly more difficult than than finding T in a set of non-targets coloured A (Figure 2). The target–non-target distances $\overline{TA}$ and $\overline{TB}$ are equal; Kawai suggests the difference in performance is due to the fact that both T and B are located in the blue colour region, but A is not. Similar search time asymmetries for separate experiments testing green, purple, and red targets were also explained by a difference in colour regions.

3 COLOUR SELECTION TECHNIQUE

Results discussed in the background section suggest that choosing effective colours for data visualization depends on at least three separate criteria:

- *colour distance:* the Euclidean distance between different colours as measured in a perceptually balanced colour model
- *linear separation:* the ability to linearly separate targets from non-targets in the colour model being used
- *colour category:* the named colour regions occupied by both the target and non-target elements

We began our investigation by controlling colour distance and linear separation, but not colour category. 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 visualization unfolds. Our requirement meant that the colour selection technique had to allow for rapid and accurate identification of any of the elements being displayed.

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:

- *three-colour study:* each display contained three different colours (*i.e.,* one colour for the target and two for the non-targets) which we named red, green-yellow, and purple-blue
- *five-colour study:* each display contained five different colours: red, yellow, green-yellow, blue, and purple
- *seven-colour study:* each display contained seven different colours: red, yellow, green-yellow, green, blue-green, purple, and red-purple
- *nine-colour study:* each display contained nine different colours: red, yellow-red, yellow, green-yellow, green, blue-green, purple-blue, purple, and red-purple

Every colour in a given study was tested as a target. For example, the three-colour study was run three times, first with a red element acting as a target (and green-yellow and purple-blue elements acting as non-targets), next with a green-yellow target (and red and purple-blue non-targets), and finally with a purple-blue target (and red and green-yellow 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 which 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

Colours for each of the four studies were chosen such that the colour distance between pairs of colours and the linear separation for each colour were fixed to constant values. Our results showed that detection was rapid and accurate for all colours from both the three-colour and five-colour studies. Results from the seven-colour and nine-colour studies were mixed; some colours gave better performance than others. This difference was explained when we examined the colour regions occupied by each of the colours.

3.1 Method

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 visualization tasks based on colour [4]. 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 3a).

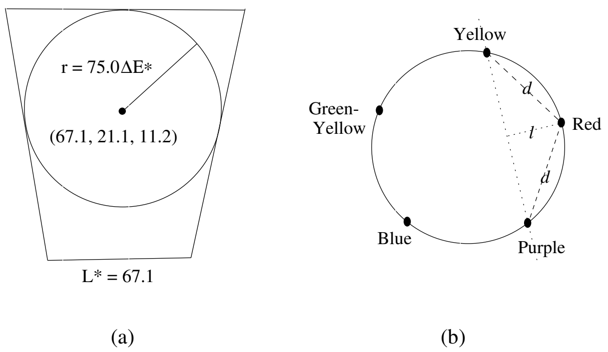


Figure 3: Choosing colours from the monitor's gamut: (a) the boundary of the gamut at $L^* = 67.1$, along with the maximum inscribed circle centered at $(L^*, u^*, v^*) = (67.1, 21.1, 11.6)$, radius 75.0 ΔE^*; (b) five colours chosen around the circle's circumference; each element has a constant colour distance d with its two neighbours, and a constant linear separation l from the remaining (non-target) elements

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 3b). This method ensured that neighbouring colours had a constant colour distance. It also ensured that any colour acting as a target had a constant linear separation from every other (non-target) colour. A similar technique was used to select colours for the three-colour, seven-colour, and nine-colour studies. This gave us the following colour distances d and linear separations l:

Study	d	l
three-colour	129.9 ΔE^*	112.5 ΔE^*
five-colour	88.2 ΔE^*	51.9 ΔE^*
seven-colour	65.1 ΔE^*	28.4 ΔE^*
nine-colour	51.3 ΔE^*	17.6 ΔE^*

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. A target block consisted of 360 displays testing a single colour target from one of the four studies. A total of 66 target blocks were completed. 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.

3.2 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 4a and 4b; note that the y-axis ranges from 400 msec to 800 msec in these graphs). 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.

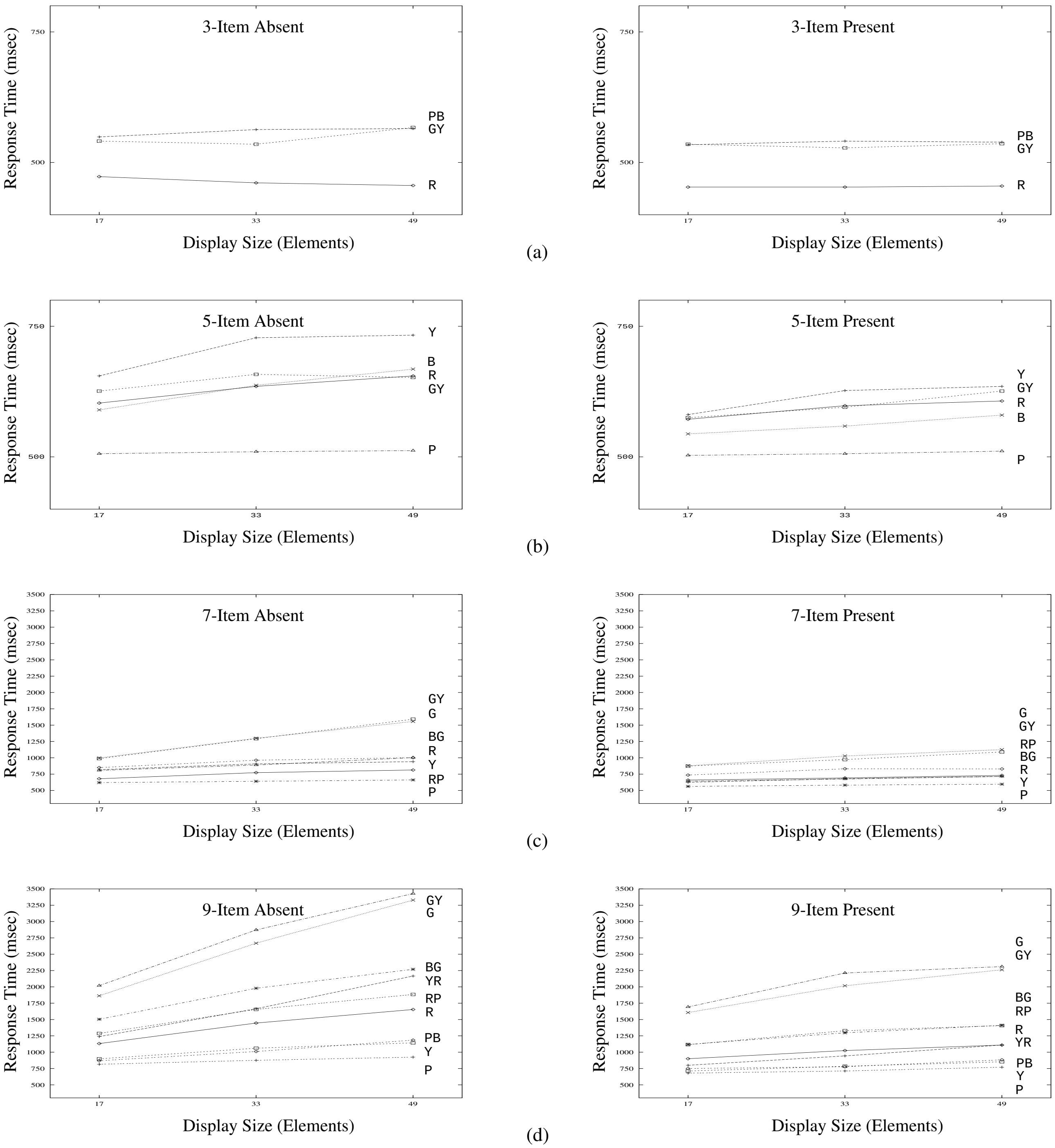


Figure 4: Response time graphs for the four studies, the graph on the left represents displays where the target was absent, while the graph on the right represents displays where the target was present: (a) response time as a function of display size (*i.e.,* total number of elements shown in the display) for each target from the three-colour study; (b) response times for each target from the five-colour study; (c) response times for each target from the seven-colour study; (d) response times for each target from the nine-colour study

Target identification became significantly more difficult for certain colours during the seven-colour and nine-colour studies (Figures 4c and 4d; note that the y-axis ranges from 100 msec to 3500 msec in these graphs). Mean response error during the seven-colour study was still low, at approximately 3.3%. The purple, yellow, red, blue-green, and red-purple targets all exhibited relatively flat response time graphs. Mean response time for these elements ranged from 611 msec to 870 msec. The green and green-yellow 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 green-yellow, and 17 msec per element for green) was roughly twice that for target-present displays (7 msec and 8 msec per element for green-yellow and green, 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 purple, yellow, and purple-blue targets (approximately 3.4%), and highest for the green, green-yellow, and red-purple targets (approximately 14%). The purple, yellow, and purple-blue targets displayed relatively flat response time graphs during target-present displays. The remaining targets showed some form of serial search. The effect was weakest for the red and yellow-red targets, and strongest for the green and green-yellow targets.

3.3 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 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.

In order to do this, we needed a method of placing individual colours within a named colour region. We used a simple technique described by Healey and Enns [10]. This technique places individual colours within user-named colour regions as follows:

- *Step 1:*
  Individual colours are mapped into the Munsell colour model. Colours are specified in the Munsell model using the three dimensions hue, chroma (which controls saturation), and value (which controls luminance) [3]; the hue dimension is divided into ten different 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 for short). A colour is assigned the name of its hue dimension in Munsell space.

- *Step 2:*
  The above technique is used to subdivide all or part of a colour space into named colour regions. We wanted to divide our maximum inscribed circle into named colour regions. Healey and Enns divided a similar circle at $L^* = 71.6$ (this corresponds closely to our circle from the $L^* = 67.1$ slice), as shown in Figure 5.

- *Step 3:*
  Observers are asked to name representative colours from each colour region. This converts Munsell names into user-chosen names; it also provides a way of measuring how strongly a particular name represents a given colour region. Healey and Enns used the colour at the middle of each colour region in Figure 5 to represent the region. Table 1 shows the names chosen for each of the ten representative colours. For example, 5.3% of the observers called the YR region's representative colour "red", 86.8% called it "orange", and 7.9% called it "brown".

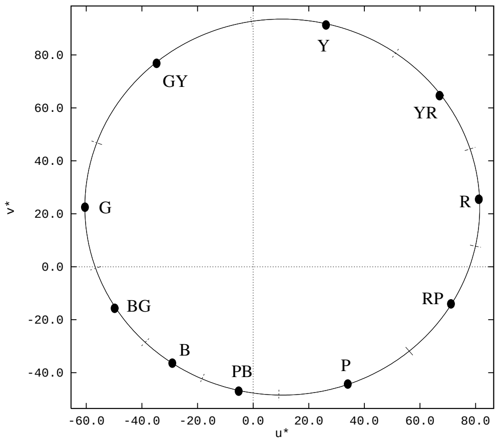


Figure 5: Example of segmenting the circumference of a maximum inscribed circle from a u^*v^*-slice through CIE LUV at $L^* = 71.6$; ticks along the the circle mark the boundaries between the ten named regions, points are the representative colours for each region

Each of the colours from our four studies was assigned a name using the method described in Step 1 above; this also placed each colour within one of the ten named colour regions. The user-chosen colour names shown in Table 1 allow us to measure the similarity of different colour regions. This similarity is dependent on:

- the range of user-chosen colour names assigned to a given region
- how "strongly" a region covers any user-chosen name in its range
- how the ranges of the two regions overlap

As an example, consider the P and R regions, which have ranges:

	purple	*magenta*	*pink*	*red*
P	86.9%	2.6%	5.2%	
R			26.3%	71.0%

Colours from P and R overlap only at the "pink" name. Their overlap is not that strong, since neither P nor R are strongly classified as pink. The amount of overlap is computed by multiplying the percentages for the common name. This gives a P–R overlap of $5.2\% * 26.3\% = 0.014$. A closer correspondence of user-chosen names for a pair of regions results in a stronger similarity, for example, G and GY have an overlap of $97.3\% * 100.0\% = 0.973$. Colours that overlap over multiple names are combined using addition, for example, YR and Y overlap in both the "orange" and the "brown" name, giving a YR-Y overlap of $(86.8\% * 2.6\%) + (7.9\% * 44.7\%) = 0.058$. These overlap values can be used to explain why certain colours were difficult to use for target identification.

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 Table 2 suggests the following explanation:

	purple	*magenta*	*pink*	*red*	*orange*	*brown*	*yellow*	*green*	*aqua*	*blue*	*other*
P	86.9%	2.6%	5.2%	–	–	–	–	–	–	–	5.2%
RP	15.7%	28.9%	55.3%	–	–	–	–	–	–	–	–
R	–	–	26.3%	71.0%	–	–	–	–	–	–	2.6%
YR	–	–	–	5.3%	86.8%	7.9%	–	–	–	–	–
Y	–	–	–	–	2.6%	44.7%	47.4%	–	–	–	5.2%
GY	–	–	–	–	–	–	–	97.3%	–	–	2.6%
G	–	–	–	–	–	–	–	100.0%	–	–	–
BG	–	–	–	–	–	–	–	26.3%	57.8%	15.8%	–
B	–	–	–	–	–	–	–	–	7.9%	89.4%	2.6%
PB	5.2%	–	–	–	–	–	–	–	–	92.1%	2.6%

Table 1: Responses for the representative colours from the ten named regions, shown as the percentage of observers who chose a given name

- targets R, Y, P, and RP have a weak similarity to the other colours used during the study; Y has no similarity to the other colours, R and RP have a similarity (*i.e.,* an R-RP overlap) of 0.145, and P and RP have a similarity of 0.173

- target BG has a moderate similarity to GY and G (0.256 and 0.263, respectively)

- targets G and GY have a moderate similarity to BG, and a very strong similarity (0.973) to one another

For each target in the seven-colour study, an increase in its similarity to the other colours corresponded closely to an increase in its mean response time. We can measure this correspondence by computing Spearman's correlation coefficient on the rank order of our colours in terms of total similarity and mean response time.

	R	Y	GY	G	BG	P	RP
R	–	0	0	0	0	.014	.145
Y	0	–	0	0	0	0	0
GY	0	0	–	.973	.256	0	0
G	0	0	.973	–	.263	0	0
BG	0	0	.256	.263	–	0	0
P	.014	0	0	0	0	–	.173
RP	.145	0	0	0	0	.173	–
Total	.159	0	1.229	1.236	.519	.187	.318

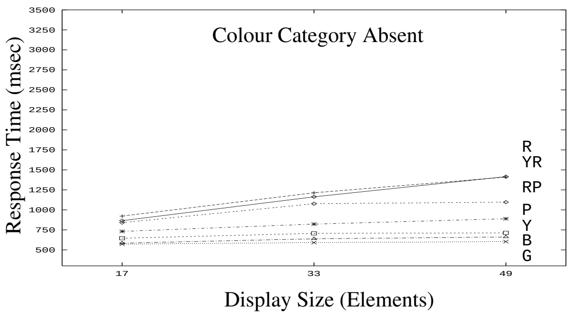
Table 2: Similarity table, showing individual and total similarity values for each of the colours used during the seven-colour study

Table 2 sums the similarity measures for each of the seven colours we used. The rank order of total similarity from lowest to highest is Y, R, P, RP, BG, GY, and G. Ranking our colours based on total mean response time (including both present and absent trials across all three display sizes) from smallest to largest gave an order of P, R, Y, BG, RP, GY, and G. The Spearman correlation between these rankings is $r = 0.821$, confirming that higher mean response times for a given target correspond to a higher colour category similarity.

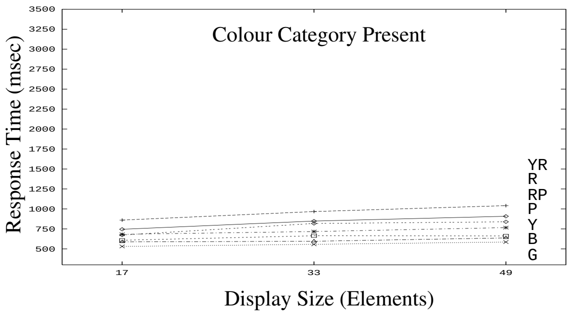
A similarity–response time correlation can also be observed during the nine-colour study. Targets P, Y, YR, and PB had a weak similarity to other colours used during the study. Targets R, RP, and BG had a moderate similarity to each other and to G and GY. Targets G and GY had a moderate similarity to BG, and a strong similarity to one another. The Spearman correlation between the colours' total similarity and mean response time ranking was $r = 0.762$. Results for YR and BG are somewhat anomalous; similarity values suggest we should have observed better performance for the YR target, and worse performance for the BG target. Results for the other colours correspond closely to their similarity measures.

3.4 Colour Category Study

Results from investigating colour category integration might imply that effective colours can be selected by controlling colour category alone. This was tested by selecting seven colours from colour regions that had low similarity with one another. An examination of observer responses in Table 1 shows that colours chosen from the R, YR, Y, G, B, P, and RP regions satisfy our restriction; the largest similarity overlap occurs between RP-R (0.145) and RP-P (0.173). The colours we used were exactly those chosen by Healey and Enns to represent the above seven regions. A single observer (the author) completed 360 displays for each target (we address the use of only one subject following the description of our experiment results).



(a)



(b)

Figure 6: Response time graphs for the first colour-category study: (a) target absent displays; (b) target present displays

Mean response error was 3.8%, which shows responses were accurate. Although response times are somewhat better than those from the original seven-colour study, several colours (in particular

R and YR) still exhibit poor search performance. This can be explained by examining the distance between neighbouring pairs of colours, and the linear separation for each colour when it acted as a target (Table 3). Colours that gave the worst search performance (R, YR, and RP) had the smallest neighbour distances and linear separation; in fact, the 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 (G, B, and Y) had the largest neighbour distances and linear separation.

	CW d	CCW d	l
R	42.0 ΔE^*	42.0 ΔE^*	12.2 ΔE^*
YR	42.0 ΔE^*	49.8 ΔE^*	14.6 ΔE^*
Y	49.8 ΔE^*	112.6 ΔE^*	39.0 ΔE^*
G	112.6 ΔE^*	67.5 ΔE^*	52.9 ΔE^*
B	67.5 ΔE^*	63.2 ΔE^*	29.1 ΔE^*
P	63.2 ΔE^*	48.6 ΔE^*	20.4 ΔE^*
RP	48.6 ΔE^*	42.0 ΔE^*	14.2 ΔE^*

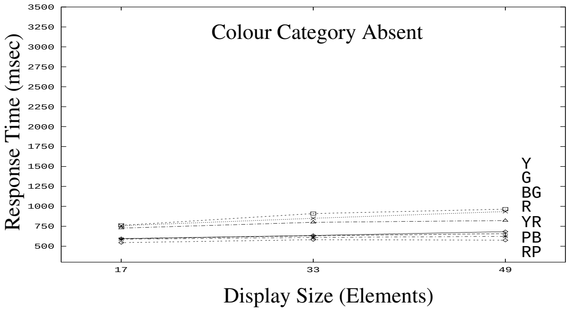
Table 3: The exact distance in CIE LUV between a colour, its clockwise neighbour (CW d), and its counter-clockwise neighbour (CCW d), as well as the linear separation from the other colours when it acted as a target

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 have an affect on search performance. We used a more systematic colour selection technique to choose another set of seven colours:

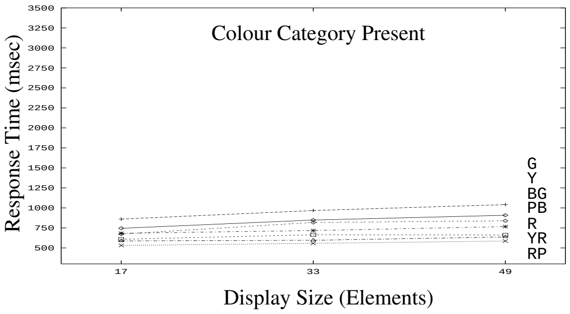
- a single green was chosen from the combined G-GY colour region; observers see the entire region as green, which means it can be used for only a single colour
- the clockwise neighbour of our green was a yellow, chosen to lie on the border between the GY and Y colour regions
- the counterclockwise neighbour of our green was a blue-green, chosen to lie in the center of the BG colour region (we did not use the colour from the border between G and BG, because it was too difficult to differentiate from our green)
- the remaining four colours (chosen from the YR, R, RP, and PB colour regions) were at equal steps between our yellow and our blue-green

This gave us a constant neighbour distance d and linear separation l ($59.4\ \Delta E^*$ and $24.6\ \Delta E^*$, respectively) between the Y, YR, R, RP, PB, and BG colours (G had a larger d and l than the other colours). Results from displays using these colours as targets are shown in Figure 7. 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 still gave mixed results during target-absent displays. We could have further differentiated these elements by choosing a yellow from the center of the Y colour region (rather than at the GY-Y border). This might have resulted in poorer performance for the other targets due to the reduction in colour distance and linear separation between our Y, YR, R, RP, PB, and BG colours, however. 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.

Although it would have been preferable to test additional subjects, using only one observer does not invalidate the generality of our results. Our experiments tested the ability of the low-level visual system to identify target colours in parallel; numerous experiments in cognitive psychology have shown that the low-level visual system functions in a similar manner across subjects. This does not preclude the ability of an observer to learn and improve their performance for a given task. However, the subject who completed the colour category experiment (the author) was well versed in the particular task being investigated. The low-level visual system was capable of detecting any of the seven colours used during the second colour category experiment in times independent of the display size. This was not the case during the first colour category experiment. The visual system had to resort to serial scanning to detect the R, YR, and RP targets. Moreover, the experience of the subject who completed the experiment suggests the results would not benefit from additional "practice" sessions.



(a)



(b)

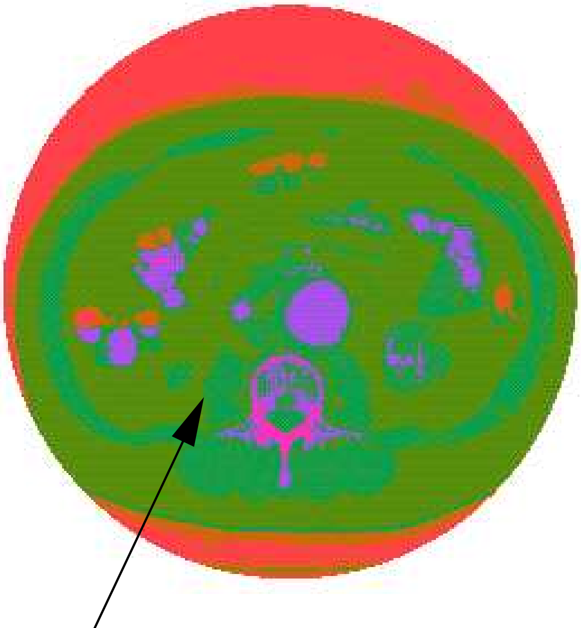
Figure 7: Response time graphs for the second colour-category study: (a) target absent displays; (b) target present displays

4 VISUALIZING MEDICAL IMAGES

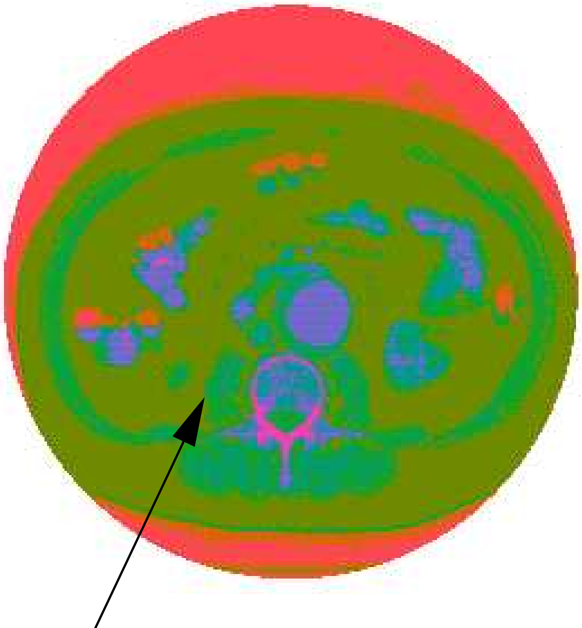
As we discussed in the introduction, our long-term goal is a set of guidelines that help users maximize the total number of colours available during visualization, while still allowing for the rapid and accurate detection of any one of the given colours. Experimental results showed how we can control colour distance, linear separation, and colour category to pick seven isoluminant colours that satisfy our requirements. Extending these results to a practical application is a non-trivial problem. However, even the specific results reported here can help to design useful visualization applications. To this end, we discuss how colours from our experiments were used to visualize slice data from a Computerized Tomography (CT) scan. The scan was taken to locate an aneurism in a patient's brain. This resulted in 116 individual slices, each of which had a 512×512 resolution.

We began by displaying slices as greyscale images (Figure 8a); values in each slice had been scaled to range from $(0 \dots 255)$ for

(a)



(b)



(c)

Figure 8: Examples of a single CT slice: (a) the slice displayed using a greyscale intensity ramp; (b) colours from the original seven-colour experiment, although the location and borders of each region are clearly visible, the makeup of certain regions is difficult to determine, compared to; (c) colours from the final seven-colour experiment, which show clearly that the region referenced by the arrow is made up of two different groups of elements

exactly this purpose. Higher values in a slice correspond to denser material detected during the CT scan. Greyscale images are a standard method for displaying medical slice data [21, 2], since they allow a user to easily locate individual regions of similar density. Through consultation we identified seven intensity ranges that corresponded to seven individual regions of interest. Slices were then viewed using seven different colours. Figure 8b displays the slice using colours from our original seven-colour experiment. Figure 8c displays the slice using colours from our final colour category experiment.

There are three important points to note. First, as with the greyscale representation, it is easy to locate individual regions in both colour slices. The densities (and therefore the intensity values) of different regions of interest are consistent across slices. Therefore, our colour displays work well for any individual slice chosen by the user.

Second, as predicted by our experiment results, the colours used in Figure 8b do not always provide good differentiation between regions, compared to the colours used in Figure 8c. Consider the region reference by the arrow. In Figure 8b, it appears to made up of a single type of element. In fact, it is made up of two types of elements; this is clear when examining the same region in Figure 8c. Notice there is no corresponding tradeoff between Figure 8c and Figure 8b, that is, there are no pairs of elements that are difficult to differentiate in Figure 8c but easy to differentiate in Figure 8b.

Finally, colour images can provide visual cues that might not be immediately apparent in a greyscale display. For example, consider the small spots of purple at the top of the region in the center of the image. These spots represent areas of high relative density, and are quickly identified as different from the surrounding material. The same information is present in the greyscale image, but the visual system is not immediately drawn to it, since the contrast between the region and its neighbours is not as pronounced. One method of visualizing slice data is to "fly" through the slice stack, rapidly displaying individual slices one after another in a movie-like fashion [8]. In this context, we expect users could more accurately detect areas with a specific density range using colour rather than greyscale.

5 DISCUSSION

This paper has presented a simple method for choosing effective colours for use during data visualization. We accomplish this by controlling colour distance, linear separation of colours, and colour category similarity during colour selection. Results from target identification studies show that all three of these effects need to be considered in order to ensure that the colours used provide consistently good differentiation from one another.

We found that an isoluminant slice through a monitor's gamut yielded up to seven different colours, any one of which could be rapidly and accurately detected, even in the presence of all the others. Each colour needed a minimum colour distance, linear separation, and colour category differentiation to guarantee good performance. One of our future goals is a set of investigations that explicitly measure the tradeoff in observer performance when these effects are varied. This will also allow us to identify the minimum required value for each effect.

The perceived intensity of each colour was held constant during our studies to eliminate the potential for visual interference. It might be possible to use intensity in a controlled manner to increase the number of colours we can display simultaneously. For example, we might be able to choose multiple groups of colours, with each group being restricted to an isoluminant slice through the monitor's gamut. A colour group would be assigned to a conceptually related set of attribute values in the dataset to be visualized. Proper division by intensity might allow a user to select individual colours either from within a group (*e.g.,* a dull red element or a bright green element) or across groups (*e.g.,* all red elements). This type of technique is not possible if colours of varying intensity are chosen from the monitor's gamut in an ad hoc fashion.

References

[1] BAUER, B., JOLICOEUR, P., AND COWAN, W. B. Visual search for colour targets that are or are not linearly-separable from distractors. *Vision Research* (1996), (in press).

[2] BERGMAN, L. D., ROGOWITZ, B. E., AND TREINISH, L. A. A rule-based tool for assisting colormap selection. In *Proceedings Visualization '95* (Atlanta, Georgia, 1995), pp. 118–125.

[3] BIRREN, F. *Munsell: A Grammar of Color*. Van Nostrand Reinhold Company, New York, New York, 1969.

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

[5] CAMBELL, W. J., SHORT, JR., N. M., AND TREINISH, L. A. Adding intelligence to scientific data management. *Computers in Physics 3*, 3 (1989), 26–32.

[6] CIE. *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, 1976.

[7] D'ZMURA, M. Color in visual search. *Vision Research 31*, 6 (1991), 951–966.

[8] HEALEY, C. G., BOOTH, K. S., AND ENNS, J. T. Real-time multivariate data visualization using preattentive processing. *ACM Transactions on Modelling and Computer Simulation 5*, 3 (1995), 190–221.

[9] HEALEY, C. G., BOOTH, K. S., AND ENNS, J. T. High-speed visual estimation using preattentive processing. *ACM Transactions on Computer-Human Interaction* (1996), (in press, scheduled for July 1996 issue).

[10] HEALEY, C. G., AND ENNS, J. T. A perceptual colour segmentation algorithm. Tech. Rep. TR-96-xx, Department of Computer Science, University of British Columbia, 1996.

[11] HITT, E. F. Retrofit of flat panel color displays in the F-16A/B. In *Proceedings IEEE/AIAA 11th Digital Avionics Systems Conference* (Seattle, Washington, 1992), pp. 323–328.

[12] KAWAI, M., UCHIKAWA, K., AND UJIKE, H. Influence of color category on visual search. In *Annual Meeting of the Association for Research in Vision and Ophthalmology* (Fort Lauderdale, Florida, 1995), p. paper #2991.

[13] LEVKOWITZ, H., AND HERMAN, G. T. Color scales for image data. *IEEE Computer Graphics & Applications 12*, 1 (1992), 72–80.

[14] RHEINGANS, P., AND TEBBS, B. A tool for dynamic explorations of color mappings. *Computer Graphics 24*, 2 (1990), 145–146.

[15] ROBERTSON, P. K. Visualizing color gamuts: A user interface for the effective use of perceptual color spaces in data displays. *IEEE Computer Graphics & Applications 8*, 5 (1988), 50–64.

[16] ROGOWITZ, B. E., AND TREINISH, L. A. An architecture for rule-based visualization. In *Proceedings Visualization '93* (San Jose, California, 1993), pp. 236–243.

[17] ROSENBLUM, L. J. Research issues in scientific visualization. *IEEE Computer Graphics & Applications 14*, 2 (1994), 61–85.

[18] TREINISH, L. A., FOLEY, J. D., CAMPBELL, W. J., HABER, R. B., AND GURWITZ, R. F. Effective software systems for scientific data visualization. *Computer Graphics 23*, 5 (1988), 111–136.

[19] TRIESMAN, A. Search, similarity, and integration of feature between and within dimensions. *Journal of Experimental Psychology: Human Perception & Performance 17*, 3 (1991), 652–676.

[20] VAREY, C. A., MELLERS, B. A., AND BIRNBAUM, M. H. Judgments of proportions. *Journal of Experimental Psychology: Human Perception & Performance 16*, 3 (1990), 613–625.

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

[22] WARE, C., AND BEATTY, J. C. Using colour dimensions to display data dimensions. *Human Factors 30*, 2 (1988), 127–142.

[23] WYSZECKI, G., AND STILES, W. S. *Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Edition*. John Wiley & Sons, Inc., New York, New York, 1982.