Alpha, contrast and the perception of visual metadata

Maureen Stone, StoneSoup Consulting and Lyn Bartram, Simon Fraser University

Abstract

Visual elements such as grids, labels, and contour lines act as "reference structures" or "visual metadata" that support the primary information being presented. Such structures need to be usefully visible, but not so obtrusive that they clutter the presentation. Our goal is to determine the physical, perceptual and cognitive characteristics of such structures, ideally in a way that enables their automatic computation. We present the result of a set of experiments to determine effective display ranges, described in terms of transparency (alpha), for thin rectangular grids over scatterplot data. These show that an effective range can be defined in terms of alpha. In an effort to create a display-independent set of metrics, we analyze these results in terms of luminance contrast, with mixed results. We conclude that the appearance of transparency is an important aspect of subtle visualization.

Introduction

Visual elements such as grids act as *reference structures* or *visual metadata* that support the primary information being presented. Such structures need to be usefully visible, but not so obtrusive that they clutter the presentation. Other static examples include labels and contour lines. Interactive techniques like smart cursors and object handles also create reference structures.

Visual designers expertly manipulate properties such as color, line weight and transparency to create a balance between reference structures and the critical data. The broad goal of our research is to create engineering metrics and models that enable dynamic, algorithmically generated displays to be similarly effective.

Our approach to this problem is not to characterize "ideal" or "best," but instead to define boundary conditions, outside of which the result is clearly bad. We reason that the best solution will always be influenced by both context and taste. Boundary conditions, however, are more likely to have simple rules that can easily be incorporated by engineers and researchers. By eliminating, or at least reducing, the most objectionable cases, we can more easily raise the overall quality of computer-generated presentations.

This paper will summarize the results from a first set of experiments to characterize the boundary conditions for rectangular grids. In these experiments, the subjects manipulated the transparency (alpha) of thin-line grids overlaid on scatterplots of different complexities rendered on backgrounds of different lightnesses (all grayscale imagery). The goal was to find a range of alpha values that create acceptably subtle grids.

Our results show that a statistically acceptable range of alpha values can be established for our experimental conditions, which used the same calibrated display for all subjects. To create a display-independent model, we need to tie our results to perceptual metrics such as luminance contrast, which is used, for example, to provide metrics for text legibility. Our analysis shows that contrast alone is insufficient to explain our results, suggesting that the degree of transparency may be the more critical metric.

This paper will first discuss reference structures from a design perspective, then summarize our experiments and their results. We then provide an analysis of our results in terms of contrast. We conclude with our directions for future work, focusing on image complexity and further explorations of transparency metrics.

Subtle Design

Designers create subtle reference structures by vary visual contrast, typically manipulating color, line weight and transparency [1]. Figure 1 shows a grid overlaid on a map. The lines that define the grid in Figure 1(b) appear lighter (actually, more *transparent*) and are thinner than those in Figure 1(a), resulting in a more subtle appearance. The overall goal of the designer is to achieve a well-balanced composition of visual layers, in which whatever constitutes the "figure" is well defined with respect to "ground". Grids and other visual metadata live somewhere in the middle of these layers, where sometimes the grid needs to be more *figure* (visually accessible for search or reference) and sometimes more *ground* (relegated to the background and not intrusive.)

Figure 1(a). *A badly designed grid that obscures the underlying information.*

Figure 1 (b). *This grid is more subtle, allowing the viewer to focus on the map.*

We seek a way to characterize these design decisions in terms of quantities that can easily be computed from computer displays. That is, we seek metrics for "legible, but not obtrusive."

Contrast, which is a function of both relative lightness and spatial frequency, has long been used to predict text and symbol legibility [2]. A recent paper on text legibility [3], repeated some of this work using text on web pages. In it, they propose the use of ΔL^* as a legibility metric. Their threshold (between 25 and 30) corresponds to a Michelson contrast of around 50%, or a contrast ratio of 3:1, which are generally accepted as the minimum contrast needed for text legibility. ΔL^* is an attractive metric for design both because of its integration with design software such as Adobe Photoshop®, and because it offers a potentially straightforward way to include color contrast.

The use of transparency is also critical to subtle design. Because the grids in Figure 1 are transparent, they are always darker than the background. Making the grid a constant light gray, for example, would create a structure that alternated between being lighter and darker than the underlying graphic. This makes the grid itself less coherent.

The perception of transparency, which enables the visual system to separate a scene into overlapping layers, is clearly important to this work. Metelli's model of a spinning disk [4], whose open sectors blend the light reflecting from the background with the light reflecting from the disk's surface, seems a direct analogy to the computer graphics alpha blending used in our experiments. There is extensive literature on transparency, including [5,6,7], whose relevance we are only beginning to explore.

The Experiments

In two experiments we investigated the effects of grid colour, image background and complexity on these boundary condition levels. We used the same method and metrics in all experiments. Participants were asked to adjust the alpha value of a grid with a constant line weight of one pixel over a set of images with different background colors (gray values), and different levels of visual complexity. We chose alpha as our control because this is how experienced visual designers design grids, and because it is a parameter common to most modern display controllers. Alpha is simply a linear blending between foreground and background pixel values, where 1.0 is opaque and 0 is invisible.

Using only one variable enables a relatively simple interaction based on the motion of the mouse. Holding down the left mouse button increased the strength of the grid (increased alpha); holding down the right button made the grid fainter (decreased alpha).

We set the participants two different tasks. The first was to specify the point where "the grid is usably perceptible without being unnoticeable "(*faint grid*). The second was to adjust the grid "to meet your best judgment of how obvious it can be before it becomes too intrusive and sits in front of the image; some users have called this a fence" (*strong grid*). This terminology came from observations of previous participants in pilot studies.

The participants performed the tasks as two separate tests; that is, they did all of one task on all of the images, then the other task. Participants could practice on a set of training images for an unlimited time, though all users were comfortable with both tasks after a few practice images. All users performed the experiments on the same, calibrated display under the same viewing conditions.

Experimental design

We created four images of varying complexity: a flat field and three scatter plots at different levels of density: sparse, medium and dense (Figure 2). The images were designed to cover progressively more of the background so we could explore whether contrast with the background was an important factor in setting the boundaries.

In the first experiment, the subjects manipulated black grids over relatively light backgrounds (Figure 3). In a second experiment, they manipulated white grids over dark backgrounds (Figure 4). Grid spacing was fixed at 86.5 pixels in x and 118 in y to align with the x and y axis values in the scatterplots.

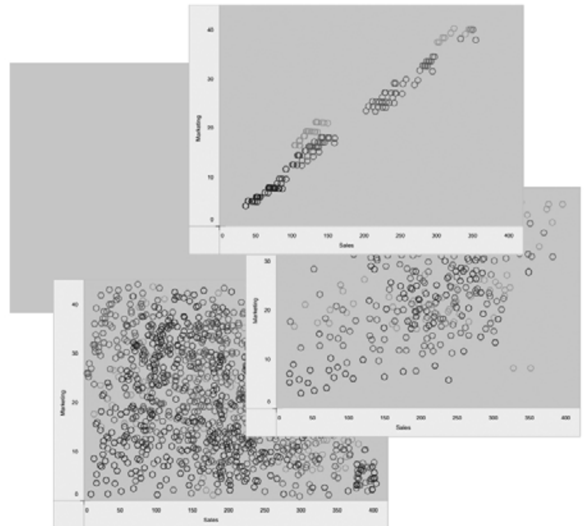
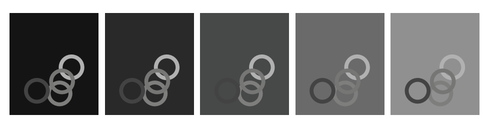


Figure 2. *Images were generated at 4 densities: flat, sparse, medium and dense. These increased both in visual complexity, and in the amount of background that was covered.*

The gray values of the foreground circles in the plots were chosen to supply some visual variety across a range of lightness levels, and to be visibly different from all of the different background levels.

L* value	96	87	78	69	60
PC pixel	243	217	191	167	143
Mac pixel	241	209	180	152	126

Figure 3. *Background gray values, Experiment 1*



L* value	4	15	30	45	60
PC pixel	23	42	72	106	143
Mac pixel	14	28	55	87	126

Figure 4. *Background gray values, Experiment 2.*

Each plot was rendered as a jpg image and displayed at a spatial resolution of 800x600 pixels on an Apple Cinema liquid crystal Display. All experiments were carried out on the same display in the same dimly lit room. The display was calibrated using a Gretag Eye-One Pro (10 nm resolution) and Gretag's profiling package to achieve a gamma of 1.8 (typical for Macintosh systems) and the native display colour temperature. Subjects sat 59 cm (24") from the display, which had a spatial resolution of 1920x1200 across a 23" diagonal viewing area. All grids displayed were a single pixel wide, or 1.5 minutes of arc.

Experimental method

For experiment 1, a 4(complexity) x 5(background) factorial design yielded 20 experimental conditions. Each subject performed two separate task blocks, one for the strong boundary, and one for the faint. Each task block had 3 repetitions of 20 images resulting in 60 trials/block.

Experiment 2 was similar to experiment 1, except the subjects manipulated a white grid over a dark plot. A 4 (complexity) x 5 (background) factorial design yielded 20 experimental conditions. Each grid task block had 3 repetitions of the 20 conditions resulting in 60 trials/block.

In all cases, the trial ordering was randomized and block ordering was counterbalanced. All subjects were university students that had normal or corrected to normal vision and were paid. The first experiment had 12 subjects, and the second had 15. No subject participated in more than one experiment.

Hypotheses

We had several hypothesises in these experiments. The first four refer to each individual experiment. The last refers to commonality between the experiments.

- H1. The faint boundary for the usable grid would show less variation than the "fence" setting.
- H2: Alpha for the faint setting would be less than 0.5.
- H3: Background would have an effect on alpha settings in both the faint and strong cases.
- H4: image density would affect alpha settings in both the faint and strong cases.
- H5: Results would be symmetric for the light and dark grids.

Results

The results for the first experiment are summarized in Figure 5, which shows the range between the average faint results and the average strong boundaries, separated by the four levels of image density. As can be seen from the figure, most subjects found the grid to be usably legible at very light alpha values, even for a complex image. H2 was confirmed; we were surprised by how low the faint setting actually was. Even accommodating the more variable strong settings, the grid was still considered usable in all densities at around 0.2 alpha. Refuting H3, background lightness was not significant in either the faint or the strong setting. Density, on the other hand, did have a significant effect, confirming H4: F(3,228) =60.0112, p<.0001 (faint); F(3,228) = 11.9789, p<.0001 (strong). The range defined by our boundary conditions, which is plotted offset by the minimum alpha, increases with complexity, as does the minimum usable alpha for the dense image. Finally, as we expected (H1), there is more variability in the strong results than in the faint, although the patterns are still obvious.

The results for experiment 2 (light on dark) are shown in Figure 6. Again we see that there is more variability in the strong than the faint grid (H1). The overall statistics were noisier also, especially for the strong boundary, as can be seen in Figure 7. Once again, the faint setting was lower than we expected but well within the expected range (H2). However, in this experiment, unlike experiment 1, alpha showed a statistically significant dependence on the background (H3). Both density [F(3,332=87.82,p<0.1] and background [F4(4,232)=14.3,p<0.1] had significant effects and there was an interaction between them: [F(12,145)=2.28,p<0.1].

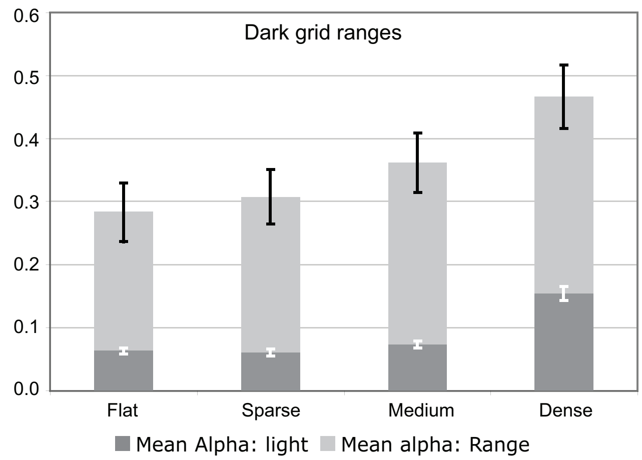


Figure 5. *Mean faint alpha and range for Exp. 1, showing that there is no statistically significant dependency on background, only on density. Error bars show the 95% confidence intervals.*

As one might expect from previous experience, this was mostly due to the dense plot. When we removed this case from analysis, we still see the same effects and an interaction. Density, however, is no longer a very large effect, leaving background as the dominant effect, though neither is very large. Against our expectations, H5 was not borne out. Light grids on a darker background are different from dark grids on a lighter background.

These results have practical implications, especially since we are looking for a "safe" range. For three out of the four cases, a light but useful grid could be created with an alpha value around 0.1, and in all cases, an alpha value of 0.2 falls in the "not bad" range. This is much lighter and more subtle than the solid black grid (alpha=1.0) common in many visualization systems and technical illustrations. However, these results are specified in terms of alpha, which a display-specific variable. We want to tie them to a perceptual metric such as luminance contrast.

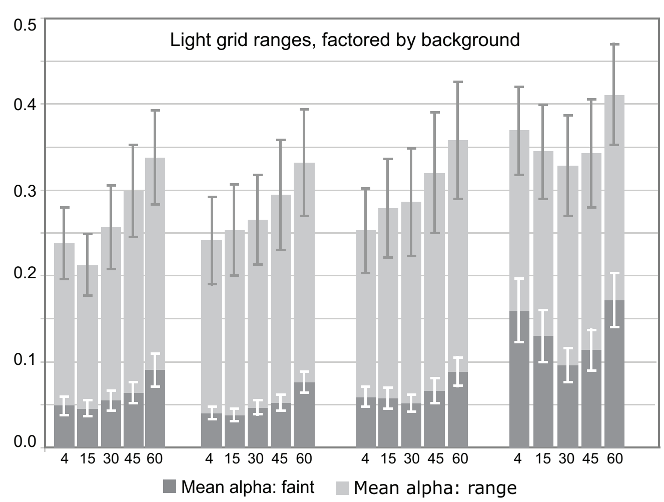


Figure 6. *Mean faint alpha and range for Exp. 2, showing that there is a small but significant dependency on background for this case.*

Alpha and Contrast

We expected that contrast of the grid with the background would correlate well with our boundaries, at least for the flat and sparse cases, where the background dominates the visual field. Luminance contrast is a perceptual measure that defines the difference between the perceived lightness of two colors. Contrast, more broadly, can include differences in hue, which is not addressed in this work but will be of eventual interest.

Luminance contrast is often used to specify legibility thresholds for text and small symbols. It can be specified as a luminance ratio, or in terms of Weber or Michelson contrast [8]. A commonly stated threshold for minimum text legibility is 3:1, which is equivalent to a Michelson contrast of 50%, where Michelson contrast is defined in terms of luminance values (Y) as:

$$C_m = (Y_{max} - Y_{min})/(Y_{max} + Y_{min})$$

The faint boundary in our experiments is a legibility boundary, so we would expect that contrast would be a useful metric for this case. To test this, we computed the effective luminance of the line, and then computed the contrast with the background.

Alpha blending is a simple linear interpolation between a foreground and background pixel. This is usually written:

$$c = \alpha f+(1-\alpha)b$$

where f is the foreground pixel value (ranging from 0 to 1), b is the background pixel value, and α ranges from 0 to 1. An α value of 1 is opaque (equals f), and a value of 0 is fully transparent, rendering the background colour (b). Since all of our colors are gray, we can approximate luminance simply as: $Y(c) = p^{1.8}$, then apply the formula for Michelson contrast above.

Both alpha and contrast are relative metrics, so we would expect them to be similarly independent of background. In Figure 7, the raw data values are plotted vs. the background L* value. The trendlines show the mean and the slope the variation with respect to the background. Alpha is relatively flat for both the faint and strong boundaries, as discussed above. For the dark on light case, Michelson contrast is similarly flat. However, for the light grid data, it depends strongly on the background, which was a surprise.

Looking a bit harder at the relationship between alpha and luminance, we see that for a black grid (f=0), $c = (1-\alpha)b$. Working through the mathematics, we get a formulation for contrast that can be stated entirely in terms of alpha:

$$C_m = (b^\gamma - (1-\alpha)^\gamma b^\gamma)/ (b^\gamma + (1-\alpha)^\gamma b^\gamma)$$
$$C_m = (1 - (1-\alpha)^\gamma)/ (1+ (1-\alpha)^\gamma)$$

Put another way, a black grid is simply reducing the brightness of the background color as a function of alpha, whereas for any other color of grid, it is a blend of the foreground and background.

Similar results can be found for Weber contrast, both in the experimental data, and in its expression in terms of alpha:

$$C_w = 1 - Y_f/Y_b$$
$$C_w = 1 - ((f-b)\alpha + b)^\gamma/b^\gamma$$
$$C_w = 1 - ((f-b)\alpha/b + 1)^\gamma$$

If f=0 (black): $C_w = 1 - (1-\alpha)^\gamma$

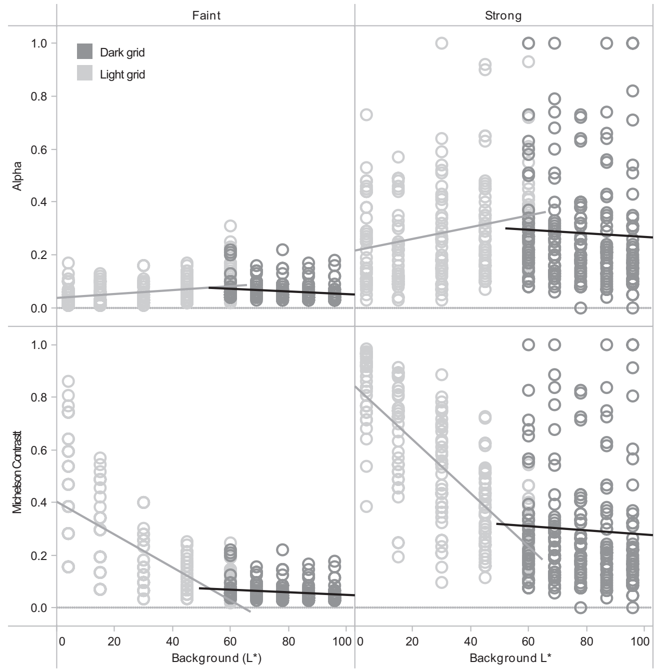


Figure 7. Trend lines indicating the mean value for alpha and Michelson contrast for each experiment, showing both faint and strong cases. Dark grid results are in dark gray, light grid results in light gray.

We also explored ΔL* as a form of contrast metric. Figure 8 shows ΔL* to be relatively independent of background for the faint boundary, but not for the strong one.

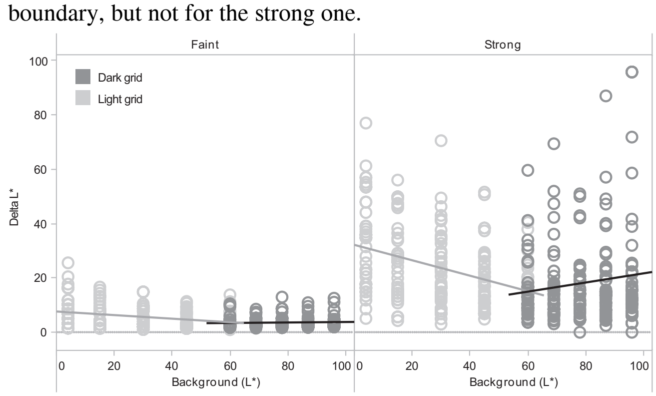


Figure 8. Trend lines showing how ΔL varies with background for our three experiments. As before, dark grid results are in dark gray, light grid results in light gray.*

Our hypothesis was that contrast metrics based on the background color would be most effective for predicting our boundaries for the flat case, and become increasingly less so as density increases (more of the background is covered). Factoring our analysis by density, we have discovered that this is basically true for the faint boundary, but not the strong one. As in the analysis for alpha, we have found that the dark and light grids are not symmetric.

In summary:

- For the dark grid, alpha and luminance contrast are equivalent. This is predicted by the derivation of luminance contrast in terms of alpha for black grids.
- For the light grid, luminance contrast metrics are highly dependent on the background, much more so than alpha.
- Like alpha, ΔL^* provides a useful metric for describing the faint boundary for both the dark and light grids, but is more sensitive to density and background than is alpha.
- Alpha provides a simpler, more consistent metric for defining the strong boundary than any of the contrast metrics

Alpha and Transparency

Alpha creates the appearance of transparency by using a weighted blend of the foreground and background colors. Metelli's episcotister model for transparency [4] describes a linear blend of reflectance values, equivalent to spinning a disk with reflectance R_f containing some percentage of open sectors over a background. If we use α to indicate the percentage of the disk that is open, and R_b to indicate the reflectance of the background, we can create the following equation to describe the reflectance of the perceived transparent surface (R_t):

$$R_t = \alpha R_b + (1-\alpha)R_f$$

The similarity to the equation for alpha is unmistakable, and is not surprising, given that both models are, in essence, linearly modulating and blending the amount of light striking the visual system. Brill [5] describes a similar result in terms of additive mixture of tristimulus values. Other researchers have shown that the reversal of contrast polarity along overlapping edges (called an X-junction) is a primary cue for transparency [6,7].

The primary focus of Metelli's work, and much other work in the field, is to establish the conditions under which a set of adjacent colored patches will induce the appearance of a transparent surface overlaying an opaque background. Our experiments use a physical model already known to give a good impression of transparency. What we have discovered is that we can create a relationship between our desired appearance and the degree of transparency (alpha).

Classically, transparency is described as a transparent surface covering an edge, but our boundaries can be seen even on a flat background. Furthermore, our very narrow grid lines make it difficult to focus locally on the appearance of edges seen through the transparent grid. We speculate that factors such as the consistent darkening of the overlapped colors and Gestalt continuity may be significant in the perception of our grid. The higher alpha values set for the dense case may be explained by interference with continuity.

As alpha increases from fully transparent, the grid first appears integrated with the surface. As it becomes more opaque, the grid "pulls forward" and becomes a separate object that sits in front of the background (our "fence"). This may be related to the fact that people tend to see objects rather than layers in transparency situations [6] so as the grid becomes stronger, it increasingly disassociates with the underlying image or objects.

Conclusions and Future Work

We have presented a set of experiments that establish a usable range, defined by alpha, for grids that are neither too faint nor too strong. For the large body of images that are not very dense, a light but useful grid could be created with an alpha value around 0.1, and in all cases, an alpha value of 0.2 falls in the "not bad" range. This is much lighter than the solid black grid (alpha=1.0) common in many visualization systems and technical illustrations.

To tie this work to perception, we look first at simple luminance contrast and color difference metrics, computed globally as a function of the background and grid colors. While there is a clear correlation between contrast and the faint boundary, transparency (alpha) appears to dominate the perception of the strong one. Furthermore, transparency performs as well as contrast for the faint boundary.

Our results strongly suggest that transparency is an important factor in subtle visualization, and that a metric for the degree of transparency, such as alpha, may continue to be a robust predictor of grid quality. In future work, we intend to examine these effects in more conditions, in particular when color is involved. We also plan to investigate interference from underlying patterns and textures (image complexity).

Acknowledgements

We would like to thank Billy Chen for his help running the experiments, and the reviewers for their insightful suggestions.

References

[1] E. R. Tufte, Envisioning Information. Graphics Press (Cheshire, Connecticut. 1998)

[2] G. E. Legge, G. S. Rubin, A. Luebker, "A. Psychophysics of Reading V. The role of contrast in normal vision," Vis. Res. 27, 1165-1171 (1987)

[3] S. Zuffi, C. Brambilla, G. Beretta, P. Scala, "Human Computer Interaction: Legibility and Contrast," 14th International Conference on Image Analysis and Processing (ICIAP 2007)

[4] F. Metelli, "The Perception of Transparency," Sci. Am., 230, 90-98 (1974)

[5] M. Brill, "The Perception of a Colored Translucent Sheet on a Colored Background." Color Res. & Appl., 19, 1, pp 34-36 (1994)

[6] P. Cavanagh, "Top-down Processing in Vision.," In R. A. Wilson, & F. C. Keil (Eds) MIT Encyclopedia of Cognitive Science, (MIT Press, Cambridge, MA, 2001) pp. 844-845.

[7] A. Kitaoka, "A New Explanation of Perceptual Transparency Connecting the X-junction Contrast-Polarity Model with the Luminance-based Arithmetic Model," Jap. Psych. Res., 47, 3, pp. 175-187.

[8] L. Arend, NASA Color Usage website (colorusage.arc.nasa.gov).

Author Biographies

Maureen Stone (StoneSoup Consulting) is an independent consultant working in the areas of digital color, information visualization and perception. She is an adjunct professor in the School of Interactive Art and Technology at Simon Fraser University, Editor-in-Chief of IEEE Computer Graphics & Applications, and is a member of ACM, IEEE and IS&T

Lyn Bartram earned her M.Math (1989) from the University of Waterloo and her Ph.D. in Computer Science (2001) from Simon Fraser University. Since 2005, she has been an Assistant Professor in the School of Interactive Art and Technology at Simon Fraser University. Her work focuses on the application of perception principles to visualization.