Discriminability measures for predicting readability of text on textured backgrounds

Lauren F. V. Scharff, Alyson L. Hill

Stephen F. Austin State University, Department of Psychology, Nacogdoches, Texas, 75962

lscharff@sfasu.edu and alyson_hill@hotmail.com

and Albert J. Ahumada Jr.

NASA Ames Research Center, Moffett Field, CA 94035-1000

aahumada@mail.arc.nasa.gov

Abstract: Several discriminability measures were examined for their ability to predict reading search times for three levels of text contrast and a range of backgrounds (plain, a periodic texture, and four spatial-frequency-filtered textures created from the periodic texture). Search times indicate that these background variations only affect readability when the text contrast is low, and that spatial frequency content of the background affects readability. These results were not well predicted by the single variables of text contrast (Spearman rank correlation = -0.64) and background RMS contrast (0.08), although a global masking index and a spatial-frequency-selective masking index led to better predictions (-0.84 and -0.81, respectively).

©2000 Optical Society of America

OCIS codes: (330.4060) Modeling of vision; (330.5000) Pattern; (330.5510) Psychophysics; (330.6110) Spatial filtering; (999.9999) Readability (100.5010) Pattern recognition and feature extraction; (110.2960) Image analysis;

References and links

1. G. Legge, G. Rubin, and A. Luebker, "Psychophysics of reading V. The role of contrast in normal vision," Vision Res. 27, 1165-1177 (1987).
2. K. Knoblauch, A. Arditi, and J. Szlyk, "Effects of chromatic and luminance contrast on reading," J. Opt. Soc. Am. A. **7**, 1976-1984 (1990).
3. G. Legge, and G. Rubin, "Psychophysics of reading IV. Wavelength effects in normal and low vision," J. Opt. Soc. Am. A. **3**, 40-51 (1986).
4. S. Pastoor, "Legibility and subjective preference for color combinations in text," Human Factors, **32**, 157-171 (1990).
5. G. Legge, D. Pelli, G. Rubin, and M. Schleske, "Psychophysics of reading I. Normal vision," Vision Res. **25**, 239-252 (1985).
6. J. Farrell and A. Fitzhugh, "Discriminability metric based on human contrast sensitivity," J. Opt. Soc. Am. A. 7, 1976-1984 (1990).
7. J. A. Solomon and D. G. Pelli, "The visual filter mediating letter identification," Nature **369**, 395-397 (1994).
8. D. H. Parish and G. Sperling, "Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination," Vision Res. 31, 1399-1415 (1991).
9. D. Regan and X. H. Hong, "Recognition and detection of texture-defined letters," Vision Res. **34**, 2403-2407 (1994).
10. P. Kember and D. Varley, "The legibility and readability of a visual display unit at threshold," Ergonomics **30**, 925-931 (1987).
11. B. Parker and L. V. Scharff, *Influences of contrast sensitivity on text readability in the context of a GUI*, (1997) http://hubel.sfasu.edu/research/agecontrast.html.
12. A. Hill and L. V. Scharff, "Readability of screen displays with various foreground/background color combinations, font styles, and font types," Proceedings of the Eleventh National Conference on Undergraduate Research, **II,** 742-746 (1997).

13. M. Youngman and L. V. Scharff, "Text width and border space influences on readability of GUIs," Proceedings of the Eleventh National Conference on Undergraduate Research*,* III*,* 786-789 (1998).
14. A. L. Hill and L. F. V. Scharff, "Readability of computer displays as a function of colour, saturation, and background texture*"* in *Engineering Psychology and Cognitive Ergonomics Volume Four*, D. Harris, ed. (Ashgate, Aldershot, England, 1999).
15. L. F. V. Scharff, A. J. Ahumada, Jr., and Alyson L. Hill, "Discriminability Measures for Predicting Readability," Human Vision and Electronic Imaging IV, SPIE Proc. 3644, 270-277 (1999).
16. Scharff, L. V. Example stimuli using different text contrast levels and spatial-frequency-filtered backgrounds. (2000), http://hubel.sfasu.edu/research/stim/exfiltstim.html
17. A. J. Ahumada, Jr., A. M. Rohaly, and A. B. Watson, "Models of human image discrimination predict object detection in natural backgrounds," in B. Rogowitz and J. Allebach, eds., Human Vision, Visual Processing, and Digital Display IV, SPIE Proc. 2411, 355-362 (1995).
18. A. M. Rohaly, A. J. Ahumada, Jr., and A. B. Watson, "Object detection in natural backgrounds predicted by discrimination performance and models," Vision Res. **37**, 3225-3235 (1997).
19. B. L. Beard and A. J. Ahumada, Jr., "Image discrimination models predict detection in fixed but not random noise," J. Opt. Soc. Amer. A **14**, 2471-2476 (1997).
20. A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking." J. Opt. Soc. A. A. **14**, 2379-2391 (1997).
* Links to some of the above-referenced research may be found on webpages maintained by Lauren Scharff (http://hubel.sfasu.edu/research/reslvs.html) and Al Ahumada (http://vision.arc.nasa.gov/~al/ahumada.html).

1. Introduction

The readability of text displays affects multiple aspects of daily life, and in cases such as air traffic controllers, it can influence the safety of many individuals. The increased use of the Internet means that such readability issues affect more people every day. There have been numerous studies of factors that influence the legibility and readability of computer text displays. Common factors that have been studied include luminance and/or chromatic contrast [1, 2], wavelength [3,4], blur [5,6], the addition of noise [7, 8, 9], case [10], and polarity [5, 11].

There have been fewer empirical studies of how webpage design factors influence the readability of a webpage. Most design manuals include general recommendations (i.e. use high contrast or specific colors) without explanation or reference to empirical work. Many of their recommendations are obviously subjective, and thus vary across manuals (for a review see Hill and Scharff [12]). This subjectivity can be problematic because several studies have indicated that there is a low correlation between subjective preference for text displays and empirical measures of their readability [4, 11, 13, 14].

An additional factor that has received little attention is how the use of a textured background will influence the readability of text presented on top of it. More and more webpages use textured backgrounds, many of which are obviously detrimental to readability, while others seem to have little effect. The purpose of this current work is first, to measure readability (search times) for texts of different contrasts when presented on textured backgrounds containing different spatial frequency bandwidths, and second, to investigate different approaches to predicting the readability of such text displays.

Research into noise effects on text readability may be more useful for predicting readability when the display is noisy or degraded than for predicting the effect of background textures, because in the latter case the text is placed on top of the background and the text itself is not noisy or degraded. However, research on the effects of noise and blur does indicate that specific noise and text spatial frequencies ranges are particularly important for discriminating letters or reading text. Solomon and Pelli [7] measured the effect of high- and low-passed noise on lower-case letter recognition and concluded that recognition is most dependent upon text contrast in a mid-frequency range (between 1.5 and 6 cycles/letter). Using upper case, band-passed letters masked by band-passed noise, Parish and Sperling [8] concluded that maximum human efficiency for letter identification occurs between 0.42 and 1.5 cycles/object. Because the results were largely independent of viewing distance, both these studies also concluded that object frequency rather than retinal frequency determines discrimination. Similarly, Legge, et al. [5] determined a critical cut-off frequency for reading low-pass filtered text to be approximately 2 cycles/letter, independent of character size. Overall, these results suggest that background textures will mask only to the extent that they contain spectral energy in a critical frequency range.

Continuing our previous work [15] this paper will assess the use of image measures and two indices developed from discrimination models to predict readability. This earlier work measured correlations using readability data from text-on-textured-background images [14] that were not specifically created to assess differences between the different approaches. Specifically, the textured backgrounds all contained a broad range of spatial frequencies, and since all used black text, there was little range in text contrast. As a result, there was little difference in predictability using the image measures or either of the discriminability indices (the global masking index or the spatial-frequency-selective index). The current work uses three levels of text contrast and a range of backgrounds (plain, a periodic texture, and four spatial-frequency-filtered textures created from the periodic texture).

2. An experiment measuring readability

The experiment used a 6 (background) X 3 (text contrast) design, minus two conditions that were not readable. The text in these two conditions was detectable on the backgrounds; however, the task required reading and finding a target word, and since the text could not be read, we eliminated those conditions from the experiment. Each remaining condition was repeated three times, for a total of 48 trials. There were also three practice trials to familiarize participants with the procedure.

2.1 Apparatus and stimuli

Macintosh Power PC 7200/120 computers were used to create and run the experiment. The text portions of the stimuli were created in B/C Power Laboratory (an experiment application), which was also used to present the stimuli and collect the data. The average luminance and spatial frequency bandwidths of the textured backgrounds were manipulated using MATLAB. A chin rest controlled viewing distance (475 mm).

Three text shades (medium grey, dark grey, black) resulted in three text contrast levels (0.15, 0.35, and 0.95) given the average background luminance of 62.5 cd/m^2. These contrasts were chosen because they represent the wide range of contrasts seen on webpages (fortunately the majority use high contrast), and they resulted in a range of search times which allowed us to better test our various approaches to predicting readability. There were six background textures: plain, a periodic texture, and four spatial-frequency-filtered textures created from the periodic texture. The periodic texture was taken from a popular webpage dedicated to supplying free graphical backgrounds to designers, and it was one of the textures used by Hill and Scharff [14]. The textures had a period of 72 pixels horizontally and vertically.

The frequency-filtered textures were created using four filters with a rectangular spatial-frequency response and a uniform orientation response. The filters selected adjacent octaves, with the high-frequency cutoff for the highest spatial frequency band equal to the Nyquist limit (0.5 cycles per pixel = 12 cycles/deg = 3 cpl). Thus, the highest spatial frequency band (Band 4, 1.5 - 3 cycles per letter, cpl) most closely corresponded to the critical range for identification of letters as determined by Solomon and Pelli [7]. The remaining spatial frequency ranges were as follows: Band 3 (0.75-1.5 cpl), Band 2 (0.375-0.75 cpl), and Band 1 (0.1875-0.375 cpl). See Figure 1 for examples of each of these filtered background textures, the original periodic texture and a plain background of equal average luminance. Pilot testing revealed that for all conditions the text was detectable on the background textures, it just was not readable for two conditions: those using the lowest contrast text placed on the periodic texture containing all frequencies and the Band 3 filtered texture. Thus, these two conditions were excluded from the experiment.

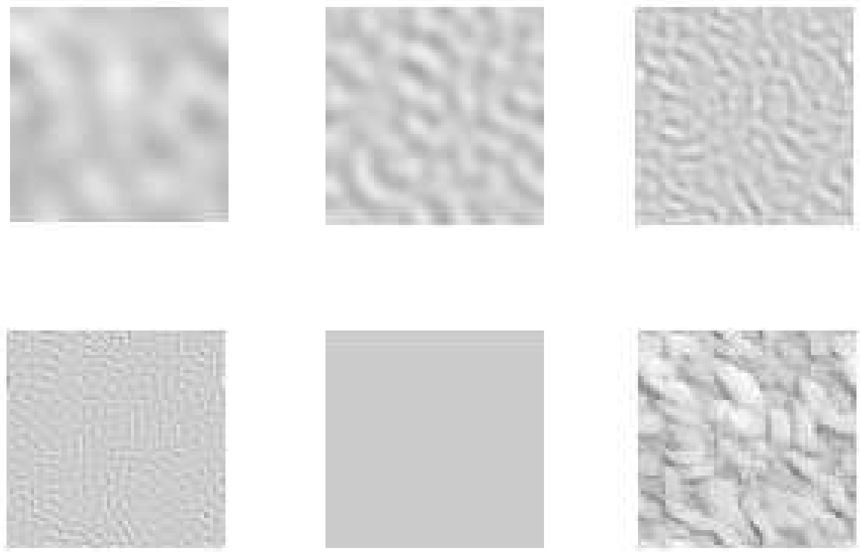

Fig. 1. Six texture samples used to create the backgrounds for the experiment. The lower right pattern shows the original , unfiltered texture. Starting from the top left are the band-pass filtered versions named: Band 1 (0.1875-0.375 cpl), Band 2 (0.375-0.75 cpl), Band 3 (0.75-1.5 cpl), Band 4 (1.5-3 cpl), and a plain texture of the same mean luminance.

The final, textured background size was created by tiling six of the periodic textures horizontally and vertically, leading to a 15.5 cm square texture (18.36 deg /side). Each textured background was centered at the top of the screen. Heavy black lines on the left and right separated each textured background from the surrounding white background. Text was placed on top of the textured backgrounds. Other variables were set to maximize readability within ranges commonly viewed on webpages: 12 point (6 pixels per letter) Times New Roman font, and the text blocks (10.2 cm x 12.7 cm) were centered at the top of the screen, leaving a 2.5 cm textured margin on either side [12, 13]. As a result, at our viewing distance, each letter was 0.25 deg in height.

The text excerpts were from a newspaper and were the same as those used by Hill and Scharff [14]. The text blocks to be read each contained 99-101 words. A target word ("triangle", "circle", or "square") was placed in a counterbalanced manner within each text block. At the bottom of each screen there were three black, geometric shapes (circle, square, and triangle) that corresponded to each of the three possible target words. These 1 cm x 1 cm shapes were spaced 3.5 cm apart and centered below the textured area. Full sized, example stimuli can be viewed on the Internet [16].

2.2 Procedure

Sixty undergraduate participants completed the experiment; however, data was only used from 47 participants due to high error rates (greater than 10% overall) which indicated that the excluded participants did not attend to the task. All participants were naive to the hypothesis.and had self-reported 20/20 or corrected to 20/20 vision. At Stephen F. Austin State University the great majority of undergraduate students are between the ages of 18 and 21.

Participants were instructed to scan the text and find a target shape word ("triangle", "square", or "circle"). Once they found the target word, they clicked (using the mouse pointer) on the corresponding shape at the bottom of the screen. The start of each trial was self-paced, and each trial ended when the participant clicked the target-word shape. Participants were instructed to respond as quickly and accurately as possible. Total time to complete the experiment varied between 20 and 45 minutes.

2.3 Results

The search time data were sorted by each condition for each participant and the median for each was calculated. The data from participants with an overall accuracy rate of at least 90% were used in the analyses, and of those, only reaction times from correct responses were used.

Because two conditions were not used in the experiment, the design was not complete. Therefore, two 2-way ANOVA's were performed, one with all text contrast levels but only four background textures, and the second with all background textures but only the darker grey and black text levels.

Results of the 2-way ANOVA using all three text contrast levels showed significant main effects and an interaction. See Figure 2 for means of all conditions (including those from the second analysis). A Tukey HSD analysis of the main effect for text contrast ($F(2,92) = 49.58$, $p<0.01$) indicated that the lighter grey text (0.15 contrast) was read significantly slower than the other two contrast levels. Background texture also significantly affected reading times ($F(3,138) = 9.37$, $p<0.01$), in that the Band 2 filtered texture was read more slowly than the Band 1 and Band 4 filtered textures and the plain texture. These main effects were modified by a significant interaction ($F(6,276) = 9.29$, $p<0.01$), in that the effect of background texture was only shown for the lighter grey text contrast. Further, for the plain background, the black text was read significantly faster than both the dark grey and medium grey text.

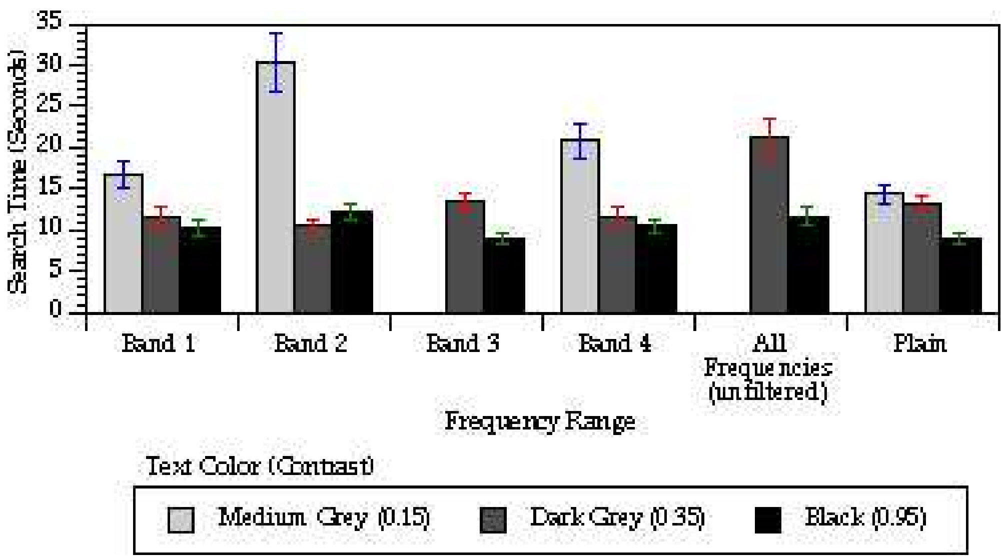

Fig. 2. Search times (sec) for all conditions that were readable. The effect of spatial frequency filtering was greatest when the text contrast was low. The text was not readable for the low text contrast using the unfiltered background (all frequencies) and the Band 3 filtered background. Vertical bars represent standard errors of the means.

Results of the 2-way ANOVA using all six background textures but only the black and dark grey text also showed significant main effects and an interaction. Black text (0.95 contrast) was read significantly faster than the darker grey text (0.35 contrast), ($F(1,46) = 37.47$, $p<0.01$). Background texture also significantly affected reading times ($F(5,230) = 9.6$, $p<0.01$), in that the periodic texture containing all frequencies was read significantly slower than any of the other background textures. These main effects were modified by the significant interaction ($F(5,230) = 7.74$, $p<0.01$), in that the effect of background texture was more strongly displayed when using the dark grey text than with the black text.

The significant difference between the black and dark grey text seen for the second analysis and not the first is due to the increased search times in the conditions not included in the first analysis (i.e. the Band 3 filtered texture and the "all frequencies," or unfiltered, texture). Note that the pattern of search times seen for the darker grey text (i.e. longer search times for the Band 3 filtered and "all frequencies" textured backgrounds) also was seen with the lighter grey text, which was often unreadable and thus not used in the experiment. Probably because the amount of background RMS contrast even in the case including all frequencies was fairly low (0.15), there was no effect of background texture when using high contrast (black) text.

2.4 Discussion

The above results indicate that the use of a textured background can influence readability of text displays (as indicated by an increase in search times), and that the effect will depend upon both the text contrast and the spatial frequencies contained in the texture. More specifically, we show a moderate effect of contrast at 0.35 and strong effect of contrast at 0.15 when some textured backgrounds are used. We found no effect on search times when using black text (0.95 contrast). Finally, there was a relatively small effect of contrast when using plain backgrounds in that both the medium and dark grey text were read more slowly than the black text. When they used 0.25 deg letters and plain backgrounds, Legge et al. [1] found a strong reduction in reading as contrast was reduced, although data were only reported from two participants, and they showed large individual differences. For 1 deg letters, however, they showed little effect of contrast until it fell below 0.1. Our letters subtended 0.25 deg and our results were similar to those of Legge et al. using the same sized letters.

The effect of background texture seems to be the greatest for the Band 3 (0.75-1.5 cpl) filtered texture or the unfiltered texture (i.e. the "all frequencies" texture). This finding relates fairly well to the previous research on letter recognition using noise and blur. Specifically, this range falls within the range (0.42-1.5 cycles/object) specified by Parish and Sperling [8] and the top end of this range matches the low end of the range (1.5-6 cpl) specified by Solomon and Pelli [7]. It is slightly below the critical spatial frequency bandwidth for reading (2 cycles/character) reported by Legge et al. [5]. Some of the discrepancy may be explained by differences in case and task (letter recognition versus reading). Kember and Varley [10] found that single letters are more legible when presented in upper case, while there was no difference between upper and lower case when reading words. Parish and Sperling used upper case letters while Solomon and Pelli used lower case letters, which might require higher frequencies for recognition, especially since they were presented in isolation rather than in words.

3. Predicting readability

The above results suggest that, at least for the textures tested in this experiment, lower contrasts and some spatial frequencies may be more detrimental to reading than others. Although these results may generalize to other textures, most webpage designers would not know how to determine the spatial frequency content of their background textures. Further, use of other fonts or font sizes, etc. may alter the critical spatial frequencies. What would be most useful is to create an application that would input the text and potential background and output some measure of readability. As a first step toward that, we will investigate how well both image measures and two discriminability indices predict the search times for the reading task above. The two indices are modified from those used by Scharff, et al. [15].

3.1 Image measure regressions

As in our previous work [15], the specific image measures used in the analyses included text contrast, background RMS contrast, and background RMS contrast in the four spatial frequency bands described above: Band 1 (0.1875 – 0.375 cpl), Band 2 (0.375 – 0.75 cpl), Band 3 (0.75 – 1.5 cpl), and Band 4 (1.5 – 3 cpl). The text contrast was defined as Eq. (1)

$$C_T = (L_B - L_T) / L_B , \quad (1)$$

where L_B is the average background luminance and L_T is the text luminance. The background RMS contrast was defined as Eq. (2)

$$C_{RMS} = L_{RMS} / L_B , \quad (2)$$

where Eq. (3)

$$L_{RMS} = ((\Sigma(L_i - L_B)^2) / n)^{0.5}, \quad (3)$$

and where the summation (Σ) is over all pixels, L_i is the luminance of the *ith* pixel, and n is the number of pixels.

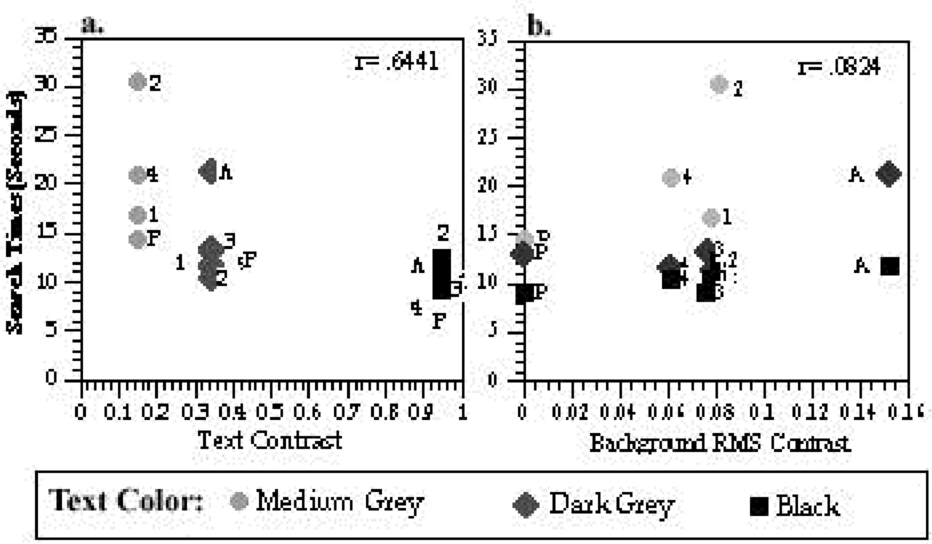

Figure 3 (a,b). Mean search times for each background plotted against text contrast (left) and background RMS contrast (right) and the two rank correlation coefficients (non-parametric). Numbers/letters to the side of each symbol indicate the frequency band (1-4), A (all frequencies), or P (plain texture).

Figure 3 shows search times with respect to text contrast and background RMS contrast and the resultant Spearman rank correlation coefficients. The text contrast had a negative correlation with search time (r = -0.64) and background RMS contrast showed a negligible correlation (r = 0.08). RMS contrast energy in the four spatial frequency bands also each led to a negligible correlation. Note however, that the two unreadable low-text-contrast conditions are not included in these correlations. If they could have been, the correlation between background RMS contrast and search times would undoubtedly been much higher since these two conditions would have led to long search times. Further, if there had been more variation in the background RMS contrast, it would have had a greater effect.

Ideally, we would like to investigate the combined influence of both text contrast and background RMS contrast. However, because we have so few conditions, is not possible to investigate the combined influence of multiple variables without estimating additional parameters. Therefore, we investigated two different discriminability indices based on models of the visual system that allow us to include both text contrast and background RMS contrast.

3.2 Metrics based on image discriminability models

Image discriminability models have been developed to predict the visibility of the difference between two similar images. They take two images as input, and output a prediction of the number of Just Noticeable Differences (JNDs) between them [17, 18]. Although our task was quite different from a typical discriminability task (ours was a search task with a target always present, rather than a target present/target absent decision), we hoped that model predictions of the text detectability on the background might predict search times.

To use the image discrimination models, the background-only image is one image and the other is the background-with-text image. Scharff et al. [15] outlines the derivation of the difference image based on equivalent contrast and our simplification such that it depends only upon the difference of the mean background and the text level (change in luminance). Although the difference of the background and its mean (change in texture) can contribute to detectability [9], we assume that it does not significantly contribute to readability. We also include the simplifying assumption of a flat contrast sensitivity function, since the readers sat close enough to the display that the frequencies relevant to reading were in the optimal visual range (about 6 cpd) or lower.

3.3 A global masking index

An index combining text contrast and background RMS contrast can be generated using a single filter, image-discrimination model with global RMS contrast masking. Such a model has been used to predict the detectability of targets in both natural and noisy backgrounds [17, 18, 19]. It assumes that the masking contrast energy is uniform over the target region and similar to the target in spatial frequency.

As derived in Scharff et al. [15] for binary text, the discriminability index turns out to be

$$d = s\, n_T^{\,0.5} C_T / (1 + (C_{RMS} / C_2)^2)^{0.5}, \tag{4}$$

where n_T is the number of text pixels and s is a contrast sensitivity parameter representing the discriminability of a single pixel at unit contrast. We define the readability index as the text contrast that would give the same discriminability on a uniform background. For this model, the readability index is independent of the size of the text target and the contrast sensitivity, giving the equivalent luminance contrast C_M of the masked text as

$$C_M = C_T / (1 + (C_{RMS} / C_2)^2)^{0.5}\ . \tag{5}$$

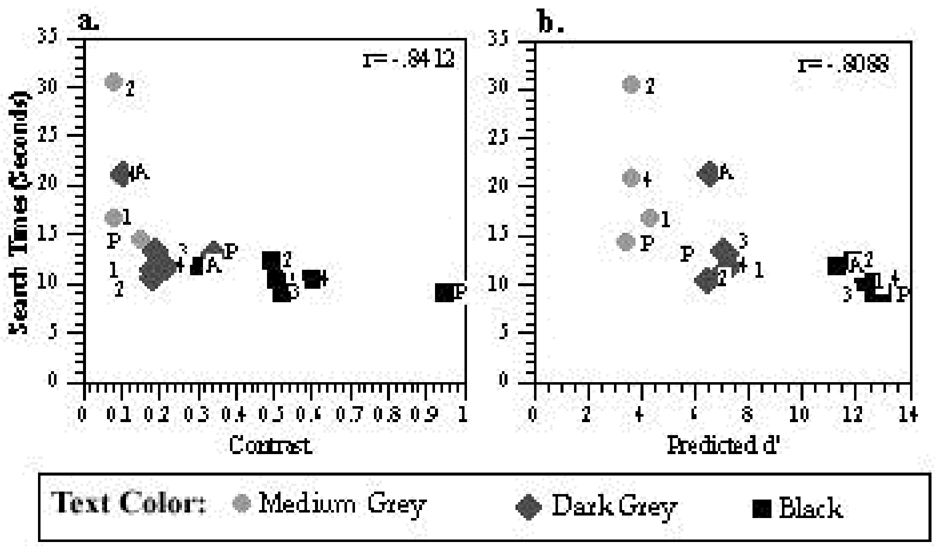

Fig. 4 (a,b). Mean search times for each background plotted against the global masking model equivalent contrast (left) and the frequency-selective model d' index (right), and the two rank correlation coefficients (non-parametric). Numbers/letters to the side of each symbol are as indicated in Figure 3.

Figure 4a shows the nonlinear relationship between search times and the global masking index. The Spearman rank correlation coefficient between this index and search times (r = -0.84) was higher than that for either the text contrast or the background RMS contrast alone. Since Solomon and Pelli [7] had indicated that the frequencies between 1.5 and 6 cpl were the most crucial for letter identification, we also calculated a Global Masking index using only our Band 4 frequencies. It resulted in a rank correlation of r = –0.83. Since our results indicated that the Band 3 frequencies were just as important for our reading task (although including all frequencies in the background was the most detrimental), we also calculated the Global Masking index using only that range of frequencies (r = -0.84). These results suggest that an index based on either global masking or particular ranges based on prior reading research do equally well.

3.4 A frequency-selective masking index

To predict the effect of background masking when the spatial frequency content of the background varies, or when the orientation of the pattern varies, a spatial-frequency-and-orientation-selective masking model can be used to compute the readability index C_M. Scharff et al. [15] found that a Cortex Transform model [18] did not make better predictions than the image measures or the global masking model. This may have been partially the result of using backgrounds whose contrast variations did not vary in their spatial frequency content.

Because the current backgrounds do vary in spatial frequency and orientation, we decided to again calculate a frequency and orientation selective index. However, this time we modified the model developed by Watson and Solomon [20] instead of using the Cortex transform. This model gives essentially the same results, but uses Gabor filters rather than complex filters which are not easily described. The complex filters of the Cortex model facilitate reconstituting the image from the Cortex model representation, a feature that is useful for image compression applications, but not for our application.

More specifically, the Watson and Solomon [20] model takes as input two images. Each image passes through a CSF filter and an array of Gabor filters, varying in phase, spatial frequency, orientation, and spatial position. The filter array outputs then pass in parallel through both an excitatory and inhibitory nonlinearity; the inhibitory path passes through a linear pooling filter, and then it divisively inhibits the excitatory signal. The resulting array representation from each of the two images is compared (subtracted) and subjected to Minkowski pooling to obtain the prediction of the distance between the images in JND units (d').

For our implementation, we first created a Gabor filter array that included the four spatial frequency ranges of our stimuli and four orientations (horizontal, vertical, and 45 deg. diagonals to the left and right). The spatial frequency ranges of the four background frequency ranges at our viewing distance were 6-12 cpd (Band 4), 3-6 cpd (Band 3), 1.5-3 cpd (Band 2) and 0.75-1.5 cpd (Band1). The one-octave-wide Gabor filters were centered at the midpoint of each of the above ranges. Finally, rather than using a global CSF filter, we adjusted the gain of each channel so that it matched the sensitivity curve for Gabor targets used by Watson and Solomon [20]. Our inhibitory pooling filter summed only over phase; unlike Watson and Solomon we did not pool over spatial frequency or orientation. See Table 1 for the parameter values.

Table 1. Parameter values.

Parameter	Symbol	Value
CSF peak amplitude *	a	40
CSF peak frequency *	f	1
CSF log bandwidth *	w	1.2
Excitatory exponent *	p	2.3
Inhibitory exponent *	q	2.0
Saturation constant	b	0.05
Pooling width in x or y *	s	0.0
Octave bandwidth of Gabor filters *	k	1.0
Minkowski exponent	B	4.0
Masking contrast threshold	r	0.05
Pixels per degree		24
Pixels per image in each dimension		72

* indicates a parameter value taken from Watson and Solomon [20].

The Spearman rank correlation coefficient for this index (r = -0.81) was essentially the same as for the global masking model index. See Figure 4b for a plot of the search times versus the index value. Using only the channel corresponding to Band 4 resulted in a correlation coefficient of r = -0.76, and, using only the Band 3 channel, r = -0.76. Thus, in this case, the combination of inputs from all frequencies (and orientations) led to a better prediction of the search times. Possibly, if the background textures themselves had not been pooled over frequency when they were created, we would have seen better predictions using this model as compared to the Global Masking Model.

3.5 Discussion of predictability

Unlike the results of Scharff, et al. [15], we did find better predictability using the discriminability model indices as compared to the single image measures of text contrast and background RMS variation. For textures such as those we used (a tiled pattern with a fairly uniform pattern within each tile) the Global Masking index predicts readability just as well as the more complex Spatial Frequency Model index. Since it is much simpler computationally, at this point we would recommend its use over the Spatial Frequency Model. However, some background textures may show more spatial variation. Thus, for an application to aid web designers in their choices of background textures, it might ultimately be more useful to implement the more complex model.

One issue not addressed in the current work is that of invariance. Although previous research [5, 7, 8] has suggested that frequency-selective noise or blur effects are independent of distance or character size, we did not manipulate either of these variables. Further, all three of our target words (circle, square, and triangle) are similar in length. It is possible that the effect of background frequency might vary with size of the target word. If the actual target word rather than the text sample had been used as the target for the model, the models might have been able to predict effects of varying the size of the target word.

4. Summary

Both text contrast and background contrast variation affect text readability. Background variation effects were only seen when the text contrast was low. Greater effects of background variation would be expected if larger background contrasts were used. The spatial frequency content of the background has some effect on the readability of the text. Of the four spatial frequency ranges we used, backgrounds with frequency content restricted to 0.75 – 1.5 cycles per letter had effects similar to those of unfiltered backgrounds. This finding is similar to those in previous research measuring the effects of noise and blur on letter recognition and readability [5, 7, 8].

The results of the experiment were fairly well predicted by the single variable, image measure regression using text contrast but not the regression using background RMS variation. The two indices based on image discrimination models lead to better predictability. Specifically, the simple Global Masking index predicted the search times just as well as a more complex model which included spatial frequency and orientation channels.

Acknowledgements

This work was in part supported by grant NCC2-1095 to the San Jose State University Foundation from NASA RTOP 548-51-12.