COMPUTER TEXT SHADING ALGORITHM BASED ON PERCEPTION OF LUMINANCE

Dr. Sehchang Hah
Federal Aviation Administration
Atlantic City International Airport, New Jersey

Dr. Thomas Hickey
Online Computer Library Center OCLC
Dublin, Ohio

With raster display systems, images on a monitor show jaggedness, because they are defined by integer coordinates of the monitor. This jaggedness is called an aliasing effect. To reduce this, engineers developed algorithms. One well-known algorithm is supersampling. This is accomplished by sampling at a higher resolution and reducing the sampled data to a lower resolution. We overlaid arrays of a 3x3 matrix on a text (supersampling) and reduced each array to a pixel on a monitor. In this process, engineers determined the level of the pixel luminance-intensity linearly to the number of elements painted in each array. Instead of using such direct, linear transformation, we determined the level of brightness, not the level of the pixel intensity, to produce better shading. Brightness, not luminance, is subjective. We created nine sets of gray levels based on this algorithm. We ran two experiments to choose the optimal gray-level set. In Experiment 1, participants chose the more legible of two letters shaded with different gray-level sets. In Experiment 2, participants counted target-letters in a string of letters as fast as they could. The experimental results did not favor gray-level sets that were close to the traditional linear transformation from the number of painted elements in arrays to pixel luminance-intensity. The best set was actually the third brightest as identified with this procedure using human participants. This perceptual algorithm can be used for any monitor to reduce aliasing effect.

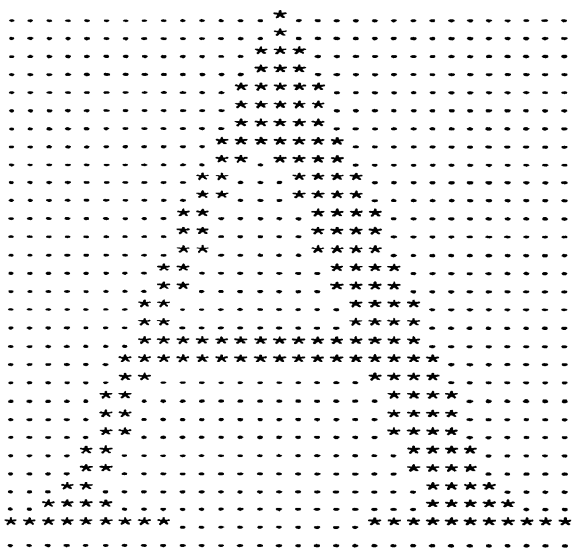
INTRODUCTION

For a raster display, objects to be displayed are defined by integer coordinates of a monitor. This digitization creates jaggedness of images. This effect is called aliasing. The increase of digital sampling enhances image quality (Gonzalez & Wintz, 1977), and luminance provides higher contrast which benefits visual acuity (Shurtleff, 1982; Snyder & Taylor, 1979).

However, aliasing is inevitable and pronounced, especially when resolution is low. Researchers have shown that shading using gray-levels is effective as an anti-aliasing method to enhance text display on a monitor (Crow, 1977 & 1981; Rogers, 1985; Schmandt, 1980; Warnock, 1980).

One popular anti-aliasing method is supersampling. This is done by sampling an object at a higher resolution and reducing the sampled data to a lower resolution. For instance, if the supersampling resolution is nine times higher than the lower resolution, grids of 3x3 data points will be overlaid on an object to be displayed, and each grid will be reduced to a single pixel. In this conversion process, engineers used a few anti-aliasing algorithms (Hearn & Baker, 1997). For example, they gave an intensity level of a pixel proportional to the number of data points in a 3x3 grid (Figure 1).

As an alternative anti-aliasing method, we used Stevens' power function of brightness, $p = k s^b$, where p is brightness, s is luminance, $k = 1$, and $b = 0.33$, to assign luminance to gray-levels. This



```
0000410000
0000850000
0002880000
0006084000
0015057000
0051019200
0076668500
0330003900
0600000840
4630002563
```

Figure 1. Anti-aliasing conversion.

is because the brightness function truly represents how we perceive luminance (Stevens, 1961). Instead of using luminance intensity directly to determine the luminance of gray-levels, we used brightness and divided the range of brightness to assign a luminance to each gray-level.

We created nine sets of gray- levels, and each set had nine equidistant gray-levels in brightness. To determine the optimal set of gray-levels, we ran two experiments and collected performance data. The goal was to optimize shading as an anti-aliasing method to make text more readable.

EXPERIMENT 1

Method

Participants. The participants were 30 volunteers employed by Online Computer Library Center (OCLC). They had 20/20 vision and were not colorblind.

Apparatus. We generated stimuli and presented them in an office setting by using an Apollo color computer (DN600). To measure the luminance of displayed text, we used a Tektronix J16 digital photometer with a J6523 probe.

Stimuli. We presented two lower-case letters of 10-point Roman font side by side for each trial. They were the same letters but were shaded with different gray-scaling sets. We presented them randomly, and each letter appeared at both sides equally often. They were presented at the center of the monitor screen, approximately 50 centimeters away from the participants, at eye level.

Method of shading. We formed 3x3 non-overlapping arrays to cover font images. The number of elements painted in each array determined the brightness level of a pixel to be displayed on the monitor (Figure 1). For example, if only one element was painted in an array, Level 1 brightness was given. In this way, we generated nine levels of brightness.

We chose Stevens' power function for our experiments over the logarithmic function. Wasserman, Felsten, and Easland (1979) argued in favor of using a logarithmic function for responses to brief stimulus presentations and using a power function for responses to long stimulus presentations.

We created nine gray-level sets, and each set had nine gray levels (Figure 2). For all of the sets, the brightest gray-level (Level 1) was anchored at the brightest luminance, 86.25 cd/m^2. This was the maximum luminance of the monitor we used. Set 9 was created by making nine equidistant levels of brightness based on the brightest and dimmest luminance, 86.25 cd/m^2 and 13.40 cd/m^2, respectively. The luminance of the blank screen was 13.30 cd/m^2.

To create other sets, we calculated a *subinterval* by dividing the gray-level interval of Set 9 by eight. We added this *subinterval* to the ninth gray-level of Set 9 to create Level 9 of Set 8. For other levels of Set 8, the interval between Level 9 and Level 1 of this Set 8 was divided by eight. This segment was used to assign the equidistance brightness between levels.

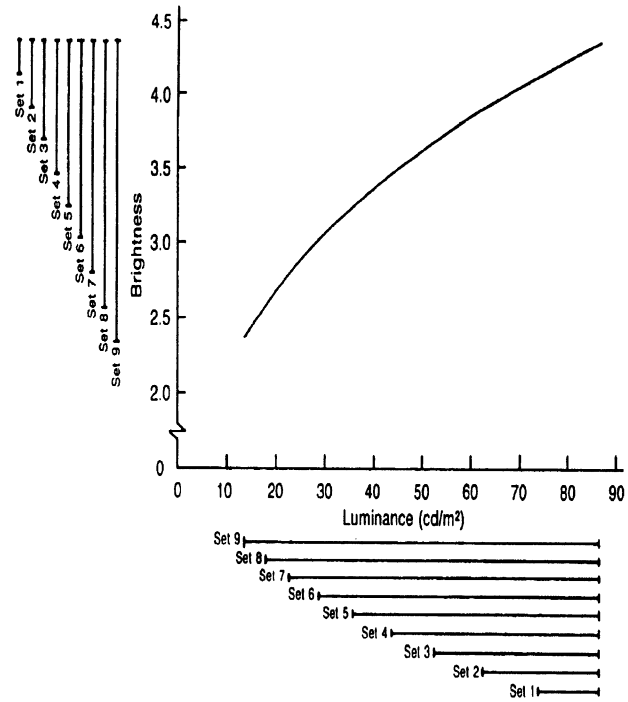


Figure 2. Gray-level sets used.

For Set 7, the *subinterval* was added twice to Level 9 of Set 9 to create Level 9. Then, we divided the interval between Level 9 and Level 1 brightness of the set by eight and used the segment to create other gray levels. We used the same procedure for other sets. We then converted all of the brightness values of the sets into luminance values using Stevens' function to display pixels on a monitor.

Procedure. Before the experiment started, we instructed the participants to choose the more legible of two letters displayed on the monitor. We told them that they were not constrained by time.

They had 8 practice trials and 72 data trials. The stimuli were presented for 10 seconds and disappeared. Then, a prompt message appeared on the monitor. After they responded, the monitor screen stayed blank for 9 seconds, and then the next trial started.

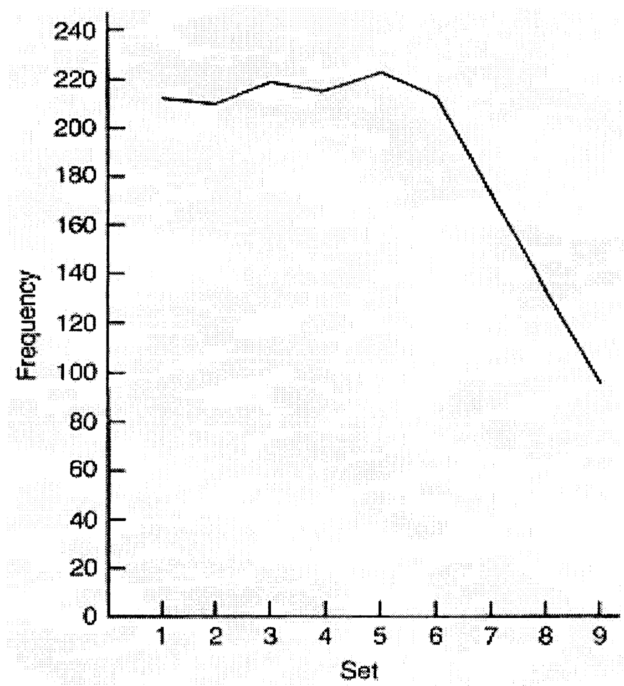


Figure 3. Frequency of preferences.

Results

There was a large variance among participants' preferences, but their overall preference was the upper middle sets of the luminance intensity continuum (Figures 2 and 3).

We tested all 36 pairs of sets (72 trials/2) using Wilcoxon Matched-pairs Signed-ranks Test (Kirk, 1968). We held the error rate of alpha at 0.10 per experiment. This corresponded to an alpha of 0.0028 for each of 36-pair comparisons, and its z value was 2.77 for two-tailed tests.

Set 9 was significantly different from all other sets except Sets 1 and 2 because the trend to choose either Set 1 or 2 over Set 9 was not consistent across participants. Only Set 5 was significantly different from Set 8 ($z = 3.31$). There were no statistically significant differences among the other sets.

EXPERIMENT 2

Method

Participants. The participants were 24 volunteers from OCLC. Nineteen had participated in Experiment 1, conducted about four months previously. All had 20/20 vision and were not colorblind.

Apparatus. The apparatus was the same as in Experiment 1.

Stimuli. We presented a string of 24 letters (lower case) for each trial.

Procedure. The participants' task was to count how many times a target-letter occurred in a displayed string of letters, as quickly as possible without errors. As soon as they finished counting, they pressed the left-most mouse button. This caused stimuli to disappear from the screen, and a prompt appeared asking them to type in the number of target-letters without time constraint.

After their response, the screen stayed blank for 14 seconds. Then, a different string of letters was displayed for the next trial. They received randomly-selected, different target-letters. They had 2 practice trials and 18 data trials. We measured the elapsed time between presentation of the stimuli and participants' responses. If their count was not correct, we counted it as an error.

Results

Wilcoxon Matched-pairs Signed-ranks Test (for two-tailed, error rate per experiment at 0.10, $z = 2.77$ to be significant) showed that the scanning times of Sets 1, 2, and 3 were significantly superior to those of Sets 7, 8, and 9 (Figure 4). The participants performed better with Sets 1 and 3 than with Set 5 and also better with Set 4 than Sets 7 and 8. Other pairs were not significantly different. Error rates did not show any significant differences among sets.

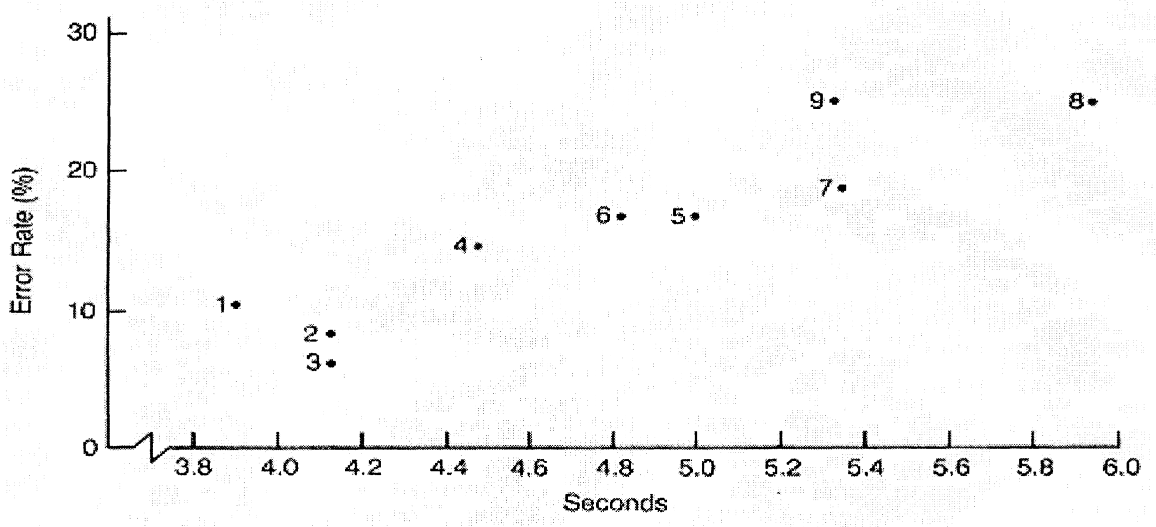


Figure 4. Error rate and scanning time for each gray-level set.

GENERAL CONCLUSION

In Experiment 1, statistical analysis showed that participants preferred all sets, except Sets 1 and 2, over Set 9. However, the overall frequency showed that participants preferred Sets 1, 2, 3, 4, 5, and 6 over Sets 7, 8, and 9 (Figure 3). In Experiment 2, tests of scanning time showed Sets 1, 2, and 3 were significantly superior to Sets 7, 8, and 9 (Figure 4). Error rate was not sensitive in detecting a significant difference between sets (Figure 4).

Based on the results of these two experiments, we chose Set 3 as the best. Set 9, which was closest to the linear transformation from the number of painted elements in arrays to pixel luminance-intensity, was one of the least preferred sets.

The results of Experiment 1 showed a large variance in participants' preferences that did not match their actual scanning performance in Experiment 2. This phenomenon must be heeded by engineers who sometimes follow subjective and intuitive judgment in designing hardware and software. Also, in the past, psychophysical characteristics were not considered seriously in formulating algorithms to display text and graphics.

In the environments like the air traffic control tower where traffic controllers read safety-critical information under time pressure, the quality of a monitor is very important. The experiments demonstrated that this algorithm can be used to identify the optimal gray-level set for any monitor to display more readable text.

ACKNOWLEDGEMENTS

We would like to thank the OCLC employees who volunteered to participate in the experiments. We also wish to thank Georgia K. M. Tobin for creating the fonts used and Richard Tobin for help in programming the gray-fonts display.

REFERENCES

Crow, F. (1977). The aliasing problem in computer-generated shaded images. *Communications of the ACM, 20(11)*, November, 1977, 799-805.

Crow, F. (1981). A comparison of anti-aliasing techniques. *IEEE Computer* graphics *and applications*, November, 1981, 40-47.

Gonzalez, R. C., & Wintz, P. (1977). *Digital image processing.* MA: Addison-Wesley.

Hearn, D. & Baker, M. P. (1997). Computer graphics: C version (2^{nd} ed.). Upper Saddle River, NJ: Prentice Hall.

Kirk, R. E. (1968). *Experimental design: Procedures* for *the behavioral sciences.* Belmont, CA: Brooks/Cole.

Rogers, D. F. (1985). *Procedural elements for computer graphics.* NY: McGraw-Hill.

Schmandt, C. (1980). Soft typography. In S. Lavington (ed.), *Information processing 80 (pp. 1027-1031).* NY: North-Holland.

Shurtleff, D. A. (1982). *How to make displays legible.* La Mirada, CA: Human interface Design.

Snyder, H. L. & Taylor, G. B. (1979). The sensitivity of response measures of alphanumeric legibility to variations in dot matrix display parameters. *Human Factors, 21 (4)*, 457-471.

Stevens, S. S. (1961). The psychophysics of sensory function. In W.A. Rosenblith (ed.), *Sensory communication.* NY: John Wiley & Sons.

Warnock, J. E. (1980). The display of characters using gray level sample arrays. *Computer graphics (SIGGRAPH 1980 Proceedings), 14(3)*, 302-307.

Wasserman, G. S., Felsten, G., & Easland, G. S. (1979). The psychological function: Harmonizing Fechner and Stevens. *Science, 204*, 85-87.