22.1: Modeling the Display and Perception of Gray-Scale Characters

Avi Naiman
University of Toronto, Toronto, Ontario, Canada

Joyce Farrell
Hewlett-Packard Laboratories, Palo Alto, CA

Abstract

To measure character quality objectively, font designers need automated tools for character generation and image-quality evaluation. This problem is addressed by modeling the entire viewing environment, from generation of grayscale characters through the visual system's response to their display on various output devices.

Introduction

Character fonts are displayed on a wide variety of devices, ranging from low-resolution instrument panels to high-resolution laser printers. Due to the characteristic differences between the various display technologies, it is not possible to design a single set of characters that will have acceptable image quality on all devices. Quite often, the only approach to manufacturing suitable character sets for a particular device is to have a font designer iteratively modify and evaluate the characters' bitmaps until the results are satisfactory. Ideally, one would like automated tools for character generation and an objective means to evaluate the perceived image quality of the characters when displayed on a particular device.

Until recently, most text on raster displays used characters represented as binary matrices, the ones and zeros corresponding to the black and white dots to be displayed. With the advent of grayscale technology, one method for improving the quality is to incorporate gray pixels in the character description, leading to a perceived improvement when comparing the discrete version of a character with its analog predecessor.

While many systems have been developed for generating characters, sorely lacking are utilities to evaluate them. We address this problem in a comprehensive manner by modeling the entire viewing environment, from generation of the characters through the visual system's response to the display of those characters on various characterized output devices.

Modeling the Viewing Environment

We model the entire viewing environment as three successive and linear stages of character processing: generating the grayscale characters, displaying them on a raster-scan device, and viewing the characters with the human visual system. If two images which differ in some aspect (e.g., the method used to generate the grayscale character) are equivalent at some particular stage, they can never differ after further processing. Distinguishing between these stages allows us to independently consider different character generation techniques, display architectures, and models of the visual system's response to various stimuli.

The Grayscale Characters

Traditional characters are analog in nature, their shapes defined by smoothly-varying boundaries separating figure from ground. With the advent of raster-scan displays and printers, the analog letterforms — produced by optical and mechanical methods — have been replaced with digital representations which can only approximate their predecessors. This will always be so, since character edges have infinite frequencies that can never be exactly reproduced with discrete devices. On the other hand, since the visual system is band-limited, we need only match the quality of the transmitter (i.e., the display device) to the capabilities of the receiver (i.e., the visual system). Unfortunately, although we know how to theoretically sample and reconstruct a band-limited signal perfectly [Pratt 78], current display device resolutions pale in comparison to the resolving power of the visual system. Furthermore, the pixel point spread functions differ substantially from the ideal reconstruction kernel (i.e., the *sinc* function).

An alternative to higher resolution is the use of grayscale technology, where in addition to black and white pixels, a multitude of gray levels are realizable. In general, if each pixel is represented with n bits, 2^n different grayscales are available at each pixel (subject to possible limitations of the display technology). Using gray pixels at the edges of characters can achieve a more faithful representation of the master character than any bi-level version could on the same grayscale device.

The common method for generating a grayscale character is *filtering*, whereby a high-resolution bi-level master character M is convolved with a digital filter F and sampled to yield a lower-resolution grayscale character G (figure 1) [Kajiya 81, Warnock 80]. For a particular grayscale display, the spatial resolution and number of intensity levels available is predetermined for the grayscale character generation process. However, many different filters can be used to generate a

ISSN0097-0966X/88/0000-424-$1.00 + .00 © 1988 SID

character. Furthermore, even with a single filter, different versions of the same character can be generated by shifting the sampling grid of the filtered character relative to the origin of the master.

Figure 1. Grayscale character G (left: at size; right: exploded view).

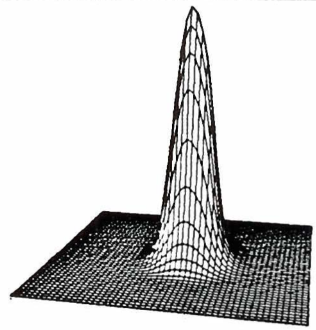

Figure 2. A two-dimensional pixel point spread function for a monochrome grayscale display device D.

Modeling the Display

Until now, we have ignored the relationship between pixel settings and intensities generated on a display's surface (i.e., the reconstruction kernel), tacitly assuming that each pixel value corresponds in a linear fashion to a point of light on the display, independent of neighboring pixel settings. In fact, luminance levels on the display are related in a monotonic, but nonlinear, fashion to pixel values, and, for each pixel, an *area* on the screen is illuminated, where the intensity profile — called the *point spread function* — is centered on the corresponding location, and decreases monotonically from the center. While techniques for compensating for intensity non-linearities have been developed [Cowan 83], a single compensation table may not be suitable for the entire display surface [Hosokawa 87, Brainard 87], although it is usually adequate for any localized area.

Modeling the pixel point spread function is even more difficult. Fortunately, the spectral power distribution of screen phosphors is invariant over emission levels [Brainard 87, Cowan 87] and the intensity profile is scale invariant (i.e., it retains the same shape at different settings, modulo a multiplicative factor) [Brainard 87]. However, it too is spatially variant (i.e., it may change shape in different regions of the screen), and, since pixel point spread functions are designed to overlap with their neighbors', they may not be spatially independent. Put another way, we may not be able to simply add the contribution of one pixel's point spread function to that of its neighbor's for all intensity settings.

Our first attempts at device modeling have been focused on a monochrome grayscale display, allowing us to ignore some of the added complexities imposed by a color system (e.g., shadow mask interference). Making the assumption of spatial invariance and independence, we have characterized the display with a luminance linearization function L (implemented as a look-up table) and a pixel point spread function D (figure 2).

To determine the light pattern on the display surface when a grayscale character G is presented, we first apply L to each of the pixels in G, and convolve the result with D. Note that by interposing L^{-1} between G and the display device, we can compensate for the display's intensity non-linearities, and thereby consider G to be a linear specification in luminance. Hence, we can directly convolve G with D to produce a representation S of the light stimulus coming off the display device in terms of the grayscale pixel values of the character and the point spread function of the display's pixels (figure 3). By specifying the resolution at which D is measured — or, alternatively, using an analytic representation of the pixel point spread function [Infante 86] — we can control the precision at which we represent S; while higher resolutions take longer to process, they reduce the error inherent in such discrete representations.

Figure 3. A representation of the stimulus S resulting from the convolution of a grayscale character G with the pixel point spread function D.

Modeling the Visual System

Once we have a useful representation of a character displayed on a grayscale device, we would like to describe what the eye actually sees (in terms of the pattern imaged on the retina when viewed from a given distance) as well as how the visual system responds to the stimulus (in terms of sensitivity to the incoming frequencies when in a specified state of adaptation). The former involves the optics of the ocular media, while the latter deals with psychophysical measurements of cortical image processing.

We have modeled the optical aspects with a visual blur function V_O that describes how a point light source is imaged onto the retina, in terms of visual angle (figure 4). While the appropriate point spread function depends on the diameter of the pupil and the spectral power of the light [Campbell 66], a single point spread function suffices for monochromatic, broadband light sources, when the eye is in good focus and has a pupil diameter of 2 mm [Westheimer 86]. Convolving the stimulus S with V_O yields a description of the lens-blurred character image I_O on the retina (figure 5). Once again, we can control the error in this discrete representation with the precision at which we define V_O.

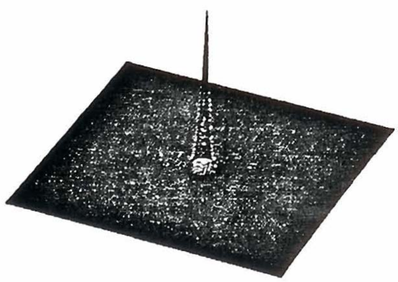

Figure 4. The two-dimensional optical blur function V_O representing the filtering of a point light source passing through the lens of the eye. The dimensions of the grid are 30×30 minutes of arc; the blur function approaches zero at approximately 2 minutes of arc from the center.

Figure 5. A representation of the image I_O resulting from the convolution of a stimulus S with the optical blur function V_O.

Additional filtering occurs due to photoreceptor sampling and cortical image processing. The combined effects of optical and psychophysical filtering are captured by the human contrast sensitivity function [Campbell 68, Robson 66], which describes the band-pass spatial filtering properties imposed upon every stimulus the visual system encounters. Similar to the case for optical blur, contrast sensitivity depends on the amount of light entering the eye (i.e., retinal illuminance) as well as the temporal frequency of the stimulus S [Kelly 72]. For a particular contrast sensitivity function, we derive a band-pass point spread function V_C (figure 6), which, when convolved with S, yields I_C, a representation of the image after luminance contrast that the visual system cannot detect has been filtered out.

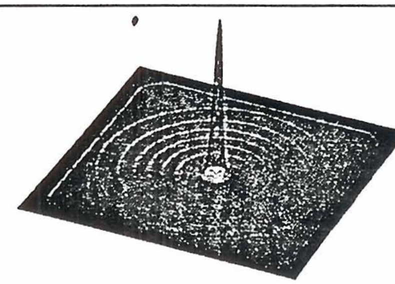

Figure 6. The two-dimensional cortical blur function V_C derived from a human contrast sensitivity function. The dimensions of the grid are 30×30 minutes of arc; the blur function becomes negative at approximately 5 minutes of arc from the center and returns to zero at approximately 15 minutes of arc.

Thus, given the optical point spread and the contrast sensitivity functions appropriate for the viewing conditions, we can represent the visual system's response to a stimulus S by convolving with filters V_O and V_C, to yield I_O and I_C, the images perceived by the visual system, either in terms of the physical mapping at the retina, or the psychophysical response to the stimulus, respectively.

Image Quality Evaluation

In modeling the viewing environment, we have identified three stages in the processing of grayscale characters: creation, display, and viewing. Two stimuli which are the same at one stage cannot differ after further processing. This fact gives us the means to evaluate, for example, whether two different filters F_1 and F_2, when applied to the same master character M, produce different grayscale characters G_1 and G_2 to be displayed. If $G_1 \approx G_2$, then we can select one filter over the other by criteria other than image quality, e.g., the cost of the filters. Similarly, if S_1, the result of displaying G on D_1, is found to be optimal (in some sense), and S_2, the result of displaying G on D_2 is (approximately) the same as S_1, then we can use G on D_2 as well as on D_1, and need not search for a better character for that display.

Note that not only can we compare master characters which have been filtered, displayed, and/or viewed, but we can bypass certain stages as appropriate. For example, if we have M_I defined as the ideal character to present to the visual system (e.g., the resolution of M_I is greater than the resolving power of the visual system), we can compute I_I, the response of the visual system to M_I, by convolution with V_O. Then, if a particular grayscale character G, when displayed on a device D, has an effect on the visual system I_G that is identical to the effect I_I, we can consider G to be optimal with respect to D.

By modeling the creation, display and viewing of grayscale characters, we are able to characterize the information available to the visual system for visual discrimination. This is the basis upon which we can develop and test metrics for evaluating grayscale characters. If two stimuli are identical at any of the processing stages, they must by treated identically by viewers. Of course, if two characters are different after passing through the visual filtering stage, but that difference is small, viewers may be unable to detect that the characters are in fact not the same. However, we may still be able to predict discrimination performance if we assume that there is a threshold difference criteria which viewers use in determining whether characters are different.

In order to test these predictions, we need to define a method for estimating the psychophysical difference between two characters. One approach is to compute the *filtered contrast energy* [Mannos 74], by adding the squared differences between corresponding points in two images I_1 and I_2. The next step is to check whether the psychophysical similarity is monotonically related to the difference metric, by presenting subjects with pairs of grayscale characters and measuring their ability to detect differences between the stimuli.

Conclusions

Modeling the viewing environment allows us to generate a description of the character image as perceived by the visual system. This provides the basic platform needed to develop and test metrics for judging the perceived quality of a grayscale character generated from a particular filter, displayed on a characterized device, and viewed under specified conditions. By evaluating the generation, display and visual filtering stages with appropriate metrics, we will be able to determine what filters and display architectures will optimize the quality of grayscale characters.

References

Brainard 87 Brainard, D. H., "Calibration of a Computer-Controlled Color Monitor," Submitted for Publication, *Color Research and Application*, 1987.

Campbell 66 Campbell, F. W. and R. W. Gubisch, "Optical Quality of the Human Eye," *Journal of Physiology*, Volume 186, 1966, pp. 558-578.

Campbell 68 Campbell, F. W. and J. G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," *Journal of Physiology*, Volume 197, 1968, pp. 551-566.

Cowan 83 Cowan, W. B., "An Inexpensive Scheme for Calibration of a Colour Monitor in Terms of CIE Standard Coordinates," *Computer Graphics*, Volume 17, Number 3, July 1983, pp. 315-321. SIGGRAPH 1983 Proceedings.

Cowan 87 Cowan, W. B., "Colorimetric Properties of Video Monitors," Notes for a Short Course, *Optical Society of America*, 1987.

Hosokawa 87 Hosokawa, H., Y. Nishimura, R. Ashiya and S. Odumura, "Digital CRT Luminance Uniformity Correction," *SID 87 Digest*, 1987, pp. 412-415.

Infante 86 Infante, C., "Ultimate Resolution and Non-Gaussian Profiles in CRT Displays," *Proceedings of the SID*, Volume 27, Number 4, 1986, pp. 275-280.

Kajiya 81 Kajiya, J. and M. Ullner, "Filtering High Quality Text for Display on Raster Scan Devices," *Computer Graphics*, Volume 15, Number 3 August 1981, pp. 7-15. SIGGRAPH 1981 Proceedings.

Kelly 72 Kelly, D. H., "Adaptation Effects on Spatio-Temporal Sine-Wave Thresholds," *Vision Research*, Volume 12, 1972, pp. 89-101.

Mannos 74 Mannos, J. L. and D. J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images," *IEEE Transactions on Information Theory*, IT-20, Number 4, July 1974, pp. 525-536.

Pratt 78 Pratt, W. K., *Digital Image Processing*, John Wiley and Sons, New York, 1978.

Robson 66 Robson, J. G., "Spatial and Temporal Contrast-Sensitivity Functions of the Visual System," *Journal of the Optical Society of America*, Volume 56, August 1966, pp. 1141-1142.

Warnock 80 Warnock, J. E., "The Display of Characters Using Gray Level Sample Arrays," *Computer Graphics*, Volume 14, Number 3, July 1980, pp. 302-307. SIGGRAPH 1980 Proceedings.

Westheimer 86 Westheimer, G., "The Eye as an Optical Instrument," *Handbook for Perception and Human Performance, Volume 1. Sensory Processes and Perception*, Eds. Boff, K. R., Kaufman, L., and Thomas, J. P., John Wiley & Sons, 1986, pp. 4.1-4.20.