

Improving Color Text Sharpness in Images with Reduced Chromatic Bandwidth

Scott Daly[♦]*, Jack Van Oosterhout, and William Kress*
Digital Imaging Department, [♦]*Digital Video Department*
Sharp Laboratories of America
Camas, WA USA

Abstract

Current digital imaging systems often reduce the chromatic bandwidth to speed transmission or save storage. While this has become routine for most video applications, it is also becoming true of digital camera systems. Chromatic bandwidth is usually reduced by converting the RGB image into a single luminance and two color difference images, and then subsampling the color difference images. The motivation lies in the reduced bandwidths of the visual system for the color signals, as evidenced by contrast sensitivity function (CSF) data. Systems using chromatic subsampling can exhibit visual equivalence to otherwise identical non-subsampled systems, despite the physical color blur introduced by subsampling. While this equivalence holds true for images of real scenes and optically captured scenes, problems arise with synthetically generated imagery, such as color geometrics and type fonts. The possible sources of these problems include: 1) deviation from natural image chromatic power spectra, 2) nonlinear masking effects between the luminance and chrominance visual mechanisms, and 3) filtering processes in the luminance-chrominance domain create signals outside of the RGB gamut. By taking advantage of nonlinear perceptual effects and the gamut clipping issues, we have developed a simple point-processing algorithm that accomplishes a reduction in the visibility of chromatic blur.

Introduction

Since the early days of color television it has been known that the spatial bandwidths of human color mechanisms are substantially less than that for the luminance [1]. This is most dramatic when the color signals are formed as isoluminant color differences, generally having a red-green (R/G) and a blue-yellow (B/Y) modulation. By recombining these color difference signals with the luminance signal at the receiver, the original RGB image can be reconstructed. Modern data is shown below in Figure 1 for the contrast sensitivity functions (CSF) of the visual system[1] to luminance and chrominance signals [2]. The maximum perceived frequencies, or cut-off frequencies, of this data suggest that the bandwidths of luminance to R/G to B/Y are nearly: 1 : 0.5 : 0.5.

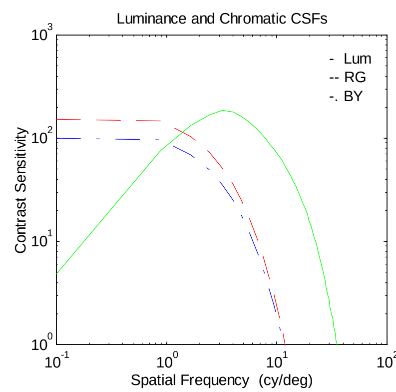

Figure 1: Luminance and Chromatic CSFs of the visual system

Different color spaces for the chromatic modulations have been used in both vision science and imaging system design. Examples include YIQ, YUV, ATD, and Y/R-Y/B-Y. One of the most predominant is the L*a*b* representation, where a* is the red-green modulation and b* is the blue-yellow. Although these are not the exact color directions currently believed to be used by the visual system, it is still a useful space within which to work. Although the L*a*b* representation was intended for color reproduction issues and has zero frequency components, an extension referred to as Spatial-CIELAB [3], incorporates spatial frequency. In this extension, the three CSFs shown in Figure 1 are modeled as Gaussian low-pass filters and applied to the L*, a*, and b* layers of an image. This is considered a perceptually uniform image space from which spatial ΔE values can be computed. However, there are still unresolved issues with using the a* and b* layers of an image for the main color difference signals in human vision. These relate to the order of

[1] The CSF is roughly analogous to a spatial frequency response.

nonlinearity between the LMS cone action spectra (≈XYZ) and the cones' static nonlinearity (≈ the cube-root in L*) prior to the formation of the opponent color signals (≈a* and b*) used to represent color in Spatial CIELAB.

Engineering uses of the bandwidth effects of the luminance and two chrominance CSFs existed before the effects were well understood. Although early use was made by an analog low-pass filter for video, the predominant approach now is by subsampling the color difference layers of a digital image. The bandwidth reductions shown in Figure 1 suggest a scheme that subsamples the chromatic layers by 2 in each direction (with appropriate low-pass filtering). This commonly used technique works well on natural images or optically captured contone images. In such images, the distortions due to subsampling the chrominance[2] are not visible to the eye. In addition, when high resolution text (i.e., greater bandwidth than the contone image) is inserted into the continuous tone images, the sharpness of the black, grey and white text is not affected.

Chromatic Bandwidth Reduction Problems

Unfortunately, there are two main problems caused by chromatic subsampling (or other chromatic spatial bandwidth reduction). These occur on images containing color text or graphical objects with binary edge transitions. One problem manifests itself as a visible chromatic blur (also referred to as color bleeding) on text[3] of saturated colors. This visibility is dependent on the grey level, where it is most visible on white, and not visible at all with chromatic text on black backgrounds. The second problem is that of overshoots or undershoots around the edges (often referred to as haloes). These are similar to ringing artifacts but consist of only one "ring".

One could ask why the visual theory derived from Figure 1 fails in this case of colored text or graphical objects. This problem occurs even if the Spatial-CIELAB domain is used, in which case it fails both along the luminance axis[4] as well as the spatial frequency axis[5].

[2] We used a* and b* of CIELAB, but any color difference signal would give similar results.

[3] The term "text" will refer to both text and graphics objects.

[4] Because the chromatic blur is not equally visible throughout the L* grey scale.

[5] Because the chromatic blur is visible with a spatial subsampling of only a factor of 2.0.

Secondary Visual System Attributes

The failure of the visual theory derived from the three CSFs is that it is over-simplified. The under-prediction of color bleeding is primarily due to viewing distance issues, the center of chromatic modulation, and cross luminance-chrominance facilitation. These three aspects will be discussed below. The other distortion, the overshoot adjacent to sharp edges, is primarily a gamut-clipping effect and results from the low-pass filtering step which creates values outside the range of the original RGB (or XYZ). Unless special precautions are taken into account when converting from a color transform such as L*a*b* back to a directly displayable color format, orthogonal projections of the out-of-gamut color to the display space results in local luminance and hue shifts. These appear as halo artifacts. Due to the limited space, we will only expound on the perceptual issues.

Viewing Distance and Nyquist Mapping

The common engineering misunderstanding relating to the CSFs of Figure 1 is to believe that since the cut-off spatial frequencies are in the ratio of approximately 1 : 0.5 : 0.5 for luminance to R/G to B/Y, a system designed with those relative bandwidths will not exhibit chromatic blur. This is only true if the system's luminance Nyquist is high enough that it exceeds the luminance cut-off visual frequency. The mapping of the system's Nyquist to visual frequency in cy/deg depends on both the maximum physical frequency as well as the viewing distance. If the Nyquist maps to less than the visual cut-off frequency (in such a chromatically subsampled system), the individual Nyquists of the luminance and two chrominance planes all intersect their respective CSFs. The consequence is that the undersampling with respect to luminance will result in luminance blur, the undersampling with respect to the R/G CSF will cause a R/G chromatic blur, and similarly for the B/Y plane. However, in such systems which have perceivable luminance blur, the color blur is usually not noticed for natural or optically captured images. This is probably due the lack of spatial power spectra at the higher frequencies for the color difference images as well as the secondary visual effects of luminance masking chrominance at high luminance contrasts. However, with images containing digitally generated text, the chromatic blur can be quite visible.

Center of Chromatic Modulation

Most of the physiologically-motivated studies [2,4] of the chromatic CSFs modulate the R/G signals about a yellow mean level; assumed to be the best isolation of the visual luminance signal. However, since the space-time average of most images is grey, engineering-motivated studies [5,6] modulate about grey. Substantial differences result between these two approaches. Modulation about grey results in R/G and B/Y CSFs which are bandpass rather than lowpass. Also, the bandwidths of the R/G and B/Y signals are much less relative to luminance[6]. While these effects can be partly explained by luminance contamination of the R/G and B/Y signals due to chromatic adaptation and gamma effects, there is still debate over which is the most appropriate for practical imaging. The bandpass shapes may represent an increased R/G and B/Y sensitivity over the physiological studies' CSFs obtained using a yellow mean.

Cross Luminance-Chrominance Masking & Facilitation

As mentioned earlier in this section there are masking effects that act across luminance and chrominance channels of the visual system [7]. While high contrast luminance signals can mask the chrominance signals, low- to mid-luminance contrasts actually facilitate (reduces the threshold) chrominance signals of similar frequencies, making them easier to see. This is shown in Figure 2 below, where the R/G threshold is plotted as a function of a luminance masking contrast. For luminance mask contrasts in the range of 1% to 20%, the R/G threshold is actually reduced.

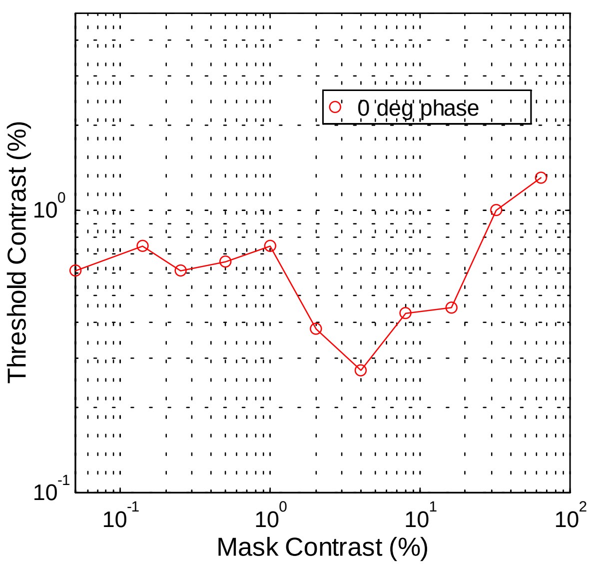

Figure 2: R/G Threshold vs. Luminance Mask, data from [7]

We believe this is the primary cause of the chromatic blur problem. Since isoluminant text are rarely used due to their illegibility, the chromatic text contains both luminance and chromatic contrast. The areas around a color edge have low luminance contrast for the frequencies higher than the chromatic CSF cut-off frequency. The difference between the original and the subsampled chromatic signal (i.e, the color blur) also consists of frequencies higher than this cut-off. Without any interaction from the luminance signal, this chromatic difference signal would not be detected. However, the facilitation of the chrominance signal by the lower contrast luminance signals of similar frequency, as shown in Figure 2, reduces the threshold of the chrominance signal. The chrominance signal due to the color blur now lies above this reduced threshold, and is visible.

Chromatic Resolution Enhancement Algorithm

Our goal is to correct for the subset of imagery where chromatic subsampling based on the simpler CSF visual model fails due to the reasons described in the previous section. Consequently, we concentrated on text and graphical objects. This acted to differentiate our work from many television applications toward increasing color resolution. The basic idea of our algorithm is to use the luminance signal as a control signal for the processing of the chrominance signals. In particular, we limit the range of the magnitude of the chrominance signals according to the luminance level, as shown in Figure 3. This is motivated by our observation that the chromatic blur was most visible on a white background. Also, the color gamut in the luminance-chrominance space narrows to a point at the minimum and maximum luminance values. That is, the values should be zero for both chrominance signals at the luminance minimum and maximum.

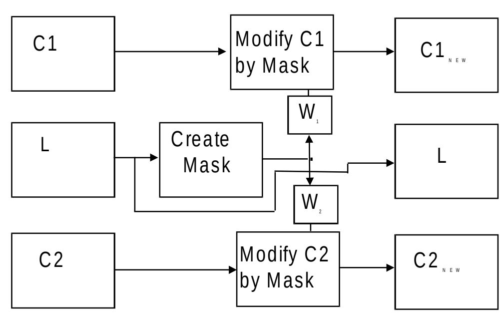

Figure 3: Algorithm with luminance-controlled mask function

Since we are mostly concerned with chromatic text on a white surrounding area, which is the most visible as well as most common case, we can remove the color bleeding from the chromatic characters with the following algorithm. Since the luminance signal is not distorted by the chromatic subsampling and we know that both chrominance signals should be zero in the white area surrounding the

[6]One study [6] has luminance to RG to BY bandwidths of 1:0.3:0.1

text, our algorithm forces them to zero if the control signal (mask) from the pixel's luminance indicates that it is white or near white. The consequence is that we can achieve a spatial resolution enhancement in the chromatic text without resorting to spatial operators, such as filtering and other local spatial operators. The available high resolution luminance signal is used to invoke a higher resolution into the chrominance signal. The scheme is essentially a point-processing scheme across the L*, a*, and b* images, generically indicated in the block diagrams as L, C_1, C_2. Consequently, it is computationally inexpensive, especially as compared with other spatial approaches.

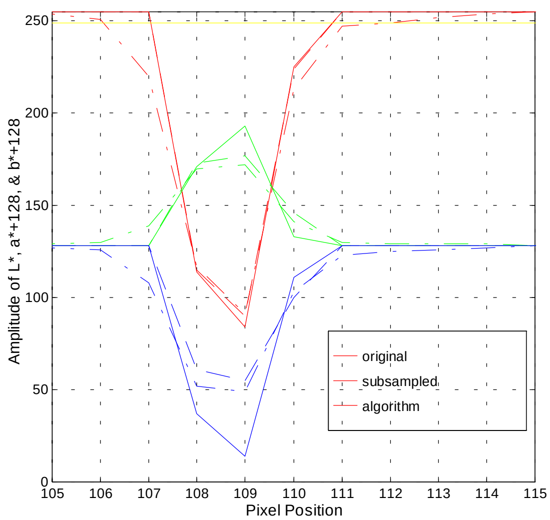

Figure 4: Algorithm effects on L, a*, and b* fontstroke profiles*

The resulting resolution enhancement is exemplified in Figure 4. In this figure, the curves are the cross-sections of a chromatic text fontstroke. The uppermost three curves are L* signals, the middle three are a* signals, and the lowest are b* signal traces. The original L*, a*, and b* chromatic text signals are shown as solid lines, while the subsampled (and upsampled back to original resolution) L*, a*, and b* signals are shown as the long-dash short-dash curves. The result of the algorithm for the L*, a*, and b* signals are shown as the dashed curves[7]. One can see that the algorithm's a* and b* signals have steeper slopes than the subsampled signals, resulting in a sharper appearance. They also return to 0 (i.e., 128 in this plot) nearer to the original stroke edge than the subsampled signals, corresponding to less visible color spread (or color bleeding). Note that these signals are taken after the L*a*b* image has been converted to RGB (then converted back to L*a*b* for this analysis). This is why the figure shows a blurring of even the non-subsampled luminance signal if the algorithm is not used. The algorithm strongly reduces any distortions introduced into the luminance signal by the L*a*b* to RGB mapping.

The algorithm is generally applied to the L*, a* and b* images after a* and b* are interpolated to the L* size from a compressed image buffer. It is also possible, however, to use the algorithm as a postprocessing technique to clean up color text from any source, such as digital video or digital still cameras that may output in the RGB format.

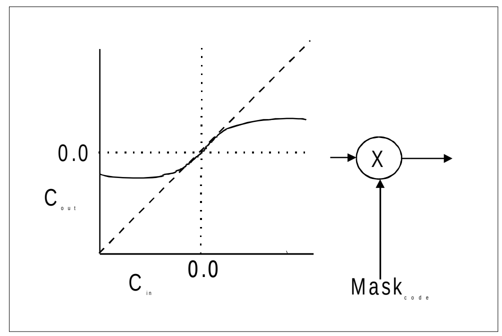

Figure 5: Modification of chrominance signal by the mask signal

The details of the modification of the chrominance images are shown in Figure 5. The simpler version use a linear transfer of C_{in} to C_{out} (dashed line) and multiplies this result by the mask signal. This merely scales the chrominance signal about zero (where the image is achromatic) by the mask. In this approach, the mask essentially controls the resulting slope as governed by the equation,

$$C_{out} = C_{in} \bullet C_{mod} \bullet Mask \qquad (1)$$

where C indicates the chromatic images from either a* or b*, C_{out} is the chromatic image's pixel value, C_{in} is the pixel's input value having the chromatic bandwidth reduction, C_{mod} is the dashed line in Figure 4, and Mask is a pixel dependent value to be described shortly. Independent control over the way the mask signal affects the C1 and C2 chrominance images signals is accomplished with the weights shown in Figure 3. A more detailed approach is to use the mask to control the degree of limiting of the C_{mod} function, shown as the solid line in Figure 5. This is accomplished by using the solid line transfer function on the chrominance signal which is subsequently multiplied by the mask. The net effect is that the output chrominance value at each pixel is given by the following equation:

$$C_{out} = C_{in} \bullet C_{mod}(Mask) \qquad (2)$$

[7] This is hard to see for L*, since it is nearly superimposed on the original L* curve.

where C_{mod} is the nonlinear transfer curve shown solid in Figure 5, whose shape is determined from the mask value, which directly controls its limiting values.

Figure 6 shows the details of the generation of the masking control signal generated from the luminance image, L. Again a simple version as well as a more sophisticated version are shown. The dashed line creates a simple threshold signal with binary values of 0.0 and 1.0. The threshold value is chosen so that the chrominance signal will only be affected in the brightest parts of the image, where the chromatic blur is most visible. It is chosen to be high enough that it does not affect the large-area tinted white regions, since the visual system is insensitive to the actual color of small specular highlights due to the chromatic CSFs. Lower threshold values will act to increase the magnitude of sharpening[8]. The result is that the C1 and C2 signal values are set to zero since the mask value is zero and the chrominance is scaled by the mask.

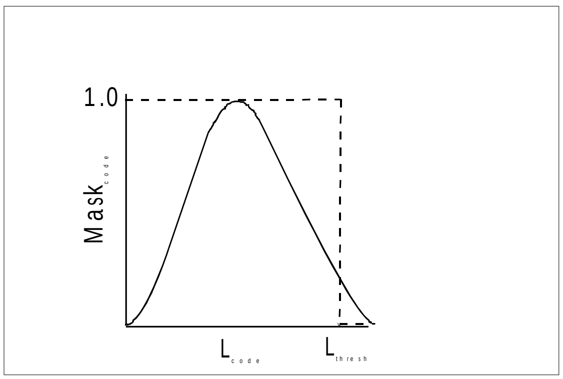

Figure 6: Generation of mask signal from luminance image

The more sophisticated version of the mask signal attempts to track the maximum chromatic extent as a function of luminance, in accordance with the overall color gamut possible at the printer. This uses a threshold as a function of grey level, shown as the solid line in Figure 6.

Conclusions

Images displayed with this algorithm demonstrate a significant decrease in the visibility of color blur. It has some difficulty with yellow text, due to its high luminance confounding the mask threshold setting. Fortunately, the use of yellow text on white backgrounds is generally avoided by the experienced user due to its illegibility.

[8] In Figure 4, a threshold value of 250 was used.

References

1. O. Schade, "On the quality of color-television images and the perception of color detail", SMPTE Journal V. 67, pp 801-819 (1958).
2. K. T. Mullen, J. "The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings," Physiol V. 359, pp. 381-400 (1985).
3. X. M. Zhang and B. A. Wandell, "A spatial extension to CIELAB for digital color image reproduction ," Proceedings of the SID Symposiums (1996)
4. D. Williams, N. Sekiguchi, and D. Brainard, "Color, contrast sensitivity, and the cone mosaic," Proc. Natl. Acad. Sci. USA V. 90, pp. 9770-9777 (1993)
5. A. Watanabe, H. Sakata, and H. Isono, "Chromatic spatial sine-wave responses of the human visual system," NHK Laboratories Note, Serial # 198, March (1976)
6. W. E. Glenn, K. G. Glenn, and C. J. Bastian, "Imaging system design based on psychophysical data," Proceedings of the SID, Vol 26, pp. 71-78 (1985).
7. E. Switkes, A. Bradley, and K. De Valois, "Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings", JOSA A. V. 5, pp1149-1162