A visual model for predicting chromatic banding artifacts

Gyorgy Denes[1], George Ash[1], Huameng Fang[2], Rafał K. Mantiuk[1]
[1] University of Cambridge (UK), [2]Huawei Technologies Co., Ltd. (CN)

Abstract

Quantization of images containing low texture regions, such as sky, water or skin, can produce banding artifacts. As the bit-depth of each color channel is decreased, smooth image gradients are transformed into perceivable, wide, discrete bands. Commonly used quality metrics cannot reliably measure the visibility of such artifacts. In this paper we introduce a visual model for predicting the visibility of both luminance and chrominance banding artifacts in image gradients spanning between two arbitrary points in a color space. The model analyzes the error introduced by quantization in the Fourier space, and employs a purpose-built spatio-chromatic contrast sensitivity function to predict its visibility. The output of the model is a detection probability, which can be then used to compute the minimum bit-depth for which banding artifacts are just-noticeable. We demonstrate that the model can accurately predict the results of our psychophysical experiments.

Introduction

Novel display technologies, such as HDR monitors and VR headsets, require the design of new color spaces that can accommodate larger dynamic ranges and wider color gamuts. It is a significant challenge to ensure that these newly-designed color spaces interact well with image and video compression.

To fit within a discrete, digital representation, color signals are quantized to a target bit-depth using some color space and electro-optical transfer function. Quantization of this sort can introduce banding artifacts, also known as quantization artifacts or false contours. These banding artifacts are most prominent in images containing smooth, low texture regions, such as skin, or the sky and water in Figure 1. With insufficient bit-depth, smooth gradients in luminance and chrominance are perceived as wide, discrete bands.

One particularly desirable property of a color space intended for compression is perceptual uniformity; i.e. when quantized to a certain bit-depth, perceivable banding artifacts are equally likely to occur in all parts of the color space. A visual model that can predict banding artifacts robustly could aid to assess the uniformity of existing and new color spaces.

Commonly used visual difference predictors (VDP, HDR-VDP) operate on luminance information only, therefore cannot measure chromatic artifacts. Other quality metrics, including PSNR and SSIM, do not model the perceptual phenomena that are responsible for the visibility of banding.

In this paper we introduce a visual model for predicting the visibility of banding artifacts for both luminance and chrominance gradients. The model analyzes the error signal from quantizing gradients in the target color space, comparing it with the contrast sensitivity function of the human eye. The output of the model is a detection probability that can be then used to compute the minimum bit-depth for which the contouring artifacts are just-noticeable. We conducted multiple experiments and found that the model could well predict the collected data. Additionally, when analyzing the bit depth predictions of our model, we found a similarity to DeltaE 2000 color differences and MacAdam Ellipses. To emphasize the applicability to novel display techniques, we performed all experiments using a virtual reality headset.

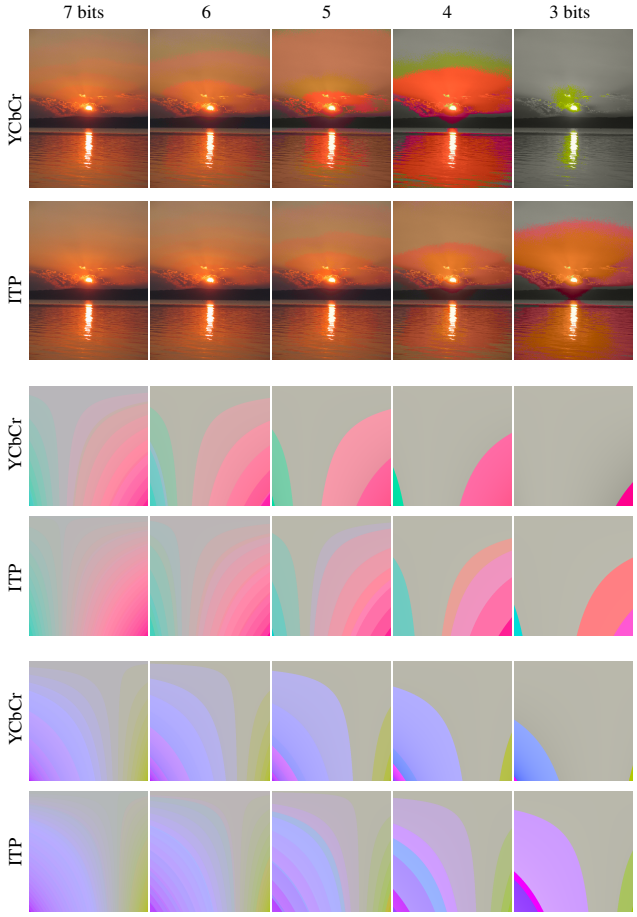

Figure 1. *Increasing severity of banding artifacts when quantized to a range of bit-depths in different color spaces (YCbCr, ITP). Consistently with our results, ITP results in less severe banding artifacts at the same bit-depth.*

In the next section we review the existing literature related to banding artifacts and quantization, give some background on the topic of contrast sensitivity, then we describe the details of our model, the experiments and the parameter fitting.

Related work

A number of published works address the visibility and subsequent correction of banding artifacts [1, 4, 2, 3].

Some authors consider the problem from an image processing points of view [1] without a perceptual calibration. Wang et al. [4] demonstrate that such image processing methods can be fine-tuned to better correlate with subjective (mean opinion score) measurements, however, existing works take no account of the complex structure of visual perception, and are unlikely to generalize well to a color space agnostic setup.

Daly et al. [2, 3] present a technique for achieving bit-depth extension via spatio-temporal dithering. Their proposed technique utilizes the contrast sensitivity limitations of the human visual system to evaluate and recommend perceptually-ideal dithering patterns. In particular, [3] analyzes the error arising from the quantization of a smooth gradient (Figure 2) in both the spatial and the frequency domains. Our visual model builds on this analysis and extends it for chromaticity.

Previous studies have highlighted the importance of quantization-related banding artifacts when considering novel encoding schemes [5, 6]. Boitard et al. [7] conducted extensive experiments to establish the minimum bit-depths in different color spaces. However, these studies do not provide a visual model for predicting the visibility of banding artifacts.

Our visual model is inspired by perceptually-motivated visual difference predictors, such as VDP and HDR-VDP [8, 9]. As described in the following section, we similarly rely on the contrast sensitivity function, and employ probability summation to aggregate different frequency bands. However, the existing models do not account for chromatic differences and are therefore not suited to predict banding artifacts in color images.

Background

When considering monochrome signal detection, the threshold of human visual system (HVS) in known to change with the average background luminance, the contrast, and the spatial frequency of the signal among other factors. Contrast in this context is commonly defined as $\Delta L/L$, i.e. the luminance difference between the signal and the background divided by the background luminance. The contrast threshold is the smallest amount of luminance difference that can be detected by human visual system given an average background luminance. The contrast threshold (or inversely the contrast sensitivity) is dependent on the background luminance, and is non-uniform across spatial frequencies. A number of authors have measured the contrast sensitivity function (CSF) for a variety of visual conditions as a function of spatial frequency . In this paper we rely on Barten's parametric model [10] to describe the achromatic CSF (CSF_A).

Existing chromatic contrast sensitivity functions do not account for changes in sensitivity with luminance, which are essential for our modelling. For that reason, we fitted a custom chromatic contrast sensitivity function to the data from [11] using our custom cone contrast definition for C_R and C_Y.

Model for monochromatic banding

Banding artifacts are the most visible when quantizing smooth gradients. Therefore, to simplify the problem formulation, we devise a model that targets this the worst-case scenario of smooth gradients. Initially we will consider the monochromatic case using the achromatic CSF_A, then generalize to the chromatic model in the following section.

Our predictor models the detection of banding based on the

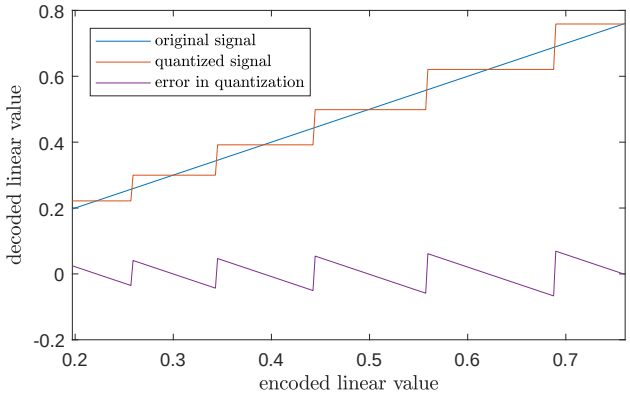

Figure 2. *Illustration of the error signal (purple) between the original (blue) and the quantized (red) signals. The error signal approximates a saw-tooth function.*

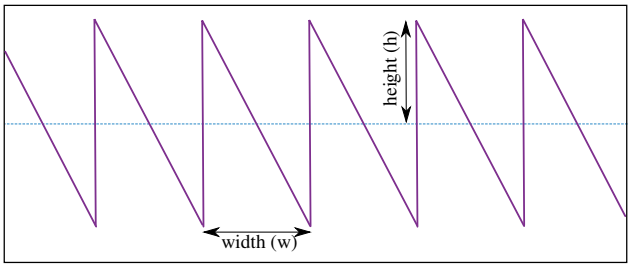

Figure 3. *Analytical model of the error signal. With a known width (w) and height (h) we can find the frequency components of the signal.*

CSF with probability summation in a manner similar to VDP and HDR-VDP. Our model does not include the effect of contrast masking, which could further elevate the detection threshold. Masking was excluded, as our goal was to predict the worst-case scenario for any kind of content.

The predictor consists of the following steps: (1) we first determine a set of spatial frequencies that are present in the contouring signal, (2) then we use those frequencies and CSF_A to determine how sensitive the eye is to each spatial frequency component. (3) We use a psychophysical function to convert such sensitivities into probabilities of detection. (4) Finally, we pool probabilities across spatial frequency components using probability summation. In the following paragraphs we explain the above steps in more detail.

To determine spatial frequencies of the contours, we analyze the Fourier transform of the difference signal between the smooth and contoured gradients. We observe that banding artifacts over a smooth gradient can be approximated with a saw-tooth pattern (see purple line in Figure 2). The Fourier transform of a saw-tooth pattern can be expressed analytically as:

$$f(x) = \frac{1}{2} - \frac{1}{\pi}\sum_{k=1}^{\infty}\frac{h}{k}\sin\left(\frac{k\pi x}{w}\right), \tag{1}$$

where w is the width (period) and h is the amplitude as in Figure 3. The amplitude of the frequency components is then:

$$\alpha_k = -\frac{h}{k\pi} \qquad \text{for } k = 1,2,\ldots \tag{2}$$

and the frequency of each component is:

$$\omega_k = -\frac{k\rho}{w}, \tag{3}$$

where ρ denotes the angular resolution of the device in pixels per visual degree and w is the width of the saw-tooth in pixels. We found that Fourier components for $k > 16$ were insignificant and did not improve the model's accuracy.

To compute the probability of detecting each Fourier component of the contour, we first determine the sensitivity to that component from Barten's CSF:

$$S = \frac{L_b}{\Delta L_{det}} = CSF_A(\omega, L_b), \tag{4}$$

where ω is the spatial frequency, L_b is background luminance, and ΔL_{det} is the detectable amplitude of that frequency component. Then, we normalize the contrast of the contouring pattern (a_k/L_b) by multiplying by the sensitivity so that the normalized values are equal to 1 when the k-th frequency component is just detectable. The normalized contrast is given by:

$$c_k = \frac{a_k}{L_b} CSF_A(\omega_k, L_b). \tag{5}$$

Next, we transform the normalized contrast into probabilities using a psychometric function:

$$P_k = 1 - \exp(\ln(0.5) c_k^{\beta}), \tag{6}$$

where β is the slope of the psychometric function. We found that with our model, $\beta = 2$ produced more uniform results than the typical estimate of $\beta = 3.5$.

Finally, we pool the probabilities P_k across all Fourier components using probability summation

$$P = 1 - \prod_k (1 - P_k) \tag{7}$$

To determine the minimum bit-depth that does not result in contouring artifacts, we perform binary-search for a bit-depth that would result in $P = 0.5$.

Model for chromatic banding

The model, as described above, accounts only for banding due to changes in luminance. While changes in luminance can be a large contributor to the visibility of banding, changes in chromaticity can also have an impact. Here we extend the model to account for this using a chromatic contrast sensitivity function, and probability summation across all visual channels. Our chromatic color discrimination model takes a color gradient, specified in the CIE XYZ (1931) color space, and predicts the probability of observing a banding artifact when the smooth gradient is quantized to a given bit-depth.

First, we convert both colors from XYZ into LMS space (assuming CIE 1931 color matching functions). Each channel of this tri-stimulus space is proportional to the response of the long, medium and short cones of the retina. It should be noted that there is no standard way to scale the absolute response of each cone type and the response values are only relative. To convert CIE XYZ trichromatic values into LMS responses we use the following linear transform:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.15514 & 0.54312 & -0.03286 \\ -0.15514 & 0.45684 & 0.03286 \\ 0 & 0 & -0.00801 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \tag{8}$$

The cone responses are further transformed into opponent responses of color vision mechanism: one achromatic (black-to-white) and two chromatic: red-to-green and yellow-green-to-violet. The exact tuning color directions of those mechanisms are unknown, but such knowledge is not relevant for our considerations. We use one of the simplest formulae commonly used to compute color opponency:

$$\begin{bmatrix} A \\ R \\ Y \end{bmatrix} = \begin{bmatrix} L+M \\ L-M \\ S \end{bmatrix}, \tag{9}$$

where A is achromatic (luminance) response, R is the red-to-green response and Y is the yellow-green-to-violet response.

Given two colors to be discriminated, we need to compute contrast between them. Since there is no single way to represent color contrast, we experimented with a number of expressions to find the most suitable for our model:

$$C_A = \frac{|A_2 - A_1|}{A_1}, C_R = \frac{|R_2 - R_1|}{\alpha|R_1| + (1-\alpha)A_1}, C_Y = \frac{|Y_2 - Y_1|}{\alpha|Y_1| + (1-\alpha)A_1}, \tag{10}$$

where α is a free variable to be optimized for the experiment data. Given the color contrast components C_A, C_R, C_Y, we follow the same steps as for the prediction of luminance banding: (1) we multiply each color contrast by the corresponding contrast sensitivity function from [11] and the Fourier coefficients of the saw-tooth pattern, a_k, (2) transform normalized contrast to detection probability, (3) then apply probability summation across all frequencies and the A, R, Y color channels.

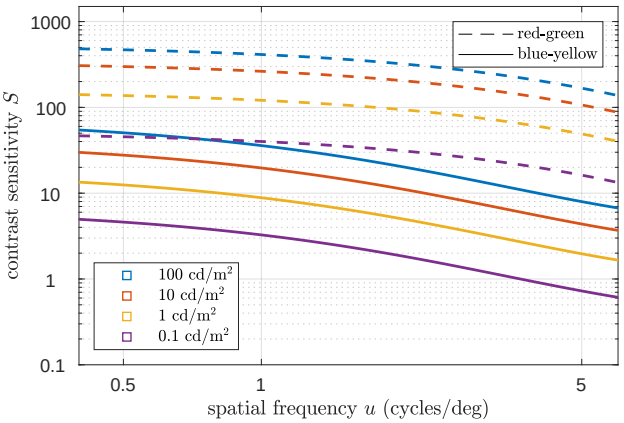

Figure 4. *Chromatic contrast sensitivity function based on [11]. Dashed and solid lines show the sensitivity to red-green and blue-yellow change respectively at different background luminance levels.*

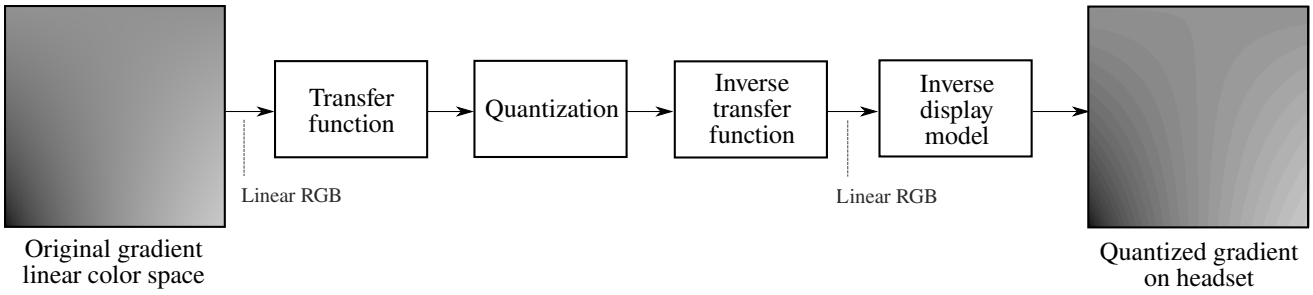

Figure 5. *Stimulus generation pipeline for the experiments. Smooth gradient images with varying contrast levels are generated in linear space, then transferred to any arbitrary color space using a transfer function before quantization. After quantization, the inverse of the transfer function is applied. The inverse display model (GOG) is used to compensate for the display characteristics before submitting to the device.*

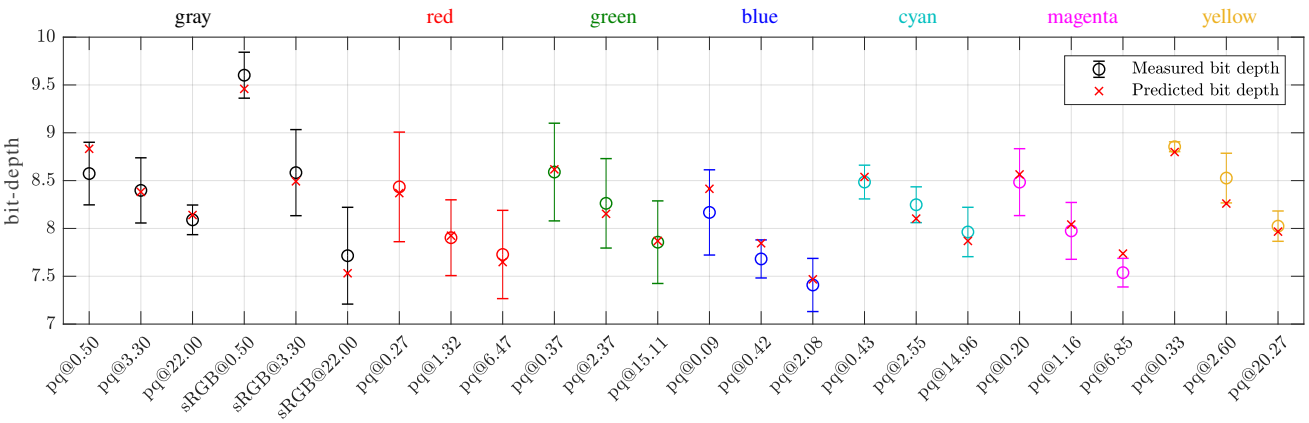

Figure 6. *Results of the monochromatic quantization experiment. Each label on the x axis denotes a stimulus as [encoding]@[mean_luminance]. E.g. pq@0.50 indicates a gradient with a mean luminance of 0.5 cd/m^2 quantized after PQ encoding. The color of the error bars indicate the chroma of the stimulus. The number of required bits decreases with increased luminance. Our model (red crosses) predicts the values accurately regardless of the color of the stimuli.*

Experiment 1: monochromatic quantization

The exact shape of CSF_A depends on a number of free variables in Barten's model [10] including the luminance of the adapting field, the angular size of an object, the background luminance and the viewing angle besides other factors. We ensured that our contrast sensitivity function matches our application scenario's viewing conditions by conducting a psychophysical experiment. As we target novel display technologies, we used a Huawei Mate Pro 9 with a DayDream VR headset (peak luminance of 44 cd/m^2). We used a spectroradiometer to measure the display properties, then fitted a gamma-offset-gain model [12]. The experiment was performed in a dark room to minimize the effect of external light sources.

Table 1: Chromaticity co-ordinates of the monochromatic smooth gradient images in CIE L'UV.

Color	u	v
Gray	0.1950	0.4581
Red	0.4242	0.5186
Green	0.1044	0.5642
Blue	0.1768	0.1651
Cyan	0.1240	0.4281
Magenta	0.2827	0.3168
Yellow	0.1969	0.5529

Stimuli

We used a number of monochrome smooth-gradient images; an example of which is shown on the left of Figure 5. Each image consisted of rows of gradients, where contrast varied from 0 (top) to 1 (bottom). Given the mean luminance level L, the luminance of all pixels in the top row were equal L, and the pixels in the bottom row varied in luminance from 0 to $2L$ on a linear scale. We measured the effect of luminance quantization at 7 chromaticities: the white point, close to the three primary colors (red, green and blue), as well as their opponent colors (cyan, magenta and yellow). The exact color co-ordinates are listed in Table 1. The stimuli were created following the processing steps shown in Figure 5. First, each linear gradient image was converted to the 0-1 range using one of the two transfer functions: *PQ* or *sRGB*. Then, the values were quantized to the sample bit-depth and converted back to linear space. Quantization levels beyond the maximum bit-depth of the display were achieved with spatio-temporal dithering. Three mean luminance levels were sampled for each chromaticity spanning the available dynamic range. The size of each square with the gradient was 20 visual degrees. The background wall had in the virtual environment had the same color as the mean value of the gradient stimulus.

Procedure

Each observer was shown four squares with smooth gradients, from which only one was quantized. For each trial, the position of the quantized gradient was randomized. The observer's task was to indicate the quantized gradient using a remote controller. The QUEST procedure [13, 14] was used with 30 trials to select consecutive bit-depths and to compute the final threshold. To minimize the effect of dark adaptation, the luminance levels were shown from darkest to brightest. Observers were given 2 minutes before the experiment to adapt to the conditions. Each observer was allowed to take any amount of time before making a decision. They could freely move their head in the VR environment. The experiment involved 9 paid participants, aged 19-45, with normal or corrected to normal color vision.

Results

The results in Figure 6 indicate that lower mean luminance levels require higher bit-depths. As expected, *PQ* offers better uniformity than *sRGB*; i.e. the deviation in bit-depth requirement for different luminance levels is smaller.

Model fitting

We formulated the CSF_A as a simplified parametric Barten model with five free variables including a relative scaling factor:

$$S(u) = p_2 \frac{5200 e^{-0.0016u^2}(1+100/L)^{0.08}}{\sqrt{\left(1+\frac{144}{p_1^2}+p_4 u^2\right)\left(\frac{p_5}{L^{p_3}}+\frac{1}{1-e^{-0.02u^2}}\right)}}, \qquad (11)$$

where u is the spatial frequency, L is the mean luminance and $p_{1..5}$ denotes the five free parameters. We optimized $p_{1..5}$ such that the error (the weighted sum of squared deviations from the mean) between our model's prediction and the observed data is minimized. We found the optimum at $p = (39.9565, 0.1722, 0.4864, 120.3724, 0.8699)$. The crosses in Figure 6 show that the model can closely approximate the observed data, with the predictions always lying within the error bars. An interesting observation to make is that luminance gradients with different chromaticities can be well predicted by this luminance-only detection model. Attempts to optimize PSNR for the same dataset yielded errors an order of magnitude greater than our model.

Experiment 2: chroma quantization

The second experiment measured the maximum amount of chroma quantization that does not result in detectable differences. We investigated the effect of chroma quantization using two colorspaces: $Y'C_bC_r$ and IC_tC_p [15]. Both of these aim to decorrelate luminance from chrominance, and although Lu et al. [15] claim IC_tC_p offers better luminance-chrominance decorrelation and a more perceptually uniform chroma encoding, the former is more widely adopted at the time of writing.

Our setup and experimental procedure were identical to the first experiment. The stimuli consisted of equiluminant smooth image gradients at three fixed luminance levels in the CIE $L'u'v'$ color space. Two line segments were selected in the $u'v'$ plane (Figure 8). The line segments were parallel to the chroma axis, bound by the device's color gamut, and intersected at the white point. For both line segments, the smooth gradient stimuli were generated in a similar manner to the achromatic gradient used in Experiment 1. Color saturation was zero at the top of the image, and increased linearly down the image, where it was maximal at the bottom (see rows 3-6 in Figure 1).

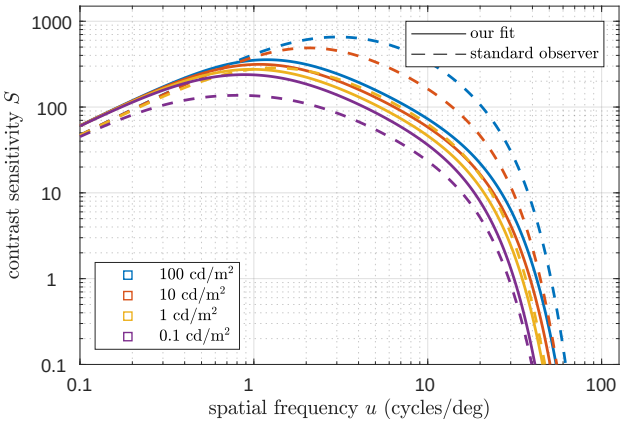

Figure 7. *Fitted achromatic contrast sensitivity function compared with the standard observer.*

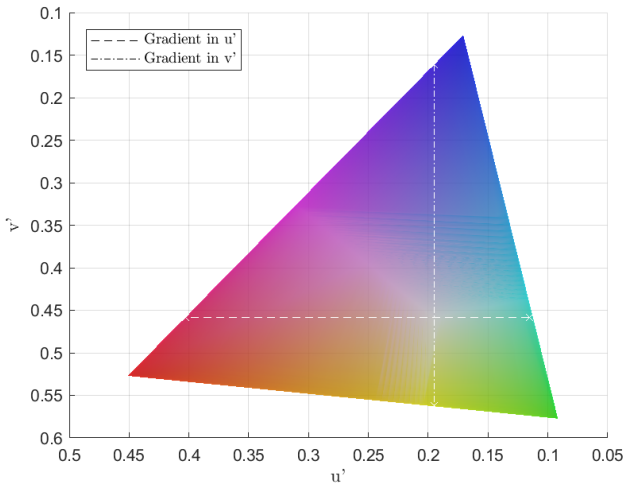

Figure 8. *Visualization of the color lines (u: horizontal, v: vertical) for Experiment 2. Both lines cross the white-point of the display, and are orthogonal to the chroma dimensions of CIE L'u'v'.*

Results

The results in Figure 9 indicate that both color spaces are approximately perceptually uniform, as the number of required bits does not change with luminance levels or across color channels. However, IC_tC_p requires consistently fewer bits than $Y'C_bC_r$.

Model fitting

The only free parameter of the chromatic model was the α variable determining the exact cone-contrast formula. We found that $\alpha \approx 2/3$ provided the best fit. The crosses in Figure 9 show that the model can closely approximate the observed data, with the predictions always lying within the error bars.

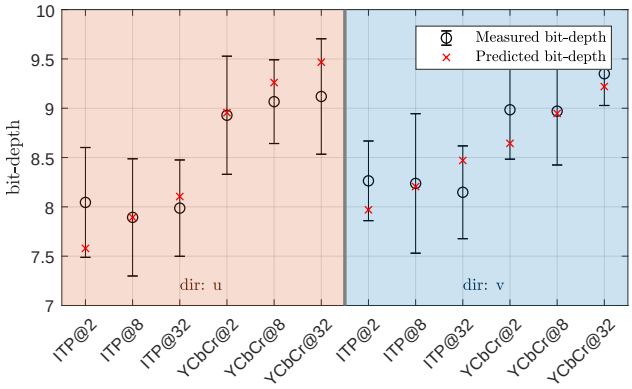

Figure 9. *Results of the chromatic quantization experiment for two color line segments: u (left) and v (right). Labels denote [color_space]@[mean_luminance]. E.g. ITP@2 is quantization in the IC_tC_p color space for a stimulus with a mean luminance of 2 cd/m^2. $Y'C_bC_r$ requires more bits than IC_tC_p. Our model predictions (red crosses) are within the error bars (95% confidence) of the observed data.*

Model predictions

The model can be trivially extended to take a starting color and a color direction vector as input (instead of the smooth image gradient). Binary search can then establish the color along the color direction vector for which the probability of detectable banding artifacts is equal to 0.5. We use such extension of the model to establish a detection threshold and to plot color uniformity ellipses akin to MacAdam ellipses. Figure 10 compares the predictions of our model to CIE DeltaE 2000 difference and to the original MacAdam ellipses. Note that our model is meant to provide better predictions for banding rather than predicting traditional color patch difference; hence it is an interesting observation that the resulting shapes are comparable.

Conclusion

We presented a visual model to predict the visibility of banding artifacts for arbitrary color spaces taking both luminance and chrominance into account. The model parameters were adjusted for the results of a subjective user experiment containing monochrome and color gradients. Our model provided a good fit for the observed data. We also used quantization artifacts as a proxy for color space uniformity and found a close match to MacAdam ellipses. One of the main limitations of this current model is that it only considers the limited dynamic range of the test device. We wish to address this in the future, then utilize the model to design a more uniform, efficient color space.

Acknowledgments

This work was supported by Huawei Innovation Research Program (HIRP HO2016050002BR) and EPSRC research grant EP/P007902/1.

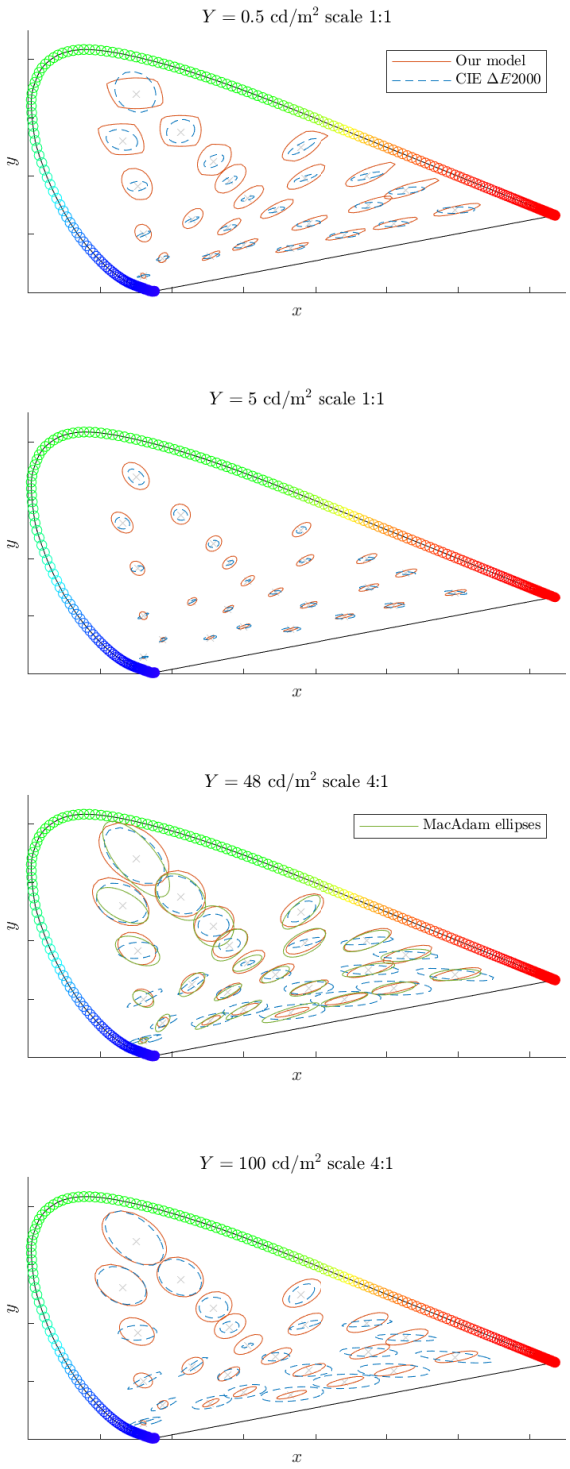

Figure 10. *Predictions of our model compared to CIE DeltaE 2000 and MacAdam ellipses [16]. Each plot corresponds to different luminance level. Note that MacAdam ellipses were measured only for the background luminance of 48 cd/m^2*

References

[1] Ji Won Lee, Bo Ra Lim, Rae-Hong Park, Jae-Seung Kim, and Wonseok Ahn. Two-stage false contour detection using directional contrast and its application to adaptive false contour reduction. *IEEE Transactions on Consumer Electronics*, 52(1):179–188, Feb 2006.

[2] Scott J. Daly and Xiaofan Feng. Bit-depth extension using spatiotemporal microdither based on models of the equivalent input noise of the visual system. page 455, jan 2003.

[3] Scott J. Daly and Xiaofan Feng. Decontouring: prevention and removal of false contour artifacts. *Proceedings of the SPIE, Human Vision and Electronic Imaging IX.*, 5292:130–149, 2004.

[4] Yilin Wang, Sang Uok Kum, Chao Chen, and Anil Kokaram. A perceptual visibility metric for banding artifacts. In *Proceedings - International Conference on Image Processing, ICIP*, volume 2016-Augus, pages 2067–2071. IEEE, sep 2016.

[5] Rafal Mantiuk, Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Perception-motivated high dynamic range video encoding. *ACM Transactions on Graphics*, 23(3):733, aug 2004.

[6] Scott Miller, Mahdi Nezamabadi, and Scott Daly. Perceptual Signal Coding for More Efficient Usage of Bit Codes. In *The 2012 Annual Technical Conference & Exhibition*, pages 1–9. IEEE, oct 2012.

[7] Ronan Boitard, Rafal K. Mantiuk, and Tania Pouli. Evaluation of color encodings for high dynamic range pixels. page 93941K, mar 2015.

[8] Scott J. Daly. Visible differences predictor: an algorithm for the assessment of image fidelity. page 2, 1992.

[9] Rafał Mantiuk, Kil Joong Kim, Allan G. Rempel, and Wolfgang Heidrich. HDR-VDP-2. *ACM Transactions on Graphics*, 30(4):1, jul 2011.

[10] Peter G. J. Barten. Formula for the contrast sensitivity of the human eye. pages 231–238, dec 2003.

[11] Kil Joong Kim, Rafal Mantiuk, and Kyoung Ho Lee. Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance. page 86511A, mar 2013.

[12] Roy S. Berns. Methods for characterizing CRT displays. *Displays*, 16(4):173–182, may 1996.

[13] Andrew B. Watson and Denis G. Pelli. Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33(2):113–120, 1983.

[14] Mario Kleiner, David H Brainard, Denis G Pelli, Chris Broussard, Tobias Wolf, and Diederick Niehorster. What's new in Psychtoolbox-3? *Perception*, 36:S14, 2007.

[15] Taoran Lu, Fangjun Pu, Peng Yin, Tao Chen, and Walt Husak. ITP colour space and its compression performance for high dynamic range and wide colour gamut video distribution. *ZTE Communications*, 1:32–38, 2016.

[16] W. R. J. Brown. Color Discrimination of Twelve Observers*. *Journal of the Optical Society of America*, 47(2):137, feb 1957.

Author Biography

Gyorgy Denes is a PhD student at the University of Cambridge. He previously received his BA and MA from this same institution, and also worked as a Software Engineer at Microsoft.

George Ash received his BA in Computer Science from the University of Cambridge where he currently works as a research assistant.

Rafał K. Mantiuk is a Senior Lecturer in the Computer Laboratory at the University of Cambridge. Previously he worked at Bangor University, the University of British Columbia and the Max-Planck-Institut (MPI) for Computer Science