https://doi.org/10.2352/CIC.2022.30.1.08

The perceptibility of color differences between thin lines and its application to printing imaging pipelines

Ján Morovič, Peter Morovič, HP Inc.

Abstract

Imaging is enabled by the limitations of the human visual system, which is blind to certain physical differences. The trichromacy of human vision, e.g., allows for very different materials to be combined in a way that results in the same signals triggered by the eye's cones. As a result a print can elicit the same response as a display, or a projection can yield colors like those in a painting. The limits of spatial acuity too allow for discrete patterns, e.g., those resulting from halftoning, to appear continuous. This paper turns its attention to the limits of color difference perception in stimuli with a very small subtense, such as thin lines or fine features of 3D objects. A first set of psychophysical data, obtained in an on-line visual experiment, indicates a dramatic relaxation of perceptibility thresholds when comparing very thin with thicker lines. The second half of the paper then presents printer imaging pipeline strategies that take advantage of these experimental findings to successfully render fine lines while taking advantage of the more limited sensitivity with which their specific colors are perceived.

Introduction

Color matching is important in many applications, including the imaging of architectural and engineering drawings. A key requirement there is the accurate representation of line colors, which in turn relies on knowing how they and their differences are perceived. More specifically, the question arises of how much of a difference in the color of thin lines can be perceived and there is anecdotal evidence for sensitivity being lower than when thick lines or objects with solid areas have their colors compared. This phenomenon is also related to how the colors of fine features, e.g., in 3D printed objects, are seen. In both cases the question about perceptibility of differences sets limits to what changes an imaging pipeline can make in such features before those changes result in visible differences.

While there is rich literature on color perception varying with a multitude of factors, including scale and angular subtense (Xiao *et al.*, 2010), and color differences depending on different parameters, e.g., spatial separation (Mirjalili *et al.*, 2019), there is no published work on line color differences.

As a result, an on-line visual experiment was launched in February 2022 to explore the phenomenon of color difference perception between lines of different thicknesses and the present paper reports its findings. This is followed by the presentation of various printing imaging pipeline strategies that benefit from a characterization of color difference perceptibility in these image features.

Experimental design

General approach

The experiment set out here is a first, pilot exploration of the impact of line thickness on color different perceptibility, with a hypothesis that sensitivity to the color of thinner lines is lower than that to thicker lines. The objective is to understand whether such a hypothesis is consistent with empirical data and whether there it is a robust phenomenon rather than one that is only detectable under narrow and carefully controlled conditions. As a result, a very quick experiment was designed that could be deployed on-line to a large number of observers under uncontrolled and unknown conditions. If a significant phenomenon is detected, it could then be characterized in a follow-up study with a larger and more systematically–structured set of stimuli presented under controlled conditions.

Stimuli

Using the method of constant stimuli (Gescheider, 1985), 22 line pairs were generated for two thickness pairs (two thick lines versus one thin and one thick line) and along three color change directions (lightness from a mid-gray, mostly chroma from a cyan and mostly hue also from a cyan). The colors were assigned and varied in HSB terms, which is common for the applications where line content is created (Fig. 1).

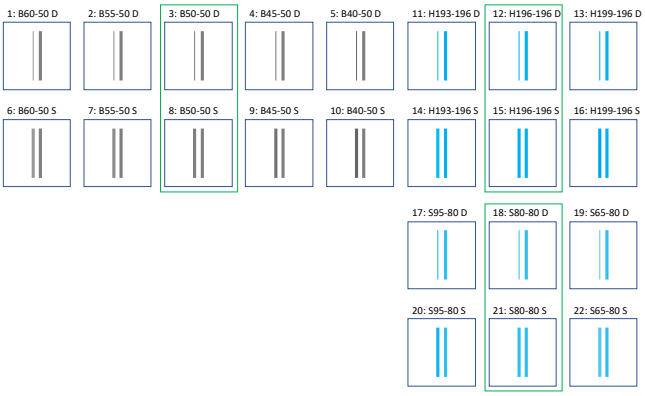

Figure 1. Line pairs varying in HSB difference dimensions and magnitudes. A green box indicates the central line pairs that had a strict color match in each of the three cases.

Tab. 1 shows the colorimetries corresponding to each of the lines from Fig. 1, assuming an sRGB display. Tab. 2 then shows ΔE2000 color differences for each of the line pairs under those assumptions.

Table 1: Line colorimetries

Line color label	Colorimetry		
	L*	a*	b*
B40	43.19	0	0
B45	48.44	0	0
B50	53.19	0	0
B55	58.25	0	0
B60	63.22	0	0
H193	70.48	-27.25	-37.43
H196	66.79	-21.49	-43.31
H199	63.11	-16.35	-49.21
S65	74.80	-24.20	-31.00
S80	71.26	-24.77	-36.70
S95	68.09	-23.07	-41.23

Table 2: Line pair color differences

Stimulus #	Label	ΔE00
1	1: B60-50 D	9.05
2	2: B55-50 D	4.74
3	3: B50-50 D	0.00
4	4: B45-50 D	4.74
5	5: B40-50 D	9.90
6	6: B60-50 S	9.05
7	7: B55-50 S	4.74
8	8: B50-50 S	0.00
9	9: B45-50 S	4.74
10	10: B40-50 S	9.90
11	11: H193-196 D	4.76
12	12: H196-196 D	0.00
13	13: H199-196 D	4.42
14	14: H193-196 S	4.76
15	15: H196-196 S	0.00
16	16: H199-196 S	4.42
17	17: S95-80 D	3.16
18	18: S80-80 D	0.00
19	19: S65-80 D	3.42
20	20: S95-80 S	3.16
21	21: S80-80 S	0.00
22	22: S65-80 S	3.42

Procedure

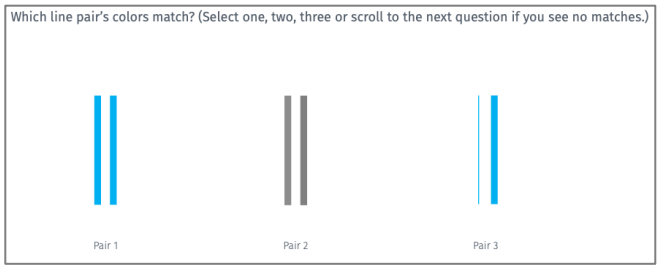

Figure 2. Stimulus presentation and instructions

11 of the 22 stimuli (2, 3, 7, 8, 11, 12, 14, 15, 17, 18 and 20) were included twice in the experiment to allow for quantifying intra-observer repeatability, resulting in a total of 33 stimuli. The order of stimuli was randomly scrambled and shown to each observer in that same, scrambled order with three stimuli displayed at any one time and the observer being instructed as shown in Fig. 2.

To obscure the structure of the underlying stimuli and the inclusion of repetitions, the 33 stimuli were labelled sequentially in the scrambled order. The experiment was made available to observers via using the *questionpro.com* platform.

Viewing conditions

Unlike in a laboratory-based experiment, viewing conditions were entirely uncontrolled and unknown in this experiment. Aspects like display size, brightness, contrast, gamut or calibration state; viewing distance, angle, temporal duration; observer color vision status, adaptation state, age, sex, gender or color matching experience were all unknown. While this does not allow for conclusions to be drawn about the effect of independent variables, or to relate the findings of this experiment to a particular set of conditions, it does provide a view of how color difference perceptibility depends on line thickness over the variety of conditions under which participants performed the present visual task.

Observers

Key to the usefulness of this experiment's findings is to have a large number of observers participate in it, as has been shown in previous visual experiments where data from large-scale, on-line, uncontrolled execution compares well with data from laboratory-based, carefully controlled conditions (cf Moroney, 2003; Mylonas and MacDonald, 2010; High *et al.*, 2021).

478 observers have completed this experiment, which is significantly higher than the typical level of in-person participation rate, which tends to be in the low tens. Fig. 3 shows the locations from which the experiment was accessed.

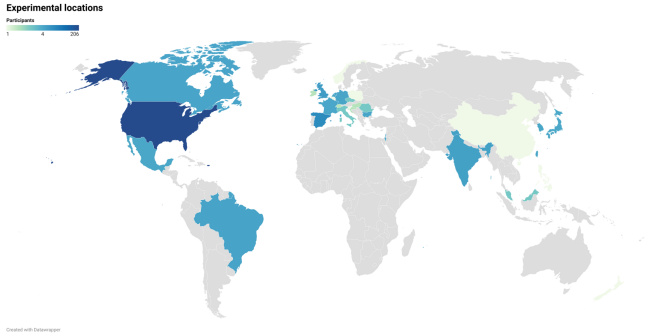

Figure 3. Locations from which observers participated in on-line experiment.

Data analysis

The raw data from this experiment are 478 binary, nominal, yes-no responses for each of the 33 stimuli for which observers had to answer the question of whether the pair of lines of a given stimulus match.

First, this data allows for a quantification of intra–observer repeatability, thanks to the inclusion of 11 repeated stimuli in the experiment. Counting how many of the 11 stimuli were judged differently the first versus the second time results in an intra-observer repeatability error percentage. E.g., an error of 9% here would mean that only one of the 11 stimuli were judged differently on the two occasions.

Second, the binary responses to each of the 22 stimuli (pooling the repetitions for 11 of them) allow for the computation of the percentage of times that the pair of lines in a given stimulus was judged to match. E.g., an 80% value here would mean that 4/5 of the responses indicate that the lines match, while the other 1/5 state that they do not match.

Third, the match percentages can be split into two sets – one for stimuli where line thicknesses are the same and the other for where they differ. Fitting a psychometric function (here the Weibull distribution was used (Watson, 1979)) to each of these sets of percentages and identifying the ΔE value at which the fitted function has a value of 50% then identifies the perceptibility threshold, and therefore the value of the just noticeable difference between no difference being seen and some, just-noticeable difference appearing (Engeldrum, 2000).

Experimental results

Intra-observer repeatability

The histograms in Fig. 4 show intra–observer repeatability error distributions for all stimuli, only the stimuli with same line thicknesses and only the stimuli with different line thicknesses. They show a clear spread in repeatability overall, with a median of 24%, and a clear difference between the same-thickness (ST) and different-thickness (DT) cases. For ST stimuli 72% of the observers made zero or one mistake, while for DT stimuli 76% of observers made up to 3 mistakes. Given the uncontrolled nature of the experiment, the overall 24% intra-observer error rate seems reasonable and since the data set is large, future analysis could compare thresholds derived from subsets of observers with different intra-observer error levels.

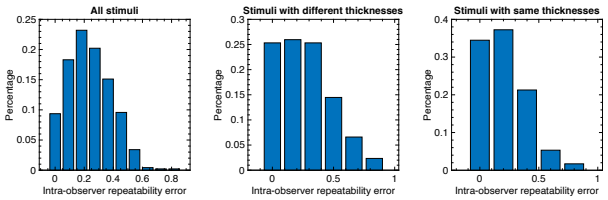

Figure 4. Intra–observer repeatability error distributions for all stimuli and per stimulus type.

Perceptibility threshold

Finally, the experiment yields data to address the question that stood at its beginning: is sensitivity to color differences lower when a thin and a thick line are compared than when two thick lines are viewed? Fig. 5 shows the ST (blue squares) and DT (red circle) match detection percentages, the psychometric functions fitted to the two data sets and the just-noticeable difference (JND) values derived from them.

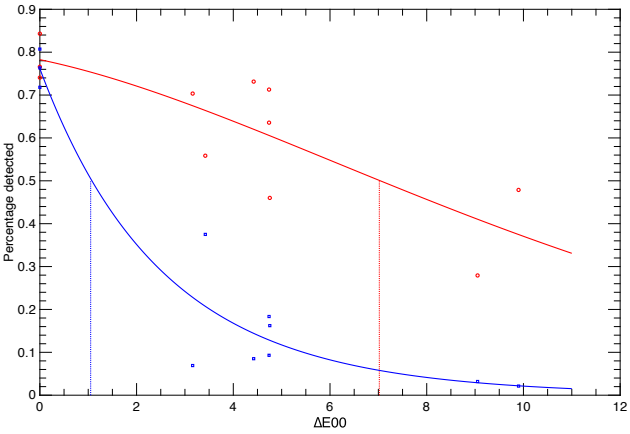

Figure 5. Perceptibility thresholds for same– and different–thickness color line pairs.

One of the most striking results here is that fitting the psychometric function to the same-thickness stimuli under these uncontrolled and unknown conditions, but for a significant number of observations, yields a JND threshold of 1.05 ΔE00. Bearing in mind that ΔE equations are derived from square color patches viewed against a gray background under carefully-controlled conditions and for about one ΔE to correspond to a JND (Hunt and Pointer, 2011). This is a startling level of agreement and a strong anchor to established color difference data.

Turning to the different-thickness case results in a JND of 7.02 in ΔE00 terms. In other words, where an only-just noticeable difference is seen between a thin and a thick line, a difference of 7 ΔE00s would be seen if those same colors were applied to two thick lines. That is, it would be possible to place six colors between the 7 ΔE00 pair, such that each color would be distinguishable from its neighbors and be a similar color difference from both.

The answer to the question of whether sensitivity to color differences between a thin and a thick line is lower than for two thick lines is therefore an emphatic yes.

What this data does not tell us is what that difference in sensitivities depends on, how it relates to specific thicknesses, brightnesses, colors, contrasts, color acuity, normal versus impaired color vision, etc. It does give a strong indication though that the effect is large and robust (showing up under these varied and uncontrolled conditions) and a good candidate for further study.

Implications for imaging pipelines

One of the consequences of this perceptibility threshold study is that imaging pipelines have opportunities for treating fine lines (or similar, small–detail content) differently to, say, photographic content, signage or spot colors, where perceptibility thresholds are much lower. In this section, alternative ways to use these insights will be explored, focusing especially on choices that can be taken in color pipeline resources, such as color look-up-tables (LUTs) and their composition as well as halftoning resources and parameters.

The examples used in the following sections will be based on content that is, to begin with, challenging to reproduce, specifically single-pixel light gray lines that, while discouraged and rarely used, can be present in some CAD applications such as architectural drawings. Such lines will, necessarily, be reproduced in a way where printed colorants form discontinuous structures, even though the input is a continuous, light line. How to best represent the

continuous input under such constraints is a key challenge here. Below is an example of such content and what it would look like printed on plain paper with black ink only.

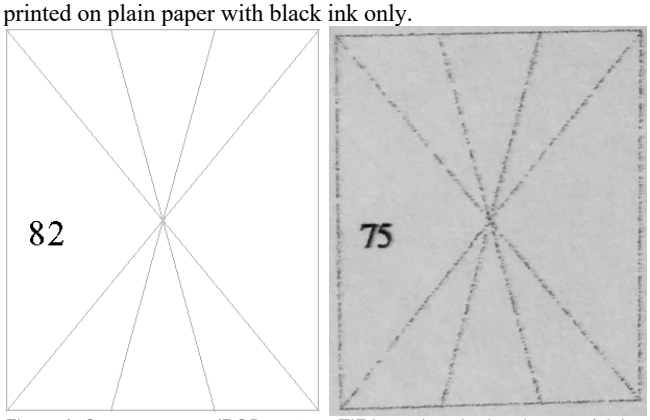

Figure 6: Source content (RGB contone TIF image) and printed output (plain paper, black ink, baseline PARAWACS halftoning using a blue-noise pattern).

As can be appreciated from Fig. 6, single pixel light gray lines are difficult to reproduce well (while additional inks such as gray can help here, even their use has limits – e.g., when reproducing a line lighter than such a gray ink) due to the sparsity of the drops and due to the irregularity of their placement. The lighter the line, the lower the number of drops to be fired and the harder it is to maintain a good distinguishability of the lines and a differentiation between solid, light lines and darker, dashed lines. One possible improvement here is to use composite ink combinations instead of black ink only, e.g., using higher amounts of cyan, magenta and yellow instead of a lower amount of black, but that itself may not be enough. The following section looks at some further options for improvement.

Quantization of Color Resources

A natural consequence of higher perceptibility thresholds is to translate these into coarser color resources by means of quantization. First a brief description of resources follows.

Without loss of generality, the focus here will be on the Halftone Area Neugebauer Separation (HANS) pipeline paradigm (Morovic, 2011) where imaging instructions are represented in terms of Neugebauer Primary (NP) area coverages (or NPacs). These are probability distributions of NPs at any one pixel, so that, e.g., an NPac of [CC:0.3, CM:0.5, w:0.2] would have a 30% chance of a 2-drop Cyan NP, 50% chance of a Cyan and Magenta drop overprinting and a 20% chance of being left blank. A color pipeline using the HANS paradigm would then define its continuous-tone, device color nodes as NPacs, resulting in a N^3 or M^4 LUT for RGB or CMYK respectively, that is used with interpolation in the process of printing to arrive at the per-pixel NPacs that then define the per-pixel probability of the possible NPs.

The first opportunity for quantization presents itself in terms of the resolution of each per-pixel NPac: the bit-depth of the area coverages (or probabilities) and the number of possible NPs for a given NPac. In a typical graphics or photo context the bit-depth of area coverages is at least 8 bits, but can be as high as 16 bits, while the number of NPs per NPac can also be high. Such NPac characteristics then result in smooth transitions and fine granularity (Morovic, 2018). Instead, for fine line content, the opposite choices can be made: lower bit-depth and lower numbers of NPs per NPac. Furthermore, as mentioned above, choices can also be made at the ink level, such as whether to use black ink for grayscale content or to combine it (especially at lower densities) with the use of composite gray (made up of cyan, magenta and yellow instead of black ink). Using composites means a higher amount of ink is used, which in turn may result in more drop positions being occupied. This has advantages and disadvantages: more drops means potentially better line definition, but varying the inks means that the visual impact may vary in lightness, hue and chroma.

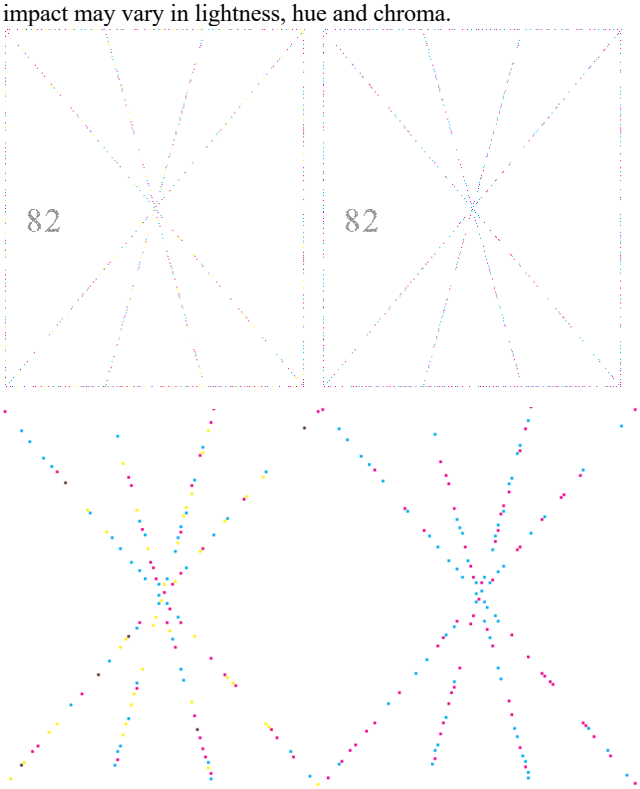

Figure 7: A comparison of two halftones, using 8 bit area coverage and unconstrained number of NPs per NPac (left top and bottom) versus a 6 bit area coverage with at-most 3 NPs per NPac (right top and bottom). Both halftones use the same generic blue-noise halftone pattern. The bottom of the figure corresponds to zoomed-in regions of the top of the figure.

For a light gray color with a continuous tone device color value of 173 in 8 bits (a light grey density of around 30% corresponding to the image in Figure 6 left) on a CMYK-ink system, an NPac of [w:0.75, C:0.08, M:0.08, Y:0.075, CMY:0.015] could be used. Here, the smallest area coverage is 1% of a CMY overprint and the NPac allows a choice of 5 NPs to be used (with their respective probabilities). Quantizing the above NPac to 6 bit precision (whereby the smallest area coverage permitted is $1/2^6$=0.0156 or 1.56%) and constraining the maximum number of NPs to be used to 3, results in an NPac of [w:0.75, C:0.125, M:0.125]. Fig. 7 shows the impact on the halftone pattern that shows a clear improvement in the quantized case (right hand side), even though a slightly lower amount of ink is used. The reason for this is that in this case, the CMY NP is removed (since its area coverage is below the threshold) and Y is also removed (since it's the smallest coverage that exceeds the 3 NP constraint), resulting in only C and M NPs being used (other than the blank NP) which also have closer lightnesses. The combination of these factors results in a better, visually more regular pattern with apparently more pixels (but at a lower total ink use). In

this example, colorimetry was not explicitly taken into account but doing so could further improve choices (such as explicitly favoring darker NPs or NPs of similar lightnesses for example).

Quantization of Custom Halftone Patterns

Another area where quantization can be applied is halftoning. In the context of HANS, the role of halftoning is purely spatial. Unlike in colorant-channel based pipelines, where halftoning occurs per-channel and the total halftone for all colorants is the combined result of halftones for each channel individually, for HANS pipelines all colorants are halftoned at once since this step amounts to selecting an NP at a pixel from the corresponding NPac at that pixel. Broadly, two approaches are possible here: some form of error diffusion (selecting an NP at a pixel and diffusing the error of not being able to fully represent an NPac to the neighboring pixels) (Morovic, 2011) or PARAWACS (Parallel Random Area Weighted Coverage Selection) which is a selector-based approach where the selector value is used to pick an NP out of the NPac (Morovic, 2017). Both error diffusion (ED) and PARAWACS lend themselves to quantization, however PARAWACS has been shown to have additional advantages such as smooth transitions and overall high IQ (Morovic, 2018).

PARAWACS can use a pre-generated selector value matrix, which may be generated to have particular Fourier–spectral properties such as blue-noise or green-noise. In general, selector value matrices should satisfy the condition of equal likelihood of all its values (e.g., a matrix of NxN size and a resolution of M values, should have N*N/M pixels for each of the [1 to M] values). In the case of lines, especially single-pixel lines, a selector value matrix should in addition be built in a way that does not penalize any line angle. Here, a checkerboard (in itself a blue-noise pattern) is often used as the starting point. Fig. 8 shows an example of such a checkerboard pattern. An analysis of this matrix shows that each odd row does not contain values below 15% while each even row does not contain values above 85%. This means that depending on the spatial position of certain densities of lines, these may even be completely lost in the halftoning stage (e.g., if a <15% density line falls on an odd row).

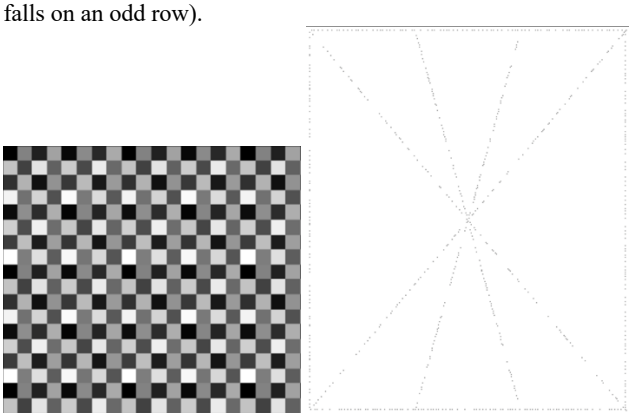

Figure 8: An 8 bit checkerboard pattern used as a selector-value matrix (left) and a test pattern for low-density single pixel lines (right) halftoned with the selector value matrix.

One way to improve over the naïve checkerboard pattern is to introduce constraints on the nature of the checkerboard in order to have all densities in all rows/columns of the matrix well represented. This can be achieved by dividing the full range of [1 to M] values in K sub-ranges and then picking random values from each sub-range in sequence. For example, for K=2, alternate values from sub-range along each row are picked and alternation is changed from row to row (i.e., always starting with a different sub-range). Fig. 9 shows an example of such a matrix at the same bit-depth as the previous example.

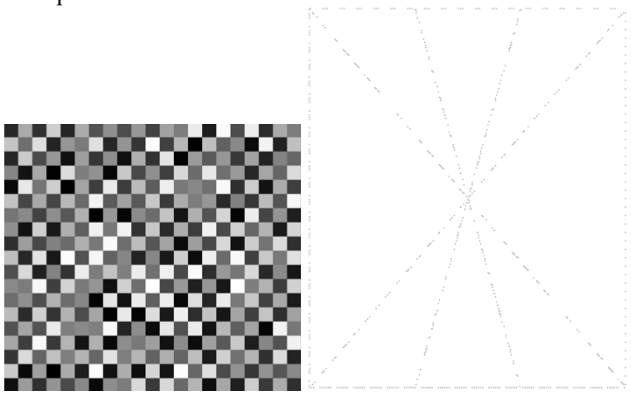

Figure 9: An 8bit sub-range alternating checkerboard pattern used as a selector-value matrix (left) and a test pattern for low-density single pixel lines (right) halftoned with the selector value matrix.

The halftone in Fig. 9 shows some signs of improvement as it results in more regularity, especially in the horizontal and vertical lines, but where it underperforms are the diagonals and also in terms of how different the regularity is in different locations (e.g., the left and right vertical lines). Here quantization can help. Fig. 10 shows two examples where selector matrices were generated similarly to Fig. 9, but at 10 bits and 6 bits respectively.

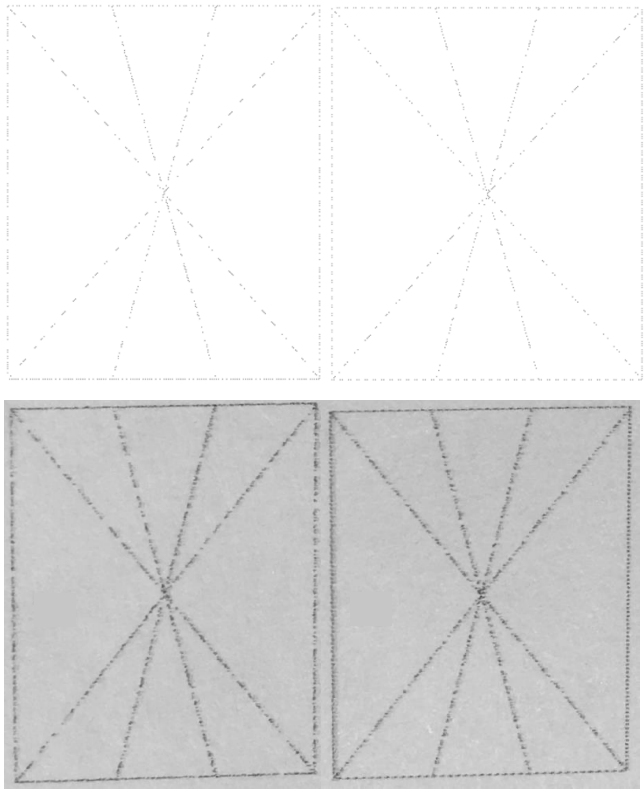

Figure 10: A test pattern for low-density single pixel lines halftoned with the alternating sub-range selector value matrix approach in 10 bits (top left) and 6 bits (top right), and the printed output showing the 6 bit pattern (bottom right) against the baseline blue noise pattern from earlier (bottom left).

As can be seen, the lower bit-depth version in Fig. 10 (right) shows an advantage over the higher bit-depth version (and over the default 8 bit version).

Finally, another improvement to the matrix construction method can be achieved by further introducing regularity and alternation in the randomness inherent in these matrices, for example by reversing the alternation and introducing a zig-zag strategy as shown in Fig. 11.

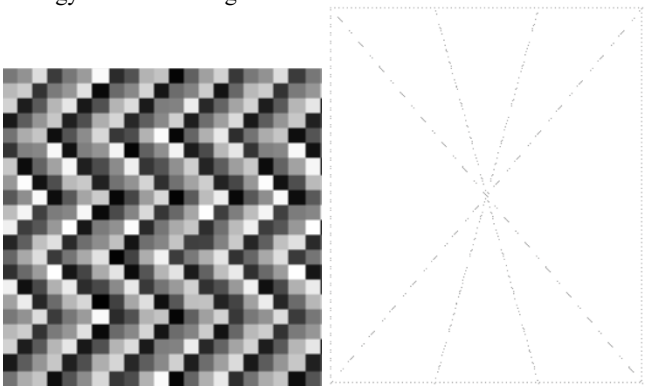

Figure 11: An 8 bit 4-sub-range alternating zig-zag checkerboard pattern used as a selector-value matrix (left) and a test pattern for low-density single pixel lines (right) halftoned with the selector value matrix.

Conclusions

Fine details such as single pixel light gray lines may pose a challenge in color reproduction pipelines, however the thresholds of sensitivity to differences and even departures from a target colorimetry are significantly higher to those of content occupying larger areas of the visual field.

An extensive online study of a limited set of samples showed how differences as high as 7 ΔE2000 are acceptable for such content, while also confirming the 1 ΔE2000 JND threshold for larger areas of the visual field. This insight led to the exploration of choices in color pipeline parameters that relax the tight requirements for matching color or colorant amounts to benefit pattern regularity and printed content distinguishability.

Two broad areas of such choices have been described here. First, quantization in the domain of color resources: the resolution to which differences in colorant amounts and Neugebauer Primaries are represented and second, quantization in the domain of the halftone pattern generation process. Furthermore, modifications to traditional baseline blue-noise patterns have been explored that also aid better single-pixel light line reproduction.

These choices – taking both the specific needs as well as the looser thresholds into account – have been shown to enable better performance in printed content. Further investigations are needed to fully benefit from the new thresholds, such as taking the JND requirements for lines explicitly into account in resource generation as well as colorimetric choices (such as lightness preservation or lightness contrast minimization, as shown in the composite gray example).

Acknowledgements

The authors would like to acknowledge the help of Alex Campa and Martí Rius, and the support of Elizabeth Zapata and Andreu Gonzalez.

References

P. G. Engeldrum, *Psychometric Scaling: A Toolkit for Imaging System Development*, Imcotek Press (2000).

G. Gescheider, *Psychophysics: Method, Theory, and Application*, Lawrence Earlbaum Associates, Inc., Mahwah, NJ (1985).

G. High, P. Nussbaum, P. Green, Estimating visual difference between reproduction gamuts: moving our pilot study from the lab to online delivery, *29th IS&T Color and Imaging Conference* (2021), 317-322.

R. W. G. Hunt, M R. Pointer, *Measuring Colour*, 4th Ed., John Wiley & Sons, Ltd, Chichester, UK (2011) 57.

F. Mirjalili, M. R. Luo, G. Cui, J. Morovic, Color-difference formula for evaluating color pairs with no separation: ΔE_{NS}, *Journal of the Optical Society of America A*, 36 (2019), 789-799.

N. Moroney, Unconstrained web-based color naming experiment. *Color Imaging VIII: Proc. Hardcopy and Applications. Edited by Eschbach, Reiner; Marcu, Gabriel G. Proc. of the SPIE*, 5008 (2003) 36-46.

J. Morovič, P. Morovič, J. Arnabat HANS – Controlling Inkjet Print Attributes Via Neugebauer Primary Area Coverages, *IEEE Transactions on Image Processing*, 21, 2, (2011)

P. Morovič, J. Morovič, J. Gondek, R. Ulichney, Direct Pattern Control Halftoning of Neugebauer Primaries, *IEEE Transactions on Image Processing*, 26(9):4404-4413, (2017)

P. Morovič, J. Morovič, V. Diego, J. Maestro, H. Gomez, S. Etchebehere, X. Fariña, P. Gasparin, Halftone structure optimization using convex programming, *Proceedings of IS&T 26th Color and Imaging Conference*, Vancouver Canada, (2018)

D. Mylonas, L. W. MacDonald, Online Colour Naming Experiment Using Munsell Samples, *5th European Conference on Colour in Graphics, Imaging, and Vision–CGIV,* Joensuu, Finland (2010) 27–32.

A. B. Watson, Probability summation over time, *Vision Research* 19 (1979), 515–522.

K. Xiao, M. R. Luo, C. Li, G. Hong, Colour appearance of room colours, *Color Research and Application*, 35 (2010), 284-293.