

Linear Color-Separable Human Visual System Models for Vector Error Diffusion Halftoning

Vishal Monga, *Associate Member, IEEE*, Wilson S. Geisler, III, and Brian L. Evans, *Senior Member, IEEE*

Abstract—Image halftoning converts a high-resolution image to a low-resolution image, e.g., a 24-bit color image to a three-bit color image, for printing and display. Vector error diffusion captures correlation among color planes by using an error filter with matrix-valued coefficients. In optimizing vector error filters, Damera-Venkata and Evans transform the error image into an opponent color space where Euclidean distance has perceptual meaning. This letter evaluates color spaces for vector error filter optimization. In order of increasing quality, the color spaces are YIQ, YUV, opponent (by Poirson and Wandell), and linearized CIELab (by Flohr, Kolpatzik, Balasubramanian, Carrara, Bouman, and Allebach).

Index Terms—Color image display, color quantization, image quality, multilters.

I. INTRODUCTION

IMAGE halftoning converts a high-resolution image to a lower resolution image, e.g., for printing and display. Common examples are converting an eight-bit-per-pixel grayscale image to a binary image and a 24-bit color image (with eight bits per pixel per color) to a three-bit color image. Applying grayscale halftoning methods separably to color images does not take the correlation among color planes into account and can lead to artifacts such as spurious color impulses in the resulting halftones.

In grayscale halftoning by error diffusion, each grayscale pixel is thresholded to white or black, and the quantization error is fed back, filtered, and added to the neighboring grayscale pixels [1]. Although an error filter is typically lowpass, the feedback arrangement causes the quantization error to be highpass filtered, i.e., pushed into high frequencies where the human eye is least sensitive. The feedback arrangement sharpens the original image by passing low frequencies and amplifying high frequencies. Traditional grayscale error diffused halftones appear sharper than the original and contain highpass noise [2].

Vector error diffusion, first proposed in [3], represents each pixel in a color image as a vector of values. The thresholding step would threshold each vector component separately. The vector-valued quantization error (image) would be fed back, filtered, and added to the neighboring (unhalftoned) color pixels. A matrix-valued error filter could take the correlation among color planes into account. For an RGB image, each error filter coefficient would be a 3×3 matrix. For RGB vector error diffusion, matrix-valued error filter coefficients are adapted in [4]

Manuscript received June 3, 2002; revised August 14, 2002. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Scott T. Acton.

The authors are with the Center for Perceptual Systems, The University of Texas, Austin, TX 78712 USA (e-mail: vishal@ece.utexas.edu; geisler@mail.utexas.edu; bevans@ece.utexas.edu).

Digital Object Identifier 10.1109/LSP.2002.806708

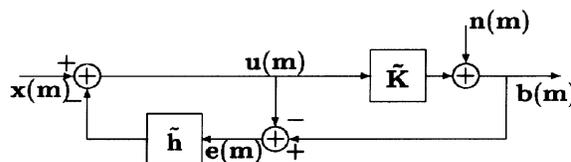


Fig. 1. Linearized model of vector color error diffusion. The quantizer has been replaced by a linear transformation by $\tilde{\mathbf{K}}$ plus additive noise $\mathbf{n}(\mathbf{m})$ that is uncorrelated with $\mathbf{u}(\mathbf{m})$. The original image is $\mathbf{x}(\mathbf{m})$, and the halftone is $\mathbf{b}(\mathbf{m})$.

to reduce the mse between the halftone and original. However, the mse does not have perceptual meaning in RGB space.

A linear color human visual system (HVS) framework [5] designs matrix-valued error filters by transforming the error image into luminance and chrominance components and minimizing a linear model of the human visual response to the quantization noise. In this framework, the mse has perceptual meaning. Based on this framework, this letter evaluates four color spaces in which to optimize matrix-valued error filters. We find that the objective and subjective rankings of color spaces agree. The color spaces in order of increasing quality are 1) YIQ space, 2) YUV space, 3) opponent color space [6], [7], and 4) linearized CIELab color space [8]. The first two color spaces are common in video, where Y refers to the luminance component, and the other two letters refer to the two chrominance components. The opponent color space is a part of the pattern-color-separable vision model [6], [7] that forms the basis for the industry standard Spatial-CIELab (S-CIELab) color space [7]. We did not consider the S-CIELab color space because we are restricting our attention to color spaces based on a linear transformation from RGB. However, the HVS model with the linearized CIELab color space under consideration [8] closely approximates the S-CIELab framework.

Section II linearizes vector error diffusion and defines the vector error filter optimization problem. Section III generalizes the linear color HVS model [5] to compute the vector error filter coefficients. Section IV describes color space transformations and the associated spatial filters. Sections V and VI report objective and subjective quality measures, respectively. Section VII concludes the letter.

II. VECTOR COLOR ERROR DIFFUSION

We use \mathbf{m} to denote a two-dimensional (2-D) spatial index (m_1, m_2) , and \mathbf{z} to denote the z domain variables (z_1, z_2) .

A. Matrix Gain Model

Fig. 1 shows a model of vector color error diffusion halftoning after linearizing the system by replacing the quantizer with ma-

trix gain $\tilde{\mathbf{K}}$ and an additive white-noise image $\mathbf{n}(\mathbf{m})$ [5]. The matrix gain is related to the amount of sharpening, and the noise image models the quantization error. $\tilde{\mathbf{K}}$ is chosen to minimize the error in approximating the quantizer with a linear transformation, in the linear minimum mse sense

$$\tilde{\mathbf{K}} = \arg \min_{\tilde{\mathbf{A}}} E \left[\|\mathbf{b}(\mathbf{m}) - \tilde{\mathbf{A}} \mathbf{u}(\mathbf{m})\|^2 \right] \quad (1)$$

where $\mathbf{b}(\mathbf{m})$ is the quantizer output process (halftone), and $\mathbf{u}(\mathbf{m})$ is the quantizer input process. When $\mathbf{b}(\mathbf{m})$ and $\mathbf{u}(\mathbf{m})$ are wide-sense stationary [9], the solution for (1) is

$$\tilde{\mathbf{K}} = \tilde{\mathbf{C}}_{\mathbf{bu}} \tilde{\mathbf{C}}_{\mathbf{uu}}^{-1} \quad (2)$$

where $\tilde{\mathbf{C}}_{\mathbf{bu}}$ and $\tilde{\mathbf{C}}_{\mathbf{uu}}$ are covariance matrices. The linearized vector error diffusion system (Fig. 1) has two inputs (original signal and quantization noise) and one output (the halftone). Using (2), the signal and noise transfer functions are [5]

$$\mathbf{B}_s(z) = \tilde{\mathbf{K}} \left[\tilde{\mathbf{I}} + \tilde{\mathbf{H}}(z)(\tilde{\mathbf{K}} - \tilde{\mathbf{I}}) \right]^{-1} \mathbf{X}(z) \quad (3)$$

$$\mathbf{B}_n(z) = \left[\tilde{\mathbf{I}} - \tilde{\mathbf{H}}(z) \right] \mathbf{N}(z). \quad (4)$$

The overall system response is given by

$$\mathbf{B}(z) = \mathbf{B}_s(z) + \mathbf{B}_n(z). \quad (5)$$

B. Error Measure

We form an objective function J that measures the average visually weighted noise energy in the halftone. The output noise is computed by inverse transforming (4)

$$\mathbf{b}_n(\mathbf{m}) = \left[\tilde{\mathbf{I}} - \tilde{\mathbf{h}}(\mathbf{m}) \right] \star \mathbf{n}(\mathbf{m}). \quad (6)$$

We weight the noise energy by a linear spatially invariant matrix-valued HVS model $\tilde{\mathbf{v}}(\mathbf{m})$ and form

$$J = E \left[\left\| \tilde{\mathbf{v}}(\mathbf{m}) \star \left[\tilde{\mathbf{I}} - \tilde{\mathbf{h}}(\mathbf{m}) \right] \star \tilde{\mathbf{n}}(\mathbf{m}) \right\|^2 \right]. \quad (7)$$

Given a linear spatially invariant HVS model $\tilde{\mathbf{v}}(\mathbf{m})$, the problem is to design an optimal matrix-valued error filter

$$\tilde{\mathbf{h}}_{\text{opt}}(\mathbf{m}) = \arg \min_{\tilde{\mathbf{h}}(\mathbf{m}) \in \mathcal{C}} J \quad (8)$$

where the constraint \mathcal{C} enforces the criterion that the error filter diffuses all quantization error [10]

$$\mathcal{C} = \left\{ \tilde{\mathbf{h}}(\mathbf{i}), \mathbf{i} \in \mathcal{S} \mid \sum_{\mathbf{i}} \tilde{\mathbf{h}}(\mathbf{i}) \mathbf{1} = \mathbf{1} \right\}. \quad (9)$$

\mathcal{S} is the set of coordinates for the error filter support, i.e., $\mathcal{S} = \{(1, 0), (1, 1), (0, 1), (-1, 1)\}$ for Floyd–Steinberg.

For each of the four HVS models in Section IV, we obtain the optimum filter coefficients by solving a matrix version of the Yule–Walker equations, based on the constraint \mathcal{C} [5]. We use a descent algorithm based on the quasi-Newton Broyden–Fletcher–Goldfarb–Shanno update [11]. This reduces the number of iterations required to converge to the minimum by about half for the same effort per iteration versus steepest descent [5].

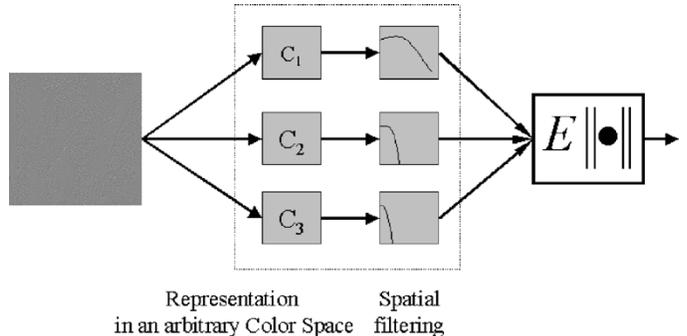


Fig. 2. Generalized linear color model for the human visual system.

III. GENERALIZED PERCEPTUAL MODEL

The linear color model employed in [5] is based on the pattern–color-separable model in [6] and [7]. Damera-Venkata and Evans [5] transfer device-dependent RGB values into an opponent representation [7], [12]. The three opponent visual pathways are white–black (luminance pathway) and red–green and blue–yellow (chrominance pathways). By x - y , we mean that in value x is at one extreme, and y is at the other.

We generalize this model as a linear transformation $\tilde{\mathbf{T}}$ to a desired color space, which is not necessarily the opponent representation [6] but any one that satisfies pattern–color separability, followed by appropriate spatial filtering in each channel. A complete HVS model is uniquely determined by the color space transformation and associated spatial filters. This generalization provides a platform for evaluation of different models in perceptual meaning and error filter quality obtained by minimizing (7).

The linear color model Fig. 2 consists of 1) a linear transformation $\tilde{\mathbf{T}}$ and 2) separable spatial filtering on each channel. Each channel uses a different spatial filter. The filtering in the z domain is a matrix multiplication by a diagonal matrix $\mathbf{D}(z)$. In the spatial domain, the linear HVS model $\tilde{\mathbf{v}}(\mathbf{m})$ is computed as

$$\tilde{\mathbf{v}}(\mathbf{m}) = \tilde{\mathbf{d}}(\mathbf{m}) \tilde{\mathbf{T}}. \quad (10)$$

The following section discusses the transformation $\tilde{\mathbf{T}}$ and the spatial filters $\tilde{\mathbf{d}}(\mathbf{m})$.

IV. COLOR TRANSFORMATIONS AND SPATIAL FILTERS

Since we are targeting color halftones for display, we first account for the nonlinear response of a CRT to frame buffer values. We pass the RGB values of original image through this nonlinearity to obtain the RGB values of the colors displayed on the monitor before halftoning. This ensures that the colors in the halftone are closest to the colors actually rendered on the monitor. This process is the inverse of gamma correction. We use “gamma uncorrection” as specified by the sRGB standard [13] on the original image prior to halftoning.

We employ transformations to the following color spaces: linearized CIELab [8], opponent [7], YUV, and YIQ. These color spaces are chosen because they all score well in perceptual uniformity [14] and approximately satisfy the requirements for pattern–color separability [6]. Since RGB values are device dependent, we perform the color transformations based on an sRGB monitor. The transformation to opponent color space is

$$\text{sRGB} \rightarrow \text{CIEXYZ} \rightarrow \text{opponent representation.}$$

The standard transformations from sRGB to CIEXYZ and from CIEXYZ to opponent representation are taken from the S-CIELab [7] code,¹ which is also the source for transformations to the YUV and YIQ representations. The linearized CIELab color space is obtained by linearizing the CIELab space about the D65 white point [8] in the following manner:

$$Y_y = 116 \frac{Y}{Y_n} - 16 \quad (11)$$

$$C_x = 500 \left[\frac{X}{X_n} - \frac{Y}{Y_n} \right] \quad (12)$$

$$C_z = 200 \left[\frac{Y}{Y_n} - \frac{Z}{Z_n} \right]. \quad (13)$$

Hence $\tilde{\mathbf{T}}$ is sRGB \rightarrow CIEXYZ \rightarrow Linearized CIELab. The \mathbf{Y}_y component is similar to the luminance, and the \mathbf{C}_x and \mathbf{C}_z components are similar to the R-G and B-Y opponent color components. The original transformation to the CIELab from CIEXYZ is a nonlinear one

$$L^* = 116 f \left(\frac{Y}{Y_n} \right) - 16 \quad (14)$$

$$a^* = 500 \left[f \left(\frac{X}{X_n} \right) - f \left(\frac{Y}{Y_n} \right) \right] \quad (15)$$

$$b^* = 200 \left[f \left(\frac{Y}{Y_n} \right) - f \left(\frac{Z}{Z_n} \right) \right] \quad (16)$$

where

$$f(x) = \begin{cases} 7.787x + \frac{16}{116} & \text{if } 0 \leq x \leq 0.008856 \\ x^{1/3} & \text{if } 0.008856 < x \leq 1. \end{cases}$$

The values for X_n , Y_n , and Z_n are as per the D65 white point [13].

The nonlinearity in the transformation from CIELab distorts the spatially averaged tone of the images, which yields halftones that have incorrect average values [8]. The linearized color space overcomes this and has the added benefit that it decouples the effect of incremental changes in (Y_y, C_x, C_z) at the white point on (L, a, b) values

$$\nabla_{(Y_y, C_x, C_z)}(L^*, a^*, b^*)|_{D65} = \frac{1}{3} \mathbf{I}. \quad (17)$$

In the opponent color representation, data in each plane are filtered [7] by 2-D separable spatial kernels

$$f = k \sum_i \omega_i E_i \quad (18)$$

where $E_i = k_i \exp(-(x^2 + y^2)/\sigma_i^2)$. The parameters ω_i and σ_i are based on psychophysical testing and are available in [7]. The spatial filters for Linearized CIELab and the YUV and YIQ color spaces are based on analytical models of the eye's luminance and chrominance frequency response.

Nasanen and Sullivan [15] chose an exponential function to model the luminance frequency response

$$W_{(Y_y)}(\tilde{\rho}) = K(L) e^{-\alpha(L)\tilde{\rho}} \quad (19)$$

¹<http://white.stanford.edu/~brian/scielab/scielab1-1>

TABLE I
NOISE GAIN (IN DECIBELS) OF THE FLOYD-STEINBERG ERROR FILTER
OVER THE OPTIMUM ERROR FILTER FOR FOUR HVS MODELS

<i>Image</i>	<i>Linear CIELab</i>	<i>Opponent</i>	<i>YUV</i>	<i>YIQ</i>
lena	3.3511	3.2912	3.2731	3.2778
pasta	1.4623	1.4566	1.4472	1.4383
fruits	1.3832	1.3730	1.3682	1.3644
peppers	1.9562	1.8766	1.8633	1.8545

where L is the average luminance of display; $\tilde{\rho}$ is the radial spatial frequency; $K(L) = aL^b$; and

$$\alpha(L) = \frac{1}{c \ln(L) + d}. \quad (20)$$

The frequency variable $\tilde{\rho}$ is defined [8] as a weighted magnitude of the frequency vector $\mathbf{u} = (u, v)^T$, where the weighting depends on the angular spatial frequency ϕ [15]. Thus

$$\tilde{\rho} = \frac{\rho}{s(\phi)} \quad (21)$$

where $\rho = \sqrt{u^2 + v^2}$ and

$$s(\phi) = \frac{1 - \omega}{2} \cos(4\phi) + \frac{1 + \omega}{2}. \quad (22)$$

The symmetry parameter ω is 0.7, and $\phi = \arctan(v/u)$. The weighting function $s(\phi)$ effectively reduces the contrast sensitivity to spatial frequency components at odd multiples of 45° . The contrast sensitivity of the human viewer to spatial variations in chrominance falls off faster as a function of increasing spatial frequency than does the response to spatial variations in luminance [16]. The chrominance model we use reflects this [17]

$$W_{(C_x, C_z)}(\rho) = A e^{-\alpha\rho}. \quad (23)$$

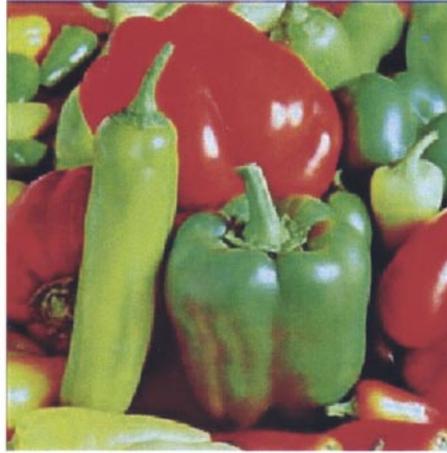
Here, α is determined to be 0.419, and $A = 100$ [17]. Both the luminance and chrominance response are lowpass in nature, but only the luminance response is reduced at odd multiples of 45° . This will place more luminance error across the diagonals in the frequency domain where the eye is less sensitive. Using this chrominance response as opposed to identical responses for both luminance and chrominance will allow more low-frequency chromatic error, which will not be perceived by the human viewer.

V. SIMULATION RESULTS

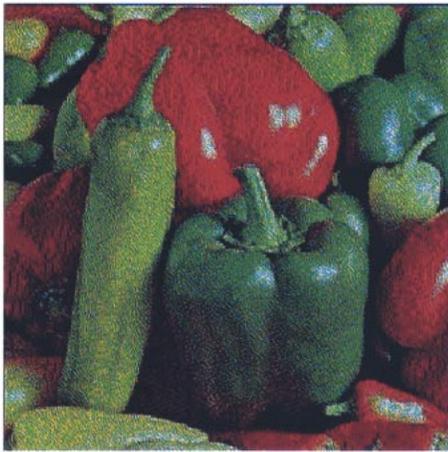
We evaluate four models based on linearized CIELab, opponent, YUV, and YIQ color spaces using four images. Table I gives the noise shaping gain NG of the optimal filter over the Floyd-Steinberg filter in decibels [5]

$$NG = 10 \log_{10} \left(\frac{J_{fs}}{J_{opt}} \right). \quad (24)$$

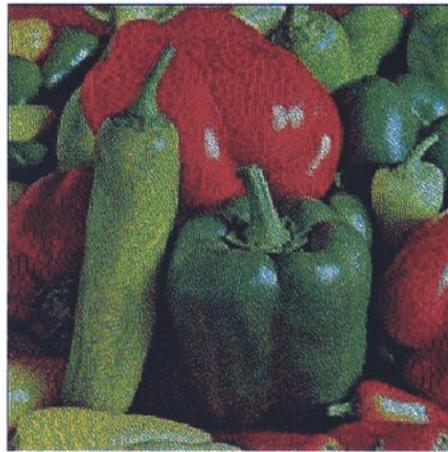
Here, J refers to the value of the objective function given by (7). The linearized CIELab model outperforms the other three models. The other three models have similar performance. Matrix-valued error filters for the four HVS models are available at <http://www.ece.utexas.edu/~vishal/halftoning.html> (Fig. 3).



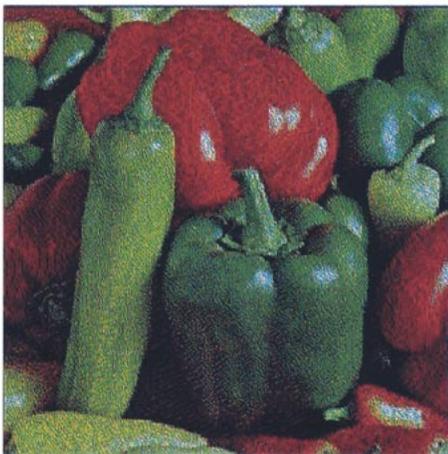
(a) Original peppers image



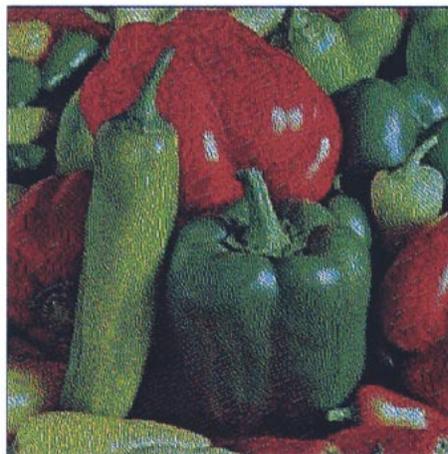
(b) using linearized CIELab



(c) using opponent color space



(d) using YUV



(e) using YIQ

Fig. 3. Results of halftones of peppers generated by optimizing the vector error filter in the four color spaces.

VI. SUBJECTIVE TESTING

We develop a subjective assessment procedure for comparing the models based on a pair-comparison task. On each trial, the

observer saw an original image above a pair of halftone images positioned to the left and right. The images were displayed on an sRGB monitor. The observer was forced to pick the halftone that looked most similar to the original. For each original image,

observers made six preference judgments, one for each possible pairing of the four halftone images generated using the four different HVS models.

The relevant data from the assessment procedure is the proportion of trials where one given halftoning method is preferred to the other. If one method is not preferred over the other, then the measured proportion should (ideally) be 0.5. Further, to conclude that one method is significantly better, we require it to be picked over the other 60% or more of the time. Using the binomial probability distribution, the total number of trials in the study was picked so that the binomial proportion could be estimated to within a tolerance of 0.03 with 95% confidence. This requirement resulted in a total of 960 comparisons [18], [19] for each of the six possible pairings of the four halftoning methods. The total comparisons/trials were obtained from 60 subjects who judged 96 comparisons each.

Let A , B , C , and D be the four halftoning methods (or HVS models) to be compared. Based on the subjective test data, we build

$$\mathbf{P} = \begin{bmatrix} 0.5 & p_{AB} & p_{AC} & p_{AD} \\ p_{BA} & 0.5 & p_{BC} & p_{BD} \\ p_{CA} & p_{CB} & 0.5 & p_{CD} \\ p_{DA} & p_{DB} & p_{DC} & 0.5 \end{bmatrix}$$

where p_{AB} is the proportion that method A was preferred to method B . The entries in \mathbf{P} satisfy $P(k, l) = 1 - P(l, k)$. The data from the subjective assessment procedure give \mathbf{P} as

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.625 & 0.833 & 0.854 \\ 0.375 & 0.5 & 0.563 & 0.621 \\ 0.167 & 0.437 & 0.5 & 0.521 \\ 0.146 & 0.379 & 0.479 & 0.5 \end{bmatrix}.$$

This results in a unique rank order of the methods i.e., $A \geq B \geq C \geq D$, where A refers to the HVS model based on Linearized CIELab; B refers to the model based on the opponent color space; C refers to the YUV model; and D refers to YIQ. Note that if a rank order of the form $A \geq B \geq C \geq D$ exists and we have $p_{AB} \geq 0.6$, then we would automatically have $p_{AC} \geq 0.6$ and so forth, and we can conclude that A is significantly better than or preferred to B , C , and D . Similarly, we can find if B is better than C and so on. From the entries in \mathbf{P} , we hence conclude that the HVS model based on the linearized CIELab [8] color space is the clear winner. In descending subjective quality

Linearized CIELab \gg Opponent $>$ YUV \approx YIQ.

The objective rankings in Table I and subjective rankings given by \mathbf{P} agree. In Table I, YUV is slightly better than YIQ in three of the four images. Likewise, the “perceptual” distinction in the subjective test between YUV and YIQ is not appreciable. The subjective test is available online at <http://www.ece.utexas.edu/~vishal/cgi-bin/test.html>.

VII. CONCLUSION

For grayscale images, contrast sensitivity functions are linear models of the HVS response versus spatial frequency. These linear HVS models have been explicitly used in grayscale halftoning for more than a decade to improve halftone quality [20]. For color images, similar linear HVS models exist for luminance channels and certain chrominance channels, which can be applied after a color space transformation. In this letter, we evaluate four color HVS models for color halftoning by vector error diffusion using objective measures and subjective testing. The color HVS model based on transformation to the linearized CIELab [8] color space and spatial filters for the luminance frequency response in [15] and the chrominance frequency response in [17] yields the best error filter for vector error diffusion.

REFERENCES

- [1] R. W. Floyd and L. Steinberg, “An adaptive algorithm for spatial grayscale,” in *Proc. SID*, vol. 17, 1976, pp. 75–77.
- [2] R. Ulichney, *Digital Halftoning*. Cambridge, MA: MIT Press, 1987.
- [3] H. Haneishi, H. Yaguchi, and Y. Miyake, “A new method of color reproduction in digital halftone image,” in *Proc. IS&T Conf.*, Cambridge, MA, May 1993.
- [4] L. Akarun, Y. Yardimci, and A. E. Cetin, “Adaptive methods for dithering color images,” *IEEE Trans. Image Processing*, vol. 6, pp. 950–955, July 1997.
- [5] N. Damera-Venkata and B. L. Evans, “Design and analysis of vector color error diffusion halftoning systems,” *IEEE Trans. Image Processing*, vol. 10, pp. 1552–1565, Oct. 2001.
- [6] A. B. Poirson and B. A. Wandell, “Appearance of colored patterns: Pattern-color separability,” *J. Opt. Soc. Amer. A*, vol. 10, pp. 2458–2470, Dec. 1993.
- [7] X. Zhang and B. A. Wandell, “A spatial extension of CIELab for digital color image reproduction,” *SID Tech. Dig.*, pp. 731–734, 1996.
- [8] T. J. Flohr, B. W. Kolpatzik, R. Balasubramanian, D. A. Carrara, C. A. Bouman, and J. P. Allebach, “Model based color image quantization,” in *Proc. SPIE Human Vision, Visual Proc. and Digital Display IV*, vol. 1913, 1993, pp. 270–281.
- [9] H. Stark and J. W. Woods, *Probability, Random Processes and Estimation Theory for Engineers*. Englewood Cliffs, NJ: Prentice-Hall, 1986.
- [10] N. Damera-Venkata, “Analysis and design of vector error diffusion systems for image halftoning,” Ph.D. thesis, Dept. ECE, Univ. Texas, Austin, Dec. 2000.
- [11] D. P. Bertsekas, *Nonlinear Programming*: Athena Scientific, 1995.
- [12] M. D. Fairchild, *Color Appearance Models*. Reading, MA: Addison-Wesley, 1998.
- [13] What is sRGB?—Introduction to the standard default RGB color space developed by Hewlett-Packard and Microsoft (1999). <http://www.srgb.com/aboutsrgb.html> [Online]
- [14] Frequently asked questions about color, C. A. Poynton. (1999). <http://www.inforamp.net/~poynton/ColorFAQ.html> [Online]
- [15] J. R. Sullivan, L. A. Ray, and R. Miller, “Design of minimum visual modulation halftone patterns,” *IEEE Trans. Syst. Man. Cybern.*, vol. 21, pp. 33–38, Jan. 1991.
- [16] D. H. Kelly, “Spatiotemporal variation of chromatic and achromatic contrast thresholds,” *J. Opt. Soc. Amer. A*, vol. 73, pp. 742–750, 1983.
- [17] B. Kolpatzik and C. Bouman, “Optimized error diffusion for high quality image display,” *J. Electron. Imaging*, vol. 1, pp. 277–292, Jan. 1992.
- [18] G. E. P. Box, J. S. Hunter, and W. G. Hunter, *Statistics for Experimenters*. New York: Wiley, 1978.
- [19] StatPoint, LLC, “5.4.3 Binomial Proportion – for estimating the parameter of a binomial distribution”, in *STATLETS User Manual*, Englewood Cliffs, NJ, 2002 [Online]: Available http://www.statlets.com/usermanual/sect5_4_3.htm.
- [20] S. H. Kim and J. P. Allebach, “Impact of human visual system models on model based halftoning,” *IEEE Trans. Image Processing*, vol. 11, pp. 258–269, Mar. 2002.