Determination of Significant Color Map

Teck Ping Sim, and Perry Y. Li; Department of Mechanical Engineering, University of Minnesota; 111 Church Street S.E., Minneapolis, MN 55108 USA

Abstract

*This paper propose an approach to extract the most significant color information that takes into account 1) the inherent sensitivity of the Human Visual System (HVS) to color variations and 2) the sensitivity of the HVS to specific colors of a given image. By giving higher importance to the significant colors in the stabilization control formulation, it is possible to retain the consistency of only the most significant colors to a human viewer. Such restriction is critical due to the limited actuation authorities of typical xerographic color printing process. To account for color-specific sensitivity, we make use of the CIE L*a*b* color formulation. To account for the image-specific sensitivity, we make use of the fact that the HVS have lower sensitivity to high frequency components (detail) of high chrominance image. Moreover, typically "smooth" images are visually more irrelevant compared to textured and detailed images. These effects in the color and spatial domain are taken into account to reduce the redundancy in the image representation by following similar procedure to that used in wavelet transform coding decomposition.*

Introduction

The color reproduction characteristics(CRC) function describes how a printer maps the desired colors into actual printed colors. These printed colors vary with time due to printer dynamics as well as environmental and material variations, resulting in inconsistent color prints. In our current and previous works [1, 2, 3, 4], we have looked into control and sensing issues of maintaining tone and color consistency. The key control challenge is the high dimensional print-map (in that a color printer can reproduced high numbers of colors) with limited actuation authorities. This means that it is not possible to maintain the consistency of all the printed colors. In this paper, we developed a technique to stabilize only colors that are significant to a human viewer.

Specifically there are two main redundancies of a printed image. The first is the inherent sensitivity of the Human Visual System (HVS) to color variations. In fact it is well known that human does not differentiate two colors with a $\Delta E^*_{94} < 1$ [5]. The second redundancy can be found on the printed color image itself. For a color image, the redundancies can be found in the color and spatial representation of the image. Human visual system (HVS) has lower sensitivity to high frequency components (details) of high chrominance image. Moreover, typically "smooth" images are visually more irrelevant compared to textured and detailed images. These effects in the color and spatial domain are taken into account to reduce the redundancy in the image representation by following similar procedure to that used in wavelet transform coding [6]. The multi-resolution properties of wavelet decomposition is particularly well suited to the characteristics of the HVS because psycho-visual studies show that the human visual system can be modeled using multiple perceptual channels, which are octave-wise spread over the spatial frequency range. From one octave to the next, the summation area (area of the retina contributing to the perceived signal) also changes by a factor of four. This behavior match that of the dyadic structure of the discrete wavelet transform.

Problem Formulation

In both sensing and control for maintaining color consistency, it is convenient to take the input colorspace as the CMY colorspace and the output colorspace as the CIE $L^*a^*b^*$ colorspace [1]. Uniformly discretizing the CMY colorspace domain, by M_1, M_2 and M_3 points in each of the C, M, Y coordinates respectively, the CRC can be adequately approximated by its response at a finite number of $M_t = M_1M_2M_3$ color combinations. Let $L^*_i a^*_i b^*_i(k)$ denotes the i-th values of the M_t discretized points time-varying output color gamut of the printer. In this context, the time dependent error vector, $\bar{\mathbf{e}}(k) \in \Re^{3M_t}$ of discretized output of the spatial signal is given by:

$$\bar{\mathbf{e}}(k) = \begin{bmatrix} \Delta L^*_1(k), \Delta L^*_2(k), \ldots, \Delta L^*_{M_t}(k), \\ \Delta a^*_1(k), \Delta a^*_2(k), \ldots, \Delta a^*_{M_t}(k), \\ \Delta b^*_1(k), \Delta b^*_2(k), \ldots, \Delta b^*_{M_t}(k) \end{bmatrix}^T \tag{1}$$

where $\Delta L^*_i(k) := L^*_i(k) - L^{**}_i$, $\Delta a^*_i(k) := a^*_i(k) - a^{**}_i$, $\Delta b^*_i(k) := b^*_i(k) - b^{**}_i$. $L^{**}_i a^{**}_i b^{**}_i$ gives the i-th value of the nominal color gamut. In the color stabilization control formulation [1], we can then specify the relative importance of the CRC errors at different colors for their stabilization given by the performance index term:

$$J(k) = \bar{\mathbf{e}}^T(k)\mathbf{W}_e\bar{\mathbf{e}}(k) \in \Re \tag{2}$$

where the error weighting is given by $\mathbf{W}_e \in \Re^{3M_t \times 3M_t}$. The error weighting matrix is made up of the color and image specific weightings i.e.:

$$\mathbf{W}_e = \mathbf{W}_{\wp}\mathbf{W}_{\Delta E^*} \tag{3}$$

where

1. $\mathbf{W}_{\Delta E^*} \in \Re^{3M_t \times 3M_t}$ is a diagonal matrix which gives the color specific weighting. It accounts for the sensitivity of the human visual system (HVS) to colors variations.
2. $\mathbf{W}_{\wp} \in \Re^{3M_t \times 3M_t}$ is also a diagonal matrix with diagonal element of either 1 or 0 - 1 indicating nominal colors that are significant in an image, $\wp$ and 0 indicating non-significant colors.

Color specific weighting, $\mathbf{W}_{\Delta E^*}$

In the color-science literature [7], the color difference formulation, $\Delta E^*_{94,i}$ is typically used to quantify the difference (distance

measure) between i-th index value of the printed color gamut and the corresponding i-th index value of the nominal color gamut. This color difference formulation given as follows, can be used to specify the color specific weighting, $\mathbf{W}_{\Delta E^*}$.

$$\Delta E^*_{94,i}(k) = \left[\left(\frac{\Delta L^*_i(k)}{k_{L,i} S_{L,i}} \right)^2 + \left(\frac{\Delta C^*_{ab,i}(k)}{k_{C,i} S_{C,i}} \right)^2 + \left(\frac{\Delta H^*_{ab,i}(k)}{k_{H,i} S_{H,i}} \right)^2 \right]^{\frac{1}{2}} \tag{4}$$

with

$$\Delta C^*_{ab,i}(k) = (a^{*2}_i(k) + b^{*2}_i(k))^{\frac{1}{2}} - (a^{**2}_i + b^{**2}_i)^{\frac{1}{2}} \tag{5}$$

$$\Delta H^*_{ab,i}(k) = \left[\Delta a^{*2}_i(k) + \Delta b^{*2}_i(k) - \Delta C^{*}_{ab,i}{}^2(k) \right]^{\frac{1}{2}} \tag{6}$$

$$S_{L,i} = 1 \tag{7}$$

$$S_{C,i} = 1 + 0.045 C^{**}_{ab,i} \tag{8}$$

$$S_{H,i} = 1 + 0.015 C^{**}_{ab,i} \tag{9}$$

$$C^{**}_{ab,i} = (a^{**2}_i + b^{**2}_i)^{\frac{1}{2}} \tag{10}$$

k_L, k_C, k_H are positive real-valued scaling parameters. As noted in [7], $k_L = k_C = k_H = 1$ at reference conditions [1].

Assuming reference conditions and substituting (6) into (4) we have:

$$\Delta E^{*}_{94,i}{}^2(k) = \left(\frac{\Delta L^*_i(k)}{S_{L,i}} \right)^2 + \left(\frac{\Delta a^*_i(k)}{S_{H,i}} \right)^2 + \left(\frac{\Delta b^*_i(k)}{S_{H,i}} \right)^2 + \Delta C^{*}_{ab,i}{}^2(k) \left(\frac{1}{S^2_{C,i}} - \frac{1}{S^2_{H,i}} \right) \tag{11}$$

Since the actual colors to be stabilized are expected to be closed to the nominal, it is assumed here that $\Delta C^*_{ab,i} \approx (\Delta a^{*2}_i + \Delta b^{*2}_i)^{\frac{1}{2}}$. Hence, from (11) we have:

$$\Delta E^{*}_{94,i}{}^2(k) \approx \left(\frac{\Delta L^*_i(k)}{S_{L,i}} \right)^2 + \left(\frac{\Delta a^*_i(k)}{S_{C,i}} \right)^2 + \left(\frac{\Delta b^*_i(k)}{S_{C,i}} \right)^2$$

Therefore, the performance index given in (2) assuming for the time being that $\mathbf{W}_{\wp} = \mathbf{I}$ (i.e. picking all colors in the discretized gamut) is given by:

$$J(k) = \sum_{i=1}^{M_t} \Delta E^{*}_{94,i}{}^2(k) = \tilde{\mathbf{e}}^T(k) \mathbf{W}_{\Delta E^*} \tilde{\mathbf{e}}(k) \tag{12}$$

where

$$\mathbf{W}_{\Delta E^*} = \mathrm{diag} \begin{pmatrix} 1/S^2_{L,1} \\ \vdots \\ 1/S^2_{L,M_t} \\ 1/S^2_{C,1} \\ \vdots \\ 1/S^2_{C,M_t} \\ 1/S^2_{C,1} \\ \vdots \\ 1/S^2_{C,M_t} \end{pmatrix} \in \Re^{3M_t \times 3M_t} \tag{13}$$

[1] The reference conditions assume that a pair of nontextured specimens in edge contact are placed on a background with $L^* = 50$, illuminated by D65 with an illuminance of 1000 lux, and viewed at a distance so that the pair subtends a visual angle of at least 4^o. Although not explicitly stated, a perceptibility judgment is assumed

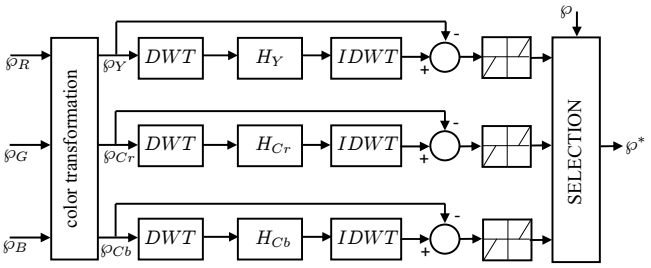

Figure 1. *Finding the significant color map, $\wp^*$. Here DWT denotes the discrete wavelet transform and $IDWT$ denotes the inverse wavelet transform*

Image specific weighting, $\mathbf{W}_{\wp}$

The image specific weighting, $\mathbf{W}_{\wp} \in \Re^{3M_t \times 3M_t}$ is achieved by a two-step procedure:

Step 1: Determine the most significant colors

Let $C_{\wp} \subset \Re^3$ denote the color gamut of the image, $\wp$. In this step, we determine $C^*_{\wp} \subset C_{\wp}$ such that $C^*_{\wp}$ gives the collection of colors of image, $\wp$ that are visually most significant to a human viewer. The basic idea of the proposed approach to determine $C^*_{\wp}$ is shown in Figure 1 and described as follows:

- A RGB image $[(\wp_R, \wp_G, \wp_B)$ image plane] is transformed to the YCrCb format $[\rightarrow (\wp_Y, \wp_{Cr}, \wp_{Cb})$ image plane]. This representation is used because the Y(luminance), Cr(chrominance-red) and Cb(chrominance-blue) color planes by themselves are widely de-correlated [5], and hence provide a suitable platform for the development of the algorithm to follow.
- The discrete wavelet transform (DWT) with *embedded HVS model* is then performed on each image plane individually. The wavelets coefficients of this decomposition are then used to synthesize a color "enhanced" image with the inverse discrete wavelet transform (IDWT).
- By comparing the original image at each color plane with that of the reconstructed image, we find areas where there are significant differences determined by a threshold value, ε_T. These areas indicate the location where there is significant color or spatial information on that color plane. By picking these areas on the original image, $\wp$ we have the significant color image, $\wp^*$ where $C^*_{\wp}$ can be obtained.

To embed the HVS into the 2D wavelet decomposition, the contrast sensitivity function (CSF) is used. The CSF quantitatively describes the human eye's sensitivity to changes in luminance and chrominance as a function of spatial frequency. The shapes of the chrominance CSF's are different from that of the luminance CSF's. The following luminance and chrominance CSF's models due to Marcus [6] are used:

Luminance's CSF:

$$H_Y(f_s) = 0.997 f_s^2 \exp(-0.970 f_s^{0.758}) + 0.22 \exp(-0.800 f_s^{1.999}) \tag{14}$$

Chrominance's CSFs:

$$H_{Cr}(f_s) = \exp(-0.1521 f_s^{0.893}) \tag{15}$$

$$H_{Cb}(f_s) = \exp(-0.2041 f_s^{0.900}) \tag{16}$$

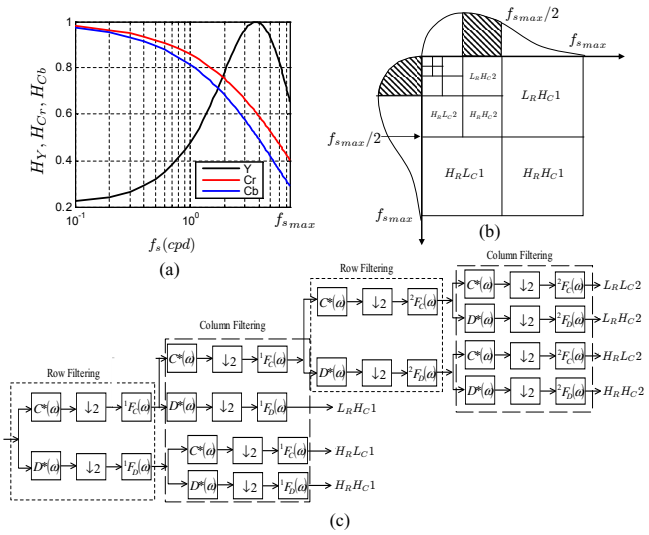

Figure 2. *(a)The CSF for luminance(Y) and chrominance(Cr,Cb) channels; (b) Relation between the CSF and a 4-level 2D wavelet decomposition. As an illustration, the shaded areas of the CSF function at the vertical and horizontal frequency axis correspond to the L_RH_C2 of the wavelet decomposition. The weighting that must be used for the wavelet coefficients at each sub-band is described by the corresponding horizontal and vertical frequency range in the CSF functions; (c) 2-level 2D wavelet decomposition with embedded human visual system.*

where f_s is spatial frequency in cycles per degree [cpd]. The frequency representation in cpd makes the CSF independent of the view distance. The CSFs for each Y, Cr and Cb color channels are given in Figure 2(a).

The relationship between the CSF and a 4-level 2D wavelet decomposition is given in Figure 2(b) [6]. The 2D wavelet decomposition gives the corresponding horizontal and vertical details of the image at each level of decomposition. The frequency range of the sub-band corresponds to the horizontal and vertical frequency range of the CSF functions. The maximum frequency range, $f_{s_{max}}$ of the CSFs are determined by how much the human eye can tolerate at a certain viewing distance, v [m] for a given image resolution, r [pixel/inch], given as follows:

$$f_{s_{max}} = v\tan(0.5^o)r/0.0254 \text{ [cpd]} \tag{17}$$

Hence,CSF functions to be embedded into the wavelet decomposition depend on the viewing distance and the quality of image to be viewed. Typically images with $r = 72$ pixel/inch resolution are viewed from a distance of $v = 0.3$ m.

To perform the 2D wavelet transform, the typical 1D wavelet transform can be extended using separable wavelet filters[2]. With separable filters, the 2D wavelet decomposition can be computed by applying a 1D transform to all the rows of the input, and repeating this on all of the columns as shown in Figure 2(c).

[2] $C^*(\omega), D^*(\omega)$ shown in Figure 2(c) are the low-pass and the corresponding high-pass 9/7 biorthogonal separable wavelet filters system respectively. This basis is used because for image transform coding it gives the best distortion rate performance. The 9/7 biorthogonal wavelet provides an appropriate trade-off between vanishing moments, support and regularity requirements. This biorthogonal wavelet basis is nearly orthogonal and thus introduces no numerical instability.

The CSF's embedding with this decomposition is also shown in Figure 2(c). The embedded CSF model in this decomposition tree, at each decomposition level, ℓ for the low pass sub-band denoted by ${}^{\ell}F_C(\omega)$ and high pass sub-band denoted by ${}^{\ell}F_D(\omega)$ are given as follows (see [6] for details):

$$ {}^{\ell}F_C(\omega) = H_{cc}\left(\frac{|\omega| f_{s_{max}}}{\pi 2^{\ell}}\right) \tag{18}$$

$$ {}^{\ell}F_D(\omega) = H_{cc}\left(\frac{f_{s_{max}}}{2^{\ell-1}} - \frac{|\omega| f_{s_{max}}}{\pi 2^{\ell}}\right) \tag{19}$$

where $\omega \in [-\pi, \pi]$ and $cc = \{Y, Cr, Cb\}$ denotes the color channels with H_{cc} given by equations (14)-(16). The embedding of these HVS filters makes the wavelet decomposition as shown in Figure 2(c) different from that of standard wavelet decomposition. These filters appropriately weights the spatial frequency band of the image that are critical to a human viewer in each of the de-correlated Y, Cr and Cb color planes.

To implement ${}^{\ell}F_C(\omega), {}^{\ell}F_D(\omega)$ filters and embed them in the wavelet decomposition, they are approximated here by FIR-filters. Note that ${}^{\ell}F_C(\omega)$ and ${}^{\ell}F_D(\omega)$ have even symmetry(i.e. ${}^{\ell}F_C(\omega) = {}^{\ell}F_C(-\omega)$ and ${}^{\ell}F_D(\omega) = {}^{\ell}F_D(-\omega)$), hence they can be expressed by real and even Fourier series:

$$ {}^{\ell}F_C(\omega) = {}^{\ell}b_{C0} + 2\sum_{k=1}^{n_F} {}^{\ell}b_{Ck}\cos(k\omega) \tag{20}$$

$$ {}^{\ell}F_D(\omega) = {}^{\ell}b_{D0} + 2\sum_{k=1}^{n_F} {}^{\ell}b_{Dk}\cos(k\omega) \tag{21}$$

where

$$ {}^{\ell}b_{D0} = \frac{1}{\pi}\int_0^{\pi} {}^{\ell}F_D(\omega)d\omega \qquad {}^{\ell}b_{Dk} = \frac{1}{\pi}\int_0^{\pi} {}^{\ell}F_D(\omega)\cos(k\omega)d\omega$$

$$ {}^{\ell}b_{C0} = \frac{1}{\pi}\int_0^{\pi} {}^{\ell}F_C(\omega)d\omega \qquad {}^{\ell}b_{Ck} = \frac{1}{\pi}\int_0^{\pi} {}^{\ell}F_C(\omega)\cos(k\omega)d\omega$$

The coefficients of the HVS embedded wavelet decomposition are used to synthesize a new visually enhance image, $\wp_{cc}^S$ at each $cc = \{Y, Cr, Cb\}$ de-correlated color channels. To locate the areas where significant color and spatial information reside, the enhance image is compared to the original image at each color planes. The difference between the the original and visually enhanced synthesized image at pixel (x,y) is given as follows:

$$ \delta_{cc}(x,y) = \left(\frac{\wp_{cc}(x,y) - \bar{w}}{w_{max} - w_{min}}\right) - \left(\frac{\wp_{cc}^S(x,y) - \bar{z}}{z_{max} - z_{min}}\right) \tag{22}$$

where $\bar{w} = \text{mean}(\wp_{cc})$, $w_{max} = \max(\wp_{cc})$, $w_{min} = \min(\wp_{cc})$, $\bar{z} = \text{mean}(\wp_{cc}^S)$, $w_{max} = \max(\wp_{cc}^S)$, and $w_{min} = \min(\wp_{cc}^S)$. To obtain the binary significant map, $\Im_{cc}$ at the particular cc color plane, thresholding is performed.

$$ \Im_{cc}(x,y) = \begin{cases} 1 & \text{for } |\delta_{cc}(x,y)| > \varepsilon_T \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

where $\Im_{cc}(x,y)$ gives the binary significant map value at (x,y) pixel and ε_T is the threshold setting.

Figure 3 shows the results of this algorithm. The original image is shown on the left with the corresponding color gamut of this image in term of the printer's output color gamut (or equivalently the expected output colors in CIE L*a*b* for a nominal

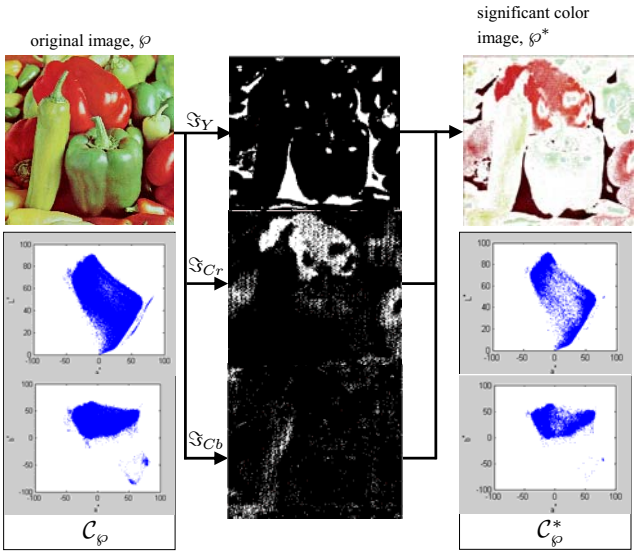

Figure 3. *The original image, $\wp$ with the corresponding desired color gamut in the printer's $L^*a^*b^*$ colorspace, $C_\wp$ (left); the binary significant map at each YCrCb color planes using $\varepsilon_T = 3\sigma, \sigma = std(|\delta_{cc}|)$ (middle); the resulting significant color image, $\wp^*$ with the corresponding desired color gamut in the printer's $L^*a^*b^*$ colorspace, $C^*_\wp$ (right)*

printer system). Applying the algorithm yields the binary significant map for each of the Y, Cr and Cb color plane as shown in the middle. The white region gives the region of interest at each color planes. In the luminance(Y) plane, note that the dark regions of the color image have been selected because it has stark contrast to the surrounding brighter region and appears more intense. Also the red color appears intense in the picture, and this has been selected in the red chrominance(Cr) color plane. By picking the region of interest(white regions on the binary significant map of all the color planes) on the original image, $\wp$ we have the significant color image, $\wp^*$ as shown on the right of Figure 3. The corresponding significant color gamut, $C^*_\wp$ is also shown. Clearly, the required colors to consider for stabilization are less compared to the colors in the original image.

In this algorithm, a higher threshold setting, ε_T enables restriction to a smaller set of critical colors for stabilization. Appropriate setting of this threshold depends on trade-off between color consistency requirements (i.e. quality of printouts) and capability of the system to compensate for disturbances.

Step 2: Deduce the image specific weighting

Deduce the image specific weighting, $\mathbf{W}_\wp \in \Re^{3M_t \times 3M_t}$ from the significant color image gamut, $C^*_\wp$. In general if a collection of p images (i.e. $\wp_1, \ldots, \wp_p$) needs to be printed, we deduce $\mathbf{W}_\wp$ from $\bigcup_{i=1}^{p} C^*_{\wp_i}$. Let $w_\wp(i)$ be the weight assigned for the i-th value of the print-system *nominal* color gamut given by $[L_i^{**}, a_i^{**}, b_i^{**}]$. Hence, for $i = 1, \ldots, M_t$, $w_\wp(i)$ is defined by:

$$w_\wp(i) = \begin{cases} 1 & \text{if } [L_i^{**}, a_i^{**}, b_i^{**}] \in C^*_\wp \\ 0 & \text{otherwise} \end{cases}$$

Let $\mathbf{w}_\wp = [w_\wp(1), \ldots, w_\wp(M_t)]$, then the image specific weighting can now be defined as $\mathbf{W}_\wp = \text{diag}(\mathbf{w}_\wp, \mathbf{w}_\wp, \mathbf{w}_\wp)$(corresponding to L^*, a^* and b^* coordinates)

Conclusion

In this paper an approach is proposed to determine the visually significant colors of a given printed image or images. Procedures for determining a color and image specific weightings are described. These weightings are important since they focused on the limited actuation authorities of the print process to stabilize colors of a given image/images that are visually significant to a human viewer. This is an important consideration for a successful color consistency control strategy.

Acknowledgment

This research is supported by the National Science Foundation under grant: CMS-0201622 and Xerox UAC grant. Inputs from Professor Ahmed Tewfik, ECE Department, University of Minnesota are gratefully acknowledged.

References

[1] T.P.Sim and P.Y.Li, Direct color consistency control for xerographic printing, in Proceedings of the 25th International Congress on Digital Printing Technologies (NIP25), Louisville.

[2] T.P.Sim and P.Y.Li, Optimal time-sequential sampling of color reproduction characteristic function, in Proceedings of the 22th International Congress on Digital Printing Technologies (NIP22), Denver, pp. 585–591 (2006).

[3] T.P.Sim, P.Y.Li and D.J.Lee, Using time-sequential sampling to stabilize the color and tone reproduction functions of a xerographic printing process, IEEE Trans. on Control Systems Technology, 15(2), 349 (2007).

[4] T.P.Sim and P.Y.Li, Optimal time sequential sampling of tone reproduction function, in Proceedings of the 2006 American Control Conference, Minneapolis, United States, pp. 5728–5733 (2006).

[5] Gaurav Sharma, Digital Color Imaging Handbook, CRC Press, Inc., Boca Raton, FL, US (2002).

[6] M.J.Nadenau, Integration of Human Color Vision Models into High Quality Image Compression, Ph.D. thesis, École Polytechnique Federerale De Lausanne (2000).

[7] R.S.Berns, Billmeyer and Saltzman's principles of color technology, John Wiley and Sons Inc. (2000).

Author Biography

Teck Ping, Sim received the Ph.D. in Mechanical Engineering for the University of Minnesota, Twin Cities in 2009, working on the color stabilization control for the xerographic printing process.

Perry Y. Li received the Ph.D. in Mechanical Engineering from the University of California, Berkeley in 1995. From 1995-1997, he was a research staff at Xerox Corporation, Webster, NY. He has been with the University of Minnesota since 1997 and is currently Professor of Mechanical Engineering. His research interests are in mechatronics and control systems as applied to fluid power, robotics, and imaging systems.