Farhan A. Baqai, Je-Ho Lee, A. Ufuk Agar, and Jan P. Allebach

Digital Color Halftoning

Problems, algorithms, and recent trends

Digital color halftoning is the process of transforming continuous-tone color images into images with a limited number of colors. The importance of this process arises from the fact that many color imaging systems use output devices such as color printers and low-bit depth displays that are bilevel or multilevel with a few levels. The goal is to create the perception of a continuous-tone color image using the limited spatiochromatic discrimination capability of the human visual system.

In decreasing order of how locally algorithms transform a given image into a halftone and, therefore, in increasing order of computational complexity and halftone quality, monochrome digital halftoning algorithms can be placed in one of three categories: 1) point processes (screening or dithering), 2) neighborhood algorithms (error diffusion), and 3) iterative methods. All three of these algorithm classes can be generalized to digital color halftoning with some modifications. For an in-depth discussion of monochrome halftoning algorithms, the reader is directed to the July 2003 issue of *IEEE Signal Processing Magazine*. In the remainder of this article, we only address those aspects of halftoning that specifically have to do with color. For a good overview of digital color halftoning, the reader is directed to [1]. In addition, [2] contains a more in-depth treatment of some of the material found in this article.

The naive digital color halftoning approach is to apply these monochrome halftoning techniques scalarly and independently to the color [red, green, and blue (RGB)] or colorant [cyan, magenta, yellow, and black (CMYK)] planes. As expected, this scalar approach leads to color artifacts and poor color rendition because it does not exploit the correlation between color or colorant planes, which is a key element in our color perception and appreciation of the halftone quality.

Color printing presents many problems that we encounter in monochrome printing. However, it also presents some problems that are unique to color printing. For instance, if the colorants are printed on top of each other, this will result in dots that contrast undesirably with the paper in the highlights. Another problem unique to color halftoning is that of "misregistration." This is due to the difficulty of controlling the relative offsets of the different colorant

1053-5888/05/$20.00©2005IEEE

planes. If the colorants are ideal, the visual appearance is independent of small registration errors. Unfortunately, real colorants are not ideal; and registration errors can cause significant color shifts. Therefore, an important objective of the color halftoning scheme is robustness to registration errors and minimization of dot-on-dot printing.

COLOR DEVICE MODELS

Color halftoning algorithms that do not account for the nonideal characteristics of colorants and the complicated dot interactions suffer from incorrect color reproduction. Several techniques to model and characterize color printing have been developed. These techniques can be broadly classified into three groups [2]: 1) approaches that characterize the spectral distribution created by a certain combination of colorant dots at a given printer addressable location, 2) dot model based approaches in which a model accounting for nonideal dot interactions between neighboring dots is incorporated into the halftoning algorithm, and 3) empirical tone correction based approaches that precompensate for dot interactions by modifying the image data using a look-up-table (LUT) prior to halftoning and printing.

The most important spectral (colorimetric) model used in color printing is the Neugebauer color mixing model. According to the model, the spectral distribution $R(\lambda)$ as a function of wavelength λ of a color halftone obtained using three colorants can be written as an area-weighted sum of the reflectances of the one, two, or three color overprints of the three colorants and of the paper substrate. These eight colorant combinations are called the Neugebauer primaries. The spectral distributions of the eight Neugebauer primaries printed with an HP DeskJet 970C three colorant inkjet printer are shown in Figure 1.

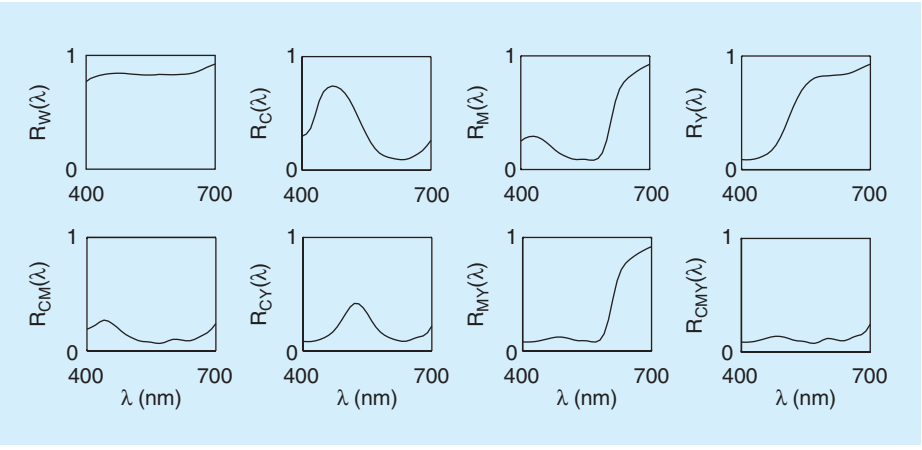
Among the dot models, the hard circular dot model is very popular. This model is an idealization of printer behavior and assumes that every printer dot is a perfect circle [3]. Another novel dot interaction model is the 2×2 centering model developed by Wang [4]. The novelty of this method is the idea that the printer grid is assumed to be shifted by (0.5, 0.5) pixels from the image grid. This has the effect of making each pixel dependent on only four printer pixels instead of nine printer pixels, as would be the case if there were no offset between the image lattice and the printer lattice. The result is that the number of possible different dot combinations is greatly reduced. This eases both measurement and storage requirements.

Empirical tone correction approaches can be applied by halftoning and printing all possible colors, making colorimetric measurements of the resulting halftones, and applying the inverse mapping of this transformation to color images prior to halftoning, to obtain halftones with desired colorimetric values. Since measuring halftone patches for a large set of colors, e.g., $256^3 \approx 16{,}750{,}000$ for an 8-b color imaging system, is infeasible; it is impossible to have a complete color correction LUT. Instead, a limited number of halftoned colors are stored in a LUT, and interpolation schemes are employed to obtain the colorimetric values for the remaining colors. In this case, the number of different halftone patches that must be printed and measured will be much more manageable.

COLOR HUMAN VISUAL SYSTEM MODELS

Psychophysical studies prove that color/tone discrimination and appearance closely depend on spatial pattern in addition to the global differences in perceptual attributes like hue, saturation, and brightness. This finding has motivated the use of color human visual system (HVS) models for design and evaluation of the performance of digital halftoning algorithms, which aim to minimize the visual differences between continuous-tone images and their halftones.

Zhang and Wandell used an opponent channel representation based HVS model called S-CIELAB to predict the perceived quality of color halftone images [5]. In S-CIELAB, a spatial extension of CIELAB, a color image is transformed into an opponent-color space and decomposed into luminance, [Red-Green $(R-G)$] and [Blue-Yellow $(B-Y)$] images. Each opponent-color image is filtered by a different two-dimensional separable spatial kernel. After filtering each opponent-color image, the filtered representation is transformed to a CIE-XYZ representation, and the S-CIELAB values are computed using the standard CIE $L^*a^*b^*$ formulae. The nonlinear transformation from CIE XYZ tristimulus values to CIE $L^*a^*b^*$ values cascaded with the linear transformation from RGB image coordinates to CIE XYZ tristimulus values does not preserve the spatially averaged tones of images, which is undesirable in halftone color reproduction. Flohr et al. [6] linearized the CIE $L^*a^*b^*$ color space about the reference stimulus (X_n, Y_n, Z_n). They denoted this linearized color space by $Y_yC_xC_z$ and defined the linearized components as



[FIG1] Spectra of Neugebauer primaries measured from prints made with an HP DeskJet 970C printer.

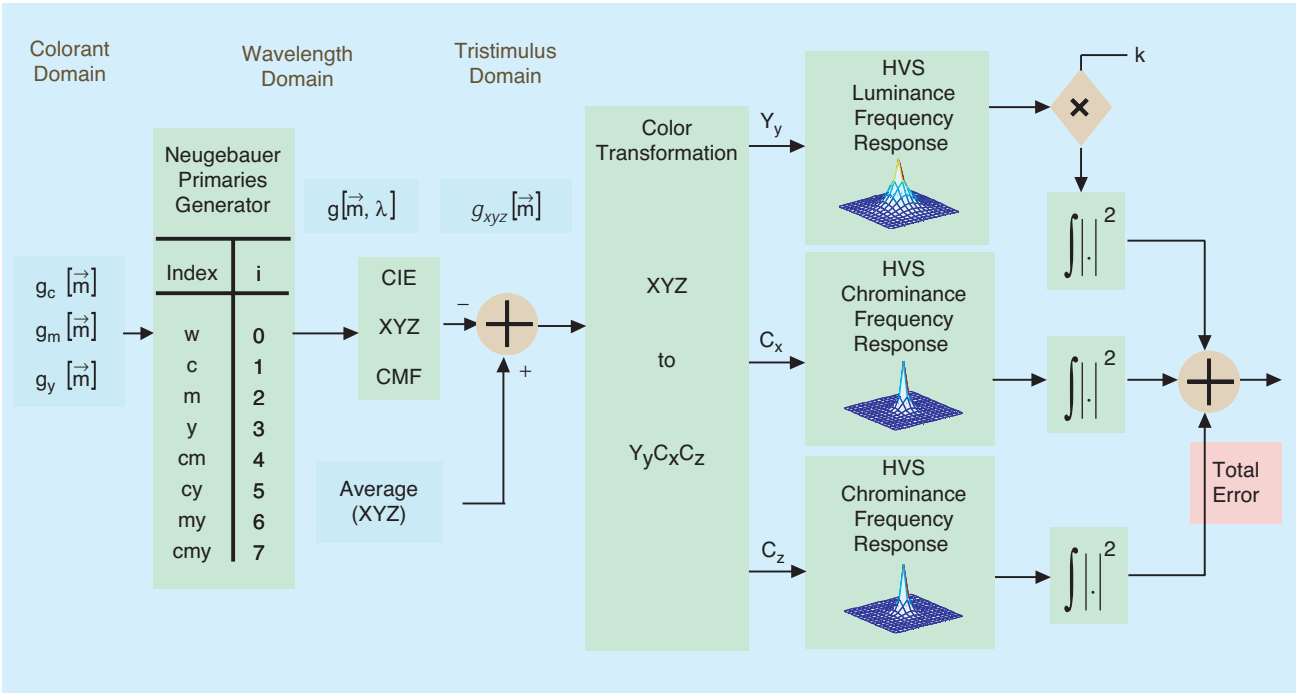
$$Y_y = 116\frac{Y}{Y_n} - 16,$$
$$C_x = 500\left[\frac{X}{X_n} - \frac{Y}{Y_n}\right],$$
$$C_z = 200\left[\frac{Y}{Y_n} - \frac{Z}{Z_n}\right]. \quad (1)$$

The Y_y component serves as a correlate of luminance, and the C_x and C_z components are similar to the $R-G$ and $B-Y$ opponent color chrominance components. In the Flohr HVS model, the luminance channel is based on the Näsänen [7] exponential model and the model for chrominance is due to Kolpatzik and Bouman [8]. These channel responses may be based on other models as well, including those proposed in [5]. Kim and Allebach [9] and Monga et al. [10] compared the performance of different monochromatic models and different color models, respectively.

The overall color imaging system model can be described in terms of a color device model and a color HVS model. One such model [11] that employs the spectral Neugebauer color mixing model, the CIE color matching functions (CMFs), and Flohr's color HVS model is illustrated in Figure 2. This model can be used for joint design of halftone textures to represent constant colors. Its input consists of digital halftones $g_i[\vec{m}]$, $i = c, m, y$ for the three colorants. These are converted to a spatial map of CIE *XYZ* tristimulus values. Since we are interested in assessing the visibility of fluctuations about the average colorant value, we subtract that average value before conversion to (Y_y, C_x, C_z). We then apply spatial filters according to the frequency response of the HVS in these three channels. Note that although the luminance and chrominance frequency responses are both low pass in nature, 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. This means that the chrominance channel has a narrower bandwidth than the luminance channel. Therefore, more chromatic error than luminance error is allowed at lower spatial frequencies. The parameter k controls the relative gain of the three channels. The possibility of increasing the perceived smoothness of halftone textures by trading errors in the luminance channel for errors in the chrominance channel was first pointed out by Mulligan [12].

COLOR SCREEN DESIGN

In screening, threshold arrays are applied to each of the colorant planes independently, and then the resulting halftones with different colorants are superimposed to generate a final color halftone. Halftone screens are of two types: periodic clustered-dot and dispersed-dot. Clustered-dot screens produce binary textures in which the individual printer dots are grouped into clusters. Dispersed dot screens spread the individual printer dots as far apart as possible. They work well for printing processes that produce consistently positioned and sized dots with little or no dot gain. Clustered-dot halftones render well on printers that have difficulty generating an isolated dot (e.g., electro-photographic printers). They are quite robust to dot gain and variations in the process parameters. Clustered-dot halftones may suffer from moiré due to the interference between periodic structures in the image. Dispersed dot screens are generally less susceptible to moiré, but may appear noisier due to the stochastic nature of the halftone textures.



[FIG2] **Overall framework for perceptual model.**

WHAT CAUSES MOIRÉ AND ROSETTE ARTIFACTS?

Moiré results from the beating of two or more periodic structures against each other, which creates the appearance of new periodic structures with frequencies given by the sum and difference of the frequencies associated with the original structures. If the difference frequencies are low enough, they will be quite visible and will result in moiré patterns. Two distinctly different kinds of moiré can occur in color halftoning. The first is subject moiré that results when periodic structures in the image content to be rendered, such as textures in fabric, beat against the periodic halftone screens in one or more colorants that are used to render the fabric.

The second is screen-to-screen moiré due to interference between two or more of the halftone screens used to represent each of the colorants. In order minimize this type of moiré, the halftone screens for the different colorants are rotated to place them as far apart in angle as possible. A typical solution is to place the yellow screen at 0°, magenta (M) at 15°, black at 45°, and cyan (C) at 75°. Thus, with the exception of yellow, the spacing between the screens is 30°. The yellow screen is placed at 0° because of the low visibility of this colorant on white paper. When the screens are rotated in this manner, the interference between them causes a particular type of moiré known as rosettes, which are circular patterns of overlapping dots that, depending on the relative phasing of the halftone patterns from the different colorants, may have a clear center, a center containing a dot, or something in between [13]. The size of the rosette structures is about twice that of the period of the individual halftone screens; so this type of moiré is minimally visible.

DIGITAL COLOR HALFTONING IS THE PROCESS OF TRANSFORMING CONTINUOUS-TONE COLOR IMAGES INTO IMAGES WITH A LIMITED NUMBER OF COLORS.

DISPERSED-DOT APERIODIC SCREENS

A number of researchers have advocated the use of blue noise patterns for generating dispersed dot screens [14], [15]. The spectra of blue noise halftones are composed entirely of high-frequency components. These components are least visible to the human observer. To suppress moiré, Wang and Parker [15] jointly designed screens in which, not only each individual colorant screen produces halftones with blue noise properties, but the combination of the screens produce images that have blue noise characteristics, i.e., the dots from different colorant planes are uncorrelated. Shu et al. [16] used a modified void and cluster method for adaptive screening. They first applied Voronoi tessellation and its region areas to determine the locations of largest voids and tightest clusters, and then applied the void and cluster algorithm to redistribute the dots in the screen. In this manner, they achieved a homogeneous dot distribution for all tonal levels. Lin and Allebach [17] used the HVS model described earlier within a least-squares search algorithm to jointly design artifact-free color screens.

The screens discussed thus far are very useful for marking devices that have no trouble in rendering an isolated dot, e.g., inkjet printers. However, for electrophotographic devices (laser printers), which cannot reliably produce an isolated dot, these halftones are not very stable. Periodic clustered-dot color screens (to be discussed in the following) are widely used for such cases. They form the basis for commercial printing. To retain the homogeneity of blue noise textures and achieve the stability of periodic clustered-dot screens, Lau et al. proposed green noise halftoning [18]. Green-noise halftones have an aperiodic clustered-dot blue noise texture and consist of mid-frequency spectral components, i.e., they have dot clusters on blue-noise centers. The visibility of stochastic moiré is minimized by varying the coarseness between colorants such that black has the smallest cluster size, followed by M, C, and Y [19]. Lau et al. conclude that since the overlapping green-noise screens minimize the visibility of stochastic moiré without correlation, registration errors do not affect green-noise halftones as much as they do blue-noise halftones.

PERIODIC CLUSTERED DOT SCREENS

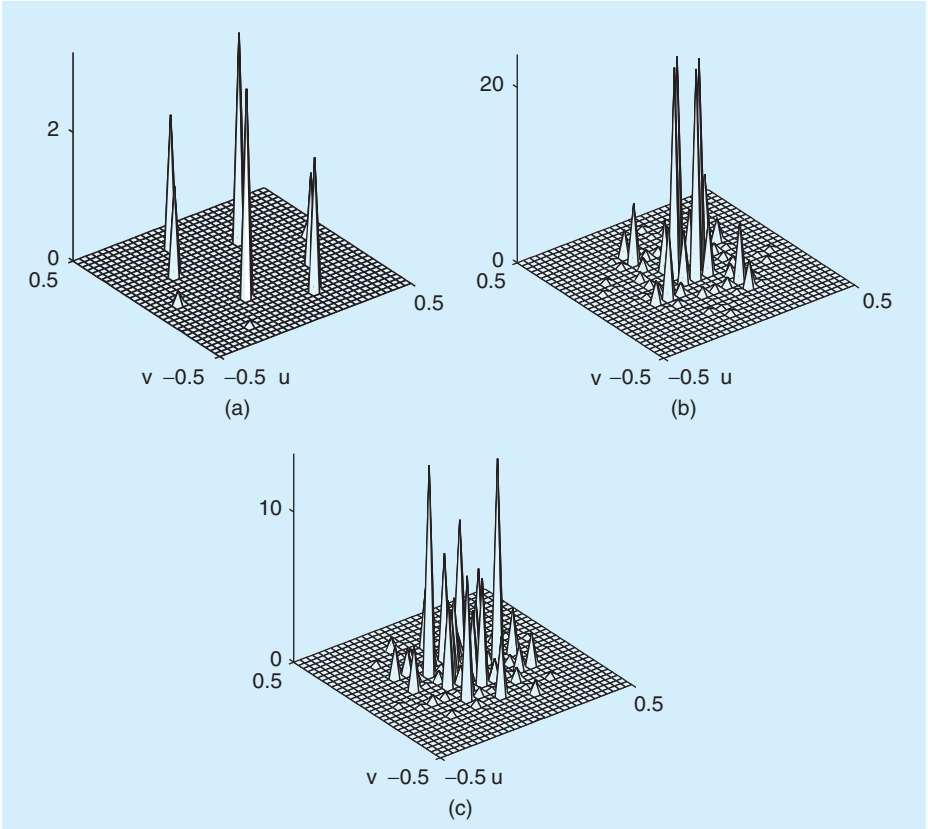
A number of approaches to analyzing moiré due to superposition of periodic clustered-dot screens for color halftoning have been reported. The geometric model [13], [20] is based on periods and angles of the superposed layers. It provides equations that, under certain limitations, predict the geometric properties of the moiré patterns. The algebraic approach [21] also gives the same result. Amidror et al. [22] analyzed moiré formation in the frequency domain while Kaji et al. [23] proposed a method to minimize moiré at the expense of rosette formation.

As mentioned previously, the conventional solution in the graphic arts industry to minimizing moiré is to use a single screen rotated to angles that are 30° apart for the C, M, and black colorants. The angles have to be quite accurate to avoid the appearance of moiré. These rotation angles are achieved by using the following three approaches [24]: rational tangent, irrational tangent, and rational supercell. The rational tangent method (conventional approach) only allows integer elements in the periodicity matrices, thereby restricting the available screen angles. The irrational tangent technique [25] is used in high-end graphic arts plotters. It employs floating point arithmetic and registers to attain precise angles. The drawback is that the screen cells are different from location to location, which can cause moiré [24]. In the rational supercell approach, a large cell that contains an array of cells is fit to the raster grid. Screen angles can be achieved to any desired precision by increasing the supercell size. This method requires less processing than the irrational tangent method. However, it requires large memory and the possibility of moiré still exists.

Delabastita [26] and, more recently, Wang et al. [27] derived conditions for moiré-free screening that are based on a first-order analysis of frequency domain aliasing. The basic concept is

to force the vector sum of the fundamental frequencies associated with two or more screens to be zero, corresponding to a moiré pattern with an infinitely long period. This solution leads to nonorthogonal screens for each colorant that are generally not rotated versions of each other. Baqai and Allebach [11] also proposed a framework based on nonorthogonal screens. However, they based their design on the model shown in Figure 2 that accounts for the spectral characteristics of the colorants, a detailed spatial model for the halftone pattern formed by the overlapping colorants, and the sensitivity of the human viewer to this halftone pattern. Using this model, they performed a numerical search for the jointly optimal parameters of the individual halftone screens.

A CMYK halftone screen is completely specified by the periodicity matrix, the offset vector, and the dot profile for each of the four colorants. The objective is to choose these parameters to minimize the visibility of moiré and rosette artifacts, and yield screens that are robust to registration errors.



[FIG3] Magnitude of the Fourier transform of the error in Y_y, weighted with the HVS luminance spatial frequency response, obtained for various screens for absorptance = 0.25. (a) Best, (b) worst, and (c) conventional.

The magnitude of the Fourier transform, for the error in Y_y weighted with the HVS luminance spatial frequency response for the best, worst, and conventional screens obtained by using their technique, for an absorptance of 0.25, are shown in Figure 3. The spectra for the two chrominance components have similar properties. It can be seen that the spectra for the best case has a few low-amplitude high-frequency spectral peaks, which cause no moiré or rosette artifacts, while the spectra for the worst and conventional cases contains high-amplitude low-frequency peaks, which can cause moiré or rosette artifacts. Magnified scanned textures of the corresponding color patches are shown in Figure 4. It can be seen that the best case gives smooth regions with no moiré or rosette artifacts. The worst case contains objectionable moiré patterns. Rosette patterns exist in the image halftoned by the screen rotated at conventional angles.

COLOR ERROR DIFFUSION

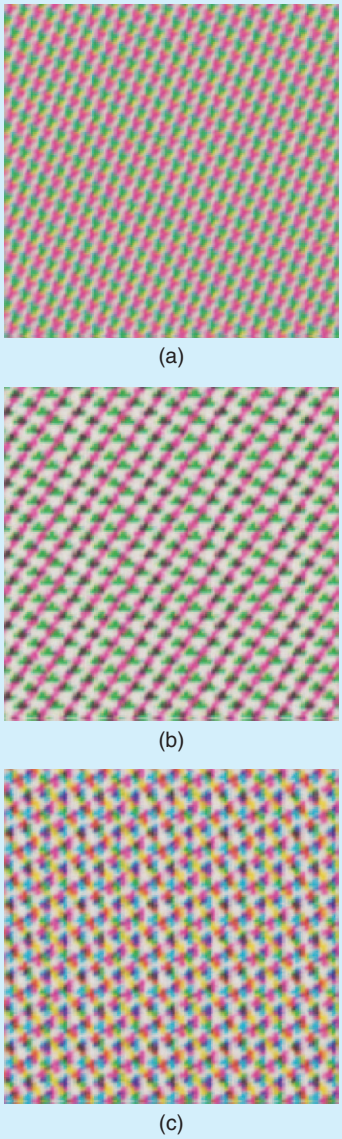
SCALAR VERSUS VECTOR APPROACHES

Error diffusion is a halftoning scheme that has been widely used in desktop printing. Even though it was originally developed and used for gray-scale images [28], error diffusion lends itself naturally to digital color halftoning, where the color components are halftoned either independently or jointly, and generally yields a higher level of image quality than does screening.

Typically, the diffusion part of the algorithm is done independently for each color channel using identical diffusion filters. The quantization step may be performed independently or jointly for the entire color vector. The first method is called scalar error diffusion, while the latter is called vector error diffusion. Haneishi et al. [29] compared scalar and vector error diffusion and analyzed the origins of the artifacts caused by each method. The scalar method often leads to color artifacts and poor color rendition because it does not exploit the correlations between color or colorant planes, which, as discussed earlier, is a key element in our perception of color and judgment of halftone quality. In order to overcome this problem while keeping the benefit of the simplicity of the scalar approach, Akarun et al. [30] proposed a scalar error diffusion scheme based on color coordinate systems that have less correlation between color components. They suggested the use of a pseudo-Karhunen-Loeve (KL) coordinate system, which is an approximation of the KL color coordinate system. Kolpatzik and Bouman [8] used separable error filters in a luminance-chrominance space to account for correlation among the color planes.

Vector error diffusion methods greatly reduce the interference between color planes and hence generally yield better

color halftone textures than do the scalar versions. However, for vector error diffusion methods, severe quantizer overload can occur, leading to other significant color artifacts such as smear and slow response, which do not appear in the case of the scalar approach. This happens when colors outside the gamut of the rendering device are continuously input to the system. In [29], Haneishi et al. proposed a conditional error propagation method to resolve the problem. Namely, if the difference between the input color and output colors of the neighboring pixels exceeds a predetermined threshold, this is regarded as a transition area, and the errors are omitted in the diffusion process. As one might expect, a drawback of this simple remedy is that the smear is reduced at the expense of lower spatial quality in transition regions. Kim et al. [31] also proposed two different techniques to reduce smear artifacts. The first one uses clipping, and the other one uses a window-based error minimization technique. In the first method, the optimal values for the error clipping are determined experimentally in order to generate a color image without color distortion caused by clipping. At the same time, the lightness of the modified input vector is also clipped if it falls outside the device gamut. In the window-based method, the quantization error is diffused over 2×2 blocks of pixels. This enables the algorithm to populate the output color space with many more directly reproducible colors, and to thereby reduce the maximum color errors that must be diffused.

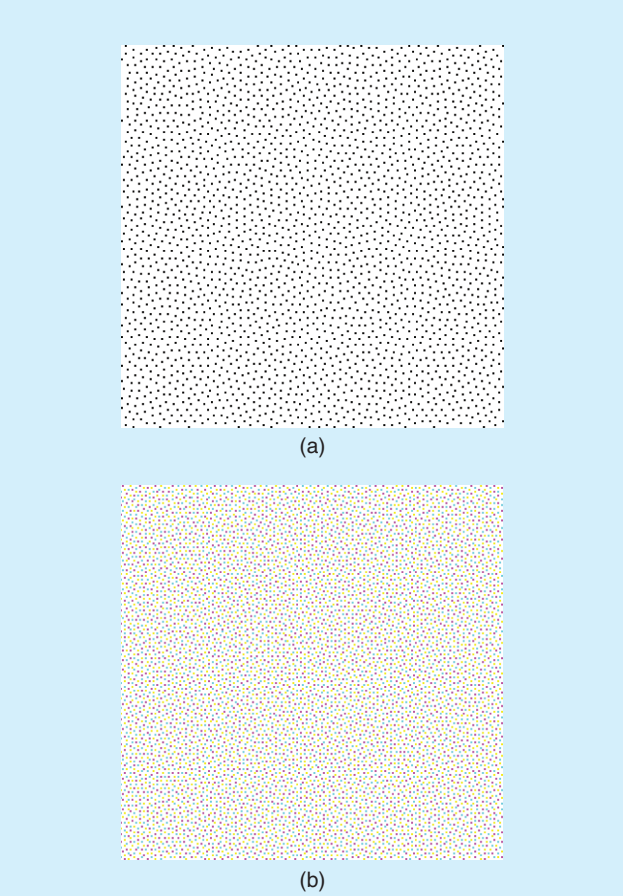


[FIG4] Magnified scanned textures for various screens for absorptance = 0.25. To be viewed from a distance of 72 in (printed at 300 dpi on an HP DeskJet 970Cxi, magnified by 6×). (a) Best, (b) worst, and (c) conventional.

COLOR SPACES FOR ERROR DIFFUSION

Color reproduction is an important problem in color science. To achieve colorimetric color reproduction between different rendering devices, it is necessary to use a common color space such as CIE-XYZ, or CIELAB rather than the device dependent RGB or CMYK representation. Haneishi et al. [29] performed vector error diffusion in several different color spaces such as the device independent color spaces CIE-XYZ and CIELAB, and pointed out that the color space chosen for the error diffusion made a significant difference in the accuracy of the color reproduction. They concluded that vector error diffusion with CIELAB space yielded the best performance compared with other color spaces. In the current approach to color printer design, color accuracy is normally addressed after selecting a halftoning scheme. In other words, the halftoning algorithm is developed first, with relatively little concern for the colorimetric average that results from a particular arrangement of a fixed number of CMY dots. This stage is followed by a calibration stage in which the printer mechanism and the halftoning algorithm are together treated as a black box. A large number of patches are printed with combinations of amounts of CMY, which correspond to points on a three-dimensional grid. Each patch is measured to determine its colorimetric value. These data then form the basis for developing a transformation from tristimulus values expressed in a standard color space, such as CIE-XYZ to the amounts of CMY that will result in the printing of a halftone texture that yields that average colorimetric value.

Therefore, the rationale for the use of color spaces other than the RGB or CMYK spaces has been more related to halftone quality than colorimetric reproduction issues. In particular, many authors have used a luminance-chrominance opponent color space to exploit the fact that human viewers are less sensitive to lower frequency chromatic errors than to luminance errors of the same frequency [8], [10], [32]–[34]. The result is to move the luminance quantization errors to the chrominance band, where errors are less visible. As we will see in the next section, this space has served as a basis for model-based approaches which exploit the characteristics of the human visual system.



[FIG5] Two halftone patterns illustrating (a) dot-on-dot and (b) dot-off-dot printing for a 7% gray level printed at 100 dpi.

VARIATIONS ON VECTOR ERROR DIFFUSION

Damera-Venkata and Evans [34] proposed a linearized vector error diffusion algorithm based on a matrix gain model for the quantizer, which accounts for correlations among the components of the error vector being diffused. The matrix gain model describes vector color diffusion in the frequency domain, and predicts noise shaping and linear frequency distortion produced by halftoning. Lau et al. [19] generalized the error diffusion algorithm with an output-dependent feedback algorithm, which was introduced by Levien [35]. In order to take into account color interactions, they added an interference term (matrix) in the feedback loop so that the overlap of pixels of different colors can be regulated for increased color control. Negative off-diagonal terms in the interference matrix inhibit the overlap, and positive terms encourage the overlap, of corresponding colorants.

Considering the fact that there may be significant differences in the statistics of different regions in a typical image, Akarun et al. [30], Lee et al. [33], and Bozkurt et al. [36] proposed adaptive schemes in which the diffusion coefficients are varied locally to minimize the mean square error between the average color value of the original image and the halftone image. Akarun et al., in particular, compared a vector adaptive algorithm with a scalar adaptive error diffusion algorithm, and demonstrated the superiority of the vector version. The appearance of color impulses was considerably reduced and smoother color transitions were achieved.

Much activity in the area of color digital halftoning has centered on the use of models for the HVS or rendering system. There has been a great deal of research considering how models may be incorporated into iterative color halftoning algorithms, which halftone the image in multiple passes. For color error diffusion, however, relatively little research has been done due to the limited controllability of halftone textures. Sullivan et al. [32] used a visual model that is similar to those discussed previously. The visual model is incorporated into an error feedback loop so that perceived errors are distributed to neighboring pixels. Lee et al. [33] took a similar approach to account for spatial and spectral sensitivities of the HVS.

Kolpatzik and Bouman [8] used a visual model that combines the models from Sullivan et al. and Näsänen [7]. In their method, the optimal error filter was determined so that a cascade of the quantization system and the observer's visual modulation transfer function yields a whitened error spectrum. The error diffusion is performed in the luminance-chrominance space, while the quantization is done in the RGB space. In [34], Damera-Venkata and Evans employed a linear color model, which is based on the pattern color separable model by Zhang and Wandell [37], in an opponent color space. Monga et al. [10] generalized this linear color model as a linear transformation to a desired color space, which is not necessarily the opponent representation, but any one that satisfies pattern color separability, followed by appropriate spatial filtering in each channel. Kolpatzik and Bouman [8] incorporated the modulation transfer function of the monitor into the computation of optimal diffusion filter coefficients. Pappas presented a circular-dot-overlap model for color printers in [3], which was an extension of the parametric model for the grayscale printer [38]. Kim et al. [31] also adopted this parametric model to implement a variation of error diffusion whereby the quantization occurs in CIELAB space.

RECENT TRENDS

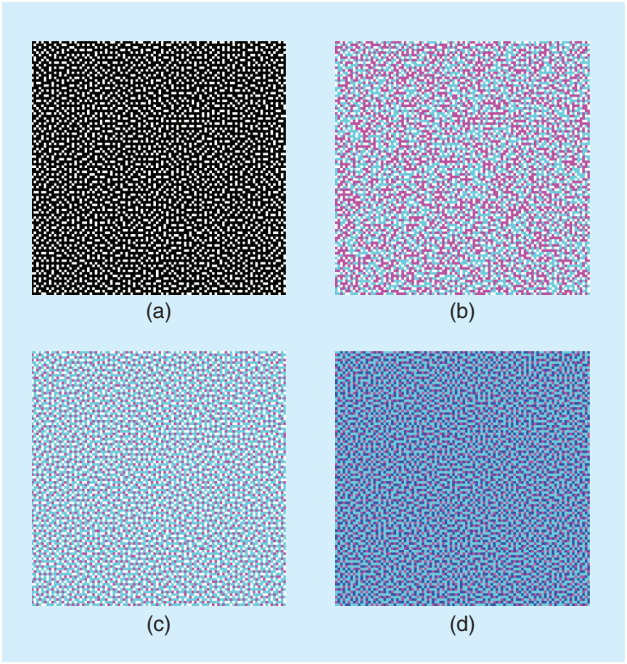
A recent trend in color halftoning research has been toward approaches that explicitly account for the differences in pattern visibility between textures composed of different sets of color planes to reduce artifacts that are unique to color halftoning. Consider a situation in which we wish to print a 100×100 pixel patch of constant gray with 7% absorptance. To achieve this level of gray, we will need to print 700 dots each of CMY. If we constrain the printing so that the CMY dots are printed on top of each other, we will have 700 black dots that can be arranged in a uniform pattern with 7% fill. This situation is maximally dot-on-dot. At the other extreme, we constrain the textures so that no dots of different colorants may be printed on top of each other. In this case, we have a texture consisting of a mixture of 2,100 CMY dots. The overall fill is now 21%, which will be perceived as a more uniform texture than a 7% fill, regardless of the color of the individual dots. In addition, each of

the CMY dots will contrast less with the white substrate than will the black dots, further reducing the overall visibility of the halftone texture. We refer to this case as maximally dot-off-dot. Figure 5 compares dot-on-dot and dot-off-dot printing for a 7% gray level. In addition to producing a smoother, more homogeneous texture, dot-off-dot printing yields a larger color gamut.

In order to yield smoother, less visible textures, several authors have considered approaches to constraining the primaries that are locally rendered to yield a particular visually averaged color [39]–[41]. Klassen et al. [39] exploited the limited spatial frequency response of the HVS by suggesting the selection of low-contrast color combinations wherever possible to generate the finest possible mosaic pattern without large clumps. This reduces the contrast of the halftone texture by replacing black pixels by dot-off-dot printing of CMY pixels. Shu [40] proposed an error diffusion algorithm combined with a color clustering method to enforce the exclusion of nonharmonic colors in a local area of an image. Shaked et al. [41] argued that brightness variation between adjacent dots in the print is a major cause of color noise, and a better halftone quality can be obtained by limiting the range of brightness in local areas in a halftone. This criterion, which is known as the minimum brightness variation criterion (MBVC), is based on the observation that the HVS is more sensitive to changes in brightness than in chrominance. Based on this criterion, the authors derived a simple postprocessing technique called ink relocation [41] for arbitrary color halftoning algorithms, where pairs of pixels with distinct brightness values are replaced with ones having a smaller brightness difference. They later proposed a color error diffusion algorithm, which provides full compliance with the MBVC, and optimal halftone patterns.

Lau et al. [19] proposed a metric which can quantitatively measure the visibility of the stochastic moiré, and showed how this metric can be used to identify the better of two arrangements of dots. Another suggested improvement by Levien [42] extended his output-dependent error diffusion algorithm [35] to support multiplane optimization. In this algorithm, color planes are ordered from darkest to lightest, and for each location in the scan order, a pixel for each plane is processed successively such that the area of overlapping dots between planes is minimized.

COLOR HALFTONING ALGORITHMS THAT DO NOT ACCOUNT FOR THE NONIDEAL CHARACTERISTICS OF COLORANTS AND THE COMPLICATED DOT INTERACTIONS SUFFER FROM INCORRECT COLOR REPRODUCTION.



[FIG6] Steps of Lee and Allebach's [44] color halftoning algorithm for the case where $C = 80\%$, $M = 50\%$, and $Y = 0\%$: (a) set overall dot arrangement, (b) color the dots randomly, (c) refine the halftones for C' and M', and (d) fill holes by B.

ITERATIVE APPROACHES TO COLOR HALFTONING

In iterative color halftoning algorithms, several passes over the halftone are made to minimize an error metric or satisfy certain constraints before the halftoning process is completed. The general approach is to start with an initial halftone and sequentially visit the individual halftone pixels to modify the halftone to reduce the error metric. We then iterate in this manner and make several passes over the halftone pixels until the error metric reaches a local minimum or until certain constraints are met.

Iterative color halftoning algorithms generally are capable of creating halftones resulting in noticeably superior visual quality than the previously discussed two groups of color halftoning algorithms, i.e., the point processes such as screening and the neighborhood algorithms such as error diffusion. However, this superior performance comes with a cost. Iterative algorithms are, generally, computationally significantly more expensive than the point processes or neighborhood algorithms for color halftoning. Due to their computational cost, iterative algorithms cannot be implemented in current printing devices for real-time halftoning of documents. However, they provide a gold standard for judging the quality of color halftone images, and serve as powerful tools for the design of simpler, more efficient derivative pixel-based or neighborhood-based halftoning algorithms. They can also be used to create and prestore halftones for printing of static documents or other content where high quality is extremely important.

The goal of the iterative color halftoning algorithms is to iteratively modify the halftone to minimize the perceived difference between the continuous-tone original image and its halftone. This goal is generally attained by minimizing a frequency weighted measure of halftoning error where the weights depend on a human visual model, e.g., the luminance and chrominance spatial sensitivity functions shown in Figure 2. To account for

the nonideal characteristics of colorants and complex dot and media interactions, iterative color halftoning algorithms may also include a color hardcopy model. We briefly discuss two examples in the following.

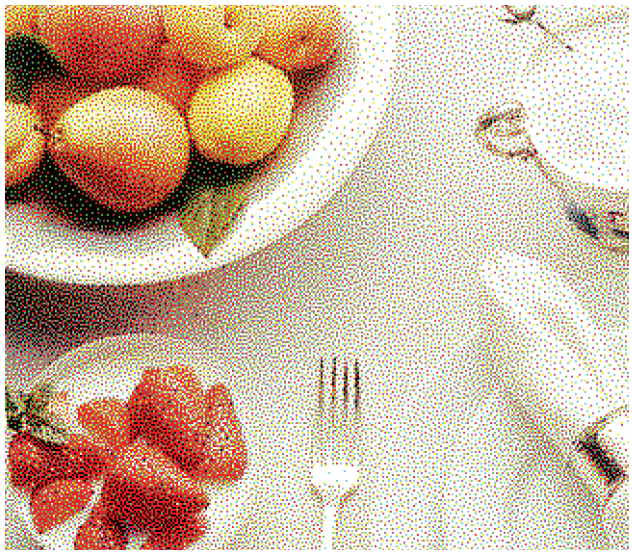
Pappas [3] developed a least-squares, model-based color halftoning algorithm that employed both a color hardcopy model and a HVS model. He used the circular dot overlap model described previously and a 2-D filter to model the HVS characteristics. He applied the same filter to all the color components (RGB) in his halftoning algorithm and minimized the least squares error metric independently for each color component.

Another example of HVS model-based iterative color halftoning algorithms is the model based color halftoning initially developed by Flohr et al. [6] which uses the Flohr HVS model [6] described previously. Flohr et al. [6] used the HVS model discussed previously as the basis for computing the total squared error in a luminance/chrominance space as the metric for their model based color halftoning algorithm. They minimized this error metric using a coordinate descent approach in which the state of each pixel in the halftone image is optimized independently and sequentially. Agar and Allebach [43] employed the same metric, but they modified the search algorithm to include swaps between each pixel and each one of its eight nearest neighbors. They also incorporated the 2×2 centering model developed by Wang [4], described previously, into this algorithm to prevent artifacts due to dot overlap and to improve color texture quality. They proposed a halftoning calibration method to ensure accurate colorant distributions in the resulting halftones.

Lee and Allebach [44] also used the direct binary search (DBS) algorithm but operated directly in the CMY colorant space instead of the linearized version of the CIELAB uniform color space. They employed only a luminance HVS model to enforce exclusion of visually nonhomogeneous patterns for the total dot distribution, as well as for each individual colorant texture. Their algorithm enables explicit control over the individual colorant textures and the way in which these textures interact to determine the overall visual appearance of the halftone. It also allows one to directly minimize dot-on-dot printing. Figure 6 illustrates the halftoning process in case of a constant-tone image, which consists of 80% C and 50% M. The sum of $C + M$ is greater than 100%, which implies that at least some portions of C and M have to be printed on top of each other thereby introducing the color blue (B). Let $C' = C - B$ and $M' = M - B$ denote the portions of pure C and M colors in the halftone image, respectively. To avoid dot-on-dot printing as much as possible, C', M', and B are given by

$$\begin{aligned} C' &= 1 - M, \\ M' &= 1 - C, \\ B &= C + M - 1. \end{aligned} \tag{2}$$

The halftoning algorithm first sets a homogeneous binary pattern for the total dot arrangement, which is a $C' + M' = 70\%$ fill monochrome dot pattern. Then the algorithm colors the dots so that each colorant plane has uniform texture, without altering the total dot distribution. Finally, the holes are simply filled with B. To illustrate the detail rendition provided by an iterative method, the result of colorant-based DBS for a natural image is shown in Figure 7.



[FIG7] Halftone generated by the colorant-based DBS algorithm. The image is printed at 100 dpi.

AUTHORS

Farhan A. Baqai received the M.Eng.Sc. degree in electrical and electronic engineering from the University of Melbourne, Australia, in 1993, and the M.S. and Ph.D. degrees in electrical and computer engineering from Purdue University in 1997 and 2000, respectively. He was a research engineer with Carrier Telephone Industries and a Research Associate with GIK Institute of Engineering Sciences and Technology. In 2000, he joined Xerox Corporation in Webster, New York. Currently, he is with Sony Electronics, San Jose, California. He is a Senior Member of the IEEE and a member of the IEEE Signal Processing Society and the Society for Imaging Science and Technology (IS&T). His current research interests include color imaging, digital photography, and statistical signal processing.

Je-Ho Lee received his B.S. degree from the Hanyang University, Seoul, Korea, in 1990 and M.S. degree from the Pennsylvania State University, State College, in 1992, both in electrical engineering. From 1992 to 1997, he was with the Korea Institute of Science and Technology as a research staffer. He received his Ph.D. degree in electrical engineering from Purdue University in 2002. He is an imaging and color scientist with Hewlett-Packard Company, Vancouver, Washington. His research interests include digital halftoning, image enhancement, color image processing, color management, and image compression.

A. Ufuk Agar received the B.S. degrees in electrical engineering and mathematics from Bogazici University, Istanbul, Turkey, in 1992 and the M.S. and Ph.D. degrees in electrical and computer engineering from Purdue University in 1995 and 1999, respectively. She was a Fulbright scholar from 1993 to 1995. From 1999 to

2004, she has worked in the Color Imaging and Printing Technologies Department of Hewlett Packard Laboratories in Palo Alto, California, first as a research scientist and then as a senior research scientist. During the 2002–2003 academic year, she was an assistant professor in the Department of Electronics Engineering at Isik University in Istanbul, Turkey. She is currently with Garanti Technology, Istanbul, Turkey. Her research interests include color imaging, color halftoning, spectral modeling, colorimetric calibration and interpolation of multidimensional functions.

Jan P. Allebach received his B.S.E.E. from the University of Delaware in 1972 and his Ph.D. from Princeton University in 1976. He was with the University of Delaware from 1976 to 1983. Since 1983, he has been at Purdue University where he is Michael J. and Katherine R. Birck professor of electrical and computer engineering. His current research interests include image rendering, image quality, color imaging and color measurement, and digital publishing. He is a member of the IEEE Signal Processing Society, the Society for Imaging Science and Technology (IS&T), and SPIE. He has been especially active with the IEEE Signal Processing Society and IS&T. He is a Fellow of both these societies, has served as distinguished/visiting lecturer, an officer, and on the board of directors of both societies. He is a past associate editor for *IEEE Transactions on Signal Processing* and *IEEE Transactions on Image Processing*. He is editor for the IS&T/SPIE *Journal of Electronic Imaging*. He received the Senior (best paper) Award from the IEEE Signal Processing Society and the Bowman Award from IS&T. In 2004, he was named Electronic Imaging Scientist of the Year by IS&T and SPIE.

REFERENCES

[1] C. Haines, S. Wang, and K. Knox, "Digital color halftones," in *Digital Color Imaging Handbook*, G. Sharma, Ed. Boca Raton, FL: CRC, 2003, ch. 6, pp. 385–490.

[2] A.U. Agar, F.A. Baqai, and J.P. Allebach, "Human visual model-based color halftoning," in *Digital Color Imaging Handbook*, G. Sharma, Ed. Boca Raton, FL: CRC, 2003, ch. 7, pp. 491–557.

[3] T.N. Pappas, "Model-based halftoning of color images," *IEEE Trans. Image Processing*, vol. 6, no. 7, pp. 1014–1024, Jul. 1997.

[4] S. Wang, "Algorithm-independent color calibration for digital halftoning," in *Proc. 4th IS&T/SID Color Imaging Conf.*, 1996, pp. 75–79.

[5] X. Zhang, D.A. Silverstein, J.E. Farrell, and B.A. Wandell, "Color image fidelity metric s-CIELAB and its application on halftone texture visibility," in *IEEE COMPCON97 Symp. Dig.*, 1997, pp. 44–48.

[6] 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 Processing, Digital Display IV*, 1993, vol. 1913, pp. 270–281.

[7] R. Näsänen, "Visibility of halftone dot textures," *IEEE Trans. Syst., Man, Cybern.*, vol. 14, no. 6, pp. 920–924, 1984.

[8] B.W. Kolpatzik and C.A. Bouman, "Optimized error diffusion for image display," *J. Elect. Imaging*, vol. 1, no. 3, pp. 277–292, Jul. 1992.

[9] S.H. Kim and J.P. Allebach, "Impact of HVS models on model-based halftoning," *IEEE Trans. Image Processing*, vol. 11, no. 3, pp. 258–269, Mar. 2002.

[10] V. Monga, W.S. Geisler, and B.L. Evans, "Linear color-separable human visual system models for vector error diffusion," *IEEE Signal Processing Lett.*, vol. 10, no. 5, pp. 93–97, Apr. 2003.

[11] F.A. Baqai and J.P. Allebach, "Computer-aided design of clustered dot color screens based on a human visual system model," *Proc. IEEE*, vol. 90, no. 1, pp. 104–122, Jan. 2002.

[12] J.B. Mulligan, "Digital halftoning methods for selectively partitioning error into achromatic and chromatic channels," *Proc. SPIE*, vol. 1249, pp. 261–270, 1990.

[13] J.A.C. Yule, *Principles of Color Reproduction, Applied to Photomechanical Reproduction, Color Photography, and the Ink, Paper, and Other Related Industries*. New York: Wiley, 1967.

[14] R.A. Ulichney, "Dithering with blue noise," *Proc. IEEE*, vol. 76, no. 1, pp. 28–38, Jan. 1988.

[15] M. Wang and K.J. Parker, "Properties of jointly-blue noise masks and applications to color halftoning," *J. Imaging Sci. Technol.*, vol. 44, no. 4, pp. 360–370, July/Aug. 2000.

[16] J. Shu, C-H. Li, H. Ancin, and A. Bhattarcharjya, "Color stochastic screening with smoothness enhancement," in *Proc. IS&T's NIP 13: 1997 Int. Conf. Digital Printing Technologies*, 1997, pp. 522–525.

[17] Q. Lin and J.P. Allebach, "Color FM screen design using DBS algorithm," *Proc. SPIE*, vol. 3300, pp. 353–361, 1998.

[18] D.L. Lau, G.R. Arce, and N.C. Gallagher, "Green noise digital halftoning," *Proc. IEEE*, vol. 86, no. 12, pp. 2424–2442, Dec. 1998.

[19] D.L. Lau, A.M. Khan, and G.R. Arce, "Minimizing stochastic moiré by means of green noise masks," *J. Opt. Soc. Amer. A*, vol. 19, pp. 2203–2217, Nov. 2002.

[20] I. Amidror, "The moiré phenomenon in colour separation," in *Proc. 2nd Int. Conf. Raster Imaging Digital Typography*, 1991, pp. 98–119.

[21] G. Oster, M. Wasserman, and C. Zwerling, "Theoretical interpretation of moiré patterns," *J. Opt. Soc. Amer. A*, vol. 54, no. 1, pp. 169–175, 1964.

[22] I. Amidror, R.D. Hersch, and V. Ostromoukhov, "Spectral analysis and minimization of moire patterns in color separation,'' *J. Elect. Imag.*, vol. 3, no. 3, pp. 295–317, Jul. 1994.

[23] M. Kaji, Y. Satou, and J. Tajima, "A construction method of digital screen sets: realization of moiré-free rational tangent screens by using the multi-unit area design method," in *Proc. IS&T PICS*, 1998, pp. 349–357.

[24] M. Rodriguez, "Graphic arts perspective on digital halftoning," *Proc. SPIE* vol. 2179, pp. 144–149, 1994.

[25] J. Schoppmeyer, "Screen systems for multicolor printing," U.S. Patent 4 537 470, 1985.

[26] P.A. Delabastita, "Screening system and method for color reproduction in offset printing," U.S. Patent 5 155 599, 1992.

[27] S. Wang, Z. Fan, and Z. Wen, "Non-orthogonal screens and its application in moiré free halftoning," *Proc. SPIE,* vol. 5008, pp. 399–408, 2003.

[28] R.W. Floyd and L. Steinberg, "An adaptive algorithm for spatial grayscale," in *Proc. SID Int. Symp., Dig. Tech. Papers*, 1976, pp. 36–37.

[29] H. Haneishi, T. Suzuki, N. Shimoyama, and Y. Miyake, "Color digital halftoning taking colorimetric color reproduction into account," *J. Electron. Imaging*, vol. 5, no. 1, pp. 97–106, Jan. 1996.

[30] L. Akarun, Y. Yardimci, and A.E Cetin, "Adaptive methods for dithering color images," *IEEE Trans. Image Processing*, vol. IP-6, no. 7, pp. 950–956, July 1997.

[31] C.Y. Kim, I.S. Kweon, and Y.S. Seo, "Color and printer models for color halftoning," *J. Electron. Imaging*, vol. 6, no. 2, pp. 166–180, Apr. 1997.

[32] J. Sullivan, R. Miller, and G. Pios, "Image halftoning using a visual model in error diffusion," *J. Opt. Soc. Amer. A,* vol. 10, no. 8, pp. 1714–1724, 1993.

[33] J. Lee, Y. Kwon, and H. Kim, "A color halftoning algorithm for low-bit flat panel displays," in *Proc. IEEE Int. Conf. Image Processing*, Washington, DC, Oct. 1995, vol. II, pp. 346–349.

[34] N. Damera-Venkata and B.L. Evans, "Design and analysis of vector color error diffusion halftoning systems," *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1552–1565, Oct. 2001.

[35] R. Levien, "Well tempered screening technology," in *Proc. IS&T 3rd Tech. Symp. Prepress, Proofing and Printing*, Chicago, IL, Nov. 1993, pp. 98–101.

[36] G. Bozkurt, Y. Yardimci, O. Arikan, and E. Cetin, "Qr-rls algorithm for error diffusion of color images," in *Proc. IEEE Int. Conf. Image Processing*, Chicago, Oct. 1998, pp. 49–53.

[37] X. Zhang and B.A. Wandell, "A spatial extension of CIELAB for digital color image reproduction," *J. Soc. Inform. Display*, vol. 6, no. 7, pp. 61–63, 1997.

[38] T.N. Pappas and D.L. Neuhoff, "Printer models and error diffusion," *IEEE Trans. Image Processing*, vol. 4, no. 1, pp. 66–80, Jan. 1995.

[39] R.V. Klassen, R. Eschbach, and K. Bharat, "Vector diffusion in a distorted colour space," in *Proc. IS&T's 47th Annu. Conf.*, 1994, pp. 489–491.

[40] J. Shu, "Error diffusion with vivid color enhancement and noise reduction," *Proc. SPIE,* San Jose, CA, vol. 2657, pp. 464–470, 1996.

[41] D. Shaked, N. Arad, A. Fitzhugh, and I. Sobel, "Ink relocation for color halftones," in *Proc. IS&T Image Processing, Image Quality, Image Capture Syst. Conf.*, Portland, OR, 1998, pp. 340–343.

[42] R. Levien, "Practical issues in color inkjet halftoning," *Proc. SPIE* vol. 5008, pp. 537–541, 2003.

[43] A.U. Agar and J.P. Allebach, "Model-based color halftoning using direct binary search," in *Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts V*. Proc. SPIE 3963, pp. 521–535, 2000.

[44] J. Lee and J.P. Allebach, "Colorant-based direct binary search halftoning," *J. Electron. Imaging*, vol. 11, no. 4, pp. 517–527, Oct. 2002.

SP