Digital multitoning with overmodulation for smooth texture transition

Qing Yu
Kevin J. Parker
University of Rochester
Department of Electrical and Computer Engineering
Rochester, New York 14627

Kevin Spaulding
Rodney Miller
Eastman Kodak Company
Imaging Research and Advanced Development
Imaging Science Division
Rochester, New York 14650-1816

Abstract. *Multilevel halftoning (multitoning) is an extension of bitonal halftoning, in which the appearance of intermediate tones is created by the spatial modulation of more than two tones, i.e., black, white, and one or more shades of gray. In this paper, the conventional multitoning approach and a previously proposed approach, both using stochastic screen dithering, are investigated. A human visual model is employed to measure the perceived halftone error for both algorithms. The performance of each algorithm at gray levels near the printer's intermediate output levels is compared. Based on this study, a new overmodulation algorithm is proposed. The multitone output is mean preserving with respect to the input and the new algorithm requires little additional computation. It will be shown that, with this simple overmodulation scheme, we will be able to manipulate the dot patterns around the intermediate output levels to achieve desired halftone patterns. Implementation issues related to optimal output level selection and inkjet-printing simulation for this new scheme will also be reported. © 1999 SPIE and IS&T.*
[S1017-9909(99)00203-2]

1 Introduction

Recently we have seen a newer and expanding role of stochastic screen dithering[1–4] in digital printing because of its implementation simplicity and visually pleasing output. The implementation of screening employs a simple pointwise comparison, as shown in Fig. 1. An input image value is thresholded by a corresponding screen value to turn the output pixel on or off. Halftone patterns like those from stochastic screens contain ''blue noise''[5] characteristics, similar to those of error diffusion methods.[6] In the academic literature, the nature of the noise is often described by a color name; i.e., white noise is so named because of its flat power spectrum. Blue noise, on the other hand, has most of its energy located at high spatial frequencies with very little low-frequency content. Typical white-noise and

blue-noise radial-average power spectrums are shown in Fig. 2. Since the human visual system (HVS) is less sensitive to high-frequency content (blue noise), halftone patterns generated from stochastic screening are less visible to a human observer.

When devices have multilevel outputs, such as a multilevel inkjet printer, the stochastic screen technique can easily be generalized to utilize this new capability,[3] as shown in Fig. 3. It can be seen that this is equivalent to the binary implementation in Fig. 1, except that the screen is first scaled to an intermediate range before the comparison is taken, and the output is set to one of those output levels based on the comparison result and corresponding intermediate range. Assume an input image $i(x,y)$ has p different levels and the device has q possible output levels; also assume the screen $M(x,y)$ is 256 by 256 in dimension and has 256 levels. Then, the multitoning process can be given by the following equation:

$$o(x,y) = \text{INT}\left[\frac{i(x,y) + \dfrac{M(x_m,x_y)}{256}\left(\dfrac{p-1}{q-1}\right)}{\left(\dfrac{p-1}{q-1}\right)} \right], \qquad (1)$$

where $o(x,y)$ is the output value, $X_m = x \text{ MOD } 256$ and $Y_m = y \text{ MOD } 256$ (MOD stands for the modulo operation), and INT( ) indicates an integer truncation. This simple extension of this halftone process for devices with multiple output levels is defined as the conventional multitoning scheme.

Figure 4 illustrates the bilevel halftoned version of a gray ramp and the multitoned (four levels) version of the same ramp.

In this paper, the conventional multitoning approach and a previously proposed approach, both using stochastic
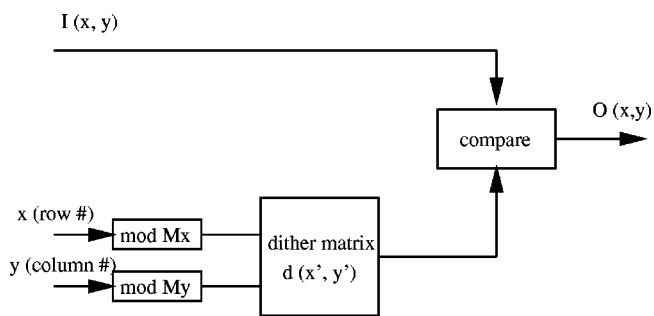
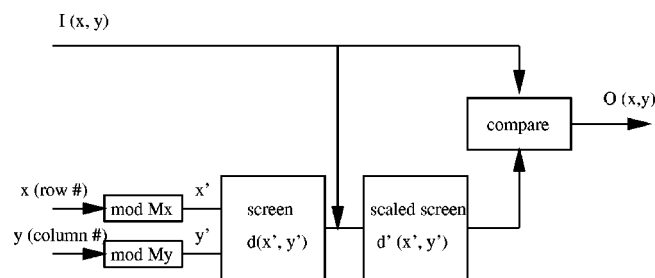Fig. 1 Bilevel halftoning with stochastic screen.

Fig. 3 Conventional multilevel halftoning with stochastic screen.

screen dithering, are investigated regarding texture variation across the tone scale. A human visual model is employed to measure the perceived halftone error for both algorithms. The performance of each algorithm at gray levels near the printer's intermediate output levels is compared. Based on this study, a new overmodulation algorithm is proposed that consists of a preprocessing stage where the pixels in the input image are modified, followed by conventional multitoning with specially design stochastic screens. It will be shown that, with this computationally efficient overmodulation scheme, we will be able to manipulate the dot patterns around the intermediate output levels to achieve desired halftone patterns. The special screen generation process will be outlined, and implementation issues related to optimal output level selection and inkjet-printing simulation for this new approach will also be reported.

2 Evaluation with a Human Visual Model

To evaluate the multitoning result, a HVS model could be utilized to study the frequency-weighted mean square error (FWMSE) between the original image and its multitoned version. Assume an N by N multitone pattern $b(i,j)$ for gray level g, and a point spread function $h(i,j)$ for the HVS system. The perceived error $e(i,j)$ due to the multitone process at pixel location (i,j) is expressed as

$$e(i,j) = b(i,j) * h(i,j) - g \qquad (2)$$

where ''*'' stands for circular convolution. Therefore, with Parseval's theorem, the FWMSE of the multitone pattern is given as

$$
\begin{aligned}
\text{FWMSE} &= \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |e^2(i,j)|}{N^2} \\
&= \frac{\sum_{f_i=0}^{N-1} \sum_{f_j=0}^{N-1} |B(f_i,f_j)|^2 |H(f_i,f_j)|^2}{N^2},
\end{aligned} \qquad (3)
$$

where $B(f_i,f_j)$ and $H(f_i,f_j)$ are discrete Fourier transform of $b(i,j)$ and $h(i,j)$. The final summation should exclude the (0, 0) point to account for the ''$-$'' term in Eq. (2).

A model of the low-contrast photopic modulation transfer function was used to characterize the human visual system:[7]

$$
H(f_{ij}) = \begin{cases} a(b+cf_{ij})\exp(-(cf_{ij})^d), & \text{if } f_{ij} > f_{\max} \\ 1.0, & \text{otherwise} \end{cases} \qquad (4)
$$

where f_{ij} is radial frequency with unit of cycles/degree, and the constants a,b,c,d take on values as 2.2, 0.192, 0.114, and 1.1, respectively. To apply this model for a specific viewing distance and resolution, a conversion from cycles/degree to cycles/inch is also carried out. Please refer to Sullivan and Miller[7] for the construction of $H(f_{ij})$ under
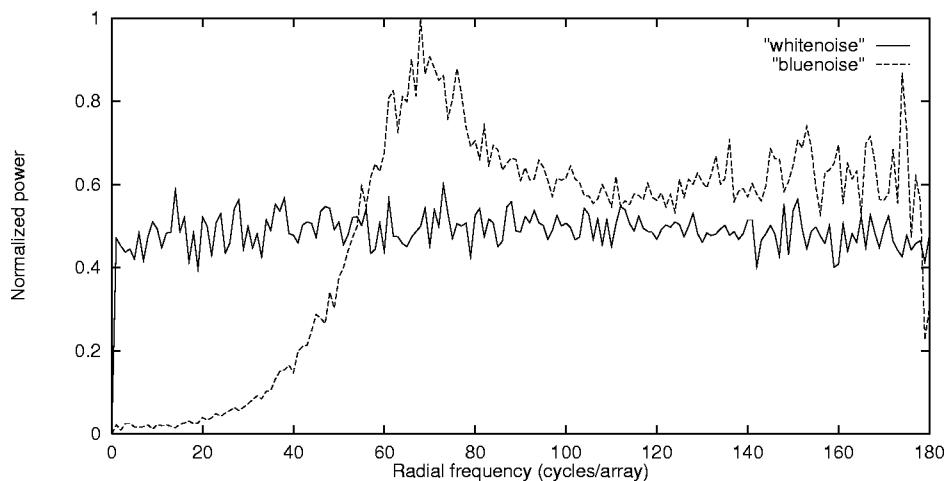

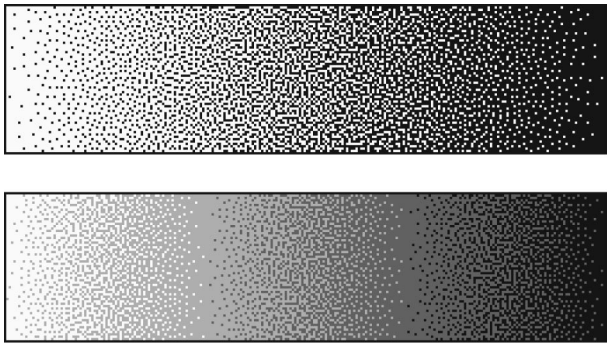Fig. 2 Typical white-noise and blue-noise radial-average power spectrum.

Fig. 4 Gray ramp: halftoned (above) and four-level multitoned (below).

different conditions, as well as the extension of $H(f_{ij})$ to a two-dimensional function $H(f_i, f_j)$ that is required in Eq. (3).

A plot of this visual model is shown in Fig. 5, which illustrates the lowpass nature of the visual system and the reduced sensitivity at 45 deg.

Uniform gray patches for each level are generated and then multitoned with the conventional scheme, and the FWMSE is calculated for each patch. Figure 6 shows the FWMSE versus gray level for both bilevel halftone (solid line) and four-level multitone (dashed line). In the latter case, we assume the device has four output states, say, gray level 0, 85, 170, and 255, respectively (0 is pure black while 255 is pure white). The stochastic screen used is 128 by 128 in dimension and has 256 levels. The viewing distance is set as 20 in. and resolution as 400 dpi (these two parameters are used throughout this paper).

A few observations are warranted. First, there is a great similarity between these two curves; each segment of the multitone curve could actually be perceived as a scaled version of the bilevel curve with compressed tone scale range. This agrees well with the implementation of conventional multitoning as a simple extension of the bilevel halftoning process. Second, for the multitone curve, around

each intermediate output level (85 and 170 in this case), there are distinct local minimum and overshoot of the FWMSE, which could also be found for the bilevel curve at both ends of the tone scale. The explanation for this is given as follows. Right at the output states, there is no halftone (multitone) error introduced, which leads to zero visual error. Plus or minus a few code values from these levels, the resulting halftone patterns have sparse minority dots over a uniform background, which contain significant low frequency content. This results in high visual error. In practice, screens are usually ''punched'' at both ends of the tone scale; for example, level 240 and beyond are set as white and level 15 and below as black. This works fine for bilevel halftoning, since most image details do not fall in those levels, and punch will also increase the contrast of image. For multilevel halftoning, however, this is not an option. It could be quite possible that a typical image region contains a smooth transition just around a certain intermediate level; then with the conventional multitoning scheme, there will be a distinct texture change in the output, which will be visible within certain viewing distances, as shown in Fig. 7. Another potential problem with the conventional scheme is that some devices, such as electrophotographic printers, do not produce uniform density regions very well. For these devices, having a uniform region in the output will actually degrade the image quality.

One thing we should keep in mind is that this conventional multitoning scheme is just a simple extension of the bilevel halftoning scheme; we have not fully taken advantage of having multiple output levels. A natural refinement of the conventional scheme would be to introduce dots of adjacent output states at gray levels near a typical intermediate output state so as to achieve a smoother texture transition.

3 Earlier Approach with Special Bitmap Design

An earlier multitoning scheme[8] proposed by Miller and Smith could be a solution for this problem. In this scheme, a modularly addressed screen is used to store pointers to a series of screen look-up tables (LUT), rather than storing
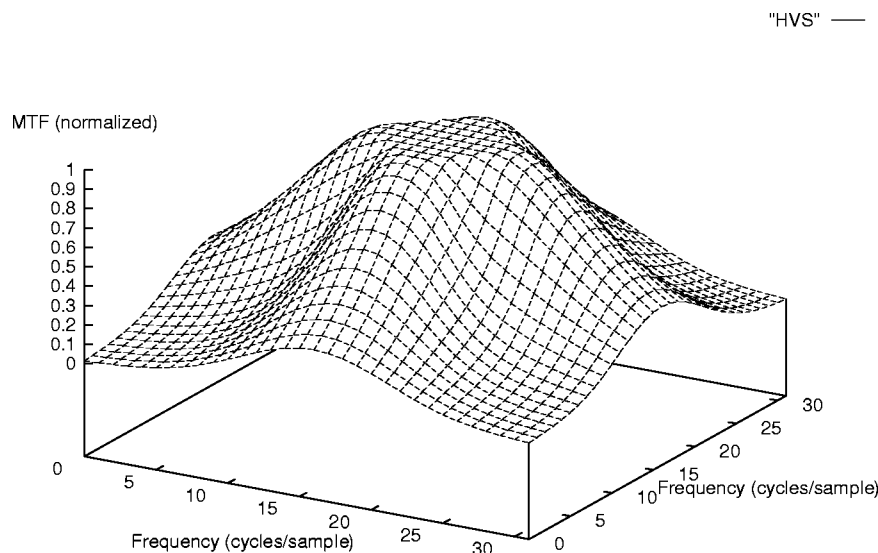

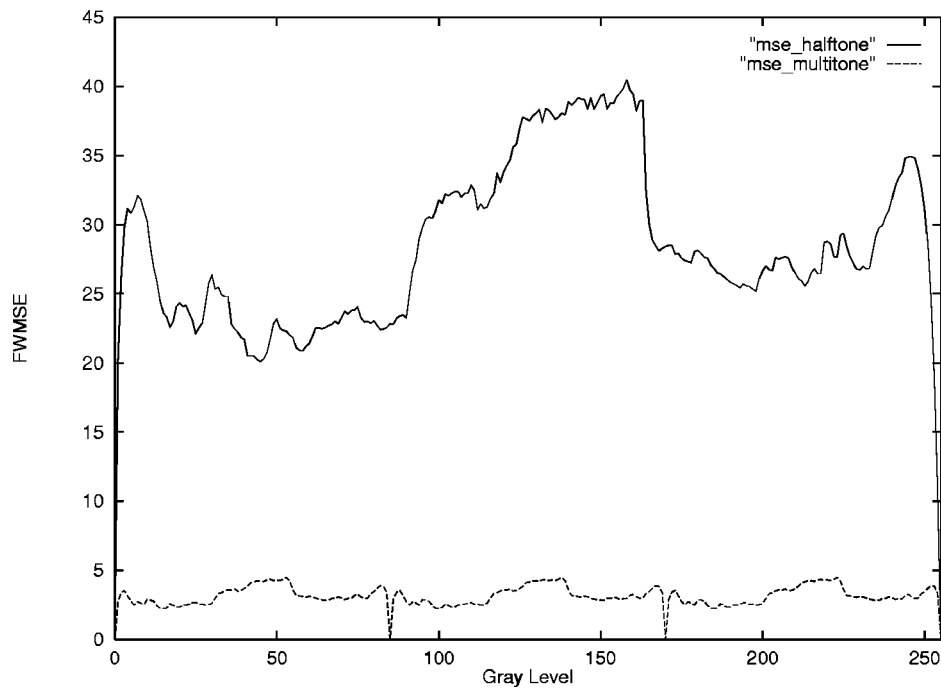**Fig. 5** HVS model used for evaluation.

Fig. 6 FWMSE for bilevel and multilevel halftoning.

actual screen values. The results of the screening process for each of the possible input levels are precalculated and stored in these LUTs. The algorithm can now be executed with only table look-ups rather than the adds and multiplies in the conventional scheme. Obviously, this trades off memory requirements for a faster execution.

One specific advantage of this LUT-based approach is that any conceivable dot growth can be specified for output patterns. With the conventional scheme, as the input gray level is increased, all of the pixels in the halftone pattern are generally increased to the second output level before increasing any of the pixels to the third output level. With the LUT approach, we gain the flexibility to increase the value of one pixel in the halftoned pattern through multiple output levels before starting to increase the value of a second pixel. Specifically, there is a texture parameter (T) in this algorithm that controls the design of LUTs with a variety of texture characteristics from a regular halftone screen. The value of T varies from 0 to 1. When T is set to 0, the LUT approach is equivalent to the conventional scheme that applies minimum amplitude modulation using nearest output levels, and when T is set to 1, this leads to maximum amplitude modulation. Please refer to Miller and Smith[8] for further details.
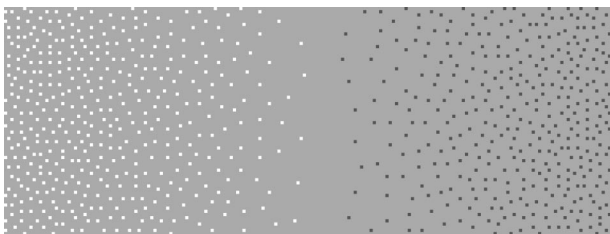

Fig. 7 Texture changes around output level 170.

This algorithm is tested with the same output-level setting (4) and the same screen (128 by 128 in size) used in previous evaluation, but with different texture parameters. The plots of FWMSE versus gray level for several texture parameter values are shown in Fig. 8. It can be seen that smoother texture transitions are achieved with increasing texture parameter value. However, for large texture values, the perceived errors are raised all over the tone scale. This tradeoff is not surprising since the LUT scheme was not intended and optimized for our typical study. What is found from this study is that a small texture value such as 0.0015 gives a good tradeoff. For this value, when the gray value is far from the intermediate levels, the conventional scheme (bilevel) is used to design corresponding LUTs so that only two levels of dots are involved; when the values are within a certain neighborhood of the intermediate output levels, dots of adjacent output states are gradually introduced.

4 A New Overmodulation Approach

In this section, a novel overmodulation scheme is proposed to handle the same problem with the same goal, that is, to achieve smooth transition at intermediate output states by introducing pixels of adjacent states. However, it is desirable to keep the conventional multitoning structure; thus a preprocessing step is introduced instead.

4.1 Overmodulation Scheme Overview

Figure 9 shows the flow chart of this new algorithm. A screen-guided modulation operation is added before the conventional multitoning process. For each input pixel I, we first decide if it is inside a predetermined gray scale range near the intermediate output levels. If not, this input pixel is directly sent to the conventional multitoning process; if the pixel is within a range R (i.e., $[X-R, X+R]$) near an intermediate output level X, a modulation function
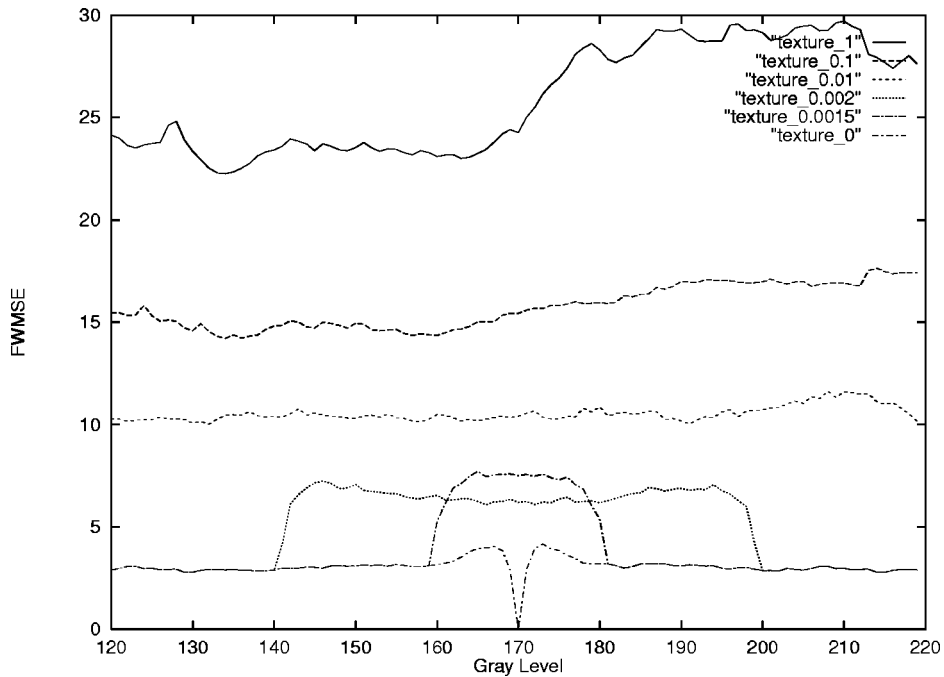
Fig. 8 FWMSE vs gray for different texture values.

is called, which requires as inputs the pixel value *I*, the specific level *X*, and corresponding screen value *S* for that pixel location. The modified input pixel value (I') is then used as the input for the conventional multitoning scheme to get the final output value *O*. The modulation process is nonlinear, and could be best described with the following pseudo computer code:

$D = I - X;$
$A = \mathrm{MAP}(D);$
if $(D >= 0)$
        if $(S >= 128)$
                $I' = X - A;$
        else
                $I' = I + A;$
else
        if $(S >= 128)$
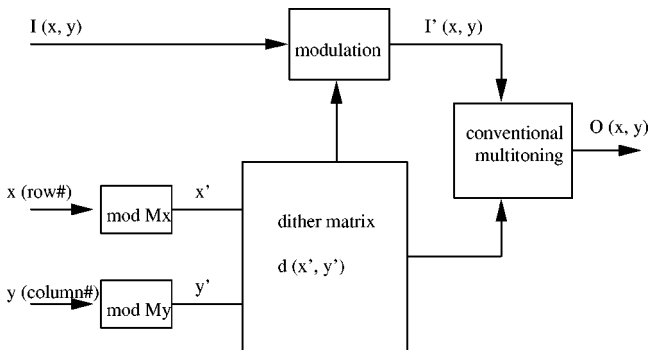 $I' = I - A;$
        else

$$I' = X + A;$$

where MAP( ) is a mapping function determining which pixels are to be modified during the preprocessing stage and also how much those pixels are modified. The MAP function should peak when the input pixel is right at the output level so as to reduce the difference of texture appearance between this typical level and rest of the tone scale; on the other hand, the MAP function should gradually decrease as input pixel values deviate from the output level to avoid adding additional texture to the adjacent gray levels significantly while ensuing a smooth transition. Both the peak of the MAP function and the value of *R* should be related to the total output levels of a typical system. For a four-output-level case, an empirically derived MAP( ) is shown in Fig. 10 with *R* set to 3.
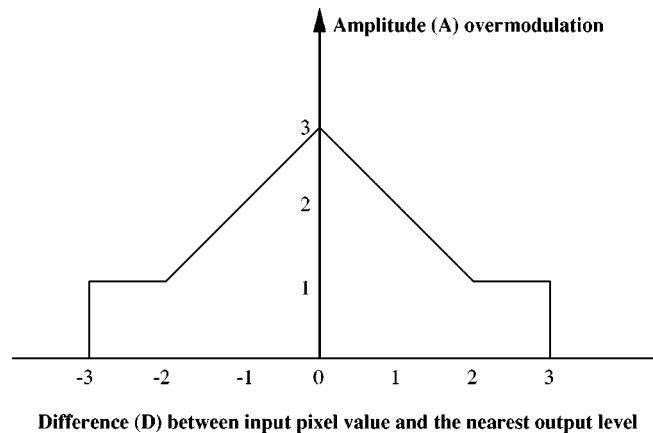

Fig. 9 Flow chart of overmodulation scheme.

Fig. 10 Modulation function.
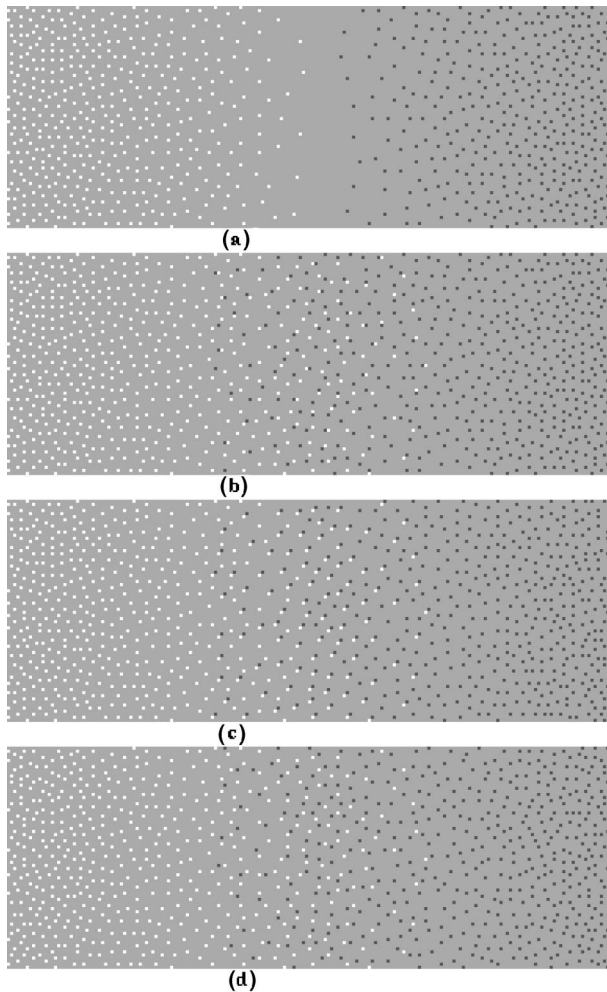
(a)

(b)

(c)

(d)

Fig. 11 Overmodulation with different dot patterns: (a) conventional scheme, (b) regular screen, (c) the binar screen, and (d) the "maximal-dispersed" screen.

Fig. 12 An image multitoned using a conventional screen with no overmodulation.

This overmodulation multitone approach is a meaning-preserving process, and an example is given in the Appendix for demonstration.

A gray ramp (around output level 170) multitoned using a regular stochastic screen with the overmodulation scheme is shown in Fig. 11(b). For visual comparison, the multitone output from conventional scheme is also shown in the same figure [Fig. 11(a)].

4.2 Overmodulation Dot Patterns Design

The overmodulation scheme introduces minority dots from adjacent output states to smooth the texture transition around intermediate output levels. These newly added dots belong to dot patterns at both ends of tone scale in a halftone screen. However, during the regular screen design process, the dot patterns at individual gray levels are normally built from dot patterns of adjacent levels (please refer to Refs. 1–4 for detailed descriptions of regular screen design procedures). Thus, the dot patterns at both ends are remotely correlated such that they are essentially independent of each other; therefore, the output from the overmodulation scheme normally has a noisy appearance, as shown in Fig. 11, when a regular screen is employed. Thus, if the

overmodulation scheme is considered during screen design, we should be able to add special correlation or special characteristics between those dot patterns to reduce the noisy appearance and achieve a smoother texture transition. Two methods have been proposed so far:

1. We enforce the condition that all the minority pixels from neighboring output levels are maximally dispersed from each other, as illustrated the gray ramp in Fig. 11(d). This can be achieved in the following way. Assume we know that the top 5% and bottom 5% dot patterns will be involved in the overmodulation process, which can be estimated from the setting of R and number of intermediate levels. In the case of an 8-bit screen that has 256 output levels, the top and bottom 5% of the patterns correspond to those of level 243 and up and those of level 13 and down. From a regular screen, we first identify locations where screen values are in the range of 242–229, and set these locations as forbidden ones. Then, starting from level 243, we construct another screen in the normal way, except that the pixels at those forbidden locations are reserved for dot patterns between level 0 and 13.

2. In the second method, we enforce the condition that all the minority pixels from neighboring output levels are paired, as illustrated the gray ramp in Fig. 11(c). We define these pairs as "binars." Obviously, we could have three arrangements for these binars; they could be lined up in vertical, horizontal, or diagonal directions. The screen design is very similar to the first method. With the same assumption we made for the first method, from a regular screen, we first identify those locations with screen value S in the range of 243–255, then shift the locations diagonally (assume diagonal binars) and set the screen value at that location as $255-S$. After this is done, we have built up the binar correlation between the top 5% and bottom 5% dot patterns. The rest of the dot patterns for the screen are designed in the normal way.

Figures 12 and 13 show a real image multitoned without overmodulation and with overmodulation using the binar

Fig. 13 An image multitoned using the binar screen with overmodulation.

screen. It can be seen that the texture transition in the sky region is smoother for the binar image.

4.3 Deficiency of FWMSE and Introducing a New Halftone/Multitone Texture Visibility Metric

Figure 14 shows the curves of FWMSE versus gray level for multitoning outputs from the three different screens (regular, maximally dispersed, and binars). One observation is that, if FWMSE is used as the texture visibility metric, then the output from the ''maximally dispersed'' screen has the lowest texture visibility around the intermediate level (level 170) when the proposed overmodulation approach is employed.

A simple psychophysical experiment was further carried out to evaluate the actual perception of multitone texture by HVS. In this experiment, multitone patches from the three screens were displayed side by side on the wall, and subjects were asked to slowly walk away from these patches until the multitone texture around the intermediate levels is invisible for one of the patches. Both real images and gray ramps were used in the experiment, and subjects had no prior knowledge which screen technique was used on each patch or image. The general observation from the experiment is that textures in the multitone output from the binar screen are dissolved more quickly when viewing distance is increased. However, this does not agree with the rankings in Fig. 14.

Both Yao[9] and Wong[10] had shown that FWMSE-type metrics fail in some cases when predefined Gaussian-type filters are used to approximate the HVS sensitivity. A new metric for color halftone texture visibility has been proposed recently that correlates well with a series of psychophysical experiments. Readers are referred to Ref. 11 for a detailed description. In brief, this metric combines a progressively adjusted ideal lowpass filtering with a detection model based on Weber's law,[12] and it searches for the minimum value of F_{cf} such that

$$\frac{SD(Y(x,y)*LPF(x,y|F_{cf}))}{Y} \geqslant \text{Webfactor}, \tag{5}$$

where ''*'' stands for the circular convolution, $Y(x,y)$ is the luminance image of the color halftone pattern, Y is the average value of $Y(x,y)$, SD( ) is the standard deviation function, LPF( ) is an ideal lowpass filter with cutoff radial frequency F_{cf}, and Webfactor is the famous Weber constant (typical value of 1/30).[12] When F_{cf} is used as the metric value, the higher the F_{cf}, the lower the texture visibility for the color halftone pattern.

The same evaluation process as in Sec. 2 is carried out with the FWMSE metric replaced by the new texture vis-
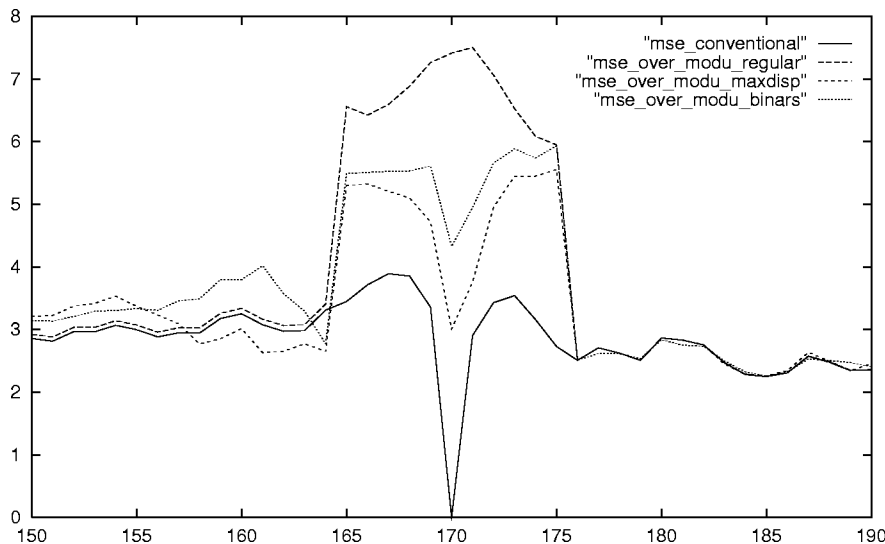

Fig. 14 FWMSE vs gray level for different overmodulation screens (400 dpi and 20-in. viewing distance).
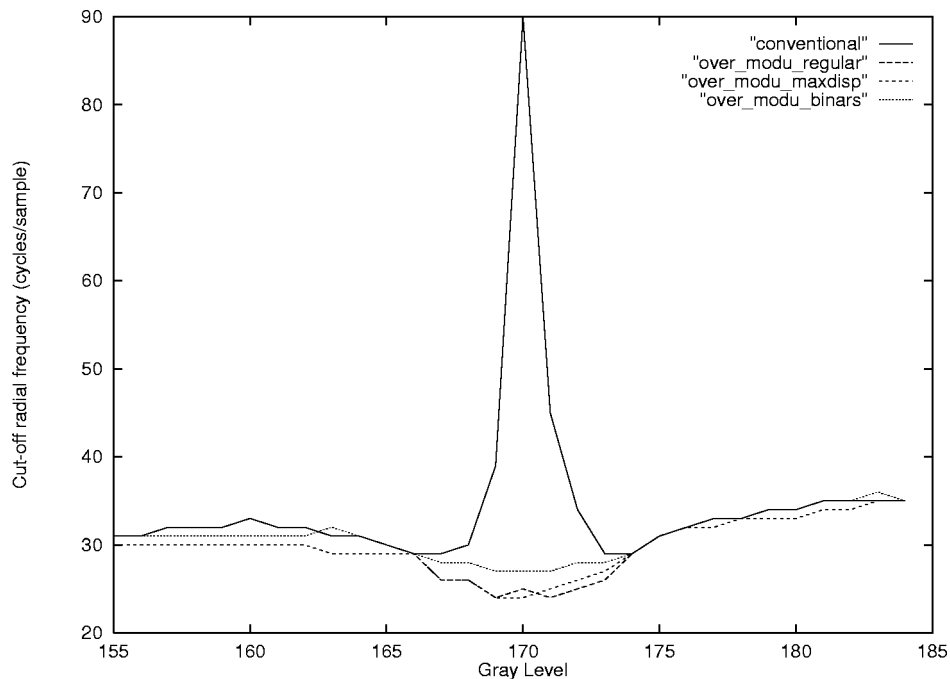
Fig. 15 Cutoff frequencies vs gray level for different overmodulation screens.

ibility metric where the cutoff frequency is identified for each multitoned patch of constant gray. The results are shown in Fig. 15. (Note: When no overmodulation is applied, the multitone output for level 170 is a constant gray patch; thus the cutoff frequency for this patch should be infinite. For easy demonstration, a cutoff frequency of 91, the largest realizable radial frequency in this evaluation, is assigned to that patch instead.)

It can be seen that the multitone output from the binar screen has the highest cutoff frequencies within the overmodulation range (level 165–175), thus the lowest texture visibility. The transition of texture visibility around the intermediate level (level 170) is also the smoothest for the binar case. These predications agree well with the experimental observations. Therefore, the new metric for color halftone texture visibility is applicable to multitone texture evaluation.

5 Implementation Issues

In this section, a couple of implementation issues that are critical to applying the overmodulation scheme to multilevel printing are further explored.

5.1 Optimal Output Levels Setting

For multilevel halftoning, one would expect that equal spacing of the halftone levels in lightness space ($L*$ in CIELAB), rather than in reflectance space (or some RGB code value space), will result in overall lower visibility of the halftone patterns and therefore higher level of image quality. However, by evaluating multilevel prints from the thermal printer used in this study, we found that the visibility of the resulting halftone patterns was not uniform with density when the output levels were equally spaced in $L*$.

Instead the halftone patterns were imperceptible for most of the density range, but were easily visible at the lower density (higher lightness) values.

There are two possible explanations for this result. First, it is possible that the density levels that were reproduced in the halftone patterns were different than the expected values. This could be due to differences between the density produced for solid areas and that produced for isolated pixels from halftoning process. The second possibility is that the $L*$ function is not a good measure for the perceived lightness difference between the output levels of a halftone pattern. The $L*$ function was developed by CIE using large area constant patches to estimate the perceived lightness. However, the current result suggests that perceived lightness differences may be a function of the spatial frequency content of the signal as well. In particular, the effective lightness difference for light signals is apparently larger than that for dark signals. Based on these assumptions, a new effective perceived lightness function can be derived as follows.

Define a new perceived lightness difference term ΔL^*_{new} as a simple function of ΔL^* in CIELAB space:

$$\Delta L^*_{\text{new}} = f(L^*, \Delta L^*). \tag{6}$$

In a simple form of this function, the ratio $\Delta L^*/\Delta L^*_{\text{new}}$ varies linearly with lightness. Assuming the value of this ratio $\Delta L^*/\Delta L^*_{\text{new}}$ at the dark end of the lightness scale is L, and the value of this ratio at the light end of the lightness scale is H, Eq. (6) is then given by

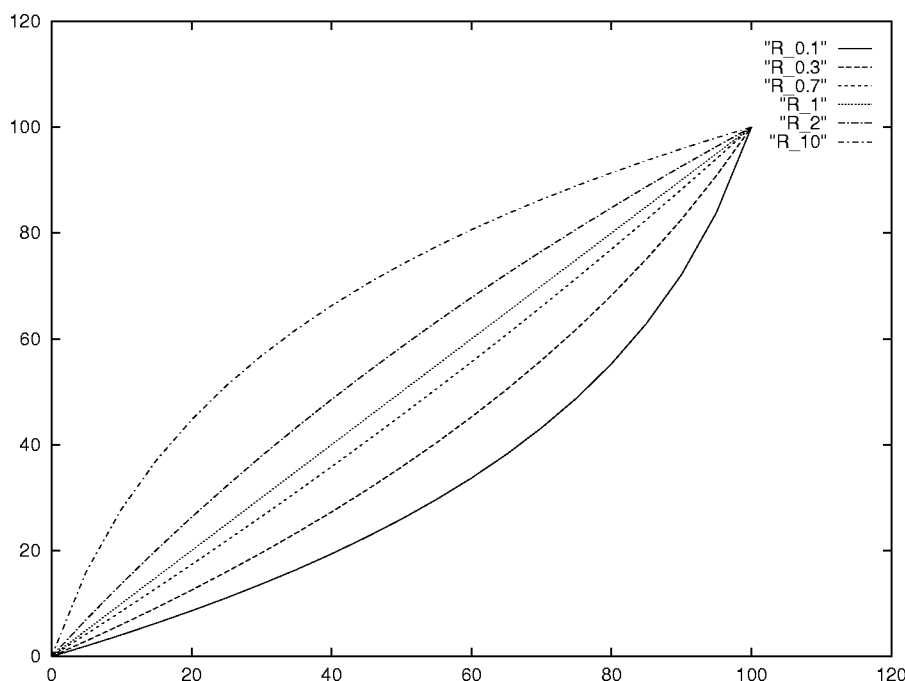$$\frac{\Delta L^*}{\Delta L^*_{\text{new}}} = a + bL^*, \tag{7}$$

Fig. 16 Modified lightness values vs original lightness values for various H/L values.

where $a = L$ and $b = (H - L)/100$.

In order to derive a new expression for the perceived lightness function, we let the lightness difference become very small in Eq. (7); then the following differential equation can be obtained:

$$\frac{dL^*}{dL^*_{new}} = a + bL^*.$$ (8)

Solving this equation for L^*_{new}, we get

$$L^*_{new} = \frac{\ln(a + bL^*)}{b} + c,$$ (9a)

where c is a constant. If we apply the constraint that $L^*_{new} = 0$ when $L^* = 0$, we get

$$L^*_{new} = \frac{\ln(a + bL^*) - \ln a}{b}$$

$$= \frac{\ln\left(1 + \frac{b}{a}L^*\right)}{b} = \frac{\ln\left(1 + \frac{(H-L)}{100L}L^*\right)}{\frac{H-L}{100}}.$$ (9b)

Scaling the result so that $L^*_{new} = 100$ when $L^* = 100$, we have the following relationship:

$$\ln\frac{H}{L} = H - L.$$ (9c)

Plug this relationship back to Eq. (9b) and define $R = H/L$, we get the following equation for the new perceived lightness function:

$$L^*_{new} = 100 \frac{\ln\left[\frac{L^*}{100}(R-1)+1\right]}{\ln R}.$$ (9d)

R actually represents the relative weight of lightness differences for dark tones to those for light tones. Notice that

$$R = \frac{H}{L} = \frac{\left(\frac{\Delta L^*}{\Delta L^*_{new}}\right)_{at\_light\_end}}{\left(\frac{\Delta L^*}{\Delta L^*_{new}}\right)_{at\_dark\_end}}.$$ (9e)

As R approaches 1.0, L^*_{new} approaches L^*. For value of R smaller than 1.0, the lightness difference in the light end of the tone scale will get weighted more heavily. Likewise, for values of R larger than 1.0, lightness at the dark end of the tone scale will get weighted more heavily. The fact that the multitoned patterns were more visible in the light end of the tone scale suggests that a value of R that is less than 1.0 will be appropriate. Figure 16 illustrates the mapping from the original lightness space to the modified lightness space for various R values.

For the overmodulation scheme, a series of multitoned gray ramps was created using a series of values of R (the number of output levels were set to 4). Based on our visual assessment, an R value of 0.35 seemed to be a good compromise, giving reasonably equal visibility of the halftone patterns across the entire tone scale.

5.2 *Inkjet Multilevel Printing Simulation*

To this point, all the simulations have been done on a thermal continuous tone printer; however, the overmodulation scheme is initially intended for a wide range of printers. Therefore, experiments with different printing engines are worthwhile. The one with the inkjet is reported as follows.

Inkjet output devices have traditionally been binary engines that rely on halftone techniques to render continuous tone input images. Recent technological advances in the design and manufacture of inkjet printheads have paved the way for the development of multilevel inkjet devices. The appearance of multiple output density levels is achieved by modulating the volume of ink that leaves the nozzle, which translates into a modulation of the area of the ink dot on the printed page. Recent studies have shown that the overall image quality of a particular rendering device or algorithm increases more rapidly with bit depth than resolution.[13] Therefore, inkjet systems with limited resolution may be able to offer high image quality through the addition of multiple output levels. Since we are unable to drive currently available inkjet printers with multilevel output capability, a high-resolution laser printer was used to simulate a multilevel inkjet printer model.

The printer for the Kodak Approval digital color proofing system is a digital printer that was designed and built as a color proofer for graphic arts application. The printer creates images on an intermediate media via laser thermal dye sublimation, and the final image is then created by laminating the intermediate image onto desired media. With a maximum addressable printing resolution of 1800 dpi, the printer is well suited as an output device for generating simulations of multilevel inkjet printers. For example, an image to be printed at 300 dpi can be simulated with a 1800 dpi bitmap printed on the approval printer so that each pixel in the original image can be modeled by a 6 by 6 square of pixels. This capability allows us to simulate various phenomena in the inkjet printing process on a pixel by pixel base. Currently, we are able to predict characteristics that are related to ink–media interactions such as tone and color reproduction, as well as to model typical artifacts such as dot placement accuracy variation, dot size variation, dot density variation, and duty cycle variation that took place in a real printing environment.

For this study, simulations were carried out at printer resolutions of 75, 150, and 300 dpi for the conventional multitone scheme as well as the two proposed overmodulation schemes (maximally dispersed and binars). Both gray ramps and real images were used. Subjectively, we found that the binar scheme had the best visual quality, which agrees with the results from thermal printing. We also found that by setting the output states in the modified lightness space (as discussed above), the resulting halftone patterns were most uniform with code value, although a different R value had to be used in this case because of the different maximum-density of the printer along with the dot-overlapping characteristics of the inkjet model.

6 Future Research

There are several parameters that are very important for this overmodulations scheme, such as neighborhood range

R and modulation function MAP( ). Testing on different values for R and various mapping functions are worthwhile to find the optimal settings.

This overmodulation scheme could also be extended to multilevel color rendering, where different color dots will interact with each other around certain boundaries.

Another area worth looking into is the visibility of multilevel halftoning patterns as a function of the chosen intermediate gray levels. Currently, some *ad hoc* adjustment of the levels has to be made in order to approximately equalize the visibility in the dark and light regions. A more scientific study of this phenomenon should be carried out to develop a model that would predict the visibility of the patterns as a function of the gray levels and the frequency content of the halftone pattern. This would be very useful for the design of multilevel output devices such as inkjet printers. In particular, it could be used to select the gray levels that would minimize the pattern visibility.

7 Conclusion

In this paper, a novel overmodulation scheme is presented to improve multilevel rendering around intermediate output levels with stochastic screening. Investigations on screen design, multitone texture evaluation, as well as implementation issues are also reported.

Appendix

In the following, we will show that this overmodulation is a mean-preserving process. Assume output levels are uniformly distributed over the tone scale (i.e., levels 0, L, $2L$, $3L$, and so on up to 255). If an input pixel I falls in the range between L and $2L$, and it is within the neighborhood of $2L$, then with the conventional scheme, the expected value $E\{O\}$ for output O is given by

$$E\{O\} = P_{2L}^* 2L + P_L^* L. \tag{A1}$$

The notation P_X represents the probability of the output value of X. Since there is no prior screen value information, it is easy to see that $P_{2L} = (I-L)/L$ and $P_L = (2L-I)/L$; thus, $E\{O\} = I$.

With the overmodulation scheme, $D = I - 2L$, where $D < 0$; then $A = \text{MAP}(D)$. If the screen value $S >= 128$, since $D < 0$, then according to the pseudo computer code in last section, the modified input value $I' = I - A$. In this case, the expected value $E_1\{O\}$ for output O is given by

$$E_1\{O\} = P_{2L}^* 2L + P_L^* L = 2I - 2A - 2L \tag{A2}$$

since $P_{2L} = ((I-A) - 3/2L)/(L/2)$ and $P_L = (2L - (I-A))/(L/2)$.

On the other hand, if $S < 128$, then $I' = 2L + A$; thus, the expected value $E_2\{O\}$ for O is given by

$$E_2\{O\} = P_{3L}^* 3L + P_{2L}^* 2L = 2L + 2A \qquad (A3)$$

since $P_{3L} = ((2L + A) - 2L)/(L/2)$ and $P_{2L} = (5/2L - (2L + A))/(L/2)$. Since there is a 50% chance for each screen value to go beyond and below 128, the overall expected value $E\{O\}$ for output O is

$$E\{O\} = 0.5^* E_1\{O\} + 0.5^* E_2\{O\} = I. \qquad (A4)$$

This shows that the overmodulation process is a mean-preserving process.

Note that only half of the intermediate range $L/2$ is used to calculate the probability in the overmodulation case; this is due to the fact that we have prior knowledge of the screen value at each pixel location from the overmodulation procedure (either $S >= 128$ or $S < 128$).

References

1. T. Mitsa and K. J. Parker, ''Digital halftoning using a blue noise mask,'' in *ICASSP 91: 1991 International Conference on Acoustics, Speech, and Signal Processing*, Toronto, Canada, Vol. 2, pp. 2809–2812, IEEE, New York (May 1991).
2. T. Mitsa and K. J. Parker, ''Digital halftoning using a blue noise mask,'' *J. Opt. Soc. Am. A* **9**, 1920–1929 (1992).
3. K. E. Spaulding, R. Miller, and J. Schildkraut, ''Methods for generating blue-noise dither matrices for digital halftoning,'' *J. Electron. Imaging* **6**, 208–230 (1997).
4. Q. Yu and K. J. Parker, ''Stochastic screen halftoning for electronic devices,'' *J. Visual Commun. Image Represent.* **8**, 423–440 (1997).
5. R. Ulichney, *Digital Halftoning*, MIT Press, Cambridge, MA (1987).
6. R. W. Floyd and L. Steinberg, ''An adaptive algorithm for spatial greyscale,'' *Proc. Soc. Inf. Display* **17**(2), 75–77 (1976).
7. J. Sullivan, L. Ray, and R. Miller, ''Design of minimal visual modulation halftone patterns,'' *IEEE Trans. Syst. Man Cybern.* **21**, 33–38 (1991).
8. R. Miller and C. Smith, ''Mean-preserving multilevel halftoning algorithm,'' in *Human Vision, Visual Processing, and Digital Display IV*, San Jose, CA, *Proc. SPIE* **1913**, 367–377 (Feb. 1993).
9. M. Yao, ''Blue Noise Halftoning,'' PhD thesis, University of Rochester, 1996.
10. P. W. Wong, ''A mixture distortion criterion for halftones,'' in *Proceedings, Optics & Imaging in the Information Age*, Rochester, NY, pp. 187–191, IS&T/OSA (Oct. 1996).
11. Q. Yu and K. J. Parker, R. Buckley, and V. Klassen, ''A new metric for color halftone visibility,'' in *Proceedings, IS&T's 51st Annual Conference*, Portland, OR, pp. 226–230 (May 1998).
12. G. Gescheider, *Psychophysics Method, Theory, and Application*, Erlbaum, Hillsdale, NJ, 1985.
13. J. E. Farrell, ''Greyscale and resolution tradeoffs in photographic image quality,'' in *Human Vision and Electronic Imaging II*, San Jose, CA, *Proc. SPIE* **3016**, 148–153 (Feb. 1997).

Kevin J. Parker received his BS degree in engineering science, *summa cum laude*, from SUNY at Buffalo in 1976. Graduate work in electrical engineering was done at MIT, with MS and PhD degrees received in 1978 and 1981. From 1981 to 1985 he was an assistant professor of electrical engineering and radiology. Dr. Parker has received awards from the National Institute of General Medical Science (1979), the Lilly Teaching Endowment (1982), the IBM Supercomputing Competition (1989), and the World Federation of Ultrasound in Medicine and Biology (1991). He is a member of the IEEE Sonics and Ultrasonics Symposium Technical Committee and serves as a reviewer and consultant for a number of journals and institutions. He is also a member of the IEEE, the Acoustical Society of America, and the American Institute of Ultrasound in Medicine. He has been named a fellow in both the IEEE and AIUM for his work in medical imaging. In addition, he was recently named to the Board of Governors of the AIUM. Dr. Parker's research interests are in medical imaging, linear and nonlinear acoustics, and digital halftoning.

Kevin Spaulding received his BS in imaging science from the Rochester Institute of Technology in 1983, and MS and PhD degrees in optical engineering from the University of Rochester in 1988 and 1992, respectively. He has been with Eastman Kodak Company since 1983 where he is currently a research associate in the Imaging Science Technology Lab. His research interests include color reproduction, digital halftoning, and image processing algorithms for digital camera and printers, and image quality metrics.

Rodney Miller received his BS degree in electrical engineering from the Rose–Hulman Institute of Technology, Terre Haute, Indiana, in 1981. He received his MS degree in electrical engineering from the Rochester Institute of Technology in 1994. He has been employed by Eastman Kodak Company since 1981. He began in the product development group for Business Imaging Systems where he applied image enhancement and halftoning techniques to micrographics imagery. Since then he has joined the research labs as group leader for the I/O Processing Group in the Imaging Science Technology Lab. His current interests continue to include halftone algorithms, color space transformations for accurate color reproduction, and systems analysis.

Qing Yu received his BS degree in physics from the University of Houston in 1994, and his MS and PhD degrees in electrical engineering from the University of Rochester in 1996 and 1998, respectively. He is currently a senior research scientist at Eastman Kodak Company. His research interests include digital printing, color image processing, and image quality. Dr. Yu is a member of IEEE and IS&T.