A Mixture Distortion Criterion for Halftones

Ping Wah Wong
Computer Peripherals Laboratory
Hewlett Packard Laboratories, Palo Alto

Abstract

Frequency weighted mean squared error (FWMSE) is often used for measuring image quality, in generating halftones, and in considerations of lossy compression of halftones. We construct examples to show the weakness of FWMSE when applied to halftones. Specifically, we show that FWMSE does not directly reflect the uniformity in the spatial distribution of halftone dots. We then propose and examine a mixture distortion criterion that consists of FWMSE and a dot distance term to explicitly account for the spatial arrangement of dots.

1 Introduction

The goal of halftoning is to generate bi-level images from continuous tone images so that they appear similar to the human visual system. To give an indication of the quality of halftones, one often uses a distortion criterion $d(x_{m,n}, h_{m,n})$ to measure the differences between the halftone $h_{m,n}$ and its continuous tone counterpart $x_{m,n}$ [1]. A distortion measure is essential for optimization based halftoning algorithms, [2–8], where one finds a halftone $h_{m,n}$ from a continuous tone image $x_{m,n}$ so that the average distortion $d(x_{m,n}, h_{m,n})$ is minimized. In lossy compression of halftones [9] as well as in designing entropy constrained halftoners [4,8], a distortion measure is essential in the consideration of trade-off between output quality and bit rate.

Frequency weighted mean squared error (FWMSE) [1, 10, 11] is perhaps the most popular distortion criterion that is used in practice, partly because of its simplicity and tractability. The recently proposed S-CIELAB [12, 13], an extension of the color metric CIELAB, is particularly useful for measuring the fidelity of color halftones. In S-CIELAB, the input color images are converted into the LAB color space, and then processed by a set of linear filters in the LAB space. Finally, a mean squared error (Euclidean distance in LAB) is calculated using the filtered components to give the distortion. Hence, S-CIELAB is a form of FWMSE calculated in the CIELAB color space.

Let the pixel values of a size N by N continuous tone image $x_{m,n}$ to be real numbers between 0 (black) and 1 (white), and the bi-level halftone $b_{m,n}$ to take on values in $\{0, 1\}$. Let the instantaneous frequency weighted squared error at pixel location (m, n) be

$$w_{m,n} = \Big(x_{m,n} - \sum_{k,l} v_{k,l} b_{m-k,n-l}\Big)^2 \tag{1}$$

where $v_{k,l}$ is an impulse response that approximates the characteristics of the human visual system. The FWMSE is given by

$$\mathcal{W}(x, b) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} w_{m,n} \tag{2}$$

where the sum is taken over all the pixels in the image. Since $b_{m,n}$ is a halftone rendition of $x_{m,n}$, the two images satisfy

$$\frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x_{m,n} = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} b_{m,n}, \tag{3}$$

i.e., the halftone pattern has the same average graylevel as the continuous tone image. This implies the DC components of $x_{m,n}$ and $b_{m,n}$ are identical, or equivalently $X_{0,0} = B_{0,0}$ where $X_{k,l}$ and $B_{k,l}$ are the discrete Fourier transforms (DFT) of $x_{m,n}$ and $b_{m,n}$, respectively. Then

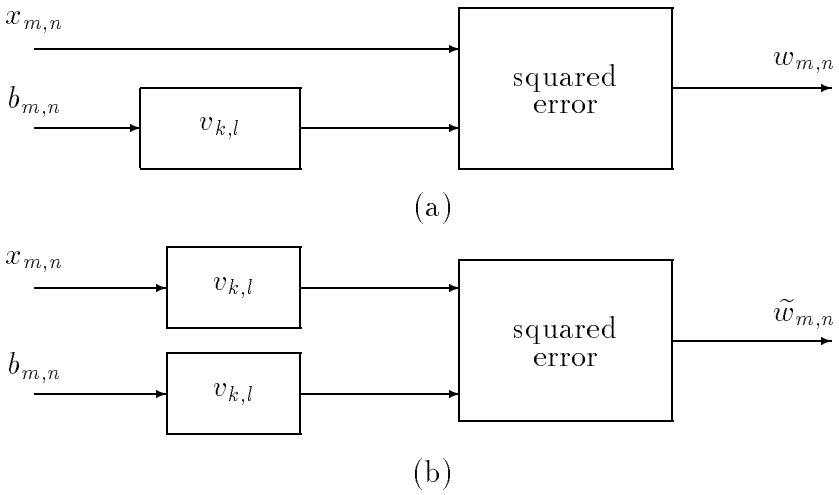

Fig. 1. Two different forms of frequency weighted mean squared error.

we can apply Parserval theorem on (2) to give

$$\mathcal{W}(x,b) = \frac{1}{N^2} \sum_{(k,l)\neq(0,0)} |V_{k,l}|^2 |B_{k,l}|^2. \tag{4}$$

This expression displays an explicit notion of weighting the halftone spectrum by the frequency response of a filter.

The operation of (1) can be represented by the block diagram in Fig. 1 (a). It makes good intuitive sense as it suggests that we measure the difference between an original continuous tone image and its corresponding halftone image as the halftone is perceived by the human visual system. Another form of FWMSE that is also frequently used in the literature is obtained by replacing $w_{m,n}$ with

$$\tilde{w}_{m,n} = \Big(\sum_{k,l} v_{k,l}(x_{m-k,n-l} - b_{m-k,n-l})\Big)^2, \tag{5}$$

which can be represented by Fig. 1 (b). In this form, both $x_{m,n}$ and $b_{m,n}$ are low pass filtered by $v_{k,l}$. Both (1) and (5) are used in the literature, and have been shown to produce good results in halftoning. For the rest of this paper, we will use the form given in (1). We note that similar results and conclusions in the paper can be applied to both forms with suitable modification.

For a halftone to be perceived as high quality, it is essential that the spatial distribution of halftone dots in smooth areas to be as uniformly distributed as possible. This is consistent with the *blue noise* (high frequency noise) characteristic [14], meaning that the error spectra between continuous tone and halftone images should preferably be concentrated in the high

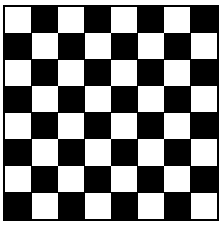

Fig. 2. A checker-board halftone rendition for a constant gray patch at mid-tone, i.e., $x_{m,n} = 0.5$. Using a visual filter with monotonically decreasing magnitude frequency response from low to high frequency, this pattern results in the smallest possible FWMSE among all possible halftone renditions of mid-tone.

frequency range. FWMSE, however, does not explicitly address the spatial distribution of halftone dots.

In this paper we examine FWMSE in detail, and give examples to show that a low FWMSE is not always consistent with a smooth spatial distribution of halftone dots. The idea of considering the distances between halftone dots has been successfully used in error diffusion to generate high quality results [15]. Using this idea, we propose a new distortion criterion that explicitly takes the spatial uniformity of halftone dots into account. Such a distortion criterion has been used in conjunction with a tree coding algorithm to generate halftones of very high quality [7, 8].

2 Pattern Uniformity and Frequency Weighted Mean Squared Error

Digital halftoning, by its nature, relies on the spreading of black and white pixels to give a perception of gray levels. For high visual quality, one prefers the spatial distribution of black and white pixels to be as "uniform" as possible, since uniformly spaced dots generally gives visually smooth renditions of graylevels. This is consistent with designing halftones that have a *blue noise* (high frequency noise) characteristic [14], meaning that the energy in the error spectra between continuous tone and halftone images should preferably be concentrated at the high frequency region.

It is evident from (1) that the spatial distribution of the black and white pixels is not *explicitly* reflected by the FWMSE. In other words, a halftone where the black and white dots are spatially distributed "more uniformly" can incur a larger FWMSE than a more irregularly distributed dot array. In the following, we will consider four examples that illustrates the spatial distribution of halftone dots may or may not be reflected by the FWMSE. Example 1 illustrates a situation where FWMSE is appropriate, while Examples 2 to 4 demonstrates situations where FWMSE is inappropriate.

Example 1: Consider a constant gray patch at mid-tone with $x_{m,n} = 0.5$, and a checker-board halftone rendition $c_{m,n}$ as shown in Fig. 2. The two-dimensional DFT (size 8 by 8) of $c_{m,n}$ has only two non-zero frequency components

$$|C_{0,0}| = |C_{4,4}| = 32.$$

For a size 8 by 8 DFT, $C_{4,4}$ is the frequency component at the highest possible frequency. We have from (4) that

$$\mathcal{W}(x, c) = 16|V_{4,4}|^2.$$

Consider any other halftone rendition $d_{m,n}$ for the same gray patch $x_{m,n} = 0.5$. Since both $c_{m,n}$ and $d_{m,n}$ have the same number of black and white pixels, we have from (3) and Parserval theorem that

$$\sum_{(k,l)\neq(0,0)} |D_{k,l}|^2 = C_{4,4}.$$

Hence

$$\begin{aligned}
\mathcal{W}(x, d) - \mathcal{W}(x, c) &= \frac{1}{N^2} \sum_{(k,l)\neq(0,0)} |V_{k,l}|^2 |D_{k,l}|^2 - \frac{1}{N^2} |V_{4,4}|^2 |C_{4,4}|^2 \\
&= \frac{1}{N^2} \sum_{(k,l)\neq(0,0)} |V_{k,l}|^2 |D_{k,l}|^2 - \frac{1}{N^2} |V_{4,4}|^2 \sum_{(k,l)\neq(0,0)} |D_{k,l}|^2 \\
&= \frac{1}{N^2} \sum_{(k,l)\neq(0,0)} (|V_{k,l}|^2 - |V_{4,4}|^2) |D_{k,l}|^2.
\end{aligned}$$

For any visual (low pass) filter with a monotonically decreasing magnitude frequency response from low to high frequency, and for any halftone pattern $d_{m,n}$ ($\neq c_{m,n}$) differing from the checker board pattern, we have

$$\mathcal{W}(x, d) > \mathcal{W}(x, c).$$

In this case, the FWMSE says that the checker board pattern is the best possible for any pattern that renders a gray patch at mid-tone. In other words, the FWMSE gives us what we expected; the most regular pattern (checker board) for rendering mid-tone yields the smallest error. ∎

Since the checker-board pattern is the most regular one for rendering the gray patch at mid-tone ($x_{m,n} = 0.5$), Example 1 shows that FWMSE is consistent with the regularity of halftone patterns for *this particular case*. We will show next using one-dimensional and two-dimensional examples that FWMSE is not always in agreement with the regularity of patterns. The one-dimensional example (Example 2) and its spectral plots are very illuminating in revealing what can go wrong in FWMSE. The two-dimensional examples are constructed in similar fashion and give realistic demonstrations of the problem.

Example 2: Consider a one-dimensional constant gray "image" of length 64

$$x_n = 0.25 \qquad n = 0, 1, \ldots, 63.$$

Two possible halftone representations for x_n are

$$t_n = \begin{cases} 1 & \text{if } n = 4k \\ 0 & \text{otherwise} \end{cases} \qquad 0 \le n < 64$$

and

$$u_n = \begin{cases} 1 & \text{if } n = 8k \text{ or } n = 8k+3 \\ 0 & \text{otherwise} \end{cases} \qquad 0 \le n < 64,$$

as shown in Fig. 3. Note that the ratios of black pixels to white pixels for both t_n and u_n are 3:1, giving an average intensity of 0.25. It is evident that the spatial distribution of black and white pixels is more "uniform" in t_n than in u_n. As a result, t_n would appear to be visually smoother than u_n, and hence t_n is preferred over u_n as a halftone for the constant graylevel 0.25. Plotted in Fig. 3 are the magnitude square of the 64-point discrete Fourier Transforms of t_n and u_n given by

$$|T_m|^2 = \begin{cases} 256 & \text{if } m = 0, 16, 32, 48 \\ 0 & \text{otherwise} \end{cases} \qquad 0 \le m < 64$$

and

$$|U_m|^2 = \begin{cases} 256 & m = 0 \\ 37.5 & m = 8, 56 \\ 128 & m = 16, 48 \\ 218.5 & m = 24, 40 \\ 0 & \text{otherwise} \end{cases} \qquad 0 \le m < 64.$$

Consider the frequency weighted mean squared error between the halftone patterns and the dc signal x_n using a filter with a frequency response (64-point DFT of an impulse response h_n)

$$H_m = \begin{cases} 1 & \text{if } m = 0 \\ \alpha & \text{if } m = 8, 56 \\ \beta & \text{if } m = 16, 48 \\ \gamma & \text{if } m = 24, 40 \\ 0 & \text{if } m = 32 \\ \text{arbitrary} & \text{otherwise.} \end{cases} \qquad 0 \le m < 64 \tag{6}$$

Note that $m = 32$ corresponds to one half of the sampling frequency, and that we have only specified in this example the response of H_m at a small number of frequencies (5 frequencies out of 32, taking the symmetry into account). We will choose the parameters of H_m to be consistent with a typical low pass filter that models the human visual response.

The frequency weighted mean squared errors are

$$\mathcal{W}(x, t) = \beta^2(|T_{16}|^2 + |T_{48}|^2) = 512\beta^2$$

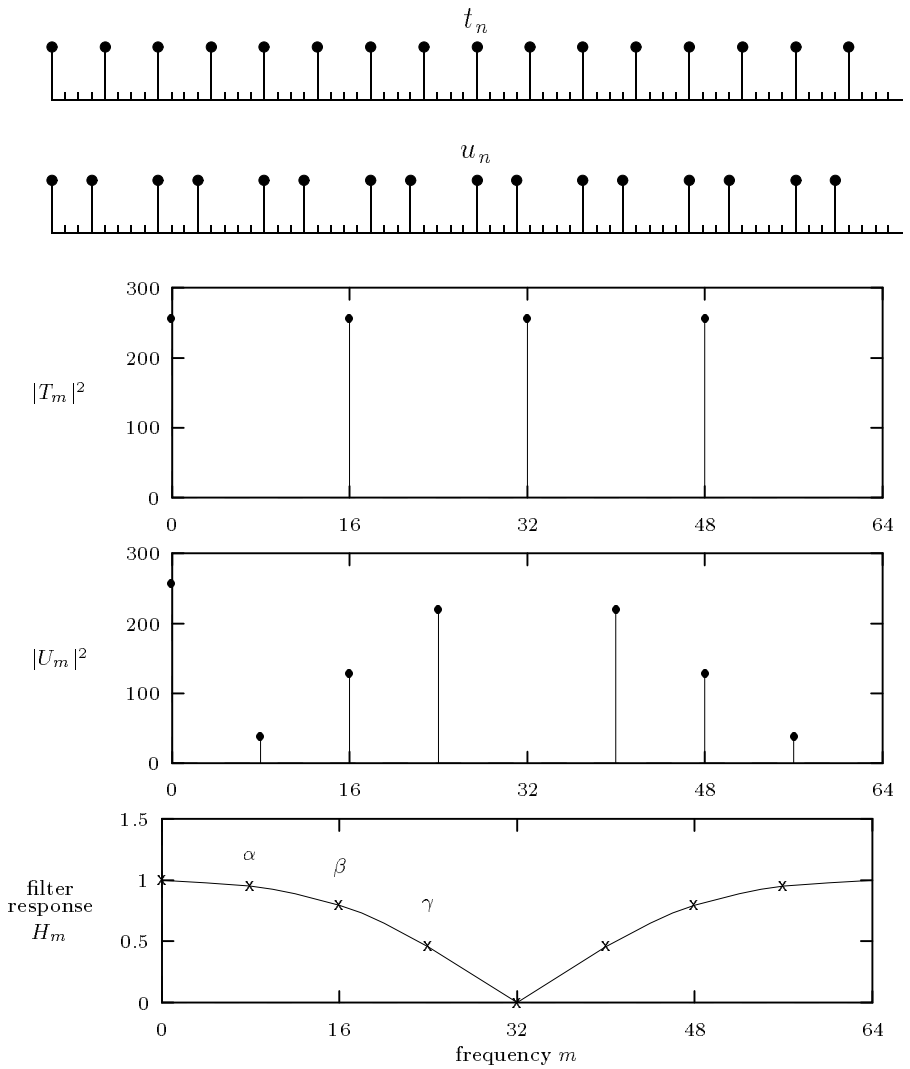

Fig. 3. Plots of the two halftones t_n and u_n, the magnitude squared spectra $|T_m|^2$ and $|U_m|^2$, and a filter response H_m that leads to the conclusion $\mathcal{W}(x,t) > \mathcal{W}(x,u)$. Notice that the frequency response of the filter is not squared.

Fig. 4. Two possible halftone dot patterns for a constant gray patch at the graylevel $g = 0.125$. This example leads to the conclusion that FWMSE can favor a less visually preferred pattern, i.e., $\mathcal{W}(x,p) > \mathcal{W}(x,q)$.

and

$$\begin{aligned}\mathcal{W}(x,u) &= \alpha^2(|U_8|^2 + |U_{56}|^2) + \beta^2(|U_{16}|^2 + |U_{48}|^2) + \gamma^2(|U_{24}|^2 + |U_{40}|^2) \\ &= 74.98\alpha^2 + 256\beta^2 + 437.02\gamma^2.\end{aligned}$$

We like to choose α, β and γ so that $\mathcal{W}(x,t) > \mathcal{W}(x,u)$, i.e., we want to satisfy the inequality

$$256\beta^2 > 74.98\alpha^2 + 437.02\gamma^2. \tag{7}$$

There are infinite number of choices of the parameters that corresponds to H_m being a low pass filter, and that (7) is satisfied, implying $\mathcal{W}(x,t) > \mathcal{W}(x,u)$. An example is plotted in Fig. 3, with $\alpha = 0.95$, $\beta = 0.8$, and $\gamma = 0.46$. The specific values of H_m that we have specified are marked by crosses in Fig. 3. Any frequency response that passes through these points will result in $\mathcal{W}(x,t) > \mathcal{W}(x,u)$. The smooth curve drawn in Fig. 3 is an example of such a response. ∎

Example 2 shows that although t_n is visually preferred over u_n as a halftone, t_n incurs a bigger FWMSE than u_n. If we examine the spectral plots in Fig. 3, we notice that by shifting certain pixels with value "1" in t_n to new locations in u_n (resulting in a less regular pattern), we have also changed the spectra so that some energy in T_m (at $m = 16, 48$) has been pushed to the high frequency region in U_m (at $m = 24, 40$). These high frequency components in U_m are attenuated by the visual model in the calculation of the FWMSE, resulting in $\mathcal{W}(x,t) > \mathcal{W}(x,u)$.

In the following, we first construct a two-dimensional example (Example 3) in a similar fashion, using a somewhat contrived filter. Example 4 then illustrates the same problem using a filter derived from a well-known visual model in halftoning and printing.

Example 3: Consider a constant gray patch of size 8 by 8 at graylevel $g = 0.125$

$$x_{m,n} = 0.125 \qquad m = 0, 1, \ldots, 7; n = 0, 1, \ldots, 7.$$

The arrays $p_{m,n}$ and $q_{m,n}$ in Fig. 4 represent two possible halftone dot patterns for $x_{m,n}$. Note that the average graylevels for both $p_{m,n}$ and $q_{m,n}$ are $1/8$, as they both contain 8 entries

of 1's out of 64. The difference between $p_{m,n}$ and $q_{m,n}$ is that the 1's at locations $(2,2)$ and $(6,6)$ in $p_{m,n}$ has been moved to locations $(1,1)$ and $(7,7)$ in $q_{m,n}$[1]. It is perhaps obvious that the dot arrangement in $p_{m,n}$ is more "regular" compared to that in $q_{m,n}$, and hence $p_{m,n}$ is usually considered to be a better halftone rendition of $x_{m,n}$. We demonstrate in the following FWMSE can actually favors $q_{m,n}$ than $p_{m,n}$.

We first calculate the magnitude squares of the DFT's of $p_{m,n}$ and $q_{m,n}$ as

$$|P_{k,l}|^2 = \begin{matrix} 64 & 0 & 0 & 0 & 64 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 64 & 0 & 0 & 0 & 64 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 64 & 0 & 0 & 0 & 64 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 64 & 0 & 0 & 0 & 64 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \qquad |Q_{k,l}|^2 = \begin{matrix} 64 & 2 & 4 & 2 & 16 & 2 & 4 & 2 \\ 2 & 4 & 2 & 16 & 2 & 4 & 2 & 0 \\ 4 & 2 & 16 & 2 & 4 & 2 & 64 & 2 \\ 2 & 16 & 2 & 4 & 2 & 0 & 2 & 4 \\ 16 & 2 & 4 & 2 & 64 & 2 & 4 & 2 \\ 2 & 4 & 2 & 0 & 2 & 4 & 2 & 16 \\ 4 & 2 & 64 & 2 & 4 & 2 & 16 & 2 \\ 2 & 0 & 2 & 4 & 2 & 16 & 2 & 4 \end{matrix}$$

Note again that $|P_{0,0}|^2 = 64$ and $|Q_{0,0}|^2 = 64$, each corresponds to the entry at the upper left corner of the respective array.

For the sake of simplicity, let us assume that the filter $v_{k,l}$ has a symmetric frequency response of the form

$$V_{k,l} = \begin{matrix} 1 & \alpha & \beta & \gamma & 0 & \gamma & \beta & \alpha \\ \alpha & \alpha & \beta & \gamma & 0 & \gamma & \beta & \alpha \\ \beta & \beta & \beta & \gamma & 0 & \gamma & \beta & \beta \\ \gamma & \gamma & \gamma & \gamma & 0 & \gamma & \gamma & \gamma \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma & \gamma & \gamma & \gamma & 0 & \gamma & \gamma & \gamma \\ \beta & \beta & \beta & \gamma & 0 & \gamma & \beta & \beta \\ \alpha & \alpha & \beta & \gamma & 0 & \gamma & \beta & \alpha \end{matrix}$$

Note that the responses $|P_{k,l}|^2$, $|Q_{k,l}|^2$ and $V_{k,l}$ are all arranged in a *typical DFT fashion*. For example, the value $V_{0,0} = 1$ at the upper left hand corner corresponds to dc, while the row and column of zeros in $V_{k,l}$ correspond to one half of the sampling frequencies in the "vertical" and "horizontal" directions. The assumption here that $V_{k,l} = 0$ for $k = 4$ or $l = 4$ is not necessary but it simplifies our calculations a little bit. Since we want the filter to be low pass, the parameters should satisfy $1 \geq \alpha \geq \beta \geq \gamma \geq 0$.

Using (4), the FWMSE between $x_{m,n}$ and $p_{m,n}$ are

$$\mathcal{W}(x,p) = 4\beta^2 \quad \text{and} \quad \mathcal{W}(x,q) = 0.25\alpha^2 + 3\beta^2 + 1.75\gamma^2.$$

[1] The coordinate system is defined such that the origin $(0,0)$ is located at the upper left corner of the pattern.

Fig. 5. Two 8 by 8 halftone renditions, $r_{m,n}$ and $s_{m,n}$, of a constant gray patch of graylevel $g = 1/16$. Using a modified Mannos-Sakrison visual model at 300 dpi and 12 inches viewing distance, we have $0.6911 = \mathcal{W}(x, r) > \mathcal{W}(x, s) = 0.6764$, illustrating a less visually pleasing pattern can incur a smaller frequency weighted mean squared error.

We like to choose α, β and γ subject to the constraint $1 \geq \alpha \geq \beta \geq \gamma \geq 0$, so that $\mathcal{W}(x, p) > \mathcal{W}(x, q)$. That is, we like to satisfy the inequality

$$\beta^2 > 0.25\alpha^2 + 1.75\gamma^2. \tag{8}$$

There are infinite number of choices of the parameters that are consistent with $v_{k,l}$ being a low pass filter and that (8) is satisfied. The parameters

$$\alpha = 0.8, \qquad \beta = 0.5, \qquad \gamma = 0.2 \tag{9}$$

give an example where we have $\mathcal{W}(x, p) > \mathcal{W}(x, q)$. ∎

We have shown that although $p_{m,n}$ is visually preferred over $q_{m,n}$ as a halftone, $q_{m,n}$ incurs a smaller frequency weighted mean squared error than $p_{m,n}$. We have used for simplicity an 8 by 8 example here, where the frequency response of the visual filter $V_{k,l}$ may appear to vary rather abruptly from the pass band to the stop band. If we want an example where the visual filter would have a finer frequency resolution, e.g., specified by a 16 by 16 point or bigger DFT, we can replicate the dot patterns $p_{m,n}$ and $q_{m,n}$ to the desired size before we take the DFT. The conclusion will remain unchanged. Finally, we demonstrate in the next example the problem with FWMSE using a well-known visual model with realistic parameters for printing at 300 dpi.

Example 4: Consider the two 8 by 8 bit patterns $r_{m,n}$ and $s_{m,n}$ of Fig. 5, that are possible renditions of a constant grayscale patch of graylevel $g = 1/16$. The Mannos-Sakrison's visual model [11] is

$$\tilde{A}(f_r) = 2.6(0.0192 + 0.114 f_r) \exp[-(0.114 f_r)^{1.1}]$$

where f_r is the number of cycles of a sinusoid subtended per degree angle at the retina. This response peaks at $f_r = 7.891$ with a peak value 0.9809. As is usually done, we use a modified Mannos-Sakrison response given by

$$A(f_r) = \begin{cases} \tilde{A}(f_r) & \text{if } f_r \geq 7.891 \\ L(f_r) & \text{if } 0 \leq f_r < 7.891 \end{cases}$$

where $L(f_r)$ is a straight line segment joining the points $A(0) = 1$ and $A(7.891) = 0.9809$. Using 8 by 8 point discrete Fourier transform, the frequency response of the visual filter is

$$V_{k,l} = A(\sqrt{f_k^2 + f_l^2}) \qquad k = 0, 1, \ldots 4; l = 0, 1, \ldots 4$$

and by symmetry

$$V_{8-k,l} = V_{k,8-l} = V_{8-k,8-l} = V_{k,l} \qquad k = 0, 1, \ldots, 4; l = 0, 1, \ldots, 4$$

where

$$f_i = 2DR \tan(0.5^\circ) i/8 \qquad i = k, l,$$

D is the viewing distance in inches, and R is the resolution in dots per inch. For $D = 12$ inches and $R = 300$ dots per inch, the magnitude squared response of the visual filter $|V_{k,l}|^2$ is

$$\begin{bmatrix} 1 & 0.962 & 0.497 & 0.132 & 0.026 & 0.132 & 0.497 & 0.962 \\ 0.962 & 0.836 & 0.378 & 0.103 & 0.021 & 0.103 & 0.378 & 0.836 \\ 0.497 & 0.378 & 0.170 & 0.050 & 0.011 & 0.050 & 0.170 & 0.378 \\ 0.132 & 0.103 & 0.050 & 0.017 & 0.004 & 0.017 & 0.050 & 0.103 \\ 0.026 & 0.021 & 0.011 & 0.004 & 0.001 & 0.004 & 0.011 & 0.021 \\ 0.132 & 0.103 & 0.050 & 0.017 & 0.004 & 0.017 & 0.050 & 0.103 \\ 0.497 & 0.378 & 0.170 & 0.050 & 0.011 & 0.050 & 0.170 & 0.378 \\ 0.962 & 0.836 & 0.378 & 0.103 & 0.021 & 0.103 & 0.378 & 0.836 \end{bmatrix}$$

The matrix is organized in the typical DFT fashion where the dc energy $|V_{0,0}|^2$ is at the upper left hand corner, and the highest frequency corresponds to the fifth column ($k = 4$) from the top and the fifth row ($l = 4$) from the left. For this example, we have

$$\mathcal{W}(x, r) = 0.6911 \qquad \text{and} \qquad \mathcal{W}(x, s) = 0.6764.$$

Once again this illustrates that a less regular pattern $s_{m,n}$ can incur a smaller frequency weighted mean squared error than a more regular pattern $r_{m,n}$. ∎

We emphasize that it is *not* true that we always have a smaller frequency mean squared error with a less regular halftone pattern than with a more regular one. It is true, however, that there are many examples that can lead to the same conclusion as in Examples 2 to 4. Hence it is *correct* to say that frequency weighted mean squared error does not generally reflect the uniformity in the distribution of black and white dots in a halftone.

3 A Mixture Distortion Criterion

The example in the previous section demonstrates a shortcoming of FWMSE in that it does not always reflects the spatial distribution of the halftone dots. Hence if we use the FWMSE with an optimization based halftoning algorithm, we can obtain suboptimal results

in the sense that the dot patterns in the output halftones may not be the most subjectively pleasing. One way to improve on the output halftone quality is to use an adaptive FWMSE where we allow the spatial width of the visual filter (hence also the frequency response) to change according to the local graylevel [16–18], i.e., use a filter with a larger width for highlight and shadow regions where the dots are spaced far apart, and use a filter with a smaller width for mid-tone regions. In error diffusion one can also explicitly spread out the halftone dots according to the optimum distance using an output feedback mechanism [15]. Here we consider a new distortion criterion that explicitly incorporates information on the spatial distribution of halftone dots. This distortion criterion has been used with a tree coding algorithm [7,8] and an entropy constrained error diffusion algorithm [19] to generate high quality halftones.

We use the concept of *minority pixels* as defined by Ulichney [14]. Specifically, if the gray scale of a local smooth region in an image is between 0 and 0.5, then the number of black pixels in a halftone must be larger than the number of white pixels in the corresponding region. In such case the white pixels are called minority pixels. Similarly, the black pixels are minority pixels when the local graylevel has a value between 0.5 and 1. Let

$$\rho_{m,n} = \begin{cases} 1 & \text{if } 0 \le x_{m,n} < 0.5 \\ 0 & \text{if } 0.5 \le x_{m,n} \le 1 \end{cases}$$

be the value of the minority pixel at location (m,n). Based on an approximation using square packing, one can define the *principal distance* d_p [14] as the average distance between minority pixels in a halftone. Specifically,

$$d_p(g) = \begin{cases} \sqrt{1/g} & \text{if } 0 \le g < 0.5 \\ \sqrt{1/(1-g)} & \text{if } 0.5 \le g \le 1, \end{cases}$$

where g is the local gray level. Note that $d_p(g)$ is infinite for $g = 0$ or $g = 1$, as it should because no minority pixel should be inserted for complete black or white gray values. Let $\delta_{m,n}$ be the distance from the position (m,n) to the nearest minority pixel. We can define a distortion measure using the distances between minority pixels by

$$\mu_{m,n} = \begin{cases} 0 & \text{if } \delta_{m,n} \ge d_p(x_{m,n}) \\ & \quad \text{and } b_{m,n} = \rho_{m,n} \\ 0 & \text{if } \delta_{m,n} < d_p(x_{m,n}) \\ & \quad \text{and } b_{m,n} \ne \rho_{m,n} \\ \left(\dfrac{d_p(x_{m,n}) - \delta_{m,n}}{d_p(x_{m,n})} \right)^2 & \text{otherwise.} \end{cases} \tag{10}$$

Note that $\mu_{m,n}$ favors putting a majority pixel at (m,n) if the distance from the nearest minority pixel is less than $d_p(x_{m,n})$, while it favors a minority pixel at (m,n) if the distance from the nearest minority pixel is larger than $d_p(x_{m,n})$.

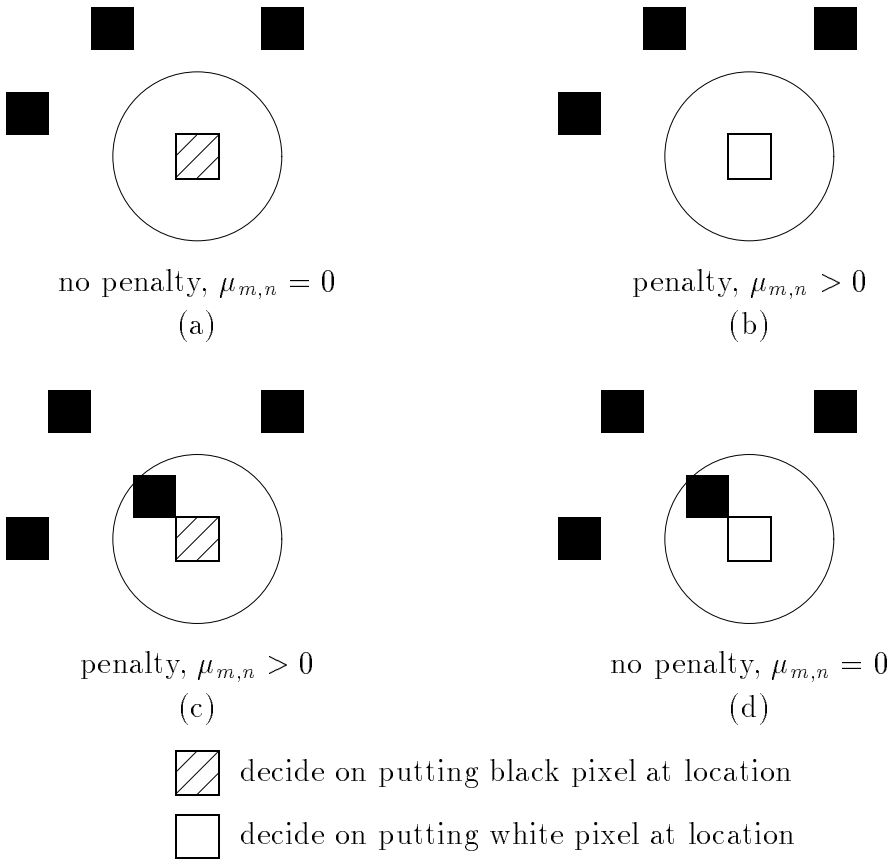

Fig. 6. Examples showing the four different situations in the dot distance based distortion measure of (10). In these examples, we have $g = 0.75$, $\rho = 0$ (minority pixel is black) and $d_p(g) = 2$. The circle in each case is of radius 2, which equals the principal distance $d_p(g)$ at the graylevel used in this example.

Consider an example with $g = 0.75$. Hence we have $\rho = 0$, i.e., the minority pixels are black pixels, and $d_p(g) = 2$. In the two cases shown in Fig. 6 (a) and (b), all the existing minority pixels in the halftone are more than a distance of 2 away from the location being considered. It would have be desirable if we could put a black pixel at some "past" locations so that the distance between minority pixels could be kept to $d_p(g)$. Since we cannot change the pixels that are already on the page, we would want to put a black pixel at the current location. Consequently we assign a penalty to case (b), and no penalty to case (a). On the other hand, the distance from (m, n) to the nearest minority pixel is only $\sqrt{2}$ in the cases (c) and (d), which is smaller than the principal distance. In such a situation, we would want to put a white pixel at position (m, n). Consequently, we put a penalty to case (c), and no penalty to case (d). The specific penalty as defined in (10) is given by the relative error between the principal distance and the actual distance to the nearest minority pixel. The criterion in (10), when incorporated into a distortion measure for an optimization based halftoning algorithm, explicitly encourages the minority pixels in a halftone to be located apart by the principal distance. As a result, this allows a smooth rendition of the continuous tone image, which leads to good subjective halftone quality.

Using the frequency weighted mean squared error and the distance from the nearest minority pixel, we define a mixture distortion measure as

$$e_{m,n} = w_{m,n} + \lambda \mu_{m,n} \tag{11}$$

where λ is an experimentally determined parameter that controls the weighting between $w_{m,n}$ and $\mu_{m,n}$. Previously, we have found the value of 0.03 to be good in a halftoning algorithm based on tree coding [7,8]. The overall mixture distortion is then given by

$$\mathcal{D}(x, b) = \sum_{m,n} (w_{m,n} + \lambda \mu_{m,n}).$$

The mixture distortion criterion has been applied in a tree coding halftoning algorithm [7,8] and an entropy constrained error diffusion algorithm [19], that have shown to be effective in generating halftones with good quality.

4 Conclusion

We have considered in this paper a deficiency of the frequency weighted mean squared error (FWMSE) for describing the quality of halftones, because it does not explicitly take into account the distance between halftone dots. Examples are given using halftone renditions of a constant gray patch that shows a FWMSE can give preference to a halftone rendition that is considered to be less visually favorable than another rendition. We have also proposed a mixture distortion criterion, that is a weighted combination of FWMSE and a measure that reflects the distance between halftone dots. Such a distortion criterion has been used elsewhere to generate high quality halftones.

References

[1] Q. Lin, "Halftone image quality analysis based on a human vision model," in *Proceedings of SPIE, vol. 1913*, pp. 378–389, February 1993.

[2] J. P. Allebach, "Visual model-based algorithms for halftoning images," in *Proceedings of SPIE, vol. 310*, pp. 151–158, 1981.

[3] J. B. Mulligan and A. J. Ahumada, Jr., "Principled halftoning based on models of human vision," in *Proceedings of SPIE, vol. 1666*, (San Jose, CA), pp. 109–121, February 1992.

[4] R. A. Vander Kam, P. A. Chou, and R. M. Gray, "Combined halftoning and entropy-constrained vector quantization," in *SID Digest of Technical Papers*, (Seattle, WA), pp. 223–226, May 1993.

[5] A. Zakhor, S. Lin, and F. Eskafi, "A new class of B/W halftoning algorithms," *IEEE Transactions on Image Processing*, vol. 2, pp. 499–509, October 1993.

[6] J. P. Allebach, T. J. Flohr, D. P. Hilgenberg, and C. B. Atkins, "Model-based halftoning via direct binary search," in *Proceedings of IS&T's* 47^{th} *Annual Conference*, (Rochester, NY), pp. 476–482, May 1994.

[7] P. W. Wong, "Image halftoning using multipath tree coding," in *Proceedings of ICIP*, (Austin, TX), pp. II 31–35, November 1994.

[8] P. W. Wong, "Entropy constrained halftoning using multipath tree coding," *IEEE Transactions on Image Processing*. To appear.

[9] R. A. Vander Kam, P. A. Chou, E. A. Riskin, and R. M. Gray, "An algorithm for joint vector quantizer and halftoner design," in *Proceedings of ICASSP*, (San Francisco CA), March 1992.

[10] J. G. Robson, "Spatial and temporal contrast sensitivity functions of the visual system," *Journal of The Optical Society of America*, vol. 56, pp. 1141–1142, 1966.

[11] J. L. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 525–536, July 1974.

[12] X. Zhang and B. A. Wandell, "A spatial extension of CIELAB for digital color image reproduction," in *SID International Symposium, Digest of Technical Papers*, pp. 731–734, 1996.

[13] Z. M. Zhang, J. E. Farrell, and B. A. Wandell, "Aplications of a spatial extension to CIELAB," in *Proceedings of SPIE/IS&T Symposium on Electronic Imaging*, (San Jose, CA), February 1997.

[14] R. A. Ulichney, "Dithering with blue noise," *Proceedings of the IEEE*, vol. 76, pp. 56–79, January 1988.

[15] R. Levien, "Output dependent feedback in error diffusion halftoning," in *Proceedings of IS&T's* 46th *Annual Conference*, (Cambridge, MA), pp. 115–118, May 1993.

[16] Q. Lin, "Improving halftone uniformity and tonal response," in *Proceedings of IS&T International Congress on Advances in Non-impact Printing Technologies*, (New Orleans LA), pp. 377–380, November 1994.

[17] Q. Lin, "Screen design for printing," in *Proceedings of ICIP*, (Washington, DC), pp. II 331–334, October 1995.

[18] P. W. Wong, "Adaptive error diffusion and its application in multiresolution rendering," *IEEE Transactions on Image Processing*, vol. 5, pp. 1184–1196, July 1996.

[19] P. W. Wong and H. Nguyen, "Error diffusion with an entropy constraint," in *Proceedings of The* 9th *IEEE IMDSP Workshop*, (Belize City, Belize), pp. 104–105, March 1996.