A Mixture Distortion Criterion for Halftones

Ping Wah Wong
Hewlett Packard Laboratories, Palo Alto, California

Abstract

Frequency weighted mean squared error (FWMSE) is often used for measuring image quality. We construct examples to show a weakness of FWMSE when applied to halftones. We then consider a mixture distortion that consists of FWMSE and a dot distance term to explicitly account for the spatial arrangement of dots.

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 $d(x_{m,n}, b_{m,n})$ to measure the differences between the halftone bm,n and its continuous tone counterpart $x_{m,n}$. A distortion measure is also essential for optimization based halftoning algorithms,[1–7] where one finds a halftone $b_{m,n}$ from a continuous tone image $x_{m,n}$ so that the average distortion $Ed(x_{m,n}, b_{m,n})$ is minimized.

Frequency weighted mean squared error (FWMSE)[8–10] is perhaps the most popular distortion criterion that is used in practice, partly because of its simplicity and tractability. Let the pixel values of the 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

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

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

$$W(x,b) = \sum_{m,n} w_{m,n} \tag{2}$$

where the sum is taken over all the pixels in the image.

The operation of (1) can be represented by the block diagram in Figure 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 popular in the literature is obtained by replacing $w_{m,n}$ with

$$\tilde{w}_{m,n} = \left(\sum_{k,l} v_{k,l} \left(x_{m-k,n-l} - b_{m-k,n-l} \right) \right)^2, \tag{3}$$

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 (3) 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, with suitable modifications, can also be applied when using (3).

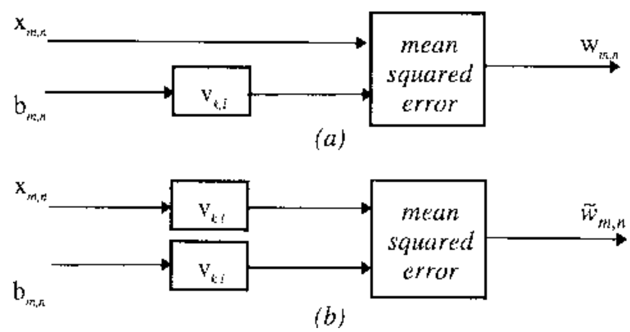

Figure 1. Two different forms of frequency weighted mean squared error.

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,[11] meaning that the error spectra between continuous tone and halftone images should preferably be concentrated in the high 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.[12] Here 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.[6,7]

A Deficiency of 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. Consider a constant gray patch of size 8 by 8 as

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

The arrays $p_{m,n}$ and $q_{m,n}$ in Fig. 2 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 $q_{m,n}$.* 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 therefore would like to have a distortion measure that favors $p_{m,n}$.

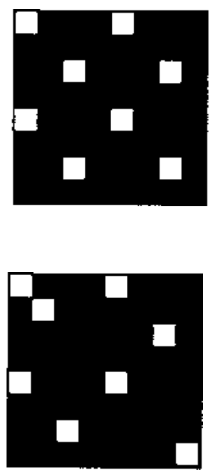

Figure 2. Two possible halftone dot patterns for a constant gray patch at the graylevel g = 0.125. Here, the graylevel values for black and white pixels are 0 and 1, respectively.

It is evident from (1) that the spatial distribution of the black and white pixels is not *explicitly* reflected by the FWMSE. It is therefore conceivable that 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 fact, $p_{m,n}$ and $q_{m,n}$ of Fig. 2 presents one such example. To see this, we calculate the magnitude squares of the DFT's of $p_{m,n}$ and $q_{m,n}$ as

$$\left|P_{k,l}\right|^2 = \begin{bmatrix} 64 & 0 & 0 & 0 & 64 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 64 & 0 & 0 & 0 & 64 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 64 & 0 & 0 & 0 & 64 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 64 & 0 & 0 & 0 & 64 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$\left|Q_{k,l}\right|^2 = \begin{bmatrix} 64 & 2 & 4 & 2 & 16 & 2 & 4 & 2 \\ 2 & 4 & 2 & 16 & 2 & 4 & 2 & 0 \\ 4 & 2 & 16 & 2 & 4 & 2 & 64 & 2 \\ 2 & 16 & 2 & 4 & 2 & 0 & 2 & 4 \\ 16 & 2 & 4 & 2 & 64 & 2 & 4 & 2 \\ 2 & 4 & 2 & 0 & 2 & 4 & 2 & 16 \\ 4 & 2 & 64 & 2 & 4 & 2 & 16 & 2 \\ 2 & 0 & 2 & 4 & 2 & 16 & 2 & 4 \end{bmatrix}$$

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

$$V_{k,l} = \begin{bmatrix} 1 & \alpha & \beta & \gamma & 0 & \gamma & \beta & \alpha \\ \alpha & \alpha & \beta & \gamma & 0 & \gamma & \beta & \alpha \\ \beta & \beta & \beta & \gamma & 0 & \gamma & \beta & \beta \\ \gamma & \gamma & \gamma & \gamma & 0 & \gamma & \gamma & \gamma \\ 64 & 0 & 0 & 0 & 64 & 0 & 0 & 0 \\ \gamma & \gamma & \gamma & \gamma & 0 & \gamma & \gamma & \gamma \\ \beta & \beta & \beta & \gamma & 0 & \gamma & \beta & \beta \\ \alpha & \alpha & \beta & \gamma & 0 & \gamma & \beta & \alpha \end{bmatrix}$$

Note that the arrays $|P_{k,l}|^2$, $|Q_{k,l}|^2$ and $V_{k,l}$ are all arranged in a *typical DFT fashion*. That is, 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 in this example. Since we want the filter to be low pass, the parameters should satisfy $1 \geq \alpha \geq \beta \geq \gamma \geq 0$.

The 8 by 8 DFT of $x_{m,n}$ is

$$X_{k,l} = \begin{cases} 8 & \text{if } k = l = 0 \\ 0 & \text{otherwise} \end{cases}$$

Using Parserval theorem with (2), the FWMSE between $x_{m,n}$ and $p_{m,n}$ can be calculated as

$$W(x,p) = \frac{1}{64} \sum_{(k,l)\neq 0} \left|P_{k,l}\right|^2 \left|V_{k,l}\right|^2 = 4\beta^2.$$

Similarly,

$$W(x,q) = 0.25\alpha^2 + 3\beta^2 + 1.75\gamma^2.$$

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

$$\beta^2 > 0.25\alpha^2 + 1.75\gamma^2. \qquad (4)$$

There are infinite number of choices of the parameters that are consistent with $V_{k,l}$ being a low pass filter and that (4) is satisfied. We can, for example, choose

$$\alpha = 0.8, \qquad \beta = 0.5, \quad \text{and} \quad \gamma = 0.2. \qquad (5)$$

We have shown that although $p_{m,n}$ is visually preferred over $q_{m,n}$ as a halftone, $q_{m,n}$ incurs a smaller FWMSE than $p_{m,n}$. We have used for simplicity an 8 by 8

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

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 still come out to be the same. More importantly, note that the sequences $p_{m,n}$ and $q_{m,n}$, as well as the responses specified in (5) only serve as a convenient example. There are many other halftone dot patterns for various continuous tone images that can lead to the same conclusion, i.e., that the FWMSE does not generally reflect the uniformity in the distribution of black and white dots in a halftone.

A Mixture Distortion Criterion

The example in the previous section demonstrates a shortcoming of FWMSE. It is evident that 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 of the highest quality. Previously a high quality error diffusion algorithm has been design that explicitly controls the distances between halftone dots.[12] We now 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 to generate high quality halftones.[6,7]

We use the concept of *minority pixels* as defined by Ulichney.[11] 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 so that the graylevel is rendered correctly. 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

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

be the value of the minority pixel at the location (m,n). Based on an approximation using square packing, one can define the principal distance d_p[11] 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 $d_{m,n}$ 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

$$u_{m,n} = \begin{cases} 0 & \text{if } d_{m,n} \ge d_p(x_{m,n}) \text{ and } b_{m,n} = \rho_{m,n} \\ 0 & \text{if } d_{m,n} < d_p(x_{m,n}) \text{ and } b_{m,n} \ne \rho_{m,n} \\ \left(\dfrac{d_p(x_{m,n}) - d_{m,n}}{d_p(x_{m,n})} \right) & \text{otherwise.} \end{cases} \qquad (6)$$

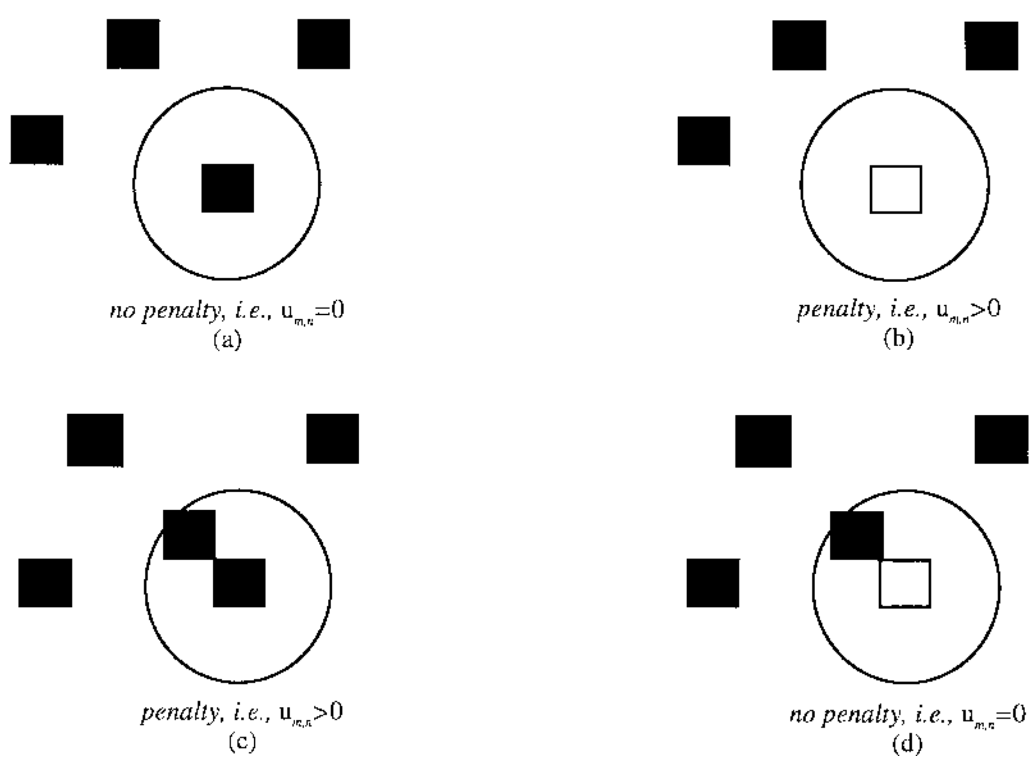

Figure 3. Examples showing the four different situations in the dot distance based distortion measure of (6). In these examples, we have g = 0.75, r = 0 (minority pixel is black) and dp(g) = 2. The circle in each case is of radius 2, which equals the principal distance dp(g) at the graylevel used in this example.

Note that $u_{m,n}$ 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 $d_p(x_{m,n})$.

Consider an example with $g = 0.75$. Hence we have $\rho = 0$, i.e., the minority pixels are black pixels, and $d_p(g) = 2$. We have drawn a circle of radius 2 in the four cases of Fig. 3 centered at the pixel location being considered. For the two cases in Fig. 3(a) and (b), all the existing minority pixels in the halftone are more than a distance of 2 away from the current location. Since the minority dots in the current neighborhood are too sparse compared to $d_p(g)$, we favor the case of having a black pixel in the center of the circle. Consequently we assign a penalty to Fig. 3 (b), and no penalty to Fig. 3 (a). On the other hand, the distance from (m,n) to the nearest minority pixel is only $\sqrt{2}$ in Fig. 3 (c) and (d), which is smaller than the principal distance. In such a situation, we favor having a white pixel at position (m,n). Consequently, we put a penalty to Fig. 3 (c), and no penalty to Fig. 3 (d). The specific penalty as defined in (6) is given by the relative error between the principal distance and the actual distance to the nearest minority pixel.

Using the FWMSE and (6), we define a mixture distortion measure as

$$e_{m,n} = w_{m,n} + \mu u_{m,n} \tag{7}$$

where μ is an experimentally determined parameter that controls the weighting between $w_{m,n}$ and $u_{m,n}$.[6,7] This mixture distortion criterion evidently contains a penalty term that explicitly depends on the distance between halftone dots. It is important to note that the principal distance is a function of the pixel value of the graylevel image, and hence it varies from locations to locations within the image. As a result, (7) encourages the minority dots to be spread out within local neighborhoods in a fashion that is consistent with the average local graylevels. A similar approach, that explicitly considers the distance between minority pixels, has been introduced to error diffusion[12] to obtain good output quality. We have successfully used (7) in conjunction with a tree coding algorithm to generate very high quality halftones.[6,7]

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. J. P. Allebach, "Visual model-based algorithms for halftoning images," in *Proceedings of SPIE*, vol. **310**, pp. 151-158, 1981.
2. J. B. Mulligan and A. J. Ahumada, Jr., "Principled halftoning based on models of human vision," in *Proceedings of SPIE*, vol. **1666**, pp. 109-121, February 1992.
3. 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.
4. 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.
5. 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 47th Annual Conference* (Rochester, NY), pp. 476-482, May 1994; ***see Recent Progress in Digital Halftoning, Vol. I, pg. 50***.
6. P. W. Wong, "Image halftoning using multipath tree coding," in *Proceedings of ICIP* (Austin, TX), pp. II 31-35, November 1994.
7. P. W. Wong, "Entropy constrained halftoning using multipath tree coding," submitted to *IEEE Transactions on Image Processing.*
8. 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.
9. 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. **20**, pp. 525-536, July 1974.
10. Q. Lin, "Halftone image quality analysis based on a human vision model," in *Proceedings of SPIE*, vol. **1913**, pp. 378-389, February 1993.
11. R. A. Ulichney, "Dithering with blue noise," *Proceedings of the IEEE*, vol. **76**, pp. 56-79, January 1988.
12. 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; ***see Recent Progress in Digital Halftoning, Vol. I, pg. 106***.

❋ Previously published in *OSA/IS&T Optics & Imaging in the Information Age*, pp. 187–191, 1996.