Digital Halftoning as 2-D Delta-Sigma Modulation

Thomas D. Kite, Brian L. Evans, and Alan C. Bovik
Laboratory for Vision Systems
Department of Electrical and Computer Engineering
The University of Texas at Austin, Austin, TX 78712-1084 USA
E-mail: {tom,bevans,bovik}@vision.ece.utexas.edu

Terry L. Sculley
ESS Technology, Inc., Austin, TX 78746 USA
E-mail: Terry.Sculley@esstech.com

Abstract— **The error diffusion algorithm for digital halftoning is equivalent in form to a noise-shaping feedback coder, a class of delta-sigma modulator. The white noise assumption of the quantizer error is known to be false; in fact, the quantizer error is seen to be highly correlated with the input image. To account for this correlation, we use a gain model for the quantizer. This model accurately predicts the edge sharpening and noise shaping caused by all error diffusion schemes. It also permits an extension of error diffusion to oversampled imagery.**

I. Introduction

In 1975, Floyd and Steinberg [1] introduced error diffusion for halftoning grey-level images. Since then, researchers have been trying to understand the mechanism by which it works, in particular, how the error filter affects the halftoned results. The choice of the filter coefficients is driven by three factors (in order of importance):

- Perceptual quality of the halftoned image
- Ease of computation
- Checkered output at mid-gray

By trial and error, Floyd and Steinberg found a filter that achieves good visual quality, has only four dyadic coefficients, and produces a checkered output at mid-gray. (The checkerboard output at midgray, however, is not a definitive indication of the efficacy of a halftoning algorithm, but simply a measure that Floyd and Steinberg used.) Larger filters [2], [3], input-dependent weights [4], non-standard scanning schemes [5] and optimal filter design [6] have been used to improve performance.

During the 1990s, the connection between error diffusion and delta-sigma modulation has been published. The equivalent circuit for error diffusion is identical to the noise shaping feedback coder, a form of delta-sigma modulator that is commonly used for wordlength reduction, for example, in digital audio [7]. Delta-sigma modulation [8], which predates the Floyd-Steinberg algorithm, has only become popular in the last decade because of the need to fabricate high-quality, low-cost data converters using VLSI technology. Delta-sigma modulation permits the use of a low-resolution converter in a high-resolution application by feeding back the quantization error to linearize the converter and reduce the in-band quantization noise. Indeed, it is possible to have a one-bit converter inside the feedback loop; this technology has become widely used in consumer digital audio products [7].

The similarity between one-bit digital-to-analog conversion and halftoning by error diffusion is evident: both accept a high-resolution input and produce a single-bit output that, as closely as possible, resembles the input in the passband. In the audio industry, simulation is used to predict the performance of delta-sigma modulators [9]. In error diffusion, however, the much larger data sets make this infeasible, and therefore a model which accurately predicts performance is required. In this paper, we show how a linear gain model for the quantizer is able to predict both the edge sharpening and the noise shaping of error diffusion schemes, allowing error filters to be designed and optimized automatically.

II. Background

The equivalent circuit of error diffusion is shown in Figure 1. The image is scanned and the current pixel, here denoted $x(n)$, is quantized to one-bit resolution. The quantization error is filtered by a (possibly two-dimensional) filter $H(z)$, and subtracted from the input. Because the system is causal, the current pixel is not an input to the filter. Analysis of the noise-shaping feedback coder proceeds by assuming that thresholding is a linear operation which adds white noise to the signal. While this is somewhat difficult to justify mathematically, it at least makes the system tractable, and can also lead to useful results.

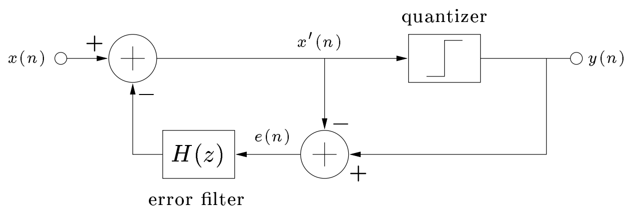


Fig. 1. Noise-shaping feedback coder.

T. Kite was supported in part by NASA Grant NAGW4247. B. Evans was supported in part by NSF CAREER Award under Grant MIP-9702707 and a UT Austin Summer Research Assignment Grant.

T. Sculley was formerly with Crystal Semiconductor Corporation, Austin, TX 78744 USA.

Referring to Figure 1, in the z-domain, we have

$$Y(z) = X(z) + N(z)\,(1 - H(z)) \tag{1}$$

where we see that the signal is passed undelayed to the output, and the noise is shaped by the factor $1-H(z)$. For oversampling systems, $H(z)$ is usually a lowpass function; thus the noise response is highpass. This action is called *noise shaping*—the quantization noise is moved from low frequencies to higher frequencies, where it can be filtered out. For Nyquist-rate systems, in which all the quantization noise falls in the signal band, $H(z)$ may be any function which shapes the noise so that it is perceptually less objectionable. The total quantization noise power does not change, but the perceptually-weighted signal-to-noise ratio improves. In error diffusion, we usually assume that the number of input and output pixels are equal; that is, the system operates at the Nyquist rate. Models of the human visual system [10] indicate that the eye-brain system is lowpass. Therefore, the function $H(z)$ should be lowpass.

In summary, a digital halftoning scheme must perform the same task as a one-bit delta-sigma modulator; namely, reduce the wordlength of a high-resolution input to one bit while shaping the quantization noise to increase the perceived signal-to-noise ratio.

III. Modeling the Quantizer

The simple analysis that results in (1), as we have seen, assumes that the quantization error is white and uncorrelated with the input. It has been shown that this is not generally true for either 1-D [9] or 2-D modulators [11]. Knox demonstrated that the quantization error in the 2-D case is in fact almost linearly correlated with the input [11]. This suggests a model for the quantizer comprising a gain block and a source of additive white noise, as shown in Figure 2. The idea of modeling a quantizer as a linear gain block has been used in the 1-D case [12], but has not been applied to error diffusion before in the open literature.

Referring to Figure 2, we see that the quantizer has been replaced by the cascade of a gain block of gain K and an additive, uncorrelated white noise source, $q(n)$. The error signal becomes

$$e(n) = (K-1)x'(n) + q(n) \tag{2}$$

where $x'(n)$ is the input to the quantizer. The system with the gain model for the quantizer is still linear, and

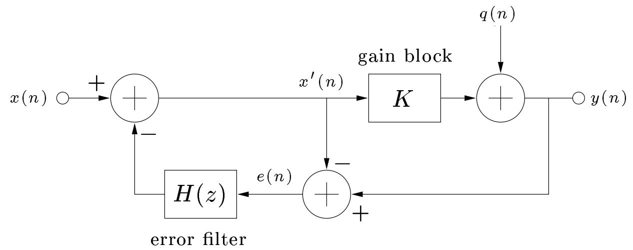


Fig. 2. Model of noise-shaping feedback coder using gain block.

TABLE I

Computed values for quantizer gain K for various schemes and images.

Input Image	Floyd-Steinberg	Jarvis	Stucki
lena	2.11	5.21	4.42
boats	2.00	4.88	4.24
barbara	2.03	4.01	3.74
mandrill	2.05	3.69	3.51
Average	2.05	4.45	3.98

we analyze it in the z-domain as before to arrive at

$$Y(z) = \underbrace{\frac{K}{1+(K-1)H(z)}}_{\text{STF}}X(z) + \underbrace{\frac{1-H(z)}{1+(K-1)H(z)}}_{\text{NTF}}N(z) \tag{3}$$

where STF and NTF are the signal transfer function and the noise transfer function, respectively. Note that (3) reduces to (1) when $K = 1$, as expected.

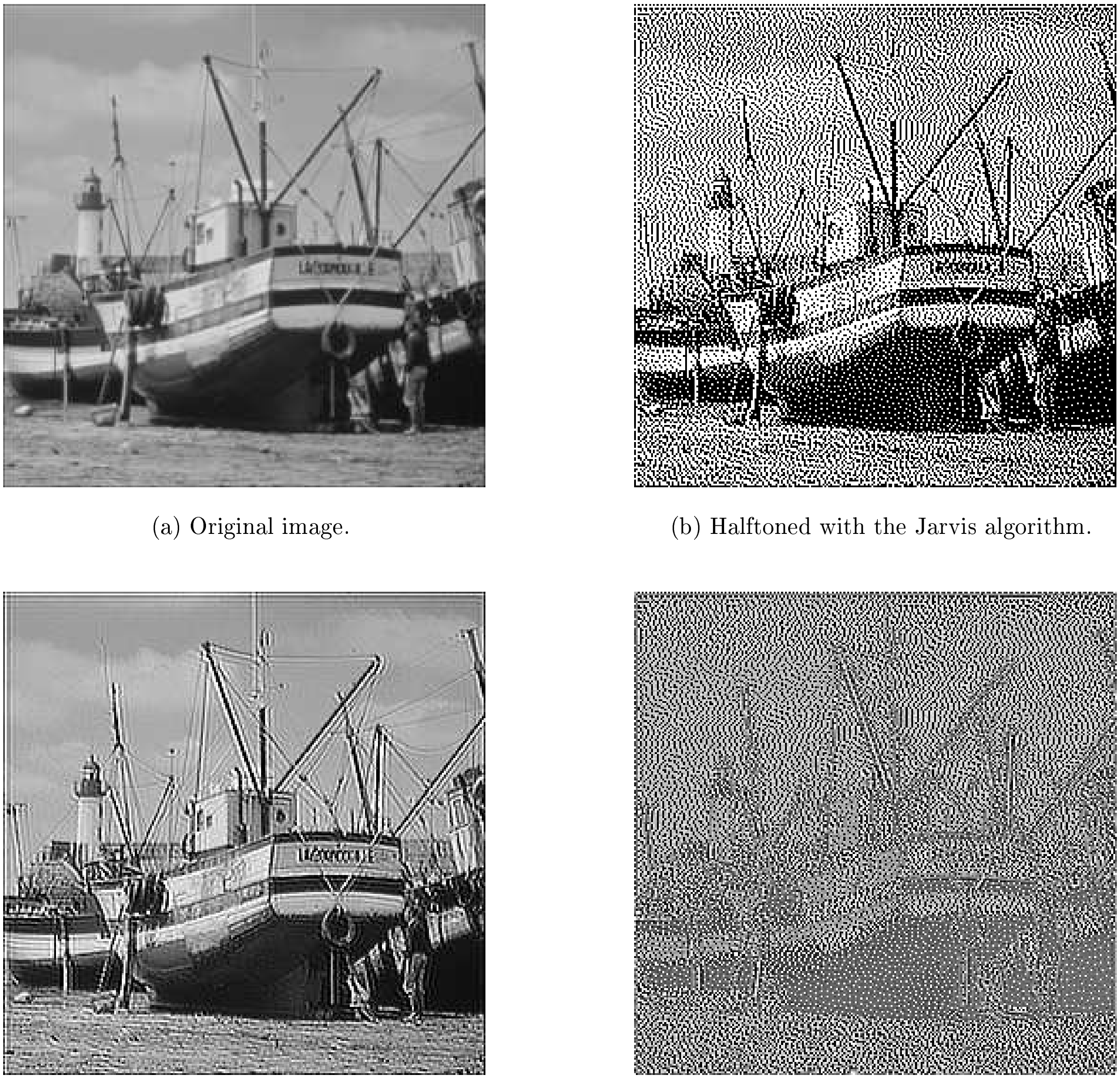
For a given error filter and image, we can measure K, allowing us to predict the STF and NTF. To measure K, we halftone an image using a particular scheme and save the error image $e(n)$. The input image to the quantizer is given by $x'(n) = y(n) - e(n)$. Since we have $x'(n)$ and $e(n)$, we can use (2) to compute K. We perform a linear least-squares fit of $x'(n)$ to $e(n)$. The linear coefficient is an estimate of $K-1$.

We conducted experiments to measure the value of the gain K for different error filters and images. We chose the original 4-tap error filter by Floyd and Steinberg, and the classic 12-tap filters by Jarvis, Judice and Ninke [2] (hereafter referred to simply as 'Jarvis' for brevity) and Stucki [3]. These results are shown in Table I. We see that K is approximately 2 for the Floyd-Steinberg filter, approximately 4.5 for the Jarvis filter, and approximately 4 for the Stucki filter. The value of K shows some variation between images. However, it appears from experiments that the value of K used to model the system is not critical.

By substituting the measured average value of K into (3), we can compute the STF. The coefficients of the error filter usually sum to one, that is, $H(z) = 1$ at DC. This ensures that the STF is unity at DC, and the NTF is zero. Because $H(z)$ is a lowpass function and appears in the denominator of the STF, the STF rises at higher spatial frequencies. Furthermore, the degree of high frequency boost increases with K. This accounts for the noticeable sharpening effect of all error diffusion schemes. The Floyd-Steinberg scheme has a maximum gain of around 4 at the Nyquist frequency, which is quite a mild boost, whereas the Jarvis scheme boosts the high frequencies by as much as a factor of 10.

IV. Evaluation of the Gain Model

By setting the values of K and $H(z)$ in the system of Figure 2, and by setting $q(n) = 0$, we can incur the edge



(a) Original image.

(b) Halftoned with the Jarvis algorithm.

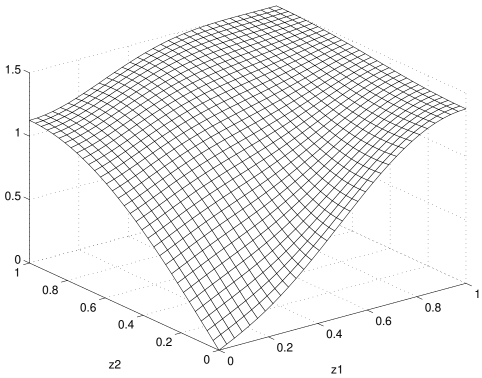
(c) Processed with the gain model, $K = 4.7$.

(d) Difference between (b) and (c).

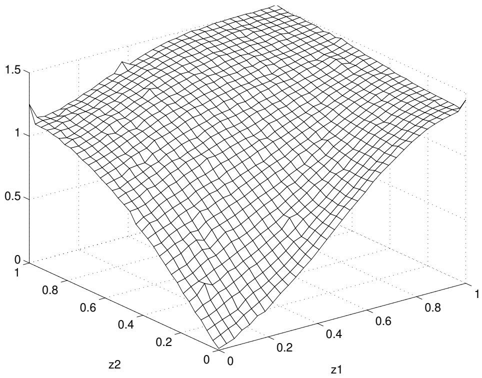
Fig. 3. Demonstration of the accuracy of the gain model using the *boats* image.

sharpening effect of a particular algorithm, without adding quantization noise. Figure 3(a) shows the original *boats* image. (Note: all four images in Figure 3 are of size 256×256 pixels to allow accurate reproduction.) Figure 3(b) shows the *boats* image halftoned using the Jarvis algorithm. There is a noticeable increase in sharpness. Figure 3(c) shows the image processed using the gain model, with $K = 4.45$ (the average value from Table I). The edge sharpening seen in this image is very similar to that of Figure 3(b), especially around the masts of the boats and at the boundary of the whole image, where ringing is evident. We further show the accuracy of the gain model by computing the difference between Figure 3(b) and Figure 3(c). We add a DC offset to make the image strictly positive. The result is shown in Figure 3(d). It is clear that the bulk of the residual is quantization noise, indicating that the gain model has accurately modeled the edge sharpening of error diffusion. Further experiments show that the residual image does not change appreciably for values of K within ± 1 of the measured value. Thus the averages shown in Table I can be used for most images without large error.

Our experiments show that, regardless of the measured value of K, the noise transfer function always behaves as if $K = 1$. That is, modeling the quantizer as a device which adds uncorrelated white noise predicts the noise shaping effect of error diffusion extremely well. We should probably not be surprised that the model has different values of K for signal and noise, since the generation of noise in the



(a) Predicted.



(b) Measured.

Fig. 4. Predicted and measured noise transfer functions for the Floyd-Steinberg algorithm.

quantizer is a separate process from the amplification of the signal. The need to deal with signal and noise separately has been noticed before [12].

For the Floyd-Steinberg scheme, the noise transfer function is given by

$$H(z_1, z_2) = \frac{1}{16}\left(z_1^{-1}z_2^{-1} + 5z_2^{-1} + 3z_1z_2^{-1} + 7z_1^{-1}\right) \qquad (4)$$

where z_1 and z_2 refer to horizontal and vertical frequency, respectively. This NTF is shown in Figure 4(a). Figure 4(b) shows the actual NTF, measured by averaging quantization noise spectra over many images. The close agreement between theory and experiment also holds for other error filters. The Floyd-Steinberg NTF exhibits several desirable properties, such as a zero at DC, smoothly increasing noise power with frequency, and a bias towards the diagonal direction, where the human visual system is less sensitive [10].

V. Summary and Future Work

The gain model for the quantizer predicts the edge sharpening seen in all error diffusion schemes. Although the value of the gain K varies somewhat with the input image, an average value can be used with little error. However, it is still necessary to measure K empirically using test images. We note that the larger error filters tend to have larger values of K; we hope to discover the connection between the filter coefficients and the value of K so that K may be predicted from the error filter alone. This will permit automatic error filter generation, since both the degree of image sharpening (which can be computed from $H(z)$ and K) and the noise transfer function (which can be computed from $H(z)$ alone) will be known. In conjunction with a model of the human visual system, such as that due to Mannos and Sakrison [13], error filters can be designed to optimize both the degree of sharpening and the noise transfer function simultaneously.

We plan to design error filters for oversampling halftoning algorithms using the model. The common practice of interpolating the original grayscale image and then halftoning it with a conventional technique is wasteful of computation time and produces sub-optimal results. The blurring effect of the interpolation filter can be counteracted by the sharpening effect of error diffusion; if the two filters are designed concurrently, higher quality images will be obtained at lower computational cost. An obvious application is printing, in which typically there are more pixels in the printed image than in the original.

References

[1] R. Floyd and L. Steinberg, "An adaptive algorithm for spatial grayscale," *Proc. SID*, vol. 17, no. 2, pp. 75–77, 1976.

[2] J. Jarvis, C. Judice, and W. Ninke, "A survey of techniques for the display of continuous tone pictures on bilevel displays," *Comp. Graph. and Image Proc.*, vol. 5, pp. 13–40, 1976.

[3] P. Stucki, "Mecca—a multiple-error correcting computation algorithm for bilevel hardcopy reproduction," Research Report RZ1060, IBM Research Laboratory, Zurich, Switzerland, 1981.

[4] R. Eschbach, "Reduction of artifacts in error diffusion by means of input-dependent weights," *J. Elec. Imag.*, vol. 2, pp. 352–358, Oct. 1993.

[5] K. Knox, "Threshold modulation in error diffusion on nonstandard rasters," *Proc. SPIE, Human Vision, Visual Processing, and Digital Display V*, vol. 2179, pp. 159–169, 1994.

[6] P. Wong and J. Allebach, "Optimum error diffusion kernel design," *Proc. SPIE/IS&T Symp. on Elec. Imag.*, Feb. 1997. Invited paper.

[7] K. Pohlmann, *Principles of Digital Audio*. New York, NY: McGraw-Hill, 3rd ed., 1995.

[8] R. DeFreitas, "The low-cost way to send digital data: delta-sigma modulation," *Electronic Design*, vol. 22, pp. 68–73, Jan. 1974.

[9] S. Norsworthy, R. Schreier, and G. Temes, eds., *Delta-Sigma Data Converters*. New York, NY: IEEE Press, 1997.

[10] B. Wandell, *Foundations of Vision*. Sunderland, MA: Sinauer Associates, 1995.

[11] K. Knox, "Error image in error diffusion," *Proc. SPIE, Image Processing Algorithms and Techniques III*, vol. 1657, pp. 268–279, 1992.

[12] S. Ardalan and J. Paulos, "An analysis of nonlinear behavior in delta-sigma modulators," *IEEE Trans. on Circuits and Systems*, vol. 34, pp. 593–603, June 1987.

[13] T. Mitsa and K. Varkur, "Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 301–304, 1993.