FM Halftone Screen Design Trading-off Grain and Mottle Performance

Jay Thornton and Bernd Kolpatzik*
Polaroid Corporation Image Science Lab
Cambridge, MA USA

Abstract

Many of the methods for design of FM (or "stochastic") halftone screens neglect the real world problems that determine whether a screen is practical for use in a graphic arts production environment or not. Problems often arise from FM screens that are too demanding on the production process because they require the reproduction of many tiny dots. Our approach fosters the clustering of dots into groups, which are more stable during the many production steps. The degree of clustering is controlled by a weighting parameter in an error function. Varying this parameter alters the coarseness of the dot pattern. The accompanying changes in the perimeter to area ratio of the clusters are comparable to varying the ruling of conventional halftone screens. Small clusters promote fine grain but, due to variation in the reproduction process across space, the small clusters are prone to large area non-uniformity (mottle) if the process control is not stringent. Large clusters are robust to mottle problems, but also appear grainy. We developed metrics for both grain and mottle in order to compare screen quality in the computer during development.

Another real world problem impacting screen design is the actual marking engine dot shape and characteristics. If these characteristics are neglected, the performance of a screen that has been optimized in a computer can be substantially degraded in the field. Incorporating the dot characteristics into the screen design process produces screens that live up to their performance expectations when used in production.

Introduction

With the introduction of computer processing, halftone imaging began to be done digitally using techniques that imitated the analog process and also new alternatives like error diffusion[1] and direct binary[2] search. Error diffusion and direct binary search are capable of producing higher image quality than digital screening, but require more processing time. Ulichney[3] coined the term "blue noise" to describe the (high-pass) shape of the radially averaged power spectrum characteristic of error diffusion. Mitsa and Parker[4] created a "blue noise mask," a screen that produced pleasing halftone textures that appeared similar to those produced by error diffusion. The industry has dubbed screens with a more random appearance "stochastic screens" or "FM screens" since they vary gray level via frequency modulation of many small dots rather than amplitude modulation of the size of large periodic dot clusters as in conventional digital screening.

FM screens have been welcomed into the graphic arts because of their robustness to subject moire and the freedom from screen angles they provide. This freedom makes them the halftoning method of choice for Hi-Fi color systems using more than the four colorants CMYK. At the same time, many printers have found FM screens to be finicky in production because of the tight process control necessary to control the tiny, dispersed dots characteristic of most FM screening methods. "2nd Order" FM screen is the term used to describe FM screens which attempt to group dots in clusters which reproduce more reliably, thereby making them more robust to process variation. Dalton[5] has pointed out that the spectral properties of band-pass FM screens are much more robust to the distortions caused by dot overlap, than are dispersed (blue noise) screens.

When Polaroid Corporation developed Dry Tech™ graphic arts film, we wanted a 2nd Order FM screen to go with it offering printers the quality and flexibility of FM screens along with the process forgiveness due to clustered dots. To develop such a screen and to optimize it for a particular film and film writer, we developed metrics allowing us to predict the performance characteristics of a screen and a means of adjusting those characteristics for the demands of a specific printing site. Below, we describe the metrics we created, a screen design procedure which allows controlled clustering of dots, and some experiments demonstrating that the approach does work in practice.

Performance Metrics

Determining an optimal degree of dot cluster can be viewed as a tradeoff between apparent uniformity at two different scales. Non-uniformity on a fine scale (less than 1 mm) that is apparent upon close inspection will be called "grain." Non-uniformity on a large scale (visible at arm's length) will be called "mottle." The tradeoff between grain and mottle can be understood in terms of the perimeter to area ratio of the dot clusters. Process variation tends to erode or dilate clusters along their perimeters. Since small clusters have larger perimeter to area ratios, large area uniformity, i.e. mottle, may appear worse in samples made up of many small clusters. On the other hand, as clusters get bigger, they begin to be resolved by the eye and cause a grainy appearance. Of course, other factors affect uniformity, and one might hope to improve both grain and mottle by improving the underlying screen design technique. To study the grain/mottle trade-off and optimize the screen design algorithm, we first required metrics linking the subjective impression of uniformity at different scales to quantities objectively measurable in film and prints.

Printer Dot Model

Grain and mottle are directly observable in most samples with the naked eye and any metrics must be consistent with those observations. But since we wanted to be able to predict performance properties of candidate screens without printing them, it was necessary to link the metrics to a model of the printer that describes the patterns actually produced on film in response to the logical bit-map sent to the printer. In our case, the printer was the DrySetter™, an internal drum imagesetter that writes separation films used to make printing plates. We assume that the DrySetter™ is configured to write at 2540 dpi (a 10 μm pitch). The DrySetter™ uses a high-powered laser to expose the DryTech™ film. Since this film is binary in its response, a simple printer model consists of (1) lowpass filtering the dot bitmap with gaussian blur to form a gray scale image representing the energy incident at each point on the film, and (2) thresholding the energy image to identify areas that will be rendered white in the final film.

Grain Metric

The grain metric consists of a human contrast sensitivity[6] weighted variance calculated from a nominally flat field (a halftone "tint"). First the input image, obtained by either writing to film and scanning or output from the printer model above, is convolved with the HVS spatial kernel. Next, the variance of the resulting gray scale image is computed from its histogram. Notice that the grain metric can be calculated from the ideal bitmap in the computer, or it can be calculated from images collected by scanning actual samples. The two measures may disagree in the case that the printer model is not sufficiently accurate. We found that the simple model of Figure 1 provided a good description of DrySetter™ and DryTech™ film.

Mottle Metric

In a mottle metric, we hoped to capture a measure of how robust a screen would be in the face of the natural variation found in all real world processes. Casting that variation into the simple printer model, many sources of variation (e.g. laser power variation from optical, mechanical or electrical sources or true variation in media threshold) can be well modeled by assuming that the threshold parameter is not constant. As the threshold varies, the large area percent dot (i.e. 1.0-transmittance) varies because the halftone dot clusters change their sizes. Our mottle metric consists of the slope of the function relating large area transmittance or reflectance to the model's threshold parameter (see Figure 2.) This metric can be easily computed from the histogram of the gray scale image representing the energy distribution falling on the film. If the histogram has a narrow peak in the vicinity of the threshold, then variation in that threshold will result in substantial variation in the average gray level. If the histogram is low and broad in the vicinity of the threshold, then average gray level will be only slightly affected by variation in the threshold parameter. The mottle metric captures the intuitive explanation based on the perimeter to

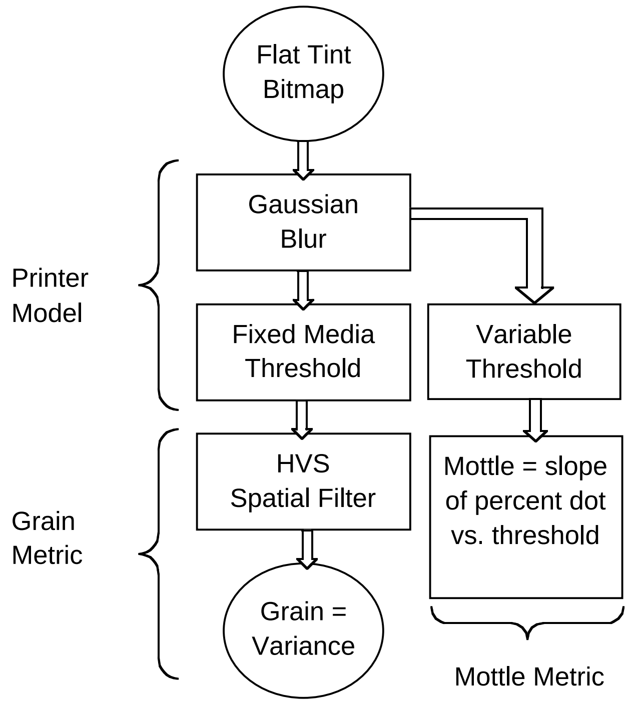

Figure 1. Computing predicted grain and mottle for a binary bitmap representing a halftoned flat field.

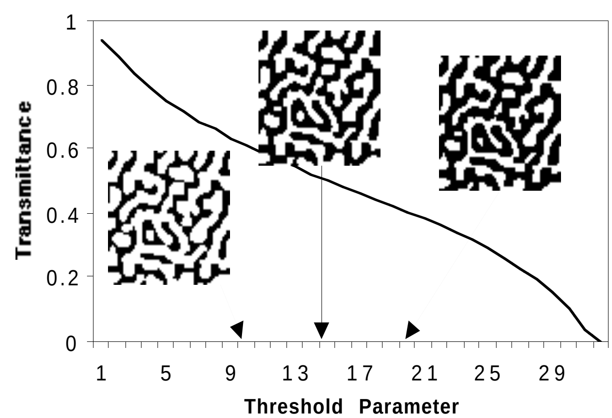

Figure 2. The mottle metric is the slope of the cumulative histogram of the gray scale energy image taken at the true threshold value of 15. The inset images show how a 50% bitmap would print if threshold or exposure varied.

area ratio because it measures the change in total white area as the threshold parameter varies incrementally. Since almost all that change occurs at the clusters' perimeters, dot configurations with reduced perimeter to area ratios produce smaller mottle metric values. Using the mottle metric, for any FM screen, one can establish a corresponding conventional screen ruling with equivalent process robustness.

Halftone Screen Design

Ulichney[7] described the "void-and-cluster algorithm" to generate a digital halftone screen. The screen design procedure we developed is most easily described as a modification of Ulichney's, so his procedure will be reviewed first. The objective is to design a digital screen function, t(n,m), of size N by M pixels to convert contone images into binary output images. The first step for void-and-cluster is the design of a prototype blue noise pattern for a fixed gray level, g. A random binary pattern, b(n,m), is generated and circularly convolved with a (typically gaussian) lowpass filter, lp(n,m). Subtracting the average gray value, g, from the filtered output signal yields the (signed) error function in Equation 1.

$$e(n,m)=b(n,m)*lp(n,m)-g \qquad (1)$$

Next, the algorithm searches for the position of the minimum of e(n,m), which determines the location of the largest cluster of black pixels. At this location, a black pixel is flipped to a white pixel. Subsequently, the modified pattern is filtered again and the position of the maximum of e(n,m) is found. This location determines the largest void of black pixels. Here, a white pixel is flipped to black. These two steps are repeated as illustrated in Figure 3, until flipping the pixel from the largest cluster generates the largest void. Then the procedure is terminated. The final pattern, b(n,m), is used to initialize the threshold array, t(n,m). To complete the threshold array for all other gray levels, he adds (subtracts) one black pixels at a time at the largest voids (clusters) of the prototype binary pattern, applying the same filtering process. In each step he updates t(n,m) with the new average gray value of the binary pattern, at the corresponding location, (n,m), where b(n,m) has been modified. After all gray levels have been generated, and the changes have been recorded, t(n,m) contains values for all gray levels of the tonescale. The screen produces patterns with highly dispersed dots and blue noise spectral properties for almost all gray values.

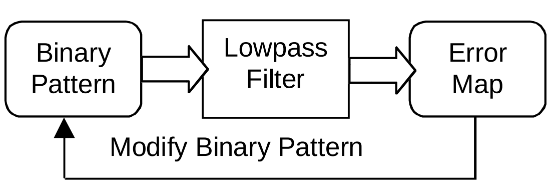

Figure 3. Ulichney's void-and-cluster iteration.

Yao and Parker[8] describe a design algorithm similar to void-and-cluster, but one which is faster and achieves a slightly more isotropic pattern. These improvements are partially due to a more sophisticated selection of the lowpass filter, lp(n,m), one which relates it to the visual contrast sensitivity function. They are also due to employing a procedure we also adopted in which multiple dots are changed simultaneously during the optimizing iterations.

Model-based screen design

The algorithms above generate extremely dispersed dot patterns dominated by isolated dots and are therefore sensitive to process variability. In our screen design algorithm, we wanted to make two improvements on the techniques above. First, we wanted to cluster dots so that the screen would be more robust to process variation. Second, we wanted to incorporate the dot model into the screen design.[9] This second factor can be quite important. The "voids" and "clusters" identified by the HVS lowpass filtering are the microscopic features that create visible grain. Filtering directly on the bitmap implicitly assumes dots are perfect squares that pack to cover the surface without having any overlap. In practice, dots are typically round and require substantial overlap to cover the surface. These differences mean that patterns optimized for low grain in the computer may not be low grain in practice. By inserting a dot model operation on the bitmap prior to the lowpass filtering in the screen design iteration, the optimization holds up under real world conditions.

Grain/Mottle knob

To foster clustering of dots, we added another term to the error function, a term we call a neighborhood filter. This new term is a highpass filter, hp(n,m), which gives a measure of how similar a pixel is to its neighbors. A reasonable neighborhood filter can again be based on a gaussian, one which is spatially narrower than the HVS lp filter above, by subtracting the (unity volume) gaussian from 1.0. This term in the error function is near 0 when pixels are surrounded by similar neighbors (i.e. in a cluster), but is near +/- 1.0 when pixels are surrounded by unlike neighbors (i.e. isolated dots.)

The two components of the error function are combined with a weighting factor, λ, allowing control of the emphasis to be placed on clustering. Thus the total error function is:

$$e'(n,m) = PM[b(n,m)]*lp(n,m)-g +\lambda[b(n,n)*hp(n,m)] \qquad (2)$$

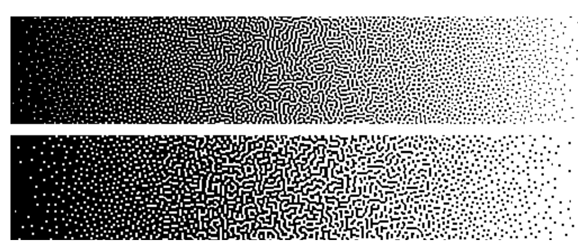

Figure 4. Variation in coarseness of the screen is determined by the weighting parameter, λ.

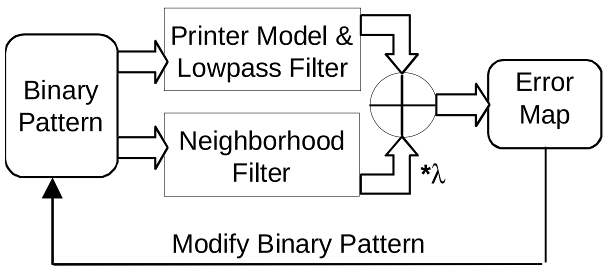

Figure 5. Modified iteration including dot model, neighborhood clustering term, and weighted summation in error function.

λ can be thought of as the grain/mottle trade-off knob. When λ is large, the resulting screen is coarser (contains larger clusters) and is more resistant to mottle. When λ is small the resulting screen is finer but more demanding of process control. Figure 4 shows wedges halftoned using screens based on different values of λ. Figure 5 diagrams the iteration in the optimization.

Notice that, while the new error function more-or-less directly incorporates the grain metric, the linkage to the mottle metric is indirect. The mottle metric applies to large areas like the entire threshold array. The neighborhood filter is a local measure that directly encourages clustering of dots. Larger clusters are associated with lower mottle metric values, but the mottle metric itself is not optimized.

Experimental Results

Grain metric validation

To verify that our proposed grain metric did in fact correlate with subjective perception of grain, we performed a simple psychophysical scaling test. Film samples of 50% tints were written based on experimental screens we had created. In addition, three commercially available screens samples were included. We performed a category rating experiment in which 18 observers assigned a value between 1 (lowest grain) and 10 (highest grain) to each sample. The observers were all shown the same reference samples representing grain values of 1 and 10. The observers saw the samples in different random orders. The measured grain value was taken to be the mean category rating. For those screens we had developed, we also calculated the grain metric by applying the dot model to a 50% tint bitmap and then performing the grain metric calculation of Figure 1.

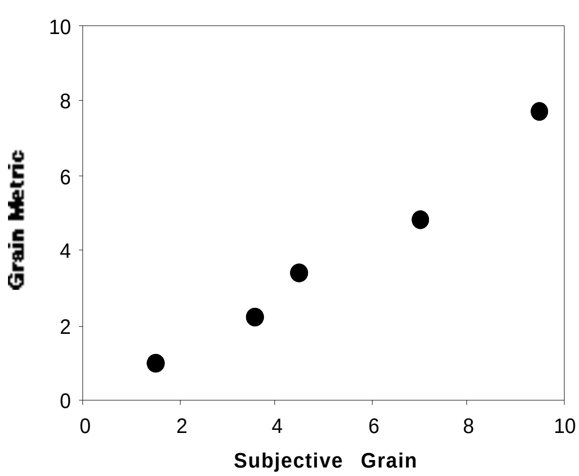

Figure 6. The grain metric is well correlated with subjective measurement. Our new screen had the lowest grain.

The graph in Figure 6 shows that there was excellent agreement between the subjective and objective measurements. This agreement bolsters both the grain metric itself and the dot model used to get from the bitmap to the simulated halftone image. When we scanned the films, and then applied the same grain metric to those digital images, we also observed good agreement between subjective and objective measurements. Among all the samples, both commercial and experimental, the one with the least grain in this experiment was our new screen.

Screen mottle performance

To evaluate the impact of dot modeling and clustering on the threshold array produced by void-and-cluster , we calculated both the grain and the mottle for 50% tints produced with and without our modifications (see Table 1). For both metrics, smaller values are better.

	Grain	Mottle
Void-and-cluster	4.9	2.7
Polaroid Screen	1.0	1.0

Table 1. Comparison of Void-and-cluster with Polaroid screen incorporating dot modeling and clustering.

The results have been normalized to 1.0 for the Polaroid screen. The mottle result should not be surprising given the high dpi (10 μm dots) and given that the original void-and-cluster procedure generates an extremely dispersed dot pattern with many isolated dots. Ordinarily, the clustering in the Polaroid screen would come at the expense of increased grain. In this case the increased grain has been more than compensated for by grain reduction due to the dot model. If the Polaroid screen in Table 1 were compared to another with a different value of λ, then the one with less grain would necessarily have more mottle. The tabled results are entirely computer calculated (and therefore somewhat circular, since the same model was used in optimizing the screen and evaluating the results,) however printed samples showed roughly the same result and we have repeatedly demonstrated good agreement between computer modeled and printed samples.

To evaluate the mottle performance of the Polaroid screen in the field, we tested the large area uniformity of samples created with the Polaroid screen and two commercial screens. Large (approximately 3'x4') flat fields of a 50% tint were prepared on the same imagesetter on the same day. The system specification requires that a 50% tint vary no more than +/- 1% dot. With a graphic arts densitometer, we measured the dot percentage using a 2mm aperture in a 3x4 grid pattern spread evenly over each of the sample films (i.e. on a rectangular grid with ~10" between sample points.) The result of these measurements showed that only the Polaroid screen, could pass the 1% specification (see Table 2.) Since these same commercial screen samples had been used in the grain experiment, the Polaroid screen had proven to have both better mottle properties and better grain properties.

Polaroid Screen	Fast Scan Position			
Slow Scan Position	0.8	0.0	0.1	0.1
	0.4	-0.1	-0.2	-0.4
	0.3	-0.1	-0.4	-0.4
Commercial Screen A				
	2.3	0.0	-1.1	0.0
	1.7	-0.6	-0.9	0.0
	1.5	-1.3	-1.0	-0.6
Commercial Screen B				
	4.3	-0.6	-1.9	-0.5
	2.6	-2.1	-2.4	-0.9
	3.2	-1.5	-0.7	0.6

Table 2. Deviations from 50% dot measured in a 3x4 grid. Acceptable performance is +/- 1% dot.

Conclusions

By grouping dots into clusters, we produced a FM screen that was significantly more robust to low frequency, mottle non-uniformity. While forming dot clusters is expected to increase grain, we counteracted that tendency by applying a model of the printer dot formation process. The application of this model during screen design anticipated and compensated for distortions in the printer that would have destroyed the optimality the computation was meant to achieve. The result is a screen with both superior grain and mottle performance. By varying a weighting parameter in the error function, we were able to vary the trade-off between grain and mottle and create a family of screens with properties matched to different applications and environments. The performance of these screens was properly predicted by metrics we devised. These metrics allowed us to do more of the system testing in the computer and rely less on laborious field testing of many screen design variations.

References

1. R.W. Floyd, L. Steinberg, "An adaptive algorithm for spatial grey scale," *Pr. SID*, vol **17/2**, p. 75-77, 1976.
2. M. Analoui and J.P. Allebach, "Model-based halftoning using Direct Binary Search," *Proc. Of the 1992 SPIE/IS&T Symp. On Electronic Imaging Science and Technology*, vol. **1666**, p. 96-108, 1992.
3. R.A. Ulichney, "Dithering with blue noise," *Proc. IEEE*, vol. **76**, p. 56-79, 1988.
4. T. Mitsa, K. J. Parker, "Digital halftoning technique using a blue-noise mask," *J. Opt. Soc. Am. A*, vol. **9**, p. 1920-1929, 1992.
5. J. Dalton, "Perception of binary texture and the generation of stochastic halftone screens," *Proc. SPIE, Human Vision, Visual Proc., Digital Display VI*, vol. **2411**, p. 207-220, 1995.
6. O. H. Schade, "Optical and photoelectric analog of the eye," *J. Opt. Soc. Am.*, vol. **46**, p. 721-739, 1956.
7. R. A. Ulichney, "the void-and-cluster method for dither array generation," *Proc. SPIE, Human Vision, Visual Proc., and Digital Display IV*, vol. **1913**, p. 332-343, 1993.
8. M. Yao, K. Parker, "Modified approach to the construction of a blue noise mask," *J. Electronic Imaging*, vol**. 3/1**, p. 92-97, 1994
9. T. N. Pappas, D. L. Neuhoff, "Model-based halftoning," *Proc. SPIE, Human Vision, Visual Proc., Digital Display II*, vol. **1453**, p. 244-255, 1991.

*Dr. Kolpatzik is currently employed at Siemens Corporation.