Stoclustic (Stochastic Clustered) Halftone Screen Design

Shen-ge Wang
Xerox Corporation, Digital Imaging Technology Center
Webster, New York

Abstract

A new optimization technique has been developed for designing various digital halftone screens. A single-value merit function, which evaluates all levels of the halftone output, is applied throughout the entire optimization process. By randomly swapping thresholds of the halftone screen and simulated annealing, an optimal solution can be approached automatically. Using this design method with different specifications and constraints, we have successfully designed stochastic screens, ordered-cluster screens, stoclustic (stochastic clustered) screens, as well as digital watermarks.

Introduction

Most printers rely on halftoning for reproducing images with continuous tones, because many printing processes are inherently binary or with a limited number of specific tones. The most common halftoning techniques include dithering and error-diffusion. The dithering method is implemented by storing a halftone screen, or dither matrix, of thresholds in memory and comparing these thresholds with the image data on a pixel-by-pixel basis to create the binary output signals. Based on the appearance of the halftone output, halftone screens can be categorized as clustered screens, stochastic screens, line screens and etc. Since the quality of halftone images is directly affected by the choice of screens, the halftone screen design has been always one of the most important issues in digital printing technology.

There are many aspects should be considered in halftone screen design. For example, screen frequency and angle, dot size and shape, dot growth, dot gain and other geometrical aspects, as well as non-geometrical aspects, such as color stability, color gamut and human visual response[1-3]. Practically, it is impossible to design a halftone screen, which meets all quality requirements; therefore, trade-off between some aspects of consideration has to be made in the screen design.

Up to date, most clustered halftone screens used in digital printing industry were "hand-tuned" by experts based on their experience and trial-and-error results. Whether or not these screens provide satisfied image quality, the design procedure is usually time-consuming and there is no guarantee of finding an optimal solution. On the other hand, most stochastic screen designs are computer-aided. For example, Parker et al use Fourier transform technique to optimize threshold values for stochastic screens, or Blue Noise Masks.[4,5] As their approach, the spatial distributions of desired outputs from a stochastic screen are specified as Blue-Noise spectra in the frequency domain and an iteration technique is used to optimize the threshold screen in a level-by-level sequence. In the design process of stochastic screens, little human interaction is involved except initial setting and final evaluation. Relative to the average size of most clustered screens, some considerably very large stochastic screens have been successfully designed. However, two major factors of above design method for stochastic screens prevent a direct extension to design of clustered screens. First, the merit functions for optimizing stochastic screens are defined in the Fourier transform domain, but many geometrical specifications for clustered screens can not be easily translated into the frequency domain. Second, the optimization process is an irreversible sequence. All threshold values associated with levels in the earlier design are fixed for the later optimization. The freedom is getting smaller and smaller while the design is reaching the end. In fact, the results of designing relatively small stochastic screens using Blue-Noise-Mask technique are not even close to the optimal ones.

In this paper we present a new versatile method for halftone screen design. By randomly swapping a pair of threshold values in the halftone screen and calculating the difference of the merit function, an optimal configuration can be approached in conjunction with a simulated annealing process. The new approach provides the following important features:

1. The design process is to determine an optimal arrangement of all thresholds in the halftone screen simultaneously, not in a level-by-level sequence. All threshold values are subject for change until the entire optimization is finished.
2. The merit function for optimization is defined in the spatial domain. Hence, most geometrical specifications for desired halftone screens could be mathematically described and combined into a single merit function. Also, the calculation involved in optimization could be much less than designing a similar screen with a merit function defined in the frequency domain.

Using this new design technique, we have successfully designed various halftone screens including stochastic, clustered, stoclustic screens, as well as halftone screens with special features, such as digital watermark screens. In the following sections we will describe the new design method and demonstrate a few examples of its applications.

Screen Design Procedure

The design procedure depends on specific requirements of each halftone screen being designed. Here, we only discuss some general consideration in the optimization process.

The overlap between adjacent pixels is an important issue in halftone screen design. Our approach is completing the design in two steps: first, optimizing a halftone screen for all levels regardless of overlapping; then, using overlap-ping calibration to map the optimized result into a halftone screen with desired gray levels. For a halftone screen with N elements, it is possible to have N different threshold values and to generate halftone images simulating $N + 1$ different gray levels. For example, when we design a 128×128-element screen, we can optimize the screen in all 16385 gray levels. The result of optimization will be device indepen-dent. A linear or non-linear mapping from the optimized design into a device-dependent halftone screen for 8-bit, or 256-level, inputs can be conducted by several overlapping correction methods. For example, an algorithm-independent calibration method based on the 2×2 concept is described in the references 6 and 7 and will not be discussed here.

The Merit Function

Now, consider a halftone screen with N different threshold values, from 1 to N. The binary output $B(x, y)$ is determined by the following dithering rule:

$$\begin{aligned} B(x,y) &= 1 \qquad \text{if } G(x,y) \le T(x,y) \\ B(x,y) &= 0 \qquad \text{elsewhere,} \end{aligned} \tag{1}$$

where x, y are the spatial coordinates, T is the threshold value at (x, y) and G is the gray-level input ranging from 0 to N. The value one of the output B represents a white pixel and zero, a black pixel. Without dot overlapping, a binary halftone pattern corresponding to a constant input level G will have G black pixels and $N - G$ white pixels within the area defined by the halftone screen. The visual appearance of this halftone pattern depends on whether the black pixels or the white pixels are minorities. If the white pixels are, i.e., $G < N/2$, the appearance of the output should be based on the distribution of white pixels. Otherwise, the appearance should be on the distribution of black pixels. We use a merit function combining evaluations of all levels, thus, a half of the merit is on distributions of black pixels and another half is on white.

The quantitative measure of visual evaluation on halftone outputs varies with specifications of the desired screens. For example, it is visually pleasant to have all minority pixels "evenly" distributed on the output of a stochastic screen. Mathematically, it could be translated as maximizing the summation of all distances between any two minorities.

In general, the contribution to merit function from each gray level can be written as a function of the gray level and locations of all minorities appearing at this level. Since being minority or majority is determined by the dithering rule given by Eq. (1), in consequence, the contribution Q at level G can be written as a function of G and locations of corresponding thresholds in the halftone screen:

$$\begin{aligned} Q(G) &= Q(G; x_1, y_1, x_2, y_2; \ldots, x_G, y_G), \\ &\qquad \text{if } G < N/2, \\ Q(G) &= Q(G; x_{G+1}, y_{G+1}; x_{G+2}, y_{G+2}; \ldots x_n, y_n), \\ &\qquad \text{elsewhere,} \end{aligned} \tag{2}$$

where x_T, y_T are coordinates of a threshold T.

The entire merit function M can be defined as a sum of contributions $Q(G)$ from all levels:

$$M = \sum_{G=0}^{N} w(G) \cdot Q(G), \tag{3}$$

where $w(G)$ is a weighting function of G. The objective of optimization is to maximize (or minimize) the single-value merit function M given above. Since the merit is used throughout the entire optimization process, the design can reach the best balance of all gray levels and the best compromise between given constraints in size, shape and other geometrical requirements. This advantage is especially significant for automatic design of relatively small halftone screens.

Swapping and Simulated Annealing

Simulated annealing is a standard optimization technique and it has many different variations and modifications.[8] In our application to halftone screen design, we use the temperature to control the acceptance of a random swapping. The acceptance is stochastically set by the temperature and compared with the change of the merit function due to swapping. When the temperature is high, even a negative result could be accepted. So, the optimization would not easily fall into a local minimum. The temperature is gradually reduced during the optimization process and the final result is saved when the temperature drops to the low end.

The flow chart in Fig. 1 demonstrates the basic optimization process for our halftone screen design. At the beginning, all threshold values of the halftone screen are randomly set or copied from a previous design. The control temperature τ is set to the starting temperature $\tau 1$. The merit function defined for the specific halftone screen is calculated.

The optimization loop begins with a random selection of two threshold values, $T1$ and $T2$. After swapping the selected thresholds, the change of the merit function due to swapping is calculated. In general, the entire merit function M does not need recalculated, instead, only contribution of levels between $T1$ and $T2$ does. If the change of the merit function meets the acceptance set by the current temperature of simulated annealing, we keep the swapped order, lower the temperature τ, and proceed to the next swapping loop. Otherwise, we restore the order before swapping and move on to the next loop. This iteration process is continued until the temperature reaches a desired low level $\tau 0$ and the final configuration of the screen is saved. Of course, if the result is not a satisfactory, we can reset the temperature and start the optimization over again.

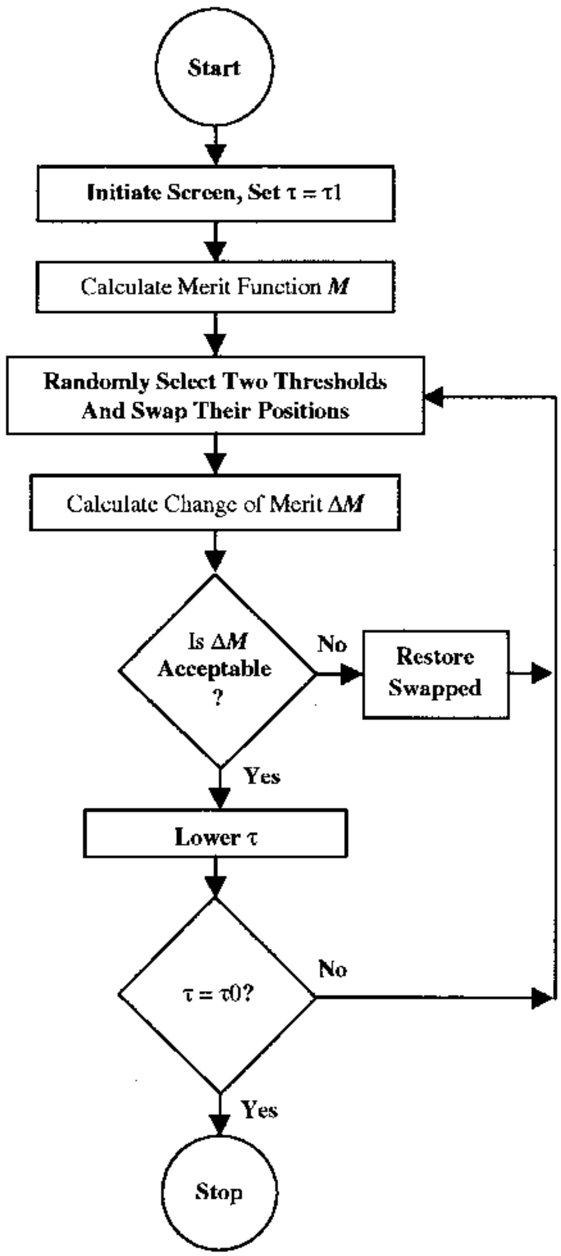

Figure 1. Optimization process for halftone screen design.

For many halftone screens, it is possible to simplify Equation (2), the merit function $Q(G)$ at each gray level G, to a summation form:

$$\begin{aligned} Q(G) &= \sum_{i=1}^{G} \sum_{j=1}^{G} q(G; x_i, y_i; x_j, y_j) \qquad \text{if } G < N/2 \\ Q(G) &= \sum_{i=G+1}^{N} \sum_{j=G+1}^{N} q(G; x_i, y_i; x_j, y_j) \qquad \text{elsewhere} \end{aligned} \tag{4}$$

where q only depends on G and locations of a pair of thresholds. Under such circumstances, calculation of the merit function could be reduced. For example, if both selected thresholds are less than the mean, i.e., $T1 < T2 < N/2$, the change of the merit function due to swapping, ΔM, can be written as the following equation:

$$\begin{aligned} \Delta M = \sum_{G=T1}^{T2} w(G) \sum_{i=1}^{G} \Big[\big(G; x_{T2}, y_{T2}; x_i, y_i\big) \\ - q\big(G; x_{T1}, y_{T1}; x_i, y_i\big) \Big]. \end{aligned} \tag{5}$$

For a large halftone screen, the saving of computation, comparing with calculation using merit functions defined in the Fourier transform domain, could be tremendous. Using the new technique, we are able to optimize a stochastic halftone screen with 16K elements and 16K gray levels within a day on a SUN Ultra1 Station. Surely, the optimization can be conducted either for all levels, determined by the screen size, or for selected levels to accelerate the process. As mentioned earlier, we separate dot-overlapping correction from the optimization process in a complete screen design. An advantage of optimizing more than 256 levels is that one optimization result can be calibrated for different output devices without losing of the levels.

Another significant advantage of this new design method is that any output of a design can be saved as an intermediate result, which could be used in the future for further improvement or modification.

Application Examples

Stochastic Screen

An idealized stochastic screen would create such output patterns with any constant input G that all minority pixels were "evenly" separated. Mathematically, above statement is equivalent to minimizing a merit function $Q(G)$ given by Eq. (4), in which $q(G; x_i, y_i; x_j, y_j)$ is a function of the distance between (x_i, y_i) and (x_j, y_j), or

$$q\big(G; x_i, y_i; x_j, y_j\big) = \frac{1}{\big(x_i - x_j\big)^2 + \big(y_i - y_j\big)^2}. \tag{6}$$

A merit function M, given by Eq. (3), summarizes $Q(G)$ of all levels with a weighting function $w(G)$ defined by

$$\begin{aligned} w(G) &= N/G, && \text{if } G < N/2; \\ w(G) &= N/(N - G) && \text{elsewhere.} \end{aligned} \tag{7}$$

Using the defined merit function above, we can design stochastic halftone screens in different sizes and shapes. Since a halftone screen is applied repetitively in dithering process, additional consideration of the periodicity should be applied to the function q given by Eq. (6). As a matter of fact, if we utilize a look-up-table with pre-calculated function q between pixels, most calculation involved in the optimization is simple addition and the simulated annealing can be proceeded very quickly.

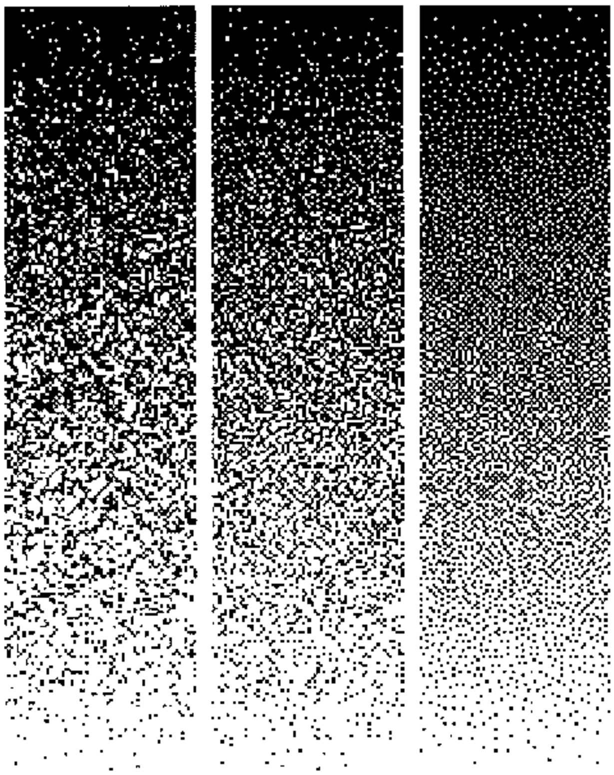

Figure 2. Halftone patterns generated by three 64x64 stochastic halftone screens, respectively. a) The left column, by the initial screen of an optimization. b) The middle one, by a screen saved after 5,000 swaps. c) The right one, by a screen optimized in ten minutes with 1,000,000 swaps.

Figure 2 shows three halftone patterns generated by three 64 × 64-element stochastic screens, respectively, to simulate a gray-scale sweep. The three stochastic screens are saved at different stages of a design process conducted on a SUN Ultra1 Station. Fig. 2a is created by a screen with the initial random setting in optimization. Fig. 2b is by a stochastic screen saved after a few second optimization with 5,000 swaps. At the last, Fig. 2c is the result optimized in ten minutes with approximately one million swaps.

Stoclustic Screen

Here, we use the word "stoclustic", or stochastic clustered, loosely. A stoclustic screen might be a halftone screen with many clusters stochastically centered; a screen with fixed centers and stochastic dot growth; a screen with clusters stochastically centered and shaped; or a screen with other possible stochastic features associated with dot size, dot shape and/or dot locations.

For example, we may design a stoclustic screen with cluster centers specified in arbitrary floating numbers. Since it is desirable to see all clusters showing the same shape and the same size at a constant input level, we can mathematically define a merit function $Q(G)$ for a level G as

$$Q(G) = \sum_{i=1}^{G} (x_i - X_{ci})^2 + (y_i - Y_{ci})^2, \quad \text{if } G < N/2;$$
$$Q(G) = \sum_{i=G+1}^{N} (x_i - Y_{ci})^2 + (y_i - Y_{ci})^2, \quad \text{elsewhere} \qquad (8)$$

where pixel (x_j, y_j) is the location of a threshold value $T = i$ and (X_{ci}, Y_{ci}) is the closest cluster center to pixel (x_j, y_j). The objective of optimization is to minimize the merit function M given by Eqs. (3) and (8). A demonstration design is shown by Fig. 3, which is a halftone pattern generated by a stoclustic screen with varying line frequencies.

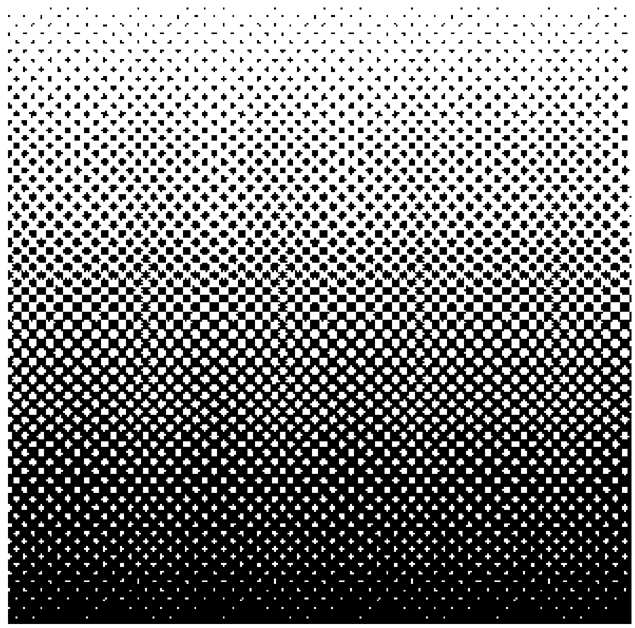

Figure 3. Halftone pattern generated a stoclustic halftone screen with varying line frequencies.

Digital Watermark

Another interesting demonstration for showing the versatility of our screen design method is the digital watermark. Invisible digital watermarks can be incorporated in printed halftone images using specially designed stochastic screens.[9] For example, consider a watermark halftone screen, which configuration is shown in Fig. 4. The 180 × 90-element screen is divided in two parts: the left half consists of a watermark region W and the outside region S; the right half consists of regions W' and S', which are identical to the left half in shape and size. All elements of this halftone screen are engaged in pairs, whether located inside the watermark regions or outside the watermarks. Any two engaged elements are separated horizontally by X_s, a half of the width, or 90 pixels in this example.

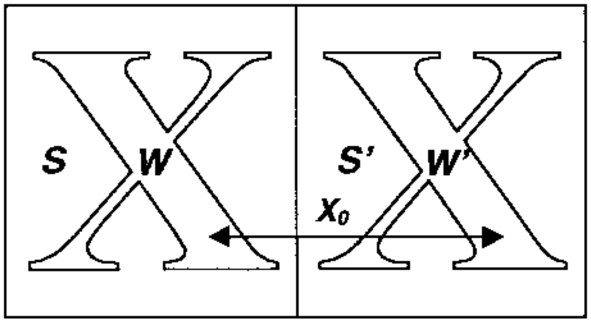

Figure 4. Configuration of a watermark halftone screen. The two watermark regions, W and W', are identical in shape and separated by a distance X_0 equal to a half of the screen width.

In the design process, besides the merit function defined for a standard stochastic screen, we also specify two additional constraints:

$$T(x, y) = T(x + X_S, y), \quad \text{if } (x, y) \in S;$$
$$T(x, y) = N - T(x + X_S, y), \quad \text{if } (x, y) \in W, \qquad (9)$$

where N is the total number of elements, or the maximal gray level. Under these constraints, the halftone output generated by the left half of the watermark screen is highly correlated with the output by the right half. When a constant input is equal to the mean $N/2$, the binary output generated by this halftone screen will have the following conjugate relations:

$$B(x, y) = B(x + X_S, y), \quad \text{if } (x, y) \in S;$$
$$B(x, y) \neq B(x + X_S, y), \quad \text{if } (x, y) \in W. \qquad (10)$$

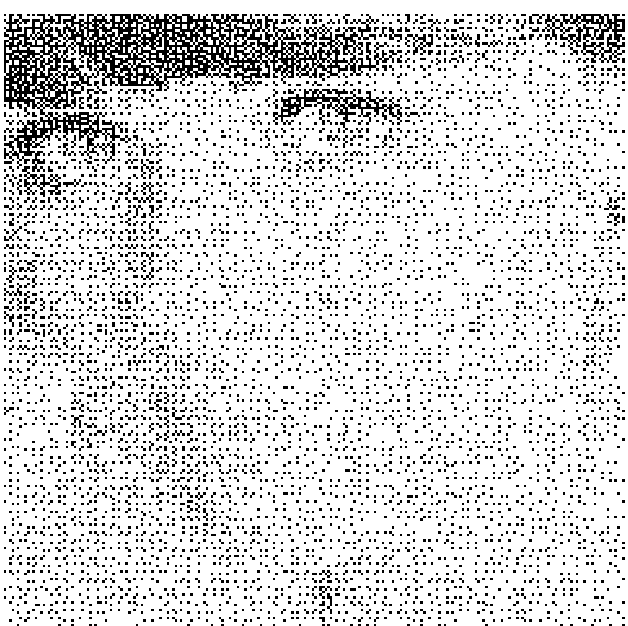

Figure 5. Halftone image generated by a stochastic halftone screen incorporated with invisible watermarks.

Figure 6. Watermarks shown by shifting and overlaying the halftone image in Fig. 5 with itself.

If the binary output pattern generated by the left-half screen is overlaid with the output by the right half, a maximal contrast between the watermark region and the background will be observed. Since the entire screen is designed under the same requirements as general stochastic screens, the general image quality will not be sacrificed by incorporating the watermark information into output images.

Figure 5 is a halftone image generated by a watermark screen. The watermarks become visible by shifting and overlaying the image with itself. The overlay result is shown by Fig. 6.

Conclusion

In our proposed screen design process, the merit function combines evaluation of all levels of the halftone output. By random swapping thresholds and simulated annealing, an initial halftone screen can be gradually reaching an optimal arrangement. Since the evaluation is based on geometrical specifications in the image domain, this optimization technique can be easily applied to designing new types of halftone screens. List a few possibilities:

- screens with special shapes;
- screens with varying line frequencies;
- screens designed with consideration of human visual response;
- screens with special correlation requirements.

Because we translate the task of screen design into an optimization problem with a defined merit function and freedom of rearranging threshold values in the screen, many other standard optimization techniques can be also adapted for halftone screen design.

References

1. R. Ulichney, *Digital Halftoning*, MIT Press, 1987.
2. P. G. Roetling and T. M. Holladay, "Tone Reproduction and Screen Design for Pictorial Electrographic Printing," *J. App. Photo. Eng.*, **5**, 179–182 (1979).
3. W. L. Rhodes and C. M. Hains, "The Influence of Halftone Orientation on Color Gamut and Registration Sensitivity," *Soc. Imag. Sci. and Tech. 46th Annual Conference*, 180–182 (1993).
4. T. Mitsa and K. J. Parker, "Digital Halftoning Technique Using a Blue-Noise Mask", *J. Opt. Soc. Am. A*, **9**, 1920–1929 (1992).
5. M. Yao and K. J. Parker, "Modified Approach to the Construction of a Blue-Noise Mask", *J. Elect. Imaging*, **3**, 92–97 (1994).
6. S. Wang, K. T. Knox and N. George, "Novel Centering Method for Overlapping Correction in Halftoning", *Proc. IS&T 47th Annual Conference*, 482–486 (1994); ***see Recent Progress in Digital Halftoning, Vol. I, pg 56***.
7. S. Wang, "Algorithm-Independent Color Calibration for Digital Halftoning", *Fourth Color Imaging Conference*, 75–77 (1996); ***see pg. 411, this publication***.
8. E. Aarts and P. van Laarhoven, *Simulated Annealing: Theory and Practice*, John Wiley and Sons, New York, 1987.
9. K. T. Knox and S. Wang, "Digital Watermarks Using Stochastic Screens", *Proc. Electronic Imaging Conference*, 316–322 (1997).

❋ Previously published in *IS&T's NIP13 Conference Proc.*, pp. 516–521, 1997.