Optimizing Halftone Masks with Genetic Algorithms and Printer Models

J.S. Arney, P. G. Anderson, Sunadi Gunawan, and Kenneth Stephens
Center for Imaging Science
Rochester Institute of Technology
Rochester, New York

Abstract

Techniques are described for determining optimum halftone masks for electrophotographic laser printers. Earlier techniques were based primarily on up-shifting the noise power of the halftone pattern in order to minimize visual granularity.[1,2] This strategy worked well with low addressable printers up to the mid 1990s. However, current addressability of EP printers is so high that too much up-shifting introduces printer instabilities that can introduce different kinds of granularity and/or mottle. Thus, the optimum halftone mask for image quality should have noise power concentrated within at frequency band between the edge of human vision and the edge of printer stability. The authors have developed search techniques for examining permutations of halftone masks of size HxW. The techniques are based on a strategy called a genetic algorithm (GA).[3]

Introduction

The search for practical improvements in print image quality (IQ) has always been an iterative process involving trial and error printing, evaluating, modifying, printing, etc. This iterative search is summarized in Fig. 1. A test pattern, (A), is selected and printed with the printing device of interest. The result is a printed sample (B). Step (II) in the process is the evaluation of image quality. Step (III) involves making changes to the system.

Each loop of iteration in the optimization of Fig. 1 can take hours or days, depending on the methods used to evaluate printed samples and to select system improvements. The objective of the current project was to shorten the time for each iteration to milliseconds in order to increase by several orders of magnitude the number of halftone screens one can test. To accomplish this, Steps (I), (II), and (III) were carried out as software simulations rather than by printing and evaluating real printed samples. Each simulation is described below, and the combined loop of Steps (I), (II), and (III) was configured as a genetic algorithm (GA).[3]

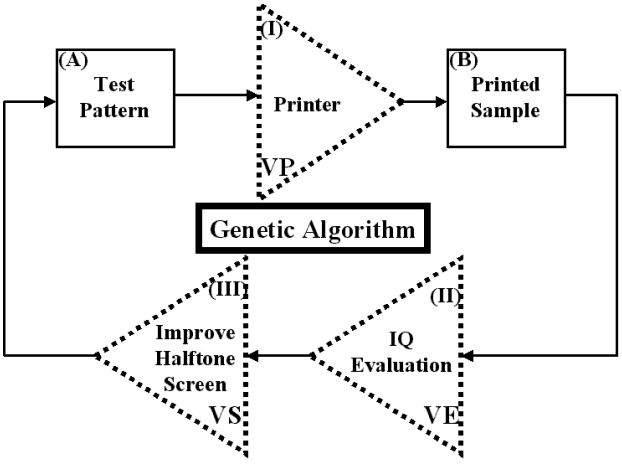

Figure 1. Iterative IQ Search

Step (I) The Virtual Printer

We selected a single printer for optimization, a laser EP printer (HP-4500 at 600 dpi), and the printer model was constructed and calibrated to this printer. Details of the model and the calibration procedure have been described elsewhere.[4] The key requirement for the printer model of Step (I) is that the simulated printed image must adequately approximate the real print regardless of the halftone screen one selects. While many printer models have been described in the literature, only those models that describe the physical mechanisms involved in the EP process have been shown to be capable of meeting this criterion.[5] Such a mechanistic model is too computationally intensive to meet the needs of the current project, so a semi-empirical model of printer non-linearity was developed. This virtual printer is summarized in Fig. 2. The assumption in the design of the virtual printer is that the overall printing process can be summarized with four key functions.

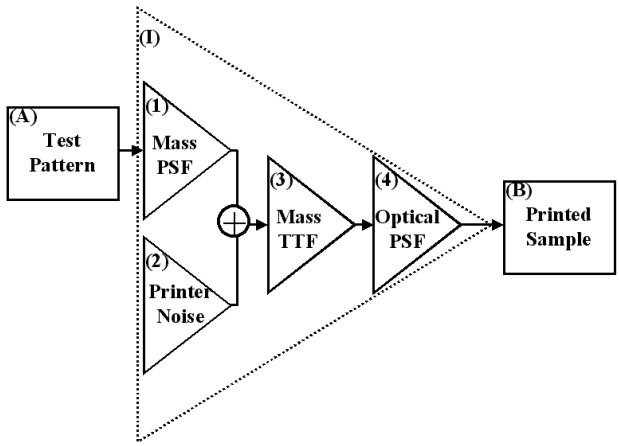

Figure 2. Step (I), the virtual printer

Process (1) is a mass spread function (PSF) that approximates the physical spread of toner in the printing process. This is analogous to modeling the physical dot gain of a printing process. It is important to keep in mind that the PSF of process (1) is not meant to be a description of the physical mechanism but a summary of the physical spread of toner in the process. The mechanism of toner spreading can involve different physical mechanisms at different points in the EP process. The overall result is modeled here by a single PSF function.

A test pattern (A) is selected and sent to the printer. The test pattern is a matrix of 0s and 1s generated by the halftone mask under evaluation. The 0s and 1s represent the idealized pattern of toner coverage (0 g/m^2 and 100% g/m^2) one intends to form on the paper. The spread process is accomplished by super-sampling the matrix of 0s and 1s in order to approximate continuous space, C(x,y). Then a point spread function, PSFm(x,y), is applied as shown in equation (1) where * is the convolution operator. The result, $C_b(x,y)$, may be thought of as a blurred distribution of toner mass.

$$C_b(x,y) = C(x,y) * PSFm(x,y) \tag{1}$$

The spread function used in this project is shown in equation (2). It contains two arbitrary parameters, p and s, that were determined by a calibration process that will be described below.

$$PSF_m(r) = \frac{k}{1+\left(\frac{r}{\sigma}\right)^p} \tag{2}$$

$$\text{where } k = \frac{p \cdot \sin(\pi/p)}{\sigma \cdot \pi} \text{ and } r = \sqrt{x^2 + y^2}$$

Process (2) in the printer model of Fig. 2 is a simulation of printer instability. This is simulated by a random number generator that produces normally distributed values for a metric we call "printer noise", ΔS, at each location (x,y) in the super-sampled image. Values of ΔS are distributed with a mean of zero and a standard deviation of σ_S. The value of σ_S also is determined by a calibration process described subsequently.

$$\Delta S(x,y) = RANDOM(x,y,\sigma_S) \tag{3}$$

The noise from equation (3) is added to the blurred toner coverage from equation (1) to produce a signal voltage as shown in equation (4).

$$V(x,y) = C_b(x,y) + \Delta S(x,y) \tag{4}$$

Caution should be exercised in the interpretation of equation (4) and of the model in general. This is not a real mechanistic model. It is an empirical model that is meant to simulate the important non-linear characteristics of the printer. Thus, the term V(x,y) is not meant as a model of the electrical potential across a photo-conductor. Rather, it is the empirical transfer function of process (4), illustrated in Fig. 3.

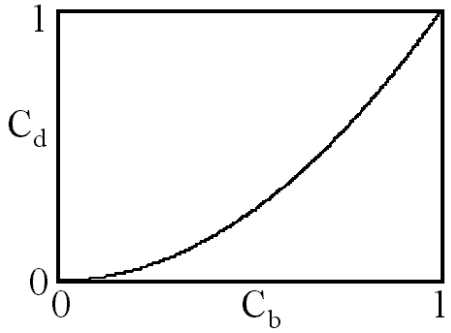

Figure 3. The Mass Tone Transfer Function, TTF

Step (3) in the printer model simulates the overall tone transfer characteristic of the printing process. A halftone imaging process is a bi-level process, and one often forgets that an underlying continuum process is involved. For an EP printer, this underlying continuum is a tone transfer function involving toner coverage, C_d in g/m^2, as a function of a voltage level, V. In Fig. 3, the bi-level voltages are intended to be 0 and 1 for no toner and for a toner halftone dot respectively. The function in Fig. 3 was modeled by equation (5). The constant, γ, is determined by a calibration process.

$$C_d(x,y) = V(x,y)^\gamma \tag{5}$$

Step (4) in the printer model describes the optical dot gain of the virtual printer. With a map of distributed toner, $C_d(x,y)$, and independent measurements of the optical properties of the substrate paper, the resulting reflectance distribution, R(x,y), may be calculated reliably, as has been reported in detail previously.[6,7,8,9]

Step (I) Calibration of The Virtual Printer

The virtual printer described above has four arbitrary constants (p, σ, σ_S, γ). Values for these constants were estimated experimentally by calibration with seven test patterns. The test patterns were clustered dot halftones of nominal dot area fraction F = 0.25, as illustrated in Fig. 4. The cluster sizes for intended toner delivery range from 1x1 to 9x9 in units of printer addressable (1/600 in^{-1}). Each pattern was printed with the HP-4500 EP printer using cyan toner, and each printed sample was analyzed quantitatively for the toner coverage, $C_m(x,y)$, using an analytical technique previously reported.[4] The same test patterns were passed through the virtual printer, and the simulated tone coverage, $C_d(x,y)$, was calculated. The values of the model constants (p, σ, σ_S, γ) were adjusted for a minimum error between $C_m(x,y)$ and $C_d(x,y)$.

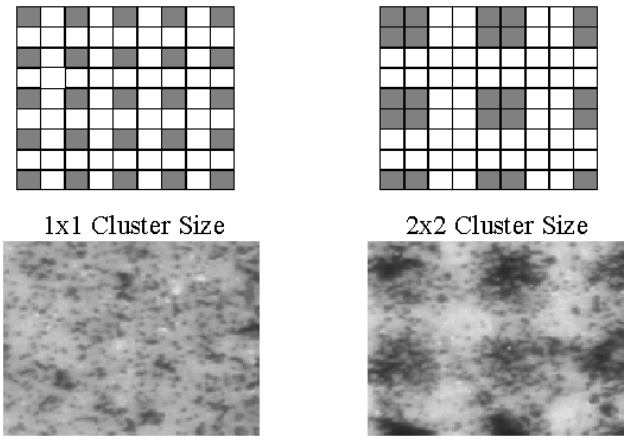

Figure 4. Example of 1x1 and 2x2 test patterns and the photomicrographs of the corresponding printed samples

Step (II) IQ Evaluation

Each halftone pattern in this study was evaluated with an overall figure of merit calculated from the simulated distribution R(x,y) calculated in Step (I). The figure of merit chosen for this study involved the RMS granularity of the simulated image. This RMS granularity was calculated by first calculating the Wiener spectrum W(x,y) as the squared modulus of the Fourier transform of R(x,y). Then the visual RMS granularity was calculated as shown in equation (6), where CSF(x,y) models the contrast transfer function of human vision.[10] A visual gamma function, dL/dR, was not included in the calculation because only relative ordering of IQ was needed in this study.

$$\sigma_V = \iint_{x,y} \mathrm{CSF}(x,y)\cdot \mathrm{W}(x,y)\,dxdy \qquad (6)$$

The visual granularity, σ_V, is caused by two factors. The first is the granularity of the halftone pattern itself. This is the granularity one tries to minimize by shifting to higher spatial frequencies so it will be filtered out by the visual CSF.[1] The second cause of σ_V is the instability of the printer, as illustrated in the micrographs of Fig. 4. The maximum possible value of σ_V would occur for a very large clustered dot pattern that is not filtered out by the CSF. This maximum value of σ_V is given by equation (7), where F is the dot area fraction. R_{max} and R_{min} are the reflectance values of the background paper and of the solid printed region.

$$\sigma_{max} = F\cdot(1\text{-}F)\cdot(R_{max} - R_{min}) \qquad (7)$$

The figure of merit, FOM, calculated in each loop of the iterative optimization process and passed to Step (III) was defined as shown in equation (8).

$$FOM = (\sigma_{max} - \sigma_v)/\,\sigma_{max} \qquad (8)$$

Step (III) Genetic Improvement

The genetic algorithm simulates evolution by selective breeding. The GA works with a population of many HxW halftone masks, determines their relative fitness as described above, and breeds children from pairs of better fit individuals replacing lesser fit individuals. The new children are slightly mutated, and the process iterates until a stopping criterion is met (such as: a sufficiently high fitness mask is discovered, the allocated time budget has elapsed, or insufficient further progress is being made).

The masks all consist of permutations of the integers {0,1,2,..., HW-1} (these are scaled appropriately when used as arrays of thresholds), and crossover and mutation operations follow the patterns described by Goldberg.[3] Full details of the genetic algorithm parameter choices are described elsewhere.[11]

We have tried several variations on this theme:

1. full permutations
2. gear wheels method
3. hybrid screens

Methods 1 and 3 involved searches for 64x64 masks; method 2 for masks approximately that size.

The full permutation method involves searching the massive (H·W)! space of all permutations. This technique works, but it is terribly slow. The other two methods exploit some things we already know about halftone screens and allow us to search considerably smaller spaces.

The gear wheel method refers to a pair of meshed gears with H and W teeth, where H and W are relatively prime. The teeth on each wheel are permutations (thus there are H!W! possible patterns here). If the teeth are clocked through time steps 0 to H·W-1, the touching pair of teeth (p,q) at time t indicates that the halftone mask matrix contains value t at position (p,q).

For computational convenience (the genetic algorithm was implemented in MatLab), we construct a standard mask using the trivial permutations (0,1,2,..., H-1) and (0,1,2,..., W-1). Then we permute the rows with the H-permutation and the columns with the W-permutation.

The hybrid HxW screens are constructed from two smaller screens, U and V, of size AxB and CxD, respectively, where H=A·C and W=B·D. The two smaller screens are also permutation matrices, so our search space is of size (A·B)!(C·D)!.

The rule for creating the HxW array from U and V is:

$$M[kA+i][lB+j] = C\cdot D\cdot U[i][j] + V[k][l] \quad (9)$$

This rule, similar in pattern to the Kronecker or tensor product of matrices, is a standard method for producing halftone masks, such as Bayer's dispersed dot mask in the case that both U and V are dispersed dot masks. In case U is a small single clustered dot and V is a large dispersed dot mask, M will be a clustered dot mask with a small cluster but a large number of thresholds.[12]

Results

Figure 5 illustrates typical results for two GA searches.[11] The FOM increases rapidly for the first few iterations and then more slowly. As one might expect, the FOM does not always increase for consecutive iterations. However, over time the overall FOM increases. The upward trend may or may not approach FOM=1, and the decision to halt the process was made based on the amount of time required to conduct the search. The time required for each iteration was between 10^{-1} and 10^{1} seconds, depending on the details of the chosen GA. Thus, 14,000 iterations typically required many hours to run, so the practical choice was made to halt the process after a convenient number of hours.

Conclusion

It is certainly true that the printer model described in this work is a highly simplified representation of the actual processes that govern the spread and delivery of toner in an electrophotographic laser printer. It is also true that the IQ evaluation and the figure of merit are simplifications, and the GA search technique still has much room for improvement. However, the work thus far strongly indicates the utility of combining the GA technique with a non-linear printer model for printer optimization. Further work is underway to improve all three steps of Fig. 1. In particular, improvements in the virtual printer are underway to account for color reproduction as well as noise power, and improved GA models are under development.

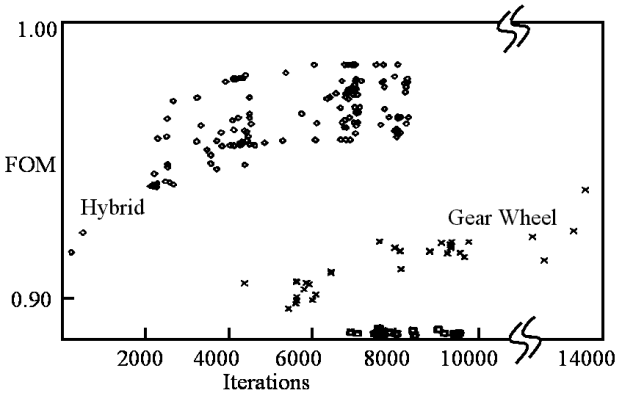

Figure 5. Figure of merit versus GA iteration number for the Hybrid and the Gear Wheel techniques.

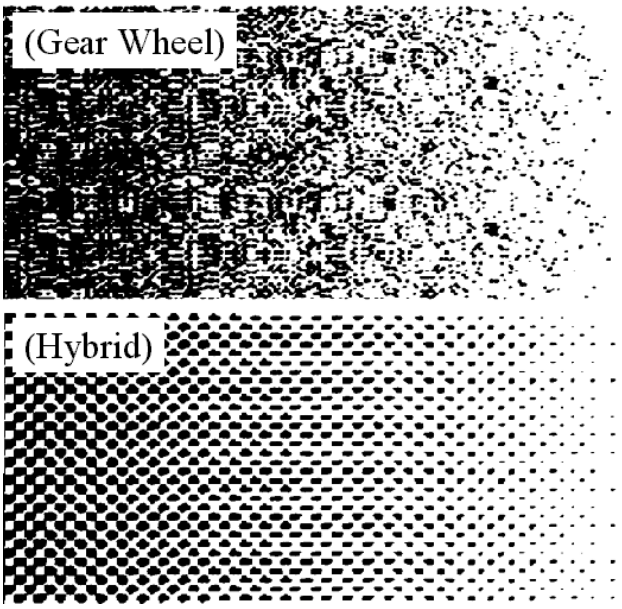

Figure 6. Gray ramps from Gear Wheel and Hybrid algorithms that produced the highest FOM values.

Another observation is that the rate of improvement of FOM is often faster for some genetic models, and slower for others. It is possible that many different genetic models may lead to similar, very high quality masks. However, some genetic models may approach an optimum more quickly and thus be preferred.

Figure 6 shows part of the gray ramps for masks with the highest observed FOM values from the Gear Wheel and Hybrid genetic models.[11] The hybrid approach began with a built-in clustering property to minimize, a priori, the noise characteristics associated with instability characteristics of the printer engine. The result was a rapid initial rise in FOM. The Gear Wheel simulation was begun with a more dispersed arrangement of threshold values, and the FOM increased much more slowly that observed for the Hybrid simulation.

Acknowledgement

This project was sponsored by a grant from the Hewlett-Packard Corporation, with guidance from Norm Burningham and Ken Lindblom. Special thanks to Bob Chin, Manager of Digital Technologies in OMNOVA Solutions, Inc., for sponsorship of summer research opportunities for undergraduate students participating in the project.

References

1. R. Ulichney, Digital Halftoning, MIT Press, Cambridge, MA, 1987.
2. Jan Allebach, Selected Papers on Digital Halftoning, MS-154, SPIE Press, 1999.
3. D.E. Goldberg, Genetic Algorithms in Search, Optimization & Machine Learning, Addison-Wesley, Reading, MA, 1989.
4. J.S. Arney, P.G. Anderson, and Sunadi Gunawan, A Printer Model that is Independent of Halftone Algorithm, J. Imag. Sci. and Technol., 47, in press (2003).
5. M.E. Scharfe, D.M. Pai, and R.J. Gruber, Electrophotography, Ch.5, Imaging Processes & Materials, Van Nostrand Reinhold, NY (1989).
6. G.L. Rogers, J. Imag. Sci. and Technol, 43, 341 (1998).
7. Shinichi Inoue, Norimichi Tsumura and Yoichi Miyake, J. Imag. Sci. and Technol, 41:6, 657(1997).
8. S. Gustavson, J. Imag. Sci . and Technol., 41, 283 (1997).
9. J.S. Arney, Cindy Scigaj, and Prashant Mehta, J. Imag. Sci and Technol., 45, 466 (2001).
10. Mark Fairchild, Color Apearance Models, Addison Wesley Longman, Inc., p. 30, 1997.
11. P.G. Anderson, J.S. Arney, S.A. Inverso, D.R. Kunkle, T.M. Lebo, and C. Merrigan, A Genetic Algorithm Search for Improved Halftone Masks, Artificial Neural Networks in Engineering Conference, St. Louis, MO, Nov. 2003.
12. H.R., Kang, Frequency analysis of microcluster halftoning, Recent Progress in Digital Halftoning, Vol. II, R. Eschback, Ed., IS&T, 1999.

Biography

Jonathan Arney is an associate professor at the Center for Imaging Science at Rochester Institute of Technology where he teaches Tone and Color Reproduction, Experimental Image Microstructure, and Properties of Imaging Materials. He also serves as faculty advisor for the Student Chapter of IS&T. In 1988 he participated in the revival of the Dayton Chapter of IS&T while working as a researcher at Mead Corporation in Ohio. He served as Programs Vice President and then as President of the Dayton Chapter prior to his move to Rochester. He is currently faculty advisor to the RIT Student Chapter of IS&T and serves as a council member of the Rochester Chapter of IS&T where he is the liaison between the Student Chapter and the Rochester Chapter. Jonathan received the IS&T Journal Award (Science) in 1990 for his publication "Radiation Attenuation Effects in Photosensitive Coatings of Microcapsules", J. Imag. Sci., 33,(1989)184. In 1995 he was presented the IS&T Service Award, and in 1996 received the Raymond C. Bowman Award for contributions to education in the fields of Photographic and Imaging Science.