An Analysis of the Blue Noise Mask Based on a Human Visual Model

Meng Yao, Lan Gao and Kevin J. Parker
Department of Electrical Engineering
University of Rochester, Rochester, New York

Abstract

The Blue Noise Mask (BNM) is a stochastic screen that produces visually pleasing blue noise. In its construction, a filter is applied to a given dot pattern to identify clumps in order to add or remove dots and thereby generate a correlated binary pattern for the next level. But up to now, all the filters were selected on a qualitative basis. There is no reported work describing precisely how the filtering and selection of dots affects the perceived error of the binary pattern. In this paper, we will give a strict mathematical analysis of the BNM construction based on a human visual model, which provides insights to the filtering process and also prescribes the locations of the dots that will result in a binary pattern of minimum perceived error when swapped. The analysis also resolves some unexplained issues noticed by other researchers.

Introduction

Stochastic screening has been an active research field in recent years. Blue noise halftone screens were first developed by Mitsa and Parker[1] in 1991. The BNM combines the blue noise characteristic of error diffusion[2] and the fast speed of ordered dither. The original BNM was constructed for one grey level at a time beginning with an intermediate starting binary pattern, or seed. Each subsequent level was constrained by the binary pattern at a preceding level, such that a single valued function, or ordered dither array was constructed with desired first order and second order statistics. At each level, a circularly symmetric filter was used to identify and eliminate low frequency structures (large "clumps") incompatible with the desired blue noise power spectrum. Implicit periodicity, or "wraparound" filtering was used so the BNM could be seamlessly tiled with itself to cover larger image spaces. Yao and Parker[3] later proposed a simpler and more efficient approach that further reduced the low frequency contents of the halftone patterns of the BNM. Ulichney[4] also used a filtering approach which he called the "void and cluster" method to generate a blue noise screen.

Using the filtering approach to generate a blue noise screen is efficient and easily implementable. When using the filtering approach to generate a blue noise pattern starting from a white noise pattern, the filter is applied repeatedly to the pattern to locate the centers of black and white clumps and then the values of the clump centers are swapped, thus diminishing the clumps. Similarly, in making a blue noise screen, the filter is used to find locations to add black or white pixels. However, all the reported work has been done on a qualitative basis. In the following section, we will derive the mathematical expressions for the filtering approach based on a human visual model and we will explain what is happening to the perceived error of the halftone pattern during the filtering and swapping process.

Analysis of Filtering Process

The Human Visual Model

Our analysis of the filtering approach is based on a human visual model, which is basically a low-pass filter. The following is a model given by Daly[5]:

$$V_{ij} = \begin{cases} a(b + c\tilde{f}_{ij})\exp\left(-(c\tilde{f}_{ij})^d\right), & \text{if } \tilde{f}_{ij} \rangle f_{max,} \\ 1.0 & \text{otherwise} \end{cases} \tag{1}$$

where a = 2.2, b = 0.1 92, c = 0.114 and d = 1. 1; $\tilde{f}_{ij}$ is the radial spatial frequency in cycles/degree and f_{max} is the frequency at which the function peaks. To incorporate the decrease in sensitivity at angles other than horizontal and vertical, the radial frequency is scaled such that

$$\tilde{f}_{ij} \rightarrow \tilde{f}_{ij} / s(\theta),$$

where

$$s(\theta) = \left(\frac{1-w}{2}\right)\cos(4\theta) + \left(\frac{1+w}{2}\right) \tag{2}$$

where w is a symmetry parameter.

Changing One Black Pixel to a White Pixel

In this section, we consider the case of changing one black pixel to a white pixel and try to minimize the perceived error between the constant grey level and the perceived binary pattern.

Given a current binary pattern b(i,j) for level g and a human visual model, filter h(i,j), find the locations of the black pixels to be converted to white pixels that will minimize the mean squared error of the level $g' = g + \Delta g$ and the perceived binary pattern for level g'. For a BNM of a size greater than 16×16, more than one pixels need to be changed to move to the next level. We will first consider the case of changing one black pixel to a white pixel. The perceived binary pattern for level g is:

$$f(i,\ j) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} h\Big((i-m)_{\mathrm{mod}\ N}, (j-n)_{\mathrm{mod}\ N}\Big) b(m,n) \qquad (3)$$

where N×N is the size of the BNM and mod means modulo. The modulo operation gives the filter a "wrap-around" property which eliminates any border discontinuity when the BNM is tiled to cover larger image space. Suppose we change a black pixel at (i_0, j_0) to a white pixel, the binary pattern for level g' will be:

$$b'\ (i, j) = b(i, j) + \delta(i - i_0, j - j_0\ ) \qquad (4)$$

and the perceived binary pattern for level g' is:

$$f'\,(i,j) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} h\Big((i-m)_{\mathrm{mod}\ N}, (j-n)_{\mathrm{mod}\ N}\Big) b'\,(m,n) \qquad (5)$$

From equations (3) through (5) we obtain:

$$f\ (i, j) = f\ (i, j) + h((i\text{-}i_0)_{\mathrm{mod}\ n}\ , (j\text{-}j_o\ )_{\mathrm{mod}\ N}\ ) \qquad (6)$$

The MSE for the perceived binary pattern for level g' is:

$$E'^2 = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left[f'\,(i,j) - g'\right]^2 \qquad (7)$$

Our goal is to find the location (i_0, j_0) which corresponds to a black pixel and which minimizes E'². Expanding the right side of the equation (7) we obtain:

$$\begin{aligned} E'^2 &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\Big[f(i,j) + h\Big((i-i_0)_{\mathrm{mod}\,N}, (j-j_0)_{\mathrm{mod}\,N}\Big) - g\Big]^2 \\ &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}[f(i,j) - g'\,]^2 + \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} h^2((i-i_0)_{\mathrm{mod}\,N}{}'\,(j-j_0)_{\mathrm{mod}\,N},) \\ &\quad +2\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} f(i,j)h((i-i_0)_{\mathrm{mod}\,N'}\,(j-j_0)_{\mathrm{mod}\,N}) \\ &\quad -2\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} g'\,h((i-i_0)_{\mathrm{mod}\,N}, (j-j_0)_{\mathrm{mod}\,N}) \end{aligned} \qquad (8)$$

There are four terms on the right side of the above equation. Since g' is a constant and f(i,j) is known, the first term in equation (8) is fixed. The second term is a summation of the shifted filter squared over the support of the BNM. Due to the "wrap-around" property, this term is constant. For the same argument, the fourth term is also a constant. Let

$$G(i_0, j_0) = \sum_{i-0}^{N-1}\sum_{j-0}^{N-1} f(i,j)h((i-i_0)_{\mathrm{mod}\ N}, (j-j_0)_{\mathrm{mod}\ N}) \qquad (9)$$

where $G(i_0, j_0)$ is a function of i_0 and j_0. It follows that minimizing E'² is reduced to minimizing $G(i_0, i_0)$. After further simplification,

$$G(i_0, j_0) = \sum_{m=0}^{N-1}\sum_{n-0}^{N-1} b(m,n)R_h((i_0-m)_{\mathrm{mod}\ N}, (j_0-n)_{\mathrm{mod}\ N}) \qquad (10)$$

where R_h is the circular autocorrelation of h. In the frequency domain, this filter is $|H|^2$.

It can be easily seen that $G(i_0, i_0)$ is the circular convolution of the binary pattern for level g with a new filter that is the autocorrelation function of h.

Examination of equation (11) tells us that to minimize $G(i_0, j_0)$, or the perceived error for level g' measured by the human visual model h, we should use R_h, the autocorrelation function of h, instead of the human visual model itself, to filter the binary pattern for level g, and find the black pixel (i_0, j_0) that gives the minimum value of the filtered pattern. To look at it some other way, the filter we use to search for the pixels to be changed will result in minimum mean squared error measured by another filter. If a Gaussian filter is assumed as the human visual model, the actual filter that should be used has a standard deviation of

$$\frac{\sqrt{2}}{2}\sigma$$

in the frequency domain, where σ is the standard deviation of the Gaussian human visual model. Mitsa and Parker[6] found empirically that using the principal frequency as the cut-off frequency in the filter design did not generate the most visually pleasing patterns. Instead, under certain experiments they found that a factor of

$$\frac{\sqrt{2}}{2}$$

should be applied to the principal frequency to obtain the best pattern. Our theoretical analysis shows that this is not a coincidence. For a symmetric Gaussian filter, the autocorrelation R_h of h is obtained by simply replacing σ with

$$\frac{\sqrt{2}}{2}\sigma,$$

which will shift the cut-off frequency by a factor of

$$\frac{\sqrt{2}}{2}.$$

Swapping Multiple Dots

In the previous section, we solved the problem of finding the single black or white pixel, which when swapped, yields the minimum MSE of the new binary pattern. Normally, the size of a BNM is larger than 16×16, which means multiple dots have to be swapped to reach the next level. Intuitively, swapping one dot each time and minimizing the MSE at every step may not give us the dots that minimize the MSE for the next level. The added dots will interact with each other as well as interact with the dots of level g. The ideal way would be to simultaneously identify and convert a group of pixels. Suppose that P black pixels need to be changed to white pixels to move from level g to g'. Taking the same approach as the last section, we have:

$$\begin{aligned} G(i_0, j_0) &= 2\sum_{p=0}^{P-1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} b(m,n)R_h\Big((i_p-m)_{\mathrm{mod}\ N}, (j_p-n)_{\mathrm{mod}\,N}\Big) \\ &\quad + \sum_{p=0}^{P-1}\sum_{\substack{p'=0 \\ p'\neq p}}^{P-1} R_h\Big((i_p - i_{p'})_{\mathrm{mod}\ N}, (j_p - j_{p'})_{\mathrm{mod}\ N}\Big) \end{aligned} \qquad (11)$$

Let's examine how we can find the P black points in f(i,j) which, when converted to white, will minimize E'^2, or $G(i_0, j_0)$. To minimize the first term of $G(i_0, j_0)$, we use the autocorrelation function R_h of h as a filter and apply it to the binary pattern for level g. We find the P black pixels with the smallest filtered values, which will give the minimum value of the first term. For the second term, assuming that we are dealing with a Gaussian low pass filter, then R_h is Gaussian too. The Gaussian filter decreases exponentially as the distance from the origin increases, which means that for the second term to be minimum, the distance between the P black pixel should be as large as possible, which requires that the P black pixels be distributed as uniformly as possible in the N×N dimension of the binary pattern. To minimize $G(i_0, j_0)$, we should take both terms into consideration. For many levels, the first term will be the dominant factor in the perceived error, because there are many more terms in its summation than in the second term. However, the second term can play an important role in some cases, especially at extreme levels. For example, when we start from level 0 and move to level 1, the first term is 0 and the minimization depends solely on the second term. When two of the P black pixels we choose are too close to each other, they can contribute considerably to the second term. We should also point out that the filter R_h cannot guarantee the relative positions of the P black pixels. It is very likely that two neighboring black pixels are picked by the filter to be swapped, which will increase the second term immediately. Equation (12) also provides guidance to our BNM construction algorithm. As we use our filter to pick the candidates to be swapped, we should make sure that any two of the candidates are not close neighbors.

Conclusion

In our analysis of the minimization problems, we demonstrate that in order to minimize the MSE of a binary pattern using a human visual model h, the autocorrelation of h should be used as the filter to choose candidates to be swapped, although some restrictions should be imposed on the candidates, for example, the P black pixels should not be close neighbors in the case of swapping multiple dots. The BNM so constructed minimizes the perceived error of a binary pattern based on the human visual model, resulting in visually pleasing patterns.

Reference

1. T. Mitsa and K. J. Parker, "Digital halftoning using a blue-noise mask," in *Image Processing Algorithms and Techniques II*, MR. Civanlar, s. K Mitra, and B. P. Mathur, eds., *Proc. Soc. Photo-Opt. Instrum. Eng.* **1452**, 47–56 (1991).
2. R. A. Ulichney, *Digital Halftoning,* MIT Press, Cambridge, Mass. (1987).
3. M. Yao and K. J. Parker, "Modified approach to the construction of a blue noise mask," *J. Elec. Imag.,* vol. **3**, no. 1, 92–97, 1994.
4. R. A. Ulichney, "The void-and-cluster method for dither array generation," *IS&T/SPIE Symposium on Electronic Imaging Science & Tech.,* San Jose, CA (1993).
5. J. Sullivan, L. Ray, and R. Miller, "Design of minimum visual modulation halftone patterns," *IEEE Trans. Syst. Man Cybern., vol.* **21**, No. 1, 33–38 (1991).
6. T. Mitsa and K. J. Parker, "Digital halftoning technique using a blue-noise mask," *J. Opt. Soc. Am.* A **9**, 1920–1929 (1992).

❋ Previously published in *IS&T NIP11*, pp. 424–426, 1995.