Morphological characterization of dithering masks

Vladimir Misic  
Kevin J. Parker  
University of Rochester  
Department of Electrical and Computer Engineering  
Rochester, New York 14627

Abstract. We present some novel tools for the analysis of blue-noise binary patterns. Unlike most of the existing methods that evaluate the frequency content of a given mask or its lower order statistics, our new metrics characterize the morphological content of a mask that is quantified using simple one-pass filtering. An analytical filter expression is given. As a result, one can balance the structural content of the mask—diagonal, vertical, and horizontal interconnections of the majority (or minority) pixels—at the same level. In addition, it is possible to improve the overall mask quality by prescribing the occurrence of morphological shapes of connected pixels. Examples of morphological analysis are given to demonstrate the different qualities of blue-noise and white-noise patterns.

1 Introduction

An important issue in blue-noise binary pattern design is the use of a quality metric or error metric, in the optimization process. The choice of metric directly affects the quality of the final patterns.

There are several metrics available that could be used to grade the quality of the binary approximation of the gray level (and, respectively, the dithering mask). However, none of them completely answers the question, “What is a good mid-tone level?” At both ends of the halftone scale—at very dark and very light color levels—it is fairly easy to say if some level is “good” or “bad” blue noise, since minority pixels are few and widely spaced. However, at gray levels closer to mid-tones, minority pixels are no longer widely spaced and must form connected morphological shapes. We propose a morphological analysis to quantify the blue-noise patterns at mid-tones (and over all gray levels), as a useful tool for analysis and design.

2 Evaluation of Binary Patterns

A number of evaluation metrics have been proposed to assess the quality of individual blue-noise binary patterns. Two of the commonly used metrics are frequency-weighted mean square error (FWMSE) and average distance between nearest-neighbor, minority-pixel pairs (AMD).

2.1 Frequency-Weighted Mean Square Error

The FWMSE is the most commonly used metric. It is also referred to as the human visual system weighted, mean square error (HWMSE), since it employs a model of the human visual system (HVS) during assessment. If one denotes b as the binary pattern (dimensions $N \times N$) to be evaluated, and h as the point-spread function for the HVS model, the perceived halftone error at the observed level $g$ could be expressed as:

$$\text{err} = b \ast h - g,$$

where $\ast$ denotes the circular convolution. Therefore, the FWMSE could be calculated in both the spatial and the frequency domain as:

$$\text{FWMSE} = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |\text{err}(i,j)|^2,$$

$$= \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |B(i,j)H(i,j)|^2,$$

where $B$ and $H$ stand for the discrete Fourier transform of $b$ and $h$, respectively. The final summation in the Fourier domain excludes the dc value (the origin point in the referent coordinate system), assuming that the expected dc value is simply $g$.

From Eq. (2) in the Fourier domain, it is clear that the FWMSE metric evaluates the binary pattern globally. Therefore, local details can be averaged out in the evaluation process.

2.2 Average Distance Between Nearest-Neighboring, Minority-Pixel Pairs

Due to the failure of the FWMSE to recognize local characteristics of the binary pattern, Yu proposed a method to quantify the graininess of the pattern, where the AMD was measured. Namely, for each minority pixel, a search was conducted for the neighboring minority pixel with the minimal Euclidean distance to the observed pixel. Wong suggested a similar approach but with an additional level-related weighting factor. The average of all such distances is the AMD value:
Morphological characterization...

\[ \text{AMD} = \frac{\sum_{i=1}^{k} D_{\text{min}}(i)}{k} = g \cdot L, \]

where \( D_{\text{min}} \) denotes the minimal distance to the nearest neighbor for the \( i \)th minority pixel, and \( k \) stands for the number of minority pixels at the gray level \( g \) in the mask that generates \( L \) color levels.

When considering two binary patterns that are both candidates for the gray-level approximation, we should choose the one with the bigger AMD value (and, consequently, the less grainy of the two patterns). The AMD describes the binary pattern in terms of graininess, but it does not reveal the exact nature (morphological shape) or position of a grainy artifact.

Although generally successful, the FWMSE and AMD (as well as other existing metrics) fail to localize (and sometimes even to recognize) problems at the mid-tone levels (output levels between 0.25 and 0.75) where the AMD is smaller than 2. Our proposed algorithm not only localizes such problems but also allows the efficient location of the exact position and morphological shape of a pixel “clump” (see Fig. 1).

3 Morphology Information Retrieval by Means of Filtering

In order to extract the morphological information from a certain gray level \( g \) as a result of the filtering process, one should construct a filter that has a unique response for each pixel configuration. For simplicity, we show as an example a very small filter size (2×2). However, since the filter construction process is generic, larger filters of this type could be used as an optimal look-up table (LUT) in a blue-noise mask (BNM) construction.

3.1 Filter Construction

Consider the binary pattern \( b \) to be filtered by a rectangular \( M \times N \) filter \( f \). The result is described by:

\[ \text{res}(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) \cdot b(x+m,y+n). \]

In order to obtain unique responses for each pixel configuration, we propose use of \( M \times N \) filter \( f \):

\[ f(m,n) = 2^{m+N+n} \quad m = 0,1,\ldots,M-1 \quad n = 0,1,\ldots,N-1. \]

Equation (4) can be rewritten now as:

\[ \text{res}(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} 2^{m+N+n} \cdot b(x+m,y+n). \]

If we reorder \( b(x+m,y+n) \) in raster scan order as \( b_{(x+m)N+y+n} \), then Eq. (5) can be rewritten as:

\[ \text{res}(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} 2^{m+N+n} \cdot b_{(x+m)N+y+n}. \]

The filter generates 16 possible output levels (values ranging from 0 to 15), one for each possible pixel configuration in a 2×2 neighborhood (Fig. 2). Thus, the filter represents a given pixel configuration as a binary number (i.e., a 4-bit integer vector).

\[ f = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}. \]

Since any number \( N \) can be uniquely written in a binary numeric system as the ordered sum of the powers of 2, Eq. (8) proves that the filter given in Eq. (5) really has unique response.

3.2 Morphology Information Retrieval

As mentioned before, our proposed metric registers the position and the number of horizontal, vertical, and diagonal connections between neighboring pixels. Since the information we are interested in is strictly local, the filter does not have to be larger than 2×2:

\[ f = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}. \]

The filter generates 16 possible output levels (values ranging from 0 to 15), one for each possible pixel configuration in a 2×2 neighborhood (Fig. 2). Thus, the filter represents a given pixel configuration as a binary number (i.e., a 4-bit integer vector).

\[ f = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}. \]
For example, if 0 represents a white pixel and 1 represents a black pixel, then an “upper” black horizontal connection will result in a filter output value of three. Thus, the filter output from a binary pattern as shown in Fig. 3 will contain unique numbers corresponding to the morphology of pixels within a sliding 2×2 window.

4 Morphological Characterization Algorithm

The extraction of morphological features using the described generic filter [see Eq. (5)] can be schematically represented as a two-step algorithm: the first step being filtering (circular convolution or correlation) and the second step being result identification (Fig. 4).

As a result of the filtering process, we have a matrix of the “morphological content” of a filtered binary pattern. Each value uniquely represents the content of the appropriate sliding window centered at the same pixel position as in the original (binary) pattern. The second step, result identification, is now simple. These values are used as pointers on the LUT with predefined actions for each pixel configuration. In mathematical morphology this approach is known as a hit-or-miss transform. 

For the purpose of calculating the LUT index (pointer), the filtering step from Fig. 4 may be replaced by direct use of the mask binary values. The n×n neighborhood of processed pixel may be reordered (vectorized) as binary value b_n−1, b_n−2, ..., b_0, and that value may be used to index into LUT. However, our goal is not to merely identify certain morphological shapes, but also to use results of the proposed filtering for further mask processing (e.g., relocation of certain morphological shapes based on their uneven spa-
tial distribution, which can be calculated as spatial distribution of appropriate filtered values).

5 Metric Analysis

When considering a dithering mask at any given gray level \( g \), it can be seen (Fig. 2.) that there are a few basic groups of local pixel configurations: zero (all white), one-pixel, two-pixel horizontal, and two-pixel vertical and diagonal connections. The L-shaped connection is actually a one-pixel configuration, with reversed minority pixels. The number and distribution of these basic configurations can give us information that is not provided by any of the metrics previously used to evaluate the quality of the individual BNM.

Using the algorithm described in Fig. 4 with \( n = 2 \) (a \( 2 \times 2 \) filter), we can easily extract these features. For the most simple assessment of a binary pattern, the required action for each output pixel is to count the number of each type of the output values (add up one to the appropriate counter). After all the pixels are taken into account, the sums are averaged to represent the actual distribution (i.e., the expectation) of respective \( 2 \times 2 \) configurations.

5.1 White-Noise Mask and Blue-Noise Mask Analysis

In the case of WNM, at any given gray level the probability of any pixel to be turned “on” (white) is \( g \), and to be turned “off” (black) is \( 1 - g \). Since pixels have independent distribution, it is obvious that in any given group of four pixels, the level-dependent probabilities for the configuration given in Fig. 1 depend simply on gray level \( g \). For example, see Table 1.

The theoretical results in Table 1 are consistent with the actual data collected, where the average was calculated from a set of experiments using 20 WNM generated independently (Fig. 5).

In the case of the BNM, it is clear that the probability of a pixel being turned on or off is not independent of neighboring pixels. It is dependent on the observed gray level \( g \) as well as on the arrangement of all other pixels in the mask. For that reason, it is not possible to give an exact mathematical expression for observed distributions of local \( 2 \times 2 \) pixel configurations.

Due to certain properties of the BNM, we prefer certain desirable connectivity relationships between pixels. At each level in a BNM, the pixels should be placed maximally distanced to neighboring pixels. This is an attempt to avoid some of the previously mentioned configurations at certain levels. For example, if there is a place to add an isolated minority pixel at a certain level, the new two-pixel connection will be avoided. If there is one predominant type of connection at a certain level (e.g., horizontal versus vertical connections), the pattern becomes unbalanced (or visually less pleasant, perhaps, even disturbing). In addition, the number of diagonal connections should be larger than the number of vertical and horizontal connections at any moment, since the HVS is more sensitive to the existence of horizontal and vertical lines.

When comparing the plots of a WNM [Fig. 6(c)] and a

<table>
<thead>
<tr>
<th>Local Configuration</th>
<th>Probability Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four white pixels</td>
<td>( E_{ww} = g^4 )</td>
</tr>
<tr>
<td>Any combination of one black and three white pixels</td>
<td>( E_{wbb} = (1-g)g^3 )</td>
</tr>
<tr>
<td>Any combination of two black and two white pixels</td>
<td>( E_{bb} = (1-g)^2g^2 )</td>
</tr>
<tr>
<td>Any combination of three black and one white pixel</td>
<td>( E_{bwb} = (1-g)^3g )</td>
</tr>
<tr>
<td>Four black pixels</td>
<td>( E_{bb} = (1-g)^4 )</td>
</tr>
</tbody>
</table>

Fig. 5 Distribution of all \( 2 \times 2 \) pixel combinations for WNM versus gray level.
typical BNM [Fig. 6(d)], it is apparent that all of the WNM distributions are intersecting at one point (output level 128). That means there are equal numbers of all types of 2×2 configurations present at output level \( g = 0.5 \). This results in visually disturbing pixel clumps (all black) and spatial voids (all white). In the case of the BNM, the mask building algorithm tends to arrange minority pixels in certain patterns, resulting in the virtual nonexistence of all black and all white 2×2 configurations at the middle of the color scale. Also, the number of L-shaped connections is significantly smaller than the number of any two-connections at the mid-tone color levels.

From Fig. 7, it is easy to locate the nature of nonoptimality of an analyzed BNM by inspection. In the mask building process, the original algorithm did not recognize that the number of horizontal and vertical rods \( \{3, 12, 5, 10\} \) in the morphological content matrix became almost the same as the number of diagonal connections \( \{6, 9\} \) at an output level of 71. That characteristic propagated in the mask building process toward the lower part of the gray scale \( g \) (gray levels: \( g < 71/256 \)). As a consequence, this particular dithering mask is better at the lighter mid-tones (output levels in the range of 180 to 220 on this scale) than at the darker mid-tones (output levels from 35 to 70). The particular mask building algorithm used in this case failed to produce a completely balanced scale, thus affecting the overall dithering mask quality.

An example of this unbalance is given in Fig. 8, where two symmetrical gray levels (205 and 50) are compared. The lighter one appears better, due to better balanced morphological content. The observed portion of the level 205 [Fig. 7(a)] has 6 vertical, 5 horizontal, and 15 diagonal connections versus 13 vertical, 15 horizontal, and 14 diagonal connections at the same-sized portion of the level 50 [Fig. 7(b)]. These numbers are consistent with the mask statistic shown in Figs. 7(a) and 7(b).

If the information about the unbalance between symmetrical levels and unbalance in number of vertical, horizontal, and diagonal connections were used, the mask building algorithm would produce a better balanced mask. All methods for constructing blue-noise masks employ quality or goodness criteria, and morphological characterization can be included into these criteria. However, the use of this proposed analysis in BNM synthesis is beyond the scope of this work.

6 Conclusions

Some novel tools for dither mask analysis are presented in this work. An analytical filter expression is given, and it is shown that the filter response is unique for each pixel combination (equivalent to a morphological hit-or-miss trans-
A more complex analysis and synthesis could be employed using the same filter type (but a larger size). However, the LUT size grows exponentially as $2^{nm}$ (where the $n$ and $m$ filter dimensions are $n \times m$). The strategy for the utilization of such a filter will be the topic of future research.

References