

A Comprehensive Halftone Image Quality Evaluation of First- and Second-order FM Halftones

Sasan Gooran

Linköping University, Media and Information Technology, Department of Science and Technology, Norra grytsgatan 10A,
Norrköping 60233, Sweden
E-mail: sasan.gooran@liu.se

Abstract. Halftoning is a crucial part of image reproduction in print. For large format prints, especially at higher resolutions, it is important to have very fast and computationally feasible halftoning methods of good quality. The authors have already introduced an approach to obtain image-independent threshold matrices generating both first- and second-order frequency modulated (FM) halftones with different clustered dot sizes. Predetermined and image-independent threshold matrices make the proposed halftoning method a point-by-point process and thereby very fast. In this article, they report a comprehensive quality evaluation of first- and second-order FM halftones generated by this technique and compare them with each other, employing several quality metrics. These generated halftones are also compared with error diffusion (ED) halftones employing two different error filters. The results indicate that the second-order FM halftoning with small clustered dot size performs best in almost all studied quality aspects than the first- and second-order FM halftoning with larger clustered dot size. It is also shown that the first- and second-order FM halftones with small clustered dot sizes are of almost the same quality as ED halftones using Floyd–Steinberg error filter and of higher quality than halftones generated by ED employing Jarvis, Judice, and Ninke error filter. © 2022 Society for Imaging Science and Technology. [DOI: 10.2352/J.ImagingSci.Technol.2022.66.1.010506]

1. INTRODUCTION

Since many of the reproduction devices, such as printers, are limited to only a few output states (printing inks), digital halftoning, which is the process to convert a continuous-tone image to a pattern of binary pixels, is one of the most essential parts of printing. Since the human eye is limited in its capacity to resolve small dots and dots close to each other, the human eye is not able to distinguish between the con-tone image and the binary one if the viewing distance is sufficiently far or/and the dots are sufficiently small.

Halftoning algorithms are commonly divided into two main sub-categories, called amplitude modulated (AM) and frequency modulated (FM). In AM halftones, i.e., periodic clustered-dot halftones, the size of the dots is varied while their spacing is constant. The FM methods, themselves, can be categorized into two sub-groups, which we refer to as first-order FM (also called dispersed-dot halftones) and second-order FM (also called stochastic clustered-dot halftones). In the first-order FM, the size of the dots is

constant while their frequency is variable. In the second-order FM, the size and the frequency are both varied [1]. Regarding the computational process, halftoning methods can be divided into a number of categories, from point-by-point ordered dithering [2] to error diffusion [3] to advanced iterative halftoning methods such as direct binary search (DBS) halftoning algorithm [4]. In point-by-point ordered dithering, or any other point-by-point halftoning, pixels in a cone-tone image are converted to black or white pixels based on a simple comparison with a threshold, making them the fastest halftoning technique. In error diffusion (abbreviated to ED in this article), the current pixel value is, at first, compared to a threshold, and the same position in the output halftone is set to 1 if the pixel value is greater than the threshold. Otherwise, a 0 is set at that position. An error is then calculated and diffused to neighboring pixels using a so-called error filter. This algorithm starts at one pixel and scans the image, usually one row and one pixel at a time, and terminates when the last pixel is processed. Therefore, this algorithm is not terminated until all pixels have been processed, making it slower than point-by-point halftoning algorithms. The original or non-modified ED is simple but suffers from correlated artifacts and directional hysteresis [2]. Hence, there have been many modifications to improve this algorithm reported in literature [5, 6]. The iterative halftoning methods, such as DBS [4] or Iterative Method Controlling the Dot Placement (IMCDP) [7], operates on the whole image and usually result in higher quality halftones but at the cost of high computational complexity.

For high resolution prints, especially large format prints, it is very important that the halftoning process is fast but still of good quality. For example, in order to produce a printout at 1200 dpi of the size of an A3 page, i.e., approximately 12×16 inches², an image of size $14\,400 \times 19\,200$ pixels is to be halftoned. If the iterative methods or even ED methods were directly applied to such a large image, it would require a large amount of data to be processed and many operations to be performed, making the computational procedure very slow. Just to give an indication of the computational complexity, consider the ED method using Floyd–Steinberg filter with four elements being applied to the image in the above example. At each pixel, in addition to the comparison to a threshold, four multiplications and four additions/subtractions (one to calculate the error and three to diffuse the error) are also

Received Mar. 15, 2021; accepted for publication July 29, 2021; published online Sep. 28, 2021. Associate Editor: Norimichi Tsumura.

1062-3701/2022/66(1)/010506/27/\$25.00

required. For the example above, it means nearly 1.1 billion additions and 1.1 billion multiplications. Furthermore, since the result is not obtained in this method until the last pixel in the scanning order is processed, they cannot be performed by parallel processing. In [1], a method, referred to as point-by-point IMCDP in this article, is proposed that generates image-independent threshold matrices for first- and second-order FM halftoning. Predetermined and image-independent threshold matrices make the proposed halftoning method point by point and thereby very fast. The only operation that is required is a comparison per pixel. In addition, this method has the benefit that it can be performed by parallel processing. The point-by-point IMCDP is thoroughly reported in [1] and briefly described in Section 2.

The main goal of this article is to report a thorough study of the properties and a quality evaluation of the first- and second-order FM monochromatic and color halftones generated by point-by-point IMCDP. The characteristics and the quality of these halftones are compared with ED and ordered dithering methods. Four important objective quality measures, including mean squared error, gray level representation, graininess, and sharpness, are going to be implemented for monochromatic halftones. Three important objective quality measures for color halftones, including spatial-CIELAB ΔE , graininess, and sharpness are also implemented. These quality measures are all applied on the perceived printed halftones, which is achieved by using a printer model and an eye model. The printer model used in this article is the dot-overlap model, and the eye model is a Gaussian function approximating the impulse response to Mannos and Sakrison's eye sensitivity function. These models are briefly described in this section. In order to apply the eye model to color halftones, spatial-CIELAB (S-CIELAB) filtering is used, which is also shortly described in this section.

1.1 Printer Model

The printer model used in this article is the dot-overlap model, thoroughly described in Refs. [8] and [9]. In this printer model, it is assumed that the printed dots are circular rather than square and they are larger than the minimal covering size, as illustrated in Figure 1. In this figure, $b(m, n)$ is the binary (halftone) image, in which the empty pixels are supposed to be 0. There are four printed dots illustrated in this figure. The printer model output, $b_p(m, n)$, is obtained by

$$b_p(m, n) = \begin{cases} 1, & \text{if } b(m, n) = 1 \\ f_1\alpha + f_2\beta - f_3\gamma, & \text{if } b(m, n) = 0, \end{cases} \quad (1)$$

where each element (m, n) in b_p is dependent on $b(m, n)$ and its eight neighbors. The parameter f_1 is the number of horizontally and vertically neighboring dots that are black, f_2 is the number of diagonally neighboring dots that are black and not adjacent to any horizontally or vertically neighboring black dot, and f_3 is the number of pairs of neighboring black

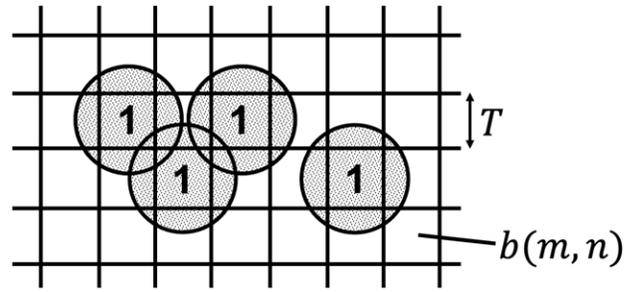


Figure 1. The dot-overlap model.

dots in which one is a horizontal neighbor and the other is a vertical neighbor. The area of each square pixel is T^2 , where T denotes the height/width of each pixel as shown in Fig. 1. What the parameters α , β , and γ are and how they are calculated are thoroughly described in Ref. [8]. There is an important parameter in this model called ρ , which is the ratio of the actual dot radius of a printed dot to the ideal dot radius, which is equal to $T/\sqrt{2}$. With the ideal radius, we mean the radius of the smallest circle that covers an entire square. Therefore, the larger the ρ , the more the impact of the printer.

1.2 Eye Model

The reason halftoning works is that the eye attenuates higher frequencies. There have been a number of models proposed in different research works to estimate the modulation transfer function (MTF) of the eye. In this article, we employ Mannos and Sakrison's estimation, thoroughly described in Refs. [9] and [10]. The MTF of the eye according to Mannos and Sakrison is

$$H(f) = 2.6(0.0192 + 0.114f)e^{-(0.114f)^{1.1}}, \quad (2)$$

where the frequency f is in cycles/degree. This MTF is plotted in Figure 2(a). As seen in this graph, the eye is most sensitive to frequencies around 8 cycles/degree according to this model. It has been found in Ref. [9] that the impulse response of the 1-D eye filters can be well approximated by Gaussian functions with an appropriate standard deviation (σ). Note that, the impulse response is specified in degrees of visual angle and the spacing of the dots is given by,

$$\tau \approx 1/RD \text{ radians} = \frac{180}{\pi RD} \text{ degrees}, \quad (3)$$

where R is the print resolution in dpi, and D is the viewing distance in inch. The frequency will then be $f = 1/\tau = (\pi RD/180)$ cycles/degree. As explained in Ref. [9], a Gaussian filter with $\sigma = 0.0095$ degrees is virtually identical to the impulse response of the 1-D eye filter. The standard deviation $\sigma = 0.0095$ degrees corresponds to $\sigma = 0.0095/\tau = 0.0095\pi RD/180$ pixels. For example, for $R = 600$ and $D = 15$, the standard deviation is $\sigma \approx 1.5$ pixels. A Gaussian function approximating the 1-D impulse response of the eye filter for $R = 600$ and $D = 15$ is shown in Fig. 2(b).

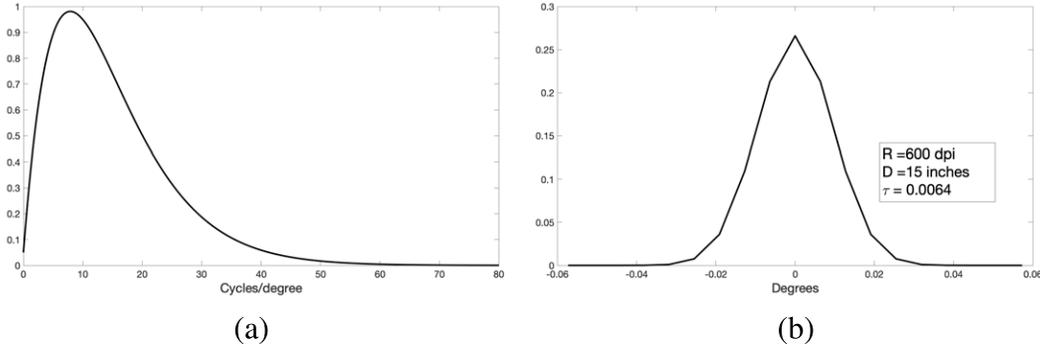


Figure 2. (a) The MTF of the eye (according to Mannos and Sakrison). (b) The impulse response of the Gaussian function representing the 1-D best-fit eye model for print resolution $R = 600$ dpi, and viewing distance $D = 15$ inches.

1.3 Spatial-CIELAB

Spatial-CIELAB (S-CIELAB) filters are used in this article to mimic the human visual system (HVS) to simulate the perceived printed color halftones [11]. To apply the S-CIELAB filtering, an image is first converted to CIEXYZ color space and then linearly transformed to an opponent color space, represented by a luminance channel (O_1), and two chromatic, i.e., red-green (O_2) and blue-yellow (O_3), channels by

$$\begin{cases} O_1 = 0.279X + 0.720Y - 0.107Z \\ O_2 = -0.449X + 0.290Y - 0.077Z \\ O_3 = 0.086X - 0.590Y + 0.501Z. \end{cases} \quad (4)$$

These three images are then filtered by three filters representing the sensitivity functions of the eye. The luminance channel is filtered by a band-pass filter already discussed in Section 1.2 and shown in Fig. 2(a), and the other two channels by two different low-pass filters [12]. As shown in Ref. [12], eye sensitivity function for the blue-yellow channel has a lower cut-off frequency than the one representing the red-green channel. The sensitivity drops to its half at around 20 cycles/degree for red-green and at around 10 cycles/degree for blue-yellow channel. The filtered opponent color space matrices are then converted back to CIEXYZ with inverse matrix operations and thereafter are transformed into CIELAB color space, resulting in three matrices representing L^* , a^* , and b^* components of S-CIELAB. The resulting CIELAB images can be used as both a full-reference and a no-reference quality metric, discussed in Sections 4.1.2 and 4.3.2, respectively.

The remainder of this article is organized as follows. Section 2 provides a brief description of the point-by-point IMCDP method. In Section 3, two important FM halftone properties; blue/green-noise characteristics and dot gain, are studied for point-by-point IMCDP and error diffusion generated halftones. In Section 4, twelve different halftones are compared with each other based on four halftone quality measures for monochromatic and three quality metrics for color halftones. In Section 5, halftone images generated by the examined halftoning methods are illustrated and Section 6 provides a brief summary and conclusion.

2. POINT-BY-POINT IMCDP, FIRST- AND SECOND-ORDER FM HALFTONES

This section starts with a subsection describing the approach to generate first- and second-order FM threshold matrices to halftone monochromatic images. This is followed by another subsection describing how different threshold matrices for different color channels should be designed to obtain independent, dot-on-dot and dot-off-dot printing.

2.1 Point-by-Point IMCDP: Monochromatic

As discussed in Section 1, it is desirable to have halftoning techniques of good quality that are fast in operation. There are different methods reported in literature that are able to generate different halftone structures, dot shapes, and alignments by creating thresholds to be applied to an input image on a point-by-point basis [1, 13].

The point-by-point IMCDP halftoning algorithm generates image-independent threshold matrices creating first- and second-order FM halftones. This point-by-point technique is based on an iterative halftoning method called Iterative Method Controlling the Dot Placement (IMCDP), thoroughly described in Ref. [7]. The IMCDP method, briefly described, starts with a blank image the same size as the original image and places the first dot at the position where the original image is darkest, i.e., holds the maximum pixel value. It ensures that this position will not be found as the maximum again by setting a very small number at this position in the original image. The effect of this dot placement is then fed back into the halftoning process by subtracting a neighborhood of the position of the found maximum by a filter. We refer to this as the feedback process from now on. By doing that, the probability to find the next maximum in that neighborhood is reduced. This process proceeds and in each iteration one dot is placed at the position of the maximum pixel value and the effect of the placed dot is fed back by an appropriate filter until a predetermined number of dots are placed. In point-by-point IMCDP, instead of an original image as the input, an image (a matrix) holding uniformly distributed pseudo-random numbers is used. The generated threshold matrix will be the same size as this input random matrix. Another difference is that, in IMCDP, a pre-decided number

of pixels are iteratively set to 1 in the initial blank image. In this point-by-point method, the initial blank matrix is iteratively filled by successive integers starting from 1 to the number of elements in the matrix. For example, for a 256×256 threshold matrix, the matrix is filled by successive integers from 1 up to $256^2 = 65,536$. In Ref. [1], it is also thoroughly described how one can generate the threshold matrix in a way that dot distributions (or dot shapes) on both sides of the mid-tone, i.e., 50%, are symmetric. Since the threshold matrices are usually smaller than the images being halftoned, they need to be tiled in order to cover the whole image. Thus, cutting the feedback filter, when outside the border of the matrix will cause boundary/tiling artifact, where the matrices are tiled. This issue is managed in this method by performing a wraparound process [1]. The most important part of the generation of a threshold matrix is the feedback filter. Different appropriate filters can generate different halftone structures, shapes, and alignments [1]. In order to generate well-formed first-order FM halftones having blue-noise characteristics, the following Gaussian function is used to perform the feedback process

$$f(m, n) = e^{-(m^2+n^2)/2\sigma^2}. \quad (5)$$

In order to find an optimal standard deviation (σ), a number of different threshold matrices were generated using different σ 's. The obtained halftone images using these threshold matrices were then compared to the original image using the mean squared error, discussed in Section 4.1.1. We found that $\sigma = 1.7$ generates threshold matrices that result in the smallest average mean squared error when tested on many different test images. Therefore, in this article, to generate first-order FM threshold matrices, we use the filter in Eq. (5) with $\sigma = 1.7$. In Section 3.1, it is shown that these first-order FM halftones have blue-noise characteristics. A number of halftone images using this threshold matrix are illustrated in Section 5.

As discussed earlier, another important aspect of point-by-point IMCDP is that, it can also generate threshold matrices obtaining second-order FM, i.e., green-noise [14], halftone structures. This is achieved by using the following filter in the feedback process

$$h(m, n) = e^{-(m^2+n^2)/2\sigma_1^2} - e^{-(m^2+n^2)/2\sigma_2^2}. \quad (6)$$

The filter in Eq. (6) is a Gaussian function subtracted from another Gaussian function with larger standard deviation, i.e., $\sigma_1 > \sigma_2$. By this filter, the pixel values around the found maximum are decreased with a radius decided by σ_1 . After the single dots have been distributed, the dots start to cluster and the maximum size of the clustered dots will depend on σ_2 . By appropriate choices of σ_1 and σ_2 , it is possible to meet a specific demand for the size of the clustered dot at a certain gray level. In this study, we use $\sigma_1 = 1.7$ as in the first-order FM. In order to generate second-order FM halftones with three different clustered dot sizes, i.e., small, medium, and large, we use $\sigma_2 = 0.5$, $\sigma_2 = 0.6$, and $\sigma_2 = 0.7$, respectively. In Section 3.1, it is shown that these generated

second-order FM halftones have green-noise characteristics. Several halftone images using these threshold matrices are illustrated in Section 5.

2.2 Point-by-Point IMCDP: Color

Since periodic clustered halftones (i.e., AM) usually suffer from moiré, second-order FM halftones provide a solution because of their stochastic nature of distributing the clustered dots. FM color halftoning is usually performed by halftoning each color channel independent of the other channels. However, dot-off-dot halftone structures have been proposed in literature, with the advantage of producing smoother halftones and a larger gamut while using less ink compared to independent color halftoning [7, 13]. In order to explain how our proposed point-by-point method is able to generate dot-off-dot halftone structures, let us for simplicity focus on two colorants, for example, cyan and magenta and call the threshold matrices for these two channels T_c and T_m , respectively. In the CMY print, the yellow channel is usually halftoned independent of the other two because of its low contrast on white paper [7, 15]. If identical threshold matrices are used for C and M, i.e., $T_m = T_c$, the dots in C and M channels will be placed precisely at the same positions producing a dot-on-dot printing. If two different threshold matrices T_c and T_m are generated and used for C and M, different colorant dots are placed independent of each other, although the same filters and parameters have been used to generate both matrices. The reason is that the input image to point-by-point IMCDP is a random matrix and it will change from time to time. On the other hand, if one of the threshold matrices, for example, T_c , is generated and the other one is computed by $T_m = 1 - T_c$, provided T_c is normalized between 0 and 1, then the overlap between the two colorants will not occur as long as the sum of their coverage does not exceed 100% [1]. Note that, in the operation $1 - T_c$, by 1 we mean a matrix of ones the same size as T_c . In order to illustrate the differences between these three strategies, a CM color patch with 50% cyan and 50% magenta is halftoned by second-order FM ($\sigma_1 = 1.7$ and $\sigma_2 = 0.6$) using dot-on-dot, independent, and dot-off-dot printing. Figure 3(a) shows the dot-on-dot printing, being generated using $T_m = T_c$. In this case, there is 50% blue (magenta on cyan) and 50% non-printed area. In Fig. 3(b), independent structure is illustrated, for which the threshold matrices were created by running the same code twice, obtaining two different threshold matrices T_c and T_m . In this halftone, there is 25% cyan, 25% magenta, 25% blue, and 25% non-printed area, according to Demichel's equations [16]. Fig. 3(c) shows the dot-off-dot printing, being generated using $T_m = 1 - T_c$. Since the sum of the coverage is equal to 100%, it is still possible to completely avoid blue dots. Consequently, in this color halftone, there is 50% cyan and 50% magenta. Note that, although the original color patch is the same, the three color halftones in Fig. 3 will result in different colors, especially when printed [7, 15]. More color halftone images will be illustrated in Section 5. In order to achieve dot-off-dot

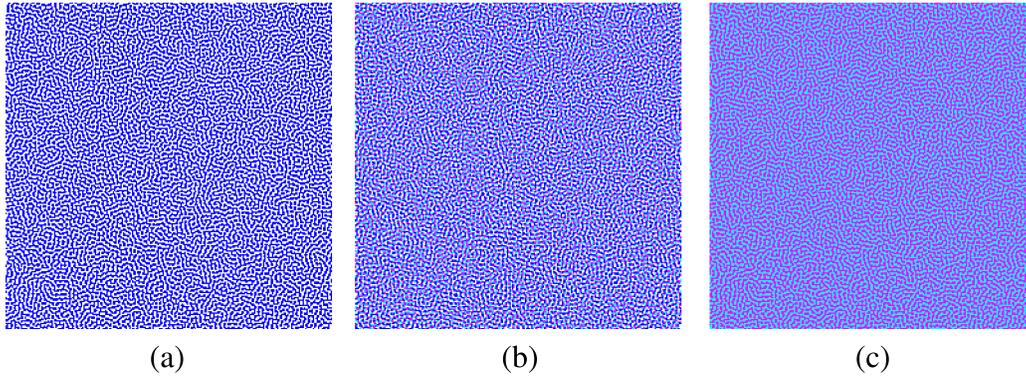


Figure 3. Second-order FM ($\sigma_1 = 1.7$ and $\sigma_2 = 0.6$) color halftones with 50% cyan and 50% magenta. The C and M channels have been halftoned (a) using dot on dot (b) independently (c) using dot off dot.

structures for three colorants, for example, C, M, and Y, we use $T_m = 1 - T_c$ and $T_y = 2 \cdot |1/2 - T_c|$ [1].

3. HALFTONE PROPERTIES

As mentioned in Section 2, well-formed first- and second-order FM halftones should have blue- and green-noise characteristics, respectively. First in this section, we study whether or not the halftones obtained by point-by-point IMCDP generated threshold matrices possess these characteristics. Another important factor that should be taken into account in printing, especially at high resolutions, is dot gain. In this section, we also study the dot gain for our proposed first- and second-order FM halftones employing the printer model discussed in Section 1.1 using different ρ 's. Comparisons in terms of these two properties between our proposed and error diffusion halftones using two different error filters are also reported in this section.

3.1 Blue- and Green-noise Characteristics

Well-formed first-order FM halftones have blue-noise characteristics, which means that the quantization noise is shifted toward higher frequencies. Well-formed second-order halftones, on the other hand, possess green-noise characteristics, meaning that the quantization noise contains smaller low-frequency and high-frequency components. One of the approaches to assess this property is to study the quantization noise spectrum (QNS), i.e., the Fourier spectrum of the difference between the original and the halftone image, as defined in Eq. (7).

$$Q(u, v) = |\text{fftshift}(\text{fft2}(g(m, n) - b(m, n)))|, \quad (7)$$

where $g(m, n)$ and $b(m, n)$ are the original and the halftone image, respectively, and fft2 denotes the two-dimensional Fourier transform. The function fftshift is used to shift the zero-frequency to the center of the spectrum, making the lower frequencies be concentrated around the center. Since the dc-term usually dominates the values of a spectrum, the log transformation, as shown in Eq. (8), is used to illustrate the QNS.

$$\text{Log}_{\text{transform}} Q(u, v) = \log(1 + Q(u, v)). \quad (8)$$

In order to illustrate the difference between first- and second-order FM halftones, a patch at 25% coverage has been halftoned by four different halftoning approaches, namely; first-order FM ($\sigma = 1.7$) and second-order FM using ($\sigma_1 = 1.7$ and $\sigma_2 = 0.5$), ($\sigma_1 = 1.7$ and $\sigma_2 = 0.6$), and ($\sigma_1 = 1.7$ and $\sigma_2 = 0.7$), which are shown in Figure 4 (top row). Fig. 4 (bottom row) shows the log transformation of the QNS for these four halftones. It is worth pointing out that, in this article, we always use $\sigma = 1.7$ to generate first-order FM halftones by point-by-point IMCDP, and therefore from now on, we skip specifying the σ value for it. To generate second-order FM halftones, in this article, we always use $\sigma_1 = 1.7$, and therefore different second-order FM halftones, i.e., small, medium, and large clustered dot size, are only referred to as by their σ_2 value. Note that, in Fig. 4 (bottom row), the darker the tone, the lower the values. There are two criteria for a well-formed FM halftone. The first one is that the quantization noise is shifted toward the higher frequencies, and the other one is that the dots are placed homogeneously which results in a circularly symmetric QNS. Thus, the first observation is that all four spectra in Fig. 4 are circularly symmetric, implying homogeneous dot placement. As seen in Fig. 4(a), the noise is shifted toward the higher frequencies where the human eye is less sensitive. Fig. 4(b)–(d) show that the strongest noise is shifted to mid-frequencies. Smaller clustered dot size makes the quantization noise be shifted toward higher frequencies, which was expected because the smaller clustered dot size means that the second-order FM approaches the first-order FM. Whether or not the second-order FM is able to shift the noise toward the frequencies where the human eye is not sensitive depends on the viewing distance and the print resolution, which will be discussed in more detail in Section 4.1.1.

In order to have a better analysis of the characteristics of the proposed first- and second-order halftones, they are compared with the well-known error diffusion halftoning. Therefore, the patch at 25% has been halftoned also by original error diffusion (ED) with Floyd–Steinberg (FS) and Jarvis, Judice, and Ninke (J) error filters, shown in Figure 5(a) and (c), respectively. The halftone created by ED

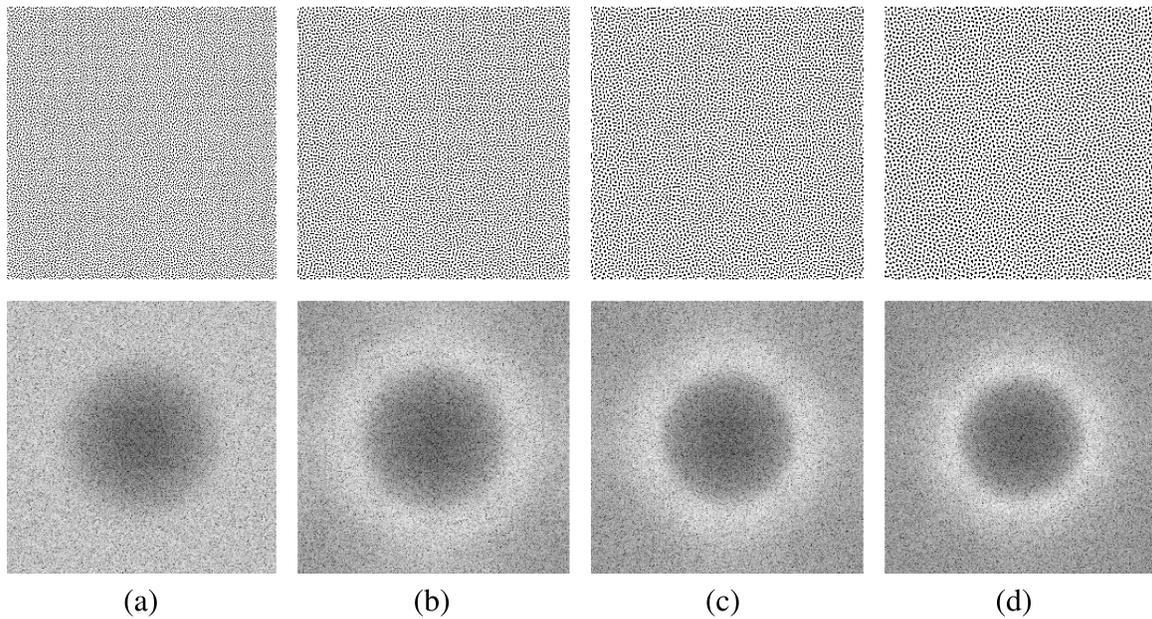


Figure 4. Halftones at 25% coverage together with their corresponding quantization noise spectrum. The halftoning methods are: (a) first-order FM, (b) second-order FM ($\sigma_2 = 0.5$), (c) second-order FM ($\sigma_2 = 0.6$), and (d) second-order FM ($\sigma_2 = 0.7$).

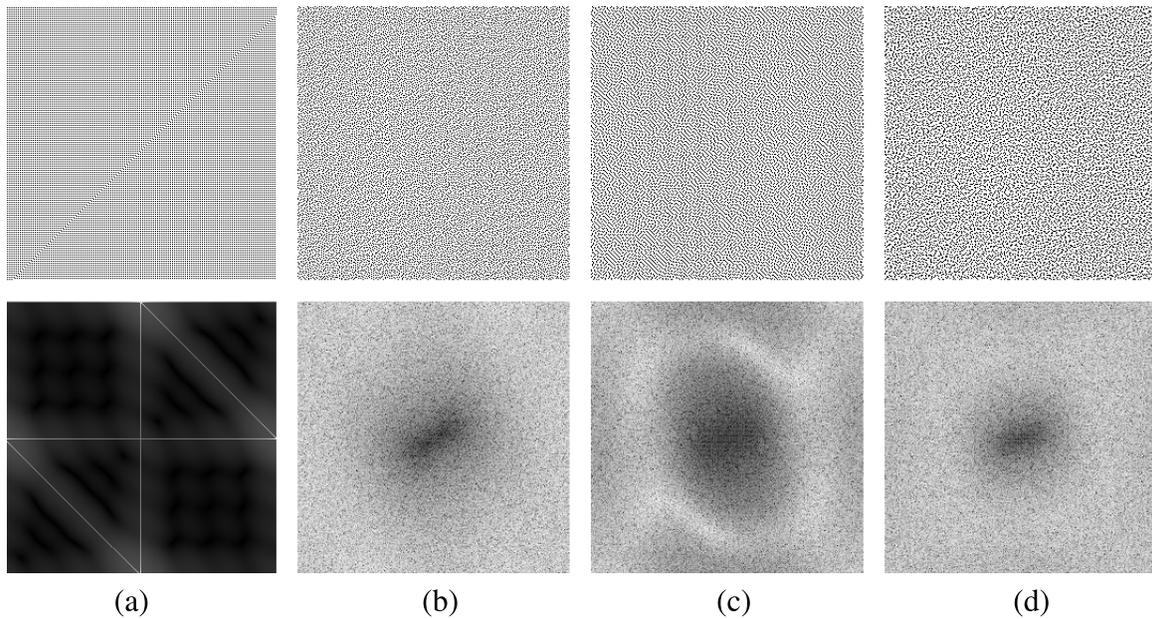


Figure 5. Halftones at 25% coverage together with their corresponding quantization noise spectrum. The halftoning methods are: (a) error diffusion (FS), (b) modified error diffusion (FS), (c) error diffusion (J), and (d) modified error diffusion (J).

(FS) is highly structured, which is also reflected in its QNS. Although, the halftone generated by ED (J) looks better than that created by ED (FS), it is still not well formed, because its QNS is not circularly symmetric. Another important observation is that the ED (J) method behaves more like a second-order FM halftoning because the quantization noise has been shifted toward mid-frequencies. In order to improve error diffusion in this aspect, we have done a simple modification. In the original error diffusion, the

threshold value is always 0.5, but in our modification we instead use a random number between 0.25 and 0.75. This modification reduces the structures but results in noisier halftones. From now on, we use the abbreviation ED for the original non-modified error diffusion and MED for this simple modified version of error diffusion. The error filters used in the method are referred to as FS (Floyd–Steinberg) and J (Jarvis, Judice, and Ninke). Fig. 5(b) and (d) show the halftone results created by MED (FS) and MED (J),

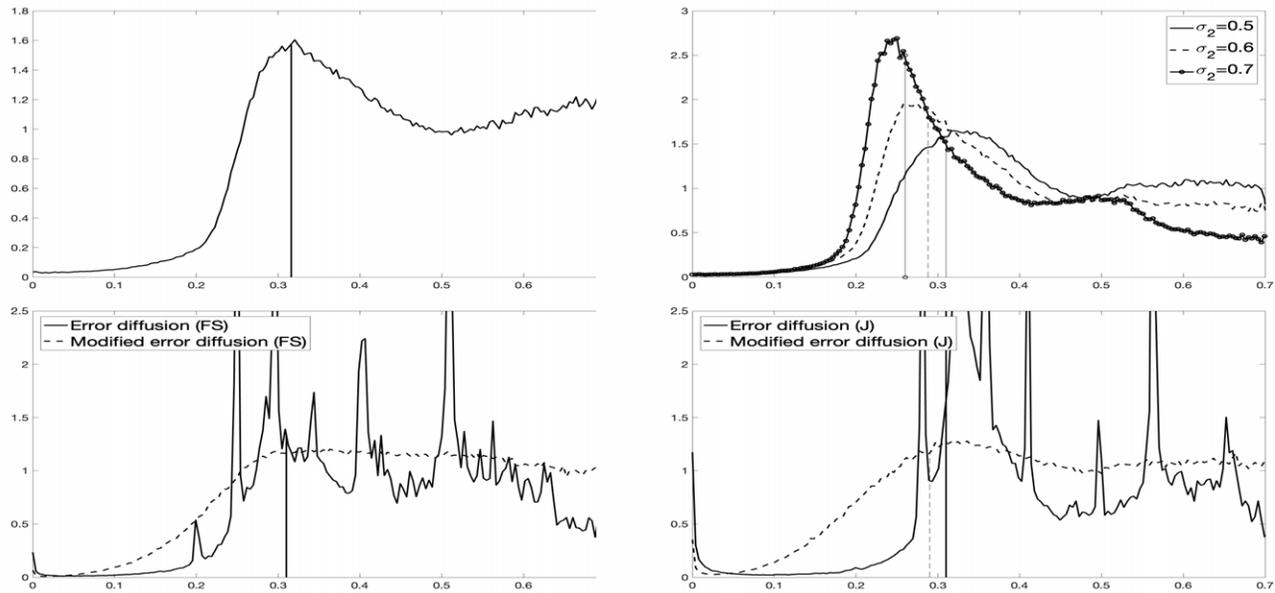


Figure 6. RAPS curves for a 10% halftone patch halftoned by: (top left) first-order FM. (Top right) Second-order FM ($\sigma_2 = 0.5, 0.6,$ and 0.7). (Bottom left) Original and modified error diffusion (FS). (Bottom right) Original and modified error diffusion (J).

respectively, and their corresponding QNS. The QNS in (b) verifies that this halftone is more well formed than ED (FS). The QNS in (d) shows that the halftone created by MED (J) behaves more like a first-order FM than created by ED (J). Note that, although the MED methods create more well-formed halftones than ED methods, their results are still not very well formed because of the lack of circular symmetry, which is also reflected in their radially averaged power spectrum, discussed shortly.

Another approach to analyze whether a first- or second-order FM halftone is well formed is to illustrate its radially averaged power spectrum (RAPS) curve. The peak of the RAPS curve of a well-formed FM halftone is at its principal frequency [2, 17]. For a first-order FM halftone, the principal frequency is $f_g = \sqrt{g}$ for $0 < g \leq 1/2$ and $f_g = \sqrt{1-g}$ for $1/2 < g < 1$, where g is the gray level [2, 17]. For a second-order FM halftone, the principal frequency is $f_g = \sqrt{g/M}$ for $0 < g \leq 1/2$ and $f_g = \sqrt{(1-g)/M}$ for $1/2 < g < 1$, where g is as before and M is the average clustered dot size [14]. Figure 6 (top row) shows the RAPS curves and the corresponding principal frequencies for first- and second-order FM halftones at 10% coverage. The top-left graph shows the RAPS curve for the proposed first-order FM halftone and the principal frequency at $\sqrt{0.1} = 0.316$. The top-right graph shows the RAPS curves for the second-order FM halftones with ($\sigma_2 = 0.5$), ($\sigma_2 = 0.6$), and ($\sigma_2 = 0.7$). The average clustered dot sizes have been computed by first labeling the binary halftones using 8-connectivity and then computing the area of each clustered dot and taking their average. The average clustered dot size for these second-order FM halftones at 10% are 1.03, 1.21, and 1.48, yielding the principal frequencies of 0.311, 0.288, and 0.260, respectively. This graph also shows that with smaller clustered dot size the second-order FM approaches first-order FM. The RAPS

curves for halftones at 10% halftoned by ED and MED methods using FS and J filters are illustrated in Fig. 6 (bottom row) together with the corresponding principal frequencies. In order to better illustrate the RAPS curves, these curves were clipped at 2.5 because the peaks for ED methods were very high. These RAPS curves reveal that MED methods generate halftones with better blue-noise characteristics, but they are still not as well formed as the proposed first-order FM.

3.2 Dot Gain

The printed dots usually become larger than their size in the corresponding digital bitmap. This phenomena is referred to as dot gain [18]. Dot gain is basically divided into two main categories; physical/mechanical and optical dot gain. Physical dot gain refers to the phenomena where the printed dots become physically larger, while the optical dot gain is because of the diffusion of the light in the paper or substrate [18]. There are many factors and parameters affecting dot gain. Some of the most important parameters are paper quality, type of printer or print press, print resolution, and halftoning. Dot gain causes the printed halftone image to be darker than its corresponding digital bitmap. Hence, the effect of dot gain is usually compensated for prior to halftoning. However, there have been halftoning methods, such as least-squares model-based halftoning reported in Ref. [9], that take into account the dot gain effect within the halftoning algorithm. Since in the first-order FM halftones, the printed dots are isolated, the dot gain effect is obviously greater than in AM and second-order FM halftones. In this article, we are going to employ the printer model shortly discussed in Section 1.1 to simulate the dot gain effect. The reason using a printer model to simulate dot gain is that we want to give a general study not necessarily

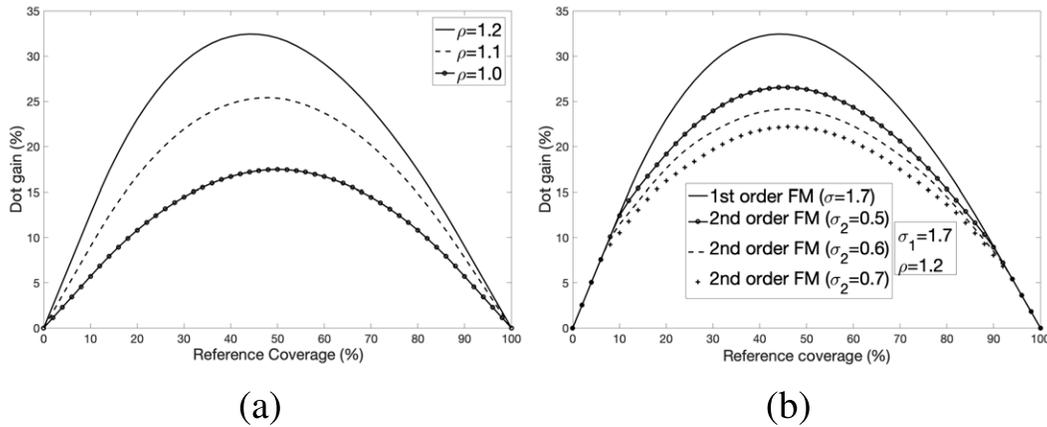


Figure 7. (a) Dot gain curves for the proposed first-order FM for $\rho = 1.0$, $\rho = 1.1$, and $\rho = 1.2$. (b) Dot gain curves for the proposed first- and second-order FM with $(\sigma_2 = 0.5)$, $(\sigma_2 = 0.6)$, and $(\sigma_2 = 0.7)$ for $\rho = 1.2$.

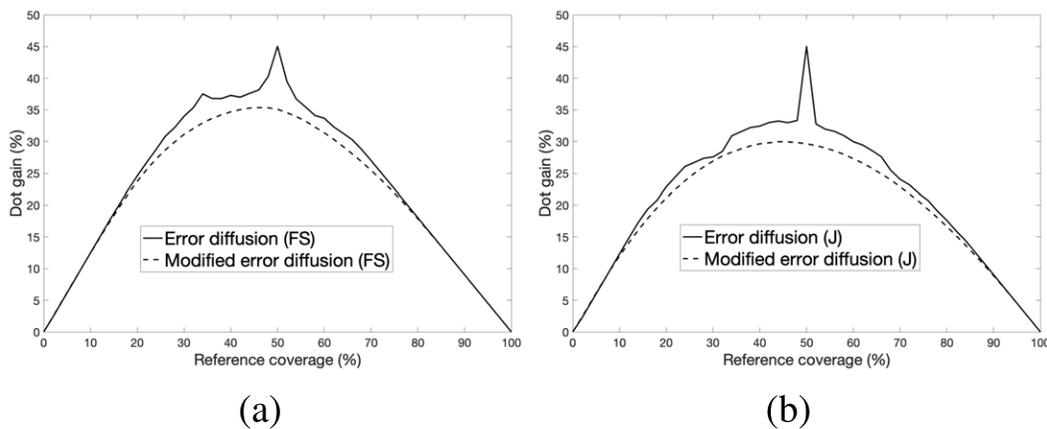


Figure 8. Dot gain curves for original and modified error diffusion for $\rho = 1.2$. (a) ED (FS) and MED (FS). (b) ED (J) and MED (J).

restricted to one type of printer or print press, paper quality etc. Figure 7(a) shows three dot gain curves for the proposed first-order FM halftones using $\rho = 1.0$, $\rho = 1.1$, and $\rho = 1.2$. As expected, the larger the ρ , the larger the dot gain. Different ρ values could represent the dot gain in different printing situations. For instance, $\rho = 1.2$ gives a dot gain at 50% very close to a first-order FM halftone printed at 1200 dpi by an offset print press on a coated paper [19]. Fig. 7(b) shows four dot gain curves for first- and second-order FM with $(\sigma_2 = 0.5)$, $(\sigma_2 = 0.6)$, and $(\sigma_2 = 0.7)$ simulated using $\rho = 1.2$. As expected, the smaller the clustered dot size, the larger the dot gain. As stated earlier, in order to take into account the effect of dot gain, the original con-tone image is commonly compensated for dot gain prior to halftoning process. Let us explain this by an example. Assume that a patch at 50% is supposed to be printed using first-order FM halftoning in a print condition represented by $\rho = 1.2$. If the coverage in the digital bitmap is 50%, according to the corresponding dot gain curve in Fig. 7, i.e., the solid-line curve, the coverage of the printed halftone will be 82% because the dot gain is 32% at 50% reference coverage. In order to get 50% coverage after print, the coverage in the digital bitmap should be 24%. Therefore, each pixel value in

the original image is compensated for dot gain the same way before being halftoned. This will result in an image lighter than the original, which will have the same tonal values as the original image after being halftoned and printed. Figure 8 shows the dot gain curves for ED and MED methods with FS and J filters for $\rho = 1.2$. As can be seen in this figure, there is a roof-like peak at 50% for original error diffusion. The reason is that at 50%, ED will result in a checkerboard structure, resulting in very high dot gain. The modified error diffusion has, on the other hand, a more symmetrical dot gain curve. Another observation is that ED (J) results in a lower dot gain than ED (FS). The reason is, as discussed in Section 3.1, that ED (J) behaves more like a second-order FM halftone.

4. QUALITY ASSESSMENTS

In this section, we analyze the quality of the halftones created by point-by-point IMCDP method. Four different such halftones are studied, including first-order FM, and second-order FM with small, medium, and large clustered dot size obtained by $(\sigma_2 = 0.5)$, $(\sigma_2 = 0.6)$, and $(\sigma_2 = 0.7)$, respectively. Besides comparing these four halftones with each other, their quality is also compared with four error

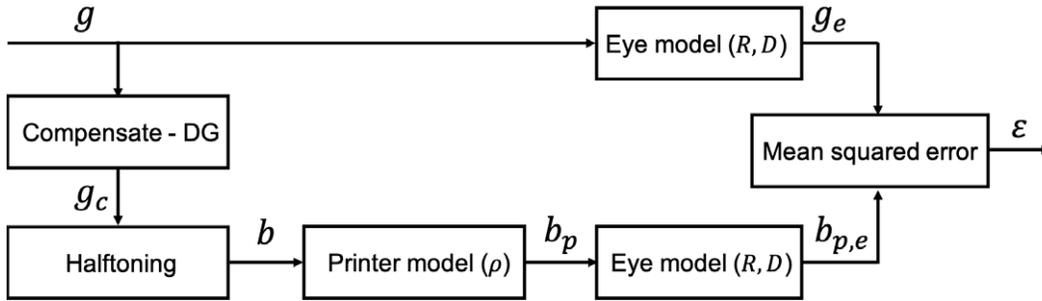


Figure 9. The mean squared error is defined as the average squared difference between g_e and $b_{p,e}$.

diffusion halftones, including non-modified error diffusion using Floyd–Steinberg filter (ED (FS)), and Jarvis, Judice, and Ninke filter (ED (J)), and a simple modification of error diffusion, discussed in Section 3.1, using these two filters. These two modified approaches are abbreviated as MED (FS) and MED (J) in the text. In order to give a better indication of the obtained quality values, we also study clustered and dispersed dot ordered dithering, using 8×8 and 16×16 threshold matrices, representing $8^2 + 1 = 65$ and $16^2 + 1 = 257$ levels of gray, respectively. The clustered dot ordered dithering using a 16×16 threshold matrix, although resulting in many gray levels, will create large halftone dots. Consequently, this halftone will mostly result in quite poor quality values, but it is still studied to give a better understanding of the quality values when compared with other halftones. For the dispersed-order dithering, Bayer’s approach to create threshold matrices is used [20]. The clustered and dispersed dot order dithering are referred to as clustered dot ordered dithering (8×8 or 16×16) and Bayer (8×8 or 16×16), respectively. Four important quality measures for monochromatic and three measures for color halftones are being used to evaluate the halftones.

4.1 Similarity

In most applications, one of the most important goals of the reproduction process is to achieve a reproduction that resembles the original image as much as possible. In this section, mean squared error is used to study how close the printed monochromatic halftones are to the original image. For the color halftones, spatial-CIELAB ΔE is used to calculate the perceived difference between the printed color halftone and the original image.

4.1.1 Mean Squared Error

One of the measures to assess halftone image quality is to calculate the squared error between the original con-tone image and the halftone image as defined in Ref. [9]. As discussed in Section 1, there are two important models that should be taken into account when calculating the squared error between a con-tone and a halftone image. The first one is the printer model to simulate how a digital halftone image will look like after being printed. The second model is the human eye model, which should also be taken into account because it is the human observer that will judge the quality.

Figure 9 shows the procedure we propose for calculating the mean squared error between the original and the halftone image. In this figure, g and g_c are the original image and the dot gain compensated original image, respectively. As discussed in Section 3.2, the original con-tone image is supposed to be compensated for dot gain prior to halftoning, which is performed using the corresponding dot gain curve. In Fig. 9, b and b_p denote the halftone image and the halftone image after the printer model has been applied, respectively. The mean squared error is then defined as the average squared difference between the perceived original and the perceived printed halftone image, i.e., g_e and $b_{p,e}$, by

$$\varepsilon = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (g_e(m, n) - b_{p,e}(m, n))^2, \quad (9)$$

where M and N are the number of the pixels in the height and the width of these images. If the image pixels are represented by 8 bits, the largest mean squared error will be $255^2 = 65,025$, which, for instance, occurs between a completely white and a completely black image. Note that, the quality measure we propose here is a bit different than the one suggested in Ref. [9]. In their strategy, they do not compensate the original image prior to halftoning, as it is supposed to be taken into account inside the halftoning algorithm. In Ref. [9], it is stated that the error criterion used in the least square model-based halftoning can be used to evaluate the quality of any halftoning algorithm. Since the dot gain is different for different halftones, we believe that this quality measure cannot be used for all halftoning algorithms without involving the dot gain compensation and therefore proposing the modification presented in Fig. 9. The printer model used is the dot-overlap model shortly described in Section 1.1. In this study, two different print parameters are used, namely; $\rho = 1.1$ and $\rho = 1.2$. Since our proposed halftoning method operates point-by-point, we believe that it is more suitable for high resolution prints, and therefore we use the following two print resolutions, namely; $R = 600$ dpi for the printer parameter $\rho = 1.1$ and $R = 1200$ dpi for $\rho = 1.2$. The eye model is the filter described in Section 1.2, using the viewing distance of $D = 15$ inches.

Figure 10 shows the mean squared error for twenty-one 512×512 pixels constant/uniform patches at 5%–95% in a

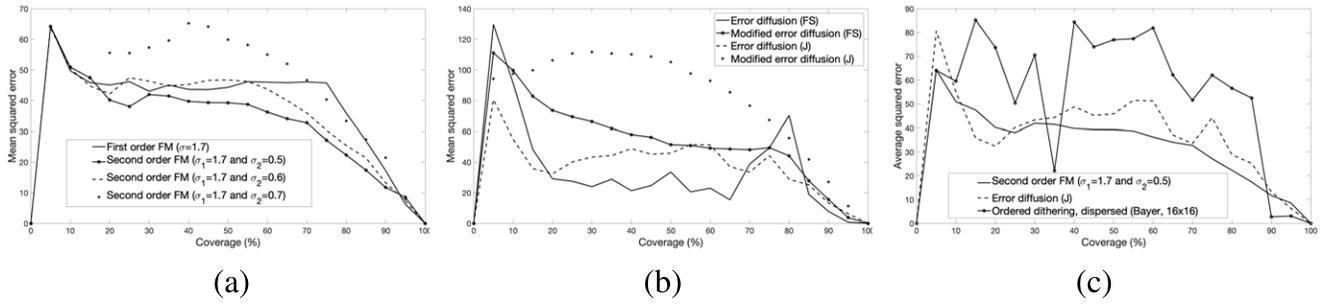


Figure 10. Mean squared error for constant patches between 5% and 95% with a step of 5% for $\rho = 1.1$, $R = 600$ dpi, and $D = 15$ inches. (a) Proposed first- and second-order FM. (b) Error diffusion and modified error diffusion with FS and J filters. (c) Comparing the best halftone in (a) and (b) and dispersed dot ordered dithering [Bayer 16×16].

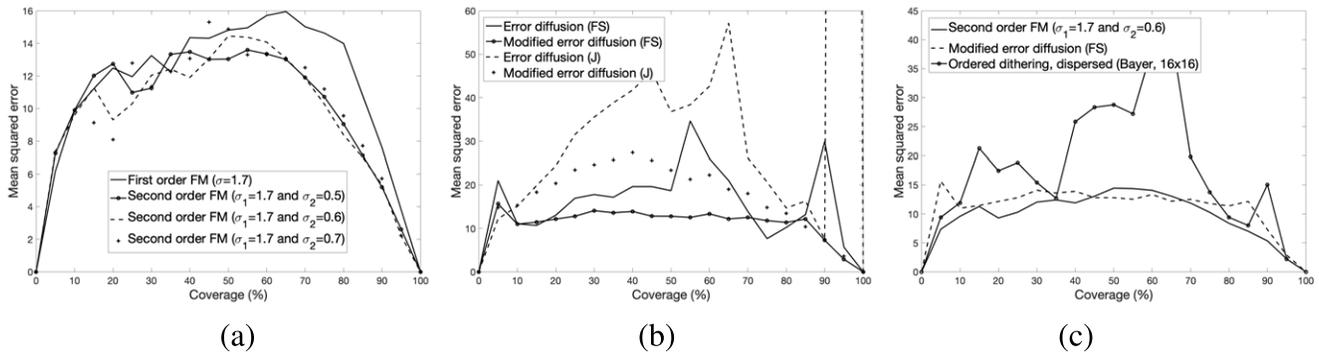


Figure 11. Mean squared error for constant patches between 5% and 95% with a step of 5% for $\rho = 1.2$, $R = 1200$ dpi, and $D = 15$ inches. (a) Proposed first- and second-order FM. (b) Error diffusion and modified error diffusion with FS and J filters. (c) Comparing the best halftone in (a) and (b) and dispersed dot ordered dithering [Bayer 16×16].

step of 5% coverage. The print parameters are $\rho = 1.1$ and $R = 600$ dpi and the viewing distance is 15 inches. In (a), the proposed first- and second-order FM halftones are compared with each other. As seen in this figure, the second-order FM with $\sigma_2 = 0.5$ results in slightly smaller error. In (b), ED and MED methods with FS and J filters are compared with each other, and in general ED (J) is better, although in the mid-tones ED (FS) is slightly better. In (c), we compare the best method among our proposed methods with the best method among the error diffusion methods with the best ordered dithering method, which is the dispersed dot ordered dithering (Bayer 16×16). The clustered dot order dithering (16×16) resulted in very large errors compared to the others. As can be seen in (c), second-order FM ($\sigma_2 = 0.5$) is better for most of the patches and Bayer (16×16) is the worst.

Figure 11 shows the mean squared error for the same patches with $\rho = 1.2$, $R = 1200$ dpi, and the same viewing distance. As seen in (a), all four methods among our proposed FM halftones result in almost the same errors, and the second-order FM ($\sigma_2 = 0.6$) results in slightly smaller errors. As seen in (b), MED (FS) is the best method among the error diffusion methods. Note that, ED (J) that was the best method for $\rho = 1.1$ is the worst one for $\rho = 1.2$. The error for ED (J) was around 900 for a patch at 95%, and therefore its graph has been clipped in Fig. 11. A look at ED (J):s dot gain curve in Fig. 8(b) reveals that 95% coverage

will be compensated to approximately 50%. Since ED at 50% results in a checkerboard structure, its dot gain has a rapid change at 50%. As ED halftones at, for example, 49% and 50% behave very differently with regards to the dot gain, small changes in reference coverage around 50% might cause large errors. In (c), it can be seen that the second-order FM ($\sigma_2 = 0.6$) results in slightly smaller error than MED (FS), while both of them result in much smaller error than Bayer (16×16), which was the best among the ordered dithering methods. The clustered dot order dithering, again, resulted in very large errors. In order to make a better comparison between the halftoning methods, we also halftoned eight different 512×512 pixels test images and calculated the mean squared errors for ($\rho = 1.1$, $R = 600$ dpi, $D = 15$ inches) and ($\rho = 1.2$, $R = 1200$ dpi, $D = 15$ inches). The test images are called *Sari1*, *Sari2*, *Contrast*, *Ramp*, *Curve*, *Lena*, *Glass*, and *Kids*. The halftone images of the first four test images are shown in Section 5. The test image *Curve* is a grayscale image containing curvy stripes of various width with two gray tones; 0.1 and 0.9. The test image *Lena* is the grayscale version of the famous image *Lena* used in image processing. The test images *Glass* and *Kids* are both demo images in MATLAB, being squared and resized to 512×512 pixels. The mean squared error for ($\rho = 1.1$, $R = 600$ dpi) and ($\rho = 1.2$, $R = 1200$ dpi) are shown in Tables I and II, respectively. In the next to the last column, titled **Av. 1**, the average value of each row is shown.

Table I. Mean squared error for eight test images ($\rho = 1.1$, $R = 600$ dpi, $D = 15$ inches).

	Sari1	Sari2	Contrast	Ramp	Curve	Lena	Glass	Kids	Av. 1	Av. 2
FM-1st	49.5	72.6	53.9	38.0	58.3	49.9	38.3	20.4	47.6	43.2
FM-2nd (0.5)	42.4	69.3	51.2	32.0	62.1	41.7	32.4	14.6	43.2	37.7
FM-2nd (0.6)	49.1	74.8	57.1	35.4	64.1	48.0	38.2	17.2	48.0	41.7
FM-2nd (0.7)	65.1	88.7	68.1	43.5	72.7	60.5	48.2	21.9	58.6	51.5
ED-FS	27.3	57.1	42.4	32.7	64.7	35.5	23.4	28.7	39.0	37.6
MED-FS	51.0	69.3	56.4	50.9	66.4	48.0	45.6	18.8	50.8	54.6
ED-J	50.3	76.5	64.0	35.4	111.3	60.8	36.9	32.8	58.5	45.6
MED-J	98.1	100.4	86.4	78.9	95.5	85.1	78.5	28.2	81.4	83.8
Clustered-8	2281	1564	1455	1499	559	1864	1633	624	1435	1538
Clustered-16	7275	5086	4870	5165	2318	6195	5453	2363	4841	5263
Dispersed-8	80.8	90.8	65.9	51.7	55.4	79.6	55.3	30.5	63.8	61.3
Dispersed-16	78.3	90.1	68.2	51.1	51.1	79.6	54.0	29.3	62.7	59.7

Table II. Mean squared error for eight test images ($\rho = 1.2$, $R = 1200$ dpi, $D = 15$ inches).

	Sari1	Sari2	Contrast	Ramp	Curve	Lena	Glass	Kids	Av. 1	Av. 2
FM-1st	16.4	21.0	18.5	9.8	23.2	17.2	13.0	7.9	15.9	13.3
FM-2nd (0.5)	13.5	19.2	19.0	8.3	28.6	14.9	11.1	6.2	15.1	12.0
FM-2nd (0.6)	13.4	20.4	18.9	8.3	32.1	15.1	11.6	6.0	15.7	12.0
FM-2nd (0.7)	15.2	22.1	20.2	8.0	36.4	15.4	12.4	6.5	17.0	12.6
ED-FS	17.9	14.6	17.0	9.2	25.8	21.4	18.6	10.7	16.9	17.1
MED-FS	12.8	12.1	13.4	8.8	17.6	14.1	11.6	6.6	12.1	12.0
ED-J	36.0	30.5	38.8	18.3	55.4	41.9	35.2	15.3	33.9	63.2
MED-J	21.7	20.2	19.9	13.5	32.1	22.7	18.4	7.4	19.5	18.7
Clustered-8	65.9	65.4	47.2	35.4	45.6	62.4	52.3	20.9	49.4	47.1
Clustered-16	2442	1667	1593	1632	687.5	2018	1776	671	1561	1682
Dispersed-8	43.8	34.6	25.1	18.0	32.0	35.5	26.4	17.8	29.2	26.2
Dispersed-16	33.0	28.7	24.2	14.7	22.4	30.3	24.1	14.7	24.0	20.6

In the last column, titled **Av. 2**, the average also includes the squared error of halftoning the constant patches, whose results were already shown in Figs. 10 and 11. As shown in these tables, among our proposed methods, second-order FM ($\sigma_2 = 0.5$) is best in average at both print resolutions. Among the error diffusion methods, ED (FS) is better for $R = 600$, but MED (FS) is better for $R = 1200$. Among the ordered dithering methods, Bayer (16×16) is the best method. In average, among all halftones, ED (FS) is the best method for $R = 600$ while MED (FS) is the best one for $R = 1200$. For the three test images *Sari1*, *Sari2*, and *Contrast*, ED (FS) is much better than second-order FM ($\sigma_2 = 0.5$) for $R = 600$. This seems to be valid for images containing most of the details in the mid-tones. When the test image contains big areas of highlights and/or shadows, ED (FS) loses its superiority, and sometimes it is even worse than its modified version and our proposed methods, for instance, for test images *Curve* and *Kids*. As for the print resolution $R = 1200$, second-order FM ($\sigma_2 = 0.5$) is slightly

better than ED (FS), and slightly worse than MED (FS). Note that, although ED (J) was the best error diffusion method for constant patches printed at $R = 600$, it is quite poor when it comes to regular test images. According to the results, it can be concluded that the error diffusion methods behave quite differently for different types of images. While one of the error diffusion methods is much better for a specific type of image, it is much worse for other types of images. Among our proposed methods, second-order FM ($\sigma_2 = 0.5$) is best for most of the test images, and whenever it is not the best, the difference to the best one is not huge. All four ordered dithering methods are worse than others, as expected, and clustered dot ordered dithering (16×16) always results in the largest mean squared error. Another observation is that Bayer (16×16) is only slightly better than Bayer (8×8), although the former one is able to represent 257, and the latter one only 65 gray levels. This shows that the number of gray levels being represented does not affect the mean squared error considerably. It is worth mentioning that we

Table III. S-CIELAB ΔE_{ab} for eight test color images ($\rho = 1.1$, $R = 600$ dpi, $D = 15$ inches).

	Sari1	Sari2	Contrast	Ramp	Curve	Lena	Greens	Peppers	Av.
FM-1st	10.8	8.0	9.1	14.3	13.6	18.1	11.8	15.7	12.7
FM-1st (dot off dot)	10.8	8.0	9.1	14.3	13.5	18.0	11.7	15.6	12.6
FM-2nd (0.5)	11.6	8.6	10.0	15.8	14.9	20.2	13.1	17.8	14.0
FM-2nd (0.6)	12.3	9.0	10.6	16.8	16.0	21.6	14.0	18.9	14.9
FM-2nd (0.7)	12.9	9.4	11.4	17.8	16.9	22.7	14.7	20.0	15.7
ED-FS	10.0	7.5	7.9	13.1	13.4	16.5	10.7	14.1	11.7
MED-FS	10.5	7.8	8.7	13.9	13.0	17.2	11.1	14.8	12.1
ED-J	11.1	8.6	9.0	14.7	15.3	20.1	13.2	16.6	13.6
MED-J	11.6	8.6	9.8	15.6	15.5	20.0	13.1	17.1	13.9
Clustered-8	17.2	12.5	15.6	24.7	23.6	32.0	20.0	27.4	21.6
Clustered-16	20.7	15.1	24.3	26.4	25.9	55.4	25.2	39.2	29.0
Dispersed-8	12.3	9.1	9.8	15.9	16.1	19.9	13.1	17.5	14.2
Dispersed-16	11.7	8.6	9.6	15.2	14.3	19.1	12.5	16.7	13.4

have also used the structural similarity (SSIM) metric to compare different halftones to the corresponding original image. The obtained results were very much in line with the results of mean squared error, and therefore we chose not to report them here.

4.1.2 Spatial-CIELAB ΔE

Section 1.3 provided a short summary of S-CIELAB filtering. By this approach, one can convert a color image into an image in CIELAB color space, taking into account the human visual system's three sensitivity functions. It is possible then to use this method to compare two different color images, for example, the reference/original and the printed halftone image. The result will then be a two-dimensional matrix, the same size as the images, where each pixel holds the ΔE color difference value between the corresponding pixels. The average of the pixel values of this matrix gives a ΔE value, indicating the color difference between the original and the printed halftone image. As the sensitivity functions of HVS have been included in the calculations, this ΔE value represents the perceived difference between the two images. In order to compare different halftoning methods, we have chosen eight test color images and compensated them for dot gain. Here, we assumed that all three color channels, i.e., C, M, and Y channels, have the same dot gain, represented by the printer model that has been employed in this article. However, we are aware that dot gain might be different for different colorant inks, but this difference will have the same impact on all halftones, and therefore, we believe this will not change the final conclusion. The compensated color channels are then halftoned by the halftoning methods, and thereafter the printer model is applied to each halftone channel, as has been done before for grayscale images. The simulation of the printed halftones together with the original image are used as the inputs to S-CIELAB metric, resulting in a ΔE value for each halftone and test image. Smaller ΔE value

indicates that the printed color halftone and the original image are more perceptually similar. The eight test color images are *Sari1*, *Sari2*, *Contrast*, *Ramp*, *Curve*, *Lena*, *Greens*, and *Peppers*. The color halftone images *Sari1* and *Sari2* are shown in Section 5. The color image called *Contrast* is the blue–yellow contrast image in Section 5. The image called *Ramp* is a color ramp only consisting of cyan and magenta, in which the C and M channels are identical. The test color image *Curve* is a color image containing curvy stripes of various width with two different colors, light blue and dark blue. The test color image *Lena* is the famous image *Lena* used in image processing. The test images *Greens* and *Peppers* are both demo images in MATLAB, being squared and resized to 512×512 pixels. Table III shows the ΔE values between these eight test color images and their corresponding halftones. The print resolution is $R = 600$ dpi and the print parameter and the viewing distance are $\rho = 1.1$ and $D = 15$ inches, respectively. The color difference is calculated using the CIE1976 color difference formula ΔE_{ab} , which is the Euclidean distance between two colors in the CIELAB color space. The color differences have also been calculated using ΔE_{94} . Although ΔE_{94} values were different than ΔE_{ab} , they did not affect the conclusion on the color halftone image quality. In Table III, besides the previous twelve halftones, we also added first-order FM using dot-off-dot printing, as explained in Section 2.2. In this dot-off-dot method, only C and M channels have been halftoned using dot off dot and the Y channel has been halftoned independent of the two others. The last column in this table, titled **AV.**, contains the average color difference value for each row. Note that, these halftones have also been examined disregarding the compensation and application of the printer model and all halftones (except for ordered dithering) gave very close ΔE_{ab} , and therefore we chose not to include those results. Note also that, the ΔE_{ab} difference between a completely black image and a

Table IV. Absolute error in gray level (%) ($\rho = 1.1$, $R = 600$ dpi).

	Mean	Max	Std
FM-1st	0.067	0.26	0.056
FM-2nd (0.5)	0.058	0.34	0.052
FM-2nd (0.6)	0.058	0.22	0.049
FM-2nd (0.7)	0.067	0.26	0.055
ED-FS	0.21	4.77	0.51
MED-FS	0.075	0.39	0.076
ED-J	0.25	10.4	0.74
MED-J	0.095	0.70	0.12
Clustered-8	0.70	2.57	0.54
Clustered-16	0.16	0.53	0.12
Dispersed-8	0.84	2.75	0.67
Dispersed-16	0.27	1.57	0.28

Table V. Absolute error in gray level (%) ($\rho = 1.2$, $R = 1200$ dpi).

	Mean	Max	Std
FM-1st	0.081	0.30	0.068
FM-2nd (0.5)	0.063	0.31	0.057
FM-2nd (0.6)	0.073	0.28	0.060
FM-2nd (0.7)	0.084	0.33	0.066
ED-FS	0.21	3.67	0.37
MED-FS	0.085	0.45	0.090
ED-J	0.28	12.5	0.84
MED-J	0.12	0.86	0.15
Clustered-8	0.79	2.95	0.65
Clustered-16	0.18	0.61	0.13
Dispersed-8	1.02	3.21	0.78
Dispersed-16	0.32	1.91	0.31

completely white image is 100. We also studied the S-CIELAB ΔE for $\rho = 1.2$ and $R = 1200$, and the obtained color differences for different halftoning methods were very close to each other. Therefore, we decided not to include them in this article.

To conclude the similarity measures, among our proposed halftones, second-order FM ($\sigma_2 = 0.5$, i.e., small clustered dot size) is the best halftone for monochromatic images printed at both $R = 600$ and $R = 1200$. This halftone is best for all types of monochromatic images, except for the test image *Curve*. When it comes to the color images, this halftone is only slightly worse than first-order FM, which is the best among our proposed halftones for color images. As for the error diffusion methods, we observed much more variations in quality. While ED (J) was best for monochromatic constant images printed at $R = 600$, MED (FS) was best for the same images printed at $R = 1200$. For regular test images, ED (FS) was best when printed at $R = 600$, but MED (FS) was best for the same images printed at $R = 1200$. For color images, ED (FS) is slightly better than MED (FS), while both of them are better than ED (J) and MED (J). The final conclusion is that our proposed halftones are more stable in terms of quality for different types of images, and second-order FM ($\sigma_2 = 0.5$) is the best halftoning between all of the studied halftones in general. The reason is that, this halftone is either better or slightly worse than the best error diffusion method considering all different types of monochromatic and color images.

4.2 Gray Level Representation

It was discussed in Section 4.1.1 that the number of gray levels was not directly reflected in the calculated mean squared error. Here, we want to examine how accurate and how fast each halftone reacts to the change of gray level. Therefore, we created a grayscale ramp in which the gray levels were varied from 0 to 255, where each gray level being 32 pixels wide. Thirty-two pixels mean 1.35 mm when printed at $R = 600$ dpi and 0.68 mm at $R = 1200$ dpi, which corresponds to $f = 4.91$ and $f = 9.82$ cycles/degree for the

viewing distance of $D = 15$ inches, respectively (see Eq. (3)). This ramp was then compensated for dot gain using the dot gain curve specified to each halftoning method and print resolution. Afterward, the compensated grayscale ramp was halftoned by the corresponding halftoning method. The printer model was then applied to the halftone ramps, using $\rho = 1.1$ for $R = 600$ and $\rho = 1.2$ for $R = 1200$. The absolute value of the difference in gray level between the input ramp and the output ramp over each 32 pixels wide area was then calculated, giving a gray level error array of size 1×256 for each halftone. There are different ways to illustrate the results, such as plotting these gray level error arrays. However, we noticed that, as the error could be quite different for different halftones, plotting them in the same graph would not give useful information. Therefore, we found it more relevant to report the results by showing the mean, max, and standard deviation of each gray level error array in two tables. Tables IV and V show these statistical data for ($\rho = 1.1$ for $R = 600$) and ($\rho = 1.2$ for $R = 1200$), respectively. First of all, there is no significant difference between the two tables and both show the same trend. As seen in both tables, our four proposed halftones behave very similarly and there is no significant difference between them. Among the error diffusion methods, ED (FS) and ED (J) result in quite large errors, especially large maximum errors. The maximum error for ED (FS) and ED (J) occurs around the reference coverage that is compensated to around 50%. Let us look at one of these maximum values. As seen in Table V, the maximum error for ED (J) is 12.5% and this occurs at the gray level of 242, corresponding to $242/255 = 94.9\%$ reference coverage. In the dot gain curve for ED (J) in Fig. 8(b), we see that the dot gain changes very rapidly at 50% coverage, and the dot gain at 50% is slightly over 45%, meaning that the coverage after print will be slightly over 95%. Consequently, the reference coverage of $242/255 = 94.9\%$ is compensated to around 49%. The original ED halftone patches at 49% and 50% behave very differently in terms of dot gain, and thus small changes in reference coverage around 50% might

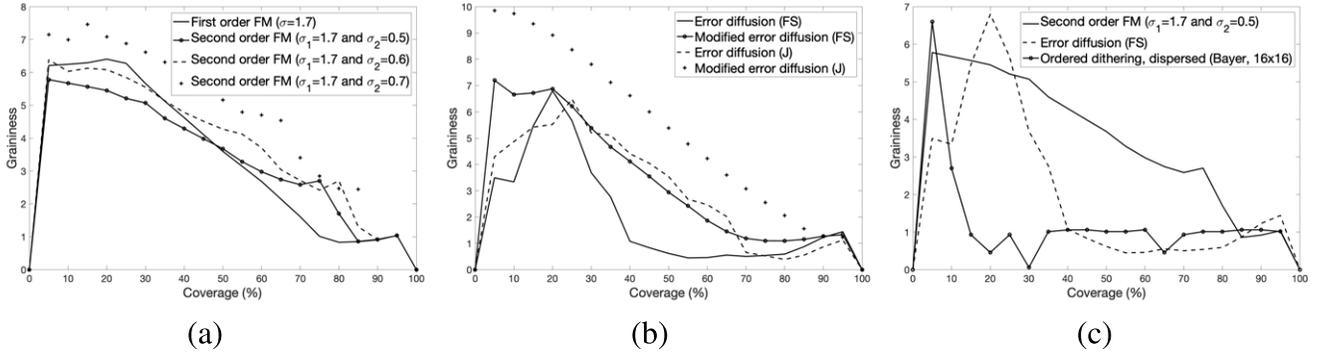


Figure 12. Graininess for constant patches between 5% and 95% in a step of 5% for $\rho = 1.1$, $R = 600$ dpi, and $D = 15$ inches. (a) Proposed first- and second-order FM. (b) Error diffusion and modified error diffusion with FS and J filters. (c) Comparing the best halftone in (a) and (b) and dispersed dot ordered dithering (Bayer 16×16).

cause large errors. The error for MED methods are much smaller but still slightly larger than our proposed methods. Not surprisingly, ordered dithering using 16×16 threshold matrices result in smaller error than using 8×8 threshold matrices, because they represent more gray levels. The reason for clustered dot order dithering being better in representing gray levels is that their dot gain is much smaller compared to dispersed dot ordered dithering. We have also carried out the same calculations on a ramp where each gray level was 16 pixels wide and noticed that the error is a bit larger but still very low for our proposed and MED methods.

In order to investigate how the gray levels are preserved in regular test images, we also examined the halftoning methods on the eight test images presented in Section 4.1.1. These test images were first compensated for dot gain and then halftoned. The printer model was thereafter applied to the halftone images. These images were afterward divided into 256 sub-images of size 32×32 pixels. The absolute value of the difference between the average value of each 32×32 pixels sub-image and its corresponding sub-image in the original test image was calculated. We studied the mean and max value of these differences for all eight test images and all halftoning methods. The average error in gray level of all 32×32 pixels sub-images in all eight test images was around 0.4% for our proposed methods. The same average error was around 0.4% even for ED (FS), while it was around 0.3% for MED (FS) which was the best among all halftoning methods. We can now conclude that our proposed methods preserve the gray levels quite well and much better than ordered dithering methods. In comparison to the error diffusion methods, the point-by-point IMCDP methods perform similar or slightly better than ED (FS) and MED (FS), which are both better than ED (J) and MED (J).

4.3 Graininess

Another important quality attribute that halftoning methods have an impact on is graininess. In this article, the employed graininess metrics for monochromatic and color halftones are based on measuring the intensity variation on an overall uniform area caused by halftoning.

4.3.1 Graininess: Monochromatic

In this subsection, different monochromatic halftones are compared to each other in terms of graininess by studying how smooth the perceived printed halftone patches are. For this purpose, patches from 5% to 95% coverage in a step of 5% were created. The patches were then compensated for dot gain and then halftoned. The printer and the eye model were thereafter applied to the halftones. This means that the graininess of the image $b_{p,e}$ in Fig. 9 is studied. Since the original image has been a constant patch, less variation in the pixel values means smoother (less grainy) printed halftones. Therefore, the standard deviation of the pixel values of $b_{p,e}$ is used as the graininess measure. The larger this standard deviation, the more grainy the halftone patch is. The graininess value was computed for each halftone using ($\rho = 1.1$ and $R = 600$), and ($\rho = 1.2$ and $R = 1200$), and the viewing distance of $D = 15$ inches. The calculated graininess values using $\rho = 1.2$ and $R = 1200$ were, as expected, smaller than using $\rho = 1.1$ and $R = 600$ for all halftones, but they followed the same pattern. Therefore, in Figure 12, we only show the graininess values using $\rho = 1.1$, $R = 600$ dpi, and $D = 15$ inches. Note that, the images were scaled between 0 and 255, meaning that the largest graininess value for an image would be $255/2 = 127.5$. As can be seen in (a), there is no significant difference in graininess values between our proposed methods, but the second-order halftones ($\sigma_2 = 0.5$) are slightly better than the others. Among the error diffusion methods, ED (FS) results in lower graininess, and there is quite big difference between ED (FS) and MED (J), which is the worst one among them. Among ordered dithering methods, Bayer (16×16) was the best. In (c), the best halftone from each class are compared to each other. Bayer (16×16) results in quite low graininess, which was expected. The reason is that, the dispersed dots are placed in a grid with equal distance and as far apart as possible, which will result in smoother halftones. The ED (FS), which is the second best method, behaves differently for different patches. The graininess is higher for lighter ED (FS) halftones, but suddenly drops to small values at around 40%. Note that, 40% coverage is compensated to around 20% at which ED (FS) behaves more like a dispersed dot order

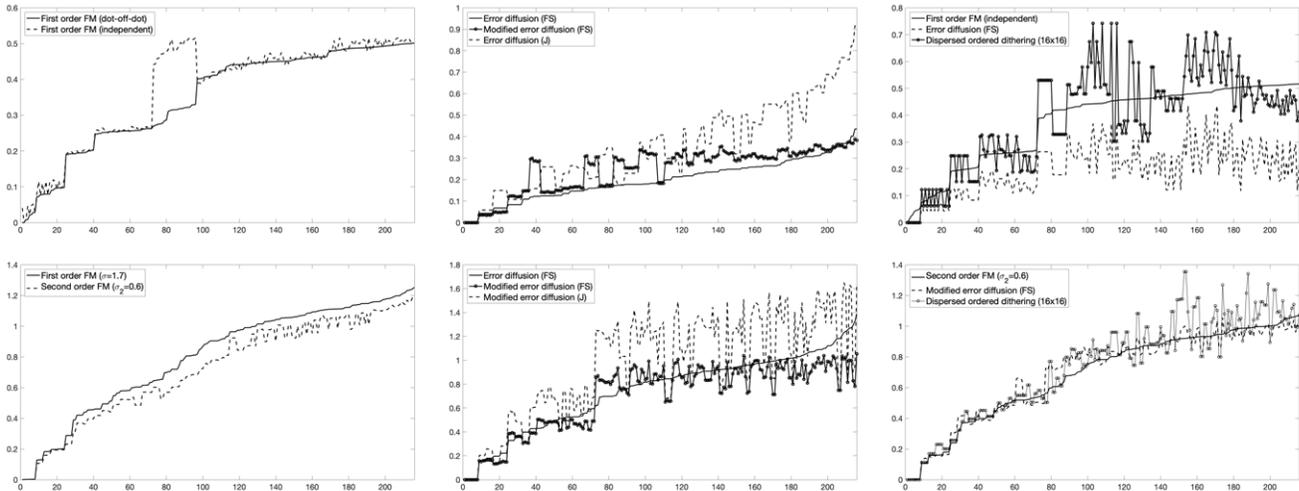


Figure 13. Graininess for 216 CMY color patches. The coverage for C, M, and Y varies from 0% to 100% in a step of 20%. (Top row) No dot gain compensation and printer model has been applied. (Bottom row) Dot gain compensation and printer model have been applied, using $\rho = 1.1$ ($R = 600$ dpi and $D = 15$ inches).

dithering than a “stochastic” first-order FM. Note also that, ED (FS) was quite bad and much worse than MED (FS) in representing the gray levels, as shown in Tables IV and V, while it is better than MED (FS) in terms of graininess. On the other hand, second-order FM ($\sigma_2 = 0.5$) that has most of the times created the best halftone among our proposed halftones, is still quite good in reproducing smooth printed halftones.

4.3.2 Graininess: Color

In order to study the graininess of color halftones, we apply the S-CIELAB filtering. The inputs are the color halftones converted to the CIEXYZ color space and the cycles per degree and the output is an image in the CIELAB color space, see Section 1.3. The graininess can then be defined as the sum of the standard deviation of the CIEL*, a^* , and b^* channels of the output image [21]. The lower this graininess value, the less grainy the color halftone. For this purpose, we have created color halftone patches with coverage for C, M, and Y channels, varying from 0% to 100% in step of 20%, i.e., six coverage for each channel. It will therefore be $6^3 = 216$ combinations and thereby 216 test color patches. We have studied the graininess once without involving dot gain compensation and printer model and once with them being taken into account. In the first study, the test color patches have been halftoned as they were and then converted into the CIEXYZ color space. Besides the halftoning methods that have been analyzed before, the graininess for our proposed halftones utilizing dot-off-dot printing is also studied. Note that, since different halftones behave differently, these resulting halftones might end up having different colors, although the input to all of them is exactly the same color patch. Figure 13 (top row) shows the results using $R = 600$ and $D = 15$ (i.e., 157.1 cycles/degree). In the top-left graph, only the graininess values for first-order FM independent and dot off dot are illustrated, because

second-order FM halftones (independent) resulted in very similar graininess as first-order FM (independent). Note that, in order to be able to see the difference between the graphs more clearly, the graininess of the first-order FM (dot off dot) has been sorted in the ascending order, and the other graph has been rearranged accordingly. As expected, the dot-off-dot printing results in less grainy halftones. In this study, dot off dot has been used only between C and M channels. The biggest differences between independent and dot off dot occur when the sum of cyan and magenta coverage is 100%. In the top-middle graph, the graininess for ED (FS), MED (FS), and ED (J) are shown. The graininess for MED (J) was much larger. The graininess of ED (FS) has been sorted in the ascending order, and the other graphs have been rearranged accordingly. Like before, the diversity among the error diffusion methods are much larger than among our proposed methods. In the top-right figure, we compare first-order FM (independent) with ED (FS), which was the best among the error diffusion methods, together with Bayer (16×16). The graininess of first-order FM (independent) has been sorted in the ascending order, and the other graphs have been rearranged accordingly. As seen in this graph, ED (FS) color halftones are less grainy and dispersed dot order dithering is the grainiest. It would, however, be more realistic to compare different halftones if we involve the dot gain compensation and the printer model. The same test color patches were, therefore, first compensated for dot gain. Although different colorant inks might have different dot gain, CMY color channels are still compensated using the same ρ value. The compensated color patches were then halftoned and the printer model was applied to its CMY channels, before being converted into the CIEXYZ color space. Fig. 13 (bottom row) shows the results using $R = 600$, $\rho = 1.1$, and $D = 15$. In general, the graininess of different color halftones are closer to each other than in the top row. The graininess of our proposed halftones are very

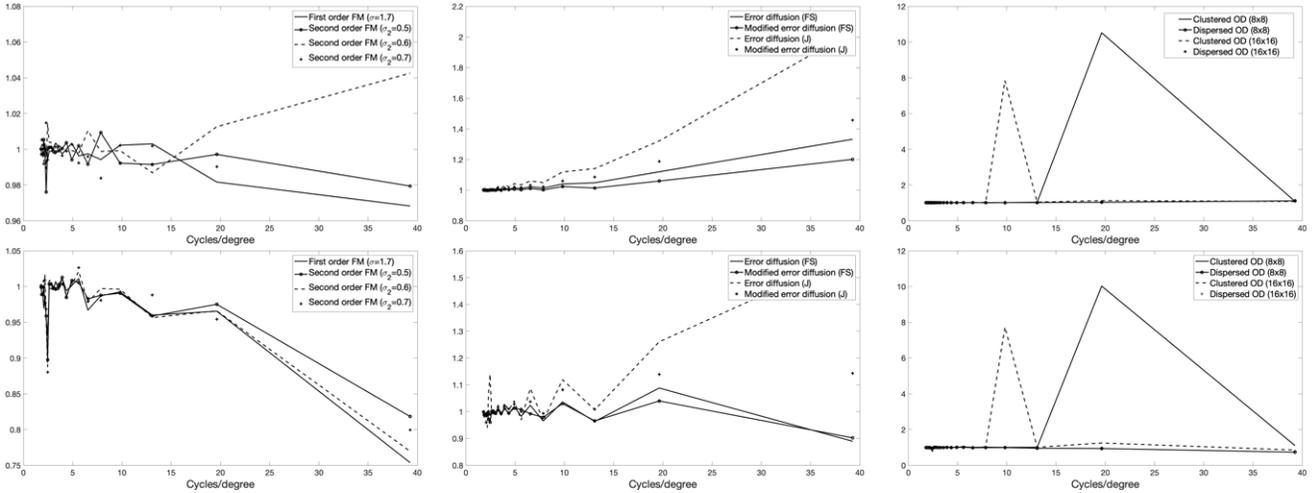


Figure 14. Frequency gain as defined in Eq. (10). (Top row) No dot gain compensation and printer model has been applied. (Bottom row) Dot gain compensation and printer model have been applied using $\rho = 1.1$.

close to each other, where second-order FM ($\sigma_2 = 0.5$ and $\sigma_2 = 0.6$) are slightly better than others and there were no noticeable difference between independent and dot-off-dot halftoning in this case. Notice that, it is not completely correct to compensate the CMY channels the same way for both independent and dot off dot, which has been done here. The same amount of cyan and magenta being printed independently and dot off dot might result in different colors [7, 15]. Among the error diffusion methods, MED (FS) was best, and the graininess for ED (J) was much higher than the others, and therefore not plotted in the bottom-middle figure. As seen in the bottom-right figure, second-order FM ($\sigma_2 = 0.6$) is better than the best error diffusion method and Bayer (16×16).

To conclude, our proposed halftones behave quite similarly, with dot off dot being better in terms of graininess. Among the error diffusion methods, ED (J) and MED (J) are quite poor in terms of graininess, and they are also worse than all of our proposed halftones. For monochromatic halftones, Bayer (16×16) is the best halftone. For color halftones, our proposed halftones and ED (FS) and MED (FS) behave quite similarly, with ED (FS) being slightly better when no dot gain compensation is performed, and second-order FM ($\sigma_2 = 0.6$) is slightly better when dot gain compensation is performed. As for the independent halftoning, second-order FM with small or medium clustered dot size are slightly better than first-order FM.

4.4 Sharpness

The last quality attribute employed in this article is to examine how different halftones affect the sharpness of an image. The most straightforward approach to obtain a sharpness metric is to calculate a sharpness value in frequency domain. In Ref. [22], an approach is proposed that returns a sharpness/blurriness value assessing the total sharpness/blurriness of an image calculated in the Fourier domain. In this article, we are more interested in studying the

sharpness at different frequencies (cycles/degree), therefore proposing another approach to evaluate the sharpness of monochromatic and color halftones in the Fourier domain.

4.4.1 Sharpness: Monochromatic

In order to obtain a measure evaluating the sharpness of monochromatic halftones, we propose the following approach. Consider an $M \times N$ pixels image with, for example, vertical stripes of two different gray tones, each being $P/2$ pixels wide, meaning a period of P . Since all rows in such an image are identical, there will be three dominant peaks in the Fourier spectrum of the image appearing on the horizontal axis. One of these three peaks corresponds to the dc-term, being located at the center of the spectrum, assuming that the zero-frequency has been shifted to the center. The other two dominant peaks correspond to the frequency of the stripes, and are located on the horizontal axis equally far from the center of the spectrum. The smallest possible period in a digital image is $P = 2$ pixels, implying the highest possible frequency of $f = 1/2$ cycles/pixel. This means that the two dominant peaks will appear at the furthest point on each side of the center, i.e., each being $N/2$ pixels away from the center on the horizontal axis. Note that, in a digital image of size $M \times N$, the position of the center is $(\text{floor}(M/2), \text{floor}(N/2))$, assuming the top-left pixel has the position of $(0, 0)$. The function $\text{floor}()$ returns the largest integer less than or equal to its argument. For simplicity, from now on, we refer to the position of the center discarding $\text{floor}()$. As an example, if the stripes are 2 pixels wide each, the period will be $P = 4$, corresponding to $f = 1/4$ cycles/pixel, implying that the two dominant peaks will appear exactly half-way between the center and ends on the horizontal axis of the spectrum [23]. This means that the dc-term will appear at the center, i.e., $(M/2, N/2)$, and the other two peaks at $(M/2, N/2 + N/4)$ and $(M/2, N/2 - N/4)$. Hence, for a general period of P , the two dominant peaks will appear at $(M/2, N/2 + N/P)$

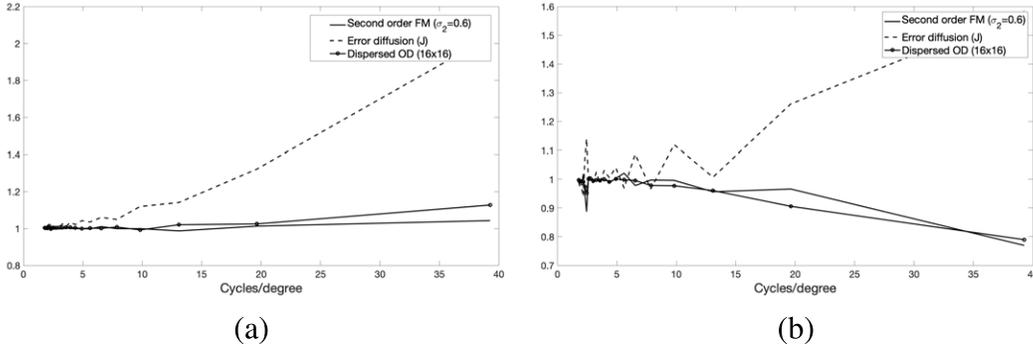


Figure 15. Comparing the sharpest halftones from each halftoning method class. (a) No dot gain compensation and printer model has been applied. (b) Dot gain compensation and printer model have been applied using $\rho = 1.1$.

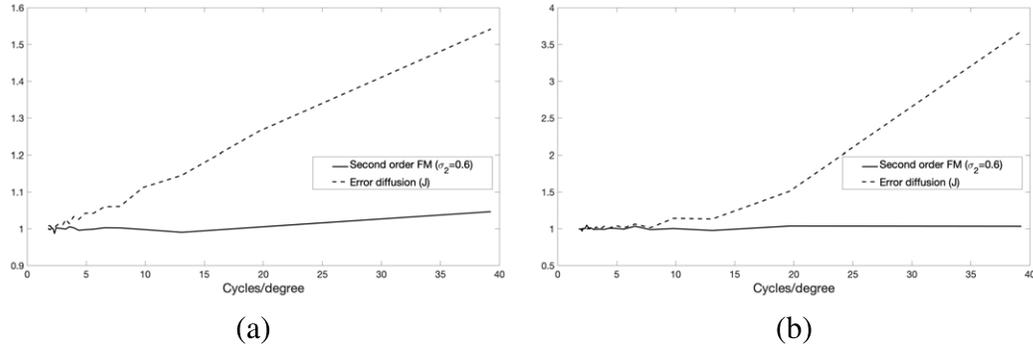


Figure 16. Comparing ED (J) and second-order FM ($\sigma_2 = 0.6$) in terms of sharpness based on the contrast. (a) On the upper half of the *Contrast* test image where the contrast is higher. (b) On the lower half of the *Contrast* test image where the contrast is lower.

and $(M/2, N/2 - N/P)$. Notice also that a period of P corresponds to the visual angle of $\tau = (P/RD)(180/\pi)$ degrees or $f = \pi RD/180P$ cycles/degree, where R and D are the print resolution (dpi) and the viewing distance (inches), respectively. For this purpose, we have created an image with vertical stripes of two gray tones. In this image, the period of the stripes (one darker and one lighter) is varied from 88 pixels down to 4 in a step of 4, representing 22 different periods and thereby 22 frequencies. We call this image *Contrast*. A part of this image is shown in Section 5. Note that, in the *Contrast* image, the contrast varies slowly along the vertical axis, being highest at the top and lowest at the bottom. However, this small change of intensity along each column will not affect the Fourier spectrum on the horizontal axis. For a print resolution of $R = 600$ and the viewing distance of $D = 15$, the corresponding frequency will change from 1.785 to 39.27 cycles/degree, which is high enough as the eye is not sensitive to frequency higher than 40, see Section 1.2. We define a so-called frequency gain measure as

$$F_{\text{gain}} = \frac{|F_{\text{out}}(M/2, N/2 + N/P)|}{|F_{\text{in}}(M/2, N/2 + N/P)|}, \quad (10)$$

where $|F_{\text{out}}|$ and $|F_{\text{in}}|$ denote the Fourier spectra of the input and the output image, respectively. Note that, if N/P is not an integer, the value of the spectrum at that specific position could be either taken from its nearest neighbor or being calculated by an interpolation between two or more

of its nearest neighbors. In this article, we use the nearest neighbor interpolation whenever N/P is not an integer. The frequency gain is then calculated for all 22 different frequencies. A frequency gain larger than 1 at any frequency, indicates that the output image is sharper than the original at that specific frequency. In order to study how different halftones behave in terms of sharpness, we halftoned the test *Contrast* image by different halftoning methods, first without involving the dot gain compensation and printer model. Therefore, in this case, F_{out} in Eq. (10) is the Fourier transform of each halftone. Figure 14 (top row) shows the frequency gain for all studied halftones without involving the dot gain compensation and printer model. As the eye is most sensitive at 8 cycles/degree and not so sensitive to frequencies higher than 20, see Fig. 2, we should mostly focus on frequencies around 8 and lower than 20 cycles/degree. We can see in the top-left graph that our proposed halftones behave very similarly and do not sharpen the original image, as the gain is very close to 1. The second-order FM halftone ($\sigma_2 = 0.6$) is slightly sharper than the others. In the top-middle graph, we can see that ED halftones behave quite differently, but all of them have a sharpening effect on the original image, with ED (J) resulting in the sharpest halftone. This was already shown in other studies that error diffusion methods have a sharpening effect and the larger the error filter, the sharper the halftone [24]. In the top-right graph, the frequency gain for ordered dithering methods

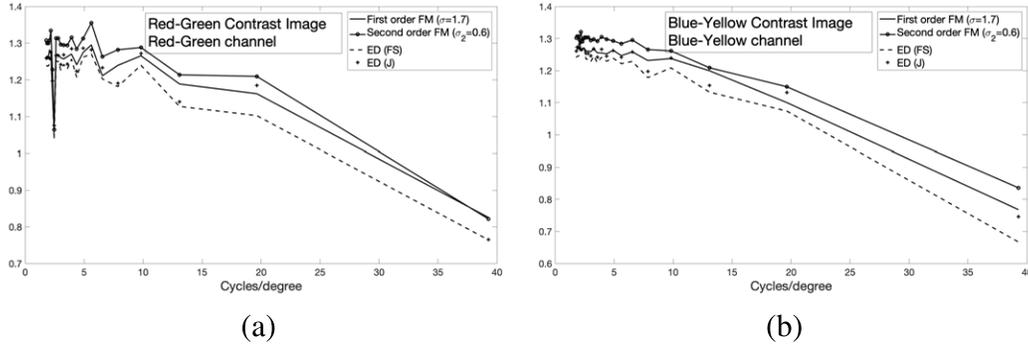


Figure 17. Frequency gain for $\rho = 1.1$, $R = 600$, and $D = 15$. (a) Red–Green contrast image and Red–Green channel. (b) Blue–Yellow contrast image and Blue–Yellow channel.

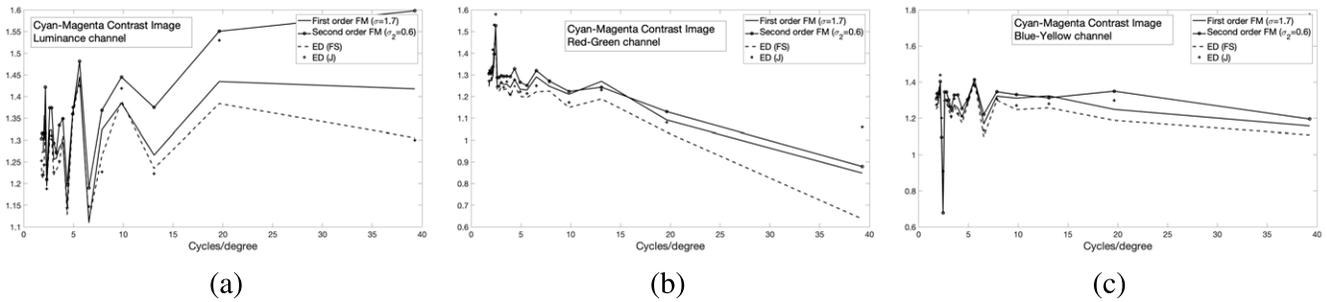


Figure 18. Frequency gain for $\rho = 1.1$, $R = 600$, and $D = 15$. The test image is the *Cyan–Magenta Contrast* image. (a) Luminance channel. (b) Red–Green channel. (c) Blue–Yellow channel.

are shown. Bayer (8×8) and Bayer (16×16) behave very similarly, and have a gain very close to 1. The clustered dot ordered dithering methods result in large peaks, coming from the frequency created by their threshold matrices. The threshold matrices of size 8×8 and 16×16 , correspond to $f = \pi RD / (180 \cdot 8) = 19.63$ and $f = \pi RD / (180 \cdot 16) = 9.82$ cycles/degree, respectively, which are actually the frequencies where the peaks occur. Figure 15(a) shows the frequency gain of the sharpest halftone among our proposed methods, i.e., second-order FM ($\sigma_2 = 0.6$), together with that of the sharpest halftone among the error diffusion methods, i.e., ED (J), and Bayer (16×16). As can be seen in this graph, ED (J) generates much sharper halftones and our proposed method and dispersed dot ordered dithering generates equally sharp halftones. The reason is that, in the iterative method called IMCDP, the goal was to obtain a halftone similar to the original image as much as possible, and since generating the threshold matrices follow the same strategy, they will result in halftones that are as sharp as the original. Another point worth noting here is that, adding the eye filter will not change the results, because it will amplify/attenuate the original image and the halftones equally, and thus the ratio will remain the same. This was also verified by our experiment. We have already discussed in previous sections that the effect of dot gain has to be taken into account. Therefore, in our next study, the test *Contrast* image is first compensated for dot gain and then halftoned and finally, the printer model is applied. Hence, in this case, F_{out} in Eq. (10)

is the Fourier transform of each halftone after printer model has been applied, i.e., b_p in Fig. 9. Fig. 14 (bottom row) shows the frequency gain for all studied halftones. The dot gain compensation and the printer model have been applied using $\rho = 1.1$, as always have been used in this article for $R = 600$. As can be seen in these graphs, all halftones are less sharp than before and the frequency gain for the halftones are much closer to each other now. Fig. 15(b) shows the frequency gain of the same halftones as in Fig. 15(a), but this time with dot gain taken into account. As can be seen, ED (J) is not much sharper than our proposed second-order halftone, but it still keeps the halftone at least as sharp as the original image, while second-order FM and dispersed dot ordered dithering are slightly attenuating the original image. Another study we have conducted is to see how much different halftones sharpen the original image based on the contrast between the stripes. As seen in the *Contrast* test image, the contrast is varied from maximum at top to minimum at the bottom. If one wants to know the sharpening effect of the halftones for a specific contrast between the stripes, one can calculate the frequency gain, not on the whole image, but in the area corresponding to that specific contrast. Here, we just illustrate the difference in the frequency gain between the upper half and the lower half of the *Contrast* test image. Figure 16(a) and (b) show the frequency gains for second-order FM ($\sigma_2 = 0.6$) and ED (J), in the upper half and the lower half of the *Contrast* test image, respectively. As can be seen, in general, the halftones behave quite similar

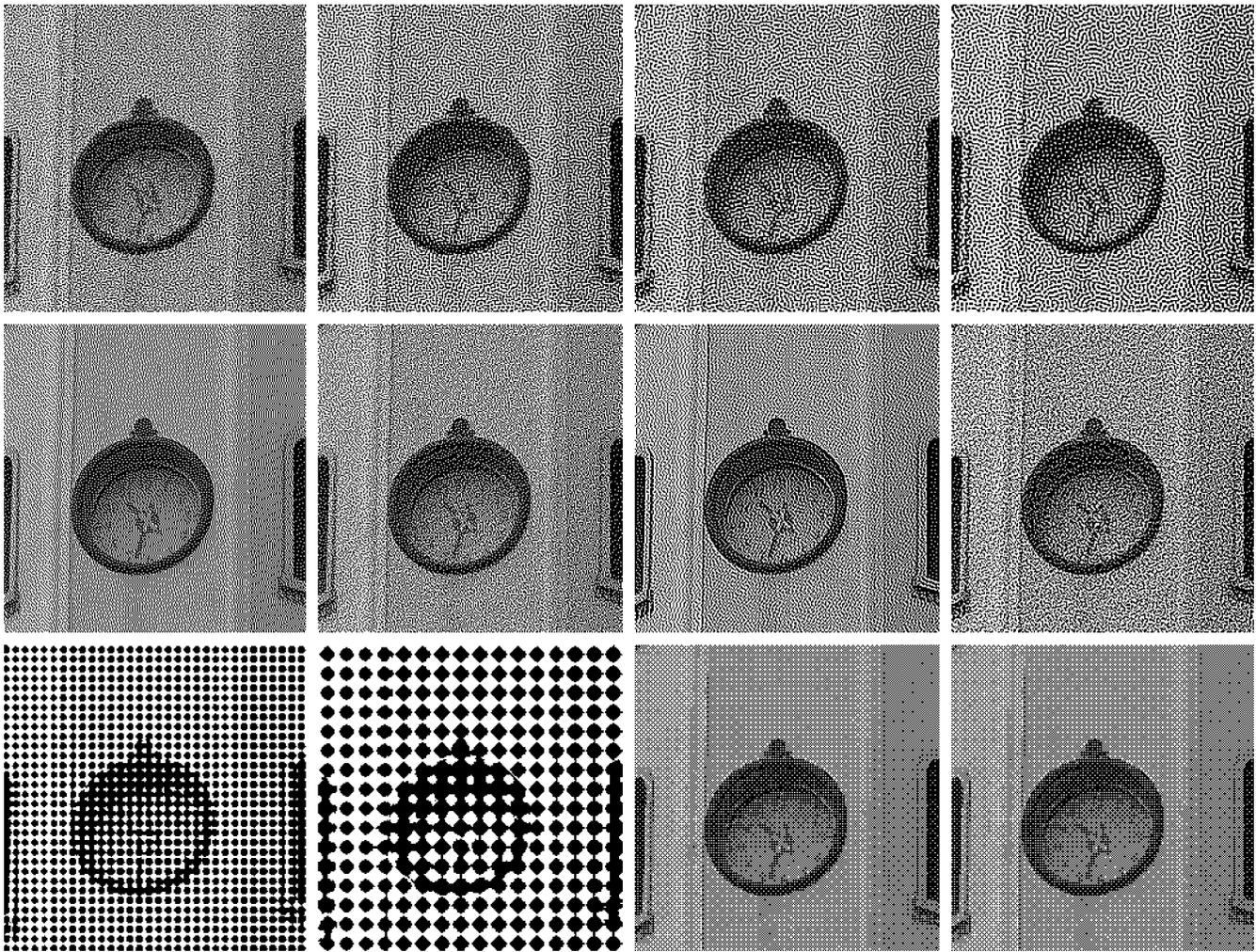


Figure 19. The test image is *Sari1* used in Section 4.1.1. These are binary halftones and no dot gain compensation or printer model has been applied. From top to bottom and left to right: (First row) First-order FM, second-order FM ($\sigma_2 = 0.5$), second-order FM ($\sigma_2 = 0.6$), and second-order FM ($\sigma_2 = 0.7$). (Second row) ED (FS), MED (FS), ED (J), and MED (J). (Third row) Clustered (8×8), Clustered (16×16), Bayer (8×8), and Bayer (16×16).

when the contrast is low and very much different when the contrast is high. It is worth pointing out that, in order to make sure that our test image is a good test image in this context, we also created 22 test images having vertical stripes, each representing periods of $P = 80$ down to 4, in a step of 4. The contrast in each test image was varied as explained above. These twenty-two test images were then halftoned and the frequency gains at these 22 frequencies were calculated, separately. The results showed a similar trend between the halftones, and although the frequency gains were not exactly the same as shown in Fig. 14, they were quite close, and more importantly, the difference between the halftones in terms of sharpness were almost identical. Therefore, we can conclude that our proposed *Contrast* image is an appropriate test image for studying the sharpness of halftones.

4.4.2 Sharpness: Color

As discussed in Section 1.3, the human visual system is represented by three sensitivity functions in the opponent

color space, one luminance and two chromatic channels, i.e., red–green and blue–yellow. The function for the luminance channel is a band-pass filter, and the other two are low-pass filters with different cut-off frequencies. In order to study the sharpness of color halftones, we calculate the frequency gain, as defined in Section 4.4.1, in these three channels. For the frequency gain in the red–green channel, we use a similar contrast image as in Section 4.4.1, but instead of having different gray tones, different shades of red and green are used in different vertical bars. This test image will be referred to as the *Red–Green Contrast* image. A similar test image using blue and yellow is created to study the sharpness in the blue–yellow channel. We refer to this test image as the *Blue–Yellow Contrast* image. The test images are compensated for dot gain by compensating their CMY channels for dot gain. The compensated images are then halftoned by the halftoning methods before the printer model being applied. For the red–green frequency gain calculation, F_{in} in Eq. (10) is the Fourier transform of the

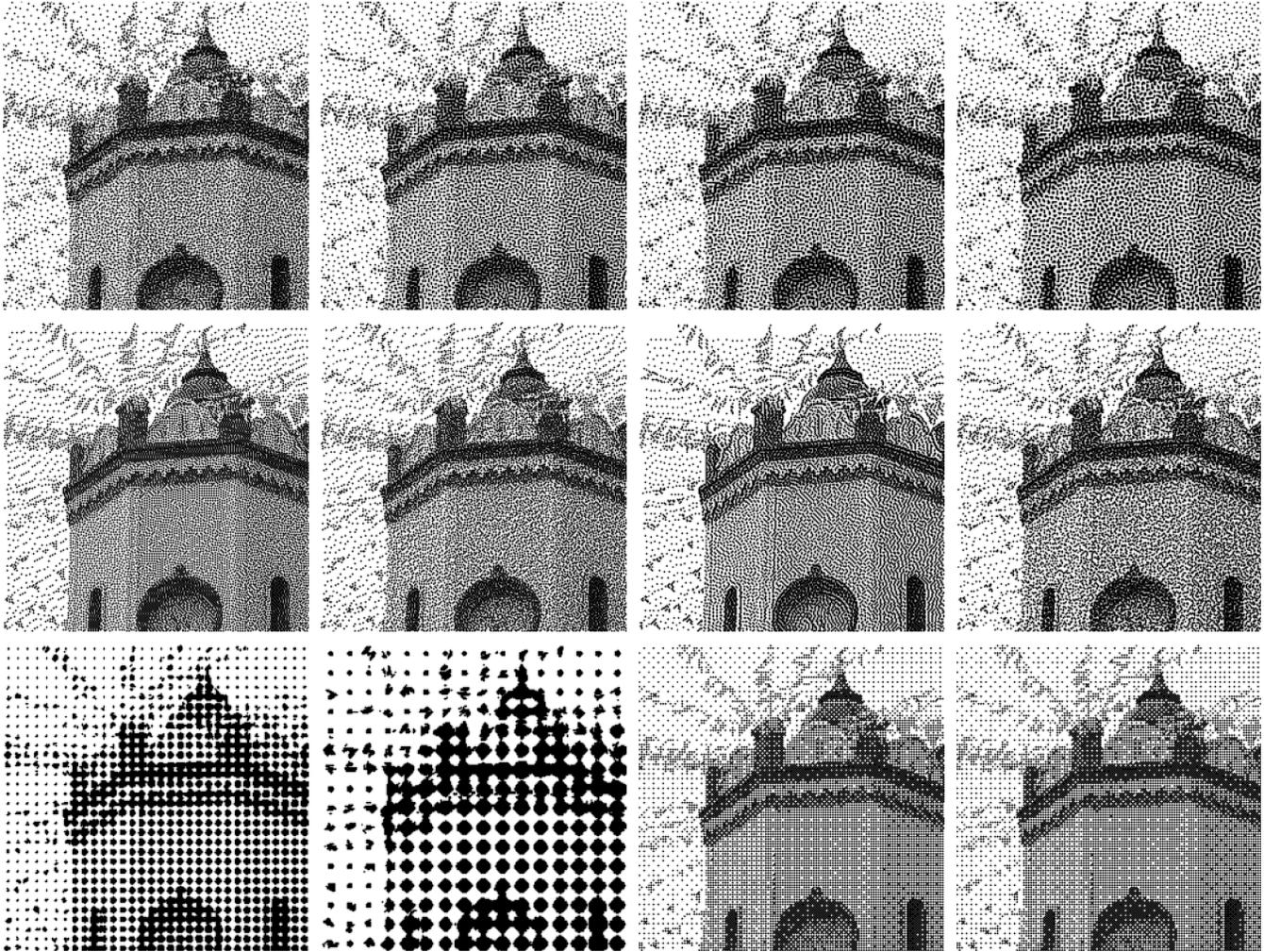


Figure 20. The test image is *Sari2* used in Section 4.1.1. These are not binary halftones. In order to obtain them, the test image has been first compensated for dot gain and then halftoned before the printer model being applied, using $\rho = 1.1$. These images are supposed to represent a magnified version of the printout. From top to bottom and left to right: (First row) First-order FM, second-order FM ($\sigma_2 = 0.5$), second-order FM ($\sigma_2 = 0.6$), and second-order FM ($\sigma_2 = 0.7$). (Second row) ED (FS), MED (FS), ED (J), and MED (J). (Third row) Clustered (8×8), Clustered (16×16), Bayer (8×8), and Bayer (16×16).

red–green channel of the *Red–Green Contrast* image, and F_{out} is the Fourier transform of the red–green channel of the corresponding printed halftones. The red–green channel of an image is obtained by O_2 in Eq. (4). For the blue–yellow channel, F_{in} is the Fourier transform of the blue–yellow channel of the *Blue–Yellow Contrast* image and F_{out} is the Fourier transform of the blue–yellow channel of the printed halftones. The results of halftoning a part of the *Blue–Yellow Contrast* image are illustrated in Section 5. The blue–yellow channel of an image is obtained by O_3 in Eq. (4). Figure 17(a) and (b) show the frequency gain for ED (FS), ED (J), first-order FM, and second-order FM ($\sigma_2 = 0.6$) in red–green and blue–yellow channels, respectively. At a first glance, one can see that all methods create halftones sharper than the original test image. As discussed in Section 1.3, the eye is most sensitive in low frequencies in both color channels, but with different cut-off frequencies. Therefore, the main focus should be on the frequencies lower than 20 cycles/degree in the red–green channel and lower than

10 cycles/degree in the blue–yellow channel. As seen in both graphs, opposite to grayscale images, our proposed methods result in sharper halftones in both chromatic channels. The MED methods, which are not shown in this figure, opposite to the monochromatic images, are slightly sharper than their corresponding ED methods. We had also a look at the frequency gains in the luminance channel of these two test images and the frequency gains of all four methods were very close to each other, with second-order FM ($\sigma_2 = 0.6$) being slightly sharper. Therefore, we do not show the frequency gains in the luminance channel for these two test images. Instead, we created another contrast image and studied the frequency gains in all three channels. The test image is very similar to previous contrast images, but in different bars different shades of cyan and magenta are used. We refer to this test image as the *Cyan–Magenta Contrast* image. This test image is compensated for dot gain as before. The compensated image is then halftoned by the halftoning methods prior to the printer model application.

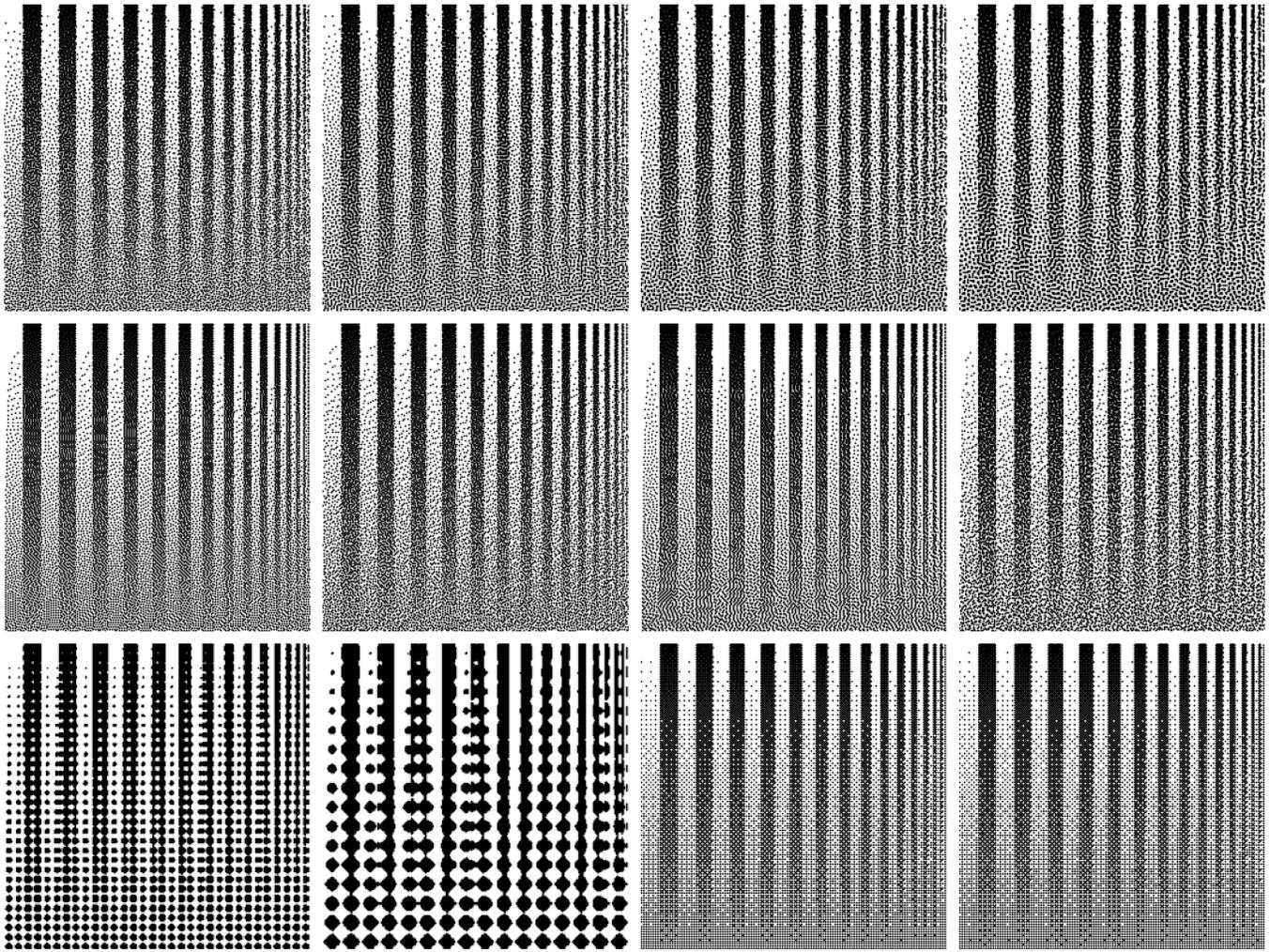


Figure 21. The test image is *Contrast* used in Section 4.1.1. These are not binary halftones. In order to obtain them, the test image has been first compensated for dot gain and then halftoned before the printer model being applied, using $\rho = 1.1$. These images are supposed to represent a magnified version of the printout. From top to bottom and left to right: (First row) First-order FM, second-order FM ($\sigma_2 = 0.5$), second-order FM ($\sigma_2 = 0.6$), and second-order FM ($\sigma_2 = 0.7$). (Second row) ED (FS), MED (FS), ED (J), and MED (J). (Third row) Clustered (8×8), Clustered (16×16), Bayer (8×8), and Bayer (16×16).

The frequency gain in all three channels, i.e., luminance, red–green and blue–yellow, for the *Cyan–Magenta Contrast* image using first-order FM, second-order FM ($\sigma_2 = 0.6$), ED (FS) and ED (J) are shown in Figure 18. This graph verifies our previous conclusion that the second-order FM is sharpest and ED (FS) is the least sharp in all three channels. One explanation to why the conclusion drawn for the sharpness of color halftones is almost the opposite to the conclusion drawn for monochromatic images could be that the ED techniques being applied to different channels will create sort of a dot-on-dot structure. Consider three identical channels being halftoned by the same ED method; the dots on different channels will be placed on top of each other. In the MED methods, this dot-on-dot printing will not occur at all positions and in our proposed methods the channels are halftoned completely independent of each other. This could be the reason for our proposed and MED methods generating

sharper color halftones than ED methods. We also studied the dot-off-dot halftones, and they behave very similarly to their counterpart using independent halftoning in terms of sharpness.

We can conclude that the error diffusion methods, as expected, generate sharper monochromatic halftones than our proposed methods. Non-modified ED halftones are generally sharper than their correspondent MED methods. It was also shown that ED (J) generated monochromatic halftones are sharper than ED (FS) halftones. The difference between our proposed monochromatic halftones and ED and MED halftones in terms of sharpness, becomes more considerable as the contrast increases. For color halftones, on the other hand, our proposed color halftones were generally sharper than ED generated halftones in all three opponent color channels.

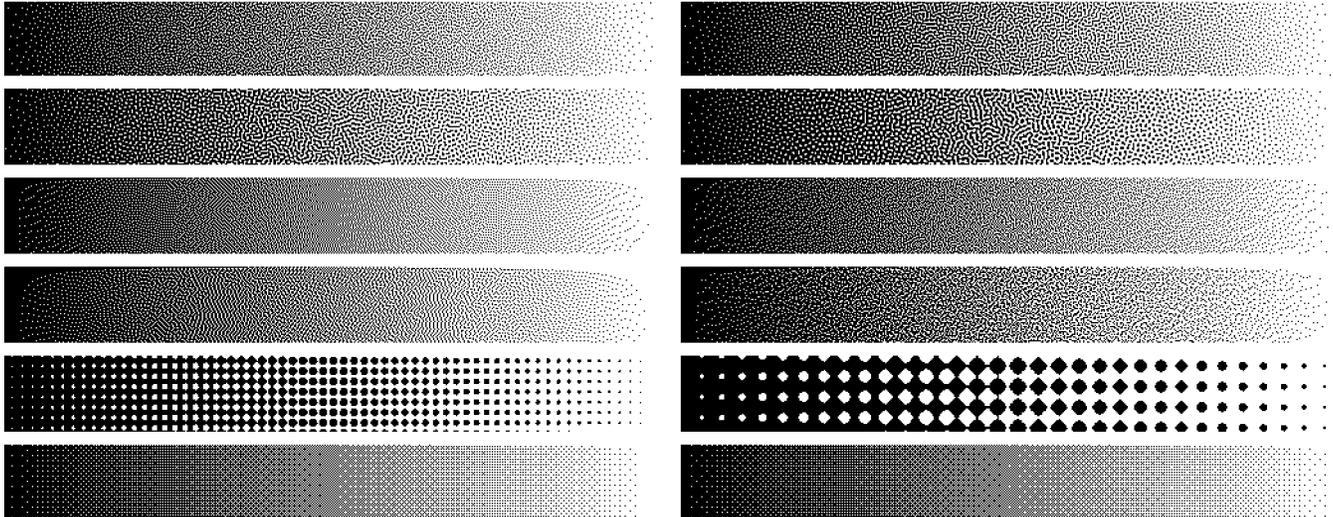


Figure 22. The test image is *Ramp* used in Section 4.1.1. These are binary halftones and no dot gain compensation or printer model has been applied. From top to bottom and left to right: (First row) First-order FM and second-order FM ($\sigma_2 = 0.5$). (Second row) Second-order FM ($\sigma_2 = 0.6$) and second-order FM ($\sigma_2 = 0.7$). (Third row) ED (FS) and MED (FS). (Fourth row) ED (J) and MED (J). (Fifth row) Clustered (8×8) and Clustered (16×16). (Sixth row) Bayer (8×8) and Bayer (16×16).

5. HALFTONE RESULTS

In this section, several monochromatic and color halftone images are illustrated. All illustrated images are 256×256 pixels, except for the grayscale and color ramps, which are 59×512 pixels. All halftones are displayed at 150 dpi in order to clearly show the halftone structures. This means that the halftone images are magnified by four considering 600 dpi. Note that, the test images used in Section 4.1 and listed in Tables I, II, and III are a larger version (512×512) of those illustrated in this section. Some of the illustrated images are the actual halftones, i.e., binary images. They were obtained by directly halftoning the test image by different halftoning methods. Some of the illustrated images are not binary. In order to obtain those, the test image has been first compensated for dot gain by each halftoning method's specific dot gain curve. The compensated images have then been halftoned and thereafter the printer model has been applied. This means, given an original image g in Fig. 9, it is first compensated for dot gain to obtain g_c , which is then halftoned to achieve the halftone image b . The printer model is thereafter applied to this halftone image to obtain the representation of the printout, called b_p in Fig. 9. The dot gain compensation process and the printer model were applied using $\rho = 1.1$. Consequently, these images are not bitmaps any more and they are supposed to represent a magnification of the printouts. Four of the illustrated halftoning methods are our proposed methods, namely; first-order FM, second-order FM with ($\sigma_2 = 0.5$), ($\sigma_2 = 0.6$), and ($\sigma_2 = 0.7$). For three of the color halftones, the dot-off-dot versions of our proposed methods are also illustrated. Note that the dot off dot has only been applied to the C and M channels, and the Y channel is halftoned independent of the two others. Four of the illustrated methods are different error diffusion methods, namely; ED

(FS), MED (FS), ED (J), and MED (J). There are also four ordered dithering methods illustrated, namely; Clustered dot (8×8 and 16×16 threshold matrices), and Dispersed dot (Bayer 8×8 and 16×16 threshold matrices) ordered dithering.

Figure 19 is the halftone results (binary images) of the image called *Sari1* in Tables I and II. In Figures 20 and 21, the representation of the printouts for the images called *Sari2* and *Contrast* are illustrated. As can be seen in these results, they are not binary anymore. Figure 22 illustrates the halftones of a grayscale ramp without dot gain compensation and printer model application.

There are also four color images that have been halftoned. In Figure 23, the representation of the printouts for the color image called *Sari1* in Section 4.1.2 and Table III are illustrated. All three channels have been compensated using $\rho = 1.1$ before being halftoned and the printer model being applied. The dot-off-dot versions of our proposed halftones are shown in the second row. Figure 24 is the halftone results (binary) of the color image called *Sari2* in Section 4.1.2 and Table III. The dot-off-dot versions of our proposed halftones are shown in the second row. Figure 25 illustrates the halftone results (binary) of a part of the *Blue-Yellow Contrast* used in Section 4.4.2. Finally, Figure 26 shows the halftone results (binary) of a color ramp only consisting of cyan and magenta, in which the C and M channels are identical. This is the same image being referred to as *Ramp* in Section 4.1.2 and Table III. The dot-off-dot versions of our proposed halftones are shown right under their corresponding independent version. As can be seen in the dot-off-dot halftones in this figure, there is no dot overlap between cyan and magenta up to the middle of the ramp, where the sum of the coverage is less than or equal to 100%.

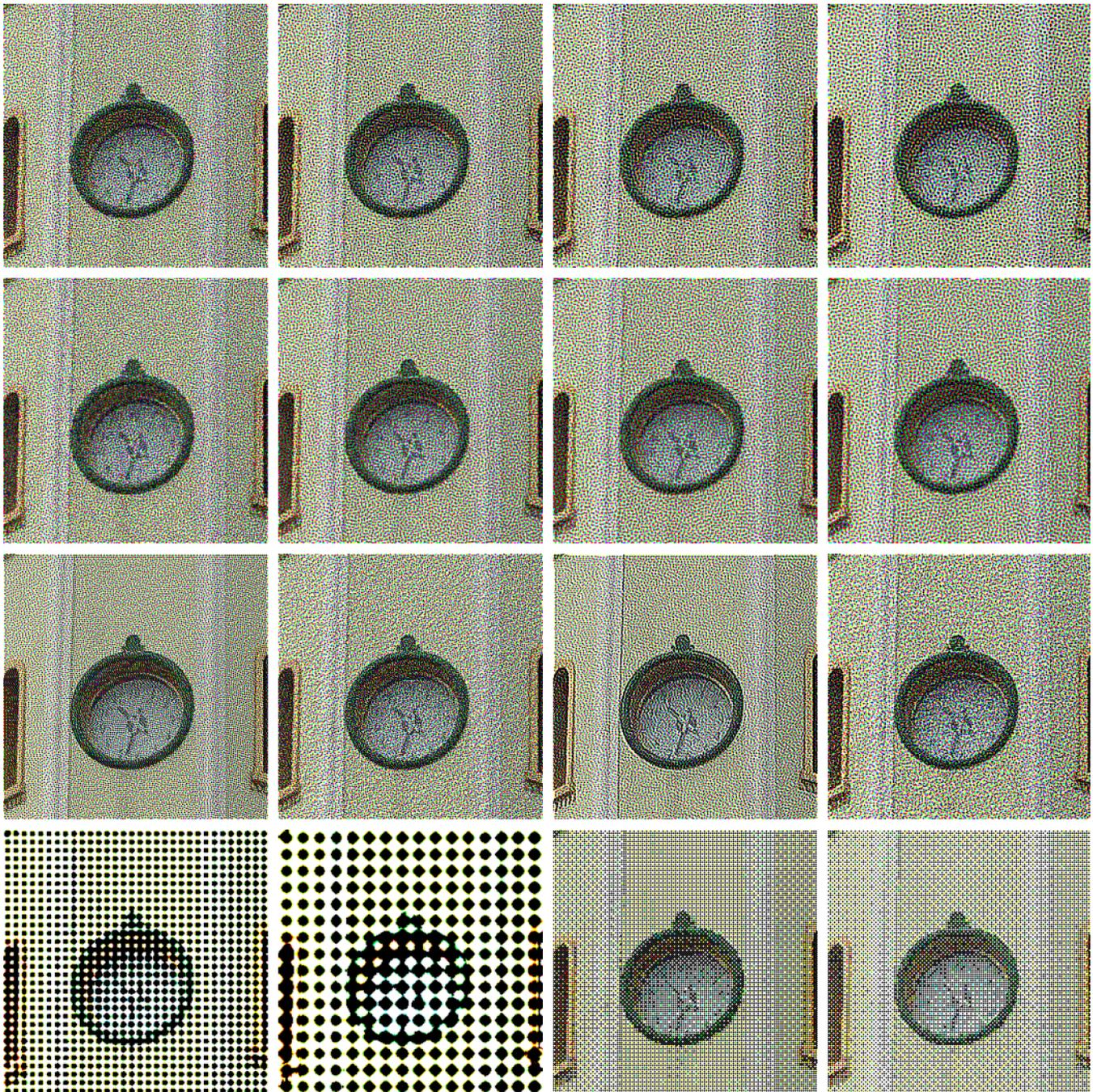


Figure 23. The test image is the color image *Sari* used in Section 4.1.2. These are not binary halftones. In order to obtain them, the CMY channels of the test image have been first compensated for dot gain and then halftoned before the printer model being applied, using $\rho = 1.1$. These images are supposed to represent a magnified version of the printout. From top to bottom and left to right: (First row) First-order FM, second-order FM ($\sigma_2 = 0.5$), second-order FM ($\sigma_2 = 0.6$), and second-order FM ($\sigma_2 = 0.7$). (Second row) The dot-off-dot version (only C and M channels) of the halftones in first row. (Third row) ED (FS), MED (FS), ED (J), and MED (J). (Fourth row) Clustered (8×8), Clustered (16×16), Bayer (8×8), and Bayer (16×16).

6. SUMMARY AND CONCLUSION

As we have studied twelve different halftoning methods employing several quality measures for monochromatic and color halftones, it is not easy to put the final conclusions into words. Therefore, we have decided to present all results in two different tables in order to make it easier to draw a final conclusion. In one of the tables, all twelve

halftones are compared, and in the other one, the ordered dithering halftones are discarded, and only eight halftones are compared. Besides the sharpness, there have been eight quality measures and print situations studied and illustrated in this article, namely; Mean Squared Error ($\rho = 1.1$, $R = 600$), Mean Squared Error ($\rho = 1.2$, $R = 1200$) (Section 4.1.1), S-CIELAB ΔE ($\rho = 1.1$, $R = 600$)

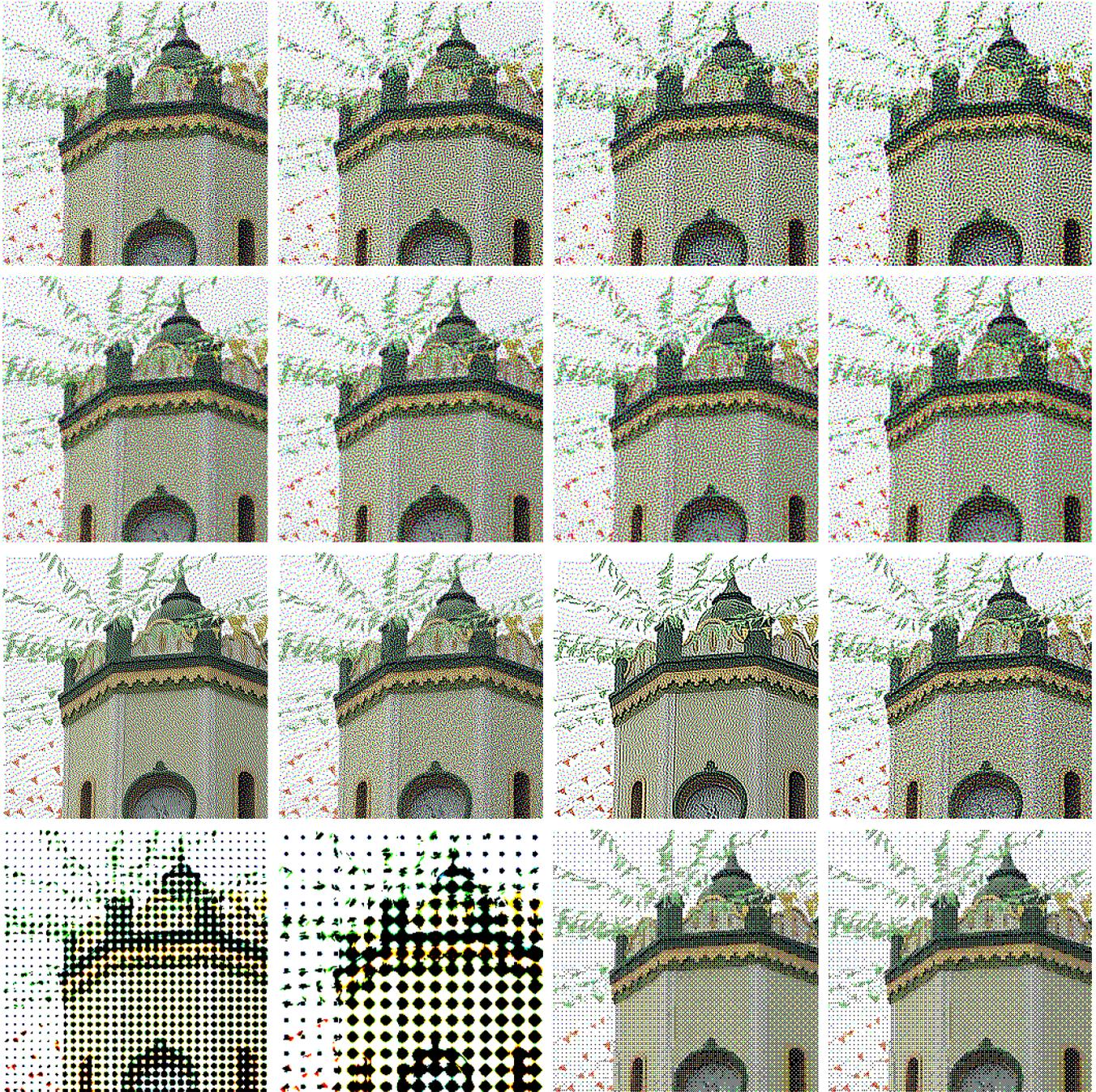


Figure 24. The test image is the color image *Sari2* used in Section 4.1.2. The CMY channels are all binary halftones and no dot gain compensation or printer model has been applied. From top to bottom and left to right: (First row) First-order FM, second-order FM ($\sigma_2 = 0.5$), second-order FM ($\sigma_2 = 0.6$), and second-order FM ($\sigma_2 = 0.7$). (Second row) The dot-off-dot version (only C and M channels) of the halftones in first row. (Third row) ED (FS), MED (FS), ED (IJ), and MED (IJ). (Fourth row) Clustered (8×8), Clustered (16×16), Bayer (8×8), and Bayer (16×16).

(Section 4.1.2), Gray level representation ($\rho = 1.1$, $R = 600$), Gray level representation ($\rho = 1.2$, $R = 1200$) (Section 4.2), Graininess (monochromatic) (Section 4.3.1), Graininess color (no dot gain compensation), and Graininess color (with dot gain compensation) (Section 4.3.2). For each of these quality measures, we rank the twelve (or eight) halftones from the best (rank 1) to the worst (rank 12 or 8) based on the

obtained quality value for each quality measure. Besides the rank, each halftone is also assigned a score. The score is scaled between 0 and 100, with 0 being the worst and 100 being the best. Let us illustrate that with an example. In Table I, there is an average value (titled *Av. 2*) specifying the mean squared error for each halftone. According to this table, ED (FS) gets rank 1, second-order FM ($\sigma_2 = 0.5$) rank 2, and finally

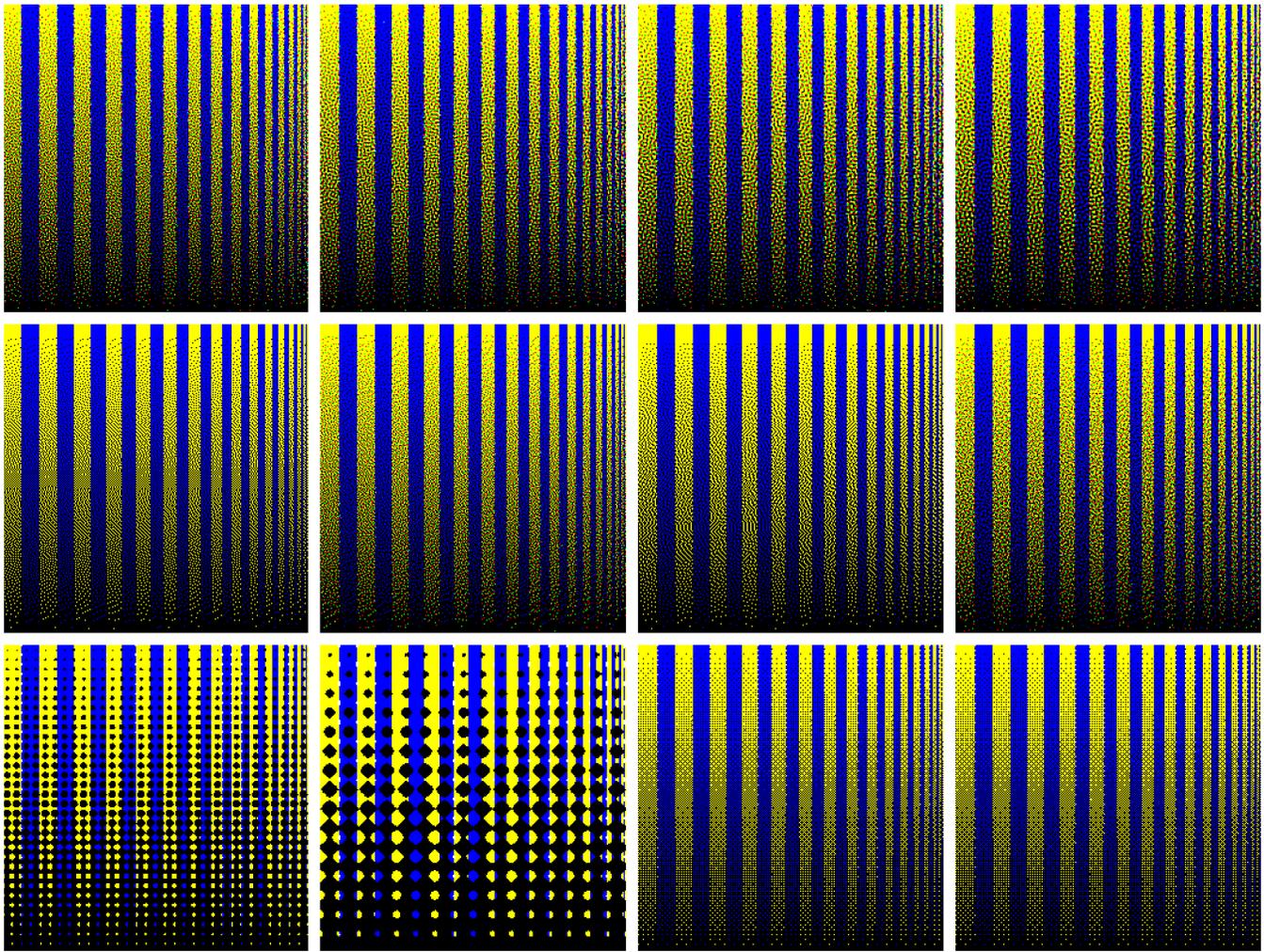


Figure 25. The test image is a part of the *Blue–Yellow Contrast* used in Section 4.4.2. The CMY channels are all binary halftones and no dot gain compensation or printer model has been applied. From top to bottom and left to right: (First row) First-order FM, second-order FM ($\sigma_2 = 0.5$), second-order FM ($\sigma_2 = 0.6$), and second-order FM ($\sigma_2 = 0.7$). (Second row) ED (FS), MED (FS), ED (J), and MED (J). (Third row) Clustered (8×8), Clustered (16×16), Bayer (8×8), and Bayer (16×16).

clustered dot ordered dithering (16×16) receives rank 12. The quality value for ED (FS) having rank 1 is 37.6, and that for clustered (16×16) having rank 12 is 5263. We linearly scale these quality values such that 5263 (the worst) becomes 0 and 37.6 (the best) becomes 100. This way all halftones are assigned a rank and a score for each quality measure.

Table VI shows the rank and the score (inside the parentheses) for the twelve halftones and eight quality measures. Table VII illustrates the rank and the score for our proposed and error diffusion halftones. For mean squared error (titled MSQ1 and MSQ2 in these tables), the last column titled **AV. 2** in Tables I and II, respectively, have been used. For S-CIELAB ΔE (titled ΔE), the last column titled **Av.** in Table III has been used. For gray level representation using ($\rho = 1.1, R = 600$) and ($\rho = 1.2, R = 1200$) titled GL (1.1) and GL (1.2), respectively, we have used the Mean values in Tables IV and V, respectively. For the three columns representing graininess (titled GR-mono,

GR-col, and GR-col2), we have used the sum of the values of the curves shown in Figs. 12, 13 (top row), and 13 (bottom row), respectively. In the next to the last column in Tables VI and VII, titled **Average**, the average score is shown. In the last column, titled **Final**, the final ranks and scores are shown, which were calculated by the same strategy using the average ranks and scores. Note that, we are aware of the fact that these scores might not have the same weight and impact because, for instance, a difference of 5 score in ΔE is probably not qualitatively equal to the same score difference in graininess. However, we believe that these scores together with the ranks can give a better indication of the halftone quality than only the ranks. Note also that, there are some halftones that get exactly the same average rank and are therefore assigned the same final rank. As seen in both tables, second-order FM ($\sigma_2 = 0.5$) is ranked the best of all twelve halftones. The same halftone, however, does not have the best score in any of the two tables and is worse than MED (FS) in

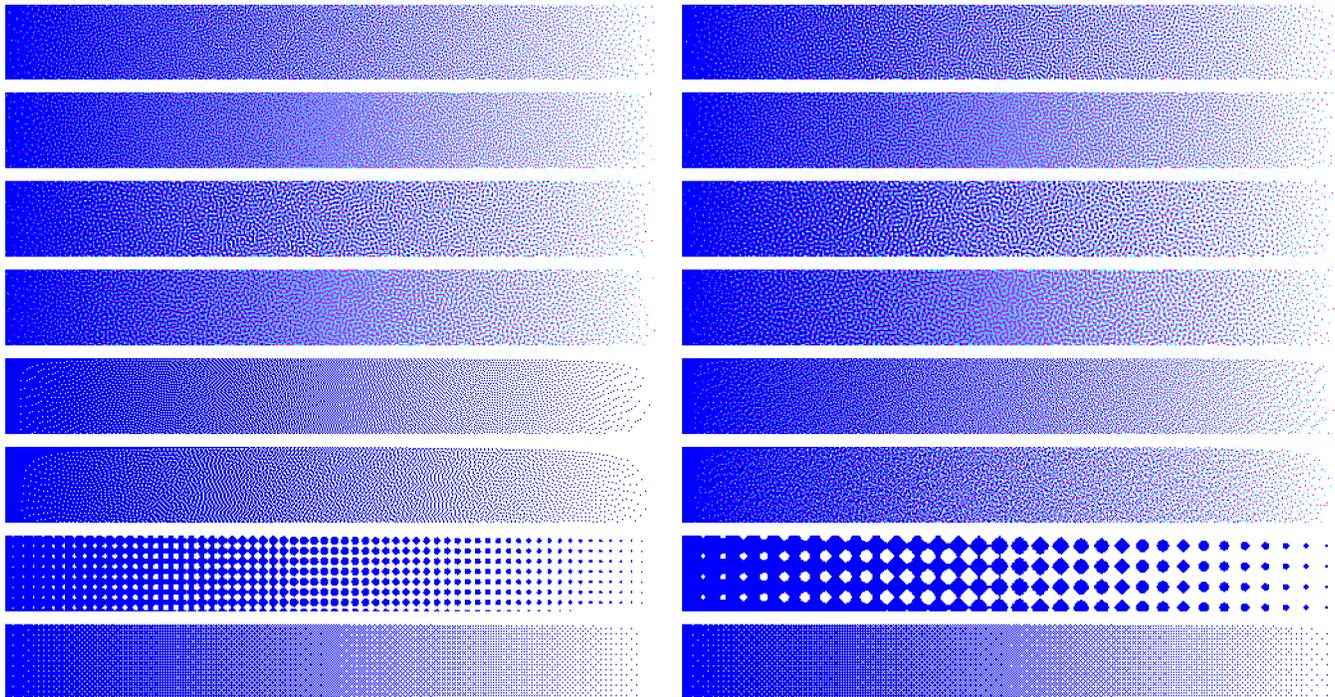


Figure 26. The test image is a color ramp only consisting of cyan and magenta, in which the C and M channels are identical (used in Section 4.1.2). The CM channels are all binary halftones and no dot gain compensation or printer model has been applied. From top to bottom and left to right: (First row) First-order FM and second-order FM ($\sigma_2 = 0.5$). (Second row) The dot-off-dot version of the halftones in first row. (Third row) Second-order FM ($\sigma_2 = 0.6$) and second-order FM ($\sigma_2 = 0.7$). (Fourth row) The dot-off-dot version of the halftones in third row. (Fifth row) ED (FS), and MED (FS). (Sixth row) ED (J) and MED (J). (Seventh row) Clustered (8×8) and Clustered (16×16). (Eighth row) Bayer (8×8) and Bayer (16×16).

Table VI. The rank and the score (inside parentheses) for all twelve halftones.

	MSQ1	MSQ2	ΔE	GL (1.1)	GL (1.2)	GR-mono	GR-col	GR-col2	Average	Final
FM-1st	4 (99.9)	5 (99.9)	3 (94.1)	3 (98.8)	3 (98.1)	7 (96.6)	5 (98.2)	8 (98.1)	98.0	5 (99)
FM-2nd (0.5)	2 (100)	1 (100)	7 (86.5)	1 (100)	1 (100)	6 (96.7)	8 (98.0)	4 (98.9)	97.5	1 (98)
FM-2nd (0.6)	3 (99.9)	2 (100)	9 (81.3)	1 (100)	2 (99.0)	8 (96.0)	9 (97.9)	2 (99.3)	96.7	4 (97)
FM-2nd (0.7)	6 (99.7)	4 (100)	10 (76.6)	3 (98.8)	4 (97.8)	9 (94.8)	10 (97.7)	7 (98.5)	95.5	6 (96)
ED-FS	1 (100)	6 (99.7)	1 (100)	8 (80.6)	8 (84.6)	3 (98.8)	1 (100)	6 (98.7)	95.3	3 (96)
MED-FS	7 (99.7)	3 (100)	2 (97.2)	5 (97.8)	5 (97.7)	5 (96.7)	2 (99.5)	3 (99.2)	98.5	2 (100)
ED-J	5 (99.8)	11 (96.9)	5 (89.0)	9 (75.4)	9 (77.3)	4 (97.3)	4 (98.4)	11 (92.5)	90.8	9 (90)
MED-J	10 (99.1)	7 (99.6)	6 (87.1)	6 (95.3)	6 (94.0)	10 (94.0)	6 (98.1)	10 (96.3)	95.4	10 (96)
Clustered-8	11 (71.3)	10 (97.9)	11 (42.5)	11 (17.9)	11 (24.0)	11 (49.4)	11 (91.9)	9 (96.4)	61.4	11 (52)
Clustered-16	12 (0)	12 (0)	12 (0)	7 (87.0)	7 (87.8)	12 (0)	12 (0)	12 (0)	21.9	12 (0)
Dispersed-8	9 (99.5)	9 (99.1)	8 (85.2)	12 (0)	12 (0)	1 (100)	3 (99.0)	1 (100)	72.9	8 (67)
Dispersed-16	8 (99.6)	8 (99.5)	4 (89.7)	10 (72.9)	10 (73.1)	2 (99.5)	7 (98.1)	5 (98.7)	91.4	7 (91)

terms of the score. Taking into account both the rank and the score of the halftones, we can conclude that among our proposed halftones, second-order FM ($\sigma_2 = 0.5$) is the best and first-order FM and second-order FM ($\sigma_2 = 0.6$) are also quite good. Among the error diffusion methods, ED (FS) and MED (FS) are both good and much better than ED (J) and MED (J). Bayer (16×16) performs similar to or slightly better than ED (J) and MED (J). Comparing our proposed

halftones with error diffusion methods, we can conclude that first and second-order FM ($\sigma_2 = 0.5$) generate halftones of almost the same quality as ED (FS) and MED (FS).

As you might have noticed, the sharpness has not been included in these two tables, because we actually did not present any value grading the sharpness, and sharper halftones do not necessarily mean better halftones. It was, however, shown that error diffusion methods generate

Table VII. The rank and the score (inside parentheses) for our proposed and error diffusion halftones.

	MSQ1	MSQ2	ΔE	GL (1.1)	GL (1.2)	GR-mono	GR-col	GR-col2	Average	Final
FM-1st	4 (87.9)	5 (97.5)	3 (74.6)	3 (95.3)	3 (91.7)	5 (55.0)	4 (22.3)	6 (81.8)	75.8	5 (86)
FM-2nd (0.5)	2 (99.8)	1 (100)	6 (42.1)	1 (100)	1 (100)	4 (56.1)	6 (13.1)	3 (93.8)	75.6	1 (86)
FM-2nd (0.6)	3 (91.1)	2 (100)	7 (20.2)	1 (100)	2 (95.4)	6 (42.5)	7 (9.2)	1 (100)	69.8	3 (75)
FM-2nd (0.7)	6 (69.9)	4 (98.8)	8 (0)	3 (95.3)	4 (90.3)	7 (18.0)	8 (0)	5 (87.9)	57.5	6 (52)
ED-FS	1 (100)	6 (90.0)	1 (100)	7 (20.8)	7 (32.3)	1 (100)	1 (100)	4 (90.6)	79.2	2 (92)
MED-FS	7 (63.2)	3 (100)	2 (88.2)	5 (91.1)	5 (89.9)	3 (57.6)	2 (78.5)	2 (98.1)	83.3	3 (100)
ED-J	5 (82.7)	8 (0)	4 (53.0)	8 (0)	8 (0)	2 (68.6)	3 (32.1)	8 (0)	29.6	7 (0)
MED-J	8 (0)	7 (86.9)	5 (44.8)	6 (80.7)	6 (73.7)	8 (0)	5 (17.0)	7 (56.0)	44.9	8 (28)

sharper monochromatic halftones than our proposed methods, while the opposite was concluded for color halftones. Regarding the blue- and green-noise characteristics, it was shown in Section 3.1 that all of our proposed first- and second-order halftones have very good blue- and green-noise characteristics, respectively. Non-modified error diffusion methods, however, did not show any blue-noise characteristics, whereas MED methods were slightly better, but still not as good as our proposed halftones. By studying the quantization noise spectra for different halftones, it was shown that ED (J) behaves more like a second-order FM than first-order FM.

One of the main goals of this study has been to figure out whether or not second-order FM halftones perform better than first-order FM in high print resolutions. The study in this article verifies that second-order FM with small clustered dot size is better than first-order FM in almost all quality aspects examined in this article. The only metric where this halftone was slightly worse than first-order FM was S-CIELAB ΔE , but if we interpret these values as ΔE_{ab} values, then the difference between these two halftones, which is 1.4 ΔE_{ab} in average, is only slightly over the just noticeable difference (JND) of $\Delta E_{ab} = 1$.

As discussed in Section 1, since the generated threshold matrices are image-independent, the only operation required to halftone an image is a comparison per pixel. Just to give an indication, assume a large format print of size 40×60 inches² (approximately 1×1.5 m²) at $R = 1200$ dpi. Thus, the image to be halftoned is $48\,000 \times 72\,000$ pixels. Halftoning such a large image only took 3 sec in MATLAB on a MacBook Pro (*Processor: 3.1 GHZ* and *memory: 16 GB, 2133 MHz*), without using any parallel processing. Hence, the speed of the proposed halftoning methods and their good quality make them feasible to be used in printing industry, especially for large format prints at high resolutions.

REFERENCES

- S. Gooran and B. Kruse, "High-speed first- and second-order frequency modulated halftoning," *J. Electron. Imaging* **24**, 023016 1–19 (2015).
- R. Ulichney, *Digital Halftoning* (MIT Press, Cambridge, MA, 1987).
- W. Floyd and L. Steinberg, "An adaptive algorithm for spatial gray-scale," *Proc. SID* **17**, 75–77 (1976).
- M. Analoui and J. P. Allebach, "Model-based halftoning using direct binary search," *Proc. SPIE* **1666**, 96–108 (1992).
- K. Kang, S. W. Kang, and C. Kim, "A modified error diffusion scheme based on the human visual model," *Recent Progress in Digital Halftoning II* (IS&T, Springfield, VA, 1999), pp. 30–34.
- R. Eschbach, "Error diffusion with homogeneous highlight and shadow response," *Recent Progress in Digital Halftoning II* (IS&T, Springfield, VA, 1999), pp. 173–178.
- S. Gooran, "Dependent color halftoning: Better quality with less ink," *J. Imaging Sci. Technol.* **48**, 354–362 (2004).
- T. N. Pappas and D. L. Neuhoff, "Model-based halftoning," *Proc. SPIE* **1453**, 244–255 (1991).
- T. N. Pappas and D. L. Neuhoff, "Least-squares model-based halftoning," *Proc. SPIE* **1666**, 165–176 (1992).
- J. L. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Trans. Inf. Theory* **IT-20**, 525–536 (1974).
- X. Zhang and B. Wandell, "A spatial extension of CIELAB for digital color image reproduction," *J. Soc. Inf. Disp.* **5**, 61–63 (1997).
- G. M. Johnson and M. D. Fairchild, "A top down description of S-CIELAB and CIEDE2000," *J. Color Res. Appl.* **28**, 425–435 (2003).
- E. A. Bernal, R. P. Loce, S. Wang, and Z. Fan, "Parametrically controlled stochastically seeded clustered halftones: design strategies and applications," *J. Electron. Imaging* **23**, 013010 1–12 (2014).
- D. L. Lau, G. R. Arce, and N. C. Gallagher, "Green-noise digital halftoning," *Proc. IEEE* **86**, 2424–2444 (1998).
- S. Gooran, "In dependent color halftoning, yellow matters," *J. Imaging Sci. Technol.* **50**, 448–457 (2006).
- J. A. C. Yule, *Principles of Color Reproduction* (John Wiley and Sons, New York, 1967).
- R. Ulichney, "Dithering with blue noise," *Proc. IEEE* **76**, 56–79 (1988).
- S. Gooran and L. Yang, "Frequency modulated halftoning and dot gain," *Proc. TAGA* (Technical Association of the Graphic Arts, Sewickley, PA, 2004), pp. 71–86.
- M. Namedanian, D. Nystrom, P. Z. Elias, and S. Gooran, "Physical and optical dot gain: Characterization and relation to dot shape and paper properties," *Proc. SPIE* **9015**, 901509 1–10 (2014).
- B. Bayer, "An optimum method for two-level rendition of continuous-tone pictures," *Proc. IEEE* **26**, 11–15 (1973).
- C. H. Son, Y. H. Cho, C. H. Lee, and Y. H. Ha, "Six color separation using additional colorants and quantitative granularity metric," *J. Imaging Sci. Technol.* **50**, 25–34 (2006).
- K. De and V. Masilamani, "Image sharpness measure for blurred images in frequency domain," *Procedia Eng.* **64**, 149–158 (2013).
- R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. (Pearson, New York, 2018).
- K. T. Knox, "Error image in error diffusion," *Proc. SPIE* **1657**, 268–279 (1992).