Digital Halftoning and Color Reproduction

Course book - TNM059

Sasan Gooran

Department of Science and Technology
Linköping University

January 9, 2018

Contents

Chapter 1

Introduction

Color reproduction is an essential part of our everyday life. When, for example, a digital camera takes a picture of a scene, this is actually the color reproduction of the original scene. Other examples are when an image is displayed using a computer monitor or a TV screen, or printed by a printer. However, different devices handle color in different ways. A monitor or LCD uses light to create colors while a printer utilizes inks or dyes. The success of a color reproduction lies in how "close" the reproduced scene is to the original scene, or how well the requirements of the user are fulfilled. Color reproduction is a complex process, as there usually are a number of different reproduction devices with completely different characteristics involved in the reproduction chain. The main goal of this course book is to introduce and describe the fundamentals of color reproduction, with an special focus on print reproduction. As digital halftoning, which transforms an original image to bitmaps to be sent to the printer, is an essential part of color reproduction in print, it makes an important part of this book.

The present book, therefore, introduces the basic concepts and models in digital halftoning and color reproduction and is supposed to cover the lectures in the course graphic arts (TNM059), given at the department of science and technology at Linköping University. The discussed methods, models and equations are thoroughly described by the use of illustrations and examples. At the end of some of the chapters, there are a number of exercises, for which you can find answers and solutions in Chapter 8 and Chapter 9, respectively.

This book consists of nine chapters as follows:

- **Chapter 1 - Introduction (1):** The present chapter.
- **Chapter 2 - Digital Image (2):** Includes very short introduction to binary numbers, digital images and memory usage.
- **Chapter 3 - Digital Halftoning - Achromatic (3):** Thoroughly describes important concepts in print reproduction and a number of important halftoning algorithms to transform a continuous tone image to

a binary image. This chapter is basically focused on achromatic reproduction. The concept of dot gain is briefly described at the end of this chapter.

- **Chapter 4 - Advanced Halftoning Methods (4):** Briefly introduces more advanced halftoning methods, by describing one iterative halftoning and discussing hybrid halftoning as well as second generation FM halftoning.
- **Chapter 5 - Halftone Image Quality (5):** Briefly introduces three uncomplicated quality metrics to objectively examine the quality of halftone reproductions.
- **Chapter 6 - Basic Color Science (6):** Describes, as its name implies, basic concepts in color science by starting with a brief explanation of what color is and how the human visual system perceives colors. This chapter also includes an introduction and description of a number of color spaces and a brief introduction of color management.
- **Chapter 7 - Color Reproduction (7):** Describes the basics of color reproduction, such as color gamut and additive and subtractive color mixing. Color halftoning, different printing strategies as well as a number of models to predict the final color of printed halftones are also described in this chapter.
- **Chapter 8 - Answers (8):** Includes the answers to the exercises at the end of Chapters 2, 3, 6 and 7.
- **Chapter 9 - Solutions (9):** Includes the solutions to the exercises at the end of Chapters 2, 3, 6 and 7.

Chapter 2

Digital Image

In order to store the information from the original photographs in computer, they have to be converted to digital format. In other words, the photographs should be digitized, which is normally done by scanning them. If you have taken a picture by a digital camera, of course you don't need to scan it because the image is already saved in digital format. In this short chapter, we firstly give a very brief introduction to binary numbers, which is followed by discussing digital images and how much memory is needed to save them. Finally in this chapter, two exercises are provided, for which you can find the answers and the solutions in Chapter 8 and 9, respectively.

2.1 binary numbers

A binary number is a number expressed in the binary numeral system or base-2 numeral system which represents numeric values using two different symbols: typically 0 (zero) and 1 (one). The binary system is used by almost all modern computers and computer-based devices. Each digit is referred to as a bit. For example, a two-digit binary number, needs 2 bits to be saved in the computer memory. Since each digit (or bit) only takes two symbols, i.e. 0 or 1, a two-digit binary number can represent $2^2 = 4$ different values, i.e. '00', '01', '10' and '11'. A three-digit binary number, for instance, can represent $2^3 = 8$ different values; namely '000', '001', '010', '011', '100', '101', '110' and '111'.
The n-digit binary number $b_{n-1}b_{n-2}\cdots b_1 b_0$ is converted to a decimal number by:

$$\textbf{decimal number} = b_{n-1}2^{n-1} + b_{n-2}2^{n-2} + \cdots + b_1 2^1 + b_0 2^0. \tag{2.1}$$

For example, the binary number 1010 represents the decimal number $1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 8 + 2 = 10$. Therefore, two-digit binary numbers 00, 01, 10 and 11, represent decimal numbers 0, 1, 2 and 3. In other words, n-digit binary numbers can represent integers from 0 up to $2^n - 1$.

Example 2.1 *What decimal numbers do the following binary numbers represent?*
a) 1111
b) 1000
c) 11001

Solution (a): *It represents:* $1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 8 + 4 + 2 + 1 = 15$
Solution (b): *It represents:* $1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0 = 8 + 0 + 0 + 0 = 8$
Solution (c): *It represents:* $1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 16 + 8 + 0 + 0 + 1 = 25$

As discussed above, a n-digit number needs n bits of memory to be stored. For example, an 8-digit number that could represent all integers from 0 to $2^8 - 1 = 255$ needs 8 bits (also denoted 8 b) to be stored. A group of 8 bits is called 1 byte, therefore an 8-digit number needs 1 byte (also denoted 1 B) to be stored. Notice that a bit is usually denoted by small b and a byte is denoted by capital B. In binary systems, a kilobyte (KB) is defined to be 1024 bytes (B) and a megabyte (MB) to be 1024 kilobytes and a gigabyte (GB) to be 1024 megabytes. However, it is quite common to assume, like in metric systems, 1 KB to be 1000 B, and 1 MB to be 1000 KB and so on. In this book, for the simplicity of the calculations, we also use the metric conversion between MB, KB and B, although not completely correct.

Example 2.2 *We have a vector containing 250000 elements. Each element in the vector holds an integer between 0 and 230. Find out how much memory (expressed in both KB and MB) is needed to store this vector.*

Solution: *Each element in the vector is an integer between 0 and 230, meaning that we need at least 8 bits (or 1 byte) to represent each element, because 8-digit numbers can represent integers from 0 to 255. Since there are 250000 elements and each element needs 1 byte, totally* $250000 \cdot 1 = 250000$ *bytes (B) is needed. Assuming a KB to be 1000 B, then the memory needed to store this vector is 250 KB (kilobytes) or 0.25 MB (megabytes).*

2.2 Digitization and memory usage

When a scanner scans a photograph or when a digital camera takes a picture, the information (color) is measured at discrete points (pixels) of the original. The number of these sample points per an inch is called the scanning resolution and is denoted by *ppi* (pixels per inch). In the grayscale photographs, the gray tone of each sample point is normally stored using eight bits (or one byte). Thus, the digital grayscale image contains 256 levels of gray tones, varying from white (0 or 255) to black (255 or 0). In the case of color photography, the color of each sample point is represented by its three primary colors, namely Red (R), Green (G) and Blue (B). In this case, each sample point needs 3×8 bits (3 bytes) to be stored in computer. The bit depth is thus 24 bits or 3 bytes and the digital color image contains 256^3 (about 16.8 million) colors. It is obvious that the higher the scanning resolution *ppi*, the more detailed the stored image and consequently the bigger the image file. Therefore, it is important to keep the resolution as low as possible. The choice of the resolution mostly depends upon the device the digital image will eventually be reproduced by. For example, if the digital image is to be reproduced by a computer screen at 1:1 scale, it is unnecessary to scan the original photograph with a *ppi* higher than the reproduction resolution of the screen. If your digital image is supposed to be printed, the choice of *ppi* depends upon a number of other factors that will be discussed in Section 3.1.1. Figure 2.1 illustrates a test image scanned using a sufficient *ppi* and a too small *ppi*. As seen in Figure 2.1 (b), the image is blurred and the details are not reproduced as clearly as in the image in (a).

(a) (b)

Figure 2.1: An original image scanned using: (a) sufficient ppi, (b) too small ppi.

Example 2.3 *A* 20×30 cm^2 *color photograph is scanned using ppi* $= 100$.
a) *How large (pixel* $\times$ *pixel) is the scanned digital image?*
b) *How much memory is needed to store this digital image in raw format?*

Solution (a): *Assuming an inch to be 2.5 cm, the photograph is* $\frac{20}{2.5} \times \frac{30}{2.5} = 8 \times 12$ $inch^2$. *ppi* $= 100$ *means 100 samples (pixels) per an inch, meaning that the digital image will be* $8 \cdot 100 \times 12 \cdot 100 = 800 \times 1200$ *pixels.*

Solution (b): *A* 800×1200 *pixels image totally contains* $800 \cdot 1200 = 960000$ *pixels. Since the photograph is color, then the digital image consists of three channels; Red (R), Green (G) and Blue (B). Assuming each pixel needs 8 bits (or one byte), the total amount of memory needed to store this digital color image is* $960000 \cdot 3 \cdot 8 = 23040000$ *bits or* $960000 \cdot 3 = 2880000$ *bytes. Since one*

Megabyte is approximately 10^6 *bytes, then we could say that the digital image is* 23.04 *Megabits (Mb) or* 2.88 *Megabytes (MB).* ***Notice*** *that, as mentioned in the example, the image was supposed to be stored in raw format, i.e. without compression. If the image was compressed and saved for example in jpg format, it would of course need less memory to be stored.*

2.3 Exercise set

2.1. A color photograph has been scanned at $ppi = 100$ and the digital image is 2000×2000 pixels. Assume an inch to be 2.5 cm.

a) How large ($cm \times cm$) is the photograph?

b) How much memory is needed to store the digital color image?

2.2. A color photograph has been scanned at $ppi = 200$ and the digital image needs 0.72 Megabytes to be stored. Find the area of the photograph in cm^2.

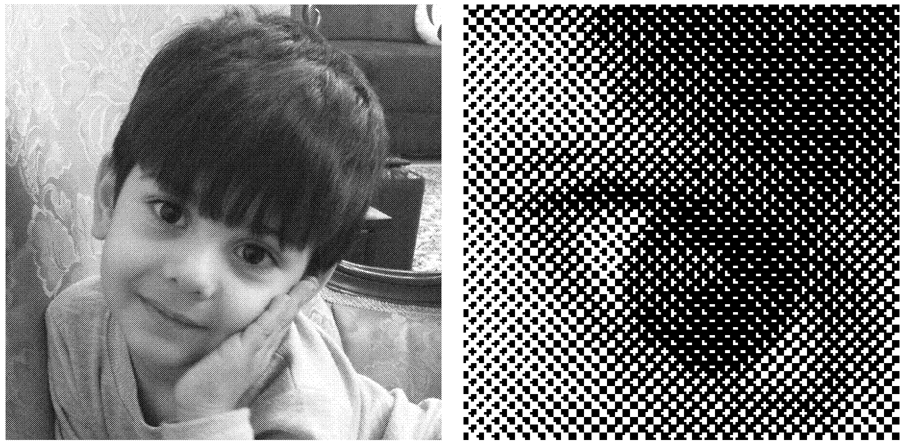
Chapter 3

Digital Halftoning - Achromatic

The digitized photographs, or the digital images, will eventually be reproduced by a device or a number of devices. Most of the image reproduction devices, particularly the printing devices, are restricted to few colors while the digital image mostly consists of millions of colors. In the grayscale case, as discussed in Chapter 2, the digital image consists of 256 different shadows of gray (assuming the bit depth to be 8 bits), while the black and white printers normally use only one colored ink, i.e. black. These 256 levels of gray should somehow be represented by the black color and the white substrate. In order to do that, the original digital image, which in the following will be called the original image, is transformed into a binary image consisting of 1's and 0's, i.e. a bitmap. A 1 at a pixel means that a black dot should be printed (or shown) at that particular position and a 0 means that the corresponding position should remain empty. This transformation from a continuous tone image to a bitmap representation is called Halftoning, also referred to as Screening. Halftoning works because of the inability of the human eye to detect the halftone dots when they are sufficiently small or/and the viewing distant is sufficiently long. If the dots are small enough, or/and the viewing distance is long enough, the human eye integrates the black halftone dots and the non-printed areas as varying shades of gray, yielding the impression of a continuous tone image. Figure 3.1 (a) shows a halftoned image and an enlargement of a part of this image is shown in (b). Since the halftone dots in the image in Figure 3.1 (a) are sufficiently small, they are hardly detected by the eye at the normal viewing distance. On the other hand, the dots in the image in Figure 3.1 (b) are easily detected by the eye at the same distance.

Example 3.1 *An achromatic halftoned image (bitmap) is* 4000×6000 *pixels. How much memory is needed to store this halftoned image?*

Solution: *The total number of the pixels in the image is* $4000 \times 6000 = 24 \times 10^6$.



(a) (b)

Figure 3.1: The image is halftoned. The halftone dots in the image in (a) are hardly detected by the human eye at the normal viewing distance while they are easily detected in the image in (b).

Since the image is halftoned (binarized), each pixel only contains either 0 or 1, meaning that only ***one bit*** *is needed for each pixel. The image is achromatic, i.e. only one channel. Therefore, the total amount of the needed memory is:* 24×10^6 *bits or* 24 *Megabits (Mb). This is of course equal to* $\frac{24}{8} = 3$ *Megabytes (MB).*

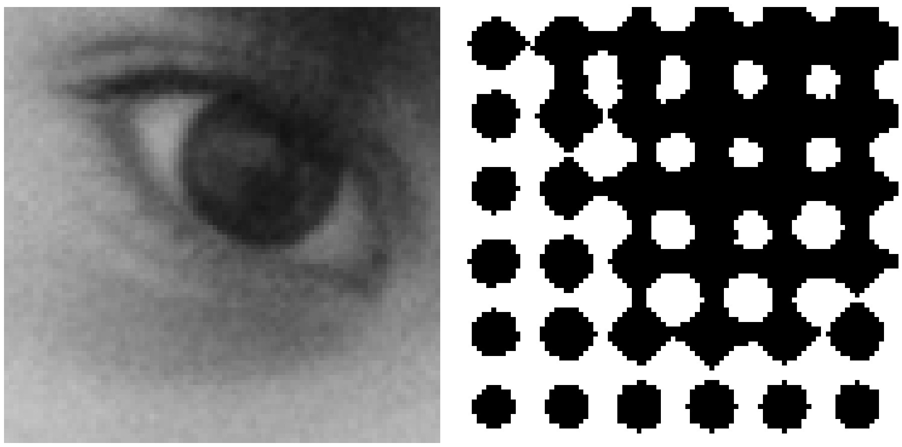
When it comes to color images, this transformation is performed for a number of color channels, normally for the color channels that the reproduction device uses. More details about color halftoning are given in Section 7.3.
It is worth mentioning that in this book, all images are supposed to be normalized between 0 and 1, 0 representing white and 1 representing black.
In this chapter, basic concepts of halftoning are firstly described, followed by the introduction of three basic and well-known halftoning methods. The two main types of halftoning methods, i.e. Amplitude and Frequency modulated, are also discussed. Finally in this chapter, a number of exercises are provided, for which you can find the answers and the solutions in Chapter 8 and 9, respectively.

3.1 Basic concepts

In this section, we introduce a number of important concepts that are fundamental for understanding halftoning. The concepts such as, print resolution, screen frequency, halftone cell are discussed in this section.

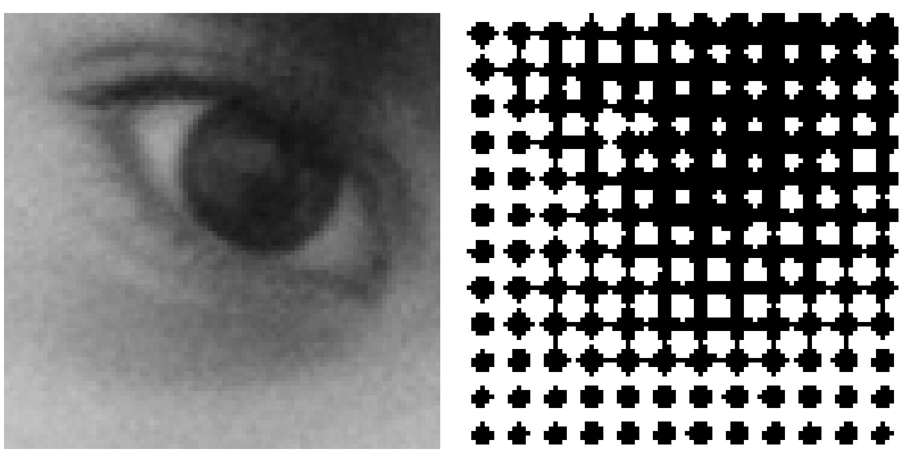
3.1.1 Screen frequency (lpi)

One of the most straightforward ways of halftoning is to represent different areas in the original image by a halftone screen (halftone cell). The fractional area covered by the ink represents the gray tone of the corresponding area in the original image. Figure 3.2 (a) shows a part of the original test image. Assume that this image is divided into 6×6 sub-images and the average gray tone of each sub-image is calculated.



(a) *(b)*

Figure 3.2: The original image in (a) is divided into 6×6 *sub-images and each sub-image is represented by a halftone cell in (b).*



(a) *(b)*

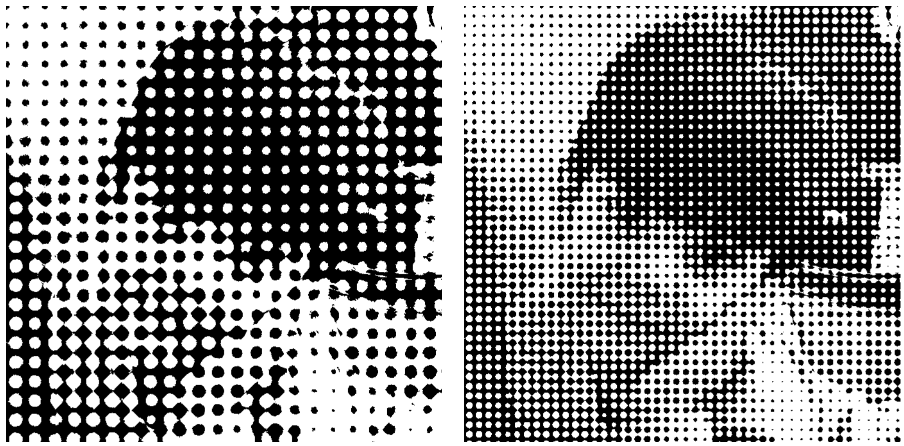
Figure 3.3: The original image in (a) is divided into 12×12 *sub-images and each sub-image is represented by a halftone cell in (b).*

Each of this sub-image is then represented by a corresponding halftone cell, in which the fractional area of the black ink represents the average gray tone of the corresponding sub-image in the original image. The halftoned image in Figure 3.2 (b), therefore, consists of 6×6 halftone cells. As seen in this figure, the size of the black halftone dots are larger in the cells corresponding to the darker parts of the original image. The same original image is now divided into 12×12 sub-images and represented by as many halftone cells, see Figure 3.3. As seen in these figures, the more the number of the halftone cells, the smaller the halftone dots, and therefore the more difficult for the human eye to detect them. One of the most fundamental concepts in conventional halftoning and printing is the screen frequency (or line screen ruling) that is the number of halftone cells (or halftone dots) per an inch. Screen frequency is denoted by *lpi* (lines per inch). It is obvious that the higher the *lpi*, the smaller the halftone cell and consequently the halftone dot and the more difficult for the eye to detect the halftone dots. Studies have shown that the halftone dots are not detected by the eye from the normal viewing distance (around 15 cm) at screen frequencies around 200 *lpi* and above. Figure 3.4 (a) and (b) show the original image being halftoned using $lpi = 10$ and $lpi = 20$, respectively. It is clearly visible in this figure that, the higher the *lpi* the smaller the halftone dots and the better the reproduction of the details of the original image. Notice that, in this figure the screen frequencies have been chosen to be very low intentionally and for the purpose of illustration. In practice, the screen frequency is much higher but it is not always possible to have it as high as one would wish. Among the important factors limiting the screen frequency are the quality of the paper and the printing technology. In newspaper print, for instance, usually the paper is uncoated and of lower quality, and therefore the used screen frequency is quite low and commonly between 65 and 100 *lpi*. In order to produce high quality prints, for example catalogs, usually coated and glossy paper substrates are used and therefore a higher screen frequency can be utilized, which is commonly between 150 to 200 *lpi*.

In Chapter 2, we had a short discussion on how to choose the scanning resolution *ppi* (pixels per inch). We mentioned that if the digital image is supposed to be displayed by for example a computer screen, it is unnecessary to scan the image with a *ppi* higher than the resolution of that computer screen, if the reproduced image is supposed to be the same size as the original. However, when it comes to print, one should take into account the screen frequency of the print before deciding the scanning resolution. There is a rule of thumb that says that the scanning resolution (*ppi*) should be about 1.5 to 2 times the screen frequency (*lpi*) used in the press at 1:1 scale. The scaling factor should also be multiplied by this factor if the printed image is not supposed to be the same size as the original photograph. This rule of thumb can be summarized in the following equation:

$$ppi = \frac{printed\ size}{original\ size} \cdot 2 \cdot lpi, \tag{3.1}$$

where *ppi* is the scanning resolution and *lpi* is the screen frequency.



(a) (b)

Figure 3.4: The original image is halftoned using (a) lpi = 10 and (b) lpi = 20.

Example 3.2 *An achromatic photograph is* 10×15 cm^2 *and you want to scan it and send the digital image to be printed in a newspaper. You want the printed reproduction to be* 20×30 cm^2*. The print shop uses a screen frequency of 85 lpi. What scanning resolution (ppi) do you need to use?*

Solution: *The scaling factor is 2, because the print size, according to the example, is twice the original size in each direction (i.e.* $\frac{20}{10} = \frac{30}{15} = 2$*). According to the rule of thumb in Equation 3.1, you need to have:* $ppi = 2 \cdot 2 \cdot 85 = 340$.

Example 3.3 *How much is it possible to maximally enlarge an image (compared to the original) and at the same time obtain the highest possible print quality if the scanning resolution is 1200 ppi and the screen frequency is 150 lpi?*

Solution: *According to the rule of thumb in Equation 3.1, we will have:* $1200 = scaling\ factor \cdot 2 \cdot 150$*, giving* $scaling\ factor = 4$*. This means that in order to obtain the highest print quality possible, you can maximally enlarge the reproduction four times in each direction (or 16 times in area).*

3.1.2 Print resolution (dpi)

As discussed in Section 3.1.1, one of the most straightforward ways to halftone an image is to represent different areas of the image by halftone cells. The coverage of the halftone dot in each halftone cell represents the average tone of the corresponding area in the original image, see Figures 3.2 and 3.3 again. Since the halftoned image is also a digital image (it is a bitmap), in order to represent different gray tones, the halftone cell itself has to be divided into smaller points (pixels), which are referred to as microdots. Two halftone cells, each consisting of 8×8 microdots are illustrated in Figure 3.5. The halftone dot in Figure 3.5 (a) is 2×2 microdots and thus representing the gray tone of $\frac{4}{64}$, because 4 out of 64 microdots in this halftone cell are covered by the ink. Recall that, we assume a filled position (dot) to be represented by 1, and an empty position by 0. The gray tone represented by the halftone cell in Figure 3.5 (b) is $\frac{44}{64}$, because 44 microdots are covered by the ink. Totally, it is possible to represent 65 ($= 8^2 + 1$) different gray tones by 8×8 halftone cells, because the number of microdots being covered by ink could be 0, 1, 2, ..., 63, 64, representing the gray tones $\frac{0}{64} = 0$, $\frac{1}{64}$, $\frac{2}{64}$, ..., $\frac{63}{64}$, $\frac{64}{64} = 1$, respectively.

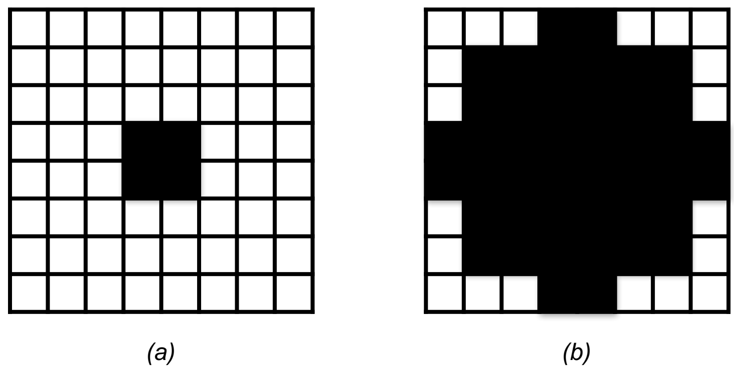


Figure 3.5: Two halftone cells (halftone screens) representing gray tones (a) 4/64 and (b) 44/64.

The number of the microdots per inch is called the print resolution and is denoted by *dpi* (dots per inch). As discussed in Section 3.1.1, the number of halftone cells per inch is the screen frequency (*lpi*). Since the halftone cells consist of the microdots, it is easy to figure out that the ratio $\frac{dpi}{lpi}$ determines the size of the halftone cells. For example, a print at 2400 *dpi* using the screen frequency $lpi = 150$ means that it uses halftone cells consisting of $\frac{2400}{150} \times \frac{2400}{150} = 16 \times 16$ microdots. This halftone cell could represent up to $16^2 + 1 = 257$ gray tones. Hence, Equation 3.2 shows how the number of gray tones can be calculated given *lpi* and *dpi*:

$$number\ of\ gray\ tones = (\frac{dpi}{lpi})^2 + 1. \tag{3.2}$$

Example 3.4 *A* 4800×3600 *pixels halftoned image is printed at* 1200 *dpi. What is the size (cm × cm) of the printed image?*

Solution: *A halftoned image is a bitmap, meaning that each pixel holds a 1 or a 0. When this image is printed, each pixel becomes a dot in print.* $dpi = 1200$ *means that 1200 pixels in the bitmap become one inch when printed, meaning* *that the printed image will be:* $\frac{4800}{1200} \times \frac{3600}{1200} = 4 \times 3$ *(inch* $\times$ *inch), i.e. around* 10×7.5 *(cm* $\times$ *cm).*

Example 3.5 *An achromatic photograph has been scanned at* $ppi = 200$ *and the digital image becomes* 2000×2000 *pixels. An inch is around* 2.5 *cm.*
a) *How large (cm* $\times$ *cm) is the photograph?*
b) *This digital image has been halftoned and the halftoned image becomes* 3000×3000 *pixels. How large (cm* $\times$ *cm) is the printed image if it is printed at* $dpi = 600$.

Solution (a): *Scanning resolution at 200 ppi means that one inch in the photograph becomes 200 pixels in the digital image. Therefore, the photograph must have been* $\frac{2000}{200} \times \frac{2000}{200} = 10 \times 10$ $inch^2$*, which is* 25×25 cm^2.

Solution (b): *Print resolution at 600 dpi means that 600 pixels in the digital halftoned image (bitmap) become one inch in print. Since the halftoned image is* 3000×3000 *pixels, then the printed image will be:* $\frac{3000}{600} \times \frac{3000}{600} = 5 \times 5$ $inch^2$*, which is* 12.5×12.5 cm^2.

Example 3.6 *Assume a print setup using* $dpi = 1200$*. What is the maximum lpi that should be utilized if the number of gray tones shall not be less than 145?*

Solution: *According to Equation 3.2:* $145 = (\frac{1200}{lpi})^2 + 1$*, i.e.* $\frac{1200}{lpi} = 12$*, giving* $lpi = 100$*. This means that if a lpi higher than 100 is used, then the number of gray tones will become less than 145. It is worth noticing that, although higher lpi is desirable to have halftone dots sufficiently small, it will however result in fewer number of gray tones to be represented.*

As can be observed from Equation 3.2 and also pointed out in the above example, a higher *lpi* will lead to a decrease of the number of gray tones when the *dpi* is constant. Choosing an appropriate *lpi* is therefore a trade off between the number of gray tones and the fine details.
Something worth mentioning here is that the straightforward way of halftoning that has been discussed so far, leads to periodic halftone structures, see for example Figures 3.2, 3.3 and 3.4, again. The screen frequency (*lpi*) is only used for this type of halftones, because the distance from the center of a halftone dot to the next one is constant. This will be further discussed in Section 3.3.

Example 3.7 *An achromatic photograph has been scanned at* $ppi = 200$ *and the digital image is* 1000×1000.
a) *The digital image is halftoned (periodic structure). What is the maximum lpi one should use for the highest possible quality if the printed image is supposed to be the same size as the photograph?*
b) *How large (pixel* $\times$ *pixel) must the halftoned image have been if it has been printed at* 600 *dpi?*
c) *How much memory is needed to store the halftoned image?*
d) *How many gray tones is possible to represent?*

Solution (a): *Since the printed image is the same size as the photograph, the scaling factor is 1. Therefore, according to the rule of thumb in Equation 3.1:* $200 = 1 \cdot 2 \cdot lpi$*, giving* $lpi = 100$*.*

Solution (b): *Since* $ppi = 200$ *and the digital image is* 1000×1000*, then the photograph must have been* 5×5 $inch^2$*, (why? see previous examples). The printed image is also supposed to be the same size, meaning* 5×5 $inch^2$*. Since* $dpi = 600$*, the halftoned image must have been* $5 \cdot 600 \times 5 \cdot 600 = 3000 \times 3000$ *pixels.*

Solution (c): *According to part (b), the halftoned image is* 3000×3000 *pixels. Since the image is halftoned (bitmap), only one bit per pixel is needed to store it. The total memory needed is therefore:* $3000 \cdot 3000$ *bits or* 9 *Mb or* $\frac{9}{8}$ *MB (Megabytes).*

Solution (d): *According to Equation 3.2, we have: number of gray tones =* $(\frac{600}{100})^2 + 1 = 37$*, because* $dpi = 600$ *(given in part (b)) and* $lpi = 100$ *(calculated in the solution for part (a)).*

3.2 Basic halftoning methods

In this section, we introduce three basic and well-known halftoning methods, namely threshold halftoning, table halftoning and error diffusion. As will be discussed, the two former halftoning methods will commonly result in periodic halftone structures (as also been discussed and shown in previous sections) while error diffusion will result in non-periodic and stochastic-like structures.

3.2.1 Threshold halftoning

As discussed, halftoning means transforming a digital image with pixels holding 256 different values (assuming the bit depth to be 8 bits) between 0 and 1 into a binary image with pixels holding either 0 or 1. One of the most basic ways of doing that is to define a threshold and replace all pixels in the original image with values greater than (or equal to) the threshold by 1, and the rest by 0. Since the original image is supposed to be scaled between 0 and 1, the most logical constant threshold value would be 0.5. Figure 3.6 (a) shows the test image being halftoned by the constant threshold of 0.5. Figure 3.6 (b)-Up shows an original digital gray-scale ramp, with smoothly varying tones from 0 (white) at the left to 1 (black) at the right. This ramp has also been thresholded by 0.5, which is shown in Figure 3.6 (b)-Down. As seen in these images, all pixels in the original image holding values greater (darker) than 0.5 has been replaced by black (1), and the rest of pixels by white (0, or no print). As expected and shown in Figure 3.6, these halftones only represent two levels of gray, either black or white.
However, it is not common to use constant thresholds but rather threshold

(a) *(b)*

Figure 3.6: (a) The test image is thresholded by 0.5. (b) Up: original gray-scale ramp. Down: the ramp thresholded by 0.5.

matrices. Threshold halftoning method can be described by following equation:

$$b(i,j) = \begin{cases} 1 & \textbf{if g(i,j)} \geq \textbf{t(i,j)} \\ 0 & \textbf{if g(i,j)} < \textbf{t(i,j)} \end{cases} , \tag{3.3}$$

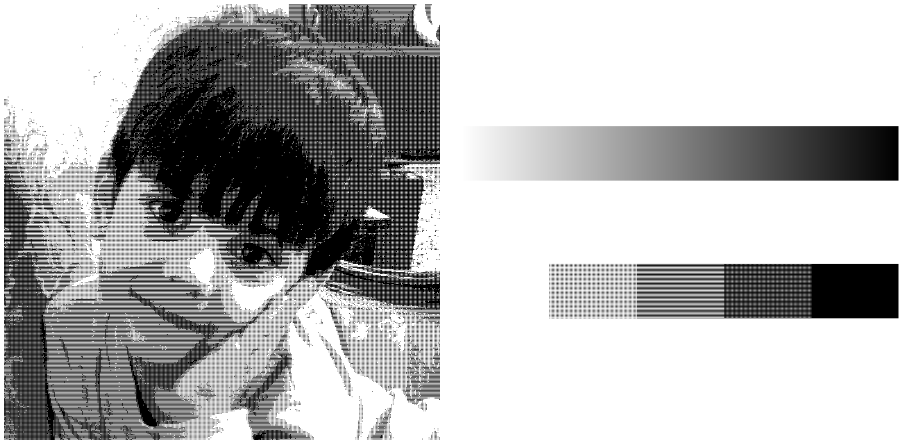
where b, g and t denote the final halftoned image, the original image and the threshold matrix, respectively. The pixel value at each position (i, j) in g is compared with the threshold value at the corresponding position in the threshold matrix t. If it is greater than or equal to the threshold, then a 1 (black dot) is set at the corresponding position in the halftoned image b. Otherwise, a 0 (white dot - no print) is set there. Observe that, in Equation 3.3 it is assumed that the threshold matrix is the same size as the original image. If the threshold matrix is smaller than the original image, which normally is the case, it should be periodically repeated in both directions to be the same size as the original image.

One of the important factors affecting the result of threshold halftoning is the size of the threshold matrix (before being repeated). Commonly, and also in this book, the threshold matrices are first filled with successive positive integers, i.e. 1, 2, 3 etc and then normalized to be between 0 and 1. Let us have a look at a simple 2×2 threshold matrix, for example $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Since the pixel values of the digital images are supposed to be between 0 and 1, then the threshold matrices also need to be normalized. It is tempting to normalize this 2×2 matrix by dividing its elements by 4 (the maximum threshold value) giving $TR_1 = \begin{bmatrix} 0.25 & 0.5 \\ 0.75 & 1 \end{bmatrix}$. The problem normalizing the threshold matrix by diving it by its largest threshold value is that you will end up having a threshold value equal to 1. Let us examine this normalized threshold matrix by using it to threshold the test image and the gray-scale ramp, see Figure 3.7. As seen in these images, there is no level to represent a completely black area. The reason is that there is a 1 in the threshold matrix, and all pixel values in the image are always less than that specific threshold value. This means that the resulting halftone will always have a 0 (white "dot") at that specific position, meaning that it will never be completely black.

(a) *(b)*

Figure 3.7: (a) The test image is thresholded by TR_1. (b) Up: original gray-scale ramp. Down: the ramp thresholded by TR_1.

However, the correct way of normalizing such a threshold matrix is to divide its elements by the largest threshold value plus 1. The 2×2 threshold matrix, $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, should therefore be normalized by dividing it by 5, giving $TR_2 = \begin{bmatrix} 0.2 & 0.4 \\ 0.6 & 0.8 \end{bmatrix}$. The results of thresholding the test image and the ramp by TR_2 are shown in Figure 3.8. As seen in this figure, both the completely white and completely black areas can be represented.
The conclusion is that, if the threshold matrices are filled by successive positive integers, the number of gray levels they represent is the largest threshold value plus 1, and the threshold matrices are also normalized by this number (i.e. the number of gray levels they represent). Another important point worth mentioning here is that, in this example, the threshold matrix has been 2×2, and since the test image is 680×680 pixels, the threshold matrix was repeated (tiled) 340 times in both directions to be the same size as the test image. Since in most common cases, the threshold matrix is repeated a number of times, then threshold halftoning will commonly result in periodic halftone structures.
Besides the size of a threshold matrix, which decides the number of the gray



(a) *(b)*

Figure 3.8: (a) The test image is thresholded by TR_2. (b) Up: original gray-scale ramp. Down: the ramp thresholded by TR_2.

levels it can represent, the arrangement of the threshold values in the threshold matrix also affects the final result. Let us illustrate that by thresholding the test image by the two threshold matrices in Equation 3.4.

$$TR_3 = \begin{bmatrix} 7 & 8 & 9 & 10 \\ 6 & 1 & 2 & 11 \\ 5 & 4 & 3 & 12 \\ 16 & 15 & 14 & 13 \end{bmatrix} \textbf{ and } TR_4 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \tag{3.4}$$

The first observation on the two matrices shown in Equation 3.4 is that since they both have been filled by successive positive integers 1 up to 16, they both represent $16 + 1 = 17$ gray levels and thereby they both need to be normalized by dividing their elements by 17. The difference between these two threshold matrices is, thought, that while threshold values in TR_3 are arranged to build a spiral form halftone dot, the values in TR_4 are arranged to build a line structure. This can be seen by following the successive integers starting from 1 and upwards to see how the halftone dot will grow in these two threshold matrices, see Figure 3.9.

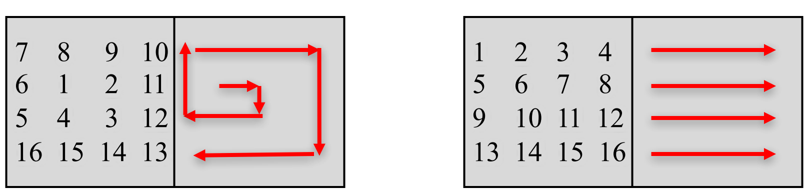
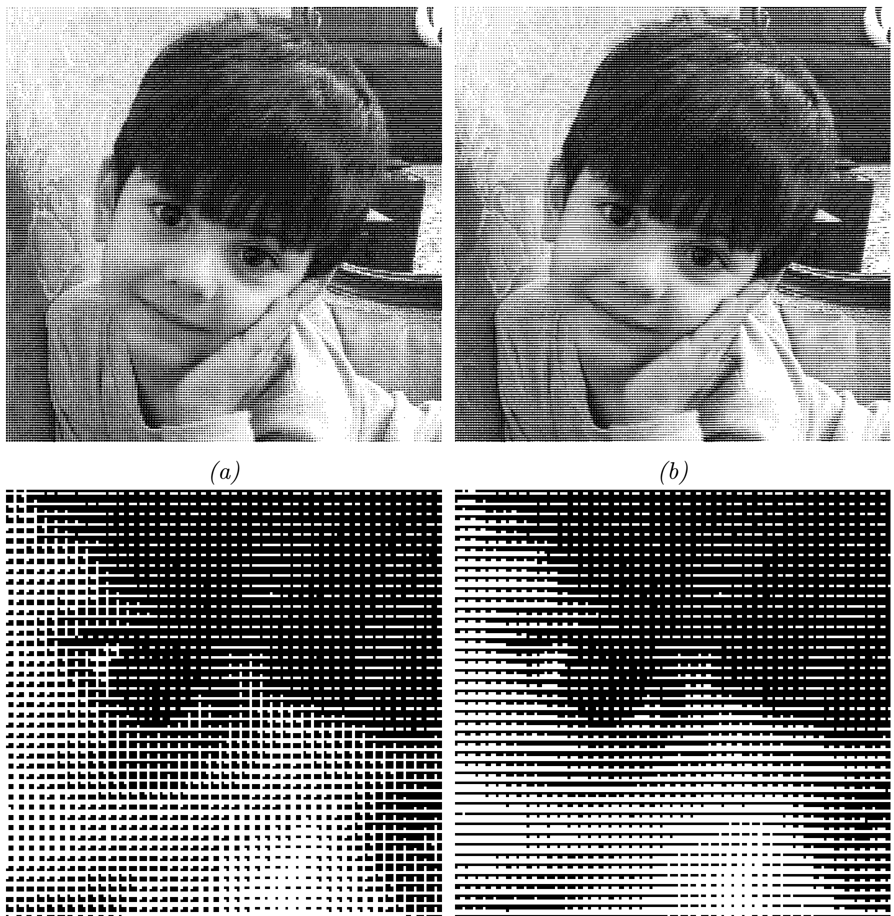


Figure 3.9: The threshold matrix to the left will result in spiral form halftone dots while the one to the right will result in a line structure.

Therefore, the threshold matrix TR_3 and TR_4 in Equation 3.4 will result in spiral and line halftone structures, respectively. The test image has been thresholded by these two matrices and the results and the enlargement of a part of them are shown in Figure 3.10. The spiral and line halftone structures can clearly be seen in Figure 3.10 (c) and (d), respectively.



(a) *(b)* *(c)* *(d)*

Figure 3.10: The test image is thresholded by: (a) TR_3 (spiral). (b) TR_4 (line). (c) Enlargement of a part of (a). (d) Enlargement of a part of (b).

Example 3.8 *As always, we assume the original image to be normalized between 0 and 1 and therefore the threshold matrices have to be normalized.*
a) *Write a 2×2 threshold matrix that represents 5 gray levels.*
b) *Write two different 4×4 threshold matrices, one representing 17 gray levels and the other one 9 gray levels.*
c) *Halftone the following image shown in Equation 3.5 by one of the threshold*

matrices in part (a) and (b).

$$image : \begin{bmatrix} 0.2 & 0.3 & 0.1 & 0.1 \\ 0.4 & 0.5 & 0.1 & 0.1 \\ 0.6 & 0.7 & 0.1 & 0.9 \\ 0.5 & 0.9 & 0.4 & 0.4 \end{bmatrix} \tag{3.5}$$

Solution (a): *An example of such a threshold matrix is* $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$*, which represents 5 gray levels and therefore need to be normalized by dividing its elements by 5, giving* $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} /5 = \begin{bmatrix} 0.2 & 0.4 \\ 0.6 & 0.8 \end{bmatrix}$. ***Notice*** *that, since in this assignment there is no mention about how to arrange the threshold values and only the number of gray levels is important, then there are many other* 2×2 *matrices that would work, for example* $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} /5$.

Solution (b): *Both threshold matrices* TR_3 *and* TR_4 *in Equation 3.4 represent 17 gray levels and should also be normalized by 17. There are of course many other* 4×4 *matrices that represent 17 gray levels. However, 17 is the maximum number of gray levels a* 4×4 *threshold matrix can represent. In order to make the threshold matrix represent fewer gray levels, some of the threshold values need to be repeated. For example,* $\begin{bmatrix} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \\ 5 & 6 & 1 & 2 \\ 7 & 8 & 3 & 4 \end{bmatrix}$ *represents* $8 + 1 = 9$ *(the largest threshold value plus 1) gray levels. This matrix has to be normalized by dividing its elements by 9 (the number of gray levels it represents). There are of course many other* 4×4 *threshold matrices that represent 9 gray levels, for example* $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} /9$.

Solution (c): *For this part, we choose the threshold matrix in part (a), i.e.* $\begin{bmatrix} 0.2 & 0.4 \\ 0.6 & 0.8 \end{bmatrix}$. *The original image is* 4×4 *and the threshold matrix is* 2×2*, therefore each* 2×2 *sub-image in the original image is compared element-by-element with this threshold matrix. If the pixel value at a specific position is greater than or equal to the threshold value at the corresponding position in the threshold matrix, then a 1 is set at the corresponding position in the halftoned image; otherwise a 0 is set there, see Figure 3.11. The halftoned image is therefore:* $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$.

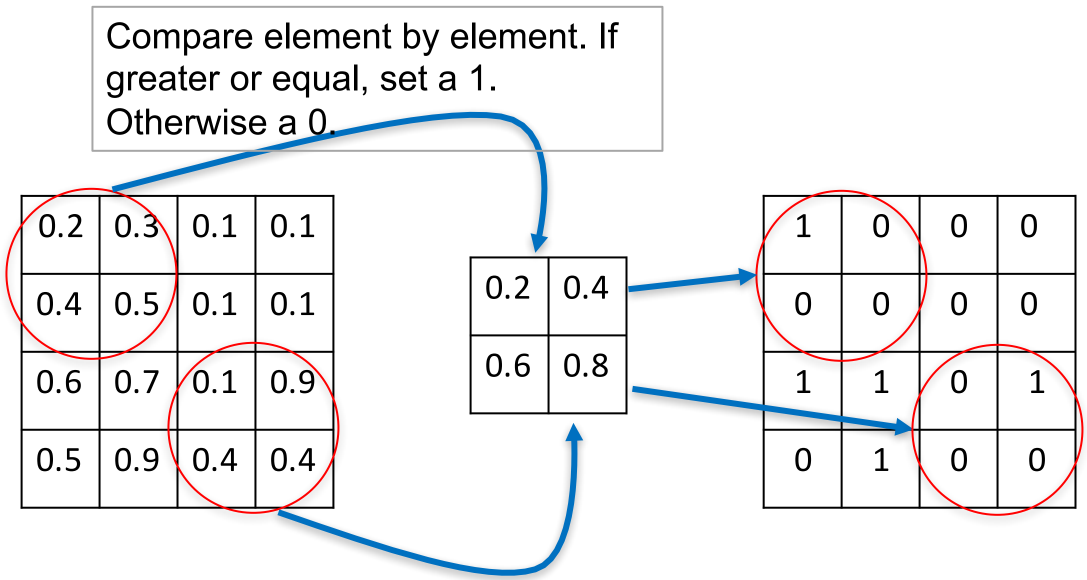


Figure 3.11: Example 3.8 (c) - How to threshold an image by a threshold matrix.

One of the well-known halftoning techniques that is based on threshold halftoning is Ordered dithering, described in the following subsection.

Ordered dithering

Ordered dithering is one of the basic and well-known halftoning techniques addressed in literature. As mentioned earlier, the design of the threshold matrix has a great impact on the characteristics of the final halftoned image. The ordered dithering techniques are divided into two parts, clustered dot and dispersed dot ordered dithering. In the clustered dot dithering, the threshold matrices are arranged in a way that the final halftone dot is a cluster of black microdots. In the dispersed dot dithering, the black microdots are dispersed. In Equation 3.6, two threshold matrices according to clustered dot and dispersed dot dithering are shown.

$$TR_{clus} = \begin{bmatrix} 7 & 8 & 9 & 10 \\ 6 & 1 & 2 & 11 \\ 5 & 4 & 3 & 12 \\ 16 & 15 & 14 & 13 \end{bmatrix} \textbf{ and } TR_{disp} = \begin{bmatrix} 1 & 9 & 3 & 11 \\ 13 & 5 & 15 & 7 \\ 4 & 12 & 2 & 10 \\ 16 & 8 & 14 & 6 \end{bmatrix} \tag{3.6}$$

By following the successive integers starting from 1 in TR_{clus}, it can be seen that the halftone dot will grow connected (coherently) and thereby resulting in a clustered halftone dot. In TR_{disp}, on the other hand, the dot will grow disconnected (separately), resulting in a dispersed halftone dot. Although both threshold matrices represent 17 gray levels, the structure of the resulting halftones will differ because of the two different arrangements. Figure 3.12 shows the test image halftoned by the two threshold matrices in Equation 3.6. Notice that both matrices are normalized by dividing their elements by 17, because both represent 17 gray levels. By looking at these halftones, specially the enlargements in Figure 3.12 (c) and (d), you can notice the difference between clustered and dispersed dot halftoning.

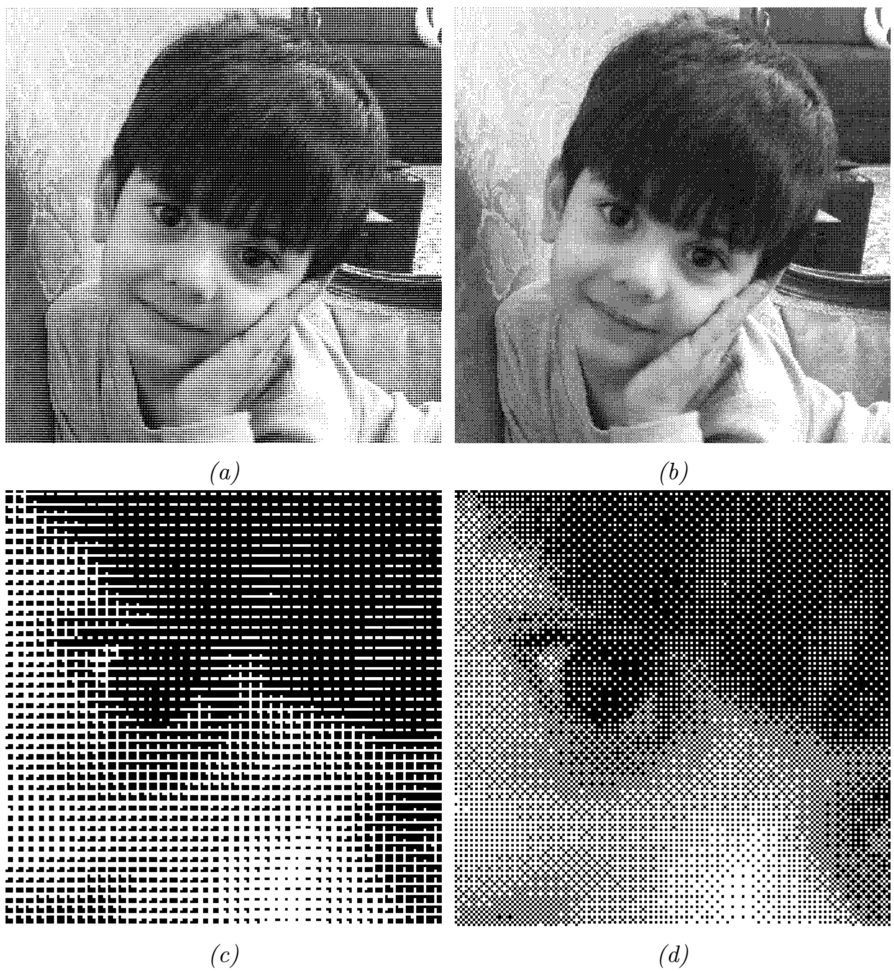


Figure 3.12: The test image is thresholded by: (a) TR_{clus} (clustered). (b) TR_{disp} (dispersed). (c) Enlargement of a part of (a). (d) Enlargement of a part of (b).

Figure 3.13 (a) and (b) show two other threshold matrices based on clustered and dispersed dot, respectively. As can be seen in the matrix in Figure 3.13 (a), the threshold values grow coherently and the halftone dot will thus be clustered. In the other threshold matrix, the threshold values grow separately, hence the halftone dot will be dispersed.

14	12	13	16	19	21	20	17
5	4	3	10	28	29	30	23
6	1	2	11	27	32	31	22
9	7	8	15	24	26	25	18
19	21	20	17	14	12	13	16
28	29	30	23	5	4	3	10
27	32	31	22	6	1	2	11
24	26	25	18	9	7	8	15

(a)

1	30	8	28	2	29	7	27
17	9	24	16	18	10	23	15
5	25	3	32	6	26	4	31
21	13	19	11	22	14	20	12
2	29	7	27	1	30	8	28
18	10	23	15	17	9	24	16
6	26	4	31	5	25	3	32
22	14	20	12	21	13	19	11

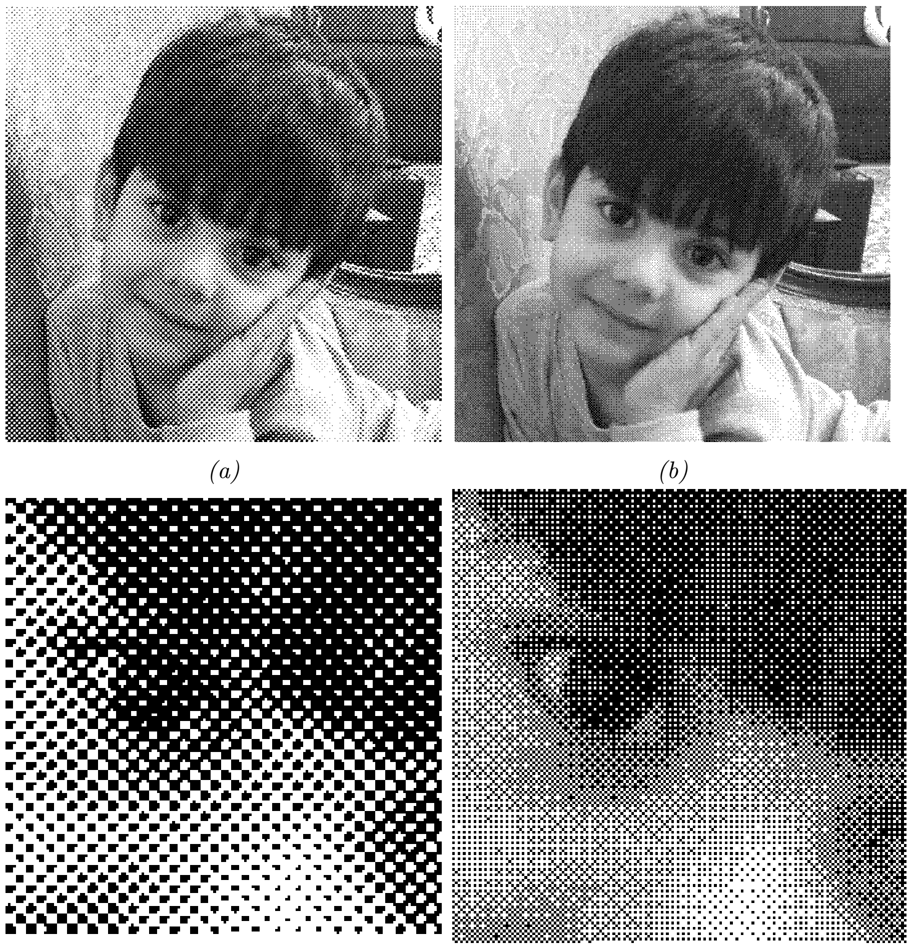
(b)

Figure 3.13: Two threshold matrices (a) Clustered dot (b) Dispersed dot.

Although both threshold matrices are 8×8 and could potentially represent 65 gray levels, they only represent 33 gray levels because the threshold values appear twice in each matrix and therefore the largest threshold value is 32, meaning $32 + 1 = 33$ gray levels. The reason for the threshold values being repeated is to make a raster angle of 45 degrees. This can be realized by noticing that the upper-left 4×4 sub-matrix is identical to the lower-right 4×4 sub-matrix in both threshold matrices in Figure 3.13. The upper-right 4×4 sub-matrix is also identical to the lower-left 4×4 in both. It is common to arrange the halftone dots aligned at 45 degrees because the human eye is less sensitive to structures at this angle. This will make the halftone dots less detectable by the human eye. Figure 3.14 shows the test image being halftoned using the threshold matrices shown in Figure 3.13. Enlargements of a part of these images are also shown in this figure. **Notice** that the threshold values in both threshold matrices shown in Figure 3.13 should be divided by 33. Notice also in both images that the halftone dots are aligned at 45 degrees, which is of course the result of arranging the threshold values in the threshold matrices in Figure 3.13 as described above.

3.2.2 Table halftoning

Another basic halftoning method is table halftoning. This method is what we in Section 3.1.1 called one of the most straightforward halftoning methods, in which different areas of the original image were replaced by halftone cells. In this method, as well, different areas (for example $m \times m$ pixels each) in the original image are replaced by their corresponding halftone tables/cells (for example $n \times n$ pixels/microdots each). If we want to follow the rule of thumb shown earlier in Equation 3.1, at least each 2×2 pixel area in the original image should be replaced by a table, i.e. $m \geq 2$. The size of the halftone table is decided by the ratio of the print resolution (dpi) to the screen frequency (lpi), as discussed in Section 3.1.2.



(a) (b) (c) (d)

Figure 3.14: The test image is thresholded by: (a) the threshold matrix in Figure 3.13 (a) (clustered). (b) the threshold matrix in Figure 3.13 (b) (dispersed). (c) Enlargement of a part of (a). (d) Enlargement of a part of (b).

Let us now explain table halftoning by Figure 3.15. For simplicity, assume that each 2×2 pixel area in the original image is going to be replaced by a 3×3 halftone table. The mean of the pixel values in each 2×2 pixel area in the original image is firstly calculated. Since the table is 3×3, only ten different gray tones can be represented by this table, see Figure 3.15 again. Among these ten possible candidates, the table with the closest mean to each 2×2 pixel area in the original image is placed at the corresponding position in the final halftoned image. In this illustration, only the table representation for the first (follow arrow A) and the last (follow arrow B) 2×2 pixel areas are shown. The 2×2 pixel area in the original image corresponding to arrow A has the mean of 0.6.

The 3×3 halftone tables, as seen in this figure, totally represent ten different gray tones, i.e. $0, 1/9, 2/9, 3/9, \cdots, 1$. Of these ten possible candidates, the table with 5 microdots filled has the closest mean to 0.6, because it represents $5/9 \approx 0.556$. Of the ten possible candidates, the table with 7 microdots filled has the closest mean to 0.8 (see arrow B), because it represents $7/9 \approx 0.778$. Corresponding calculations should be done for other 2×2 pixel areas in the original image. **Notice** that, in this example, the original image is 6×6 pixels, and since each 2×2 pixel area is replaced by a 3×3 halftone table, the final halftoned image is 9×9 pixels. Therefore, in table halftoning the halftoned image will not necessarily be the same size as the original image.

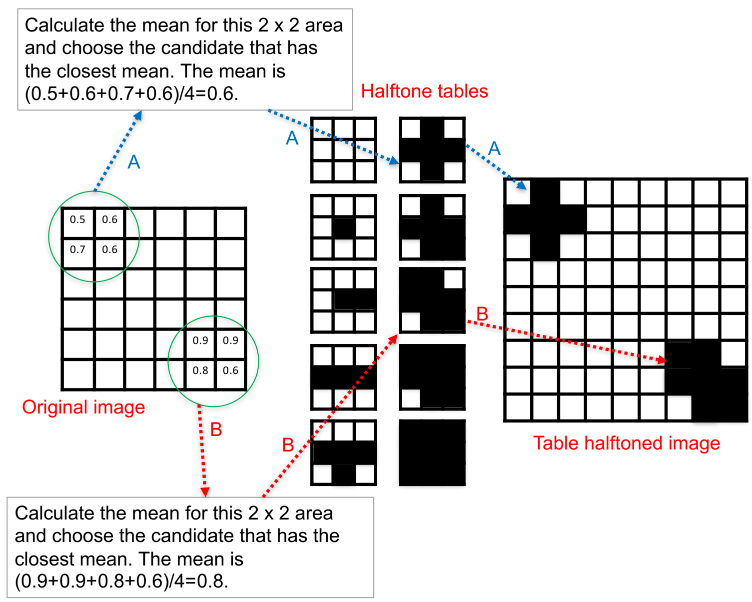
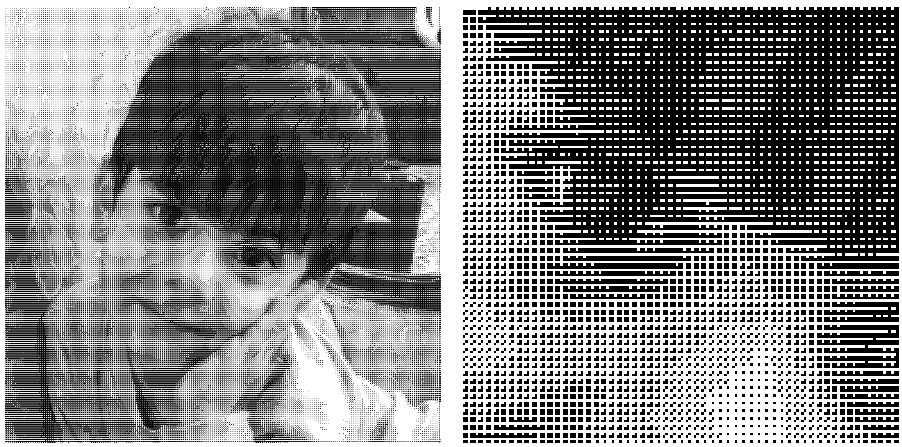


Figure 3.15: Table halftoning. Each 2×2 pixel area in the original image is replaced by a 3×3 halftone table (halftone cell). The table with the closest mean is chosen for each 2×2 region.

Observe that since the pixel values in the original image are scaled between 0 and 1, a black microdot in the halftone tables means 1 and an empty dot means 0. Figure 3.16 shows an image halftoned by a simple table halftoning. In this image 3×3 halftone tables have been used, hence, only 10 levels of gray are represented. The tables were arranged exactly as shown in Figure 3.15. An enlargement of a part of the image is also shown in this figure. Recall from Section 3.2.1 that the threshold matrices could be arranged so that they could represent either clustered or dispersed halftone dots. This is also valid for table halftoning. The arrangement of the black microdots in the halftone tables in Figure 3.15 are clustered, but they could also be arranged to be dispersed. For example, the two tables in Figure 3.17 both represent the same gray tone, i.e.

4/9, but the one to the left is arranged as a clustered and the one to the right as a dispersed halftone dot. The difference between the structures of halftones according to clustered dot and dispersed dot in table halftoning is the same as for threshold halftoning already discussed in Section 3.2.1 and illustrated in Figure 3.14.



(a) (b)

Figure 3.16: (a) The image has been halftoned by table halftoning. The tables are 3×3, hence 10 levels of gray. An enlargement is shown in (b).

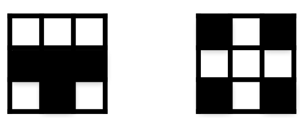
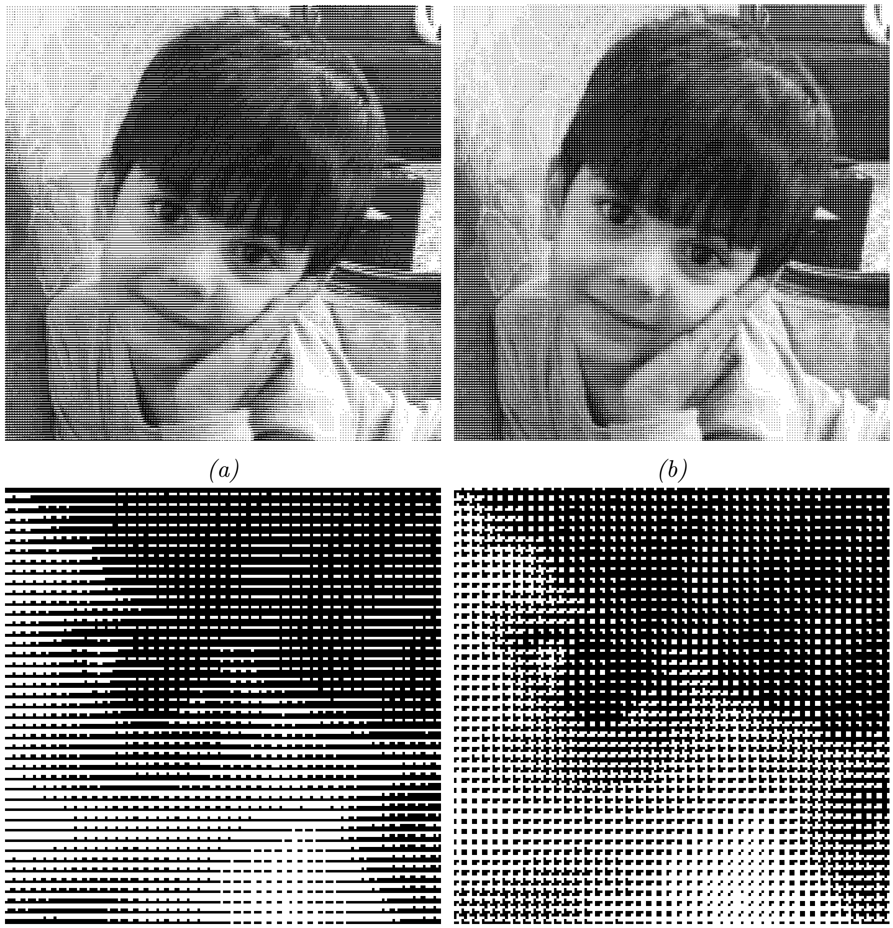


Figure 3.17: Two halftone tables both representing the gray tone $4/9$. (Left) The black microdots are clustered. (Right) The black microdots are dispersed.

Another important observation is that the shape of the halftone dots can also be changed by the arrangement of the black microdots in the halftone table. Since normally the halftone tables are much bigger than 3×3, it is not so difficult to arrange the black microdots to get a line, circular, elliptical or square halftone dot. To illustrate this, we arranged the black microdots in order to get line and circular halftone structures, see Figure 3.18 (a) and (b), respectively. In both images 4×4 halftone tables have been used, hence 17 levels of gray. Notice that since 4×4 halftone tables are still not large enough, the halftone dots in (b) don't appear as "circular" as you would expect.



(a) (b)

(c) (d)

Figure 3.18: Table halftoning using 4×4 halftone tables. (a) The black microdots are arranged to make line halftones. (b) The black microdots are arranged to make circular halftones. (c) An enlargement of a part of (a). d) An enlargement of a part of (b).

Example 3.9 *The following image is supposed to be halftoned by table halftoning and the halftoned image has to represent 10 gray levels.*

$$image: \begin{bmatrix} 0.2 & 0.3 & 0.1 & 0.1 \\ 0.4 & 0.5 & 0.1 & 0.1 \\ 0.6 & 0.7 & 0.1 & 0.9 \\ 0.5 & 0.9 & 0.4 & 0.4 \end{bmatrix} \tag{3.7}$$

a) *How large do the halftone tables (halftone cells) have to be?*

b) *Table halftone the image, where each 2×2 pixel area in the image is represented by a halftone cell/table.*

Solution (a): *Since the halftoning is supposed to represent 10 gray levels, then the halftone tables have to be at least* 3×3*.*

Solution (b): *As stated in the example, each* 2×2 *pixel area is going to be represented by a halftone table. The upper-left* 2×2 *pixel area has the mean of:* $\frac{0.2+0.3+0.4+0.5}{4} = 0.35$*. Since the halftone tables are* 3×3*, gray levels* $0, 1/9, 2/9, 3/9, \cdots, 1$ *could be represented. Of these 10 possibilities, a table with 3 black microdots giving the mean* $\frac{3}{9} \approx 0.33$ *represents the closest gray tone. Therefore, in the halftoned image the upper-left* 3×3 *area has to include 3 black microdots (or 1), see Equation 3.8. The upper-right* 2×2 *pixel area has the mean of:* $\frac{0.1+0.1+0.1+0.1}{4} = 0.1$*. Of the 10 possibilities, a table with 1 black microdot giving the mean* $\frac{1}{9} \approx 0.11$ *represents the closest gray tone. Therefore, in the halftoned image the upper-right* 3×3 *area has to include 1 black microdot. The lower-left* 2×2 *pixel area has the mean of:* $\frac{0.6+0.7+0.5+0.9}{4} = 0.675$*. Of the 10 possibilities, a table with 6 black microdots giving the mean* $\frac{6}{9} \approx 0.67$ *represents the closest gray tone. Therefore, in the halftoned image the lower-left* 3×3 *area has to include 6 black microdots. Finally, the lower-right* 2×2 *pixel area has the mean of:* $\frac{0.1+0.9+0.4+0.4}{4} = 0.45$*. Of the 10 possibilities, a table with 4 black microdots giving the mean* $\frac{4}{9} \approx 0.44$ *represents the closest gray tone. Therefore, in the halftoned image the lower-right* 3×3 *area has to include 4 black microdots. The final halftoned image is thus:*

$$\textit{Halftoned image}: \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.8}$$

Notice *first that, although the original image is* 4×4*, the final halftoned image is* 6×6*. The reason is simply the fact that in this example each* 2×2 *sub-area in the original image has been replaced by a* 3×3 *halftone table.* ***Notice*** *further that, the halftone tables in the halftoned image in Equation 3.8 make clustered halftone dots. However, since in this example there was no mention about the arrangement of the microdots, any other arrangement would be a correct answer as long as the tables include the correct number of black microdots according to the above calculations.*

3.2.3 Error diffusion

As seen in Section 3.2.1 and 3.2.2, threshold halftoning and table halftoning usually result in periodic halftone structures, organized by the utilized threshold matrix or halftone tables. In the mid 1970s, Floyd and Steinberg, invented a halftoning method, called error diffusion, resulting in non-periodic and dispersed halftone structures.

The process of error diffusion starts from a pixel in the original image and goes through all its pixels. Commonly, it starts from the pixel up to the left and goes through the pixels left-to-right and upside-down. Assume that the original image and the halftoned result are denoted by g and b, respectively. There is also an error filter involved in the technique, that has a great impact on the final result. Let us explain the error diffusion technique, assuming the error filter to be that in Figure 3.19 (a), which is the filter the inventors of the error diffusion algorithm, i.e. Floyd and Steinberg, proposed:

- At the current position (i, j), if the pixel value is greater than or equal to a threshold, set a 1 at the same position (i, j) in the halftoned image, otherwise a 0. In the non-modified error diffusion, the threshold is usually equal to 0.5. This means:

$$b(i,j) = \begin{cases} 1 & \textbf{if g(i,j)} \geq \textbf{0.5} \\ 0 & \textbf{if g(i,j)} < \textbf{0.5} \end{cases} . \tag{3.9}$$

- Calculate the error, $e(i, j) = g(i, j) - b(i, j)$.

- Diffuse the error by using the error filter. This means that the values of the pixels in the original image that have not been processed yet are going to be influenced and changed by the error. The error filter in Figure 3.19 (a) means that the error is multiplied (weighted) by the weight 7/16 and added to the pixel value of the pixel to the right of the current position. The error is also multiplied (weighted) by the weight 1/16 and added to the pixel value of the pixel to the right-down of the current position, and the same for the other two weights. This is summarized in Equation 3.10. Notice that, the position (i, j) in Equation 3.10 refers to row i and column j and position $(1, 1)$ refers to the up-left pixel in an image, as it is used to address the positions in an image/matrix in Matlab.

$$\begin{cases} g_{new}(i,j+1) = g(i,j+1) + 7/16 \cdot e(i,j) \\ g_{new}(i+1,j+1) = g(i+1,j+1) + 1/16 \cdot e(i,j) \\ g_{new}(i+1,j) = g(i+1,j) + 5/16 \cdot e(i,j) \\ g_{new}(i+1,j-1) = g(i+1,j-1) + 3/16 \cdot e(i,j) \end{cases} \tag{3.10}$$

- When all calculations are done for the current position, go to the next pixel position and do exactly the same procedure.

- The process is terminated and the final halftoned image is obtained when all pixels have been processed.

(*1/16)

	x	7
3	5	1

(a)

(*1/48)

		x	7	5
3	5	7	5	3
1	3	5	3	1

(b)

(*1/42)

		x	8	4
2	4	8	4	2
1	2	4	2	1

(c)

Figure 3.19: Three error filters. (a) Floyd and Steinberg. (b) Jarvis, Judice and Ninke. (c) Stucki.

As mentioned before, the structure of the halftones produced by error diffusion depends in a great deal on the error filter being used. In Figure 3.19, three different error filters introduced in literature are shown. **Notice** that the sum of the weights in the error filters is always equal to one and that is the reason the filters in Figure 3.19 (a), (b) and (c) are multiplied by 1/16, 1/48 and 1/42, respectively. Notice also that $\times$ in these filters shows the position of the current pixel being processed.
Before illustrating the halftone results using different error filters, let us thoroughly explain the process of error diffusion by an example.

Example 3.10 *Halftone the below image (g) by error diffusion technique using the below error filter (e).*

$$\text{Image } g\text{: } \begin{bmatrix} 0.3 & 0.4 \\ 0.3 & 0.6 \end{bmatrix} \text{ and error filter } e\text{: } \begin{bmatrix} \times & 0.8 \\ 0.2 & \end{bmatrix}$$

Solution: *Notice that, for the sake of simplicity, the error filter in this example only includes two weights. The sum of the weights is equal to 1, which should always be the case. The error filter also indicates that the error in each pixel is multiplied by 0.8 and added to the pixel value at the pixel to the right and multiplied by 0.2 and added to the pixel value at the pixel under. In this course book, as explained before, we always start with the pixel up to the left, having the position* $(1,1)$. *We call the original image g and the final halftoned image b.*

- *position (1,1):*
  - *The pixel value at this position is 0.3, i.e.* $g(1,1) = 0.3$. *Since this is less than 0.5, then* $\mathbf{b(1,1) = 0}$.
  - *Calculate the error: the error is* $g(1,1) - b(1,1) = 0.3 - 0 = 0.3$.
  - *Diffuse the error: The error, which is* 0.3, *is multiplied by* 0.8 *(the value in the error filter at the position to the* ***right*** *of* $\times$*) and added to the pixel to the* ***right****. This means* $g_{new}(1,2) = g(1,2) + 0.3 \cdot 0.8 = 0.4 + 0.24 = 0.64$. *The same way, the error needs to be diffused to the pixel* ***under*** *(this time by the weight* 0.2*, because this is the*

weight at the position ***under*** $\times$ *in the error filter), i.e.* $g_{new}(2,1) = g(2,1) + 0.3 \cdot 0.2 = 0.3 + 0.06 = 0.36$.

  - *Before going to the next position, notice that the original image is modified and* $g(1,2)$ *and* $g(2,1)$ *hold new values, which are 0.64 and 0.36, respectively.*

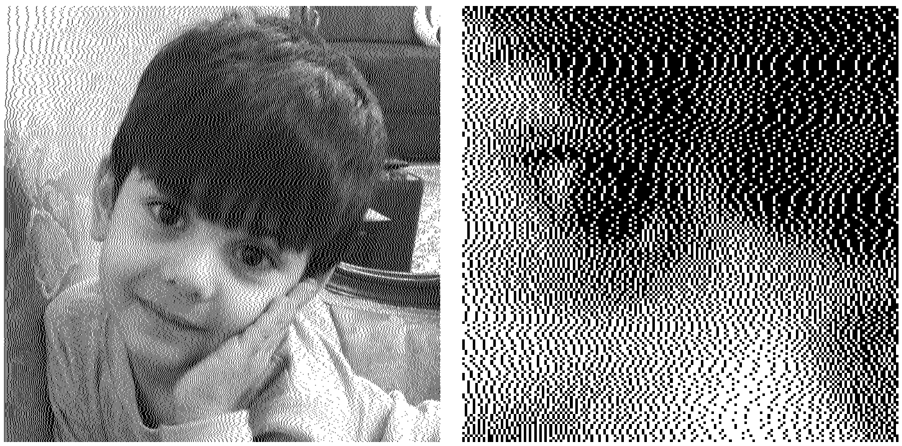
- *position (1,2):*
  - *The new pixel value at this position after the error was diffused in the previous step is 0.64, i.e.* $g_{new}(1,2) = 0.64$. *Since this is greater than 0.5, then* $\mathbf{b(1,2) = 1}$.
  - *Calculate the error: the error is* $g_{new}(1,2) - b(1,2) = 0.64 - 1 = -0.36$.
  - *Diffuse the error: The error, which is* -0.36, *should be multiplied by* 0.8 *and added to the pixel to the right. Since there is no pixel to the right of this position, this could be discarded. The error needs to be diffused to the pixel under, i.e.* $g_{new}(2,2) = g(2,2) + (-0.36) \cdot 0.2 = 0.6 - 0.072 = 0.528$.
  - *Before going to the next position, notice that the original image is modified and* $g(2,2)$ *holds a new value, which is 0.528.*

- *position (2,1):*
  - *The new pixel value at this position after the error was diffused in the previous steps is 0.36, i.e.* $g_{new}(2,1) = 0.36$. *Since this is less than 0.5, then* $\mathbf{b(2,1) = 0}$.
  - *Calculate the error: the error is* $g_{new}(2,1) - b(2,1) = 0.36 - 0 = 0.36$.
  - *Diffuse the error: The error, which is* 0.36, *should be multiplied by* 0.8 *and added to the pixel to the right, giving* $g_{new2}(2,2) = g_{new}(2,2) + 0.36 \cdot 0.8 = 0.528 + 0.288 = 0.816$. *Since there is no pixel under this pixel, then this could be discarded.*
  - *Before going to the next position, notice that the original image is modified and* $g(2,2)$ *holds a newer value, which is 0.816.*

- *position (2,2): This is the last pixel being processed. The new pixel value at this position after the error was diffused in the previous steps is* $g_{new2}(2,2) = 0.816$, *which is greater than 0.5, meaning* $\mathbf{b(2,2) = 1}$. *No error needs to be calculated because this was the last pixel.*

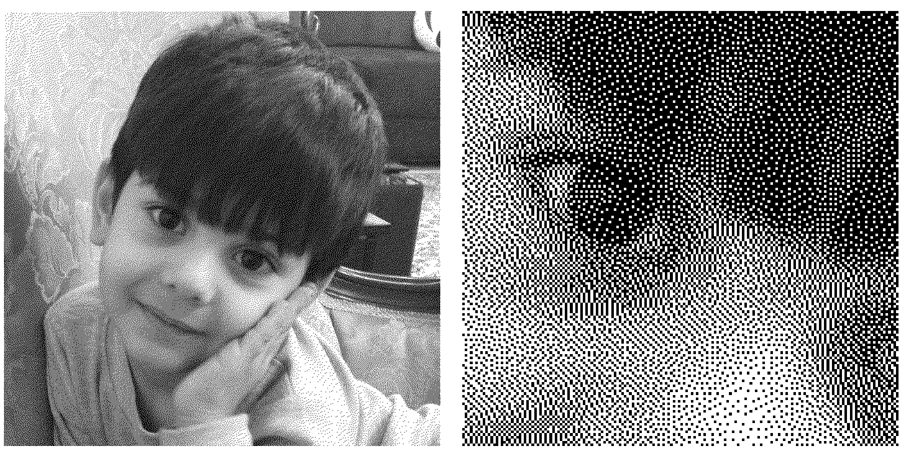
The final halftoned image is therefore: b: $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$.

In order to illustrate how the error filter used in the error diffusion algorithm impacts the halftone structure, the test image is halftoned by error diffusion using an error filter to diffuse the error to only one pixel to the right of the currently processed position, see Figure 3.20. It means that the error filter only contains one weight to the right. This weight is of course equal to 1, because the sum of the weights has to be 1. Since the error has been diffused only in one direction and only to one pixel, the result is not satisfying and the visible vertical structures are observed and easily detected by the human eye.



(a) *(b)*

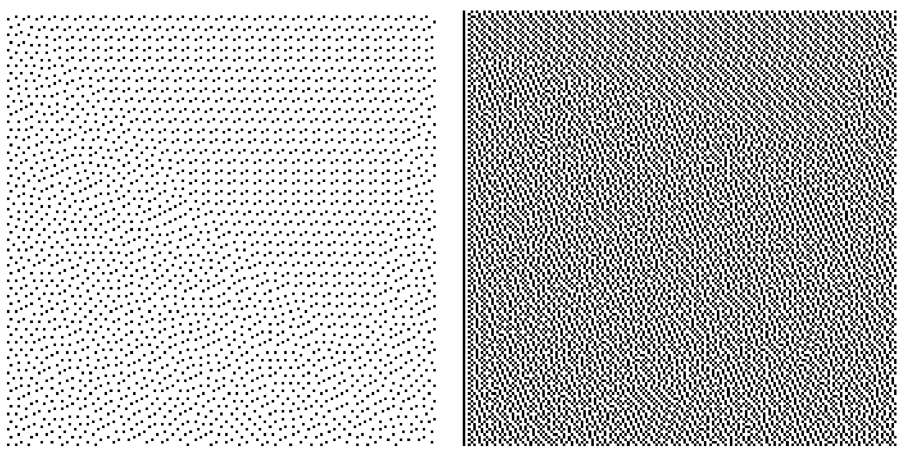
Figure 3.20: (a) The image has been halftoned by error diffusion. The error filter contains one weight, diffusing the error to the pixel to the right. An enlargement is shown in (b).



(a) *(b)*

Figure 3.21: (a) The image has been halftoned by error diffusion. The error filter is the Floyd and Steinberg filter shown in Figure 3.19 (a). An enlargement is shown in (b).

Figure 3.21 shows the test image halftoned by error diffusion using the error filter proposed by Floyd and Steinberg, shown in Figure 3.19 (a). The disturbing structures are much less evident in this figure and the details are reproduced very well. Although the non-modified error diffusion is quite simple and results in fairly good images, it has a few shortcomings. One of them is the directional hysteresis, which is best seen in the highlights and shadows of an image. Another problem is the correlated artifacts, best seen in the mid-tones of the image. Although both artifacts are visible in Figures 3.20 and 3.21, in order to illustrate these artifacts more clearly, two constant images holding 0.1 (highlight) and 0.4 (mid-tone) pixel values have been halftoned by error diffusion using Floyd and Steinberg error filter, see Figure 3.22. Notice that, in this book an achromatic constant image, is defined to be an image holding the same value at its all pixels. The directional hysteresis and correlated artifacts are clearly visible in Figure 3.22 (a) and (b), respectively. These problems associated with error diffusion are addressed and discussed and simple modifications are proposed in the following sub-section.

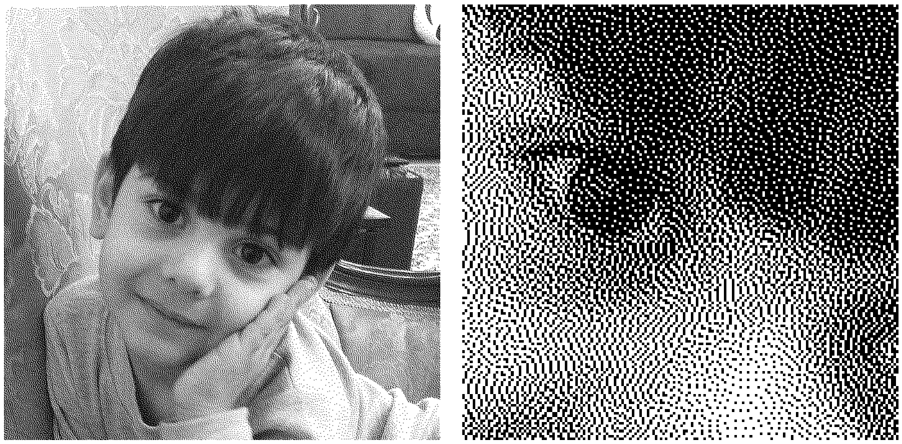


(a) (b)

Figure 3.22: Two constant images have been halftoned by error diffusion using Floyd and Steinberg error filter: (a) All pixels in the original image hold 0.1, illustrating the directional hysteresis. (b) All pixels in the original image hold 0.4, illustrating the correlated artifacts.

Modified error diffusion

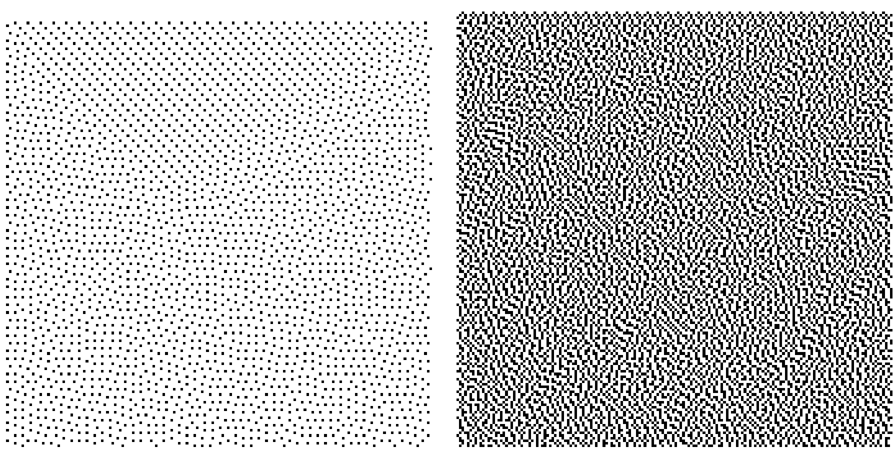
Due to the simplicity and efficiency of error diffusion, there have been many studies on this method during the recent decades and many different modifications have been suggested to reduce and eliminate the directional hysteresis and correlated artifacts. The first attempt would be to use larger error filters than the four-weight filter proposed by Floyd and Steinberg. Figure 3.23 shows the test image being halftoned by error diffusion using Jarvis et al. error filter shown in Figure 3.19 (b).



(a) (b)

Figure 3.23: (a) The image has been halftoned by error diffusion. The error filter is Jarvis et al. filter shown in Figure 3.19 (b). An enlargement is shown in (b).

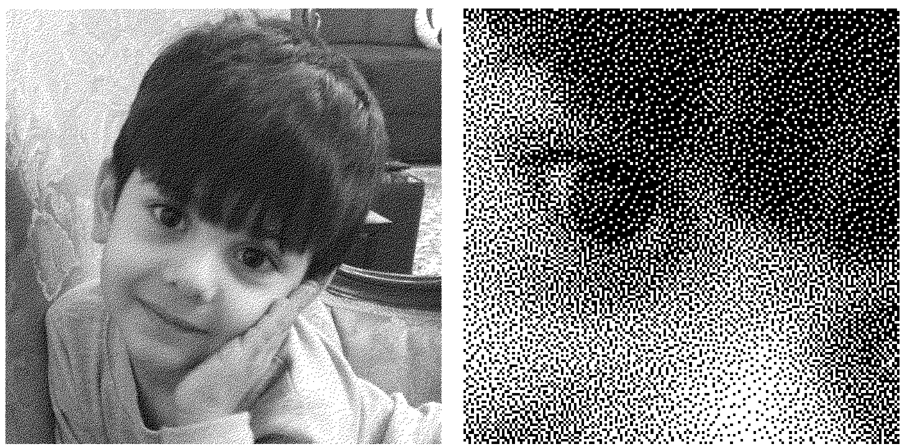
In order to better illustrate whether or not a larger error filter can eliminate the artifacts associated with error diffusion, the two constant images holding pixel values 0.1 and 0.4 are halftoned by error diffusion using Jarvis et al. filter, see Figure 3.24. As can be seen in Figure 3.24 (a), the directional hysteresis is somewhat reduced but still not satisfyingly eliminated.



(a) (b)

Figure 3.24: Two constant images have been halftoned by error diffusion using Jarvis et al. error filter: (a) All pixels in the original image hold 0.1, illustrating the directional hysteresis. (b) All pixels in the original image hold 0.4, illustrating the correlated artifacts.

The correlated artifacts are greatly reduced, as seen in Figure 3.24 (b), but the halftone structures look coarser than those in Figure 3.22 (b). Although using a bigger filter results in sharper images (seen in Figure 3.23) and reduces the artifacts, it is not a very satisfactory modification for error diffusion algorithm. It is worth mentioning that, It has been shown in literature that error diffusion techniques, in general, have a tendency of high-pass filtering (edge enhancing) the original image.
Another simple modification to reduce the mentioned artifacts is to add some noise to the original image before halftoning. Alternatively, one can change the threshold value in the error diffusion algorithm from being fixed at 0.5 to a random number between 0 and 1. Figure 3.25 shows the test image being halftoned by error diffusion with Floyd and Steinberg filter using a random threshold varying between 0.25 and 0.75. As can be seen, the result gives a noisy impression but the correlated artifacts are much less visible.

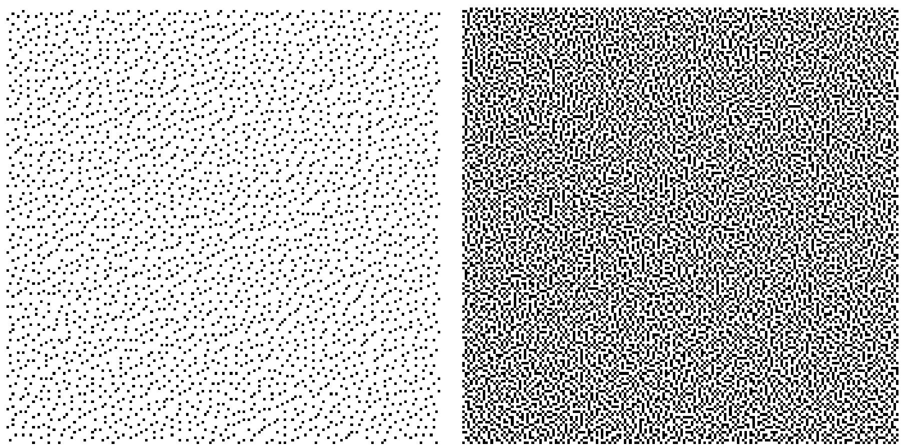


(a) *(b)*

Figure 3.25: (a) The image has been halftoned by error diffusion. The error filter is Floyd and Steinberg filter shown in Figure 3.19 (a) and the threshold value is a random number between 0.25 and 0.75. (b). An enlargement is shown in (b).

This modification has also been applied to the two constant images holding pixel values 0.1 and 0.4, see Figure 3.26. These images show that this simple modification slightly reduces the directional hysteresis, compare Figure 3.26 (a) and Figure 3.22 (a). This modification greatly reduces the correlated artifacts, compare Figure 3.26 (b) and Figure 3.22 (b). However, as mentioned above, this simple modification results in more noisy halftoned images, because of the nature of the randomness added to the error diffusion algorithm.
It is worth pointing out again that error diffusion is still one of the most popular halftoning techniques. Although the non-modified error diffusion is hardly used in practical applications, but the idea behind this technique and improved versions of the original error diffusion are, without any doubt, used in many

commercial halftoning software packages.

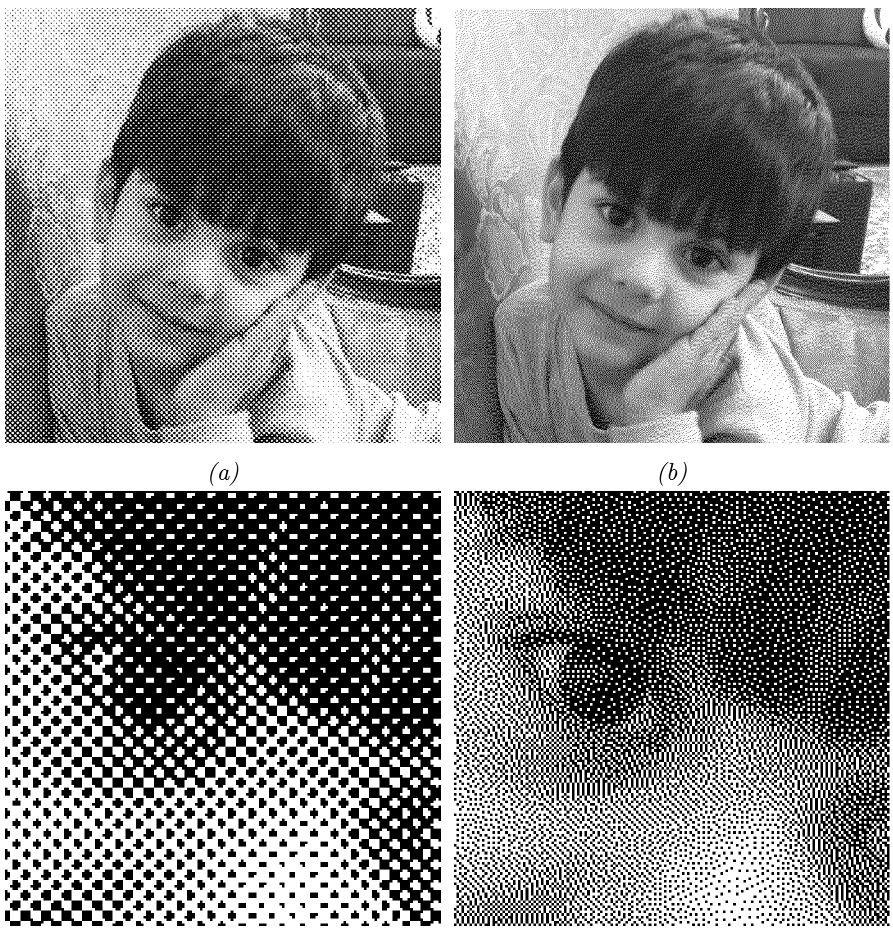


(a) *(b)*

Figure 3.26: Two constant images have been halftoned by error diffusion using Floyd and Steinberg filter with random threshold values between 0.25 and 0.75. (a) All pixels in the original image hold 0.1, illustrating the directional hysteresis. (b) All pixels in the original image hold 0.4, illustrating the correlated artifacts.

3.3 AM and FM halftoning

The halftoning methods are usually divided into two main categories, namely AM (Amplitude Modulated) and FM (Frequency Modulated). In the AM methods, the size of the halftone dots vary, while their spatial frequency is constant. This means that the size of the halftone dot becomes bigger as the tone gets darker. In the FM methods, on the other hand, the dot size is constant while the frequency (the number of microdots) varies. Notice that, this type of FM halftoning mentioned here is actually the first generation FM halftoning, but in this book when we use the term FM, we mean the first generation FM halftoning. There is also another type of FM halftoning called second generation FM halftoning, which will be discussed in Section 4.3. Another point worth mentioning here is that the terms AM and FM halftoning are sometimes incorrectly replaced by conventional and stochastic halftoning, respectively. However, in this book we rather use the terms AM and FM halftoning as defined above. According to this definition, a conventional halftoning, where we use threshold matrices or halftone tables/cells, could be either AM or FM. When the halftone dots are clustered, then they produce AM halftones and when they are dispersed they produce FM halftones. However, using periodic but dispersed halftone dots is not common in printing. Therefore, when referring to FM halftones in this book, we mean non-periodic and dispersed halftones, such as error diffusion. To summarize this, in this book, AM halftones refer to periodic and clustered halftone dots, such as threshold and table halftoning using clustered halftone dots, and (first generation) FM halftoning refers to non periodic dispersed halftone structure, such as error diffusion. Hence, the term *lpi* (the number of halftone cells/dots per inch) is only defined and used for AM halftoning. Since in FM halftoning there is no periodic halftone structure, *lpi* loses its meaning and the only term that is used is the print resolution *dpi*, which is the number of microdots per inch. Figure 3.27 illustrates two images halftoned by AM and FM methods. In the image in Figure 3.27 (a), an AM method at 50 *lpi* and 300 *dpi* has been used.



(a) (b)

(c) (d)

Figure 3.27: The test image is halftoned by AM and FM halftoning techniques. In (a) the screen frequency is 50 lpi and the print resolution is 300 dpi. In (b), the error diffusion method is used and the print resolution is 300 dpi. (c) and (d) are enlargements of a part of (a) and (b), respectively, both printed at 75 dpi.

The image in Figure 3.27 (b) is halftoned by error diffusion using Floyd and Steinberg error filter, described in Section 3.2.3, printed at 300 *dpi*. Images shown in Figures 3.27 (c) and (d) are enlargements of a part of images shown in (a) and (b), respectively, printed at 75 *dpi*. As can be seen, the AM-halftoned image has a periodical structure and the FM halftone doesn't possess any ordered structures. When the tones become darker, in the AM-halftoned image the size of the dots becomes larger. In the FM-halftoned image, on the other hand, the size of the smallest microdots, decided by the print resolution (*dpi*), are constant and when the tones become darker the number of these microdots increases.
AM halftoning has for long been the most used halftoning technique in the printing industry. The most important reason has been the inability of the printing devices to produce the small single microdots. However, since the beginning of the 1990s, when FM halftoning was used as an alternative to AM for low-cost inkjet printers, FM has started competing with AM technologies. These days, they are widely used not only in inkjet printers but also in the newspaper prints. The AM and FM halftoning have, however, their own advantages and disadvantages. The FM techniques are superior when it comes to reproducing the details, especially when the screen frequency cannot be as high as one would like because of the mechanical limitations of the print press or paper type. The AM methods, however, are better for the areas where the tones vary slowly. The FM methods generally give a "noisy" impression in these regions. The advantages and drawbacks of these methods have made researchers carry out studies to combine these two methods for different applications and purposes. This will be discussed in detail in Section 4.2.

3.4 Dot gain

The printed dots normally appear bigger than they are in the digital bitmap, i.e. the digital halftoned image. This is partly because that the dots become physically bigger due to, for example, the ink spread and other distortions produced by the printer or print press. This is what is called the physical (mechanical) dot gain. Another reason why the printed dots appear larger than their real physical size is the diffusion of the light in the paper or substrate. This is called the optical dot gain. Physical and optical dot gain together contribute to the total dot gain, which makes the printed image look darker than what it was intended to be. Therefore, in order to have a control over the print, dot gain needs to be compensated for, which is normally done before halftoning. In this section, we briefly discuss dot gain and describe a simple dot gain model and how an image can be compensated for dot gain before being halftoned and printed.

3.4.1 Physical dot gain

Due to many different factors, the printers or the print presses are not able to produce the dots exactly the same size as they are in the digital bitmap.

Normally, they become larger when printed. One of the most important reasons is that the viscosity of the ink is not high enough to avoid some spreading. The pressure from the print cylinder and the paper quality also play significant roles.

3.4.2 Optical dot gain

Unlike the physical dot gain, the optical dot gain, also called the Yule-Nielsen effect, has actually nothing to do with the mechanical processes of printing. It is simply due to the light scattering in the paper or the substrate. Some of the photons that enter the non-inked areas of the print are scattered in the paper and absorbed by the ink on their way back from the paper. There are also photons that enter the inked area but exit the print surface from the non-inked area. These two photon paths are the reason for optical dot gain. Figure 3.28 shows a simple illustration of four different paths a photon can travel when it enters the surface of a print that is partly covered by ink and partly blank. Arrow C and D illustrate the reason for optical dot gain. Arrow C, for example, enters the non-inked area but when exiting the print surface it has been affected by the color of the ink. Therefore, the non-inked point where arrow C enters the paper will be influenced by the color of the ink. Hence, optical dot gain creates an illusion of the color of the ink (a shadow) around the printed ink making the print look more saturated (darker).

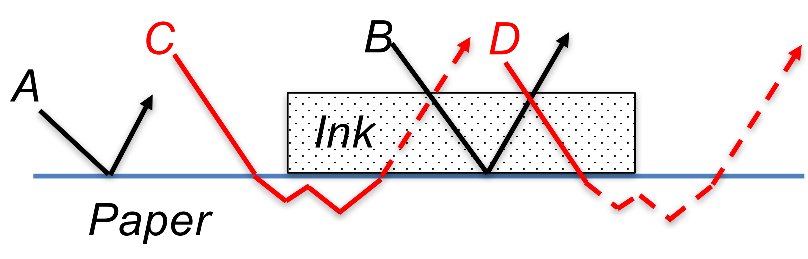


Figure 3.28: Four different paths that a photon can take when entering a print. Arrow C and D illustrate the reason for optical dot gain.

3.4.3 Dot gain and halftoning

Due to the different arrangements of the microdots in AM and FM halftoning, the impact of dot gain on these techniques is also different. It is quite obvious that the effect of the gain of each single microdot has greater impact on the total dot gain when the dots are dispersed. This fact is actually valid for both physical and optical dot gain. Since the microdots are clustered in AM halftoning and dispersed in FM halftoning, FM halftoning suffers more from dot gain than AM in the same printing conditions. Therefore, the image shown in Figure 3.27 (b), which is FM halftoned, looks darker than the image in Figure 3.27 (a), which is AM halftoned, when printed, although the print resolution is 300 *dpi* in both cases.

3.4.4 Dot gain models

Because of dot gain's impact on the print quality, many studies have been carried out in the last decades to investigate dot gain to find an appropriate and useful model for it. If the optical dot gain is neglected, only arrows A and B in Figure 3.28 describe the paths a photon can take. Hence, the total reflectance of a halftone print can easily be approximated by Murray-Davies formula:

$$R_{tot} = a \cdot R_{ink} + (1 - a) \cdot R_{pap}, \tag{3.11}$$

where R_{tot}, R_{ink} and R_{pap} denote the total/average reflectance of the halftone print, the reflectance of the full tone ink and the reflectance of the bare paper, respectively. a denotes the fractional area of the print surface covered by the ink. Note that a in Equation 3.11 corresponds to the real physical dot size in print, i.e. after the physical dot gain. If the optical dot gain is not negligible, Equation 3.11 is not valid anymore. One of the first models for optical dot gain was introduced by Yule and Nielsen. Their model is actually an extension to the Murray-Davies model and is as follows,

$$R_{tot}^{1/n} = a \cdot R_{ink}^{1/n} + (1 - a) \cdot R_{pap}^{1/n}, \tag{3.12}$$

where R_{tot}, R_{ink}, R_{pap} and a are defined as before and the parameter n is a fitting factor, which is determined by experiment. Something worth mentioning is that although Equation 3.12 is used to only model the optical dot gain, it is sometimes utilized to model both physical and optical dot gain.

3.4.5 Dot gain compensation

Since due to dot gain each dot appears bigger than its correspondence in the bitmap, the printed image will be darker than it actually is, see Figure 3.29 (a). In order to make the printed halftoned image represent the correct gray-tones, as they are in the original, the effect of dot gain has to be compensated for. This compensation is normally performed on the original image prior to halftoning by using an experimental dot gain curve or alternatively an appropriate dot gain model. The dot gain curve is normally found by experiments. Let us explain this experiment by explaining how to compensate the image shown in Figure 3.29 for dot gain.

(a) (b)

Figure 3.29: The image is halftoned by a FM method (here: error diffusion). (a) the original image has not been compensated for dot gain. (b) the original image has been compensated for dot gain prior to halftoning.

In order to carry out the experiment, a number of halftone patches with different nominal dot coverage were created. The patches were then halftoned by a FM method (in this case error diffusion) and printed out at 300 *dpi* using a laser printer (HP LaserJet 4050N). By using a spectrophotometer, we measured the spectra of all these patches. The Murray-Davies formula (Equation 3.11) was then used to find the effective dot coverage,

$$a_{eff} = \frac{R_{pap} - R_{mea}}{R_{pap} - R_{ink}}, \tag{3.13}$$

where R_{ink} and R_{pap} are defined as before and R_{mea} denotes the measured spectra. By using for example the least square errors method on Equation 3.13, one can calculate the effective dot coverage, a_{eff}. Notice that since the spectrophotometer is an optical equipment, both physical and optical dot gain are included in our measurement results and consequently a_{eff} represents both of them. Figure 3.30 shows a_{eff} versus the commanded dot coverage for the laser printer mentioned above and the utilized FM halftoning method at 300 *dpi*. The straight line in this figure shows the linear output response, i.e. when there is no dot gain. Now, by using the experimental curve shown in Figure 3.30, one can compensate the original image for dot gain prior to halftoning to achieve linear output response after print. Assume that we want to print a halftone patch with a gray tone corresponding to 30% (or 0.3) coverage. From the curve, by following the arrows in Figure 3.30, we can see that in order to have 30% coverage in print we need to have about 14% coverage in the digital halftone patch. The original image can be compensated for dot gain using the same strategy for each pixel value. The compensated image can then be halftoned with the halftoning method for which the measurements were carried out. In the image shown in Figure 3.29 (b), the original image was first compensated for dot gain using the curve in Figure 3.30 and then FM-halftoned. Notice that in order for the printed image to have the correct gray tone, the halftoned image should be printed with the same laser printer at 300 *dpi*. Furthermore, It has to be mentioned that since the dot gain compensation is a non-linear transformation the compensated image, and consequently its halftoned version, might lose some details of the original image.

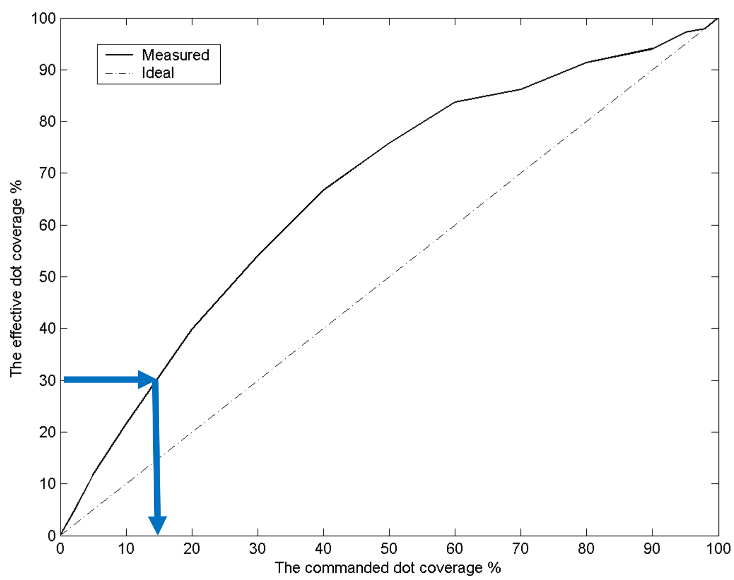


Figure 3.30: The effective dot coverage vs. the commanded dot coverage for a laser printer and FM halftoning at 300 dpi. The linear output response occurs when there is no dot gain (the straight line).

Some authors have suggested that considering the effect of dot gain within the halftoning process, if the halftoning method allows for that, can increase the quality of the final printed image compared to the case where the compensation is performed prior to halftoning. Something else worth mentioning here is that some studies have shown that although the optical dot gain is considered as an un-wanted phenomenon it can actually increase the color gamut (the set of the colors that can be reproduced by a device, see Section 7.1 for more detail) in print.

3.5 Exercise set

3.1. A 25×25 cm^2 color photograph has been scanned at $ppi = 200$. Assume an inch to be 2.5 cm.

a) How much memory is needed to store this digital image in raw-format?

b) The digital image above is AM-halftoned and the printed image is supposed to be 50×50 cm^2. What *lpi* should be used to obtain the highest possible quality?

c) The digital image above is FM-halftoned (for example by error diffusion) and the printed image is supposed to be 50×50 cm^2. What *dpi* should be used to obtain the highest possible quality?

3.2. a) Lars has taken a picture with his digital camera and it is 2592×1944 pixels. According to his friend, this digital image needs $2592 \cdot 1944 \cdot 3 \approx 14.42$ Mbytes to be stored. According to Lars, the digital image only takes ca 2.44 Mbytes. Explain which one of them is right.

b) Lars halftoned this image in Photoshop and the halftoned image is the same size ($pixel \times pixel$). How much memory is needed to store this halftoned image?

c) How large ($cm \times cm$) will this image be if printed at $dpi = 144$?

3.3. How much can you as its maximum enlarge a photograph in print if it is scanned at $ppi = 1200$ and the screen frequency is $lpi = 150$. Is it an AM or FM halftone? Why?

3.4. A color photograph has been scanned at 600 *ppi*.

a) The image is FM-halftoned. At what *dpi* should it be printed if the printed image is supposed to be half the size of the original photograph in each direction?

b) The image is AM-halftoned. At what *lpi* should it be printed if the printed image is supposed to be twice the size of the original photograph in each direction (four times in area)?

c) How many gray tones could be represented in the AM halftoning case in part (b) if the print resolution is 1200 *dpi*?

3.5. Halftone the below image by:

$$image : \begin{bmatrix} 0 & 0.3 & 0.2 & 0.5 \\ 0.4 & 0.6 & 0.4 & 0.9 \\ 0.8 & 0.8 & 0.7 & 0.1 \\ 0.8 & 0.8 & 0.1 & 0.8 \end{bmatrix}$$

a) Threshold halftoning using a 4×4 threshold matrix representing 10 gray levels.

b) Table halftoning according to AM halftoning, using 3×3 halftone tables. The halftoned image is supposed to be 6×6 pixels.

3.6. a) Halftone the below image by table halftoning. The halftoned image is supposed to be 6×6 and you can use the below halftone tables:

$$image: \begin{bmatrix} 0.3 & 0.8 & 0.1 & 0.3 \\ 0.4 & 0.5 & 0.2 & 0.2 \\ 0.2 & 0.8 & 0.1 & 0.5 \\ 0.5 & 0.9 & 0.2 & 0.4 \end{bmatrix}$$

Halftone tables:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$
$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \textit{ and } \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

b) How many gray levels is it possible to represent with these halftone tables?

3.7. Halftone the below image by table halftoning, where each 2×2 sub-area is represented by a halftone table. Assume that the screen frequency and the print resolution are $lpi = 100$ and $dpi = 300$, respectively.

$$image: \begin{bmatrix} 0.4 & 0.4 & 0.7 & 0.7 \\ 0.4 & 0.4 & 0.7 & 0.7 \\ 0.7 & 0.7 & 0.3 & 0.3 \\ 0.7 & 0.7 & 0.3 & 0.3 \end{bmatrix}$$

3.8. Halftone the below image by error diffusion, using the below error filter.

$$image: \begin{bmatrix} 0.6 & 0.5 \\ 0.8 & 0.4 \end{bmatrix}$$
$$error\ filter: \begin{bmatrix} \times & 0.6 \\ 0.4 & \end{bmatrix}$$

3.9. The image below has been halftoned by error diffusion using a filter that diffuses the error only to one pixel to the right. The resulting halftoned image is also shown. Find the conditions that a, b, c and d have to fulfill in order to make it work. Give also an example of an image that works. Recall that, all pixel values in the original image are supposed to be between 0 and 1.

$$image: \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } the\ result: \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

3.10. The image below has been halftoned by error diffusion using the shown filter. The resulting halftoned image is also shown. Find the conditions the two weights, x and y, in the error filter have to fulfill in order to make it work.

$$image: \begin{bmatrix} 0.6 & 0.6 \\ 0.8 & 0.4 \end{bmatrix},\ error\ filter: \begin{bmatrix} \times & x \\ y & \end{bmatrix} \text{ and } the\ result: \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Chapter 4

Advanced halftoning methods

The halftoning methods presented and discussed so far, have been operating either point-by-point, such as threshold halftoning, or on a small area around a pixel being processed, such as error diffusion. These methods are usually not very complex and operate quite fast. With the fast increase of the computer power and computation speed, other types of halftoning methods, here referred to as advanced halftoning methods, have become more and more popular during the last couple of decades. These types of halftoning usually result in higher print reproduction or fulfill some demands that were not possible to achieve using the conventional halftoning methods.

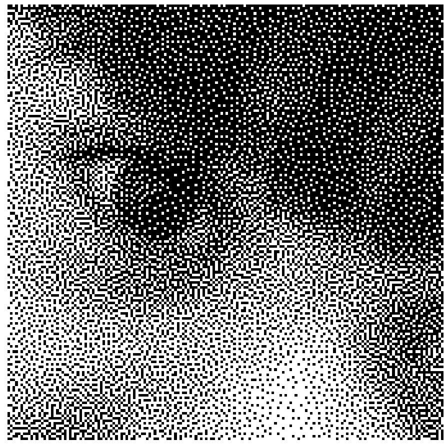
In this chapter, three different categories of advanced halftoning methods are very briefly discussed. One of them is iterative halftoning methods, which, unlike error diffusion, commonly operate over the entire image and aim at iteratively minimizing/reducing a defined error. Another category is hybrid halftoning methods, in which two or more different halftoning methods, for example AM and FM, incorporate in the same image. The third category is the second generation FM halftoning, in which, unlike first-generation FM halftoning discussed in Section 3.3, both the size and the frequency of the dots are variable.

4.1 Iterative halftoning

The need for having cheap, fast and high quality print has increased during the last decade of 20th century. The fast increase of the computer power has opened the door for the researchers to be able to introduce high quality and complex but still applicable halftoning methods. Unlike the ordered dithering methods (Section 3.2.1), which are point-by-point, and unlike error diffusion (Section 3.2.3), which operates on a neighborhood of currently processed position, the iterative methods commonly operate over the entire original image. These methods normally start with an initial binary image and iteratively decrease the difference between the binary image and the original image. The difference is usually defined as the visual difference between images. Therefore, different models of human eye are used in these methods. In this short chapter, we only describe one of these many iterative methods, i.e. "Iterative Method Controlling the Dot Placement (IMCDP).

4.1.1 IMCDP

In the halftoning method IMCDP (Iterative Method Controlling the Dot Placement), halftone dots are placed iteratively with the goal of reducing the difference between the original and the halftoned image. The generation of the halftoned image starts with a blank image the same size as the original. The total number of dots to be placed in the halftoned image (or in a number of different graytone regions of it) is dependent on the original image's overall lightness/darkness (or its average tone value in different graytone regions) and therefore is known in advance. Starting with a blank initial image, in the first iteration, the algorithm finds the position of the darkest pixel (the pixel holding the maximum value) in the original image and places the first dot at that location in the halftoned image. In the next step, the low-pass filtered version of the halftoned image is subtracted from the low-pass filtered version of the original image. The low-pass filter used is a Gaussian filter with standard deviation 1.3 truncated to 11×11 pixels. This operation is addressed as the feedback process. Subtracting the filter from the image around the found pixel reduces the pixel values in a neighborhood of that pixel, meaning that the chance of the neighboring pixels to be picked as the next maximum is reduced. Then, the location of the maximum pixel value of the subtracted image is found and at that location on the halftoned image the next dot is placed. The process continues until the known number of dots is placed, and the final halftoned image is achieved.
Using an 11×11 Gaussian filter makes the method work quite well for almost all kinds of images. However, the dots in the extreme highlights (or shadows) are not placed as homogeneously as one would expect. The reason is that the 11×11 filter is not big enough to homogeneously distribute the dots in those regions. In order to distribute the dots as homogeneously as possible in the very light (and very dark) regions, in the proposed method, the size of the Gaussian filter (or its standard deviation) is a variable of the gray level of the region where the maximum is found. The mentioned 11×11 Gaussian filter is used in the areas with tonal values between 4% and 96% and for the rest of the image a filter with varying size (or standard deviation) is used.
Figure 4.1 shows the test image being halftoned by IMCDP. Compared to the halftones created by conventional halftoning or error diffusion, the result is more visually pleasing.



(a) (b)

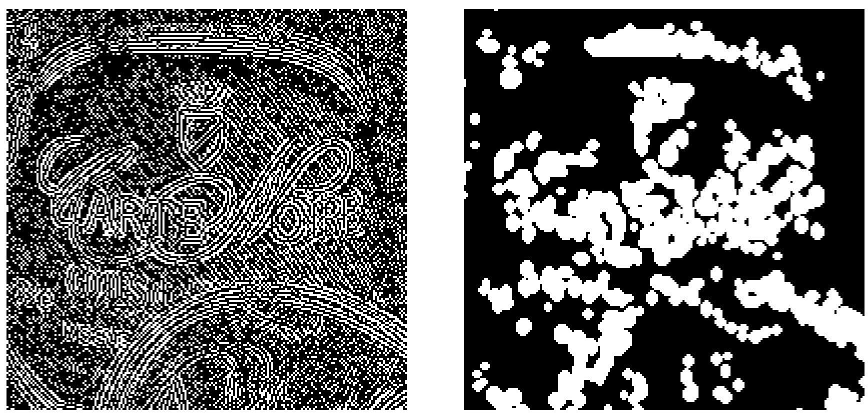
Figure 4.1: (a) The image is halftoned by IMCDP. (b) Enlargement of a part of (a).

4.2 Hybrid halftoning

Recall from Section 3.3 that halftoning methods are usually divided into two main categories, namely AM (Amplitude Modulated) and FM (Frequency Modulated). Both methods have their own advantages and drawbacks. The AM methods normally result in more homogeneous halftoned images when the gray tone of the original image varies slowly. Add to this the fact that AM methods suffer less from dot gain than FM methods do, see Section 3.4.3. The FM methods, on the other hand, are superior when dealing with heavily textured original images. Thus, these two methods can be combined so that the details of the original image are halftoned by a FM method and the rest of it by an AM method. Another possible application of hybrid halftoning is when the printing method can not produce the dots sufficiently small in order to be able to handle the very light and dark parts of the original image by just using an AM halftoning method. Flexographic printing can be mentioned as an example of such printing methods. In the two following subsections, two different ways of combining AM and FM halftones are presented.

4.2.1 Hybrid halftoning, first method

As been discussed, FM methods are preferred when the original image is a heavily textured image. The details are reproduced better when a FM halftoning method is used. However, when the gray tone of the original image varies slowly, the AM halftoning methods result in more homogeneous halftoned images. The question here is how to combine these two methods so that the details of the original image are FM-halftoned and the rest of it AM-halftoned. This is done by first high-pass filtering the original image. The absolute value of this high-pass version will have its largest density values in high-pass regions, i.e. the details of the original image. By thresholding the absolute value of this high pass version with a fixed threshold, a mask is created to be used to combine AM and FM methods. Sometimes this mask should go through a number of erosion and dilation operations before it can be used, see Figure 4.2. The details of the original image can be localized in regions where the mask, i.e. the image in Figure 4.2 (b), is white. By changing the high-pass filter and/or the fixed threshold, different masks and consequently different results can be obtained. The mask shown in Figure 4.2 (b) divides the original image into two main regions, one that should be halftoned by an AM method (black regions in the mask) and the other by a FM method (the white regions in the mask). In



(a) *(b)*

Figure 4.2: (a) The mask before dilation and erosion operations were performed. (b) The mask after erosion and dilation operations were performed.

the proposed method, we first halftone a test image by an AM-method, see Figure 4.3 (a). The dots in the AM halftoned part, which are close to the border between these two regions, should affect the parts of the original image that is supposed to be FM-halftoned. Therefore, we take into consideration the effect of the dots being placed in the AM-halftoned part on the part that should be FM-halftoned. This is done precisely as it is done in the FM method called IMCDP presented in Section 4.1.1. The image halftoned by this hybrid method is shown in Figure 4.3 (b). As can be seen, the details are produced better in the hybrid-halftoned image.

4.2.2 Hybrid halftoning, second method

Flexography is a modified form of letterpress printing method that is commonly used in the packaging industry for printing on the most varied materials. The print quality in flexography is lower than that in for example offset printing. On the other hand, flexography is the only printing method that can print on

(a) (b)

Figure 4.3: (a) The image is halftoned by an AM method. (b) The image is halftoned by the proposed hybrid method.

very thin, flexible, and solid films, thick card boards, rough-surface packaging materials and fabrics. One of the problems with the Flexographic printing is that it can not produce the dots sufficiently small in order to handle the very light and dark parts of the original image by just using an AM method. Experience on AM and FM methods showed that both technologies have their advantages and drawbacks in the Flexographic printing. AM technologies are especially useable in the mid-tones because dot gain is lower than when a FM technology is used there. FM methods, on the other hand, perform very well in the highlights where the optimized dot positioning and the choice of the minimum dot size allow for a perfect match to the technical limitations of the printing process. The possibility of combining these two technologies has been investigated for a long time. One of the biggest problems with combining these two technologies is how to handle the transition area from the AM-halftoned to the FM-halftoned parts.

Figure 4.4 (Top) shows a grayscale ramp (from 0% up to 50% coverage) halftoned by an AM halftoning method. For illustration purposes, a threshold matrix only representing 19 levels of gray was used to threshold the ramp. If the printing press is not able to produce the dots sufficiently small, the very light parts of the image will remain empty. In this example, the parts with gray values under 5.2% remain empty, marked with a solid vertical line. This problem occurs in the AM methods, because different gray tones are represented by changing the size of the dots. The lighter the gray tone becomes, the smaller the dots should be. When the dots cannot be made smaller than a specific size, the lighter gray tones cannot be represented properly. This is illustrated in Figure 4.4 (Top) to the left of the vertical solid line. The same problem occurs for the very dark part of the original image as well. The idea here is to use a FM method in those parts where the AM method is unable to represent the gray tones properly. Instead of decreasing the size of the dots, which is impossible for gray tones lower than 5.2% in this example, we can use the dots with the smallest possible size and change their frequency (number) in order to represent lighter gray tones. That is exactly what is done in a FM method, that is the size of the dots is constant while their frequency (the number of the dots) varies in order to represent different gray tones. We use the FM method IMCDP described in Section 4.1.1 and halftone the part that couldn't be handled by the AM method, see Figure 4.4 (Bottom). As seen in this image, the lighter tones can now be reproduced by using a FM method in the highlight.

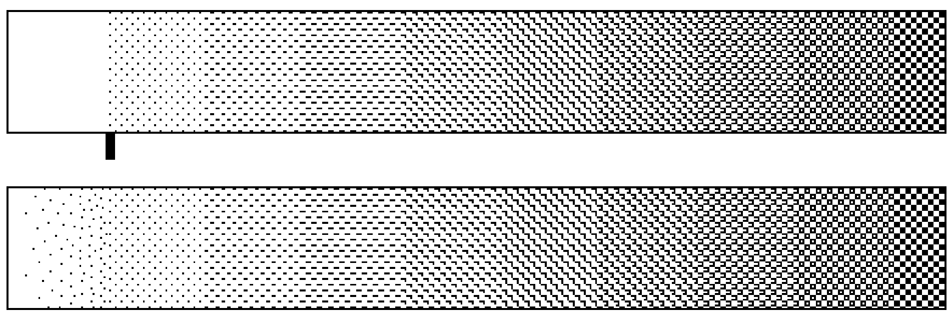


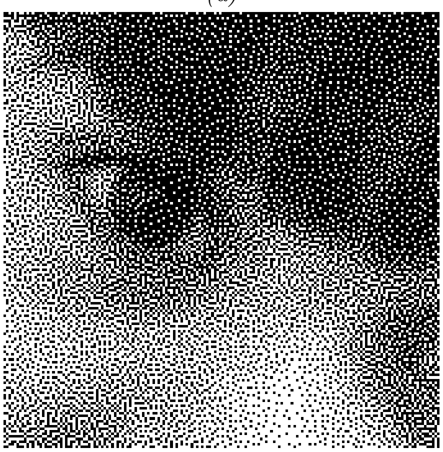
Figure 4.4: The gray scale ramp (from 0% up to 50% coverage) is halftoned by, (Top) an AM method. (Bottom) the proposed hybrid method.

4.3 Second generation FM halftoning

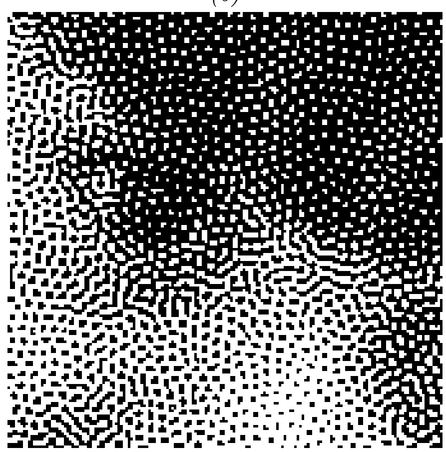
The FM halftones discussed so far in this book are the so called first-generation FM halftones (or even sometimes called stochastic halftones), in which the single dots are "stochastically" distributed. These types of halftones are widely used in printing technologies, such as inkjet, that are able to stably print isolated dispersed dots. Printers, such as laser printers, that utilize electrophotographic technology are not able to stably print the isolated dots and therefore use clustered-dot halftones (Section 3.2.1). Periodic clustered-dot, i.e. AM, halftones are commonly used in this type of printers but they suffer from undesired periodic interference pattern called Moiré (see Section 7.3.1). An alternative solution is to use second-order FM halftones, in which the clustered dots are stochastically distributed. In other words, in this type of halftones, both the size and the frequency are variable. In literature, this type of halftones is also referred to as stochastic-clustered dot halftones and even green-noise dither patterns. Figure 4.5 (a) and (b) show an image being halftoned by first generation FM halftoning (IMCDP) and a second generation FM halftoning method, respectively. As can be clearly seen in the image shown in Figure 4.5 (b) and (d) the dots are clustered but still stochastically distributed.

(a)

(b)



(c)



(d)

Figure 4.5: The test image is halftoned by: (a) First generation FM halftoning (IM-CDP). (b) Second generation FM halftoning. (c) and (d) are enlargements of a part of (a) and (b), respectively.

Chapter 5

Halftone image quality

Evaluation of image quality is difficult. It is difficult because there are many aspects that should be considered when judging the quality. In most cases, some of these aspects even contradict each other. Therefore, the image quality is normally measured subjectively. This is done by building a panel of test persons, who separately evaluate the image quality. The results are then gathered and summarized and finally an evaluation of the quality is made. This method needs time and human resources. Choosing the test panel is not so simple either. Therefore, it would be very useful if an objective quality measure was developed. Halftoning is without any doubt one of the most important tasks in most printing processes. The characteristics of the halftoning method have a great impact on the print result. Investigation of the behavior of different halftoning methods and a constant modification of the old methods, such as Error Diffusion, have been the topics of many publications. Many halftoning methods have also been proposed and developed in the past decades. Due to the increasing demand for high print quality, different measures for image quality have been proposed in different publications. However, the fact is that there won't be one simple objective quality measure to cover all the aspects of image quality.

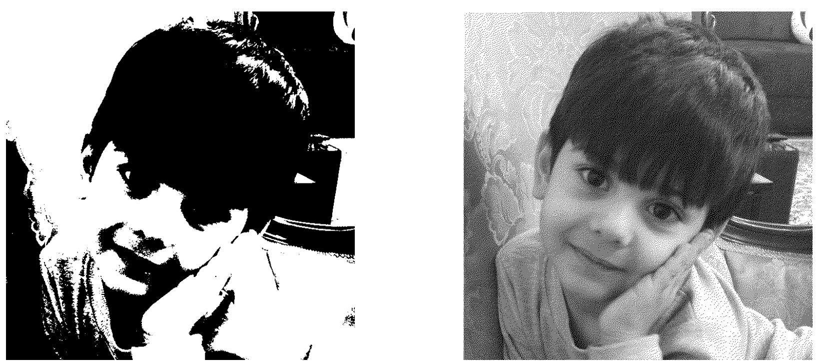
In this chapter, the focus will only be on three measures for halftone image quality without going so much into the details. These measures are based on the similarity between the original image and the halftoned image.

5.1 Signal to Noise Ratio (SNR)

SNR (Signal to Noise Ratio) is a measure used in science and engineering that compares the level of a desired signal to the level of background noise. It is defined as the ratio of signal energy to the noise energy, often expressed in decibels. A ratio higher than 1:1 (greater than 0 dB) indicates more signal than noise, meaning that the higher the SNR, the better the quality of the signal. SNR is usually applied to almost any form of signal and because of its simplicity, it is the first candidate to be used as a measure for halftone image quality. Equation 5.1 shows the equation to calculate SNR, where g and b denote the original and the halftoned image, respectively.

$$SNR = 10\log_{10}\left(\frac{\sum_{i,j}(g(i,j))^2}{\sum_{i,j}(g(i,j)-b(i,j))^2}\right). \tag{5.1}$$

As seen in this equation, the noise shown in the denominator is defined as the difference between the original and the halftoned image. The more "similar" (or closer) the two images are, the smaller the denominator, and the higher the SNR value. Notice that in this equation the numerator, i.e. the energy of the original image, is independent of the halftoned image and can be considered as constant when comparing halftoning methods with each other. By taking into account the fact that a halftone image only consists of 0:s and 1:s, it is easy to figure out that the denominator is minimized when the halftoned image b is the original image thresholded with the constant threshold 0.5. Thresholding an image by 0.5 results in an image consisting of only two levels of gray, see Figure 5.1 (a). This image gives a SNR value of 5.48. Any other halftone version of this original image will for sure give a smaller SNR value. For instance, the image shown in Figure 5.1 (b), which is the original image halftoned by error diffusion, gives the SNR value of 2.54.



(a) *(b)*

Figure 5.1: (a) The image is thresholded by 0.5. (b) The image is halftoned by error diffusion.

This example clearly shows that, this type of measures won't work satisfactorily as a halftone quality metric. The most important reason is that they don't take into account the human observer.

5.2 Modified SNR

The first idea to improve the simple SNR measure is to involve the human eye in the equation. This can be done by models developed representing the human eye. The human eye works as a low-pass filter, meaning that it is more sensitive to lower frequencies. In literature, there have been a number of models that could be used in the proposed modified SNR measure. In this short introduction, we don't go into the details of the filters representing the human eye. The SNR measure defined in Equation 5.1 is modified as follows. First, apply the human eye filter to the original image (g) and the halftoned image (b). Assume the results are denoted by g_1 and b_1, respectively. By replacing g and b in Equation 5.1 by g_1 and b_1, respectively, the simple SNR measure is modified and takes into account the human observer. Applying this modified measure to the images in Figure 5.1, gives the SNR value of 5.86 and 14.92 for the images in Figure 5.1 (a) and (b), respectively. The result of this modified SNR measure correlates much better with our perception. However, the result of the modified SNR measure doesn't always correlate well with our judgment of halftone image quality.

5.3 Quantization Noise Spectrum (QNS)

As been discussed, one of the important aspects when judging the halftone image quality is to find out how "similar" the halftoned and the original images are. The more similar they are, the better the halftone image quality. Of course, it is sometimes desirable to have a printed image that is for example sharper than the original. Therefore, we cannot state that we always want to have a printed image that completely resembles the original. This kind of demands can however be fulfilled outside the halftoning process and here we just focus on the similarity between the two images.

In this section, we examine how "similar" the halftoned and the original images are by studying the Quantization Noise Spectrum (QNS), which is defined as,

$$|Q(k,l)|^2 = |G(k,l) - B(k,l)|^2, \tag{5.2}$$

where $G(k,l)$, $B(k,l)$ and $Q(k,l)$ are the two dimensional discrete Fourier transforms of the original image, the halftoned image and the quantization noise, respectively. The smaller the quantization noise, the more similar the two images are. Since the human eye acts as a low-pass filter, the noise is desired to be as small as possible in the low-pass region of the spectrum. Equation 5.3 can therefore be used as such a measure.

$$e = \sum_{\Omega} |Q(k,l)|^2, \tag{5.3}$$

where Ω denotes a low pass region. To test this measure, let us illustrate an example. A constant image with a gray value of 1/32 is halftoned by Error Diffusion (Section 3.2.3) and IMCDP (Section 4.1.1), see Figure 5.2. The QNS for each image is shown under the corresponding image.

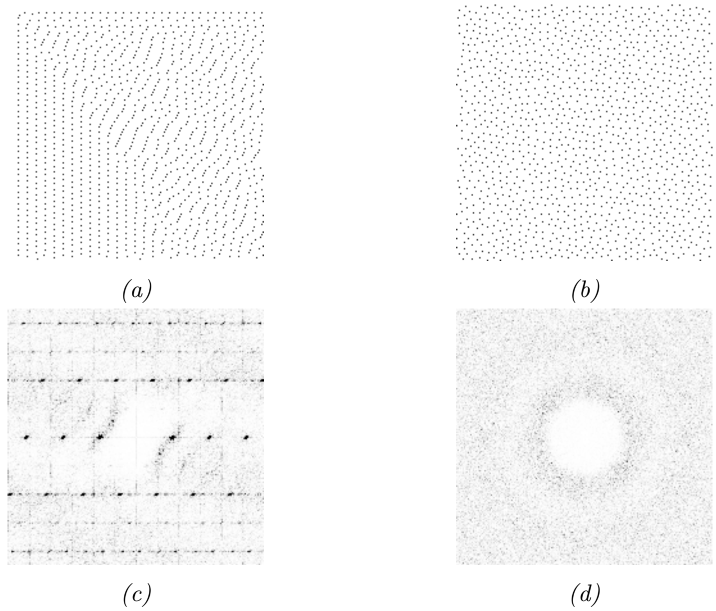
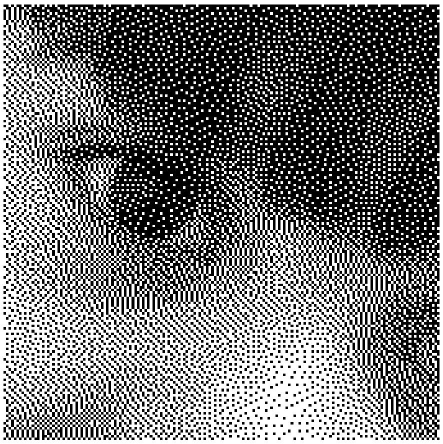
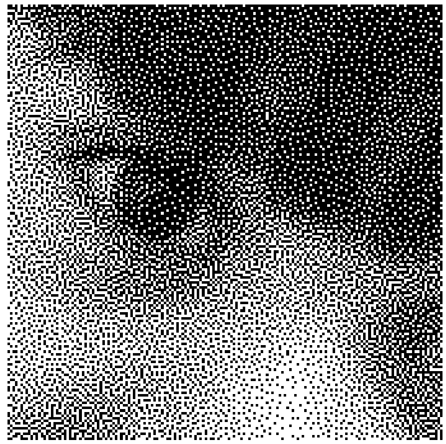


Figure 5.2: A constant image with a gray value of 1/32 is halftoned by (a) Non-modified error diffusion. (b) IMCDP. The quantization noise spectrum for each image is shown under the image.

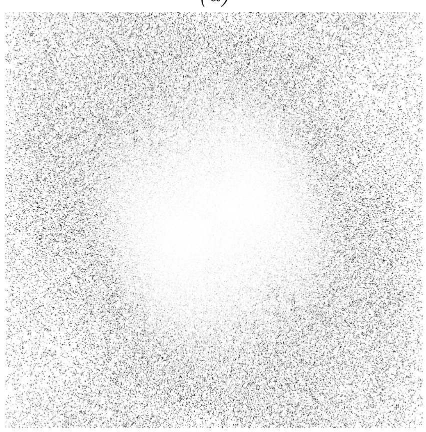
It is easy to notice that the dots in the image shown in Figure 5.2 (b) are placed more homogeneously than those in Figure 5.2 (a). These images also show how the dot placement impacts the shape of QNS. We calculated the error e, defined in Equation 5.3, when Ω is a circular low-pass region that occupies 12.5% of the spectrum. The error e is actually slightly smaller for the image shown in Figure 5.2 (a). This calculation contradicts our subjective judgment of these images. Our conclusion is that, it is not only the magnitude of QNS in the low-pass region that is of importance but also the shape of QNS in the low-pass region plays a significant role. For judging the quality of an image, we suggest that both the magnitude and the shape of QNS should be studied. Figure 5.3 shows an enlargement of a part of the test image being halftoned by error diffusion and IMCDP. The QNS for each image is shown under the image. The magnitude of QNS is approximately the same for both images, although the image shown in Figure 5.3 (b) is perceived better, especially in the highlights and mid-tones. In these cases, we wouldn't have been able to draw any certain conclusion about the quality by just calculating the magnitude of QNS as defined in Equation 5.3. On the other hand, we can see that the image whose QNS is more circularly symmetrical is perceived better. Therefore, both the magnitude and the shape of QNS have to be studied to be able to draw a certain conclusion about the halftone image quality.



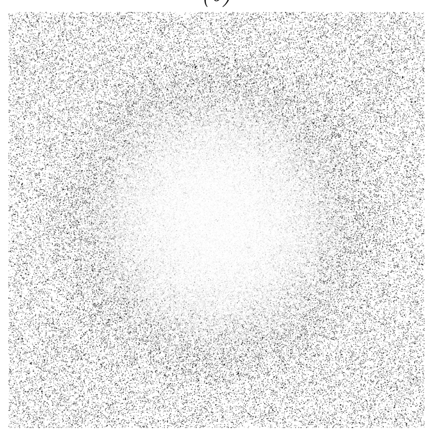
(a)



(b)



(c)



(d)

Figure 5.3: The test image is halftoned by (a) Non-modified error diffusion. (b) IM-CDP. The quantization noise spectrum for each image is shown under the image.

Chapter 6

Basic color science

The basic concepts of image reproduction and halftoning methods discussed and explained so far, have been about achromatic images but almost all of the reproductions of today are color reproductions. In order to understand the details of color reproduction, it is thus important to have knowledge about basic concepts in color science.
This chapter aims to explain concepts related to basic color science and color perception, by starting with an overview of the way humans perceive colors. This chapter also explains how to define, measure, and quantify colors, with notions such as color spaces and color difference formulas. At the end of this chapter, a number of exercises are provided, for which you can find the answers and the solutions in Chapter 8 and 9, respectively.
Parts of this chapter is taken from the lab material written by Jörgen Rydenius and Paula Zitinski Elias' PhD thesis.

6.1 What is color?

The human eye is able to detect light, i.e. electromagnetic radiation, with wavelengths in the interval between $380 - 780$ nm (nanometer). The radiant flux of the observed light at each wavelength is expressed by a Spectral Power Distribution (SPD), such as in Figure 6.1. Since the SPD shown in Figure 6.1 has its energy mostly concentrated at shorter wavelengths (around 450 nm), this radiation will be perceived as a bluish color by human eye/brain. The fact is that, this is the SPD that exists in the physical world but the color of the ray actually exists in our brain/eye. That is what Newton also meant when he stated that the rays are not colored.

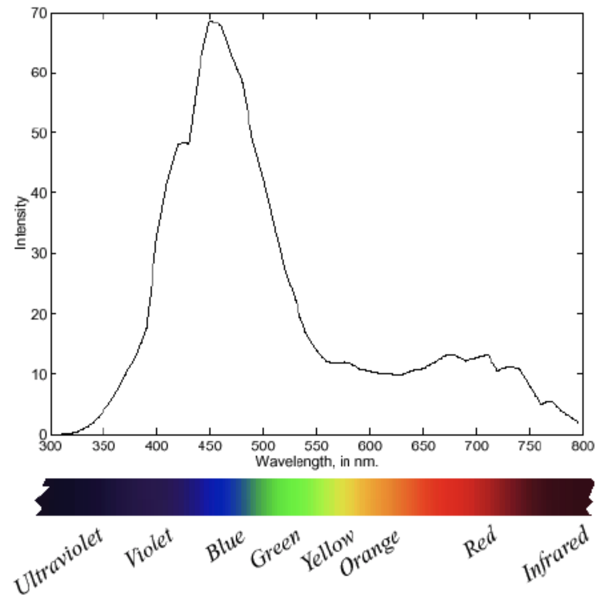


Figure 6.1: An example of a spectral power distribution. This distribution will be perceived by the human observer as a bluish color.

6.2 Human visual system

The human visual system (HVS) consists of photoreceptors located in the eye's retina that are susceptible to illumination stimuli. There are two types of photoreceptors: rods and cones. Rods are useful for vision under low light levels and do not contribute to color. When the light levels are higher, the rods become saturated and do not contribute to vision. The cones are the ones that become active under normal light levels, and are responsible for color vision.
There are three types of cones with different light susceptibility, sensitive to wavelengths approximately corresponding to red, green, and blue light. They are approximately peaking at short ($420-440$ nm), medium ($530-540$ nm), and long ($560 - 580$ nm) wavelengths of visible light. These cone types divide the visible light spectrum into three bands, accounting for the human trichromatic color vision. Any light susceptible by the cones evokes stimuli that are translated in the brain as a color sensation. These stimulus combinations account for the colors perceived by the HVS.

6.3 Measuring colors

As discussed above, the human eye has three different color receptors, i.e. cones. They are referred to as L, M, and S cones (also sometimes called r, g and b), since they are sensitive to long, medium, and short wavelengths in the visible spectrum, respectively. Since the impression of color is related to how the human eye works, the eye's three sensitivity functions are used as mathematical foundations to measure and define color values. Furthermore, light with different spectral distributions yielding the same color impression should be measured as the same color. Figure 6.2 shows the incoming light with the intensity $I(\lambda)$, reflected from an object having the reflectance spectrum $R(\lambda)$, being detected by the eye. As seen in this figure, the photon distribution detected by the eye is the product of the incoming light intensity and the spectral reflectance of the object, i.e. $R(\lambda)I(\lambda)$.

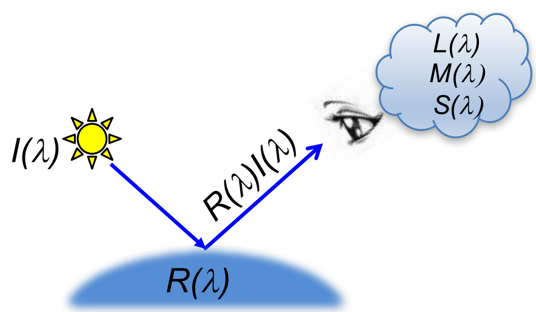


Figure 6.2: The incoming light is firstly reflected from the object and then perceived by the eye.

Equation 6.1 gives the total stimulation of L, M and S cone types.

$$\begin{cases} L = \int_\lambda R(\lambda)I(\lambda)L(\lambda)d\lambda \\ M = \int_\lambda R(\lambda)I(\lambda)M(\lambda)d\lambda \\ S = \int_\lambda R(\lambda)I(\lambda)S(\lambda)d\lambda \end{cases}, \tag{6.1}$$

where the integrals are calculated over the visible wavelength interval.
Since there are three types of color sensitive receptors, as seen in Equation 6.1, the color should be represented by three values, here L, M and S. As an example, assume an achromatic (gray/white) light source, I_{gray}, having the same intensity value at all wavelengths. If this light source hits a red object (R_r) and is reflected from that, then the reflected light, i.e. $R_r(\lambda)I_{gray}(\lambda)$, will have most concentration at longer wavelengths, because the red object reflects the longer wavelengths more than the others. Therefore, the total stimulation L in Equation 6.1 will have a larger value than M and S. As another example, consider a blue light source that hits a red object and is reflected from the object. The intensity of a blue light (I_b) includes large values at short wavelengths, and small values at medium and long wavelengths. The spectral reflectance of a red object (R_r), on the other hand, includes large values at long wavelengths, and small values at short and medium wavelengths. The reflected light, $R_r(\lambda)I_b(\lambda)$, will therefore have small values at all wavelengths, giving small L, M and S,

indicating that a dark color (black) is perceived by the eye.
Therefore, the implications of Equation 6.1 are that two objects may have the same color in one illumination and then look very different from each other in another, and that a single object is perceived as having different colors when viewed under different illuminations. These effects are called **metamerism**. Metamerism is a large problem when two colors are supposed to match each other. It is more or less unavoidable, unless the two objects have exactly the same reflectance spectrum.

6.3.1 Color matching functions

Equation 6.1 would be a useful way of expressing colors, if only the sensitivity functions of the cone types were known exactly. Since they are not, a different approach has to be taken. In 1931, the International Commission on Illumination (CIE) proposed that the sensitivity functions for the L, M, and S cones should be replaced by three other well defined sensitivity functions, called $r(\lambda)$, $g(\lambda)$ and $b(\lambda)$. Figure 6.3 shows a schematic of how the color matching experiment to find the three color sensitivity (matching) functions was carried out.

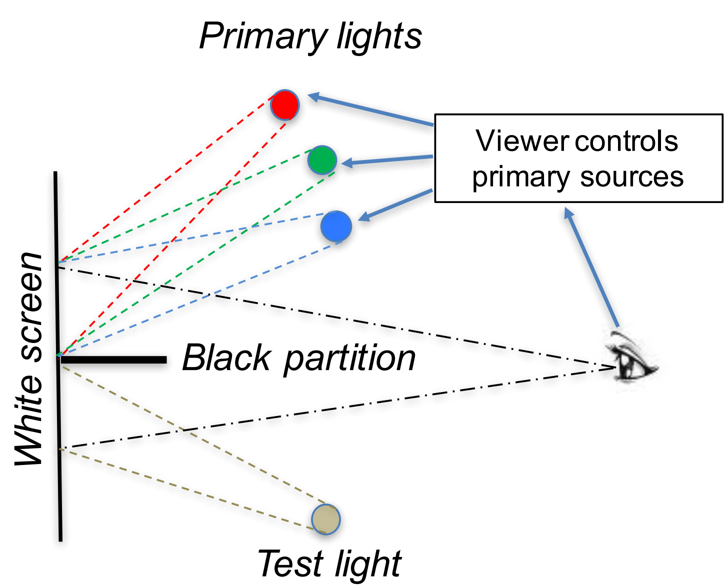


Figure 6.3: Color matching experiment.

The purpose of the experiment was to decide whether it was possible to adjust the power of three primary lights, a red, a green and a blue light with precisely known single wavelengths, until the spot had the exact same color as for example 540 nm test light. In the color matching experiment shown in

Figure 6.3, three equal energy monochromatic light sources (the primaries) red, green and blue at 700 nm, 546.1 nm and 435.8 nm, respectively, were used. A set of monochromatic colors of the spectrum were used to match each light in this set with combinations of the three primaries. As shown in Figure 6.3, each wavelength of light (test light source) illuminates the viewing screen on one side of the partition. The viewer then adjusts how much of each of the primary RGB colors are needed to produce the same color as the test source. The amount of each color was then stored for each viewer and the average value for a number of test viewers were plotted. Figure 6.4 shows the obtained rgb color matching functions. As seen in this figure, the $r(\lambda)$ color matching function holds negative values in some wavelengths, which is physically not possible. The reason for that was that some of the wavelengths of the test sources were not possible to match. In order to match those wavelengths, the red primary light was moved to the other side of the partition, i.e. to the same side as the test source. Then, it was possible to match, and therefore the intensity value for the red primary was assigned a negative value at those wavelengths.

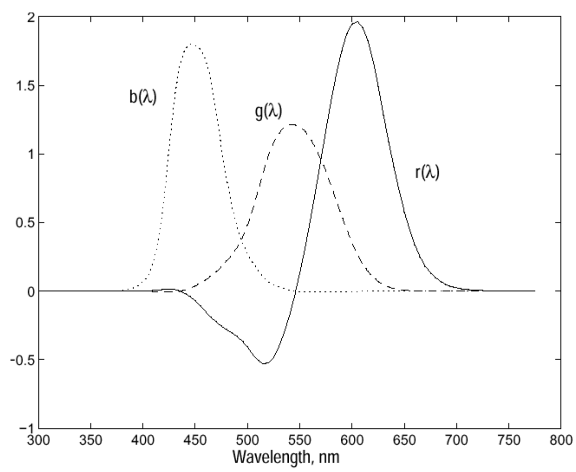


Figure 6.4: rgb color matching functions.

In order to avoid the non-physical negative sensitivity values, a linear transformation from rgb to $\bar{x}\bar{y}\bar{z}$ was performed according to Equation 6.2 to obtain useful and physically correct color matching functions, i.e. $\bar{x}(\lambda)\bar{y}(\lambda)\bar{z}(\lambda)$, representing the color matching sensitivities of a standard observer.

$$\begin{bmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{bmatrix} = \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} r(\lambda) \\ g(\lambda) \\ b(\lambda) \end{bmatrix}. \tag{6.2}$$

As noticed in the transformation matrix in Equation 6.2, the elements in the second row to obtain $\bar{y}(\lambda)$ have more decimals. The reason is that CIE wanted

$\bar{y}(\lambda)$ to approximate photopic luminous efficiency as closely as possible, which simply required more decimals to achieve a good match.
Figure 6.5 shows the obtained $\bar{x}\bar{y}\bar{z}$ color matching functions, which are positive throughout visible wavelength range. In Section 6.4.1, we will discuss how the CIEXYZ tristimulus color values can be calculated using these three functions.

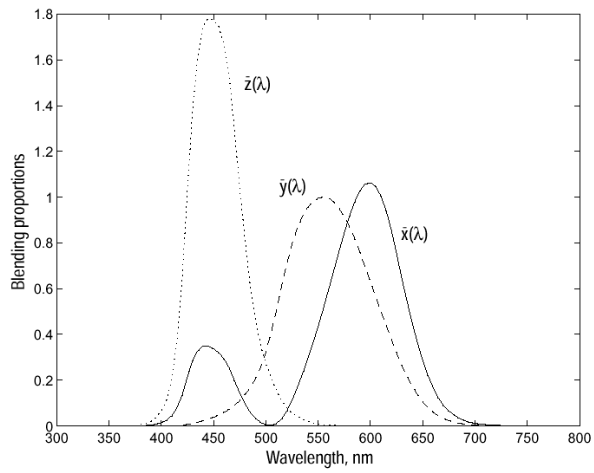


Figure 6.5: $\bar{x}\bar{y}\bar{z}$ color matching functions.

6.4 Color spaces

A color space is a defined range of colors and is a useful conceptual tool for understanding the color capabilities of a particular device. Well known color spaces include sRGB and AdobeRGB. In this section, we only discuss four different color spaces, namely CIEXYZ, CIELAB, RGB and CMYK. Some of the color spaces, such as RGB and CMYK, are device dependent, meaning that the reproduced color depends on the characteristics of the reproduction device. This means that the same RGB-color values might be displayed being different colors when reproduced by two different devices, for example two different computer screens. The device independent color spaces, such as CIEXYZ and CIELAB, on the other hand, are independent of the reproduction device and are basically based on how the human eye perceives color.

6.4.1 CIEXYZ color space

The CIEXYZ color space was created by CIE in 1931. The three CIEXYZ tristimulus values are calculated by an equation similar to Equation 6.1 as,

$$\begin{cases} X = k\int_{\lambda} R(\lambda)I(\lambda)\bar{x}(\lambda)d\lambda \\ Y = k\int_{\lambda} R(\lambda)I(\lambda)\bar{y}(\lambda)d\lambda \\ Z = k\int_{\lambda} R(\lambda)I(\lambda)\bar{z}(\lambda)d\lambda \end{cases}, \quad (6.3)$$

where $R(\lambda)$ and $I(\lambda)$ are, as defined before, the spectral reflectance of the object and the intensity of the light source, respectively and $\bar{x}\bar{y}\bar{z}$ are the color matching functions obtained by the CIE and shown in Figure 6.5. The factor k is a normalization factor for the current illumination, so that a completely white surface (reflectance function equal to one for all wavelengths, i.e. $R(\lambda) \equiv 1$) will always give $Y = 100$. Therefore, the normalization factor k is calculated by setting $Y = 100$ and $R(\lambda) = 1$ in the second equation in Equation 6.3 as shown in Equation 6.4.

$$k = \frac{100}{\int_{\lambda} I(\lambda)\bar{y}(\lambda)d\lambda}. \tag{6.4}$$

The illumination, $I(\lambda)$, is often assumed to be one out of several standard illuminations, defined by CIE. Commonly either D50 or D65 is used to represent the day light. The CIED light sources have approximately the radiation characteristics of black bodies at the temperature (in Kelvin) given by the number after D multiplied by 100. Figure 6.6 shows the spectra for a number of light sources. As seen in this figure, for example the tungsten light source (Tungsten60W) has the most concentration in the longer and medium wavelength, meaning a yellow-brownish color, while plank90k is bluish and CIED65 color neutral (representing the day-light).

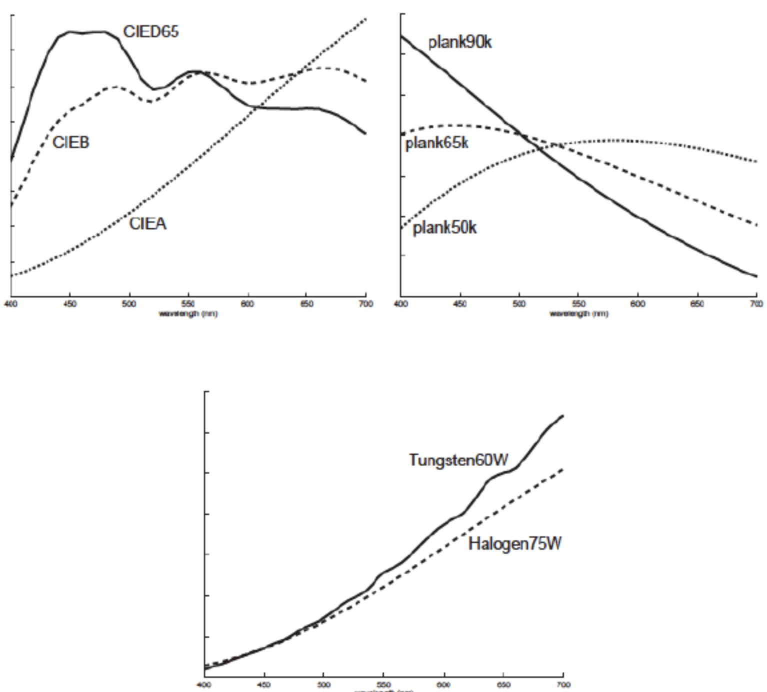


Figure 6.6: The spectra for a number of standard light sources.

Each illumination set has its own **white point**, whose CIEXYZ values are denoted by $[X_n, Y_n, Z_n]$. The white point is calculated by assigning $R(\lambda) \equiv 1$ in Equation 6.3. This means that the white point of a light source I is calculated by:

$$\begin{cases} X_n = k \int_\lambda I(\lambda)\bar{x}(\lambda)d\lambda \\ Y_n = k \int_\lambda I(\lambda)\bar{y}(\lambda)d\lambda \\ Z_n = k \int_\lambda I(\lambda)\bar{z}(\lambda)d\lambda \end{cases}, \tag{6.5}$$

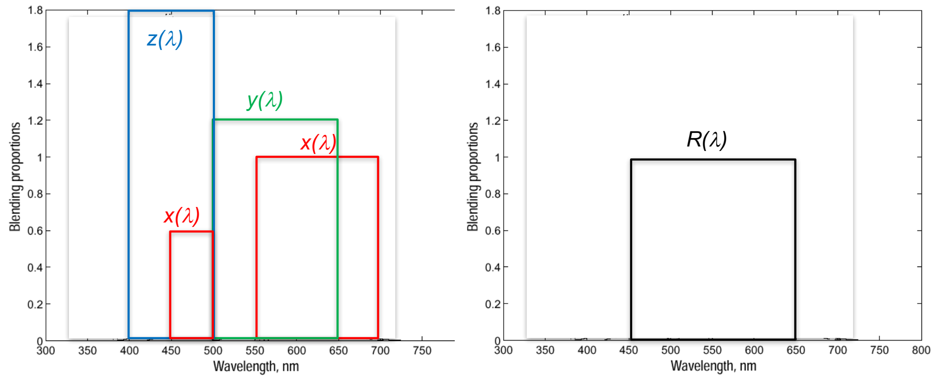
where k, the normalization factor, is calculated by Equation 6.4. The first observation looking at Equation 6.4 and 6.5, is that the CIEY value of any light source, Y_n, is 100. Also, note that the colors of all objects under the particular illumination are in the intervals, $0 \leq X \leq X_n$, $0 \leq Y \leq (Y_n = 100)$ and $0 \leq Z \leq Z_n$, if fluorescence effect is neglected. They can not be negative since the color matching functions are positive, and they can not be higher than the white point values, since no object can reflect incoming light better than an object with reflectance function equal to one at all wavelengths.
As discussed above, CIEXYZ is a device-independent color space, which means that the color representation is independent of the reproduction medium or technology of the device. The XYZ values correspond to linear transformations of the physical primaries, chosen to eliminate their negative values, and normalized to yield equal tristimulus values for the equi-energy spectrum. Furthermore, $\bar{y}(\lambda)$ is chosen to coincide with the luminous efficiency function, meaning that the tristimulus value Y represents the perceived luminance. A drawback of CIEXYZ is that it is perceptually non-uniform, meaning that the Euclidean distance between colors' CIEXYZ coordinates does not correspond to the perceived color difference, impeding a quantitative comparison between colors.

Example 6.1 *For simplicity, assume that the three color matching functions $\bar{x}\bar{y}\bar{z}$ can be approximated by the curves illustrated in Figure 6.7 (a) by red, green and blue, respectively. Assume further that the light source has an intensity equal to 2 in all visible wavelengths, i.e. $I(\lambda) \equiv 2$. The visible wavelength interval is assumed to be $[400, 700]$ nm.*
a) *Find the tristimulus values XYZ for an object having spectral reflectance shown in Figure 6.7 (b).*
b) *Find the tristimulus values of the white point of the light source.*

Solution (a): *Let us first calculate the normalization factor, i.e.* $k = \frac{100}{\int_\lambda I(\lambda)\bar{y}(\lambda)d\lambda}$.
As seen in Figure 6.7 (a), $\bar{y}$ is equal to 1.2 between 500 nm and 650 nm, and zero elsewhere. The normalization factor will therefore be,

$$k = \frac{100}{\int_{400}^{700} I(\lambda)\bar{y}(\lambda)d\lambda} = \frac{100}{\int_{500}^{650} 2{\cdot}1.2 \ d\lambda} = \frac{100}{2.4{\cdot}150} = 0.278.$$

In order to find CIEX, notice that the color matching function $\bar{x}$ is equal to 0.6 on $[450, 500]$ *and equal to 1 on* $[550, 700]$ *and zero elsewhere, giving:*



(a) (b)

Figure 6.7: Example 6.1: (a) Approximated $\bar{x}\bar{y}\bar{z}$ color matching functions. (b) The reflectance spectrum of an object.

$X = k\int_{400}^{700} R(\lambda)I(\lambda)\bar{x}(\lambda)d\lambda = 0.278(\int_{450}^{500} R(\lambda)\cdot 2\cdot 0.6\ d\lambda + \int_{550}^{700} R(\lambda)\cdot 2\cdot 1\ d\lambda).$

Since $R(\lambda)$ is equal to 1 on $[450, 650]$ and zero elsewhere according to Figure 6.7 (b), then the CIEX value is:

$X = 0.278(\int_{450}^{500} 1\cdot 2\cdot 0.6\ d\lambda + \int_{550}^{650} 1\cdot 2\cdot 1\ d\lambda) = 0.278(1.2\cdot 50 + 2\cdot 100) = 72.28.$

By the same reasoning, knowing that $\bar{y}$ is equal to 1.2 on $[500, 650]$ and zero elsewhere, the CIEY will be calculated by,

$Y = k\int_{400}^{700} R(\lambda)I(\lambda)\bar{y}(\lambda)d\lambda = 0.278\int_{500}^{650} R(\lambda)\cdot 2\cdot 1.2\ d\lambda = 0.278\int_{500}^{650} 1\cdot 2\cdot 1.2\ d\lambda = 100.$

By the same reasoning, knowing that $\bar{z}$ is equal to 1.8 on $[400, 500]$ and zero elsewhere, the CIEZ will be calculated by,

$Z = k\int_{400}^{700} R(\lambda)I(\lambda)\bar{z}(\lambda)d\lambda = 0.278\int_{400}^{500} R(\lambda)\cdot 2\cdot 1.8\ d\lambda =$ ***(Since R is zero on [400, 450])*** $= 0.278\int_{450}^{500} 1\cdot 2\cdot 1.8\ d\lambda = 50.$

Therefore, the calculated tristimulus values for this reflectance spectrum under the specified light source is $XYZ = [72.28, 100, 50]$.

Solution (b): *The normalization factor for this light source is already calculated in part (a), being $k = 0.278$. The white point of the light source, as discussed, is simply calculated by assigning $R(\lambda) \equiv 1$ in the above equations, giving:*

$X_n = 0.278(\int_{450}^{500} 2\cdot 0.6\ d\lambda + \int_{550}^{700} 2\cdot 1\ d\lambda) = 0.278(1.2\cdot 50 + 2\cdot 150) = 100.$

$Y_n = 100$, *no need for calculation, because, as discussed before, the light sources always have a CIEY-value equal to 100.*

$Z_n = 0.278 \int_{400}^{500} 2 \cdot 1.8 \; d\lambda = 0.278 \cdot 100 \cdot 3.6 = 100.$

Therefore, the white point of the light source in this exercise has the tristimulus values $[X_n, Y_n, Z_n] = [100, 100, 100]$.

Here, it is worth noticing that since the light source has been constant over the visible wavelength band in this example ($I(\lambda) \equiv 2$)*, the X, Y and Z tristimulus values of the white point of this light source are equal. The reason is that, as mentioned above, the color matching functions are normalized to yield equal tristimulus values for the equi-energy spectrum. If the light source is not constant over the visible wavelength band, only its CIEY value (i.e.* Y_n*) is equal to 100.*

As discussed in Section 6.3.1, in order to find the color matching functions a limited number of test light samples were used. Therefore, the obtained color matching functions are usually defined by vectors rather than continuous functions. For example, if the visible wavelengths is assumed to be between 380 and 780 nm, and the test light samples are presented by single wavelengths in this range with a step of 10 nm, then the color matching functions will be represented by vectors having 41 elements, where the first element represents the value at 380 nm, the second one the value at 390 nm, the third one at 400 nm, etc. The intensity of the light source and the reflectance spectra of objects are, accordingly, represented by vectors of the same size. The formula in Equation 6.3 will therefore become:

$$\begin{cases} X = k \sum\limits_{i=1}^{n} R(i) \cdot I(i) \cdot \bar{x}(i) \\ Y = k \sum\limits_{i=1}^{n} R(i) \cdot I(i) \cdot \bar{y}(i) \\ Z = k \sum\limits_{i=1}^{n} R(i) \cdot I(i) \cdot \bar{z}(i) \end{cases} , \tag{6.6}$$

where n is the size of the vectors and $R(i)$ and $I(i)$ denote element i in the vectors representing the reflectance spectrum and the light intensity, respectively. $\bar{x}$, $\bar{y}$ and $\bar{z}$ in Equation 6.6 are the vectors representing the color matching functions. The normalization factor k can be calculated, similar to Equation 6.4, by,

$$k = \frac{100}{\sum\limits_{i=1}^{n} I(i)\bar{y}(i)}. \tag{6.7}$$

Since in this course all the labs are done in Matlab, and the reader is familiar with Matlab notation and codes, it is much easier to write Equation 6.6 by

Matlab notations as,

$$\begin{cases} X = k * sum(R. * I. * \bar{x}) \\ Y = k * sum(R. * I. * \bar{y}) \\ Z = k * sum(R. * I. * \bar{z}) \end{cases}, \tag{6.8}$$

with $k = \frac{100}{sum(I.*\bar{y})}$. In Equation 6.8, *sum* denotes the Matlab function sum, which returns the sum of the elements of a vector. The operator $.*$ in Matlab denotes the element-wise multiplication between vectors.

Example 6.2 *For simplicity, assume that the visible wavelength interval has been divided into four equal sub-intervals and therefore all spectral data can be represented by vectors having four elements. Assume further that the vectors $\bar{x}$, $\bar{y}$ and $\bar{z}$ below represent the three color matching functions and vectors I_1 and I_2 represent the intensity of two different light sources. The vectors R_1 and R_2 represent the reflectance spectra of two different objects, object 1 and object 2, respectively.*

$\bar{x} = [0, 0, 1, 1]$, $\bar{y} = [0, 1, 1, 0]$, $\bar{z} = [1, 1, 0, 0]$,
$I_1 = [1, 1, 1, 1]$, $I_2 = [1, 2, 0, 3]$,
$R_1 = [1/2, 1/2, 1/4, 3/4]$ *and* $R_2 = [1, 0, 3/4, 1/4]$

Ignore the normalization factor k *and solve the following assignments.*

a) *Find the XYZ tristimulus values of object 1 under I_1.*
b) *Find the XYZ tristimulus values of object 2 under I_1.*
c) *Are R_1 and R_2 metameric under I_1?*
d) *Find the XYZ tristimulus value of object 1 under I_2.*
e) *Find the XYZ tristimulus value of object 2 under I_2.*
f) *Are R_1 and R_2 metameric under I_2?*
g) *Give an example of a reflectance spectrum that is metameric with R_1 under I_2.*

Solution (a): *Equation 6.8,* ***ignoring the normalization factor****, gives: (Recall that $.*$ denotes the element-wise multiplication between vectors)*

$X = sum(R_1. * I_1. * \bar{x}) = sum([1/2, 1/2, 1/4, 3/4]. * [1, 1, 1, 1]. * [0, 0, 1, 1])$
$= sum([1/2 \cdot 1 \cdot 0, 1/2 \cdot 1 \cdot 0, 1/4 \cdot 1 \cdot 1, 3/4 \cdot 1 \cdot 1]) = sum([0, 0, 1/4, 3/4])$
$= 0 + 0 + 1/4 + 3/4 = 1$

$Y = sum(R_1. * I_1. * \bar{y}) = sum([1/2, 1/2, 1/4, 3/4]. * [1, 1, 1, 1]. * [0, 1, 1, 0]) = sum([0, 1/2, 1/4, 0]) = 3/4$

$Z = sum(R_1. * I_1. * \bar{z}) = sum([1/2, 1/2, 1/4, 3/4]. * [1, 1, 1, 1]. * [1, 1, 0, 0]) = sum([1/2, 1/2, 0, 0]) = 1$

Therefore, the XYZ-tristimulus values for object 1 under I_1, without normalization, is $XYZ = [1, 3/4, 1]$.

Solution (b): *Equation 6.8, ignoring the normalization factor, gives:*

$X = sum(R_2.{*}I_1.{*}\bar{x}) = sum([1, 0, 3/4, 1/4].{*}[1, 1, 1, 1].{*}[0, 0, 1, 1]) = sum([0, 0, 3/4, 1/4]) = 1$
$Y = sum(R_2.{*}I_1.{*}\bar{y}) = sum([1, 0, 3/4, 1/4].{*}[1, 1, 1, 1].{*}[0, 1, 1, 0]) = sum([0, 0, 3/4, 0]) = 3/4$
$Z = sum(R_2.{*}I_1.{*}\bar{z}) = sum([1, 0, 3/4, 1/4].{*}[1, 1, 1, 1].{*}[1, 1, 0, 0]) = sum([1, 0, 0, 0]) = 1$

Therefore, the XYZ-tristimulus values for object 2 under I_1, without normalization, is $XYZ = [1, 3/4, 1]$.

Solution (c): *According to the definition of metamerism in Section 6.3, two different objects having the same color are metameric. The objects R_1 and R_2 are obviously different because they have two different reflectance spectra, but they both give the same color values, i.e. $XYZ = [1, 3/4, 1]$, under I_1. Therefore, the answers is yes, i.e. these two objects are metameric under I_1.*

Solution (d): *Equation 6.8, ignoring the normalization factor, gives:*

$X = sum(R_1. * I_2. * \bar{x}) = sum([1/2, 1/2, 1/4, 3/4]. * [1, 2, 0, 3]. * [0, 0, 1, 1])$
$= sum([1/2 \cdot 1 \cdot 0, 1/2 \cdot 2 \cdot 0, 1/4 \cdot 0 \cdot 1, 3/4 \cdot 3 \cdot 1]) = sum([0, 0, 0, 9/4])$
$= 0 + 0 + 0 + 9/4 = 9/4$

$Y = sum(R_1. * I_2. * \bar{y}) = sum([1/2, 1/2, 1/4, 3/4]. * [1, 2, 0, 3]. * [0, 1, 1, 0]) = sum([0, 1, 0, 0]) = 1$

$Z = sum(R_1. * I_2. * \bar{z}) = sum([1/2, 1/2, 1/4, 3/4]. * [1, 2, 0, 3]. * [1, 1, 0, 0]) = sum([1/2, 1, 0, 0]) = 3/2$

Therefore, the XYZ-tristimulus values for object 1 under I_2, without normalization, is $XYZ = [9/4, 1, 3/2]$.

Solution (e): *Equation 6.8, ignoring the normalization factor, gives:*

$X = sum(R_2.{*}I_2.{*}\bar{x}) = sum([1, 0, 3/4, 1/4].{*}[1, 2, 0, 3].{*}[0, 0, 1, 1]) = sum([0, 0, 0, 3/4]) = 3/4$
$Y = sum(R_2.{*}I_2.{*}\bar{y}) = sum([1, 0, 3/4, 1/4].{*}[1, 2, 0, 3].{*}[0, 1, 1, 0]) = sum([0, 0, 0, 0]) = 0$
$Z = sum(R_2.{*}I_2.{*}\bar{z}) = sum([1, 0, 3/4, 1/4].{*}[1, 2, 0, 3].{*}[1, 1, 0, 0]) = sum([1, 0, 0, 0]) = 1$

Therefore, the XYZ-tristimulus values for object 2 under I_2, without normal-

ization, is $XYZ = [3/4, 0, 1]$.

Solution (f): *The two objects give different XYZ tristimulus values (part (d) and (e) above), meaning that they are not metameric under* I_2.

Solution (g): *Assume the reflectance we are looking for is* $R_3 = [a, b, c, d]$. ***Notice*** *that since this vector represents a reflectance spectrum then its elements, i.e. a, b, c and d are between 0 and 1, because the reflectance value at any wavelength cannot be negative or larger than 1. The tristimulus values of* R_3 *under* I_2 *are:*

$X = sum([a, b, c, d]. * [1, 2, 0, 3]. * [0, 0, 1, 1]) = sum([0, 0, 0, 3d]) = 3d$
$Y = sum([a, b, c, d]. * [1, 2, 0, 3]. * [0, 1, 1, 0]) = sum([0, 2b, 0, 0]) = 2b$
$Z = sum([a, b, c, d]. * [1, 2, 0, 3]. * [1, 1, 0, 0]) = sum([a, 2b, 0, 0]) = a + 2b$

Therefore, the XYZ tristimulus values of R_3 *under* I_2 *is* $[X, Y, Z] = [3d, 2b, a + 2b]$. *In order to have it metameric with* R_1 *under* I_2*, these two should give the same color, i.e.* $[3d, 2b, a + 2b] = [9/4, 1, 3/2]$*, see part (d) above. This gives,*

$$\begin{cases} 3d = 9/4 \\ 2b = 1 \\ a + 2b = 3/2 \end{cases} \qquad \text{giving: } d = 3/4,\ b = 1/2 \text{ and } a = 1/2.$$

Notice that c is not involved in the equations and has no impact on the color of the reflectance under I_2 *and therefore can be any arbitrary value between 0 and 1. Therefore, as an example,* $R_3 = [a, b, c, d] = [1/2, 1/2, 0, 3/4]$ *is metameric with* $R_1 = [1/2, 1/2, 1/4, 3/4]$ *under* $I_2 = [1, 2, 0, 3]$.

Notice that similar equations as the ones in Equation 6.3 and Equation 6.6 can be written for a camera or a sensor. In this case, the $\bar{x}\bar{y}\bar{z}$ color matching functions in these equations are replaced by the camera's or the sensor's wavelength response (sensitivity functions).

Example 6.3 *For simplicity assume that the visible wavelength interval has been divided into four equal sub-intervals and therefore all spectral data can be represented by vectors having four elements. Assume that the vectors* s_1 *and* s_2 *below represent the two sensitivity functions for a sensor. The vectors* I *and* R *represent the intensity of a light source and the reflectance spectrum of an object, respectively. Assume that* S_1 *and* S_2 *denotes the sensor's response from* s_1 *and* s_2*, respectively.*

$s_1 = [0, 0, 1/2, 1]$, $s_2 = [1/4, 1/3, 0, 0]$, $I = [1, 0, 0, 1]$ *and* $R = [1, 1, 0, 1]$

a) *Find the normalization factor that gives* $S_2 = 1$ *for a completely white object under* I.
b) *Find the sensor's response for the reflectance* R *under* I.

c) *Write the relationships between the elements of two different reflectance spectra,* $R_1 = [a_1, b_1, c_1, d_1]$ *and* $R_2 = [a_2, b_2, c_2, d_2]$*, that make them metameric under* I *with regards to this sensor.*
d) *Give example of two reflectance spectra that are metameric with respect to the sensor under* I*.*

Solution (a): *Here, we can write a similar equation as Equation* 6.8*.*

$S_2 = k * sum(R_{white}. * I. * s_2) = k * sum([1,1,1,1]. * [1,0,0,1]. * [1/4, 1/3, 0, 0]) = \frac{k}{4}$*. Since according to the exercise we want* S_2 *to be 1, then* $k/4 = 1$ *giving* $k = 4$*.*

Solution (b):

$S_1 = k * sum(R. * I. * s_1) = 4 * sum([1,1,0,1]. * [1,0,0,1]. * [0,0,1/2,1]) = 4 * sum([0,0,0,1]) = 4$

$S_2 = k * sum(R. * I. * s_2) = 4 * sum([1,1,0,1]. * [1,0,0,1]. * [1/4,1/3,0,0]) = 4 * sum([1/4,0,0,0]) = 1$

Therefore, the sensor's response to R *under* I *is* $[S_1, S_2] = [4, 1]$*.*

Solution (c): *The sensor's response to* R_1 *under* I *is calculated by:*

$S_{1(R_1)} = 4 * sum([a_1, b_1, c_1, d_1]. * [1,0,0,1]. * [0,0,1/2,1]) = 4d_1$

$S_{2(R_1)} = 4 * sum([a_1, b_1, c_1, d_1]. * [1,0,0,1]. * [1/4,1/3,0,0]) = a_1$

giving $[S_{1(R_1)}, S_{2(R_1)}] = [4d_1, a_1]$*.*

By similar calculations, the sensor's response to R_2 *under* I *is found as* $[S_{1(R_2)}, S_{2(R_2)}] = [4d_2, a_2]$*. In order to have them metameric, they should give the same response meaning that* $4d_1 = 4d_2 \rightarrow d_1 = d_2$ *and* $a_1 = a_2$*. This means that, two different spectra having their first and last elements equal are metameric, no matter what their second and third elements are.*

Solution (d): *As discussed in part (c), two spectra having equal first and last elements are metameric under* I*. Just recall that the values in a reflectance spectrum has to be between 0 and 1. Therefore, for instance,* $[1/3, 0, 1, 2/3]$ *and* $[1/3, 1/5, 3/7, 2/3]$ *are metameric under* I *with respect to this sensor.*

6.4.2 Chromaticity

Sometimes, the relative difference between the X, Y, and Z values are of more interest than their actual values. Thus, sometimes the normalized versions, called chromaticity values, calculated by Equation 6.9 are used.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{X+Y+Z} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \tag{6.9}$$

where (X, Y, Z) and (x, y, z) denote the CIEXYZ tristimulus values and their corresponding (normalized) chromaticity values. Since $x + y + z = 1$, one of the chromaticity values can be calculated from the other two, and can thus be ignored. Usually z is ignored. The chromaticity values give a possibility to plot colors in a two-dimensional subspace, with x and y as coordinates. A plot of all monochromatic colors (pure colors of only one wavelength) in a chromaticity diagram is often referred to as the spectral locus. The spectral locus in the xy-plane has the form of a horse shoe, see Figure 6.8. Chromaticity diagrams are very common in color science literature, and are useful as long as one remembers that the colors exist in a three-dimensional space, and not two-dimensional as the chromaticity diagrams may imply.

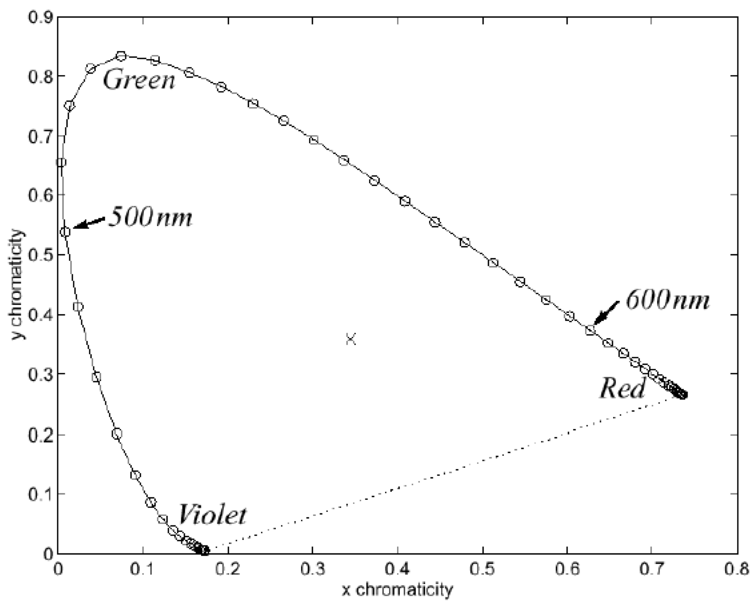


Figure 6.8: The spectral locus in a chromaticity diagram. The purple boundary is dotted. The white point of D50 is marked with an "x". The distance between the circles is 5 nm.

The dotted curve in Figure 6.8 is called the purple boundary. All colors inside this horse shoe curve, are the colors human eye can perceive. However, not all these colors can be reproduced by a specific display. Usually, the color gamut of a reproduction device is illustrated in a chromaticity diagram. For more detail, see Section 7.1.

Example 6.4 *The white point of four different CIE standard illuminants A, D_{50}, D_{55} and D_{65} are $XYZ_A = [109.85, 100.00, 35.58]$, $XYZ_{D50} = [96.42, 100.00, 82.51]$, $XYZ_{D55} = [95.68, 100.00, 92.14]$ and $XYZ_{D65} = [95.05, 100.00, 108.88]$, respectively. Find the chromaticity values for these four light sources and plot them in the chromaticity diagram and discuss the light sources' colors.*

Solution: *The chromaticity values are found using Equation 6.9, giving:*

$$[x_A, y_A] = [\frac{109.85}{109.85+100.00+35.58}, \frac{100.00}{109.85+100.00+35.58}] = [0.45, 0.41]$$

$$[x_{D50}, y_{D50}] = [\frac{96.42}{96.42+100.00+82.51}, \frac{100.00}{96.42+100.00+82.51}] = [0.35, 0.36]$$

$$[x_{D55}, y_{D55}] = [\frac{95.68}{95.68+100.00+92.14}, \frac{100.00}{95.68+100.00+92.14}] = [0.33, 0.35]$$

$$[x_{D65}, y_{D65}] = [\frac{95.05}{95.05+100.00+108.88}, \frac{100.00}{95.05+100.00+108.88}] = [0.31, 0.33]$$

Figure 6.9 shows the white points of these four illuminants.

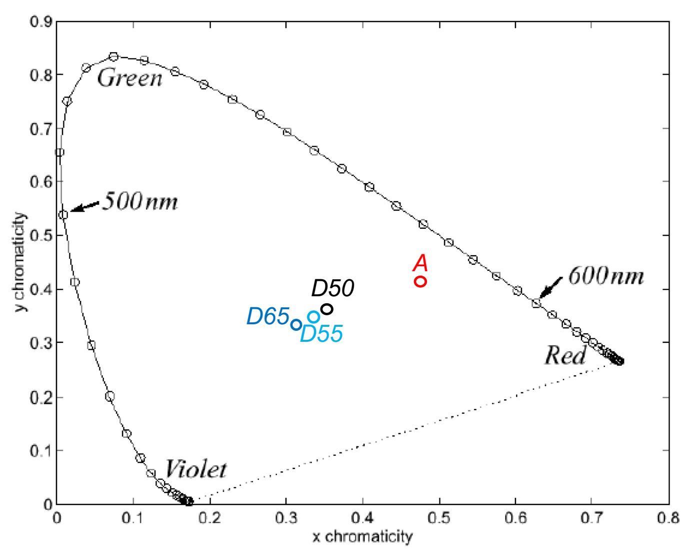


Figure 6.9: The white point of four CIE standard illuminants are plotted.

As can be seen in Figure 6.9, illuminant A is more "reddish-yellowish" than the other three, which can also be noticed by looking at A's spectral distribution shown in Figure 6.6. The other three illuminants are more or less neutral in color, representing day light, but the higher the temperature, the more bluish the light source is. For example, D65 that corresponds to a black body at approximately 6500 Kelvin is more bluish than D55 and D50 that correspond to approximately 5500 and 5000 Kelvin, respectively.

6.4.3 CIELAB color space

As discussed in Section 6.4.1, the CIEXYZ color space is not perceptually uniform, meaning that the Euclidean distance between colors' CIEXYZ coordinates does not correspond to the perceived color difference between them, impeding a quantitative comparison between colors. The CIE 1976 (L* a* b*) color space, usually written CIELAB, is derived from XYZ coordinates with aim on perceptual uniformity. Originally, it was developed to give the textile industry an accurate way to describe colors. Now, it serves as one of the most well-known device independent color spaces for all kinds of applications. The transformation between CIEXYZ values and CIELAB values is performed by:

$$L^* = \begin{cases} 116(\frac{Y}{Y_n})^{1/3} - 16, & \frac{Y}{Y_n} > 0.008856 \\ 903.3(\frac{Y}{Y_n}), & \frac{Y}{Y_n} \leq 0.008856 \end{cases}$$

$$a^* = 500(f(\frac{X}{X_n}) - f(\frac{Y}{Y_n}))$$

$$. \qquad (6.10)$$

$$b^* = 200(f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n}))$$

$$f(x) = \begin{cases} x^{1/3}, & x > 0.008856 \\ 7.787x + \frac{16}{116}, & x \leq 0.008856 \end{cases}$$

In Equation 6.10, X_n, Y_n and Z_n are the CIEXYZ values of the white point of the chosen light source. The L^* coordinate denotes lightness, where $L = 100$ is white stimulus and $L = 0$ means black stimulus. a^* is the red-green axis where positive values indicate red and negative values are green, and b^* is yellow-blue axis where positive values mean yellow and negative ones are blue. Figure 6.10 illustrates the CIELAB color space.

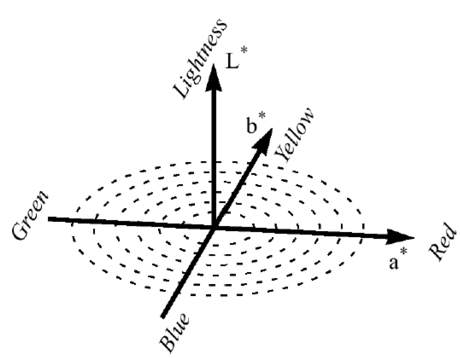


Figure 6.10: The CIELAB color space.

Let us illustrate the meaning of CIEL, CIEa and CIEb values by showing two sets of three different colors in Figure 6.11. In the top row, three different colors, having the same CIEa and CIEb, but different CIEL are illustrated. As can be seen, the hue of the three colors is the same, but their lightness is different. In the bottom row, on the other hand, the three colors have the same lightness (CIEL value) but their CIEa and CIEb are different.

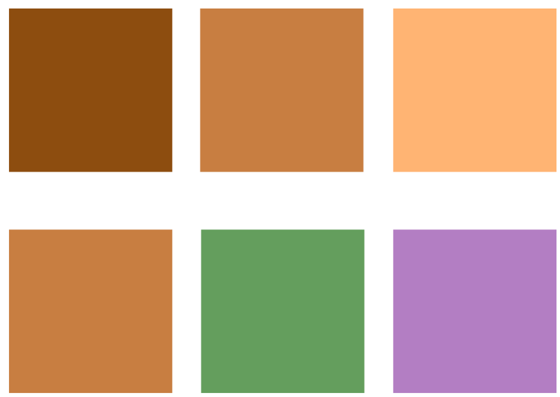


Figure 6.11: (Top) Same CIEa and CIEb-values, but different L. (Bottom) Same CIEL, but different CIEa and CIEb.

In order to find the CIELAB-values of a light source, we just need to set $[X, Y, Z] = [X_n, Y_n, Z_n]$ in Equation 6.10, which gives:

$L^* = 116(\frac{Y_n}{Y_n})^{1/3} - 16 = 116 - 16 = 100$

$a^* = 500(f(\frac{X_n}{X_n}) - f(\frac{Y_n}{Y_n})) = 500(f(1) - f(1)) = 0$

$b^* = 200(f(\frac{Y_n}{Y_n}) - f(\frac{Z_n}{Z_n})) = 200(f(1) - f(1)) = 0$

This means that, the CIELAB-values for any light source is always constant and equal to $[100, 0, 0]$, although their corresponding CIEXYZ values are different. This is another important difference between CIEXYZ and CIELAB color space.

Color difference formula

As discussed, CIELAB was created to serve as a perceptually uniform color space, in which the distance between colors could serve as a metric of their perceived difference. The CIE 1976 color difference ΔE_{ab} (Equation 6.11) is the simplest formula measuring the Euclidean distance between the coordinates of two colors in the CIELAB space,

$$\Delta E_{ab} = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2}, \tag{6.11}$$

where $[L_1, a_1, b_1]$ and $[L_2, a_2, b_2]$ are the CIELAB values of two colors.
Let us emphasize the importance of such a metric by an example. Assume that, you have a reference color with specific CIELAB-values (for example the color red in Coca-Cola) and a magazine wants to print a commercial advertisement for Coca-Cola. The printed red color has to be quite accurate to be accepted. This accuracy can be examined by using Equation 6.11 to calculate the color difference between the reference red color (given by Coca-Cola) and the measured printed red color of the advertisement in the magazine. How large ΔE_{ab} are accepted/tolerated, depends of course on the customer. If for example, Coca-Cola has a demand on not having a color difference larger than $2\Delta E_{ab}$, then the magazine has to fix it if the calculated color difference is greater than 2.
In theory, $\Delta E_{ab} = 1$ is used as the Just Noticeable Difference (JND), meaning that a color difference less than 1 means that the human eye cannot perceive any difference between the two colors. In practice, however, depending on the application, larger color differences are very well accepted. To summarize it, in some applications, $\Delta E_{ab} = 1$ to 3 is considered as good match, while $\Delta E_{ab} = 3$ to 6 are perceivable but usually accepted.
It has been found that even CIELAB color space is not perfectly perceptually uniform, and some perceptual non-linearity exists, specially around the blue area and low-chroma regions. Hence, besides the color difference formula in Equation 6.11, several other color difference formulas exist, which are out of the scope of this course and not discussed in this course book. The interested reader is referred to CIE 1994 and CIE 2000 color difference formulas.

Example 6.5 *A reference color is given by its CIELAB values as* $CIELAB = [85, 43, 65]$*. This color was created in Photoshop and printed. The CIELAB values of the printed color was measured* $[83, 44, 67]$*. Calculate the color difference and discuss whether or not this difference is acceptable.*

Solution: *The color difference between these two colors is calculated by Equation 6.11,*

$$\Delta E_{ab} = \sqrt{(85-83)^2 + (43-44)^2 + (65-67)^2} = 3$$

Whether or not this color difference is acceptable is very much dependent on the application and how accurate one wants the reproduction to be. However, a color difference of 3 is usually considered as a good and accepted color match in many practical applications.

6.4.4 RGB and CMY color spaces

The two color spaces discussed so far, namely CIEXYZ and CIELAB, are both based on the human visual system and therefore independent of the reproduction device. However, a reproduction device needs its own color space to reproduce a color. For example, a TV or a computer screen, reproduces a color by adding different shades of Red, Green and Blue. Such a device, uses the RGB color space. A color in this space is defined by giving its R, G and B value. Usually, each value is given by an integer between 0 and 255 (using 1 byte). For example, the color $[R, G, B] = [255, 0, 0]$ means that the red light source in the reproduction device emits full energy while its green and blue lights are off. Figure 6.12 shows nine different colors together with their RGB-values.

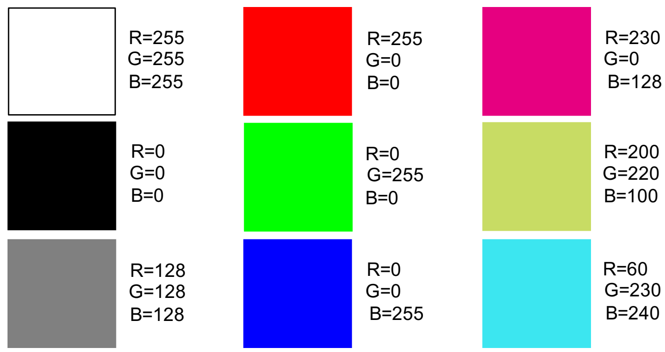


Figure 6.12: Nine different colors in RGB color space together with their RGB-values.

A printer, on the other hand, usually uses the complementary colors to red, green and blue, i.e. Cyan, Magenta and Yellow, as the primary colors. Any color is then reproduced by mixing different amount of these primary colors in a subtractive way. The color in this space is defined by giving its C, M and Y value. Each value in this space is given by a percentage number between 0% and 100%, specifying the coverage of the specific primary ink to be printed. For example, the color $[C, M, Y] = [100\%, 0\%, 0\%]$ means that the cyan's coverage is 100%, i.e. the cyan ink covers the whole surface, while magenta and yellow are not printed at all. Figure 6.13 shows nine different colors together with their CMY-values.

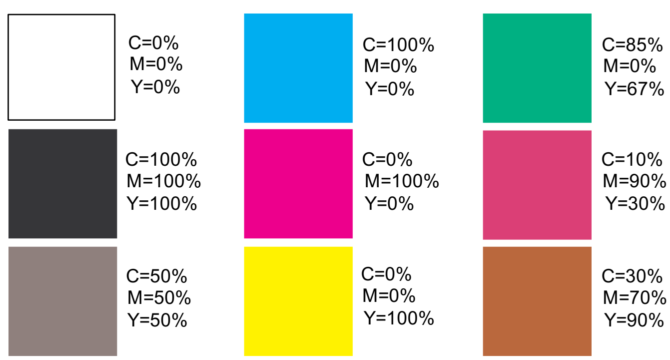


Figure 6.13: Nine different colors in CMY color space together with their CMY-values.

The RGB and CMY color spaces together with the additive and subtractive color mixing are further discussed in Section 7.2.

6.5 Color management

In a color reproduction process, there are usually different reproduction devices, such as scanners, cameras, monitors, printers etc., involved. It is therefore of crucial importance to have control over the conversion between the color representations of various devices. The primary goal of color management is to obtain a good match across color devices; for example, the colors of an image taken by a camera should appear the same on a computer monitor and as a printed poster. Color management helps to achieve the same appearance on all of these devices, as much as possible. Since different devices have different characteristics, having exact match is most of the time impossible. The color management is to control the process to make the reproduced colors match as much as possible. In this section, we briefly introduce color management system without going into the details. Let us explain it by an example. Figure 6.14 shows a monitor displaying an image (green color), which is sent to a printer that is supposed to reproduce the same color.

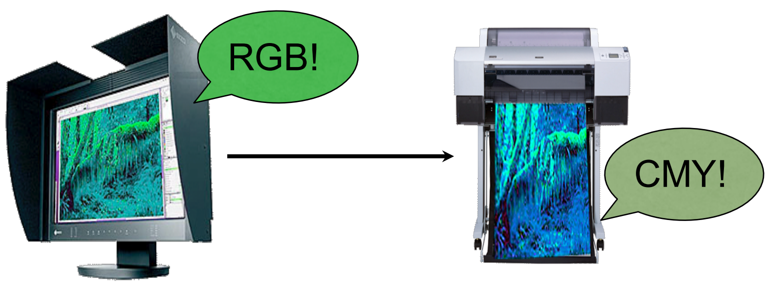


Figure 6.14: The color reproduced by the printer is supposed to match the one on the monitor.

Assume that the specific green color shown on the monitor has the RGB-values $[R, G, B] = [20, 255, 40]$. First of all, notice that, since the RGB color space is device dependent, exactly the same RGB-value might be reproduced differently by another monitor or TV-screen. The reason is that different devices have their own RGB light sources (lamps) with different characteristics. Notice also that, the printer uses another device dependent color space, i.e. CMY. In order to make these different reproduction devices communicate to make the reproduced colors match as much as possible, there is a need to use a device independent color space, such as CIELAB. One of the most used approaches to color management is to use the International Color Consortium (ICC) profiles. Briefly explained, an ICC profile is a set of data that characterizes a color input or output device, or a color space, according to standards promulgated by the ICC. For instance, in this example, the ICC-profile of the monitor and the printer, could be used to convert the RGB-values (or CMY-values) to CIELAB-values, and vice versa. Assume that this specific green color with $[R, G, B] = [20, 255, 40]$ is converted to CIELAB-value $[L, a, b] = [88, -85, 50]$ by the monitor's ICC-profile. The same color is now supposed to be reproduced by

the printer. By using the printer's ICC-profile, one could convert these CIELAB-values to the printer's CMY-values. If now, the color $[L, a, b] = [88, -85, 50]$ is inside the colors that the printer is able to reproduce, it is simply converted to its corresponding CMY-values. If it is not, then the color management will find the closest color that can be reproduced by the printer.
Figure 6.15 shows colors reproduced by three different devices, two printers and one monitor. The three reproduced colors have the same color values in the CIELAB color space, i.e. $[L, a, b] = [88, -17, 42]$ and therefore they look the same. However, when the same color is converted to the device dependent color values by the devices' ICC-profiles, they are converted to different device dependent color values as seen in this figure.

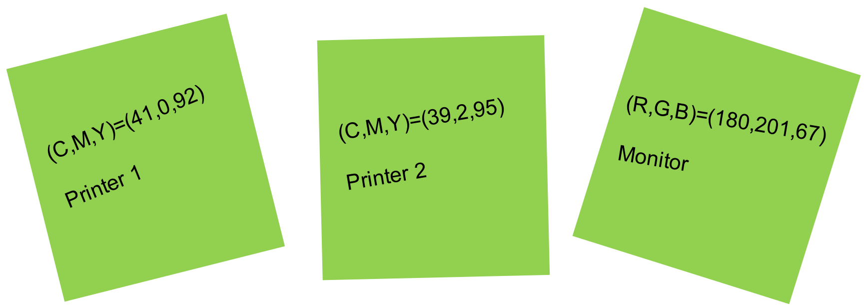


Figure 6.15: The three colors have different color values in their own device dependent color spaces but have the same color values in a device independent color space, therefore they give the same color impression.

Notice that, this section just gave a short introduction on color management and ICC-profile and was not intended to give a detailed description of color management systems and how they work. The goal here has been to emphasize the importance of device independent color spaces such as CIELAB.

6.6 Exercise set

6.1. We divide the visible wavelength interval, i.e. $[350, 750]$ nm, into four equal sub-intervals: $[350, 450]$, $[451, 550]$, $[551, 650]$ and $[651, 750]$. Therefore, all spectral data can be represented by vectors having four elements. We describe a sensor's characteristic with a 2×4 matrix as:

$$S = \begin{bmatrix} 0.5 & 1 & 0.5 & 0 \\ 0 & 0.5 & 1 & 0.5 \end{bmatrix}$$

This means that the sensor has two receptors (rows in S). Answer the following questions. There is no need for normalization.

a) Consider a light source L that produces photons with wavelength at 500 nm. Describe the light source L with a vector.

b) The photons produced by the light source L hit the sensor S. The output of S is a vector C. Find C!

c) Construct another light source L_1 that gives the same output as the one in part (b) when hitting the sensor S.

d) Find all light sources L_m that produce the same output as the one in part (c) when hitting the sensor S .

6.2. a) Use the light source vector $L = [0, 1, 2, 3, 4]$ and the reflectance vector $R_1 = [0, 1/3, 2/3, 1/3, 0]$ for an object. Calculate the spectral distribution S that is reflected from the object under the light source L.

b) Figure 6.16 shows the three color matching functions and their rough approximations.

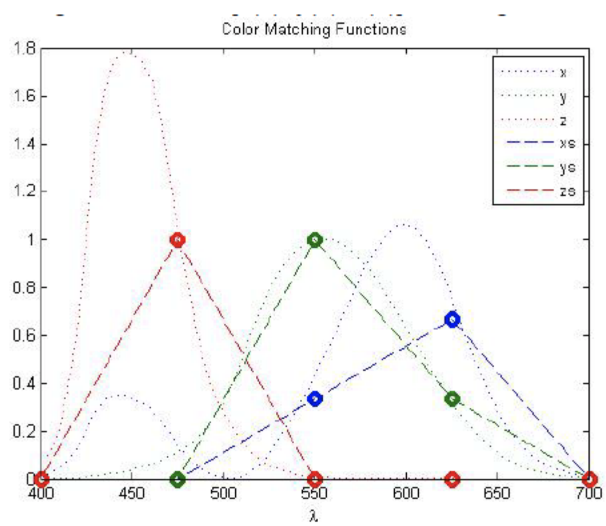


Figure 6.16: Exercise 6.2: The three color matching functions and their rough approximations.

The three approximated functions are described by the following matrix, where row 1, 2 and 3, represent $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$, respectively.

$$\begin{bmatrix} 0 & 0 & 1/3 & 2/3 & 0 \\ 0 & 0 & 1 & 1/3 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Calculate the XYZ-values for the spectral distribution S in part (a). No normalization is needed.

c) Find all reflectance spectra that are metameric with R_1 under L.

6.3. Figure 6.17 shows three sensitivity functions of a camera. As seen in this figure, there are seven marked measured spots, from 400 to 700 nm with a step of 50 nm, i.e. [400:50:700] nm.

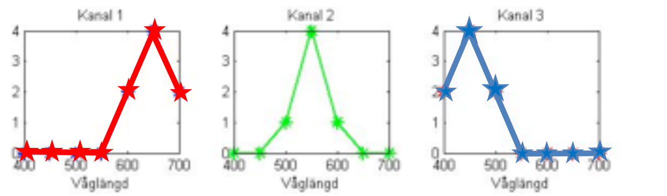


Figure 6.17: Exercise 6.3: The three sensitivity functions of a camera.

a) Describe camera's sensitivity functions with a 3×7 matrix C.

b) The vector $Q = [2, 1, 0, 1, 0, 2, 4]$ describes the reflectance spectrum of an object. Assume that all incoming light at 700 nm is reflected from this object. With what factor should we multiply Q to get the correct reflectance spectrum R? Calculate R!

c) Use the camera matrix C from part (a) and the reflectance spectrum R from part (b). Choose a light source L describing a perfect white illuminant. Calculate camera's response for the reflectance spectrum R under L. No normalization is needed.

6.4. We use the following notations:

λ: wavelength
$r(\lambda)$: a reflectance spectrum
$I(\lambda)$: spectral power of a lamp
$s(\lambda)$: the sensitivity function of a sensor

Equation 6.12 describes the relationship between the light source $I(\lambda)$, an object reflectance $r(\lambda)$, sensor sensitivity $s(\lambda)$ and a pixel value (sensor response) p.

$$p = \int_\lambda I(\lambda)r(\lambda)s(\lambda)d\lambda \tag{6.12}$$

a) Describe as a formula the following reflectance functions:

i) A mirror that reflects all incoming light

ii) A black hole that absorbs all incoming light

b) Besides the lamp $I(\lambda)$, we add an extra lamp with the power distribution $k(\lambda)$. How does Equation 6.12 change if we use both lamps at the same time?

c) Assume that we still use the same lamp $I(\lambda)$ and add an optical filter in front of the photographic objective of the camera. The filter is characterized by the function $f(\lambda)$, where $f(\lambda_0) = 1$ means that all light at the wavelength λ_0 passes through the filter and $f(\lambda_0) = 0$ means that all light at the wavelength λ_0 is absorbed (blocked) by the filter. Modify Equation 6.12 so that the modified equation takes into account the effect of the optical filter.

6.5. A two-channel camera is described by the following matrix,

$$M = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

A light source is described by $L = [1, 2, 3, 4, 5]$ and the reflectance spectrum of an object is $R = [1, 0, 1, 1, 2]/2$.

a) Calculate the camera response represented by the matrix M for the object reflectance R under the light source L.

b) Why do we need a scaling factor $1/2$ for the vector R?

c) What is the largest scaling factor x if the vector $Q = [0, 1, 3, 0, 1] * x$ is supposed to represent a reflectance spectrum?

d) Use $V = [1, 1, 1, 1, 1]$ to represent gray light. Use the camera described by M and construct two reflectance spectra R_1 and R_2 that are metameric under L but not metameric under V.

6.6. Assume that the following three vectors describe the $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ color matching functions:

$$\bar{x} = [1, 1, 1, 5, 5, 2]$$
$$\bar{y} = [0, 1, 5, 6, 3, 0]$$
$$\bar{z} = [6, 8, 1, 0, 0, 0]$$

a) Calculate the CIEXYZ tristimulus values for a spectral distribution $S_1(\lambda) = [1, 0, 1, 0, 1, 0]$ (both reflectance spectrum and the light source are included). No normalization is needed.

b) Are there spectra of the form $S(\lambda) = [a, 0, 0, 0, b, c]$ that are metameric? If yes, describe them. If no, why? Notice again that both reflectance spectrum and the light source are included in S.

Chapter 7

Color reproduction

As discussed in Section 6.5, different devices handle color in different ways. A monitor or LCD uses light to create colors while a printer utilizes inks or dyes. In addition, devices use several different systems of creating colors. A monitor usually mixes red, green and blue light to produce a particular shade while a printer might use droplets of cyan, magenta, yellow and black ink to produce the same shade. Different devices also have different color reproduction capabilities. In this chapter, our focus will mostly be on color reproduction in print. The chapter starts with explaining color gamut, which defines the set of colors that can be reproduced by a certain device. This is followed by a discussion on color mixing strategies. Color halftoning, which is a very important part of color reproduction in print, is also addressed and discussed in this chapter. A number of printing strategies to print the ink droplets as well as a number of formulas predicting the color of a print production are also brought up. At the end of this chapter, a number of exercises are provided, for which you can find the answers and the solutions in Chapter 8 and 9, respectively.

7.1 Color gamut

Color gamut defines the scope of reproducible colors of a device. As discussed in Section 6.4.2, the color gamut of a device is usually illustrated in a chromaticity diagram. The triangle in Figure 7.1 illustrates the color gamut of a typical monitor. It is noticeable that the color gamut of the display is much smaller than the spectral locus, i.e. the colors the human eye can see. This CRT-display is especially bad on saturated green colors.

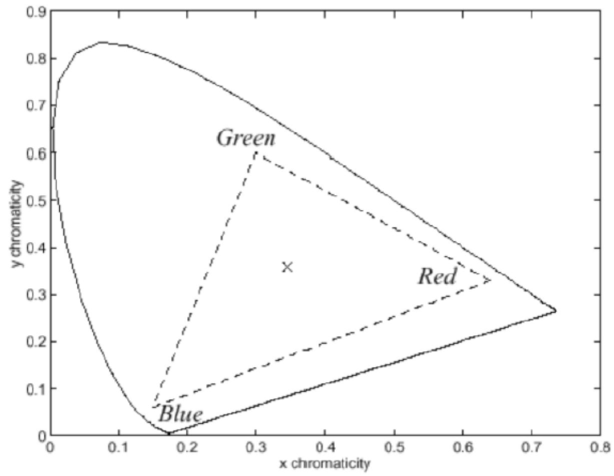


Figure 7.1: The color gamut of a typical computer monitor. The corners of the triangle are the chromaticity coordinates for red, green, and blue phosphor used in the monitor.

Figure 7.2 shows a comparison between two different color gamuts, namely sRGB and AdobeRGB. As seen in this figure, AdobeRGB has a bigger color gamut, and is commonly used in print shops. Not so many computer monitors can reproduce all these colors. sRGB, on the other hand, is standard for daily use and represents a color gamut that a usual PC monitor can show.

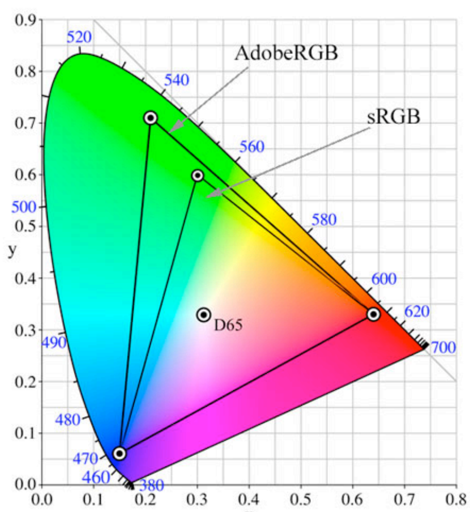


Figure 7.2: The comparison between AdobeRGB and sRGB color gamuts.

As discussed, two different monitors may have two different color gamuts.

This is also valid for different printers. There are many factors that affect the color gamut of a print setup. The most important ones are, the technology used (offset, digital, etc.), the characteristic of the ink, the type of the paper substrate (coated, uncoated, matte, etc.). Typical printers generally have smaller color gamut than typical monitors. Figure 7.3 (left) shows an example of two color gamuts, one for a monitor and one for a printer. As can be seen in this figure, the color gamut of the printer is smaller than that of the monitor. This means, there are many colors that can be reproduced by the monitor, but are impossible to be correctly reproduced by the printer, specially green colors. The opposite is only valid for very few colors. Figure 7.3 (right) shows the difference between the same colors being reproduced by the monitor and the printer.

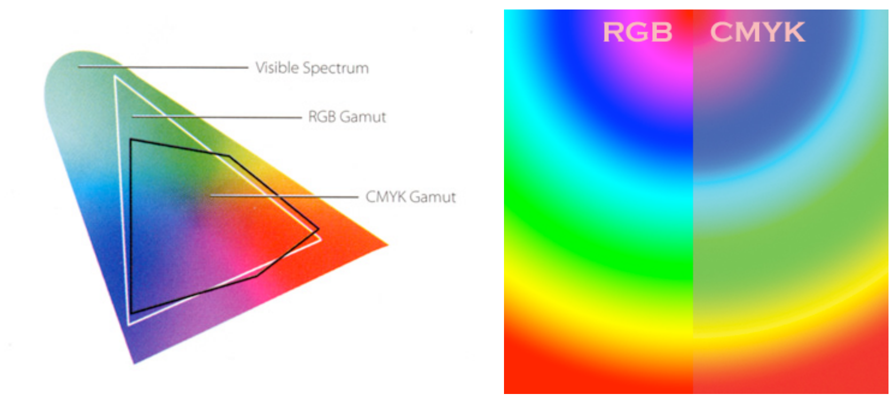


Figure 7.3: The comparison between the color gamut of a monitor and a printer.

In a reproduction chain, starting from a digital camera (or scanner) and ending in the print products, there are a number of devices involved, for example, camera/scanner, computer screens, printers etc. As discussed in Section 6.5, it is very important to have a module to control the color throughout the reproduction and reproduce the colors as accurately as possible. If there are colors that our outside the color gamut of a specific device, then there are different methods, called gamut mapping methods, to mapp them into the color gamut and reproduce them by the best possible colors. Gamut mapping is, however, out of the scope of this course and therefore not discussed in this book.

7.2 Color mixing

In Section 6.4.4, two device dependent color spaces, i.e. RGB and CMY, were very briefly discussed. The RGB and CMY color spaces use two different color mixing strategies, described in the two following subsections.

7.2.1 Additive color mixing

The idea of reproducing the full range of colors by mixing three lights with different color bands lead to the principles of some of today's color reproduction systems, such as computer monitors, TV screen, projectors, etc. The three colored lights were chosen so that their wavelengths closely matched the light susceptibility wavelengths of the three different types of cones. Therefore, red, green, and blue are the most common primary colors used in additive color spaces. The additive mixture of these three lights at their maxima renders white, labeling it additive color mixing, see Figure 7.4 (left). The combination of two of the standard three additive primary colors in equal proportions produces an additive secondary color – cyan, magenta or yellow, which, in the form of dyes or pigments, are the standard primary colors in subtractive color spaces, discussed in the following subsection.

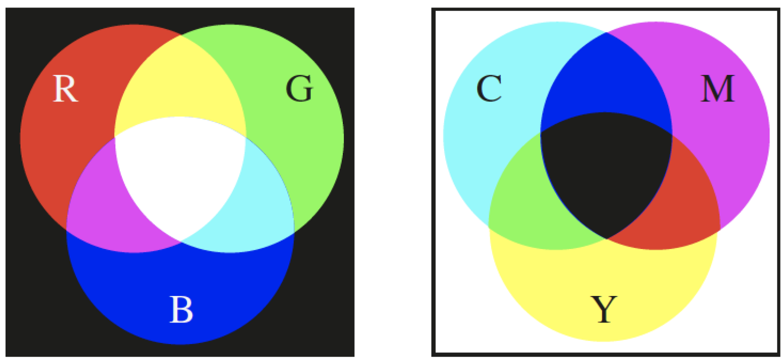


Figure 7.4: Additive (left) and subtractive (right) color mixing.

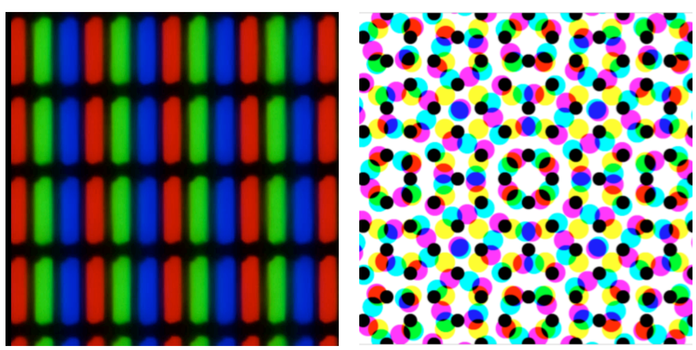


Figure 7.5: A close-up of an LCD (left) and a printer (right).

Figure 7.5 (left) shows a close-up of an LCD screen. As seen in this figure, the display is composed of red, green and blue sub-pixels, the light from which combines in various proportions to produce all the other colors as well as white and shades of gray. The colored sub-pixels do not overlap on the screen, but when viewed from a normal distance they overlap and blend on the eye's retina, producing the same result as external superimposition.

7.2.2 Subtractive color mixing

Contrary to photon emittance, color sensation could also originate from photons reflected from an object that is illuminated by a light source, e.g. when mixing ink pigments in printing reproduction. A different color mixing model is then used to reproduce colors, employing red's, green's and blue's complementary colors – cyan (C), magenta (M) and yellow (Y) – as primaries. Contrary to additive color mixing, the lack of the three primary colors creates the sensation of white (assuming a white substrate) and this model is thus called subtractive color mixing, see Figure 7.4 (right). For example, the amount of cyan (complementary to red) pigment applied to the paper will control the amount of red in the white light that will be absorbed by the ink. By applying 100% cyan coverage, in theory no red will be reflected, see Figure 7.6 (left). The same is valid between magenta and its complementary color green, Figure 7.6 (middle), and between yellow and its complementary color blue, Figure 7.6 (right). By applying 100% ink coverage of cyan, 100% magenta and 100% yellow, in theory, all light will be absorbed and thus the sensation of black will be achieved.

Figure 7.6: Each primary ink absorbs its complementary color in the white light and reflects the rest.

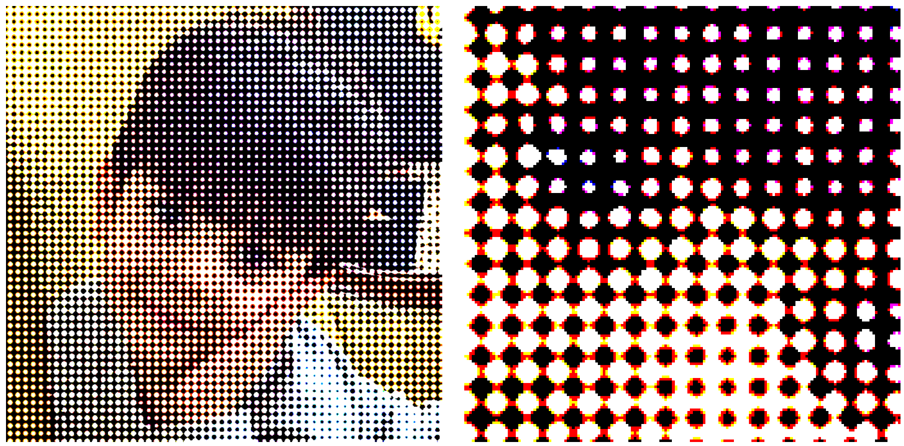
Figure 7.5 (right) shows a close-up of a print. As seen in this figure, the print is composed of cyan, magenta, yellow (and black) halftone dots, which are combined in various coverage to produce all the other colors. The colored dots partly overlap and partly not, but since the dots are usually small enough, when viewed from a normal distance they produce a continuous tone impression. It is worth mentioning that, as also seen in Figure 7.5 (right), black (K) is also added as the primary color. There are a few practical reasons for that. First of all, although cyan, magenta and yellow on top of each other, in theory, results in black, in practice they rather result in dark brown, which might lower the contrast of the print. Another reason is that, three inks have to be printed on top of each other to make a black dot, which is much more expensive than only using one single black ink instead. Add to these, the fact that having three layers of inks on top of each other might cause smearing leading to larger dot gain and unsharp edges. Therefore, CMYK printing, also called four-channel printing, is commonly used in print technologies.

7.3 Color halftoning

Digital halftoning was introduced and thoroughly explained in Chapter 3. The halftoning methods described in that chapter were all achromatic, meaning that they were only applied to achromatic images. A color image is usually defined by a number of channels, each of which can be seen as an achromatic image. For example, a color image that is going to be reproduced by an additive device usually consists of three color channels, red, green and blue. A color image that is going to be printed, has to be represented by the channels that the printer uses. In four-color printing, the color image has to be represented by its cyan, magenta, yellow and black color channels. In order to halftone such an image, usually an achromatic halftoning method is applied separately to each of these channels. Then, these halftoned channels are printed in separate printing units, each using cyan, magenta, yellow or black ink. The concept of achromatic AM and FM halftoning was introduced and described in Section 3.3. In the present section, color halftoning is discussed by describing AM and FM color halftoning in the following two subsections.

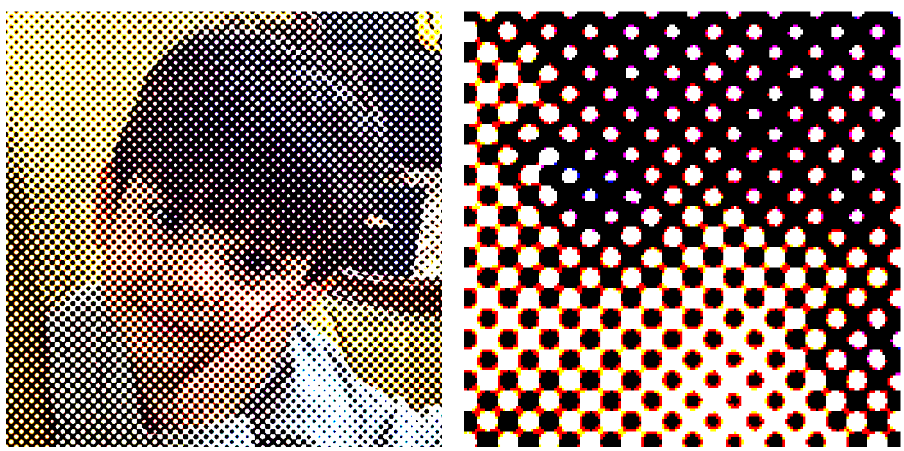
7.3.1 AM color halftoning

Recall from Section 3.3 that, in AM halftoning the size of the dots is variable while their frequency is constant. The number of halftone dots (or halftone cells) per an inch is called screen frequency (*lpi*) specifying how small/big the halftone dots are, see Figure 3.2- 3.4. The higher the *lpi*, the smaller the halftone dots and the more difficult to be detected by the eye. A color image, can simply be AM halftoned by applying achromatic AM halftoning to the image's color channels separately and then putting them together to make a halftoned color image. Figure 7.7 shows the test color image being halftoned by an AM method. The same method (i.e. same *lpi*) has been applied to all channels. As can be seen in this figure, the same screen angle, i.e. $0°$, has been used for all channels. Figure 7.8 shows the test color image being halftoned by an AM method using the same *lpi* as the images in Figure 7.7. The difference is, as seen in these figures, that the screen angle $45°$ has been used for all channels in Figure 7.8.
As long as the print press is stable and there is no miss-registration, i.e. the printed dots are printed at the same position they are supposed to, using the same screen angle for all channels, will not introduce any problem. However, in practice, there are always some kind of miss-registration, specially in conventional printing technologies such as offset, and consequently using the same screen angle can cause color shift and unwanted Moiré pattern.



(a) (b)

Figure 7.7: (a) The image is halftoned by AM color halftoning. The screen angle for all channels is 0°*. (b) Enlargement of a part of (a).*



(a) (b)

Figure 7.8: (a) The image is halftoned by AM color halftoning. The screen angle for all channels is 45°*. (b) Enlargement of a part of (a).*

Figure 7.9 illustrates the problems. The image shown in Figure 7.9 (a) is the image that is supposed to be printed. If we now have some positioning error, that means the printed dots (cyan dots in this example) are shifted in position, then it could lead to the case shown in Figure 7.9 (b), which is reproduced as a different color. Figure 7.9 (c) illustrates another problem, the visible and disturbing Moiré pattern, which is because of the screen angle error in one of the channels.

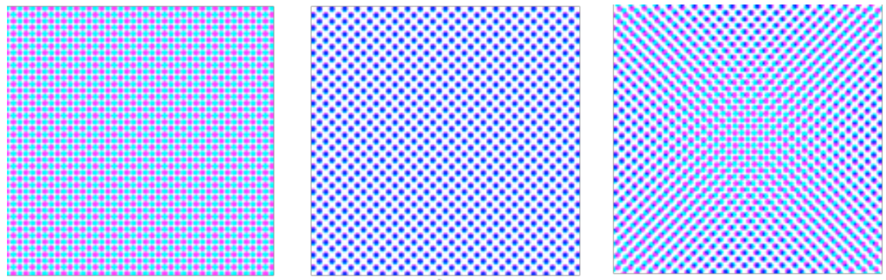
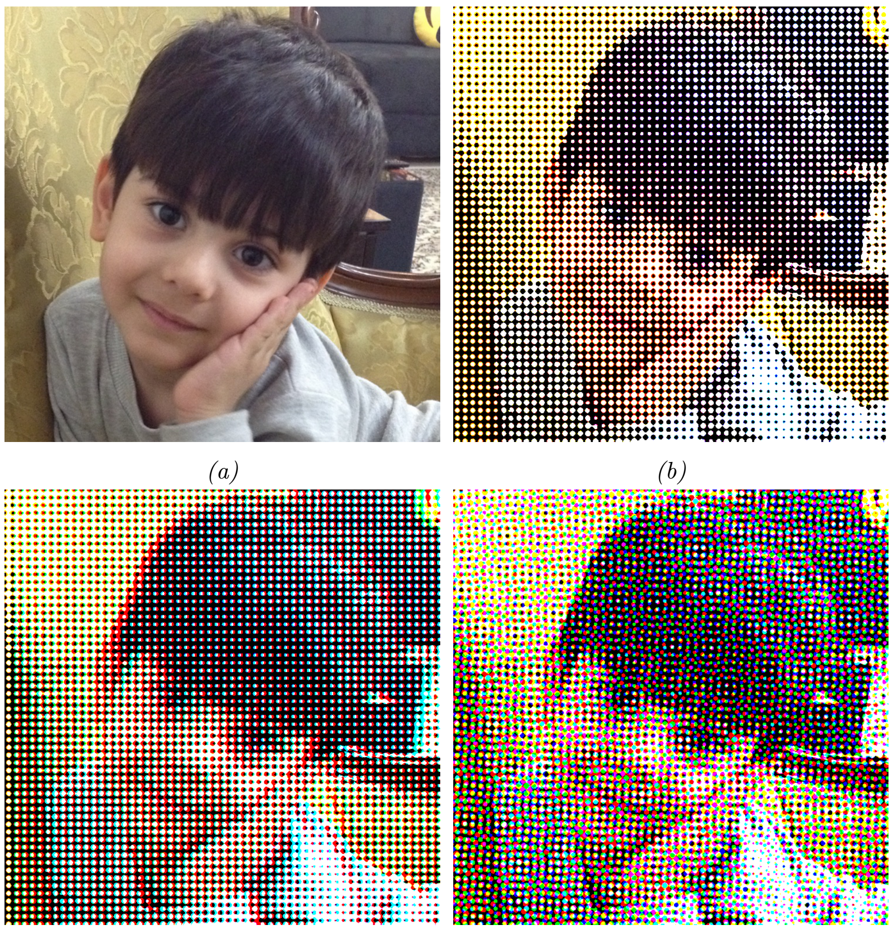


Figure 7.9: The positioning and angle error may cause color shift and unwanted Moiré pattern. (a) Correct registration. (b) Miss-registration in position causing color shift. (c) Miss-registration in angle causing Moiré.

In Figure 7.10, a simulation of what position and angle error may cause is illustrated using a regular test image. The image in (a) is the original and in (b) all channels have been halftoned at 0° and no miss-registration is simulated. As can be seen, the colors in this image looks similar to those in the original. In (c), a position error has been simulated, which shows a color shift compared to the original, look for instance at the upper right corner. In (d), an angle error has been simulated which causes visible and disturbing Moiré pattern.
The reason for miss-registration causing these problems, is that in the AM halftoning the halftone structure is periodic. If, now, the color channels are all halftoned at the same screen angle, the halftone dots in different halftoned channels will be on top of each other, causing the color halftone structure to be highly periodic (structured) as well. Now, any small shift (position error) in any of the channels might cause a big color shift (Figure 7.10 (c)), or any change in screen angle in any of the halftoned channels might cause distortions in form of Moiré pattern (Figure 7.10 (d)). In order to reduce the effect of miss-registration, the periodic structure of the color halftones has to be minimized. In practice, four different screen angles are used for Cyan, Magenta, Yellow and Black color channels. Since the human eye is less sensitive for raster at 45 degrees, the strongest color, Black (assuming a white substrate), is halftoned at this angle. The weakest color, Yellow, is printed at 0 degrees, where the human eye is most sensitive. The other two colors are placed in between making the same angle to Yellow and Black. Therefore, Cyan and Magenta are commonly halftoned at 15 and 75 degrees, respectively. Figure 7.11 shows the test image being AM-halftoned with different screen angles being used in C, M, Y and K channel. An enlargement of a part of this image is also shown.
Although using different angles for different channels reduce the effect of miss-registration, it introduces a new type of patterns, Rosette patterns, which are quite visible at lower screen frequencies, for example see Figure 7.11 (a). Two different types of Rosette patterns that may occur are shown in Figure 7.12. In (a) there is a dot at the center of the rosette while in the image shown in (b) the center is empty. It is worth noticing that, the rosette patterns usually

occur at much higher frequencies than Moiré pattern, and therefore if the screen frequency (*lpi*) is high enough, the rosette patterns are not visible, while Moiré patterns are visible. Therefore, the screen angles discussed above are commonly used in the four-color printing using AM color halftoning, if miss-registration is unavoidable.

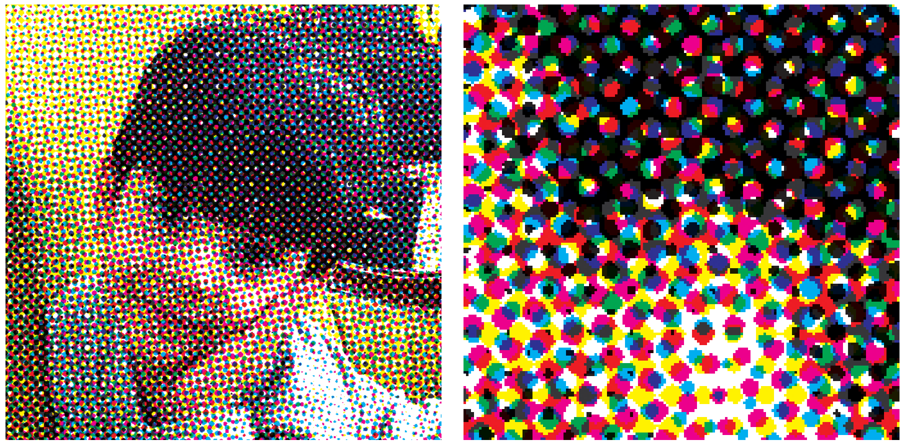


(a) (b) (c) (d)

Figure 7.10: (a) The original test image. (b) The test image is AM halftoned, the screen angle is 0° in all channels and no miss-registration has occurred. (c) The position error causing color shift is simulated. (d) The angle error causing Moiré is simulated.

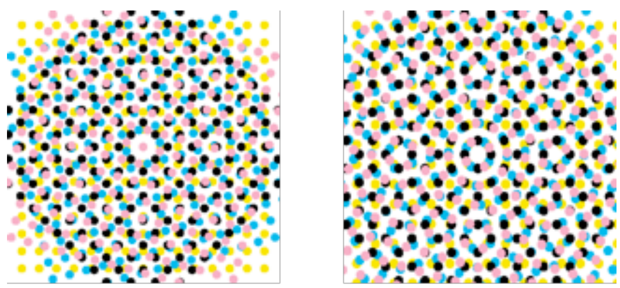
7.3.2 FM color halftoning

In Section 3.3, achromatic FM halftoning was thoroughly explained. Like AM halftoning of color images, in FM color halftoning the achromatic FM halftoning



(a) *(b)*

Figure 7.11: (a) The test image is AM-halftoned. The screen angle in the four channels (C, M, Y and K) is 15, 75, 0 and 45 degrees, respectively. (b) An enlargement of a part of (a).



a b

Figure 7.12: Two different types of Rosette patterns that may occur when different screen angles are used in AM halftoning. (a) there is a dot at the center of the rosette. (b) the center is empty.

is applied to the color channels that the image is divided into. Usually, the color channels are halftoned independently. Figure 7.13 shows two FM-halftoned color images where the channels (C, M and Y) are halftoned independently. In Figure 7.13 (a), the image has been halftoned by non-modified error diffusion using Floyd and Steinberg's error filter discussed in Section 3.2.3. In Figure 7.13 (b), the image has been halftoned by the iterative halftoning method IMCDP, discussed in Section 4.1.1. Enlargement of a part of these images are also shown. Since FM halftoning usually results in non-periodic dispersed halftone structures, FM halftoning is not very sensitive to miss-registration and therefore miss-registration doesn't cause undesirable Moiré pattern in FM color halftoning as it does in AM if the same screen angle is used in all channels.

Figure 7.13: The original test image has been FM halftoned by: (a) non-modified error diffusion using Floyd and Steinberg's error filter. (b) the iterative halftoning method, IMCDP . (c) Enlargement of a part of (a). (d) Enlargement of a part of (b).

In recent years, researchers have been looking at the possibility of halftoning different color channels in a dependent manner in order to reduce the color noise. It has also been shown that dependent color halftoning not only increases the print quality but also reduces the amount of ink needed to print the same image. This is, however, out of the scope of this course and therefore not discussed in this book.

7.4 Neugebauer equations

As discussed earlier, a device reproduces different colors by using a number of primary colors. For example, a printer usually uses C, M, Y and K primary inks. The question is now how to calculate (approximate) the resulting color when these primary colors are mixed. When several small different colored areas are averaged together by the human eye, a simple model, Neugebauer equations, can be used to find the resulting (average) spectral reflectance of a surface by,

$$R_{ave}(\lambda) = \sum_i a_i R_{i,max}(\lambda), \tag{7.1}$$

where i are the so-called Neugebauer Primaries (NPs): the substrate with no ink, full-tone single ink, and ink overlap combinations (with full coverage), summing up to a total of 2^n NPs, n being the number of the primary inks. For instance, for a three colorant example (CMY), there are three primary inks, giving $2^3 = 8$ NPs, which are: 1. white (i.e. substrate without ink), 2. cyan, 3. magenta, 4. yellow, 5. red (magenta + yellow), 6. green (cyan + yellow), 7. blue (cyan + magenta) and 8. black (cyan + magenta + yellow). $R_{i,max}(\lambda)$ in Equation 7.1 is the spectral reflectance of each NP at full coverage, and a_i is the corresponding fractional ink area coverage, which will sum to 100% or 1, i.e. $\sum_i a_i = 1$. For a four colorant example (CMYK), there will be $2^4 = 16$ NPs. Since the relationship between the CIEXYZ tristimulus values of an object and its spectral reflectance is linear (see Equation 6.3 again), the Neugebauer equations in Equation 7.1 can be rewritten as,

$$\begin{bmatrix} X_{ave} \\ Y_{ave} \\ Z_{ave} \end{bmatrix} = \sum_i a_i \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix}, \tag{7.2}$$

where $\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix}$ denotes the CIEXYZ values of the NPs at full-tone and a_i is the corresponding fractional coverage. $\begin{bmatrix} X_{ave} \\ Y_{ave} \\ Z_{ave} \end{bmatrix}$ is the average CIEXYZ values of the printed surface. Notice again that since the area of the surface is covered by the NPs then the sum of NPs' coverage must be 100% or 1, meaning that $\sum_i a_i = 1$.

It has to be pointed out here that Equation 7.1 (and 7.2) describe a simple model that is not directly applicable to all practical print situations. First of all, in these equations the effect of optical dot gain is ignored (for optical dot gain see Section 3.4.2). Secondly, in this equation it is assumed that a_i denotes the real physical coverage of NP i after print and not the coverage in the digital bitmap. Thirdly, it is also assumed that the inks are not penetrating into each other, or paper. In many practical cases, there is always some ink penetration and optical dot gain cannot be ignored. In these cases, this simple model does not apply and more complicated models should be used. However, in this book we only focus on this simple model and a_i is assumed to be the coverage in the digital bitmap (or the halftone).

Example 7.1 *Calculate the average CIEXYZ values for three different print surfaces in Figure 7.14 (a), (b) and (c). All three prints only include cyan and magenta primary inks and in all of them cyan and magenta cover 40% and 30% of the area of the surface, respectively. In surface (a), as seen in the figure, there is no dot-overlap. In surface (b), 10% of the surface is covered by both cyan and magenta, i.e. 10% dot overlap. In surface (c), there is a complete dot overlap, as seen in Figure (c). Table 7.1 shows the CIEXYZ values of the four NPs.*

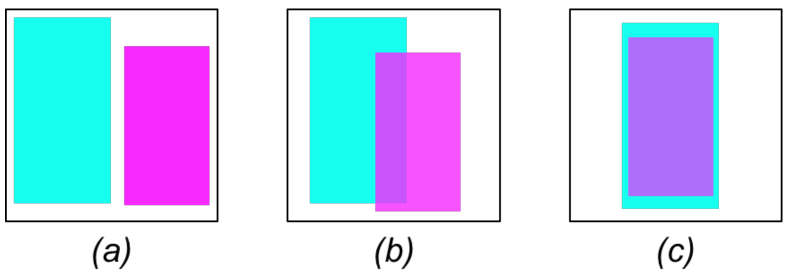


Figure 7.14: Example 7.1: three printed surfaces consisting of cyan and magenta primary inks.

Table 7.1: Example7.1: CIEXYZ values of the four NPs.

	Paper	Cyan	Magenta	Blue
X	90	50	60	40
Y	100	80	30	50
Z	110	100	100	100

Solution surface (a): *First of all, notice that there are only two primary inks, i.e. cyan and magenta, involved, resulting in $2^2 = 4$ NPs. The NPs are, paper (or no ink), only cyan, only magenta and cyan on magenta (i.e. blue). In this surface, there is no dot overlap, meaning that the coverage of blue (cyan on magenta) is 0. The coverage of pure cyan (only cyan) and pure magenta (only magenta) is 40% and 30%, respectively, as mentioned in the example. The rest of the surface is the paper, i.e. no ink, meaning that the fractional coverage of the paper is (100% - (40+30)% = 30%). To summarize, the coverage of the paper, only (pure) cyan, only (pure) magenta and blue (cyan + magenta) is 0.3, 0.4, 0.3 and 0, respectively. Notice that the sum of the coverages is 1. Equation 7.2 for surface (a) gives,*

$$\begin{bmatrix} X_{surface(a)} \\ Y_{surface(a)} \\ Z_{surface(a)} \end{bmatrix} = 0.3 \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + 0.4 \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} + 0.3 \begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} + 0 \begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix},$$

where indices p, c, m and b refer to paper, cyan, magenta and blue, respectively. The CIEXYZ values of the NPs are given in Table 7.1. Setting them in the equation above gives:

$$\begin{bmatrix} X_{surface(a)} \\ Y_{surface(a)} \\ Z_{surface(a)} \end{bmatrix} = 0.3 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.4 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

Calculating this gives the CIEXYZ values for surface (a): $[65, 71, 103]$.

Solution surface (b): *In this surface, according to the example, there is 10% dot overlap, meaning that 10% of the surface is covered by blue. The coverage of cyan is 40%, but 10% of it is overlapped meaning that the coverage of pure cyan is 30%. With the same reasoning, the coverage of pure magenta is 20%. The rest of the surface is the paper, i.e. paper covers (*$100-30-20-10 = 40\%$*). To summarize, the coverage of the paper, only cyan, only magenta and blue (cyan + magenta) is 0.4, 0.3, 0.2 and 0.1, respectively. Notice that the sum of the coverages is 1. Equation 7.2 for surface (b) gives,*

$$\begin{bmatrix} X_{surface(b)} \\ Y_{surface(b)} \\ Z_{surface(b)} \end{bmatrix} = 0.4 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.3 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.2 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.1 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

Calculating this gives the CIEXYZ values for surface (b): $[67, 75, 104]$.

Solution surface (c): *In this surface, according to the example and the figure, there is a complete dot overlap. As seen in this figure, magenta is completely covered by cyan, meaning that the coverage of pure magenta is 0 while the coverage of blue is 30% (because all magenta turned to blue). Cyan covered totally 40% of the surface, but 30% of it is now overlapped with magenta, meaning that pure cyan covers 10%. The rest of the surface is the paper: (*$100-10-0-30 = 60\%$*). To summarize, the coverage of the paper, only cyan, only magenta and blue (cyan + magenta) is 0.6, 0.1, 0 and 0.3, respectively. Notice that the sum of the coverages is 1. Equation 7.2 for surface (c) gives,*

$$\begin{bmatrix} X_{surface(c)} \\ Y_{surface(c)} \\ Z_{surface(c)} \end{bmatrix} = 0.6 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.1 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

Calculating this gives the CIEXYZ values for surface (c): $[71, 83, 106]$.

There are few interesting points worth discussing using the above example. In these three surfaces, three different printing strategies are illustrated. In surface (a), there is no dot overlap at all, which is called dot-off-dot printing (further discussed in Section 7.6). In surface (b), there is partly dot overlap and in surface

(c) there is complete dot overlap, which is called dot-on-dot printing (further discussed in Section 7.7). Another interesting observation is that in surface (c), i.e. dot-on-dot, there is more paper involved, making it lighter than the other two surfaces. This can also be verified by looking at CIEY-values of these three surfaces. The CIEY-value of surface (c) is larger than the two others, and as discussed in Section 6.4.1, the CIEY value represents the perceived luminance, and therefore the higher the CIEY the lighter the color. The CIEY for surface (a), dot-off-dot, on the other hand, is smallest of all three, meaning that it produces a darker color. Another interesting point is that, in all three surfaces the same amount of cyan and magenta was used (40% and 30% respectively), but the three surfaces resulted in three different colors, which was dependent on how they were mixed; i.e. dot-off-dot, partly dot-overlap or dot-on-dot.
In order to use Neugebauer equations, we need to figure out the coverage (a_i) for the NPs. In order to figure out that, we need to know the printing strategies being used, i.e. dot-off-dot, dot-on-dot or partly dot-overlap. These three strategies and how to find the coverage of the NPs are discussed in the following three sections.

7.5 Demichel equations

Before describing Demichel equations, notice that, in this book, a constant color image is meant to be a color image whose color channels are achromatic constant images. This means, all pixels in, for example the C channel, hold the same value, and all pixels in the M channel hold the same value (not necessarily the same pixel value as in the C channel), and so do the pixels in the other channels.
As discussed in Section 7.4, in order to figure out the coverage of the NPs, we need to know the printing strategy. In the case of partly dot-overlap, in Example 7.1, we had a figure and some explanation to figure out the coverage of the overlap. The question is now, if we for example mix 40% cyan and 30% magenta, how much overlap we will have. As discussed in Section 7.3, the color channels are usually halftoned independently. Let us now consider the same example, having a constant color image with 40% and 30% coverage in its cyan and magenta channel, respectively. The two channels in this constant image are now halftoned independent of each other and then printed. Since they are halftoned independently, we can use the probability to calculate the fractional coverage of each NP. Let us start with this question, what is the probability to have blue in the above example? To have blue, there must be both cyan and magenta printed. The probability to have cyan at a certain position on this surface is 40% and the probability to have magenta is 30%, meaning that the probability to have both (i.e. blue) is $0.4 \cdot 0.3 = 0.12$ or 12%. The probability to have pure cyan, is to have cyan but not magenta, meaning $0.4 \cdot (1 - 0.3) = 0.28$. The probability to have pure magenta, is to have magenta but not cyan, meaning $(1 - 0.4) \cdot 0.3 = 0.18$. The probability for the paper (no ink) is when neither cyan nor magenta is printed, meaning $(1 - 0.4) \cdot (1 - 0.3) = 0.42$. Therefore,

if a constant image containing 40% cyan and 30% magenta is halftoned independently and printed, the CIEXYZ values of the surface using Table 7.1 will be:

$$\begin{bmatrix} X_{independent} \\ Y_{independent} \\ Z_{independent} \end{bmatrix} = 0.42 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.28 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.18 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.12 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

Calculating this gives the CIEXYZ values for this example (the two channels are halftoned independently) as: $[67.4, 75.8, 104.2]$. Notice that the sum of the fractional coverages is $0.42 + 0.28 + 0.18 + 0.12 = 1$, as expected.

Using probability to calculate the coverage of NPs has actually a name; Demichel equations. Equation 7.3 shows how to calculate the fractional coverage of the NPs if three primary inks, cyan, magenta and yellow are involved. The coverage of cyan, magenta and yellow are denoted by c, m and y, respectively and the channels are halftoned **independently**.

$$\begin{cases} a_p = (1-c)(1-m)(1-y) \\ a_c = c(1-m)(1-y) \\ a_m = (1-c)m(1-y) \\ a_y = (1-c)(1-m)y \\ a_r = (1-c)my \\ a_g = c(1-m)y \\ a_b = cm(1-y) \\ a_k = cmy \end{cases} , \tag{7.3}$$

where a_i is the fractional coverage of NP i and indices p, c, m, y, r, g, b and k mean paper, cyan, magenta, yellow, red, green, blue and black, respectively. The Demichel equations when four primary inks are involved can be written correspondingly.

It must be pointed out again that Demichel equations can only be used assuming semi-stochastic overlap behavior. This means that, in the case of FM color halftoning, Demichel equations are very good approximation if the channels are halftoned **independently**. In the case of AM color halftoning, Demichel equations are a fairy good approximation if **different screen angles** are used for different color channels.

Example 7.2 *The color channels of a constant color image with 80%, 60% and 0% coverage in its cyan, magenta and yellow channel, have been halftoned independently. The CIEXYZ values of the four NPs are shown in Table 7.1. Calculate the CIEXYZ values for the printed halftone.*

Solution: *Since the color channels are halftoned* ***independently****, Demichel equations can be used to calculate the fractional coverage of the four NPs:*

$$\begin{cases} a_p = (1-c)(1-m) = (1-0.8)\cdot(1-0.6) = 0.08 \\ a_c = c(1-m) = 0.8\cdot(1-0.6) = 0.32 \\ a_m = (1-c)m = (1-0.8)\cdot 0.6 = 0.12 \\ a_b = cm = 0.8\cdot 0.6 = 0.48 \end{cases}$$

The CIEXYZ values of the printed halftone will be:

$$\begin{bmatrix} X_{independent} \\ Y_{independent} \\ Z_{independent} \end{bmatrix} = 0.08 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.32 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.12 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.48 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

Calculating this gives the CIEXYZ values of the printed halftone as: $[49.6, 61.2, 100.8]$.

7.6 Dot-on-Dot printing strategy

One of the printing strategies, briefly discussed in Example 7.1, is the dot-on-dot printing strategy. Dot-on-dot means complete overlap, i.e. the primary inks are printed on top of each other, as much as possible. An example of printing 40% cyan and 30% magenta using dot-on-dot was shown in Figure 7.14 (c). In dot-on-dot printing, Demichel equations can obviously not be used to approximate the coverage of NPs. Let us explain how to calculate the fractional coverage of the NPs by focusing on two-color print (cyan and magenta) for simplicity. Assume that the coverage of cyan and magenta are denoted by c and m, respectively.

7.6.1 $c \geq m$

Assume that the amount of cyan, c, is greater than or equal to the amount of magenta, m. When these two primary inks are printed using dot-on-dot, then magenta ink will be completely covered by the cyan ink, because there are more cyan, see Figure 7.15 (a). This means that, all magenta dots are overlapped by cyan dots, making the fractional coverage of pure magenta and blue 0 and m, respectively. The fractional coverage of pure cyan is $c-m$, because $m\%$ of cyan has been covered by magenta making it blue. As can be seen in figure (a), the fractional coverage of the non-inked part (paper) is $(1-c)$.
Therefore, in this case, see Figure 7.15 (a) again, the fractional coverage of paper, pure cyan, pure magenta and blue, are $1-c$, $c-m$, 0 and m, respectively. Notice specially that the sum of these fractional coverages is 1, as it should. Neugebauer equations now give the tristimulus values of this surface:

$$\begin{bmatrix} X_{ondot(c>m)} \\ Y_{ondot(c>m)} \\ Z_{ondot(c>m)} \end{bmatrix} = (1-c) \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + (c-m) \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} + 0 \begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} + m \begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix} \quad (7.4)$$

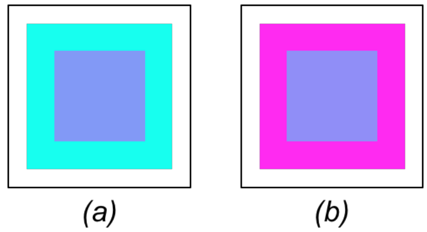


Figure 7.15: Dot-on-Dot printing using two primary inks, cyan and magenta: (a) $c \geq m$. (b) $c < m$.

7.6.2 $c < m$

The case $c < m$ is illustrated in Figure 7.15 (b). Using this figure and the same reasoning as above, it is easy to conclude that, in this case the fractional coverage of paper, pure cyan, pure magenta and blue, are $1 - m$, 0, $m - c$ and c, respectively. Notice specially that the sum of these fractional coverages is 1, as it should. Neugebauer equations now give the tristimulus values of this surface:

$$\begin{bmatrix} X_{ondot(c<m)} \\ Y_{ondot(c<m)} \\ Z_{ondot(c<m)} \end{bmatrix} = (1-m)\begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + 0\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} + (m-c)\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} + c\begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix} \quad (7.5)$$

Example 7.3 *The color channels of a constant color image with 80%, 60% and 0% coverage in its cyan, magenta and yellow channel, have been halftoned and printed using dot-on-dot. The CIEXYZ values of the four NPs are shown in Table 7.1. Calculate the CIEXYZ values for this printed halftone.*

Solution: *Since the color channels are halftoned using* ***dot-on-dot****, and since $c > m$, Equation 7.4 is used, giving:*

$$\begin{bmatrix} X_{example-7.3} \\ Y_{example-7.3} \\ Z_{example-7.3} \end{bmatrix} = 0.2\begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.2\begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0\begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.6\begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

Calculating this gives the CIEXYZ values of the printed halftone as: $[52, 66, 102]$.

If more than two primary inks are involved, similar reasoning and equations can be used to calculate the fractional coverage of the NPs. Let us show that by an example using three primary inks.

Example 7.4 *The color channels of a constant color image with 80%, 60% and 30% coverage in its cyan, magenta and yellow channel, have been halftoned and printed using dot-on-dot. The CIEXYZ values of the eight NPs are shown in Table 7.2. Calculate the CIEXYZ values for this printed halftone.*

Table 7.2: Example7.4: CIEXYZ values of the eight NPs.

	Paper	Cyan	Magenta	Yellow	Red	Green	Blue	Black
X	90	50	60	70	50	50	40	0
Y	100	80	30	80	30	80	50	0
Z	110	100	100	20	10	40	100	0

Solution: *In this example, the amount of cyan is more than the amount of magenta, which is more than the amount of yellow. This case is illustrated in Figure 7.16.*

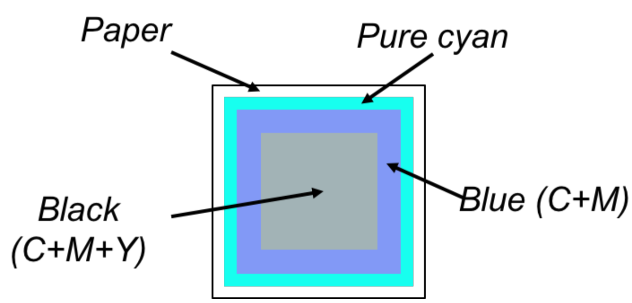


Figure 7.16: Example 7.4: 80% cyan, 60% magenta and 30% yellow are printed using dot-on-dot.

As can be seen in this figure, the fractional coverage of paper is $1 - c = 0.2$*. The fractional coverage of pure cyan is* $c - m = 0.2$*. The fractional coverage of pure magenta is* 0*, because it is all covered by cyan. The fractional coverage of pure yellow is also* 0*, because it is all covered by cyan (and magenta). The fractional coverage of red (magenta and yellow) is also* 0*, because the amount of yellow is less than the other two, and therefore there is no place where magenta and yellow occur but not cyan. The fractional coverage of green (cyan and yellow) is also* 0*, because the amount of yellow is less than the other two, and therefore there is no place where cyan and yellow occur but not magenta. The fractional coverage of blue (cyan and magenta) is* $m - y = 0.3$*, because blue occurs where cyan and magenta are present but yellow is absent. The fractional coverage of black (cyan and magenta and yellow) is* $y = 0.3$*, because it is where all three inks are printed on top of each other. Therefore, as also seen in Figure 7.16, of eight possible NPs, only four are present in this example, namely: paper (coverage* $a_p = 1 - c = 0.2$*), pure cyan (coverage* $a_c = c - m = 0.2$*), blue (coverage* $a_b = m - y = 0.3$*) and finally black* $a_k = y = 0.3$*. Notice specially that the sum of the coverages is* 1*, as expected. The Neugebauer equations give (those NPs that have zero coverage are excluded from this equation):*

$$\begin{bmatrix} X_{example-7.4} \\ Y_{example-7.4} \\ Z_{example-7.4} \end{bmatrix} = 0.2 \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + 0.2 \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} + 0.3 \begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix} + 0.3 \begin{bmatrix} X_k \\ Y_k \\ Z_k \end{bmatrix}$$

The CIEXYZ values of the NPs are given in Table 7.2. Setting them in the equation above gives:

$$\begin{bmatrix} X_{example-7.4} \\ Y_{example-7.4} \\ Z_{example-7.4} \end{bmatrix} = 0.2 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.2 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Calculating this gives the CIEXYZ values of the printed halftone as: $[40, 51, 72]$.

7.7 Dot-off-Dot printing strategy

One of the printing strategies, briefly discussed in Example 7.1, is the dot-off-dot printing strategy. Dot-off-dot means avoiding overlap, i.e. the primary inks are avoided to be printed on top of each other, as much as possible. An example of printing 40% cyan and 30% magenta was shown in Figure 7.14 (a). In dot-off-dot printing, Demichel equations can obviously not be used to approximate the coverage of NPs. Let us explain how to calculate the fractional coverage of the NPs by focusing on two-color print (cyan and magenta) for simplicity. Assume that the coverage of cyan and magenta are denoted by c and m, respectively.

7.7.1 $c + m \leq 1$

Assume that the sum of the amount of cyan and magenta is less than or equal to 100%, i.e. $c + m \leq 1$. When these two primary inks with sum less than 100% are printed using dot-off-dot, there won't be any ink overlap, meaning that there won't be any blue, see Figure 7.17 (a). This also means that, all cyan dots remain pure cyan and all magenta dots remain pure magenta. The rest will be the paper.

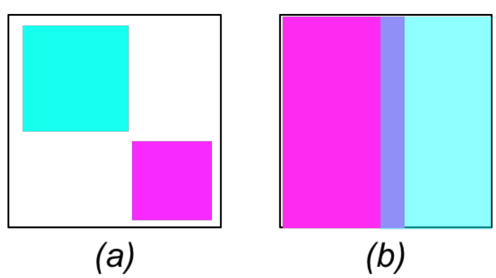


Figure 7.17: Dot-off-Dot printing using two primary inks, cyan and magenta: (a) $c + m \leq 1$. (b) $c + m > 1$.

Therefore, in this case, see Figure 7.17 (a) again, the fractional coverage of paper, pure cyan, pure magenta and blue, are $1-(c+m)$, c, m and 0, respectively. Notice specially that the sum of these fractional coverages is 1, as it should. Neugebauer equations now give the tristimulus values of this surface:

$$\begin{bmatrix} X_{offdot(c+m<1)} \\ Y_{offdot(c+m<1)} \\ Z_{offdot(c+m<1)} \end{bmatrix} = (1-c-m)\begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + c\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} + m\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} + 0\begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix} \quad (7.6)$$

7.7.2 $c+m>1$

Assume that the sum of the amount of cyan and magenta is greater than 100%, i.e. $c+m>1$. When these two primary inks with sum over 100% are printed using dot-off-dot as much as possible, it is impossible to avoid blue, see Figure 7.17 (b). Since the ink overlap is avoided as much as possible, everything above 100% coverage will cause ink overlap. This means that the coverage of blue will be $c+m-1$. The coverage of cyan was c, but $c+m-1$ of it turned to blue, so the remaining, i.e. $c-(c+m-1)=1-m$, will be the coverage of pure cyan. The coverage of magenta was m, but $c+m-1$ of it turned to blue, so the remaining, i.e. $m-(c+m-1)=1-c$, will be the coverage of pure magenta. Since the whole surface is covered by ink, then the coverage of paper will be 0. Therefore, in this case, see Figure 7.17 (b) again, the fractional coverage of paper, pure cyan, pure magenta and blue, are 0, $1-m$, $1-c$ and $c+m-1$, respectively. Notice specially that the sum of these fractional coverages is 1, as it should. Neugebauer equations now give the tristimulus values of this surface:

$$\begin{bmatrix} X_{offdot(c+m>1)} \\ Y_{offdot(c+m>1)} \\ Z_{offdot(c+m>1)} \end{bmatrix} = 0\begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + (1-m)\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} + (1-c)\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} + (c+m-1)\begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix} \quad (7.7)$$

Example 7.5 *The color channels of a constant color image with 80%, 60% and 0% coverage in its cyan, magenta and yellow channel, have been halftoned and printed using dot-off-dot. The CIEXYZ values of the four NPs are shown in Table 7.1. Calculate the CIEXYZ values for this printed halftone.*

Solution: *Since the color channels are halftoned using* ***dot-off-dot****, and $c+m>1$, Equation 7.7 is used, giving:*

$$\begin{bmatrix} X_{example-7.5} \\ Y_{example-7.5} \\ Z_{example-7.5} \end{bmatrix} = 0\begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + (1-0.6)\begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + (1-0.8)\begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + (0.8+0.6-$$

$$1) \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

Calculating this gives the CIEXYZ values of the printed halftone as: $[48, 58, 100]$.

If more than two primary inks are involved, similar reasoning and equations can be used to calculate the fractional coverage of the NPs. Let us show that by an example using three primary inks.

Example 7.6 *The color channels of a constant color image with 80%, 60% and 30% coverage in its cyan, magenta and yellow channel, have been halftoned and printed using dot-off-dot. The printing strategy is to first print cyan and magenta using dot-off-dot as much as possible. Then, the yellow dots are printed with the following priorities, First: printed on cyan as much as possible, Second: if yellow still remains, it will be printed on magenta as much as possible, and Third: if yellow still remains, it will be printed on blue to create black. The CIEXYZ values of the eight NPs are shown in Table 7.2. Calculate the CIEXYZ values for this printed halftone.*

Solution: *In this example, firstly cyan and magenta are printed dot-off-dot, resulting in* $100 - 60 = 40\%$ *pure cyan,* $100 - 80 = 20\%$ *pure magenta and* $80 + 60 - 100 = 40\%$ *blue (and of course 0% paper), see Figure 7.18 (a). The yellow dots are now firstly printed on cyan if remaining printed on magenta and finally on blue. Since the amount of yellow is 30% and the amount of pure cyan is 40%, then all yellow dots will be printed on pure cyan (making it green), see Figure 7.18 (b).*

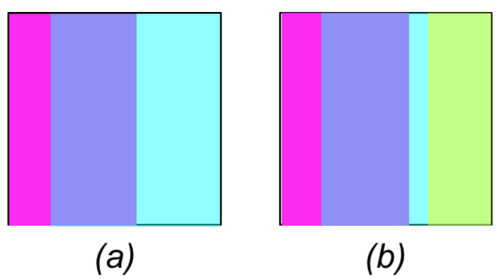


Figure 7.18: Example 7.6: (a) 80% cyan, 60% magenta are printed dot-off-dot. (b) 30% yellow are printed on the halftone in (a) according to the strategy explained in the example.

As can be seen in figure (b), the fractional coverage of paper is 0. *The fractional coverage of pure cyan is* $0.4 - 0.3 = 0.1$, *because 30% of 40% pure cyan in figure (a) is now covered by yellow. The fractional coverage of pure magenta, after adding yellow is still* 0.2, *because yellow was only printed on cyan. The fractional coverage of pure yellow is* 0, *because all of it was printed on cyan. The fractional coverage of red (magenta and yellow) is* 0, *because no yellow was* *printed on magenta. The fractional coverage of green (cyan and yellow) is* 0.3*, because all 30% yellow was printed on cyan. The fractional coverage of blue (cyan and magenta) is still* 0.4*, because no yellow was printed on blue. The fractional coverage of black (cyan and magenta and yellow) is* 0*, because no yellow was printed on blue to make black. Therefore, as also seen in Figure 7.18 (b), of eight possible NPs, only four are present in this example, namely: pure cyan (coverage* 0.1*), pure magenta (coverage* 0.2*), green (cyan + yellow) (coverage* 0.3*), and finally blue (coverage* 0.4*). Notice specially that the sum of the coverages is* 1*, as expected. The Neugebauer equations give (those NPs that have zero coverage are excluded from this equation):*

$$\begin{bmatrix} X_{example-7.6} \\ Y_{example-7.6} \\ Z_{example-7.6} \end{bmatrix} = 0.1 \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} + 0.2 \begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} + 0.3 \begin{bmatrix} X_g \\ Y_g \\ Z_g \end{bmatrix} + 0.4 \begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix}$$

The CIEXYZ values of the NPs are given in Table 7.2. Setting them in the equation above gives:

$$\begin{bmatrix} X_{example-7.6} \\ Y_{example-7.6} \\ Z_{example-7.6} \end{bmatrix} = 0.1 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.2 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 50 \\ 80 \\ 40 \end{bmatrix} + 0.4 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

Calculating this gives the CIEXYZ values of the printed halftone as: $[48, 58, 82]$.

7.8 Exercise set

7.1. The color channels of a constant color image with 50%, 70% and 0% coverage in its cyan, magenta and yellow channel, have been halftoned. The CIEXYZ values of the four NPs are shown in Table 7.3. Calculate the CIEXYZ values for this printed halftone if:

a) The channels are halftoned independently.

b) Dot-on-dot is used.

c) Dot-off-dot is used.

d) Which of the three strategies above results in the lightest color?

Table 7.3: Exercise 7.1: CIEXYZ values of the four NPs.

	Paper	Cyan	Magenta	Blue
X	90	50	60	40
Y	100	80	30	50
Z	110	100	100	100

7.2. The color channels of a constant color image with 50%, 70% and 60% coverage in its cyan, magenta and yellow channel, have been halftoned.

The CIEXYZ values of the eight NPs are shown in Table 7.4. Calculate the CIEXYZ values for this printed halftone if:

a) The channels are halftoned independently.
b) Dot-on-dot is used.
c) Dot-off-dot is used according to the following: the printing strategy is to first print cyan and magenta using dot-off-dot as much as possible. Then, the yellow dots are printed with this priority: First printed on cyan as much as possible, Second: if yellow still remains, it will be printed on magenta as much as possible, and Third: if yellow still remains, it will be printed on blue to create black.

Table 7.4: Exercise 7.2: CIEXYZ values of the eight NPs.

	Paper	Cyan	Magenta	Yellow	Red	Green	Blue	Black
X	90	50	60	70	50	50	40	0
Y	100	80	30	80	30	80	50	0
Z	110	100	100	20	10	40	100	0

7.3. The color channels of a constant color image with $c\%$, 80% and 0% coverage in its cyan, magenta and yellow channel, have been halftoned using dot-off-dot and then printed. The CIEY-value of the printed surface was measured to be $Y = 48$. Find the coverage of cyan, c! The CIEXYZ values of the four NPs are shown in Table 7.3.

7.4. The color channels of a constant color image with 60%, 40% and 40% coverage in its cyan, magenta and yellow channel are going to be printed. The cyan and magenta channels have been halftoned using AM halftoning. Figure 7.19 shows how cyan and magenta are printed. The ink overlap in the figure (i.e. blue) covers 30% of the area. Use Table 7.4 and:

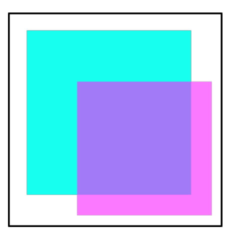


Figure 7.19: Exercise 7.4: 60% cyan and 40% magenta are printed according to this figure. The coverage of blue is 30%.

a) Calculate the color of the surface shown in Figure 7.19.

After cyan and magenta were printed according to Figure 7.19, the yellow dots are added to them. Calculate the CIEXYZ values in the following two cases:

b) The yellow dots are placed independent of the other two inks.

c) The yellow dots are placed using dot-off-dot as much as possible. Avoid black if possible and there is no priority for yellow to be printed on cyan or magenta.

7.5. A color image has been separated to three channels, C, M and Y. All pixels in channel C hold the value 0.4. All pixels in channel M holds the value 0.6. All pixels in channel Y hold the value 0. Use Table 7.3 and calculate the CIEXYZ values in the following cases. As usual, the images are supposed to be normalized between 0 and 1. The threshold matrices below have to be normalized.

a) Both C and M channels are first thresholded using the following matrix and then printed.

$$\begin{bmatrix} 1 & 9 & 2 & 10 \\ 5 & 13 & 6 & 14 \\ 11 & 3 & 12 & 4 \\ 7 & 15 & 8 & 16 \end{bmatrix}$$

b) Channel C is thresholded with the matrix in part (a) and channel M is thresholded using the following matrix. Then, the halftones are printed.

$$\begin{bmatrix} 16 & 8 & 15 & 7 \\ 12 & 4 & 11 & 3 \\ 6 & 14 & 5 & 13 \\ 10 & 2 & 9 & 1 \end{bmatrix}$$

c) Channel C is thresholded with the matrix in part (a) and channel M is thresholded using the following matrix. Then, the halftones are printed.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 6 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$$

7.6. Table 7.3 shows the CIEXYZ-values of paper, cyan, magenta and blue.

a) Is it possible to print $c\%$ cyan and $m\%$ magenta using dot-on-dot to get the color $CIEXYZ = [64, 77, 104.5]$? If no, why? If yes, find c and m.

b) Is it possible to print $c\%$ cyan and $m\%$ magenta using dot-off-dot to get the color $CIEXYZ = [64, 77, 104.5]$? If no, why? If yes, find c and m.

7.7. Table 7.3 shows the CIEXYZ-values of paper, cyan, magenta and blue. A constant image with $c\%$ cyan, $m\%$ magenta and 0% yellow has been halftoned and printed. We don't know exactly which printing strategy, i.e. independent halftoning, dot-on-dot or dot-off-dot, has been used but we know that the resulting color of the printed halftone is $CIEXYZ = [52.5, 59.5, 100]$. Find c and m and state which of these three printing strategies have been used.

7.8. The color channels of a constant color image with 80%, 60% and 30% coverage in its cyan, magenta and yellow channel, have been halftoned by FM halftoning. The CIEXYZ values of the eight NPs are shown in Table 7.4. Calculate the CIEXYZ values for this printed halftone if:

a) The channels are halftoned using dot-on-dot.

b) The cyan and magenta channels are firstly halftoned using dot-off-dot. The yellow channel is then halftoned independent of the two others.

c) The cyan and magenta channels are firstly halftoned using dot-off-dot. The yellow channel is then halftoned independent of the two others, but black is avoided as much as possible, meaning that yellow dots are avoided being printed on blue as much as possible.

Chapter 8

Answers

2.1 a) $50 \times 50\ cm^2$

b) 96 Megabits (Mb) or 12 Megabytes (MB)

2.2 The area is 37.5 cm^2

3.1 a) 12 Mbytes (MB)

b) $lpi = 50$

c) $dpi = 100$

3.2 a) Both are right, see the solution in Chapter 9

b) 15.12 Mbits (Mb) or 1.89 Mbytes (MB), if 1 Mb/MB is supposed to be 10^6 b/B

c) $45 \times 33.75\ cm^2$

3.3 4 times in each direction (or 16 times in area). The halftone should be AM.

3.4 a) $dpi = 1200$

b) $lpi = 150$

c) 65 different gray tones

3.5 a) One of the many 4×4 threshold matrices representing 10 levels of gray is:

$$\textit{threshold matrix}: \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{bmatrix} /10$$

and the thresholded image is:

$$\textit{halftoned image}: \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

b) One of the many possible correct answers:

$$\textit{halftoned image}: \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3.6 a) $\textit{halftoned image}: \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$

b) 6 levels of gray

3.7 One of the many possible correct answers is:

$$\textit{halftoned image}: \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3.8 $\textit{halftoned image}: \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$

3.9 The follwing four conditions have to be fulfilled: $a < 0.5$, $b \geq 0.5 - a$, $c \geq 0.5$ and $d < 1.5 - c$.

An example of an image that works is $\begin{bmatrix} 0.2 & 0.4 \\ 0.6 & 0.8 \end{bmatrix}$.

3.10 All values of x inside the interval $0.25 < x \leq 1 - \sqrt{3/8}$ and $y = 1 - x$.

6.1 a) $L = [0, 1, 0, 0]$

b) $C = [1, 0.5]$

c) For example, $L_1 = [2, 0, 0, 1]$

d) All $L_m = [a, b, c, d]$ with $\begin{cases} a/2 + b + c/2 = 1 \\ b/2 + c + d/2 = 0.5 \end{cases}$

6.2 a) $S = [0, 1/3, 4/3, 1, 0]$

b) $XYZ = [10/9, 5/3, 1/3]$

c) All reflectance spectra $R = [r_1, 1/3, 2/3, 1/3, r_5]$, with $0 \leq r_1 \leq 1$ and $0 \leq r_5 \leq 1$ are metameric with R_1 under L.

6.3 a) $C = \begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 4 & 2 \\ 0 & 0 & 1 & 4 & 1 & 0 & 0 \\ 2 & 4 & 2 & 0 & 0 & 0 & 0 \end{bmatrix}$

b) Q has to be multiplied by 1/4, giving $R = [1/2, 1/4, 0, 1/4, 0, 1/2, 1]$.

c) The RGB response without normalization is $RGB = [4, 1, 2]$.

6.4 a) i) $r(\lambda) \equiv 1$

ii) $r(\lambda) \equiv 0$

b) $p = \int_\lambda (I(\lambda) + k(\lambda))r(\lambda)s(\lambda)d\lambda$

c) $p = \int_\lambda I(\lambda)r(\lambda)f(\lambda)s(\lambda)d\lambda$

6.5 a) $[2, 17/2]$

b) Because the values defining a reflectance spectrum cannot be greater than 1.

c) $x = 1/3$

d) For example, $R_1 = [0, 1, 1, 1, 0]$ and $R_2 = [1, 1/2, 1, 0, 4/5]$. There are infinitely many other pairs of reflectance spectra that are metameric under L but not metameric under V.

6.6 a) $XYZ = [7, 8, 7]$

b) There are no spectra of the form $[a, 0, 0, 0, b, c]$ that are metameric.

7.1 a) $CIEXYZ_{Exerc.7.1(a)} = [56, 55, 101.5]$

b) $CIEXYZ_{Exerc.7.1(b)} = [59, 61, 103]$

c) $CIEXYZ_{Exerc.7.1(c)} = [53, 49, 100]$

d) Dot-on-dot (part(b)) results in a lighter color. Why? See the solution.

7.2 a) $CIEXYZ_{Exerc.7.2(a)} = [43.7, 42.7, 48.1]$

b) $CIEXYZ_{Exerc.7.2(b)} = [38, 36, 44]$

c) $CIEXYZ_{Exerc.7.2(c)} = [50, 49, 55]$

7.3 $c = 60\%$

7.4 a) $CIEXYZ_{Exerc.7.4(a)} = [60, 72, 103]$

b) $CIEXYZ_{Exerc.7.4(b)} = [52.4, 63.6, 69.4]$

c) $CIEXYZ_{Exerc.7.4(c)} = [53.75, 66.00, 69.25]$

7.5 a) $CIEXYZ_{Exerc.7.5(a)} = [63.75, 63.75, 103.75]$

b) $CIEXYZ_{Exerc.7.5(b)} = [56.25, 48.75, 100.00]$

c) $CIEXYZ_{Exerc.7.5(c)} = [62.50, 61.25, 103.13]$

7.6 a) Yes, $c = 55\%$ and $m = 40\%$

b) No (why? see the solution)

7.7 $c = 65\%$ cyan and $m = 45\%$ magenta have been printed using dot-off-dot printing strategy

7.8 a) $CIEXYZ_{Exerc.7.8(a)} = [40, 51, 72]$

b) $CIEXYZ_{Exerc.7.8(b)} = [42.6, 52.0, 75.4]$

c) $CIEXYZ_{Exerc.7.8(c)} = [47, 58, 79]$

Chapter 9

Solutions

2.1 a) Since the digital image is 2000×2000 pixels and $ppi = 100$, then the photograph must have been $\frac{2000}{100} \times \frac{2000}{100} = 20 \times 20\ inch^2$, which is $20 \cdot 2.5 \times 20 \cdot 2.5 = 50 \times 50\ cm^2$.

b) The total number of pixels the digital image consists of is: $2000 \times 2000 = 4 \times 10^6$. Since this is a color image, it has three channels (RGB) and since each pixel in each channel needs 8 bits, then the total amount of memory needed to store this digital image is: $3 \cdot 8 \cdot 4 \times 10^6 = 96 \times 10^6$ bits or 96 Megabits (Mb). Since 1 byte corresponds to 8 bits, you could also say that the memory needed to store this digital image is $96/8 = 12$ Megabytes (MB).

2.2 Assume the photograph to be $w \times h\ inch^2$. Then, since $ppi = 200$, the digital image will be $200w \times 200h$ pixels. Since this is a color image, the needed memory would be: $3 \cdot 200w \cdot 200h = 120000wh$ bytes (B). Setting that equal to 0.72 MB = 720000 B, we will have $wh = \frac{720000}{120000} = 6\ inch^2$, which is equal to $6 \cdot 2.5^2 = 37.5\ cm^2$, which is the area of the photograph.

3.1 a) $25 \times 25\ cm^2$, since an inch is 2.5 cm, means $10 \times 10\ inch^2$. The scanning resolution is $ppi = 200$, meaning that one inch becomes 200 pixels, therefore the digital image will be: $10 \cdot 200 \times 10 \cdot 200 = 2000 \times 2000$ pixels. This means that the digital image consists of $2000 \cdot 2000 = 4 \times 10^6$ pixels. Each pixel needs 8 bits (=1 byte) and since this is a color image, the memory needed to store this image will be: $4 \times 10^6 \cdot 3 \cdot 1 = 12 \times 10^6$ bytes (B) or 12 Mbytes (MB).

b) Since the halftoning method is AM, it is relevant to use the term *lpi*. In this exercise, we can see that the printed image is supposed to be double the size of the original photograph in each direction, i.e. $scaling\ factor = 50/25 = 2$. According to the rule of thumb, Equation 3.1, we have: $ppi = scaling\ factor \cdot 2 \cdot lpi = 4 \cdot lpi$, giving $lpi = ppi/4 = 50$.

c) Since the image is halftoned by a FM-method, as discussed in Section 3.3, there won't be any periodic structure and therefore *lpi* is not a relevant term to be used. As also mentioned in Section 3.3, when talking about FM method, we usually mean methods like error-diffusion. These techniques, if nothing else is specified, result in halftoned images that are the same size ($pixel \times pixel$) as the digital image being halftoned. This means that the halftoned image is also 2000×2000 pixels. Since each pixel in the digital halftoned image (bitmap) is going to be a dot, then the size of the printed image will be: $2000/dpi \times 2000/dpi\ inch^2$, where dpi is the print resolution. According to the exercise, this halftoned image is supposed to be $50 \times 50\ cm^2$ (or $20 \times 20\ inch^2$), meaning that $2000/dpi = 20$, giving $dpi = 100$.
However, the answer could also be found by a simple observation. Since the halftoned image is the same size ($pixel \times pixel$) as the digital image being halftoned, then $ppi/dpi = scaling\ factor$, since the scaling factor is 2 in this exercise, then $dpi = ppi/2 = 100$.

3.2 a) Both of them are right, because Lars' friend calculates it assuming the image is stored in raw-format. But the image is most probably compressed (for example to jpg-format) and stored, which will need less memory to be saved.

b) The halftoned image is the same size as the digital image according to the exercise. Thus, the total number of pixels is $2592 \cdot 1944 \approx 5.04 \times 10^6$. Since it is halftoned, each pixel only needs one **bit** to be stored. Since the image is color, assuming three color channels, the halftoned image will need: $5.04 \cdot 10^6 \cdot 3 = 15.12 \times 10^6$ bits. Assuming 10^6 bits to be 1 Mbits, it will need 15.12 Mbits (Mb), or $\frac{15.12}{8} = 1.89$ Mbytes (MB).
Notice that, as also mentioned in Section 2.1, it is more accurate to say one Mbit is $(1024)^2$ bits. If so, this halftoned image will need almost 14.42 Mbits.

c) $dpi = 144$ means that each set of 144 pixels in the digital halftoned image becomes 1 inch after print. The printed halftoned image will therefore be: $\frac{2592}{144} \times \frac{1944}{144} = 18 \times 13.5\ inch^2$ or $45 \times 33.75\ cm^2$.

3.3 According to the rule of thumb, Equation 3.1: $1200 = scaling\ factor \cdot 2 \cdot 150$, giving $scaling\ factor = 4$. This means that the printed image could be 4 times the photograph in each direction, i.e. 16 times larger in area. This should, of course, be an AM halftone, because the screen frequency, *lpi*, is only relevant for periodic structures (AM halftones), see for example Section 3.3.

3.4 a) Since nothing else is specified, the FM-halftoned image is supposed to be the same size as the digital image being halftoned. As discussed in the solution for Exercise 3.1 (c) above, $ppi/dpi = scaling\ factor$, giving $dpi = \frac{ppi}{scaling\ factor} = \frac{600}{1/2} = 1200$.

b) According to the exercise *scaling factor* $= 2$. The rule of thumb, Equation 3.1, gives $600 = \mathit{scaling\ factor}\cdot 2\cdot \mathit{lpi}$, giving $\mathit{lpi} = \frac{600}{4} = 150$.

c) According to Equation 3.2, $\mathit{number\ of\ gray\ tones} = (\frac{dpi}{lpi})^2 + 1 = (\frac{1200}{150})^2 + 1 = 65$.

3.5 a) A 4×4 threshold matrix could represent up to 17 different gray levels. Since in this exercise it is supposed to represent 10 gray levels, some of the threshold values need to be repeated, see Section 3.2.1. The threshold values should therefore start from 1, and end in 9, giving 10 levels of gray. The threshold matrix is as always normalized by diving its elements by the number of gray levels (in this case 10). The following threshold matrix is one example of such a threshold matrix but there are many other matrices that are also correct.

$$\mathit{threshold\ matrix}: \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{bmatrix} /10$$

Since the image is also 4×4, then it is halftoned by comparing each pixel value in the image with its corresponding threshold value in the threshold matrix. For example, the element at row 1 and column 1, i.e. $(1, 1)$, in the original image (has the value 0), is compared with the threshold value at the same position $(1, 1)$ that has the value of $1/10 = 0.1$. Since 0 is less than 0.1, then the halftoned image is 0 at position $(1, 1)$. At position row 1 and column 2, i.e. $(1, 2)$, the pixel value in the image is 0.3, and the threshold value at $(1, 2)$ is $2/10 = 0.2$. Since 0.3 is greater than (or equal to) 0.2, then the halftoned image is 1 at $(1, 2)$. This comparison is done element-by-element for the rest of pixel positions. The final halftoned image is therefore:

$$\mathit{halftoned\ image}: \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

b) According to the exercise, the halftone tables are 3×3 and the final image is supposed to be 6×6. This means that, each 2×2 sub-area in the original image is supposed to be represented by a 3×3 halftone table that has the closest average. 3×3 halftone tables can represent up to 10 gray levels, i.e. $0, 1/9, 2/9, \cdots, 8/9, 1$. Since in the exercise it is stated that the halftone should be AM, then the halftone dots in the halftone tables have to be clustered. The first 2×2 sub-area (up to the left) in the image has the mean $(0+0.3+0.4+0.6)/4 = 0.325$. The closest table should include 3 black microdots because it has the average of $3/9 \approx 0.33$. Therefore, the first halftone table (up to the left) include three microdots. The second 2×2 sub-area, up to the right, in the image has the average of $(0.2+0.5+0.4+0.9)/4 = 0.5$. The closest halftone table would include 5 black microdots giving the average of $5/9 \approx 0.56$ (using 4 microdots here would also give an average almost equally far from 0.5). The third sub-area (down to the left) has the average of $(0.8 + 0.8 + 0.8 + 0.8)/4 = 0.8$ and the closest halftone table should include 7 black microdots giving $7/9 \approx 0.78$. Finally, the last 2×2 sub-area has the average of $(0.7 + 0.1+0.1+0.8)/4 = 0.425$ and the closest halftone table is a table with 4 black microdots giving the average of $4/9 \approx 0.44$. Recall that in this exercise, the halftone has to be AM and therefore the microdots in each halftone table have to make a clustered halftone dot. The following halftoned image is one of the many possible correct answers:

$$\textit{halftoned image}: \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3.6 a) The halftone tables are 3×3 and the final image is supposed to be 6×6. This means that, each 2×2 sub-area in the original image is supposed to be represented by a 3×3 halftone table that has the closest average. In this exercise, there are six different halftone tables representing gray levels: $0/9 = 0$, $2/9 \approx 0.22$, $5/9 \approx 0.56$, $6/9 \approx 0.67$, $7/9 \approx 0.78$ and $9/9 = 1$. The first 2×2 sub-area (up to the left) in the image has the mean $(0.3 + 0.8 + 0.4 + 0.5)/4 = 0.5$. The closest table among the six available tables is the one with 5 black microdots giving the average of 0.56. The second 2×2 sub-area, up to the right, in the image has the average of $(0.1 + 0.3 + 0.2 + 0.2)/4 = 0.2$. The closest table among the six available tables is the one with 2 black microdots giving the average of 0.22. The third sub-area (down to the left) has the average of $(0.2 + 0.8 + 0.5 + 0.9)/4 = 0.6$ and the closest table among the six available tables is the one with 5 black microdots giving the average of 0.56. Finally, the last 2×2 sub-area has the average of $(0.1+0.5+0.2+0.4)/4 = 0.3$ and the closest table among the six available tables is the one with 2 black microdots giving the average of 0.22. The table halftoned image is therefore:

$$\textit{halftoned image}: \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

b) Of course six gray levels, because there are six halftone tables with different gray levels available.

3.7 The size of the halftone tables are not directly specified in the exercise, but we know that the ratio of *dpi* to *lpi* decides the size of the halftone cells (tables), which in this case is: $\frac{dpi}{lpi} = \frac{300}{100} = 3$. This means that the halftone tables should be 3×3. The first sub-area in the image has the average of 0.4, meaning that it should be represented by a halftone table holding 4 black microdots because $4/9 \approx 0.44$. The second and the third sub-areas have the average of 0.7, and should be represented by a halftone table holding 6 black microdots, because $6/9 \approx 0.67$ is the closest mean among the 10 available levels. Finally, the last sub-area has the average of 0.3 and should be represented by a table with 3 black microdots. The halftoned image is therefore the following. This is of course one of the many possible correct answers.

$$\textit{halftoned image}: \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Notice that, since in this exercise nothing about the structure of halftone dots is specified, the halftone dots could be arranged as you wish, and the only important aspect is their average. In the answer above, the halftone structure is according to line halftones.

3.8 Assume the original image and the final halftoned image to be denoted by g and b, respectively.

- **Position (1,1):** $g(1,1) = 0.6 \geq 0.5 \rightarrow$ **b(1,1)=1**.
  - the error is: $error = g(1,1) - b(1,1) = 0.6 - 1 = -0.4$.
  - $g_{new}(1,2) = g(1,2) + error \cdot weight\ to\ the\ right = 0.5 + (-0.4) \cdot 0.6 = 0.26$
  - $g_{new}(2,1) = g(2,1) + error \cdot weight\ under = 0.8 + (-0.4) \cdot 0.4 = 0.64$
- **Position (1,2):** The new value at $(1,2)$ is $g_{new}(1,2) = 0.26$, and $g_{new}(1,2) = 0.26 < 0.5 \rightarrow$ **b(1,2)=0**.
  - the error is: $error = g_{new}(1,2) - b(1,2) = 0.26 - 0 = 0.26$.
  - no pixel to the right
  - $g_{new}(2,2) = g(2,2) + error \cdot weight\ under = 0.4 + 0.26 \cdot 0.4 = 0.504$
- **Position (2,1):** The new value at $(2,1)$ is $g_{new}(2,1) = 0.64$, and $g_{new}(2,1) = 0.64 \geq 0.5 \rightarrow$ **b(2,1)=1**.
  - the error is: $error = g_{new}(2,1) - b(2,1) = 0.64 - 1 = -0.36$.
  - $g_{new2}(2,2) = g_{new}(2,2) + error \cdot weight\ to\ the\ right = 0.504 + (-0.36) \cdot 0.6 = 0.288$
  - no pixel under

- **Position (2,2):** The new value at $(2,2)$ is $g_{new2}(2,2) = 0.288$, and $g_{new2}(2,2) = 0.288 < 0.5 \rightarrow$ **b(2,2)=0**.
  - This was the last pixel and no error needs to be calculated.

The final halftoned image is therefore: $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$

3.9 Assume the original image and the final halftoned image to be denoted by g and h, respectively. We call the halftoned image h, in order not to be confused with the variable b in the image. The first observation is that the sum of the weights in the error filter has to be 1, and since the filter is supposed to only diffuse the error to the right, then the error filter is: $\begin{bmatrix} \times & 1 \end{bmatrix}$. The values of all four variables a, b, c and d should be between 0 and 1 because the original image is supposed to be normalized to the interval $[0, 1]$.

- **Position (1,1):** Since $h(1,1) = 0$ in the result, then $g(1,1)$ must have been less than 0.5, giving $a < 0.5$. Therefore: **Condition 1:** $a < 0.5$
  - the error is: $error = g(1,1) - h(1,1) = a - 0 = a$.
  - $g_{new}(1,2) = g(1,2) + error \cdot weight\ to\ the\ right = b + a \cdot 1 = a + b$
- **Position (1,2):** Since $h(1,2) = 1$ in the result, then the new $g_{new}(1,2)$ must have been greater than or equal to 0.5, giving: $g_{new}(1,2) \geq 0.5 \rightarrow a + b \geq 0.5 \rightarrow b \geq 0.5 - a$. Therefore: **Condition 2:** $b \geq 0.5 - a$
  - no pixel to the right, therefore no error needs to be calculated
- **Position (2,1):** Since $h(2,1) = 1$ in the result, then $g(2,1)$ must have been greater than or equal to 0.5, giving: $g(2,1) \geq 0.5 \rightarrow c \geq 0.5$. Therefore: **Condition 3:** $c \geq 0.5$
  - the error is: $error = g(2,1) - h(2,1) = c - 1$.
  - $g_{new}(2,2) = g(2,2) + error \cdot weight\ to\ the\ right = d + (c-1) \cdot 1 = c + d - 1$.
- **Position (2,2):** Since $h(2,2) = 0$ in the result, then the new $g_{new}(2,2)$ must have been less than 0.5, giving: $c + d - 1 < 0.5 \rightarrow d < 1.5 - c$. Therefore: **Condition 4:** $d < 1.5 - c$

The four conditions above together with the fact that these variables are between 0 and 1 fulfill the criteria mentioned in the exercise. To give an example of an image that works, set for example $a = 0.2$ (according to condition 1, a has to be less than 0.5). According to condition 2, $b \geq 0.5 - a \rightarrow b \geq 0.3$, for example set $b = 0.4$. According to condition 3, c has to be greater than or equal to 0.5, set for example $c = 0.6$. According to condition 4, $d < 1.5 - c \rightarrow d < 0.9$, for example set $d = 0.8$. Therefore, an example of an image that works is $\begin{bmatrix} 0.2 & 0.4 \\ 0.6 & 0.8 \end{bmatrix}$.

3.10 Assume the original image and the final halftoned image to be denoted by g and b, respectively. The first observation is that the sum of the weights in the error filter has to be 1, giving $x + y = 1 \rightarrow y = 1 - x$. Therefore, only one unknown has to be taken into account and y could be replaced by $1 - x$. The weights are also non-negative real numbers.

- **Position (1,1):** $g(1,1) = 0.6 \geq 0.5 \rightarrow$ **b(1,1)=1**, which is ok, because in the result $b(1,1)$ is 1 as well. **Notice** that since $y = 1 - x$, we replace y by $1 - x$ in the following equations.
  - the error is: $error = g(1,1) - b(1,1) = 0.6 - 1 = -0.4$.
  - $g_{new}(1,2) = g(1,2) + error \cdot weight\ to\ the\ right = 0.6 + (-0.4) \cdot x = 0.6 - 0.4x$
  - $g_{new}(2,1) = g(2,1) + error \cdot weight\ under = 0.8 + (-0.4) \cdot (1 - x) = 0.4 + 0.4x$
- **Position (1,2):** Since $b(1,2) = 0$ in the result, then the new $g_{new}(1,2)$ must be less than 0.5, giving: $g_{new}(1,2) < 0.5 \rightarrow 0.6 - 0.4x < 0.5 \rightarrow x > \frac{0.1}{0.4} \rightarrow x > 0.25$. Therefore: **Condition 1:** $x > 0.25$
  - the error is: $error = g_{new}(1,2) - b(1,2) = 0.6 - 0.4x - 0 = 0.6 - 0.4x$.
  - no pixel to the right
  - $g_{new}(2,2) = g(2,2) + error \cdot weight\ under = 0.4 + (0.6 - 0.4x) \cdot (1 - x) = 0.4x^2 - x + 1$
- **Position (2,1):** Since $b(2,1) = 1$ in the result, then the new $g_{new}(2,1)$ must be greater than or equal to 0.5, giving: $g_{new}(2,1) \geq 0.5 \rightarrow 0.4 + 0.4x \geq 0.5 \rightarrow x \geq 0.25$. Therefore: **Condition 2:** $x \geq 0.25$
  - the error is: $error = g_{new}(2,1) - b(2,1) = 0.4 + 0.4x - 1 = 0.4x - 0.6$.
  - $g_{new2}(2,2) = g_{new}(2,2) + error \cdot weight\ to\ the\ right = 0.4x^2 - x + 1 + (0.4x - 0.6) \cdot x = 0.8x^2 - 1.6x + 1$
  - no pixel under
- **Position (2,2):** Since $b(2,2) = 1$ in the result, then the new $g_{new2}(2,2)$ must be greater than or equal to 0.5, giving: $0.8x^2 - 1.6x + 1 \geq 0.5 \rightarrow 0.8x^2 - 1.6x + 0.5 \geq 0 \rightarrow 8x^2 - 16x + 5 \geq 0$, this gives the inequality: $x^2 - 2x + 5/8 \geq 0 \rightarrow (x-1)^2 - 1 + 5/8 \geq 0 \rightarrow (x-1)^2 \geq 3/8$. This inequality has two solutions:

  1: $x - 1 \geq \sqrt{3/8}$, giving $x \geq 1 + \sqrt{3/8}$, which is not possible because x cannot be greater than 1, (otherwise would $y = 1 - x$ be negative).

  2: $x - 1 \leq -\sqrt{3/8}$, giving $x \leq 1 - \sqrt{3/8}$. Therefore: **Condition 3:** $x \leq 1 - \sqrt{3/8}$

The three conditions together gives: All values of x inside the interval $0.25 < x \leq 1 - \sqrt{3/8}$ and $y = 1 - x$, will result in the halftoned image shown in the exercise.

6.1 a) The wavelength 500 nm is in the second sub-interval, meaning that the light source can be represented by $L = [0, 1, 0, 0]$.

b) The sensor has two receptors, each producing a signal when hit by light. Therefore, the output C is a vector with two elements. Notice that in this exercise the light is hitting directly the sensor and it is not reflected from any object. Using a similar equation as the one in Equation 6.8 gives:

response from row 1 in S: $sum(L.{*}[0.5, 1, 0.5, 0]) = sum([0, 1, 0, 0].{*}[0.5, 1, 0.5, 0]) = sum([0, 1, 0, 0]) = 1$
response from row 2 in S: $sum(L.{*}[0, 0.5, 1, 0.5]) = sum([0, 1, 0, 0].{*}[0, 0.5, 1, 0.5]) = sum([0, 0.5, 0, 0]) = 0.5$

Therefore, sensor's output is: $C = [1, 0.5]$.

c) The output for a general light source $L_1 = [a, b, c, d]$ is:

$sum(L_1.{*}[0.5, 1, 0.5, 0]) = sum([a, b, c, d].{*}[0.5, 1, 0.5, 0]) = sum([a/2, b, c/2, 0]) = a/2 + b + c/2$
$sum(L_1.{*}[0, 0.5, 1, 0.5]) = sum([a, b, c, d].{*}[0, 0.5, 1, 0.5]) = sum([0, b/2, c, d/2]) = b/2 + c + d/2$

We just need to construct a light source giving the output $[1, 0.5]$ as in (b). This means that we need to construct a light source $L_1 = [a, b, c, d]$ with:

$a/2 + b + c/2 = 1$ and $b/2 + c + d/2 = 0.5$. For example, $L_1 = [2, 0, 0, 1]$ gives the same output as the light source L when hitting the sensor.

d) As shown in part (c), a light source $L_m = [a, b, c, d]$ gives the output $C = [a/2 + b + c/2, b/2 + c + d/2]$ when hitting S. In order for it to have the same output as part (c), the following relationships must be fulfilled between L_m's elements:

$$\begin{cases} a/2 + b + c/2 = 1 \\ b/2 + c + d/2 = 0.5 \end{cases}$$

6.2 a) The spectral distribution reflected from the object is simply the incoming light multiplied by the reflectance of the object, meaning that: $S = L.*R_1 = [0, 1, 2, 3, 4].*[0, 1/3, 2/3, 1/3, 0] = [0, 1/3, 4/3, 1, 0]$.

b) The XYZ-values can be calculated by Equation 6.8 (ignore the normalization factor as specified in the exercise),

$X = sum(S.{*}\bar{x}(\lambda)) = sum([0, 1/3, 4/3, 1, 0].{*}[0, 0, 1/3, 2/3, 0]) = 10/9$

$Y = sum(S. * \bar{y}(\lambda)) = sum([0, 1/3, 4/3, 1, 0]. * [0, 0, 1, 1/3, 0]) = 5/3$
$Z = sum(S. * \bar{z}(\lambda)) = sum([0, 1/3, 4/3, 1, 0]. * [0, 1, 0, 0, 0]) = 1/3$

Therefore, the XYZ-values of the object under L are $[10/9, 5/3, 1/3]$.

c) Assume a general reflectance $R = [r_1, r_2, r_3, r_4, r_5]$. The XYZ-values of this reflectance under L is:

$X = sum(R. * L. * \bar{x}(\lambda)) = sum([r_1, r_2, r_3, r_4, r_5]. * [0, 1, 2, 3, 4]. * [0, 0, 1/3, 2/3, 0]) = (2/3)r_3 + 2r_4$
$Y = sum(R. * L. * \bar{y}(\lambda)) = sum([r_1, r_2, r_3, r_4, r_5]. * [0, 1, 2, 3, 4]. * [0, 0, 1, 1/3, 0]) = 2r_3 + r_4$
$Z = sum(R. * L. * \bar{z}(\lambda)) = sum([r_1, r_2, r_3, r_4, r_5]. * [0, 1, 2, 3, 4]. * [0, 1, 0, 0, 0]) = r_2$

Therefore, the XYZ-values for $R = [r_1, r_2, r_3, r_4, r_5]$ under L is $[(2/3)r_3 + 2r_4, 2r_3 + r_4, r_2]$. In order to have it metameric with R_1 (whose XYZ-values were calculated in part (b)), we have to have:

$$\begin{cases} (2/3)r_3 + 2r_4 = 10/9 \\ 2r_3 + r_4 = 5/3 \\ r_2 = 1/3 \end{cases}$$

The first two equations give $r_3 = 2/3$ and $r_4 = 1/3$. The third equation gives $r_2 = 1/3$. Therefore, all reflectance spectra $R = [r_1, 1/3, 2/3, 1/3, r_5]$, with $0 \leq r_1 \leq 1$ and $0 \leq r_5 \leq 1$ are metameric with R_1 under L.

6.3 a) From the figure, it is clearly seen that the camera matrix is the following, where row 1, 2 and 3 represent the sensitivity function for red, green and blue channels, respectively.

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 4 & 2 \\ 0 & 0 & 1 & 4 & 1 & 0 & 0 \\ 2 & 4 & 2 & 0 & 0 & 0 & 0 \end{bmatrix}$$

b) As discussed before, in a correct reflectance spectrum the reflectance value cannot be negative or greater than 1. According to the exercise, the object reflects all the incoming light at 700 nm, meaning that the seventh (last) element in Q has to be 1 instead of 4. This means that Q has to be multiplied by 1/4, giving $R = [1/2, 1/4, 0, 1/4, 0, 1/2, 1]$.

c) The camera's response can be calculated by similar equations being used in previous exercises. A perfect white light source could be described by a vector with all elements equal to 1. This gives:
The camera's response from the R channel:

$R_c = sum([1,1,1,1,1,1,1].*[1/2,1/4,0,1/4,0,1/2,1].*[0,0,0,0,2,4,2]) = sum([0,0,0,0,0,2,2]) = 4$

The camera's response from the G channel:

$G_c = sum([1,1,1,1,1,1,1].*[1/2,1/4,0,1/4,0,1/2,1].*[0,0,1,4,1,0,0]) = sum([0,0,0,1,0,0,0]) = 1$

The camera's response from the B channel:

$B_c = sum([1,1,1,1,1,1,1].*[1/2,1/4,0,1/4,0,1/2,1].*[2,4,2,0,0,0,0]) = sum([1,1,0,0,0,0,0]) = 2$

Therefore, the camera's RGB response to R under a white light source, without normalization, is $RGB = [4,1,2]$.

6.4 a) i) A mirror that reflects all incoming light, has a reflectance spectrum equal to 1 at all wavelengths, meaning that: $r(\lambda) \equiv 1$.

ii) A black hole that absorbs all incoming light, has a reflectance spectrum equal to 0 at all wavelengths, meaning that: $r(\lambda) \equiv 0$.

b) If we add an extra lamp, the power distribution of the two lamps are added together giving:

$p = \int_\lambda (I(\lambda) + k(\lambda))r(\lambda)s(\lambda)d\lambda$

c) The filter in front of the objective means that all incoming light has to be passed through the filter before hitting the sensor. This means that the distribution of the light is multiplied by the function of the filter before reaching the sensor, meaning:

$p = \int_\lambda I(\lambda)r(\lambda)f(\lambda)s(\lambda)d\lambda$

6.5 a) Similar to the previous exercises, we will have:

$S_1 = sum([1,2,3,4,5].*[1/2,0,1/2,1/2,1].*[1,1,1,0,0]) = sum([1/2,0,3/2,0,0]) = 2$

$S_2 = sum([1,2,3,4,5].*[1/2,0,1/2,1/2,1].*[0,0,1,1,1]) = sum([0,0,3/2,2,5]) = 17/2$

Therefore, the camera response is $[2, 17/2]$.

b) Because the values defining a reflectance spectrum cannot be greater than 1 at any wavelength (of course, assuming the object only being reflective and not emitting light).

c) As discussed above, the values in a reflectance spectrum cannot be greater than 1 at any wavelength, therefore the largest scaling factor is $x = 1/3$. Notice that if $x > 1/3$, then the third element in Q will hold a value greater than 1. For any $0 < x \leq 1/3$, however, Q can represent a reflectance spectrum.

d) Assume $R_1 = [a_1, b_1, c_1, d_1, e_1]$ and $R_2 = [a_2, b_2, c_2, d_2, e_2]$. The camera's response to R_1 under L is:

$S_{1(R1-L)} = sum([1,2,3,4,5].*[a_1, b_1, c_1, d_1, e_1].*[1,1,1,0,0]) = a_1 + 2b_1 + 3c_1$

$S_{2(R1-L)} = sum([1,2,3,4,5].*[a_1, b_1, c_1, d_1, e_1].*[0,0,1,1,1]) = 3c_1 + 4d_1 + 5e_1$

The camera's response to R_2 under L is:

$S_{1(R2-L)} = sum([1,2,3,4,5].*[a_2, b_2, c_2, d_2, e_2].*[1,1,1,0,0]) = a_2 + 2b_2 + 3c_2$

$S_{2(R2-L)} = sum([1,2,3,4,5].*[a_2, b_2, c_2, d_2, e_2].*[0,0,1,1,1]) = 3c_2 + 4d_2 + 5e_2$

In order to have R_1 and R_2 metameric under L, the following equations have to be satisfied,

$$\begin{cases} a_1 + 2b_1 + 3c_1 = a_2 + 2b_2 + 3c_2 \\ 3c_1 + 4d_1 + 5e_1 = 3c_2 + 4d_2 + 5e_2 \end{cases} . \tag{9.1}$$

The camera's response to R_1 under V is:

$S_{1(R1-V)} = sum([1,1,1,1,1].*[a_1, b_1, c_1, d_1, e_1].*[1,1,1,0,0]) = a_1 + b_1 + c_1$

$S_{2(R1-V)} = sum([1,1,1,1,1].*[a_1, b_1, c_1, d_1, e_1].*[0,0,1,1,1]) = c_1 + d_1 + e_1$

The camera's response to R_2 under V is:

$S_{1(R2-V)} = sum([1,1,1,1,1].*[a_2, b_2, c_2, d_2, e_2].*[1,1,1,0,0]) = a_2 + b_2 + c_2$

$S_{2(R2-V)} = sum([1,1,1,1,1].*[a_2, b_2, c_2, d_2, e_2].*[0,0,1,1,1]) = c_2 + d_2 + e_2$

In order **not** to have R_1 and R_2 metameric under V, **at least one**

of the following two inequalities has to be satisfied,

$$\begin{cases} a_1 + b_1 + c_1 \neq a_2 + b_2 + c_2 \\ c_1 + d_1 + e_1 \neq c_2 + d_2 + e_2 \end{cases} . \tag{9.2}$$

In order to have both equations in Equation 9.1 and at the same time at least one of the inequalities in Equation 9.2 satisfied, there are infinitely many options. Just notice that, all reflectance values have to be between 0 and 1. For example, $R_1 = [0, 1, 1, 1, 0]$ and $R_2 = [1, 1/2, 1, 0, 4/5]$ are metameric under L but not under V. As can be noticed, camera's response to both R_1 and R_2 under L is $[5, 7]$, while camera's responses to R_1 and R_2 under V are $[2, 2]$ and $[5/2, 9/5]$, respectively. Therefore, the given R_1 and R_2 are metameric under L but not metameric under V.

6.6 a) Similar to the previous exercises, we will have:

$X = sum([1, 0, 1, 0, 1, 0]. * [1, 1, 1, 5, 5, 2]) = sum([1, 0, 1, 0, 5, 0]) = 7$

$Y = sum([1, 0, 1, 0, 1, 0]. * [0, 1, 5, 6, 3, 0]) = sum([0, 0, 5, 0, 3, 0]) = 8$

$Z = sum([1, 0, 1, 0, 1, 0]. * [6, 8, 1, 0, 0, 0]) = sum([6, 0, 1, 0, 0, 0]) = 7$

Therefore, the CIEXYZ values for S_1 is $XYZ = [7, 8, 7]$.

b) Assume two spectra of the form mentioned in the exercise, $S_1 = [a_1, 0, 0, 0, b_1, c_1]$ and $S_2 = [a_2, 0, 0, 0, b_2, c_2]$. The CIEXYZ values for S_1 is:

$X_{S_1} = sum([a_1, 0, 0, 0, b_1, c_1].*[1, 1, 1, 5, 5, 2]) = sum([a_1, 0, 0, 0, 5b_1, 2c_1]) = a_1 + 5b_1 + 2c_1$

$Y_{S_1} = sum([a_1, 0, 0, 0, b_1, c_1].*[0, 1, 5, 6, 3, 0]) = sum([0, 0, 0, 0, 3b_1, 0]) = 3b_1$

$Z_{S_1} = sum([a_1, 0, 0, 0, b_1, c_1].*[6, 8, 1, 0, 0, 0]) = sum([6a_1, 0, 0, 0, 0, 0]) = 6a_1$

Therefore, the CIEXYZ values for S_1 are $[a_1 + 5b_1 + 2c_1, 3b_1, 6a_1]$. Very similar equations give the CIEXYZ value for S_2 as $[a_2 + 5b_2 + 2c_2, 3b_2, 6a_2]$. In order for the two spectra to be metameric, we have to have:

$$\begin{cases} a_1 + 5b_1 + 2c_1 = a_2 + 5b_2 + 2c_2 \\ 3b_1 = 3b_2 \\ 6a_1 = 6a_2 \end{cases}$$

The second and the third equations above give, $b_1 = b_2$ and $a_1 = a_2$, respectively. Setting them in the first equation gives $c_1 = c_2$. This means that the two spectra $S_1 = [a_1, 0, 0, 0, b_1, c_1]$ and $S_2 = [a_2, 0, 0, 0, b_2, c_2]$ are metameric if and only if $a_1 = a_2$, $b_1 = b_2$ and $c_1 = c_2$, which means that S_1 and S_2 have to be the same spectrum in order to be metameric. Therefore, there are no spectra of the form $[a, 0, 0, 0, b, c]$ that are metameric, because according to the definition two **different** spectra giving the same color are called metameric. It has been shown that there cannot be two **different** spectra of the form $[a, 0, 0, 0, b, c]$ that are metameric.

7.1 There are two primary inks, cyan and magenta, involved, making the number of NPs $2^2 = 4$. The NPs are: paper, pure cyan, pure magenta and blue (cyan + magenta).

a) Since the channels are halftoned **independently**, Demichel equations (Equation 7.3) can be used, giving:

$a_p = (1-c)(1-m) = (1-0.5)(1-0.7) = 0.15$, $a_c = c(1-m) = 0.5(1-0.7) = 0.15$, $a_m = (1-c)m = (1-0.5)0.7 = 0.35$ and $a_b = cm = 0.5 \cdot 0.7 = 0.35$. Notice that the sum of the fractional coverage of NPs is 1, as expected. The Neugebauer equations and Table 7.3 give then:

$$\begin{bmatrix} X_{Exerc.7.1(a)} \\ Y_{Exerc.7.1(a)} \\ Z_{Exerc.7.1(a)} \end{bmatrix} = 0.15 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.15 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.35 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.35 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.1(a)} = [56, 55, 101.5]$.

b) The two channels are halftoned according to dot-on-dot and $c < m$, see Section 7.6.2. Since the amount of cyan is less than magenta, using dot-on-dot means all 50% cyan is printed on magenta making 50% blue and 0% pure cyan. Of 70% magenta, is 50% turned to blue, remaining 20% pure magenta. The paper is the non-inked area, meaning $100 - 70 = 30\%$. Therefore, the fractional coverage of paper, pure cyan, pure magenta and blue are 30%, 0%, 20% and 50%, respectively. Notice that the sum of the fractional coverage of NPs is 1, as expected. The Neugebauer equations and Table 7.3 give then:

$$\begin{bmatrix} X_{Exerc.7.1(b)} \\ Y_{Exerc.7.1(b)} \\ Z_{Exerc.7.1(b)} \end{bmatrix} = 0.3 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.2 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.5 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.1(b)} = [59, 61, 103]$.

c) The two channels are halftoned according to dot-off-dot and $c + m > 1$, see Section 7.7.2 and Figure 7.17 (b). According to the discussion in Section 7.7.2, the amount of blue is $50 + 70 - 100 = 20\%$. The amount of cyan was 50%, having 20% of it turned to blue gives

$50 - 20 = 30\%$ pure cyan. The amount of magenta was 70%, having 20% of it turned to blue gives $70 - 20 = 50\%$ pure cyan. Since the amount of the inks is over 100%, obviously there won't be any non-inked (paper) area. Therefore, the fractional coverage of paper, pure cyan, pure magenta and blue are 0%, 30%, 50% and 20%, respectively. Notice that the sum of the fractional coverage of NPs is 1, as expected. The Neugebauer equations and Table 7.3 give then:

$$\begin{bmatrix} X_{Exerc.7.1(c)} \\ Y_{Exerc.7.1(c)} \\ Z_{Exerc.7.1(c)} \end{bmatrix} = 0 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.3 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.5 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.2 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.1(c)} = [53, 49, 100]$.

d) Since there are more non-inked (paper) area when dot-on-dot is used, it is easy to figure out that the color in part (b) is lighter than the other two. This can also be observed by looking at the CIEY-values, representing the perceptual luminance. The CIEY-value in part (b), which is 61, is greater than the two others, meaning that this is the lightest color. With the same reasoning, it is easy to figure out that dot-off-dot results in the darkest color of all three.

7.2 There are three primary inks, cyan, magenta and yellow, involved, making the number of NPs $2^3 = 8$. The NPs are: paper, pure cyan, pure magenta, pure yellow, red (magenta + yellow), green (cyan + yellow) and blue (cyan + magenta) and finally black (cyan + magenta + yellow).

a) Since the channels are halftoned **independently**, Demichel equations (Equation 7.3) can be used, giving: (notice $c = 0.5$, $m = 0.7$ and $y = 0.6$ according to the exercise)

$a_p = (1 - c)(1 - m)(1 - y) = 0.06$, $a_c = c(1 - m)(1 - y) = 0.06$, $a_m = (1 - c)m(1 - y) = 0.14$, $a_y = (1 - c)(1 - m)y = 0.09$, $a_r = (1-c)my = 0.21$, $a_g = c(1-m)y = 0.09$, $a_b = cm(1-y) = 0.14$, and $a_k = cmy = 0.21$. Notice that the sum of the fractional coverage of NPs is 1, as expected. The Neugebauer equations and Table 7.4 give then:

$$\begin{bmatrix} X_{Exerc.7.2(a)} \\ Y_{Exerc.7.2(a)} \\ Z_{Exerc.7.2(a)} \end{bmatrix} = 0.06 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.06 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.14 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.09 \begin{bmatrix} 70 \\ 80 \\ 20 \end{bmatrix} + 0.21 \begin{bmatrix} 50 \\ 30 \\ 10 \end{bmatrix} + 0.09 \begin{bmatrix} 50 \\ 80 \\ 40 \end{bmatrix} + 0.14 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix} + 0.21 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.2(a)} = [43.7, 42.7, 48.1]$.

b) The three channels are halftoned according to dot-on-dot. Since the amount of cyan is less than the other two, using dot-on-dot means all 50% cyan is printed on magenta and yellow making 50% black and 0% pure cyan, see Figure 9.1 (a). Of 60% yellow, is 50% turned to black, and the remaining 10% is printed on magenta (magenta+yellow=red), making 0% pure yellow, 10% red and 50% black. Of 70% magenta, is now 50% turned to black, and 10% turned to red and the remaining 10% is printed on paper, making 10% pure magenta, 10% red and 50% black. The paper is the non-inked area, meaning $100 - 70 = 30\%$. Therefore, the fractional coverage of paper, pure cyan, pure magenta, pure yellow, red, green, blue and black are (see Figure 9.1 (a)) 30%, 0%, 10%, 0%, 10%, 0%, 0% and 50%, respectively. Notice that the sum of the fractional coverage of NPs is 1, as expected. The Neugebauer equations and Table 7.4 give then: (only NPs with non-zero coverage are included)

$$\begin{bmatrix} X_{Exerc.7.2(b)} \\ Y_{Exerc.7.2(b)} \\ Z_{Exerc.7.2(b)} \end{bmatrix} = 0.3 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.1 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.1 \begin{bmatrix} 50 \\ 30 \\ 10 \end{bmatrix} + 0.5 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.2(b)} = [38, 36, 44]$.

c) In this exercise, firstly 50% cyan and 70% magenta are printed dot-off-dot, resulting in $100 - 70 = 30\%$ pure cyan, $100 - 50 = 50\%$ pure magenta and $50 + 70 - 100 = 20\%$ blue (and of course 0% paper), see Figure 9.1 (b). The yellow dots are now firstly printed on cyan if remained printed on magenta and finally on blue. Since the amount of yellow is 60% and the amount of pure cyan is 30%, then 30% of the yellow will be printed on pure cyan (making it green), and the rest 30% yellow will be printed on 50% pure magenta, making 30% red and 20% pure magenta, see Figure 9.1 (c). There won't be any yellow left to be printed on blue to make it black. As can be seen in figure (c), the fractional coverage of paper is 0. Therefore, as also seen in Figure 9.1 (c), of eight possible NPs, only four are present in this exercise, namely: pure magenta (coverage 0.2), red (magenta + yellow) (coverage 0.3), green (cyan + yellow) (coverage 0.3) and blue (cyan + magenta) (coverage 0.2). Notice specially that the sum of the coverages is 1, as expected. The Neugebauer equations and Table 7.4 give (those NPs that have zero coverage are excluded from this equation):

$$\begin{bmatrix} X_{Exerc.7.2(c)} \\ Y_{Exerc.7.2(c)} \\ Z_{Exerc.7.2(c)} \end{bmatrix} = 0.2 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 50 \\ 30 \\ 10 \end{bmatrix} + 0.3 \begin{bmatrix} 50 \\ 80 \\ 40 \end{bmatrix} + 0.2 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.2(c)} = [50, 49, 55]$.

7.3 We know that the channels have been halftoned using dot-off-dot, but we don't know whether $c + m \leq 1$ or not. Therefore, we need to study both cases (discussed in Section 7.7.1 and 7.7.2).

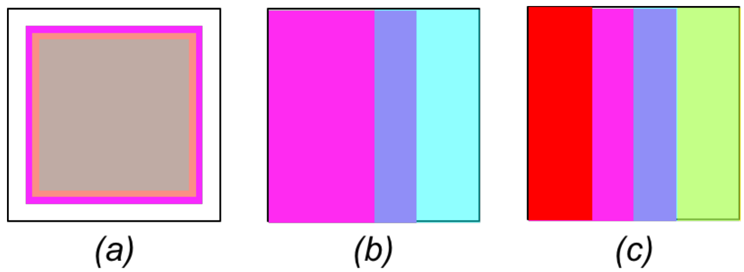


Figure 9.1: Solution-Exercise 7.2: (a) 50% cyan, 70% magenta and 60% yellow are printed dot-on-dot. (b) 50% cyan, 70% magenta are printed dot-off-dot. (c) 60% yellow is printed on the halftone in (b) according to the strategy explained in the exercise.

Case $c + m \leq 1$**:** This means that, $c + 0.8 \leq 1 \rightarrow c \leq 0.2$, and of course c cannot be negative. Therefore, in this case only those c inside $[0, 0.2]$ are accepted. Since the sum of the coverages is less than 100%, there won't be any dot-overlap, and therefore no blue. The coverage of paper, pure cyan, pure magenta and blue is $1 - (c + 0.8)$, c, 0.8 and 0, respectively. Since only the CIEY is given, we write the Neugebauer equation using Y:

$$(1 - c - 0.8)Y_p + cY_c + 0.8Y_m = 48 \rightarrow 100(1 - c - 0.8) + 80c + 30 \cdot 0.8 = 48,$$

which gives: $c = -0.2$, which is not accepted, because in this case c should be in the interval $[0, 0.2]$.

Case $c + m > 1$**:** This means that, $c + 0.8 > 1 \rightarrow c > 0.2$, and of course c cannot be over 100%. Therefore, in this case only those c inside $[0.2, 1]$ are accepted. As discussed in Section 7.7.2, the coverage of paper, pure cyan, pure magenta and blue in this case is 0, $1 - 0.8 = 0.2$, $1 - c$ and $c + 0.8 - 1$, respectively. Since only the CIEY is given, we write the Neugebauer equation using Y:

$$0.2Y_c + (1 - c)Y_m + (c + 0.8 - 1)Y_b = 48 \rightarrow 16 + 30 - 30c + 50c - 10 = 48,$$

which gives $c = 0.6$, which is inside the interval $[0.2, 1]$. Therefore, the answer is $c = 60\%$.

7.4 a) From the information given in the exercise and Figure 7.19, it is easy to figure out the coverage of NPs. The coverage of the overlap (blue) is 30% according to the exercise. Cyan's coverage is 60%, of which 30% becomes blue, making the coverage of pure cyan 30%. Magenta's coverage is 40%, of which 30% becomes blue, making the coverage of pure magenta 10%. The rest of the surface is paper, meaning $a_p = 1 - 0.3 - 0.3 - 0.1 = 0.3$. Therefore, the coverage of paper, pure cyan, pure magenta and blue is 0.3, 0.3, 0.1 and 0.3, respectively. The Neugebauer equations and Table 7.4 give then:

$$\begin{bmatrix} X_{Exerc.7.4(a)} \\ Y_{Exerc.7.4(a)} \\ Z_{Exerc.7.4(a)} \end{bmatrix} = 0.3 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.3 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.1 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.4(a)} = [60, 72, 103]$.

b) From part (a), we know that the coverage of paper, pure cyan, pure magenta and blue is 0.3, 0.3, 0.1 and 0.3, respectively. Now, 40% yellow is added independently and therefore probability formulas can be used. For example, the probability for non-inked area (paper) after addition of yellow is to have paper and not yellow.

The probability for paper is 0.3 (from part (a)) and therefore: $a_p = 0.3(1-y) = 0.3 \cdot (1-0.4) = 0.18$.

Pure cyan occurs when there is cyan and not yellow. The probability for cyan is 0.3 (from part (a)) and therefore: $a_c = 0.3(1-y) = 0.18$

Pure magenta occurs when there is magenta and not yellow. The probability for magenta is 0.1 (from part (a)) and therefore: $a_m = 0.1(1-y) = 0.06$

Pure yellow occurs when there is paper and yellow. The probability for paper is 0.3 (from part (a)) and therefore: $a_y = 0.3y = 0.12$

Red (magenta + yellow) occurs when there is magenta and yellow. The probability for magenta is 0.1 (from part (a)) and therefore: $a_r = 0.1y = 0.04$

Green (cyan + yellow) occurs when there is cyan and yellow. The probability for cyan is 0.3 (from part (a)) and therefore: $a_g = 0.3y = 0.12$

Blue (cyan + magenta) occurs when there is blue and not yellow. The probability for blue is 0.3 (from part (a)) and therefore: $a_b = 0.3(1-y) = 0.18$

Black (cyan + magenta + yellow) occurs when there is blue and yellow. The probability for blue is 0.3 (from part (a)) and therefore: $a_k = 0.3y = 0.12$

The Neugebauer equations and Table 7.4 give then:

$$\begin{bmatrix} X_{Exerc.7.4(b)} \\ Y_{Exerc.7.4(b)} \\ Z_{Exerc.7.4(b)} \end{bmatrix} = 0.18 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.18 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.06 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.12 \begin{bmatrix} 70 \\ 80 \\ 20 \end{bmatrix} +$$
$$0.04 \begin{bmatrix} 50 \\ 30 \\ 10 \end{bmatrix} + 0.12 \begin{bmatrix} 50 \\ 80 \\ 40 \end{bmatrix} + 0.18 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix} + 0.12 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.4(b)} = [52.4, 63.6, 69.4]$.

c) From part (a), we know that the coverage of paper, pure cyan, pure magenta and blue is 0.3, 0.3, 0.1 and 0.3, respectively. The yellow dots are now added using dot-off-dot and black has to be avoided as much as possible. The coverage of yellow is 40%. Since dot-off-dot is used, the yellow dots have to be first printed on paper. Since the coverage of paper is 30%, then all paper will be covered by yellow, making $a_p = 0$ and the coverage of pure yellow $a_y = 0.3$. The rest of yellow, i.e. 10%, is printed on cyan and magenta, and since the coverage of cyan and magenta together is 40%, then it is possible to avoid yellow on blue in order to avoid black (according to the exercise black has to be avoided if possible), meaning that the coverage of blue is unchanged $a_b = 0.3$ and the coverage of black $a_k = 0$. Now, the rest of yellow (10%) is printed on cyan and magenta. This means that 10% yellow is distributed to 30% cyan and 10% magenta. Notice that the probability of the rest of yellow is 10% over the whole surface, but here it is only added to cyan and magenta (sum of coverage 40%), meaning that the probability of yellow over cyan and magenta is $0.1/0.4 = 0.25$. Therefore:

The probability of paper is already discussed: $a_p = 0$

The probability of having pure cyan is equal to the probability of having cyan (0.3 from part (a)) and not yellow (1-0.25), giving: $a_c = 0.3(1 - 0.25) = 0.225$

The probability of having pure magenta is equal to the probability of having magenta (0.1 from part (a)) and not yellow (1-0.25), giving: $a_m = 0.1(1 - 0.25) = 0.075$

The probability of having pure yellow is 0.3, as discussed above: $a_y = 0.3$

The probability of having red is equal to the probability of having magenta (0.1 from part (a)) and yellow (0.25), giving: $a_r = 0.1 \cdot 0.25 = 0.025$

The probability of having green is equal to the probability of having cyan (0.3 from part (a)) and yellow (0.25), giving: $a_g = 0.3 \cdot 0.25 =$

0.075

The probability of having blue is 0.3 as discussed above: $a_b = 0.3$

The probability of having black is 0, as discussed above: $a_k = 0$

The Neugebauer equations and Table 7.4 give then: (the Nps having zero coverage are excluded from the equation)

$$\begin{bmatrix} X_{Exerc.7.4(c)} \\ Y_{Exerc.7.4(c)} \\ Z_{Exerc.7.4(c)} \end{bmatrix} = 0.225 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.075 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 70 \\ 80 \\ 20 \end{bmatrix} + 0.025 \begin{bmatrix} 50 \\ 30 \\ 10 \end{bmatrix} + 0.075 \begin{bmatrix} 50 \\ 80 \\ 40 \end{bmatrix} + 0.3 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.4(c)} = [53.75, 66.00, 69.25]$.

7.5 Threshold halftoning was discussed and explained in Section 3.2.1. All three threshold matrices in this exercise are 4×4, with threshold values starting from 1 and ending in 16. Therefore, all three matrices represent 17 levels of gray and therefore should be normalized by dividing their elements by 17. Then, thresholding means to compare each pixel value in the image with the corresponding threshold value in the threshold matrix. If the pixel value is greater than the threshold value, the halftone will have a 1 (printed dot) at that position and otherwise 0 (no print at that position).

a) According to the exercise, channels C and M are constant images holding 0.4 and 0.6 in all their pixels, respectively. Channel Y can be ignored because all its elements are 0.

If channel C is thresholded with this threshold matrix, then the result, call it C_{tr}, will hold a 1 where 0.4 is greater than the threshold value. Notice that the values in this threshold matrix are divided by 17. Since $6/17 = 0.353$ and $7/17 = 0.412$, all positions holding value less than 7 in this threshold matrix, will have 1 in C_{tr} and the rest holds 0.

If channel M is thresholded with this threshold matrix, then the result, call it $M_{tr(a)}$, will hold a 1 where 0.6 is greater than the threshold value. Notice that the values in this threshold matrix are divided by 17 as well. Since $10/17 = 0.588$ and $11/17 = 0.647$, all positions holding value less than 11 in this threshold matrix, will have 1 in $M_{tr(a)}$ and the rest holds 0.

Therefore:

$$C_{tr} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } M_{tr(a)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

As can be seen in these two halftones, the coverage of cyan and magenta will be 6/16 and 10/16, respectively. It is also easy to figure out that, since both channels are halftoned by the same matrix, the printing strategy is dot-on-dot. Now, there are two ways to figure out the coverage of the four NPs, we explain both here:

method 1: The coverage of cyan and magenta is 6/16 and 10/16 respectively. Using dot-on-dot means, (since $c < m$): the coverage of paper is $a_p = 1 - 10/16 = 6/16$, the coverage of pure cyan is $a_c = 0$, the coverage of pure magenta is $a_m = 10/16 - 6/16 = 4/16 = 0.25$, and the coverage of blue is $a_b = 6/16$.

method 2: By directly looking at the matrices C_{tr} and $M_{tr(a)}$, we can now figure out that when these two halftones are printed on top of each other we will have: Paper: those positions that hold 0 in both C_{tr} and $M_{tr(a)}$ result in no-ink and thereby paper. There are 6 such positions, giving $a_p = 6/16$. Pure cyan: those positions that hold 1 in C_{tr} and 0 in $M_{tr(a)}$ result in pure cyan. There are 0 such positions, giving $a_c = 0$. Pure magenta: those positions that hold 0 in C_{tr} and 1 in $M_{tr(a)}$ result in pure magenta. There are 4 such positions, giving $a_m = 4/16 = 0.25$. Blue: those positions that hold 1 in both C_{tr} and $M_{tr(a)}$ result in blue. There are 6 such positions, giving $a_b = 6/16$

As seen, both methods give the same results. The Neugebauer equations and Table 7.3 give then:

$$\begin{bmatrix} X_{Exerc.7.5(a)} \\ Y_{Exerc.7.5(a)} \\ Z_{Exerc.7.5(a)} \end{bmatrix} = (6/16) \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.25 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + (6/16) \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.5(a)} = [63.75, 63.75, 103.75]$.

b) Channel C is thresholded with the same matrix as in (a), giving the same halftone C_{tr}. Channel M is thresholded with the threshold matrix given in this part, giving:

$$C_{tr} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } M_{tr(b)} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

As can be seen in these two halftones, the coverage of cyan and magenta are 6/16 and 10/16, respectively (exactly as in part (a)). Now, there are two ways to figure out the coverage of the four NPs, we explain both here:

method 1: It is easy to figure out that, the matrix in part (b) is equal to (17 - matrix in part (a)), meaning that where matrix (a) holds 1, matrix (b) holds 16, and where matrix (a) holds 2 matrix (b) holds 15, and so on. This means, when a dot is printed in channel C, the position will be avoided to be filled by M as much as possible. This means that the printing strategy is dot-off-dot. Since $c + m \leq 1$, the coverage of pure cyan and pure magenta is $a_c = 6/16$ and $a_m = 10/16$, respectively. Using dot-off-dot means, (since $c+m \leq 1$): the coverage of paper is $a_p = 1-(c+m) = 0$, and the coverage of blue is also $a_b = 0$.

method 2: By direct looking at the matrices C_{tr} and $M_{tr(b)}$, we can now figure out that when these two halftones are printed on top of each other we will have: Paper: those positions that hold 0 in both C_{tr} and $M_{tr(b)}$ result in no-ink and thereby paper. There are 0 such positions, giving $a_p = 0/16 = 0$. Pure cyan: those positions that hold 1 in C_{tr} and 0 in $M_{tr(b)}$ result in pure cyan. There are 6 such positions, giving $a_c = 6/16$. Pure magenta: those positions that hold 0 in C_{tr} and 1 in $M_{tr(b)}$ result in pure magenta. There are 10 such positions, giving $a_m = 10/16$. Blue: those positions that hold 1 in both C_{tr} and $M_{tr(b)}$ result in blue. There are 0 such positions, giving $a_b = 0/16$.

As seen, both methods give the same results. The Neugebauer equations and Table 7.3 give then:

$$\begin{bmatrix} X_{Exerc.7.5(b)} \\ Y_{Exerc.7.5(b)} \\ Z_{Exerc.7.5(b)} \end{bmatrix} = 0 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + (6/16) \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + (10/16) \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.5(b)} = [56.25, 48.75, 100.00]$.

c) Channel C is thresholded with the same matrix as in (a), giving the same halftone C_{tr}. Channel M is thresholded with the threshold matrix given in this part, giving:

$$C_{tr} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } M_{tr(c)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

As can be seen in these two halftones, the coverage of cyan and magenta will be 6/16 and 10/16, respectively (exactly as in part (a) and (b)). Here, it is neither dot-on-dot nor dot-off-dot and the coverage of NPs can be directly found by the two halftones:

By direct looking at the matrices C_{tr} and $M_{tr(c)}$, we can now figure out that when these two halftones are printed on top of each other we will have: Paper: those positions that hold 0 in both C_{tr} and $M_{tr(c)}$ result in no-ink and thereby paper. There are 5 such positions, giving $a_p = 5/16$. Pure cyan: those positions that hold 1 in C_{tr} and 0 in $M_{tr(c)}$ result in pure cyan. There is 1 such position, giving $a_c = 1/16$. Pure magenta: those positions that hold 0 in C_{tr} and 1 in $M_{tr(c)}$ result in pure magenta. There are 5 such positions, giving $a_m = 5/16$. Blue: those positions that hold 1 in both C_{tr} and $M_{tr(c)}$ result in blue. There are 5 such positions, giving $a_b = 5/16$. The Neugebauer equations and Table 7.3 give then:

$$\begin{bmatrix} X_{Exerc.7.5(c)} \\ Y_{Exerc.7.5(c)} \\ Z_{Exerc.7.5(c)} \end{bmatrix} = (5/16) \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + (1/16) \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + (5/16) \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + (5/16) \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.5(c)} = [62.50, 61.25, 103.13]$.

7.6 a) As discussed in Section 7.6, there are two different cases that should be considered here: $c \geq m$ and $c < m$. Here we just show one case:

$c \geq m$**:** As discussed in Section 7.6.1, the coverage of paper, pure cyan, pure magenta and blue is $1 - c$, $c - m$, 0 and m, respectively. The Neugebauer equations and Table 7.3 give:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = (1-c) \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + (c-m) \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + m \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

According to the exercise, this should result in $CIEXYZ = [64, 77, 104.5]$, giving the following equation system:

$$\begin{cases} 40c + 10m = 26 \\ 20c + 30m = 23 \\ 10c = 5.5 \end{cases}$$

The third equation gives $c = 0.55$. Setting that in the second equation gives $m = 0.4$. Now, we need to make sure that these two also satisfy the first equation. They obviously do (check it). Since the condition $c \geq m$ is satisfied, the answer to the question is yes, with $c = 55\%$ and $m = 40\%$.

Since we have answered yes, then there is no need to check the other case ($c < m$). If we couldn't find the answer testing the previous case, we had to check this case out to make sure.

b) As discussed in Section 7.7, there are two different cases that should be considered here: $c + m \leq 1$ and $c + m > 1$. By looking at the resulting color, i.e. $CIEXYZ = [64, 77, 104.5]$, we can notice that the Z-value is larger than 100. If we look at the Z-values of the NPs in Table 7.3, we can see that only the paper has a Z-value larger than 100 and the other three have Z-value=100. This means that, there must be paper involved, otherwise the Z-value of the resulting color would have been 100. Therefore, the case $c + m > 1$ can be ignored for this exercise, because it cannot give this resulting color. Hence, only one case needs to be checked. As discussed in Section 7.7.1, since $c + m \leq 1$, the coverage of paper, pure cyan, pure magenta and blue is $1 - (c + m)$, c, m and 0, respectively. The Neugebauer equations and Table 7.3 give:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = (1 - c - m) \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + c \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + m \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

According to the exercise, this should result in $CIEXYZ = [64, 77, 104.5]$, giving the following equation system:

$$\begin{cases} 40c + 30m = 26 \\ 20c + 70m = 23 \\ 10c + 10m = 5.5 \end{cases}$$

But this equation system has no solution. Why? Find c and m by for example solving the second and the third equations. These two equations give: $c = 0.31$ and $m = 0.24$, but these two don't satisfy the first equation, check it. Therefore, the answer to this question is no. Just notice that, if it wasn't easy to figure out that $c + m > 1$ cannot be the case, this case had to be checked as well.

7.7 Like in Exercise 7.6, if we look at the Z-value of the resulting color of the printed halftone, i.e. $CIEXYZ = [52.5, 59.5, 100]$, we see that the Z-value is 100. If we look at the Z-values of NPs in Table 7.3, we can see that only the paper has a Z-value not equal to 100 and the other three have Z-value=100. This means that paper cannot be involved in the printed halftone, because if it was involved the resulting Z-value would have not been 100. Since the paper is not involved, independent halftoning and dot-on-dot cannot be the printing strategy used here (unless either c or m is 100%, which won't make any difference between independent, dot-on-dot and dot-off-dot). The only possibility is therefore, dot-off-dot, and again since paper is not involved, then only the case $c+m > 1$ is possible. As discussed in Section 7.7.2, since $c+m > 1$, the coverage of paper, pure cyan, pure magenta and blue is 0, $1-m$, $1-c$ and $c+m-1$, respectively.

The Neugebauer equations and Table 7.3 give:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = 0 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + (1-m) \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + (1-c) \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + (c+m-1) \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

According to the exercise, this should result in $CIEXYZ = [52.5, 59.5, 100]$, giving the following equation system:

$$\begin{cases} 20c + 10m = 17.5 \\ -20c + 30m = 0.5 \\ 100 = 100 \end{cases}$$

Solving the two first equations gives $c = 0.65$ and $m = 0.45$. Since $c + m > 1$ is valid, then this solution is accepted. Therefore, the answer is: $c = 65\%$ cyan and $m = 45\%$ magenta have been printed using dot-off-dot printing strategy.

7.8 a) Similar exercises have been explained (for example Exercise 7.2 (b)). By similar reasoning, we can figure out that, the fractional coverage of paper, pure cyan, pure magenta, pure yellow, red, green, blue and black are $a_p = 20\%$, $a_c = 20\%$, $a_m = 0\%$, $a_y = 0\%$, $a_r = 0\%$, $a_g = 0\%$, $a_b = 30\%$ and $a_k = 30\%$, respectively. Notice that the sum of the fractional coverage of NPs is 1, as expected. The Neugebauer equations and Table 7.4 give then: (only NPs with non-zero coverage are included)

$$\begin{bmatrix} X_{Exerc.7.8(a)} \\ Y_{Exerc.7.8(a)} \\ Z_{Exerc.7.8(a)} \end{bmatrix} = 0.2 \begin{bmatrix} 90 \\ 100 \\ 110 \end{bmatrix} + 0.2 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix} + 0.3 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.8(a)} = [40, 51, 72]$.

b) Firstly 80% cyan and 60% magenta have been printed using dot-off-dot. We have seen many similar exercises. Therefore, since $c+m > 1$, the coverage of paper, pure cyan, pure magenta and blue is 0, 0.4, 0.2 and 0.4, respectively. Now, the yellow dots, 30% coverage, are added independently and therefore probability formulas can be used:

Paper (non-inked area) cannot occur, because there was no paper after cyan and magenta was printed and therefore $a_p = 0$.

Pure cyan occurs when there is cyan and not yellow. The probability for cyan is 0.4 and therefore: $a_c = 0.4(1-y) = 0.28$.

Pure magenta occurs when there is magenta and not yellow. The probability for magenta is 0.2 and therefore: $a_m = 0.2(1-y) = 0.14$.

Pure yellow occurs when there is paper and yellow. Since there is no paper, then: $a_y = 0$.

Red (magenta + yellow) occurs when there is magenta and yellow. The probability for magenta is 0.2 and therefore: $a_r = 0.2y = 0.06$.

Green (cyan + yellow) occurs when there is cyan and yellow. The probability for cyan is 0.4 and therefore: $a_g = 0.4y = 0.12$.

Blue (cyan + magenta) occurs when there is blue and not yellow. The probability for blue is 0.4 and therefore: $a_b = 0.4(1 - y) = 0.28$.

Black (cyan + magenta + yellow) occurs when there is blue and yellow. The probability for blue is 0.4 and therefore: $a_k = 0.4y = 0.12$.

The Neugebauer equations and Table 7.4 give then: (the Nps having zero coverage are excluded from the equation)

$$\begin{bmatrix} X_{Exerc.7.8(b)} \\ Y_{Exerc.7.8(b)} \\ Z_{Exerc.7.8(b)} \end{bmatrix} = 0.28 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.14 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.06 \begin{bmatrix} 50 \\ 30 \\ 10 \end{bmatrix} + 0.12 \begin{bmatrix} 50 \\ 80 \\ 40 \end{bmatrix} + 0.28 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix} + 0.12 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.8(b)} = [42.6, 52.0, 75.4]$.

c) As stated in the exercise, 80% cyan and 60% magenta have first been printed using dot-off-dot. Therefore, the coverage of paper, pure cyan, pure magenta and blue is 0, 0.4, 0.2 and 0.4, respectively. The yellow dots are now added independently but black has to be avoided as much as possible. Since the coverage of yellow, 30%, is less than the coverage of cyan and magenta (40+20=60%), then black can be completely avoided. Therefore, 30% yellow is printed on cyan and magenta independently. This means 30% yellow is distributed to 40% cyan and 20% magenta. Notice that the probability of yellow is 30% over the whole surface, but here it is only added to cyan and magenta (sum of coverage 60%), meaning that the probability of yellow over cyan and magenta is $0.3/0.6 = 0.5$. Therefore:

The probability of having paper is equal to the probability of having paper and not yellow, but there is no paper to begin with, giving: $a_p = 0$

The probability of having pure cyan is equal to the probability of having cyan and not yellow, giving: $a_c = 0.4(1 - 0.5) = 0.2$

The probability of having pure magenta is equal to the probability of having magenta and not yellow, giving: $a_m = 0.2(1-0.5) = 0.1$

The probability of having pure yellow is equal to the probability of having paper and yellow, but there is no paper, giving: $a_y = 0$

The probability of having red is equal to the probability of having magenta and yellow, giving: $a_r = 0.2 \cdot 0.5 = 0.1$

The probability of having green is equal to the probability of having cyan and yellow, giving: $a_g = 0.4 \cdot 0.5 = 0.2$

The probability of having blue is equal to the probability of having blue and not yellow, and since yellow is avoided to be printed on blue, then the coverage of blue will be unchanged, giving: $a_b = 0.4$

The probability of having black is 0, because yellow was avoided to be printed on blue, giving: $a_k = 0$

The Neugebauer equations and Table 7.4 give then: (the Nps having zero coverage are excluded from the equation)

$$\begin{bmatrix} X_{Exerc.7.8(c)} \\ Y_{Exerc.7.8(c)} \\ Z_{Exerc.7.8(c)} \end{bmatrix} = 0.2 \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix} + 0.1 \begin{bmatrix} 60 \\ 30 \\ 100 \end{bmatrix} + 0.1 \begin{bmatrix} 50 \\ 30 \\ 10 \end{bmatrix} + 0.2 \begin{bmatrix} 50 \\ 80 \\ 40 \end{bmatrix} + 0.4 \begin{bmatrix} 40 \\ 50 \\ 100 \end{bmatrix}$$

The above equation gives $CIEXYZ_{Exerc.7.8(c)} = [47, 58, 79]$.