2 Image Enhancement in Spatial Domain

Image enhancement is a process, rather a preprocessing step, through which an original image is made suitable for a specific application. The application scenarios may vary from thermal image to X-ray image and accordingly the process of image enhancement would differ. Generally, the effect of image enhancement can be perceived visually. Even to address/handle the regular artifacts due to geometric transformations of images, image enhancement is a must. The spatial domain refers to the 2D image plane in terms of pixel intensities. When the image is enhanced by modifying the pixel intensities directly (not as an effect of some other parameter tuning in a different domain), the method is considered as spatial domain image enhancement methodology. Otherwise, the image can be transformed to some other domain—like one 2D image can be transferred to a 2D frequency domain by discrete Fourier transform (DFT). To achieve an enhanced image, the Fourier coefficients are modified. That family of image enhancement methodologies is considered as frequency domain image enhancement which is discussed in the subsequent chapters. Whatever be the domain of image enhancement (either spatial or frequency domain), by the term *image enhancement* we mean improvement of the appearance of an image (in all sense including human perception and machine perception) by increasing the dominance of some features, or by decreasing the ambiguity between different regions of the image. In most cases, the enhancement of certain features is achieved at the cost of suppressing few other features. Broadly, the image enhancement in the spatial domain is divided into four categories:

1. Contrast manipulation/intensity transformation
2. Image smoothing
3. Image sharpening
4. Image resampling

In the current chapter, processing of images through histogram, the intensity distribution is presented after a brief discussion on basic gray-level transformation. Histogram is presented here in terms of probability distribution function (PDF). For histogram-based image enhancement, the process of linearizing the cumulative

A. Das, *Guide to Signals and Patterns in Image Processing,*
DOI 10.1007/978-3-319-14172-5_2

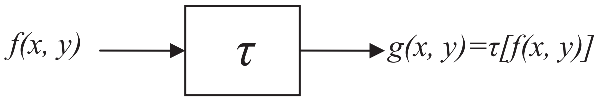

Fig. 2.1 The original image $f(x, y)$ with pixel intensity r is processed through "τ" to get the enhanced image $g(x, y)$ with pixel intensity s

density function (CDF) is described in the process of histogram equalization. Next, the process of image filtering in spatial domain is introduced with the help of convolution and correlation. Finally, resampling of images in order to achieve visually enhanced image after geometric transformations is discussed. First, the basic interpolation techniques are discussed in 1D in terms of B-splines, then the same concepts are extended to 2D for image interpolation/resampling (Fig. 2.1).

2.1 Intensity Transformations

Spatial domain refers to an aggregate of pixels consisting in an image. Spatial domain image processing is processing over the pixels directly as expressed in the following equation:

$$g(x, y) = \tau[f(x, y)] \tag{2.1}$$

where $f(x, y)$ and $g(x, y)$ are the input and the processed image through the mathematical mapping "τ" defined over (x, y). When this mapping/operator is applied on any arbitrary point of coordinate (x, y) to get the processed point at the same coordinate (x, y), this mathematical mapping "τ" is called as intensity operator or intensity mapping or gray-level transformation. If r and s be the intensity of the arbitrary point before and after transformation, Eq. 2.1 can be rewritten as

$$s = \tau[r] \tag{2.2}$$

Patterns/shapes of different intensity transformations are discussed with their effect on the images. In Fig. 2.2, different gray-level transformation characteristics are shown in terms of graphs (Fig. 2.2b, d, f, h).

2.1.1 Linear Transformation

In the transformation characteristics, the x-axes and y-axes of the graphs represent gray levels (intensity levels) of input original image and transformed image, respectively, ranging from 0 to 255. Hence, one linear characteristic (unit ramp function) just maps the intensity of the input image to the same intensity of the transformed image. It maintains all the intensities unmodified as shown in Fig. 2.2a and b.

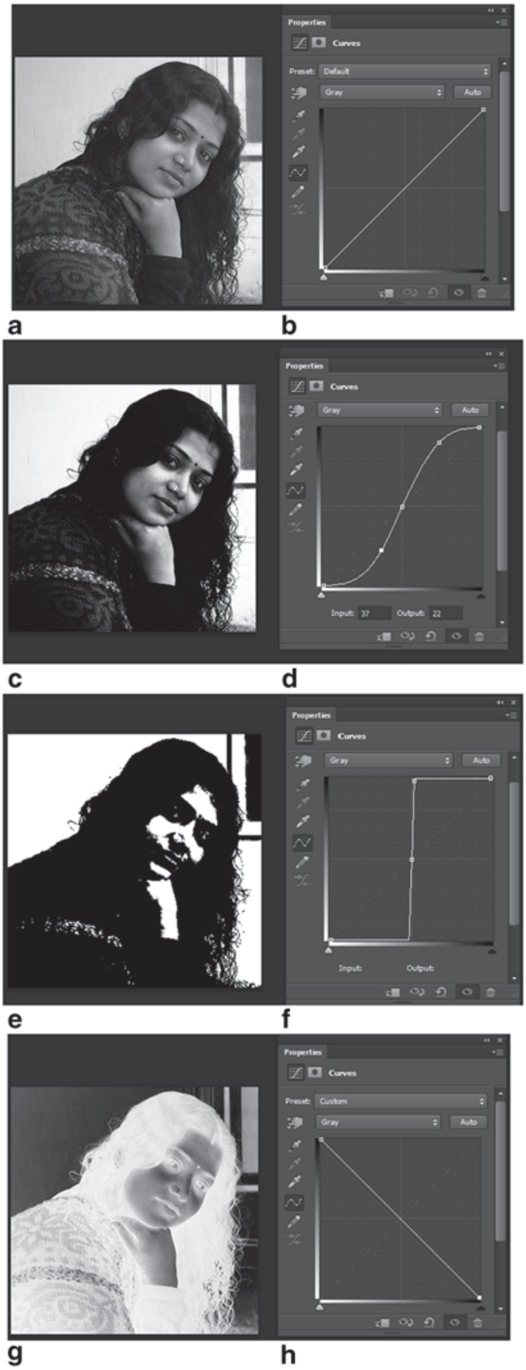

Fig. 2.2 Intensity (*gray-level*) transformations: **a** and **b** linear transformation, **c** and **d** contrast stretching, **e** and **f** thresholding for binarization, **g** and **h** negative transform

2.1.2 Contrast Stretching and Thresholding

Referring to the graph of Fig. 2.2d, the transformation clearly signifies that the contrast of the image is increased by mapping all the intensities higher than 128 to a compressed narrow range of intensities near white, and all the intensities less than 128 to a compressed narrow range of intensities near black, as shown in Fig. 2.2c. This transformation is therefore called as *contrast stretching*. In the limiting case where the higher intensities ($r > 128$) are mapped to the highest intensity ($s = 255$) and lower intensities ($r < 128$) are mapped to the lowest intensity ($s = 0$), the transformation (Fig. 2.2f) takes the form of *thresholding for binarization,* as the output image would have only two intensities (binary image)—white and black (Fig. 2.2e).

2.1.3 Negative Intensity Transform

In this case, the slope of the transformation is made −1 instead of +1 with respect to linear transformation, as shown in Fig. 2.2h. This *negative transformation* maps the intensities with the rule:

$$s = 255 - r \tag{2.3}$$

This transformation generates a complete negative image (as shown in Fig. 2.2g) by mapping the higher intensities to lower intensity and vice versa.

Negative intensity transformation is very useful in specialized applications where the intensity details are embedded/hidden in dark regions of an image. Digital mammogram is a method of analyzing lesions inside breast tissues. To analyze the breast tissue in digital mammogram, negative intensity transformation is very useful [4]. From Fig. 2.3, it is clearly observed that the tissues can be analyzed from the negative image even by visual attack. Note that here the content of information is exactly the same in both the original and transformed image, the representation of information has been changed for the ease of analysis.

2.1.4 Logarithmic Intensity Transformation

The logarithmic intensity transformation is defined by the following equation:

$$s = c\log(1 + r) \tag{2.4}$$

In the equation, c is an arbitrary positive constant, r and s are the intensities of the original and transformed images, respectively, with intensity profile 0 through 255. This intensity transformation maps a narrow range of lower intensity value in the original input image to a wider range of output levels and a wider range of higher intensity values to a narrower range of output gray levels. Hence, this transformation would be useful where expansion of dark pixel and compression of brighter

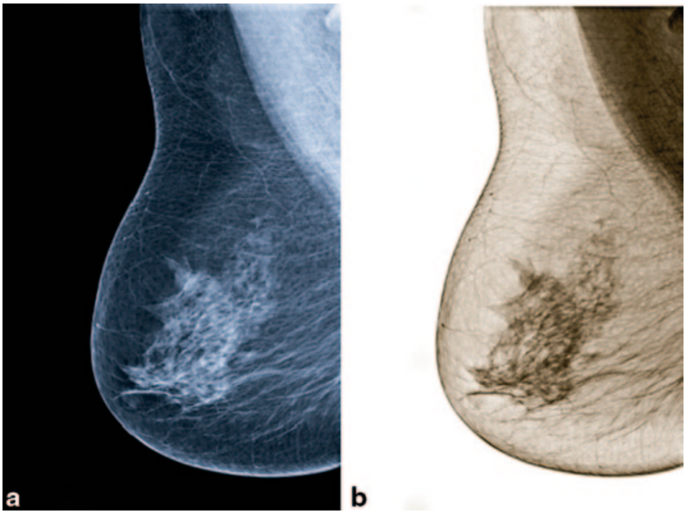

Fig. 2.3 Digital mammogram: **a** original image, **b** negative intensity transformed image

pixels are required. The behavior of inverse logarithmic intensity transformation is exactly opposite. Any transformation characteristic curve with the same pattern will behave exactly as the log transformation does, i.e., expanding one part of the intensity profile and compressing the other. Power-law transformation, which is discussed in the next subsection, is a more generalized transformation to achieve this kind of behavior in transformation

2.1.5 Power-Law Intensity Transform and Gamma Correction

The power-law intensity transformation [10] is defined by the following equation:

$$s = cr^{\gamma} \tag{2.5}$$

In the equation, c and γ are the arbitrary positive constants, r and s are the intensities of the original and transformed images, respectively, with the intensity profile 0 through 255. This intensity transformation characteristic is shown in Fig. 2.4 with varied γ. Like logarithmic intensity transformation, the power-law transformation with fractional γ value maps a narrow range of dark values (low intensity) to a wider range of output values and wider range of lighter values (high intensity) to a narrower range of output. Moreover, in the power-law transformation, we can tune the characteristic curve by tuning γ, and hence we can change the narrowness of darker intensities and wideness of the lighter intensities of the input image. As understood, with $\gamma = 1$ unit ramp characteristic would be realized which is identity transformation and $\gamma > 1$ will have exactly an opposite effect with respect to fractional $(\gamma < 1)$ values of γ.

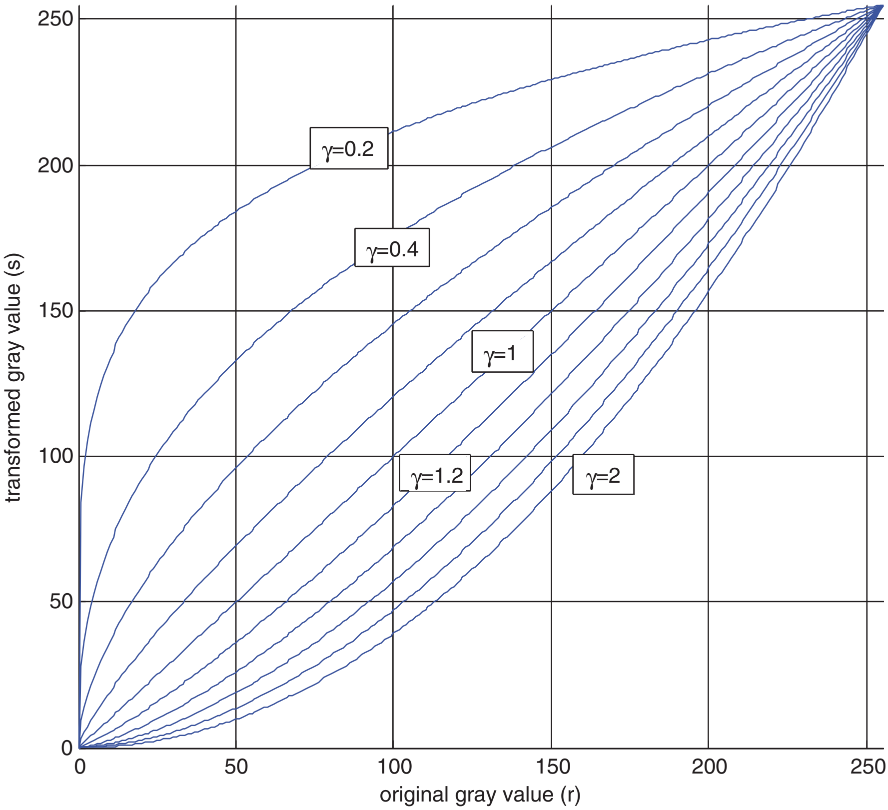

Fig. 2.4 Transfer characteristics for power-law intensity transformation

It is to be noted that, simply by applying the power-law equation, the dynamic intensity range of the transformed image would not be exactly [0 255]. To achieve this, the transformed intensities need to be normalized with respect to maximum intensity as depicted in the following equation.

$$s = s[].^{*}(255 / \max(s[])); \tag{2.6}$$

A number of devices used for image acquisition, printing, and display respond according to power law. Hence, based on the exponent or the tuning parameter γ (*gamma*), the procedure of correcting the power-law phenomena is called as *gamma correction*. The cathode ray tube (CRT) works depending on the intensity to voltage response which is a power-law relationship. For CRT, the exponent gamma varies from 1.8 to 2.5. Considering an arbitrary value of γ in this range (say, $\gamma = 2.5$), we can understand the system behavior. From the power-law characteristics (Fig. 2.4), we can understand that the wider band of lower intensity values is mapped to a narrower band of lower intensity, which generates a darker image with respect to the

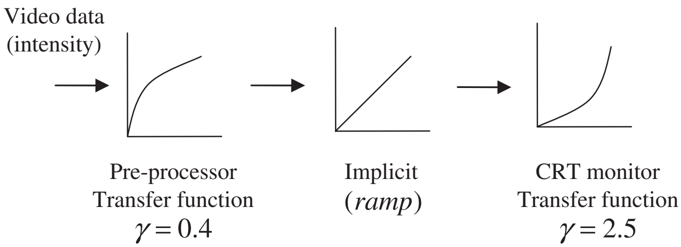

Fig. 2.5 Gamma correction for CRT monitor

Table 2.1 Histogram of a 3-bit encoded (eight levels) image

i	0	1	2	3	4	5	6	7
f[i]	10	8	18	22	14	19	0	9

original one. To handle this, we need to add a preprocessor transfer function of inverse characteristics $(i.e., s = r^{\frac{1}{2.5}} = r^{0.4})$ before sending the signal to the CRT. The combination of these two transfer functions (transfer functions of preprocessor unit and that of the CRT) will realize a unit ramp function, which ensures the realization of the original gray intensity values as depicted in Fig. 2.5.

2.2 Histogram of an Image

In the image, we can see two different illustrations of the term "frequency" one from the pattern perspective and the other from the signal perspective. Here, we discuss the concept of frequency in one form and the other in the next chapter. The current concept is very straightforward. Here, the term "frequency" signifies the number of occurrences of a particular gray-level intensity (or intensity of each color plane of a color image). If we consider an image of 256-gray-level intensity values (from 0 to 255 for 8-bit encoding), the frequency or count will be an array of 256 elements. Each array index (say i of array $f[i]$) would represent the number of occurrences of the intensity "i" in the whole image.

Let us consider an image of size 10×10 (10 rows and 10 columns; therefore 100 pixels) whose pixel intensity varies from 0 to 7 (3-bit encoding). The numbers of occurrences ($f[i]$) of each of the pixel intensities (i) are listed in Table 2.1.

Obviously, the sum of all the contents of the array $f[i]$ would be 100, the total number of pixels. The plot of this frequency of occurrence with respect to the intensity levels is called as the *histogram*.

$$\sum_{i=0}^{L-1} f[i] = M \times N \tag{2.7}$$

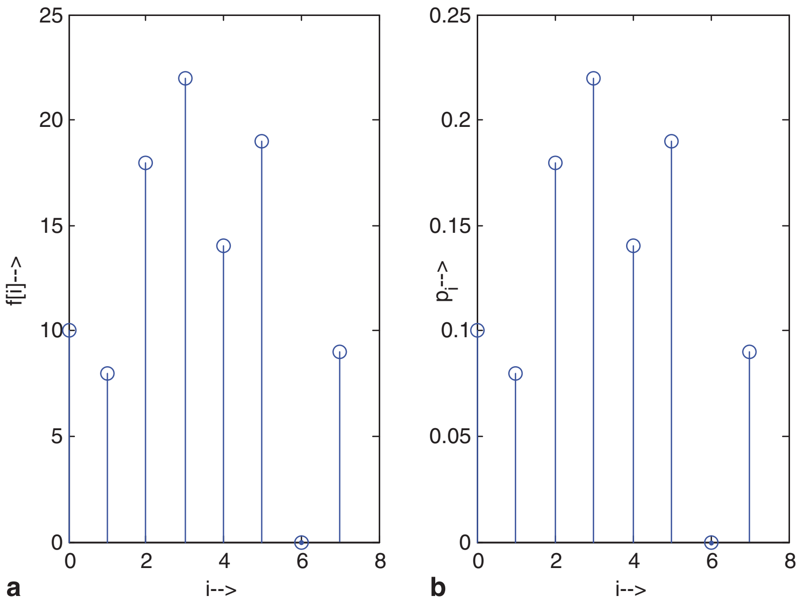

Fig. 2.6 Histogram of an image: **a** intensity versus frequency of occurrence, **b** PDF with respect to intensity as a random variable

In the equation above, L is the number of intensity levels in the image of M number of rows and N number of columns. As the total number of pixels in the image is constant, the frequency $f[i]$ can also be expressed in terms of PDF (Probability Distribution Function) and then the histogram can be judged/analyzed in terms of a statistical distribution. Moreover, the statistical moments can also be leveraged to interpret the histogram in terms of image characteristics.

Probability is defined as follows: If an event can occur in n_i different ways out of a total number of N possible ways, all of which are equally likely, then the probability of the event is $\frac{n_i}{N}$.

$$p_{\mathrm{i}} = \underset{N\to\infty}{\mathrm{Lt}} \frac{n_{\mathrm{i}}}{N} \tag{2.8}$$

In the present scenario, we can interpret the pixel intensity as a random variable, which is neither random nor variable; conversely, it can be defined as a function of the elements of a sample space S [3]. Then, the PDF can be defined as

$$p_{\mathrm{i}} = \frac{f[i]}{M \times N} \tag{2.9}$$

The plot of p_{i} would essentially be the plot of normalized $f[i]$; therefore, we can use all statistical models to analyze the histogram (Fig. 2.6).

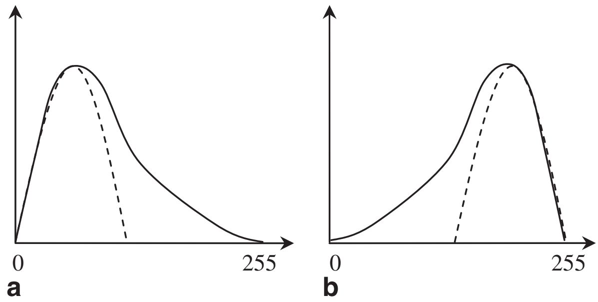

Fig. 2.7 Physical interpretation of skewness in image processing: **a** positively skewed histogram, **b** negatively skewed histogram

2.2.1 Skewness

Skewness γ_1 is the third-order statistical central moment as defined in the Eq. 2.10 and can be physically interpreted as the *measure of asymmetry:*

$$\gamma_1 = E\left\{\left(\frac{X-\mu}{\sigma}\right)^3\right\} \tag{2.10}$$

where E is the expectation, μ and σ are the mean and standard deviation of the random variable X.

In Fig. 2.7, two statistical distributions are shown, one with positive and another with negative skewness value. It can be perceived from the two histograms that, if the histogram is positively skewed, the distribution of lower gray-level intensities (toward black) is denser with respect to higher-level intensities (toward white). It signifies that the distribution is the histogram of a dark image. For the same reason, the negatively skewed distribution represents the histogram of a brighter image.

2.2.2 Kurtosis

Kurtosis α^4 is the fourth-order statistical central moment as defined in Eq. 2.11 and can be physically interpreted as the *measure of peak:*

$$\alpha^4 = E\left\{\left(\frac{X-\mu}{\sigma}\right)^4\right\} \tag{2.11}$$

where E is the expectation, μ and σ are the mean and standard deviation of the random variable X.

In Fig. 2.8, two statistical distributions are shown, one with higher and another with lower coefficient of kurtosis. It can be perceived from the two histograms that, if the kurtosis is very high, the distribution is dense toward mean. On the other hand,

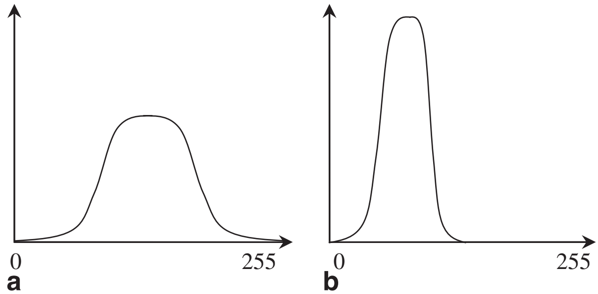

Fig. 2.8 Physical interpretation of kurtosis in image processing: **a** lower value of coefficient of kurtosis, **b** higher value of coefficient of kurtosis

lower kurtosis signifies merely distributed frequency over all intensities. We see in the section of histogram equalization how the coefficient of kurtosis relates to the information content in an image. If the frequency is uniformly distributed along all intensities from 0 through 255, the kurtosis would be significantly low, and the information content would be significantly high as details of the image in all intensities would be perceived perfectly.

2.3 Histogram Equalization and Histogram Specification

Entropy [8] is the average information per message for communication systems. In image processing, we can again correlate the concept of signal transmission with image representation in terms of different image intensities. Entropy of an image having L intensities with PDFs p_i, is defined as [8]

$$H = -\sum_{i=0}^{L-1} p_i \log p_i \tag{2.12}$$

For homogeneous image, which is having only a single intensity in the whole image, i.e., $P_1 = 1$,

$$H = P_1 \log \frac{1}{P_1} = 1 \log 1 = 0.$$

Thus, it is inferred that, in the case of homogeneous image, the information content is zero. For a binary image, the possible intensities are I_1 and I_2 with respective probabilities P_1 and P_2 with relation $P_1 + P_2 = 1$; let, $P_1 = P$; therefore, $P_2 = (1-P)$. Therefore, the entropy is

$$\begin{aligned} H &= P_1 \log \frac{1}{P_1} + P_2 \log \frac{1}{P_2} \\ &= P \log \frac{1}{P} + (1-P) \log \frac{1}{(1-P)} \end{aligned} \tag{2.13}$$

http://www.springer.com/978-3-319-14171-8

Guide to Signals and Patterns in Image Processing
Foundations, Methods and Applications
Das, A.
2015, XXIV, 416 p. 372 illus., 102 illus. in color.,
Hardcover
ISBN: 978-3-319-14171-8