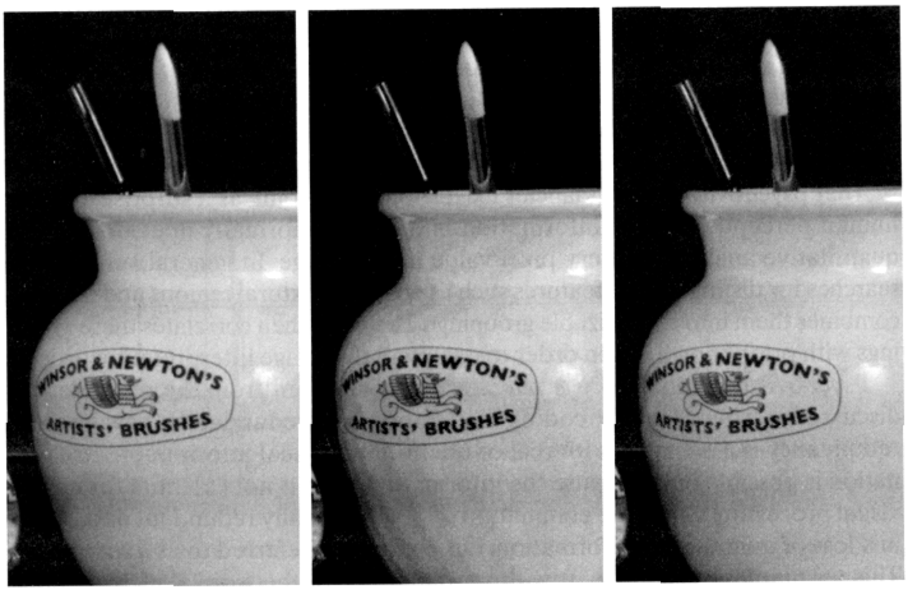
means the mapping of a broad range of input values to a limited number of output values, as discussed in Section 2.4. As it is an irreversible operation (visual information is lost), quantization results in lossy data compression.

EXAMPLE 8.3: Compression by quantization.

■ Consider the images in Fig. 8.4. Figure 8.4(a) shows a monochrome image with 256 possible gray levels. Figure 8.4(b) shows the same image after uniform quantization to four bits or 16 possible levels. The resulting compression ratio is 2:1. Note, as discussed in Section 2.4, that false contouring is present in the previously smooth regions of the original image. This is the natural visual effect of more coarsely representing the gray levels of the image.

Figure 8.4(c) illustrates the significant improvements possible with quantization that takes advantage of the peculiarities of the human visual system. Although the compression ratio resulting from this second quantization procedure also is 2:1, false contouring is greatly reduced at the expense of some additional but less objectionable graininess. The method used to produce this result is known as *improved gray-scale* (IGS) *quantization*. It recognizes the eye's inherent sensitivity to edges and breaks them up by adding to each pixel a pseudo-random number, which is generated from the low-order bits of neighboring pixels, before quantizing the result. Because the low-order bits are fairly random (see the bit planes in Section 3.2.4), this amounts to adding a level of randomness, which depends on the local characteristics of the image, to the artificial edges normally associated with false contouring.

Table 8.2 illustrates this method. A sum—initially set to zero—is first formed from the current 8-bit gray-level value and the four least significant bits of a previously generated sum. If the four most significant bits of the current value are 1111_2, however, 0000_2 is added instead. The four most significant bits of the resulting sum are used as the coded pixel value. ■



a b c

FIGURE 8.4 (a) Original image. (b) Uniform quantization to 16 levels. (c) IGS quantization to 16 levels.

Pixel	Gray Level	Sum	IGS Code
$i-1$	N/A	0000 0000	N/A
i	0110 1100	0110 1100	0110
$i+1$	1000 1011	1001 0111	1001
$i+2$	1000 0111	1000 1110	1000
$i+3$	1111 0100	1111 0100	1111

TABLE 8.2
IGS quantization procedure.

Improved gray-scale quantization is typical of a large group of quantization procedures that operate directly on the gray levels of the image to be compressed. They usually entail a decrease in the image's spatial and/or gray-scale resolution. The resulting false contouring or other related effects necessitates the use of heuristic techniques to compensate for the visual impact of quantization. The normal 2:1 line interlacing approach used in commercial broadcast television, for example, is a form of quantization in which interleaving portions of adjacent frames allows reduced video scanning rates with little decrease in perceived image quality.

8.1.4 Fidelity Criteria

As noted previously, removal of psychovisually redundant data results in a loss of real or quantitative visual information. Because information of interest may be lost, a repeatable or reproducible means of quantifying the nature and extent of information loss is highly desirable. Two general classes of criteria are used as the basis for such an assessment: (1) objective fidelity criteria and (2) subjective fidelity criteria.

When the level of information loss can be expressed as a function of the original or input image and the compressed and subsequently decompressed output image, it is said to be based on an *objective fidelity criterion*. A good example is the root-mean-square (rms) error between an input and output image. Let $f(x, y)$ represent an input image and let $\hat{f}(x, y)$ denote an estimate or approximation of $f(x, y)$ that results from compressing and subsequently decompressing the input. For any value of x and y, the error $e(x, y)$ between $f(x, y)$ and $\hat{f}(x, y)$ can be defined as

$$e(x, y) = \hat{f}(x, y) - f(x, y) \tag{8.1-7}$$

so that the total error between the two images is

$$\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[\hat{f}(x, y) - f(x, y)\right]$$

where the images are of size $M \times N$. The *root-mean-square error*, e_{rms}, between $f(x, y)$ and $\hat{f}(x, y)$ then is the square root of the squared error averaged over the $M \times N$ array, or

$$e_{\text{rms}} = \left[\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[\hat{f}(x, y) - f(x, y)\right]^2\right]^{1/2} \tag{8.1-8}$$