

## Chapter 10

# Halftones, Photos and Grayscale Printing

### 10.1 Continuous-Tone and Black-and-White

A computer screen and a photograph are both examples of *continuous-tone* media. The phosphors that are sprinkled on the inside of the glass of cathode-ray tube glow with an intensity proportional to the strength of the electron beam striking them. By varying the beam strength, the monitor can vary the glow of the smallest addressable pixel from black to white continuously. Similarly, a photograph emulsion can vary continuously on the smallest scales visible to the eye.<sup>1</sup>

In contrast, a laser printer, an offset lithographic printer, an ink-jet printer and a xerox copier are all black-and-white devices. These printers cannot vary the composition of each tiny drop of ink, but instead can only throw drops of solid black ink or toner. Nevertheless, these *monotone* printers are capable of printing gray? How?

The answer is that monotone printers organize the smallest addressable dots into larger blocks. There is no standard name for these blocks; perhaps we can call them “screen-blocks” by analogy with half-tone screens, which will be explained below. The smallest addressable area that can be targeted by a single drop of ink or toner is usually called a “pixel” by analogy with the smallest unit of a monitor screen.

If the screen-block is an 8 x 8 array of pixels, then the monotone device can represent 65 different shades of gray by putting between zero and 64 drops of ink in a given screen-block. The good news is that this more shades of gray than the eye can usually distinguish. The bad news is that the effective resolution of the grayscale image is much poorer than the dots-per-inch of the printer. A 65-shade grayscale image on a 600 dpi printer has a resolution of only 75 screen-blocks per inch, and this is poor enough to be noticeable.

One can obtain higher resolution by using few shades of gray. A 4 x 4 array of pixels allows only 17 shades of gray, but the resolution with a 600 dpi printer is doubled to 150 screen-blocks per inch. With a monotone device, there is thus a tradeoff between the number of shades of gray and resolution (Fig. 10.1).

---

<sup>1</sup>On an invisible scale, of course, the electron beam is made up of discrete quanta, the phosphor is composed of discrete molecules, and the photographic emulsion contains microscopic grains; on the nanoscale, *everything* is discrete. However, the discreteness of a CRT monitor and a photograph is on such a fine scale that it is not visible to the eye.

## 10.2 Halftones and Photographs

Many decades ago, a clever technique for transferring continuous-tone photographs to monotone printers was developed: halftone-screening. The name is a misnomer: the result of halftoning is a photolithography plate or an etched printing plate which is monotone. However, just as in modern printers, a halftoned image simulates grayscale by very fine dots which appear various shades of gray when seen from normal viewing distance.

The key idea is diffraction. The earliest screens were lattices of thin, criss-crossing wires, little different from ordinary window screens except that the mesh was finer. When light bounces off a photograph and through the mesh, it is diffracted. The diffraction translates intensity into volume. The diffracted image consists of a pattern of solid-black dots whose area is proportional to the shade of gray falling on that little cell of the screen. When very pale gray falls on a particular screen-block, a tiny black dot appears on the far side of the screen. When little light falls on a cell, a large black shadow appears within that cell. By employing a chemical process that etched away a metal plate where light fell upon it while leaving the black dot untouched, the fine network of variously sized dots in the diffracted image would be left as raised areas on the printing plate.

The philosophy is similar to a modern laser or ink-jet printer: a small area of black-and-white is perceived as a shade of gray whose lightness depends on the rate of black-to-white within the small screen-block. However, the means of achieving this ratio of black-to-white is different. There is only a single black dot within each block of a halftoned image, but the blocks are of variable area; a laser printer lays down a varying number of dots within the screen block, but the dots are all identical in size (Fig. 10.2).

For obvious reasons, the resolution of a halftone image are expressed in lines-per-inch or lpi where “line” is simply the spacing between neighboring parallel wires in the screen-mesh. Our 600 dpi laser printer, laying down 65 shades of gray through 8 x 8 arrays of pixels, each array 1/75-th of an inch across, has an effective lpi of 75.

In modern offset lithography, the etched metal plate has been replaced by a film, and the wire-mesh screen has been replaced by a film or sheet of plastic, too. Modern screens have a checkerboard of very fine, slightly diffuse loads imprinted upon them. To get the proper diffraction, the screen film needs to be as close as possible to the photolithography film. The usual technofix is use a vacuum seal so that the screen is literally sucked onto the photolithography film.

There are three types of screens in widespread use (plus additional specialty screens): “square”, “elliptical” and “stochastic”. The usual screens create a pattern of squarish dots so that the “midtone” pattern — half-black, half-white — looks like a checkerboard with the black squares just touching at the corners. A modern variation staggers and

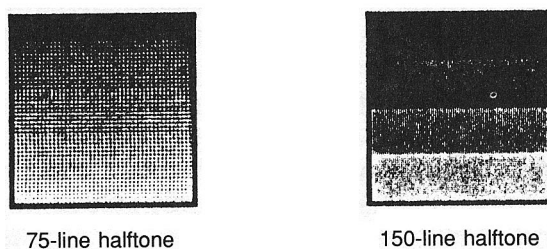


Figure 10.1: The same pattern printed on a 300 dpi printer using a 75 lpi screen (left) and a 150 lpi screen (right). The lower screen frequency actually gives a much better-looking picture.

angles the squares on the screen to produce a so-called “elliptical” screen. This helps to minimize “dot gain”, which will be discussed more below.

When high resolution images needed to be converted to grayscale, software can define the dot pattern. However, the range of patterns can be wider because the computer is not restricted by the need of a mechanical wire-mesh to hold together. The “stochastic” screen lays down a pattern of random dots whose density controls the shade of gray. In contrast to conventional screens, all the dots are the *same size*. “Stochastic” screens are also called “FM” or “frequency-modulation” screens because the frequency of dot deposition varies with the desired shade of gray. Stochastic screens are used for very high quality photographic reproduction. (A 600 dpi scanner can save effectively scan at 600 lpi resolution, far higher than can possibly be printed; the task of the software is to create a *printable* pattern of dots which preserves as much of the scan information as possible.) Because there are no regularly spaced lines in a stochastic screen, the “lines-per-inch” measurement standard does not, strictly speaking, apply. Nevertheless, the effective resolution per block of gray is still much lower than the black-dot-per-inch resolution of the printer. Some limits are inescapable!

Laser printers and ink-jet printers produce identical dots, as noted previously, whereas the dots produced by a physical screen on a photolithography plate are variable in size. Usually, the nonuniform dots of the film screen do a better job of fooling the eye than the uniform dots of the computer printers at the same lpi. This has further motivated the use of stochastic or FM screening when the final output device is a computer printer.

## 10.3 LPI and Paper

Newspapers usually employ halftone screens with resolutions of 65 to 85 lpi. Finer screens are useless because newsprint is coarse, absorbent and printed with watery inks. Even if a printer spits a dot of ink only 1/600-th of an inch in diameter at a piece of newsprint, the final printed dot will be much wider because the ink will spread out a fraction before drying on the coarse, textured paper. Printed books and glossy magazines use better paper and screens with 133 to 150 lpi. Most web presses cannot support an lpi higher than 150 because they rotate so rapidly, and the ink must therefore dry so quickly on mechanically-stretched paper, that the minimum drop size is inadequate for higher resolutions.

## 10.4 Dot Gain

**Definition 23 (Dot Gain)** *This is a printing defect in which the dots dry at a larger size than specified by the printing plate.*

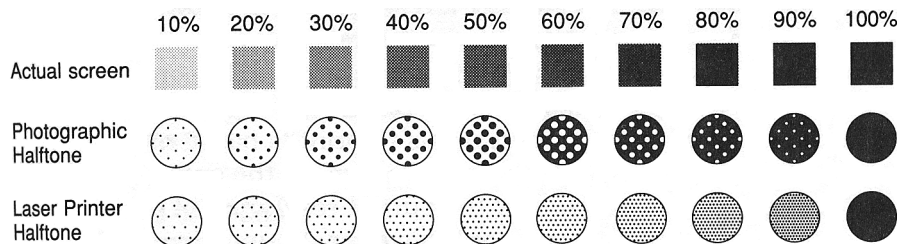


Figure 10.2: Closeup of output pixels forming a halftone screen-block from a photographic-and-physical-screen halftone (middle) and a laser printer (bottom).

Low resolution printers, aqueous inks and porous or textured paper contribute to dot gain. In a web press, the paper is stretched around a drum while being printed, and this mechanical stretching imposes its own limits on the minimum dot size.

Dot gain is a very unpleasant problem for printing halftones because when the dots are larger than intended, the whole printed image is darkened. Dot gain also tends to create noticeable jumps in shade or color.

Elliptical halftone screens and stochastic screens help some. However, the best defence against dot gain is restraint; being greedy for high resolution leads to ugly printed images when the paper or the printer aren't up to the task of printing a halftone that has been created with a very high lpi screen.

## 10.5 Xeroxing Halftones

Xeroxes of halftones usually look rather bad. The reason is that it is very uncommon for the little dots of toner to match the size of the dots of the halftone. This leads to "dot gain": the xerox has big dots where the halftone had small ones. A midtone that consisted of 32 tiny dots of black interspersed with 32 dots of white in a given screen-block will appear in the xerox as 32 larger blocks that fill almost the entire screen-block. The result is that the xerox has turned gray into block, and the image is almost unrecognizable.

## 10.6 Rescreening

It is usually a very bad idea to try to make a halftone from a halftone. The reason is that the dots of the original halftone are so small that they are probably close to the resolution limits of the scanner, and likely beyond the resolution of a xerox machine. Consequently, some dot gain will occur in the scanned image. When the scan is printed, more dot imperfection will happen because the printer is being asked to reproduce dots near the limits of its resolution.

Suppose, for example, that the scanner records dots which are each 1/300-th of an inch in diameter, and these must be reproduced on a 600 dpi printer. If the scan dots are perfectly aligned with the printer dots, then all is well: the printer can print each scan dot as a 2 x 2 array of ink-jet drops. Only all is not well: the drops of the jet are round, but the scan dots may have been bounded by diamonds or slanting ellipses. In that case, the printed array of four dots will not exactly match the shape of the scan dot, and some distortion of the gray shade will happen.

If the scan dots are not exactly aligned with rows and columns of the array of ink dots, then the printer cannot even come close to matching the scan dot. It is not possible for the printer to lay down a pattern of half-disk plus a full circle plus a half-disk to match a scan-dot that is two ink dots wide, but displaced by half a dot.

The interaction between the spacing of the scan and the spacing of the ink dots can lead to interference patterns which create so-called moiré patterns. These oscillations in shade are due to frequency-aliasing. They were first seen in the shimmering patterns of watered-silk. Because the moiré patterns usually have a spatial frequency that is small compared to that of the printer and scan frequencies — the frequency of the moiré pattern is the *difference* of the printer and scan frequencies — the watered-silk variations in shade are very obvious.

These moiré patterns can be eliminated by "rescreening". A special calculating tool is used to estimate the frequency and angle of the screen used to create the halftone.

If the new halftone matches this, then spatial frequency aliasing is suppressed. Unfortunately, the problem of dot gain and dot shape distortion cannot be suppressed. So a halftone of a halftone will still look crummy.

## 10.7 Photoshop and Halftones

A grayscale image is always printed as a pattern of black dots. Photoshop and other image-processing software allows one to control this conversion. Image  $\Rightarrow$  Grayscale  $\Rightarrow$  Bitmap initiates this process. It pops a dialogue box that offers four options for the conversion, the last of which is "Halftone". If this is selected by clicking on the button, then the Halftone Dialog Box appears.

This offers options of line frequency (lpi) and shape of dot. One must be careful; if the image is specified to have an output of only 72 dpi (monitor resolution), then specifying a high frequency will still leave a very coarse image. To successfully use a high lpi, one needs to have first scanned the image at a much higher dpi, and then specified a very high dpi for the output before the halftone conversion is attempted.