

3

Image Quality for Scanning and Digital Imaging Systems

Donald R. Lehmbeck

Xerox Corporation, Webster, New York, USA (retired)
College of Imaging Arts and Sciences, Rochester Institute of Technology,
Rochester, New York, USA (Adjunct Faculty)
Imaging Quality Technology Consulting, Penfield, New York, USA
Torrey Pines Research, Fairport, New York, USA

John C. Urbach

Portola Valley, California, USA (deceased)

CONTENTS

3.1	Introduction.....	135
3.1.1	Imaging Science for Scanned Imaging Systems.....	135
3.1.1.1	Scope.....	135
3.1.1.2	The Literature.....	136
3.1.1.3	Types of Scanners.....	136
3.1.2	The Context for Scanned Image Quality Evaluation.....	137
3.2	Basic Concepts and Effects.....	140
3.2.1	Fundamental Principles of Digital Imaging.....	140
3.2.1.1	Structure of Digital Images.....	140
3.2.1.2	The Sampling Theorem and Spatial Relationships.....	145
3.2.1.3	Gray Level Quantization: Some Limiting Effects.....	148
3.2.2	Basic System Effects.....	152
3.2.2.1	Blur.....	152
3.2.2.2	System Response.....	153
3.2.2.3	Halftone System Response.....	155
3.2.2.4	Noise.....	159
3.2.2.5	Color Imaging.....	160
3.3	Practical Considerations.....	166
3.3.1	Scan Frequency Effects.....	166
3.3.2	Placement Errors or Motion Defects.....	169
3.3.3	Other Nonuniformities.....	173
3.3.3.1	Perception of Periodic Nonuniformities in Color Separation Images.....	173
3.4	Characterization of Input Scanners that Generate Multilevel Gray Signals (Including Digital Cameras).....	174
3.4.1	Tone Reproduction and Large Area Systems Response.....	175
3.4.2	MTF and Related Blur Metrics.....	181
3.4.2.1	MTF Approaches.....	183

3.4.2.2	The Human Visual System's Spatial Frequency Response	189
3.4.2.3	Electronic Enhancement of MTFs: Sharpness Improvement.....	189
3.4.3	Noise Metrics.....	190
3.5	Evaluating binary, thresholded, scanned imaging systems.....	193
3.5.1	Importance of Evaluating Binary Scanning.....	193
3.5.1.1	Angled Lines and Line Arrays.....	193
3.5.2	General Principles of Threshold Imaging Tone Reproduction and Use of Gray Wedges.....	194
3.5.2.1	Underlying Characteristic Curve and Noise.....	194
3.5.3	Binary Imaging Metrics Relating to MTF and Blur	195
3.5.3.1	Resolving Power (A Measure for Discrimination of Fine Detail)	195
3.5.3.2	Line Imaging Interactions.....	198
3.5.4	Binary Metrics Relating to Noise Characteristics	198
3.5.4.1	Gray Wedge Noise	198
3.5.4.2	Line Edge Noise Range Metric.....	199
3.5.4.3	Noise in Halftoned or Screened Digital Images.....	200
3.6	Summary Measures of Imaging Performance.....	202
3.6.1	Basic Signal-to-Noise Ratio.....	202
3.6.2	Detective Quantum Efficiency and Noise Equivalent Quanta	204
3.6.3	Application-Specific Context.....	204
3.6.4	Modulation Requirement Measures	204
3.6.5	Area under the MTF Curve (MTFA) and Square Root Integral (SQRI)	205
3.6.6	Measures of Subjective Quality	206
3.6.7	Information Content and Information Capacity	209
3.7	Specialized Image Processing	214
3.7.1	Lossy Compression.....	214
3.7.2	Nonlinear Enhancement and Restoration of Digital Images	216
3.7.3	Color Management	218
3.8	Psychometric Measurement Methods Used to Evaluate Image Quality	219
3.8.1	Relationships between Psychophysics, Customer Research, and Psychometric Scaling.....	219
3.8.2	Psychometric Methods.....	220
3.8.3	Scaling Techniques	221
3.8.3.1	Identification (Nominal).....	222
3.8.3.2	Rank Order (Ordinal).....	222
3.8.3.3	Category (Nominal, Ordinal, Interval)	222
3.8.3.4	Graphical Rating (Interval).....	222
3.8.3.5	Paired Comparison (Ordinal, Interval, Ratio).....	222
3.8.3.6	Partition Scaling (Interval).....	223
3.8.3.7	Magnitude Estimation (Interval, Ratio)	223
3.8.3.8	Ratio Estimation (Ratio)	223
3.8.3.9	Semantic Differential (Ordinal, Interval)	223
3.8.3.10	Likert Method (Ordinal)	223
3.8.3.11	Hybrids (Ordinal, Interval, Ratio)	224
3.8.4	Practical Experimental Matters Including Statistics.....	224
3.9	Reference Data and Charts.....	226
	Acknowledgments	237
	References	238

3.1 INTRODUCTION

3.1.1 Imaging Science for Scanned Imaging Systems

This chapter presents some of the basic concepts of image quality and their application to scanned imaging systems. In this revised edition, we have added more on tonal rendition including system plots and halftones, new approximations to MTF's and revised pointers to the current industry standards in image quality, as well as more reference data and charts while reducing content on binary imaging and overall quality. New references and other technical details have been added throughout.

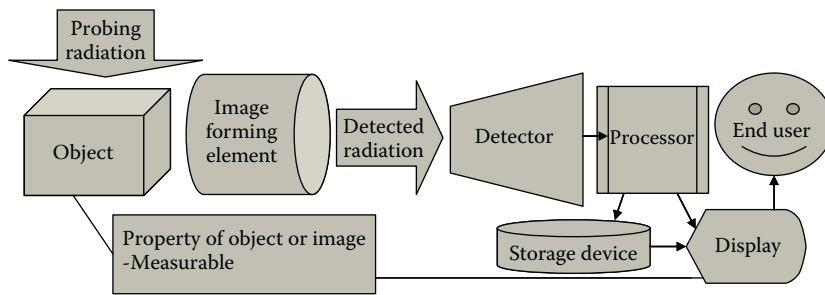
The emphasis in this chapter will be on the input scanner. Output scanners will be dealt with mainly by inference since many input scanner considerations and metrics are directly applicable to the rest of a complete electronic scanned imaging system. Expanded discussions of halftone methods, tone reproduction, and nonuniformity are examples of output scanners and systems implications. The chapter is organized as 10 major sections moving from the basic concepts and phenomena of image scanning and color, through practical aspects of image quality, to performance of input scanners that produce multilevel (gray) signals and then the special but common case of binary scanned images. This is followed by sections on very specific topics: various summary measures of imaging performance and specialized image processing. To assist the reader, psychophysical measurement methods used to evaluate image quality and some reference data and charts have been added.

3.1.1.1 Scope

We, like so many others, follow in the path pioneered over a half century ago by the classic 1934 paper of Mertz and Gray.¹ Without going into the full mathematical detail of that paper and many of its successors, we attempt to bring to bear some of the modern approaches that have been developed both in image quality assessment and in scanned image characterization. Many diverse technologies used in scanned imaging systems are addressed throughout the rest of this book. We cannot address the explicit effects of any of these on quality because they provide an enormous array of choices and trade-offs. Building on a more general foundation of imaging science, we shall attempt to provide a framework in which to sort out the many image quality engineering and technology issues that depend on these choices.

It is our intent not to show that one scanner or technique is better than another, but to describe the methods by which a scanning system can be evaluated to compare to other systems and to assess the technologies used in them. This chapter therefore deals primarily with such matters as the sharpness or graininess of an image and not with such hardware issues as the surface finish of an aluminum mirror, uniformity of a drive motor, or the efficiency of charge transfer in a charge-coupled device (CCD) imager.

Scanning is considered here in the general context of electronic imaging. An electronic imaging system can be considered as composed of 10 basic parts² illustrated in Figure 3.1 as a flow chart of fairly self explanatory terms. Both digital photography and scanning use the same type of CCD or CMOS sensors, that is, detectors. Both create images in two-dimensional pixel format. For both, the processor is on the sensor, in hardware resident on the system and also off-line in computers. Both systems generate two-dimensional prints or displays of images using one-dimensional output applying them to one-dimensionally electronic/computer stored bit streams. Both systems use optical systems and input radiation to create the captured image including arrays of color filters to create colored images.

**FIGURE 3.1**

The fundamental elements of any imaging system arranged in a flow diagram that approximates a typical scanner or digital camera.

Some input scanners use reduction optics much as a camera in macromode but some use selfoc lens arrays which nearly contact the reflection original.

The primary difference between digital photography and input scanning is that the sensor in most photography is a fixed two-dimensional array of photosites (i.e., one-pixel sensors), while in scanning the array is synthesized by moving a long line of photosites one-pixel wide (i.e., a one-dimensional array or possibly three lines one for each color) over as much of a document as is needed. This has an effect in the scanning electronics—speed of the real-time circuits and optomechanical structures—that might create errors in positioning the line of sensors. This creates a difference from two-dimensional arrays making it appear as if the synthesized array was nonuniform . . . Similar nonuniformities from one-dimensional moving arrays also occur in the printers which serve both scanners and other digital imaging devices. The nonuniformities are described in Sections 3.3.2 and 3.3.3 (pp. 167–174). Much of the chapter, therefore, applies equally well to digital photography and scanning.

3.1.1.2 The Literature

Considerable research, development, and engineering have occurred over the last two decades and only a very small portion could be referenced in the following pages. A few general references of note are provided as References 2–18 and elementary tutorials in References 19–23. Other more specific work of importance that may interest the reader includes: the vast technology of image processing,¹⁷ many papers focused on specific problems in scanner image quality (see titles),^{24–26} digital halftoning,^{27,28} color imaging,^{29–32} and various forms of image quality assessment.^{33–39}

While the focus here is on imaging modules and imaging systems, scanners may, of course, be used for purposes other than imaging, such as digital data recording, from Bar Codes for example. We believe that the imaging science principles used here are sufficiently general to enable the reader with a different application of a scanning system to infer appropriate knowledge and techniques for these other applications.

3.1.1.3 Types of Scanners

All input scanners convert one- or (usually) two-dimensional image irradiance patterns into time-varying electrical signals. Image integrating and sampling systems, such as those found in many forms of electronic cameras and electronic copying devices, have sensors such as a CCD array. The signals produced by these scanners can be in one of two

general forms, either (a) binary output (a string of on and off pulses), or (b) gray-scale output (a series of electrical signals whose magnitude varies continuously).

The term *digital* here refers to a system in which each picture element (pixel) must occupy a discrete spatial location; an analog system is one in which a signal level varies continuously with time, without distinguishable boundaries between individual picture elements. A two-dimensional analog system is usually only analog in the more rapid direction of scanning and is discrete or “digital” in the slower direction, which is made up of individual raster lines. Television typically works in this fashion. In one form of solid-state scanner, the array of sensors is actually two-dimensional with no moving parts. Each individual detector is read out in a time sequence, progressing one raster line at a time within the two-dimensional matrix of sensors.

In other systems a solid-state device, arranged as a single row of photosites or sensors, is used to detect information one raster line at a time. In these systems either the original image is moved past the stationary sensor array, or the sensor array is scanned across the image to obtain information in the slow scan direction.

Cameras in digital photography employ totally digital solid-state two-dimensional sampling arrays. In some sense they represent commonly encountered forms of input scanners. The reader should be able to infer many things about the other forms of scanners and digital cameras from examples discussed in this chapter.

3.1.2 The Context for Scanned Image Quality Evaluation

Building blocks for developing a basic understanding of image quality in scanning systems are shown in Figure 3.2. The major elements of a generalized scanning system are on the left, with the evaluation and analysis components on the right. This chapter will deal with all of these elements and it is therefore necessary to see how they all interact.

The general configuration of scanning systems often requires two separate scanning elements. One is an input scanner to capture, as an electronic digital image, an input analog optical signal from an original scene (object), shown here as a hard copy input, such as a photograph. The second scanning element is an output scanner that converts a digital signal, either from the input scanner or from computer-generated or stored image data, into analog optical signals. These signals are rendered suitable for writing or recording on some radiation-sensitive medium to create a visible image, shown here as hard copy output. The properties of this visible image are the immediate focus of image quality analysis. It may be photographic, electrophotographic, or something created by a variety of unconventional imaging processes. The output scanner and recording process may also be replaced by a direct marking device, such as a thermal, electrographic, or ink jet printer, which contains no optical scanning technology and therefore lies outside the scope of this volume. Nonetheless, its final image is also subject to the same quality considerations that we treat here.

It is to be noted that the quality of the output image is affected by several intermediate steps of image processing. Some of these are associated with correcting for the input scanner or the input original, while others are associated with the output scanner and output writing process. These are mentioned briefly throughout, with the digital halftoning process, described in Section 3.2.2.3, cited as a major example of a correction for the output writing. Losses or improvements associated with some forms of data communication, and compression are very important in a practical sense, especially for color. These are briefly reviewed in Section 3.7.1. Additional processing to meet user preferences or to enable some particular application of the image must also be considered a part of the image quality evaluation. A few examples are given throughout. A comprehensive treatment of image

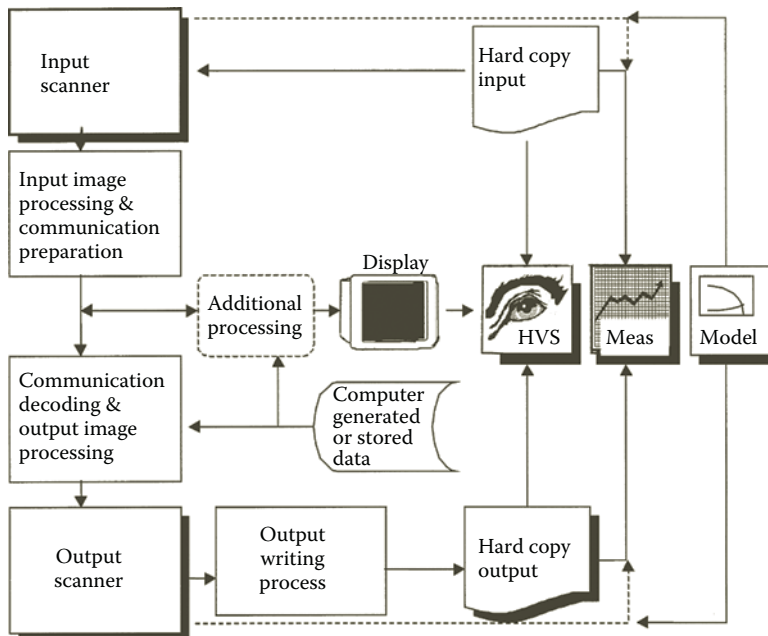


FIGURE 3.2

The elements of scanned imaging systems as they interact with the major methods of evaluating image quality. “HVS” refers to the human visual system. “Meas” refers to methods to measure both hard copy and electronic images and “Models” refers to predicting the imaging systems performance, not evaluating the images *per se*.

processing is beyond the scope of this chapter but several references are given at the end of this chapter to help the reader learn more about this critical area of scanned imaging.

The assessment of quality in the output image may take the form of evaluation by the human visual system (HVS) and the use of psychometric scaling (see Section 3.8) or by measurement with instruments as described in parts of Sections 3.3–3.5. One can also evaluate measured characteristics of the scanners and integrated systems or model them to try to predict, on average, the quality of images produced by these system elements. (Both of these hardware characterizations are also described in parts of Sections 3.3–3.5.) The description of overall image quality (Section 3.6) tends to focus on the models of systems and their elements, not the images themselves. For some purposes, for example, judging the quality of a copier, the comparison between the input and output images is the most important way of looking at image quality, whether it be by visual or measurement means. For other applications it is only the output image that counts. In some cases, the most common visual comparison is between the partially processed image, as can only be seen on the display, and either the input original or the hard copy output. In most cases, the evaluation criteria depend on the intended use of the image. A display of the scanned image in a binary (black or white) imaging mode reveals some interesting effects that carry through the system and often surprise the unsuspecting observer. These are covered in Section 3.5. Physical and visual measurements evaluate output and input images, hence the arrows in Figure 3.1 flow from hard copy toward these evaluation blocks. Models, however, are used mostly to synthesize imaging systems and components and may be used to predict or simulate performance and output. Hence the “model” arrows flow toward the system components.

The nonscanner components for electronic image processing and the analog writing process play a major role in determining quality and hence will be unavoidably included in

any realistic HVS or measurement evaluation of the quality of a scanned image or imaging system. Models of systems and components, on the other hand, often ignore the effects of these components and the reader is cautioned to be aware of this distinction when designing, analyzing, or selecting systems from the literature.

A model has been described by P. Engeldrum^{12,40-42} called the Image Quality Circle, which ties all of these evaluations together and expands them into a logical framework to evaluate any imaging system. This is shown in Figure 3.3 as the circular path connecting the oval and box shapes, along with the three major assessment categories from Figure 3.2, namely the HVS, Measurements, and Models. In his model, the HVS category above is expanded to show a type of model he calls “visual algorithms,” which predict human perceived attributes of images from physical image parameters. Examples of perceptions would include such visual subjective sensations as darkness, sharpness, or graininess (i.e., “nesses”). These are connected to physical measurements of densities, edge profiles, or halftone noise, respectively, made on the images used to evoke these subjective responses. In Engeldrum’s analysis, the rest of what we call the HVS and brain combination includes “image quality models,” which predict customer preferences based on relationships among the perceived attributes. This purely subjective dimension of individuals is often not included in the “brain” functions normally associated with HVS, therefore it is

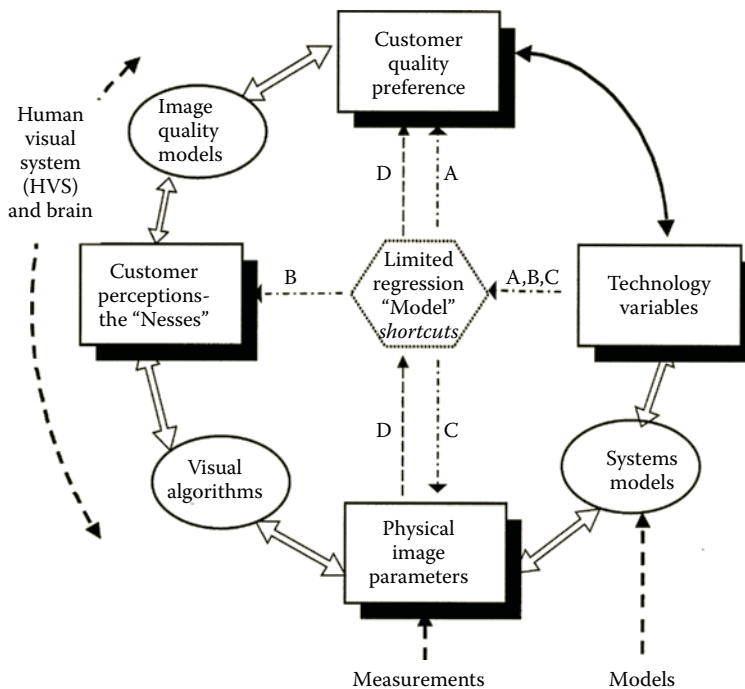


FIGURE 3.3

An overall framework for image quality assessment, composed of the elements connected by the outline arrows, known as the “Image Quality Circle” (adapted from Engeldrum, P.G. *Psychometric Scaling: A Toolkit for Imaging Systems Development*; Imcotek Press: Winchester, MA, 2000 and Engeldrum, P.G. *Chapter 2 Psychometric Scaling: A Toolkit for Imaging Systems Development*; IMCOTEK Press: Winchester, MA, 2000; 5–17.) and the inner “spokes” which illustrate four commonly used, but limited, regression model shortcuts as paths A, B, C, and D. The latter were not proposed by Engeldrum as part of the Image Quality Circle model, but added here to illustrate how selected examples given in Section 3.6 fit the framework. The connection to HVS, measurement, and model elements of Figure 3.2 are indicated by the labels and heavy dashed lines that surround the figure.

mentioned explicitly here. The methodologies to enable these types of analysis generally fall into the realm of psychometrics (quantifying human psychological or subjective reactions). They will be reviewed in Section 3.8.

Many authors (Section 3.6) have attempted to short-circuit this framework, following the dashed “spokes” we have added to the circle in Figure 3.3. These create regression models using psychometrics that directly connect physical parameters (path D) or technology variables with overall image quality models (path A) or preferences (path C). These have been partially successful, but, having left out some of the steps around the circle, they are very limited, often applying only to the circumstances used in their particular experiment. When these circumstances apply, however, such abbreviated methods are valuable. Following all the steps around the circle leads to a more complete understanding and more general models that can be adapted to a variety of situations where preferences and circumstances may be very different. The reader needs to be aware of this and judge the extent of any particular model’s applicability to the problem at hand.

3.2 BASIC CONCEPTS AND EFFECTS

3.2.1 Fundamental Principles of Digital Imaging

The basic electronic imaging system performs a series of image transformations sketched in Figure 3.4. An object such as a photograph or a page with lines and text on it is converted from its analog nature to a digital form by a *raster input scanner* (RIS). It becomes “digital” in distance where microscopic regions of the image are each captured separately as discrete pixels; that is, it is *sampled*! It is then quantized, in other words, digitized in level, and is subsequently processed with various strictly digital techniques. This digital image is transformed into information that can be displayed or transmitted, edited, or merged with other information by the *electronic and software subsystem* (ESS). Subsequently a *raster output scanner* (ROS) converts the digital image into an analog form; that is, it is *reconstructed*, typically through modulating light falling on some type of photosensitive material. The latter, working through analog chemical or physical processes, converts the analog optical image into a reflectance pattern on paper, or into some other display as the final output image.

What follows assumes optical output conversion, but direct-marking processes, involving no optics (e.g., ink jet, thermal transfer, etc.) can be treated similarly. Therefore, while one often thinks of electronic imaging or scanned imaging as a digital process, we are really concerned in this chapter with the imaging equivalent of analog to digital (A/D) and digital to analog (D/A) processes. The digital processes occur between as image processing. In fact that is where we become familiar with the scanned imaging characteristics because that is one place where we can take a look at a representation of the image, that is, in a computer.

3.2.1.1 Structure of Digital Images

Before considering all the system and subsystem effects, let us turn our attention to the microscopic structure of this process, paying particular attention to the A/D and sampling domain of the input scanner. Sampled electronic images were first studied in a comprehensive way by Mertz and Gray.¹

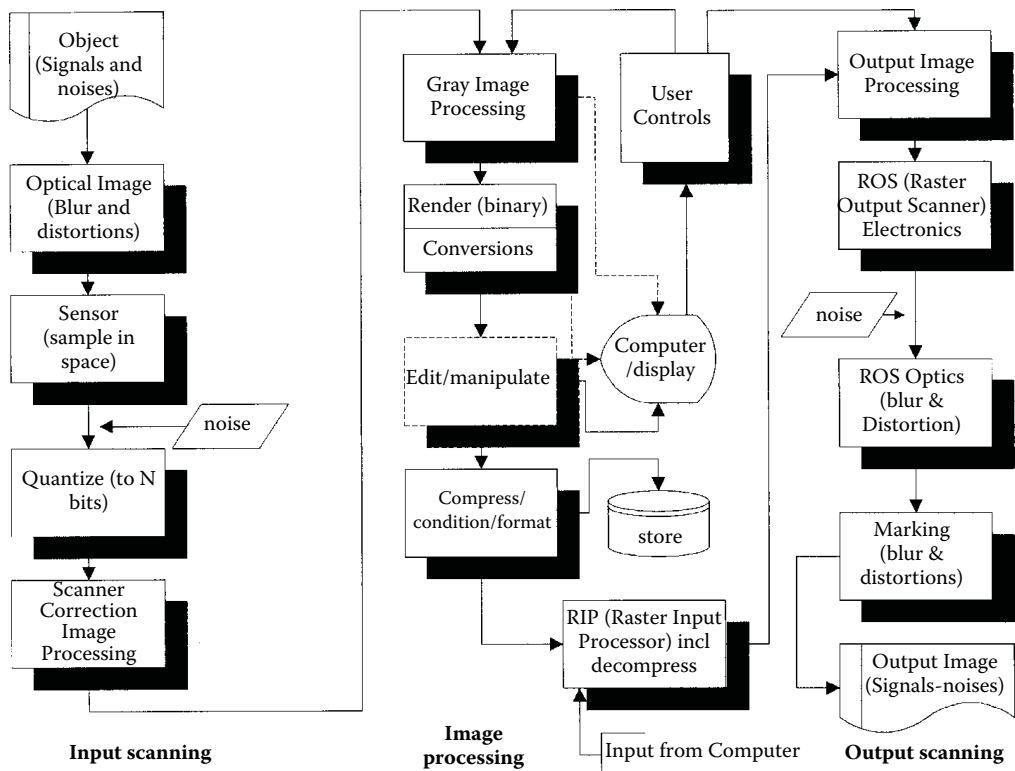


FIGURE 3.4

Steps in typical scanning electronic reprographic system showing basic imaging effects.

To understand how sampling works, let us examine Figure 3.5. It illustrates four different aspects of the input scanning image transformations. Figure 3.5a shows the microscopic reflectance profile representative of an input object: there is a sharp edge on the left, a “fuzzy” edge (ramp), and a narrow line. Figure 3.5b shows the optical image, which is a blurred version of the input object. Note that the relative heights of the two pulses are now different and the edges are sloping that were previously straight. Figure 3.5c represents the blurred image with a series of discrete signals, each being centered at the position of the arrows. This process is referred to as *sampling*.

Each sample in Figure 3.5c has some particular height or gray value associated with it (scale at right). When these individual samples can be read as a direct voltage or current, that is they can have any level whatsoever, then the system is analog. When an element in the sensor output circuit creates a finite number of gray levels such as 10, 128, or even 1000, then the signal is said to be *quantized*. (When a finite number of levels is employed and is very large, the quantized signal resembles the analog case.) Being both sampled and quantized in a form that can be manipulated by a digital processor makes the image *digital*. Each of these individual samples of the image is a *picture element*, often referred to as a *pixel* or *pel*. A sampled and *multilevel* (>2) quantized image is often referred to as a *grayscale image* (a term also used in a different context to describe a continuous tone analog image). When the quantization is limited to *two levels*, it is termed a *binary image*. Image processing algorithms that manipulate these different kinds of images can be “bit constrained” to the

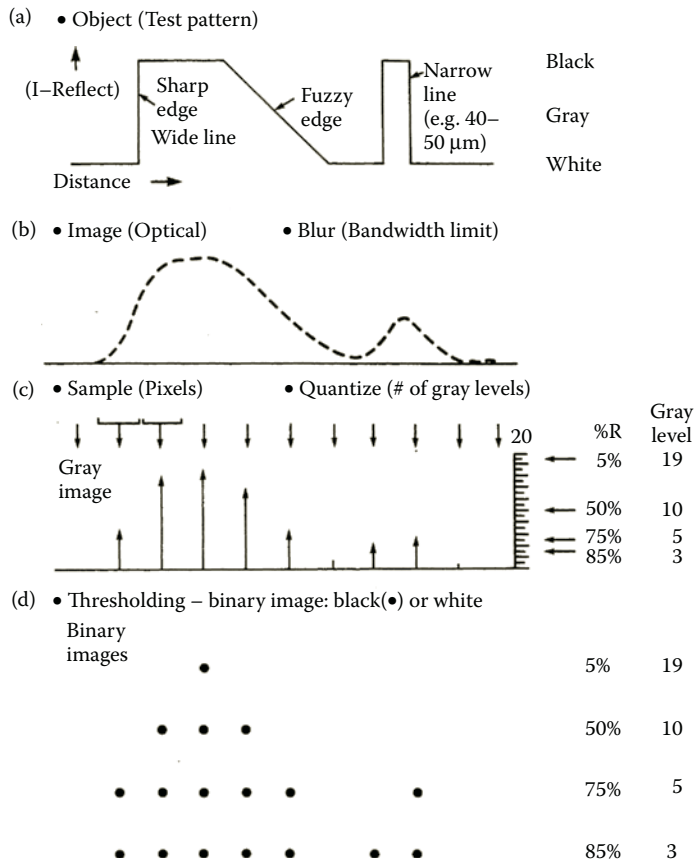


FIGURE 3.5

Formation of binary images, illustrating how a single, blurred electronic image of a small continuous tone test object could yield many different binary images depending on the threshold selected. (a) 1-reflectance profile of the test pattern, (b) blurred analog electronic image of the test pattern in relative response units (e.g. relative millivolts), (c) Sampling into pixels where each arrow at top represents a pixel location. The magnitude of each arrow at bottom represents the response at that pixel location. The % reflectance and the assigned grey level response are given at the right where the larger grey level represents the blacker (lower reflectance) parts, (d) Each row represents a different threshold for pixels shown in (c), each black dot represents a black image pixel, with thresholds for each row identified at the right.

number of levels appropriate to the image bit depth (another expression for the number of levels), that is, integer arithmetic. This is effectively equivalent to many digital image processing circuits. Alternatively, algorithms may be floating point arithmetic, the results of which are quite different from the bit constrained operations.

A common and simple form of image processing is the conversion from a gray to a binary image as represented in Figure 3.5d of Figure 3.5. In this process a threshold is set at some particular gray level, and any pixel at or above that level is converted to white or black. Any pixel whose gray value is below that level is converted to the other signal, that is, black or white, respectively. Four threshold levels are shown in Figure 3.5c by arrows on the gray-level scale at the right. Results are depicted in Figure 3.5d as four rows, each being a raster from the different binary images, one for each of the four thresholds.

In Figure 3.5d, each black pixel is represented by a dot, and each white pixel is represented by the lack of a dot. (It is common to depict pixels as series of contiguous squares in a lattice representing the space of the image. They are better thought of as points in time and space that can have any number of dimensions, attributes, and properties.)

Each row of dot patterns shows one line of a sampled binary image. These patterns are associated with the location of the sampling arrows, shown in Figure 3.5c, the shape of the blur, and the location of the features of the original document. Notice at the 85% threshold, the narrow line is now represented by two pixels (i.e., it has grown), but the wider and darker pulse has not changed in its representation. It is still five-pixel wide. Notice that the narrow pulse grew in an asymmetric fashion and that the wider pulse, which was asymmetric to begin with, grew in a symmetric fashion. These are quite characteristic of the problems encountered in digitizing an analog document into a finite number of pixels and gray levels. It can be seen that creating a thresholded binary image is a highly nonlinear process. The unique imaging characteristics resulting from thresholding are discussed in detail in Section 3.5.

Figure 3.6 represents the same type of process using a real image. The plot is the gray profile of the cross section of a small letter “I” for a single scan line. The width of the letter is denoted at various gray levels, indicated here by the label “threshold” to indicate where one could select the potential black to white transition level. The reader can see that the width of the binary image can vary anywhere from one to seven pixels, depending on the selection of threshold.

Figure 3.7a returns to the same information shown in Figure 3.5, except that here we have doubled the frequency with which we sampled the original blurred optical image. There are now twice as many pixels, and their variation in height is more gradual. In this particular instance, increased resolution is responsible for the binary case detecting the narrow pulse at a lower level (closer to 0% threshold). This illustration shows the general results that one would expect from increasing the spatial density at which one samples the image; that is, one sees somewhat finer detail in both the gray and the binary images with higher sampling frequency.

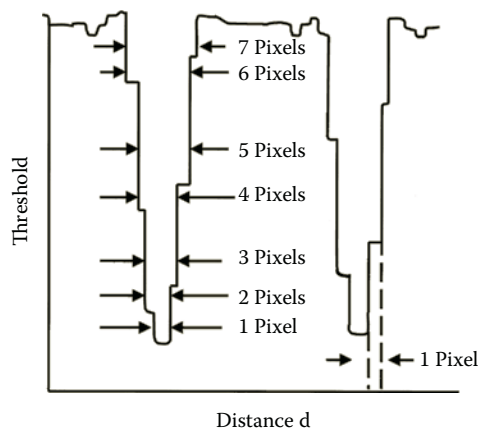


FIGURE 3.6

An actual scanned example of a gray scan line across the center of a letter “I.” A different representation of the effect shown in step (c) in Figure 3.5. Here the sample points are displayed as contiguous pixels. The width of one pixel is indicated. The image is from a 400 dpi scan of approximately a six-point Roman font.

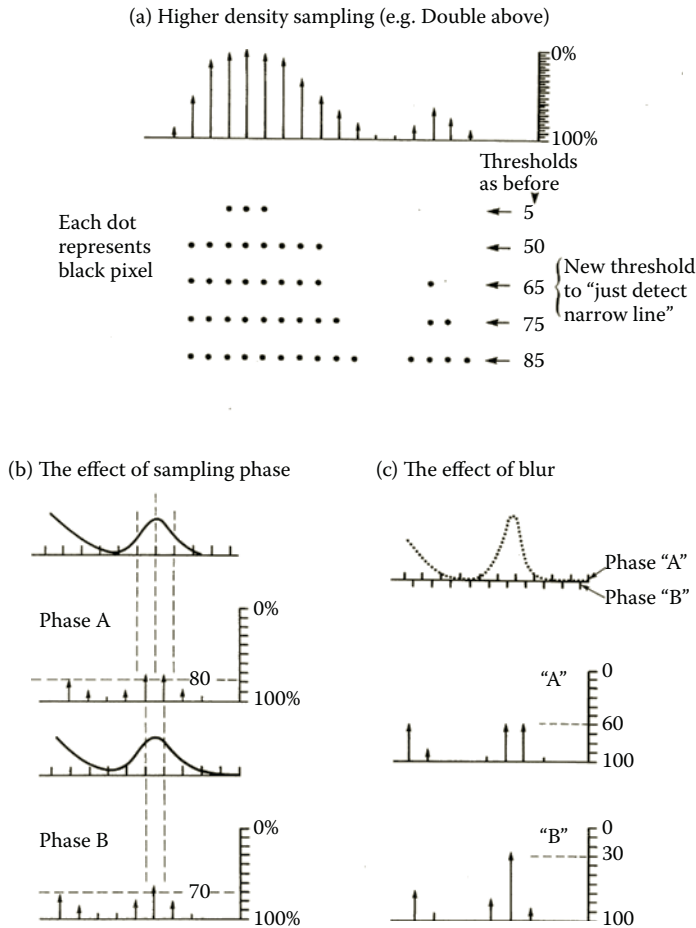


FIGURE 3.7 The effects of (a) doubling the resolution, (b) changing sampling phase, (c) sharpening the optical image.

This is, however, not always the case when examining every portion of the microstructure. Let us look more closely at the narrower of the two pulses (Figure 3.7b). Here we see the sampling occurring at two locations, shifted slightly with respect to each other. These are said to be at different sampling phases. In phase A the pulse has been sampled in such a way that the separate pixels near the peak are identical to each other in their intensity, and in phase B one of the pixels is shown centered on the peak. When looking at the threshold required to detect the information in phase A and phase B, different results are obtained for a binary representation of these images. Phase B would show the detection of the pulse at a lower threshold (closer to ideal) and phase A, when it detects the pulse, would show it as wider, namely as two pixels in width.

Consider an effect of this type in the case of an input document scanner, such as that used for facsimile or electronic copying. While the sampling array in many input scanners is constant with respect to the document platen, the location of the document on the platen is random. Also the locations of the details of any particular document within the format of the sheet of paper are random. Thus the phase of sampling with respect to detail is random and the type of effects illustrated in Figure 3.7 would occur randomly over a page.

There is no possibility that a document covered with some form of uniform detail can look absolutely uniform in a sampled image. If the imaging system produces binary results, it will consistently exhibit errors on the order of one pixel and occasionally two pixels of edge position and line width. The same is true of a typically quantized gray image, except now the errors are primarily in magnitude and may, at higher sampling densities, be less objectionable. In fact, an analog gray imaging process, sampling at a sufficiently high frequency, would render an image with no visible error (see the next subsection). Continuing with the same basic illustration, let us consider the effect of blur. In Figure 3.7c we have sketched a less blurred image in the region of the narrower pulse and now show two sampling phases A and B, as before, separated by half a pixel width. Two things should be noted. First, with higher sharpness (i.e., less blur), the threshold at which detection occurs is higher. Secondly, the effect of sampling phase is much larger with the sharper image. Highly magnified images in Figure 3.8 illustrate some of these effects.

3.2.1.2 The Sampling Theorem and Spatial Relationships

By means of these illustrations we have shown the effects of sampling frequency, sampling phase, and blur at an elementary level. We now turn our attention to the more formal description of these effects in what is known as the sampling theorem. For these purposes we assume that the reader has some understanding of the concepts of Fourier analysis or at least the frequency-domain way of describing time or space, such as in the frequency analysis of audio equipment. In this approach, distance in millimeters is transformed to frequency in cycles per millimeter (cycles/mm). A pattern of bars spaced 1 mm apart would result in 1 cycle/mm as the fundamental frequency of the pattern. If the bars were represented by a square wave, the Fourier series showing the pattern's various harmonics would constitute the frequency-domain equivalent.

Figure 3.9 has been constructed from such a point of view. In Figure 3.9a we see a single-raster profile of an analog input document (i.e., an object) represented by the function $f(x)$.

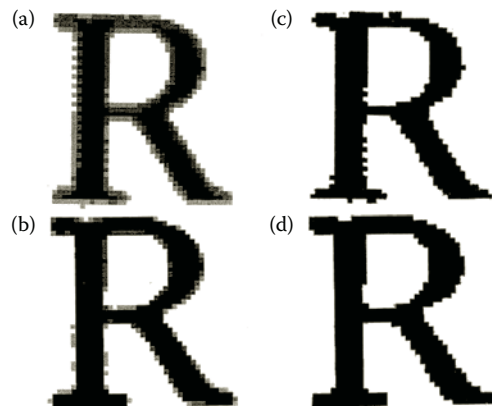


FIGURE 3.8

Digital images of a 10-point letter "R" scanned at 400 dpi showing quantization and sharpening effects. Parts (a) and (c) were made with normal sharpness for typical optical systems and parts (b) and (d) show electronic enhancement of the sharpness (see Figure 3.32). Parts (a) and (b) are made with 2 bits/pixel, that is, four levels including white, black and two levels of gray. Parts (c) and (d) are 1 bit/pixel images, that is, binary with only black and white where the threshold was set between the two levels of gray used in (a) and (b). Note the thickening of some strokes in the sharper image and the increased raggedness of the edges in the binary images. Some parts of the sharp binary images are also less ragged.

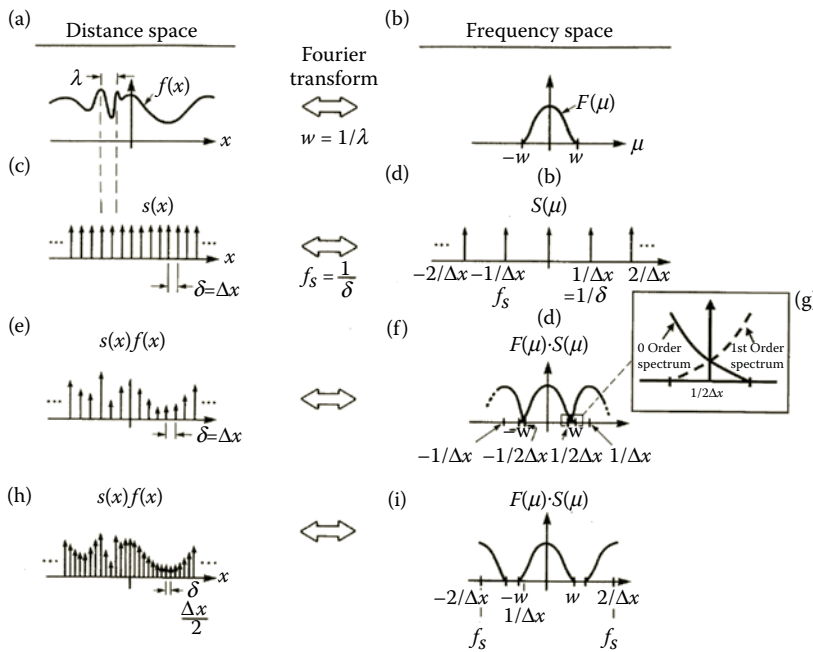


FIGURE 3.9

The Fourier transformation of images and the effects of sampling frequency. The origin and prevention of aliasing: (a) original object; (b) spectrum of object; (c) sampling function; (d) spectrum of sampling function; (e) sampled object; (f) spectrum of sampled object; (g) detail of sampled object spectrum; (h) object sampled at double frequency; and (i) spectrum of object sampled at double frequency. (Adapted from Gonzalez, R.C.; Wintz, P. *Digital Image Processing*; Addison, Wesley: Reading, MA, 1977; 36–114.)

This is a signal extending in principle to $+\infty$ and contains, upon analysis, many different frequencies. It could be thought of as a very long microreflectance profile across an original document. Its spectral components, that is, the relative amplitudes of sine waves that fit this distribution of intensities, are plotted as $F(\mu)$ in Figure 3.9b. Note that there is a maximum frequency in this plot of amplitude versus frequency, at w . It is equal to the reciprocal of λ (the wavelength of the finest detail) shown in Figure 3.9a. This is the highest frequency that was measured in the input document. The frequency w is known as the bandwidth limit of the input document. Therefore the input document is said to be band limited. This limit is often imposed by the width of a scanning aperture that is performing the sampling in a real system.

We now wish to take this analog signal and convert it into a sampled image. We multiply it by $s(x)$, a series of narrow impulses separated by Δx as shown in Figure 3.9c. The product of $s(x)$ and $y f(x)$ is the sampled image, and that is shown in Figure 3.9e. To examine this process in frequency space, we need to find the frequency composition of the series of impulses that we used for sampling. The resulting spectrum is shown in Figure 3.9d. It is, itself, a series of impulses whose frequency locations are spaced at $1/\Delta x$ apart. For the optical scientist this may be thought of as a spectrum, with each impulse representing a different order; thus the spike at $1/\Delta x$ represents the first-order spectrum, and the spike at zero represents the zero-order spectrum. Because we multiplied in distance space in order to come up with this sampled image, in frequency space, according to the convolution theorem, we must convolve the spectrum of the input document with the spectrum of the

sampling function to arrive at the spectrum of the sampled image. The result of this convolution is shown in Figure 3.9f.

Now we can see the relationship between the spectral content of the input document and the spacing of the sampling required in order to record that document. Because the spectrum of the document was convolved with the sampling spectrum, the negative side of the input document spectrum $F(u)$ folds back from the first-order over the positive side of the zero-order document spectrum. Where these two cross is exactly halfway between the zero- and first-order peaks. It is a frequency $(1/2\Delta x)$ known as the Nyquist frequency. If we look at the region in Figure 3.9g between zero and the Nyquist frequency, the region reserved for the zero-order information, we see that there is “contamination” from the negative side of the first order down to the frequency $[(1 - \Delta x) - w]$, where w is the band limit of the signal. Any frequency above that point contains information from both the zero and the first order and is therefore corrupted or mixed, often referred to as *aliased*.

Should one desire to avoid the problem of aliasing, one must sample at a finer sampling interval, as shown in Figure 3.9h. Here the spacing is one-half that of the earlier sketches, and therefore the sampling frequency is twice as high. This also doubles the Nyquist frequency. This merely separates the spectra by spreading them out by a factor of 2. Since there is no overlap of zero and first orders in this example, one can recover the original signal quite easily by simply filtering out the higher frequencies representing the orders other than zero. This is illustrated in Figure 3.10, where a rectangular function of width $\pm w$ and amplitude 1 is multiplied by the sampled image spectra, resulting in recovery of the original signal spectra. When inversely Fourier transformed, this would give the original signal back [compare Figures 3.10e and 3.9a].

We can now restate Shannon’s⁴³ formal sampling theorem, [sometimes referred to as the Whittaker–Shannon Sampling Theorem (R. Loce, personal communication, 2001)] in terms

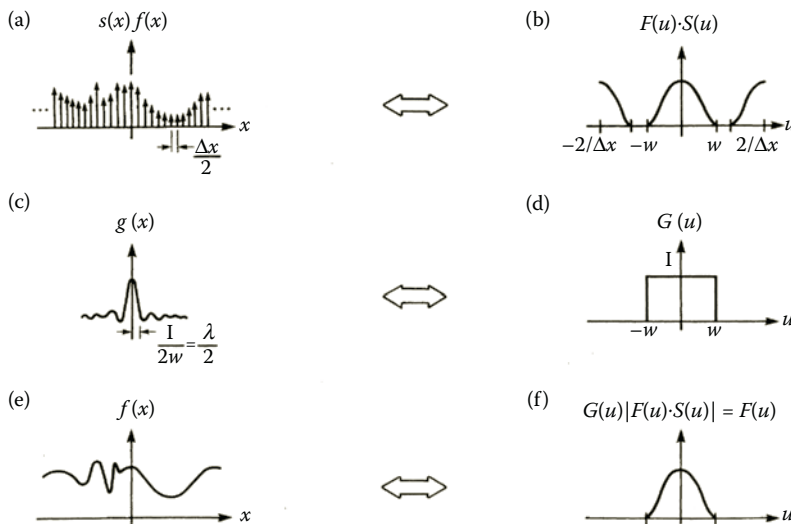


FIGURE 3.10

Recovery of original object from properly sampled imaging process: (a) object sampled at double frequency (from Figure 3.9h); (b) spectrum of “a” (from Figure 3.9i); (c) spread function for rectangular frequency filter function; (d) rectangular frequency function; (e) recovered object function; (f) recovered object spectrum. (Adapted from Gonzalez, R.C.; Wintz, P. *Digital Image Processing*; Addison, Wesley: Reading, MA, 1977; 36–114.)

that apply to sampled imaging: if a function $f(x)$ representing either an original object or the optical/aerial image being digitized contains no frequencies higher than w cycles/mm (this means that the signal is band limited at w), it is completely determined by giving its values at a series of points $<1/2w$ mm apart. It is formally required that there be no quantization or other noise and that this series be infinitely long; otherwise windowing effects at the boundaries of smaller images may cause some additional problems (e.g., digital perturbations from the presence of sharp edges at the ends of the image). In practice, it needs to be long enough to render such windowing effects negligible.

It is clear from this that any process such as imaging by a lens between the document and the actual sampling, say by a CCD sensor, can band limit the information and ensure accurate effects of sampling with respect to aliasing. However, if the process of band limiting the signal in order to prevent aliasing causes the document to lose information that was important visually, then the system is producing restrictions that would be interpreted as excessive blur in the optical image. Another way to improve on this situation is, of course, to increase the sampling frequency, that is, decrease the distance between samples.

We have shown in Figure 3.10 that the process of recovering the original spectrum is accomplished by a filter having a rectangular shape in frequency space (Figure 3.10d). This filter is known as the reconstruction filter and represents an idealized reconstruction process. The rectangular function has a $(\sin x)/x$ inverse transform in distance space (Figure 3.10c), whose zero crossings are at $\pm N\Delta x$ from the origin where $N = 1, 2, \dots$. Rectangular and other filters with flat modulation transfer functions (MTFs) are difficult to realize in incoherent systems. This comes about because of the need for negative light in the sidelobes (in distance space). A reconstruction filter need not be precisely rectangular in order to work. It should be relatively flat and at a value near 1.0 over the bandwidth of the signal being reconstructed (also difficult and often impossible to achieve). It must not transmit any energy from the two first-order spectra. If the sampling resolution is very high and the bandwidth of the signal is relatively low, then the freedom to design the edge of this reconstruction filter is relatively great and therefore this edge does not need to be as square. From a practical point of view the filter is often the MTF of the output scanner, typically a laser beam scanner, and is not usually a rectangular function but more of a Gaussian shape. A nonrectangular filter, such as that provided by a Gaussian laser beam scanner, alters the shape of the spectrum that it is trying to recover. Because the spectrum is multiplied by the reconstructing MTF, this causes some additional attenuation in the high frequencies, and a trade-off is normally required in practical designs.

3.2.1.3 Gray Level Quantization: Some Limiting Effects

Now that we have seen how the spatial or distance dimension of an input image may be digitized into discrete pixels, we explore image quantization into a finite number of discrete gray levels. From a practical standpoint this quantization is accomplished by an A/D converter, which quantizes the signal into a number of gray levels, usually some power of 2. A popular quantization is 256 levels, that is, 8 bits, which lends itself to many computer applications and standard digital hardware. There may be good reasons for other quantizations, higher or lower, to optimize a design or a system. (See Reference 21, pp. 213–227 for some practical applications and tests.)

From an overall systems engineering perspective, one needs to understand the limits on the useful number of quantization levels. This should be based upon noise in the input as seen by the system (inbound limit) or upon the ultimate output goal of how many distinguishable gray levels can be seen by the human eye (outbound limit). Both approaches

have been explored in the literature and involve complex calculations and experimental measurements.

Use of the *HVS response* with various halftoning methods represents an *outbound limit* approach to defining practical quantization limits for scanned imaging. The “visual limit” results shown in Figure 3.11⁴⁴ plot the number of visually distinguishable gray levels against the spatial frequency at which they can be seen. This curve was derived from a very conservative estimate of the visual system frequency response and may be thought of as an upper limit on the number of gray levels required by the eye. Plotted on the same curve are performance characteristics for 20 pixels/mm (500 pixels/in) digital imaging systems that produce 3 bits/pixel and 1 bit/pixel (binary) images. These were obtained by use of a generalized algorithm to create halftone patterns (see Section 3.2.2.3 and Reference 45) at different spatial frequencies. The binary limit curve, added here to Roetling’s, graph, shows the number of effective gray levels for each frequency whose period is two halftone cells wide. The 3-bit limit assumes each halftone cell contributes 2^3 gray values, including black and white.

Roetling⁴⁴ integrated the visual response curve to find an average of 2.8 bits/pixel as a good upper bound for the eye itself. Note that his general halftoning approach, using 3 bits/pixel and 20 pixels/mm (500 pixels/in) also approximates the visual limit in the important midfrequency region. Specialized halftoning techniques^{7,45} may produce different and often more gray levels per pixel at the lower frequencies.

Another approach to setting quantization limits is to examine the noise in the input, assuming in so doing that the quantization is *input bound* and not output bound by the visual process as in the foregoing approach. A range of photographic input was selected as examples of a practical lower limit (best) on input noise. The basic principle for describing the useful number M of gray levels in a photograph involves quantizing its density scale into steps whose size is based on the noise (granularity) of that photographic image⁴⁶ when scanned by the digital imaging process. In simplified terms this can be described as

$$M = \frac{L}{2k\sigma_a} \quad (3.1)$$

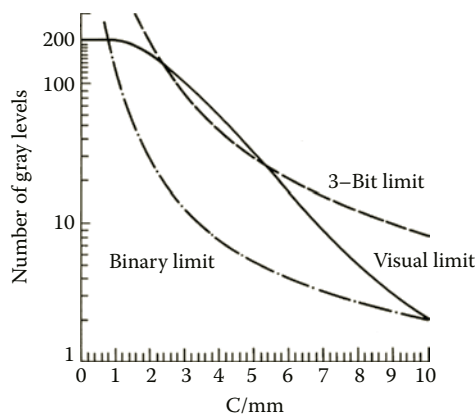


FIGURE 3.11

Example of outbound quantization limits, using visually distinguishable number of gray levels versus spatial frequency, with corresponding 1 (binary) (adapted from Roetling, P.G. Visual performance and image coding. *Proceedings of the Society of Photo-Optical Instrumentation Engineers on Image Processing*, Vol. 74, 1976; 195–199.) and 3 bit/pixel limits.

where L = the density range of the image, σ_a = measured standard deviation of density using aperture area = A , and k = the number of standard deviations in each distinguishable level.

The question being addressed by this type of quantization is how reliably one wants to be able to determine the specific tone in a given part of the input picture from a reading of a single pixel. For some purposes, where the scanned image is used to extract radiometric information from a picture,⁴⁶ the reliability must be high, for other cases such as simply copying a scene for artistic purposes it can be much lower. To precisely control a digital halftone process (see later) it must be fairly high.

Photographic noise is approximately random uncorrelated noise. To a first order, photographic noise (granularity) is the standard deviation of the density fluctuations. It is directly proportional to the square root of the effective detection area,^{47,48} a of a measuring instrument or scanner-sensor, that is, Selwyn's law:

$$s_a = S(2a)^{1/2} \quad (3.2)$$

where S is a proportionality constant defined as the Selwyn granularity. It is also proportional to the square root of the mean density, that is, Siedentopf's relationship,^{47,49} in an ideal film system. In practical cases, as is done here, the density relationship must be empirically determined. Figure 3.12 shows the number of distinguishable gray levels reported in the literature by various authors for various classes of films obtained by directly measuring granularity as a function of density. They are reported at apertures that are approximately equivalent in size to the smallest detail the film could resolve, that is, the diameter of the film spread function. For a real world example, assume that 35-mm film images are enlarged perfectly by a high-quality $3.3 \times$ enlarger. The conversion to the number of distinguishable gray levels per pixel is based on assuming Selwyn's law, a reliability of 99.7% ($\pm 3\sigma$ or $k = 6$) and that any nonlinear relationship between granularity and density scales as the aperture size changes. The actual scanner aperture is reduced by $3.3 \times$ in its two dimensions to resemble directly scanning the film.

Four specific films were selected, each representative of a different class, three of which are black and white films: (1) an extremely fine-grained microfilm; (2) a fine-grained amateur film; and (3) a high-speed amateur film.⁵⁰ A special purpose color film was also included.⁵¹ Despite now being obsolete, these films still represent a reasonable cross section of photographic materials. A $3.3 \times$ enlargement was selected as typical of consumer practice, roughly giving a $3.5'' \times 5''$ print from a 35-mm negative. The reciprocal of this magnification is used to scale the scanner aperture back to film dimensions. Two popular scanner resolutions of 600 and 300 dpi were selected. The corresponding sensor "aperture" widths in μm , scaled to the film, are noted in parentheses in the key at the top of each figure. The width is the inverse of the sampling period. A third scanner aperture, equivalent to that in the Roetling visual calculations, was used for one case, that is, a 20 samples/mm (500 samples/in) scanning system with an aperture of $50 \times 50 \mu\text{m}$ (2×2 mils). The fourth situation, called "Film @ max" describes the number of levels resulting from scanning the film with an aperture that matches the blur (spread function) for the film, given in the film category label in parentheses at the bottom of each figure. These approximate calculations are an oversimplification of the photographic and enlarging processes, ignoring significant nonlinearities and blurring effects, but they provide a rough first-order analysis.

Examination of the charts suggests that a practical range of inbound quantization limits (IQLs) for pictorial images is approximately anywhere from 2 to 4 bits/pixel (microfilm is

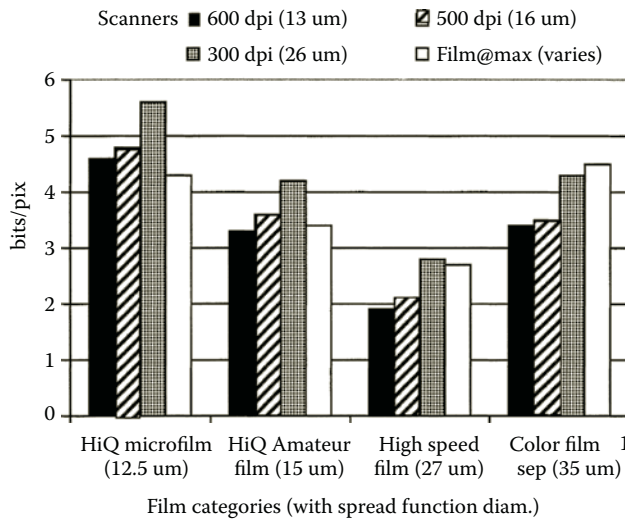


FIGURE 3.12

Example of inbound quantization limits, using the number of distinguishable gray levels, in bits/pixel, for input consisting of $3.3 \times$ enlargements (3×5 in prints of 35 mm film) from four example films (adapted from Altman, J.H.; Zweig, H.J. Effect of spread function on the storage of information on photographic emulsions. *Photog. Sci. Eng.* 1963, 7, 173–177 and Lehmbeck, D.R. Experimental study of the information storing properties of extended range film. *Photog. Sci. Eng.* 1967, 11, 270–278.) scanned by four generic types of systems indicated by their scanning resolutions. Color film is for a single separation, others are black and white films. The limiting blur in μm for the first three scanners is given in the parentheses after the scan frequency. It is the sensor aperture width scaled to the film size. The fourth scanner has variable resolution set by a scaled aperture width adjusted to equal the width of each film blur function (spread function), shown in parentheses with the film type. Assumes a 99.7% confidence on distinguishability using Equation 3.1 (i.e., with $k = 6$).

not made for pictorials). For typical high-quality reproduction, then, an input bound limit is a little over 3 bits/pixel at 600 dpi using the three standard deviation criterion. This compares with the rate of 2.8 bits/pixel found by Roetling for a visual outbound quantization limit (OQL). Recent work by Vaysman and Fairchild,⁵² limited to an upper frequency of 300 dpi by their printer selection, also found, through psychophysical studies, that 3 bits/pixel/color was a useful system optimum for reproducing color pictures.

One may ask, then, why are there so many input scanners operating at 8, 10, or even 12 bits/pixel? First of all there are many reasons to modify these calculations for specific situations such as larger tolerances on probabilities for distinguishing differences less reliably, considering larger sampling apertures for certain rendering/viewing methods, different frequency weightings and many others that would result in more inbound or outbound gray levels (Reference 53, p. 198).

A very practical reason, however, is that actual hardwired scanners cannot adapt to detail and granularity in originals and change performance striving for these optimums in the way they were calculated. For slower, computationally intensive, off-line image processing which can adapt to the information in small regions of the image (as in the case of JPEG and other lossy compressions—see Section 3.7.1) one can, in essence, approximate the limits just discussed. (As an example $10\text{--}20 \times$ JPEG compression for 8-bit images works well and is approx. 4+ bits, leaving 3+ bits for the resulting image.)

Actual hardwired real-time scanners have to assume the worst case (e.g., 200 gray levels—see Figure 3.11). This is rounded up to 256 or 8 bits. However, the 200 gray levels

are not equally spaced in linear units. They are essentially spaced as equal increments on an L^* scale (See Equations 3.5 and 3.8.)

Thus for an input density of 2.0 or an L^* of ~ 9 and a difference of $.5 L^*$ ($=1/200$ of full L^* scale) a linear difference of $\sim .0006$ is called for. This is 1 part out of 1700, or more than 10 bits (1024) and would require an 11-bit system (not a common A to D circuit). From a practical perspective that suggests 12 bits (4096 levels) which allows for some enhancement of high density areas. For those believing that a ΔL^* of 1 is just noticeable (true for certain conditions) the above situation calls for approx. 870 levels and 10 bits is satisfactory. Another alternative to stay with 8 bits and meet the visual requirements is to distort the linear sensor response via scanner electronics to approximate L^* (some digital cameras and scanners do this) prior to the final digital output

Being aware of the inbound limits, the system options and the outbound limits as an endpoint gives a framework for robust engineering and optimization of image quality in a systems context. Information capacity approaches extend these concepts (see Section 3.6.7).

3.2.2 Basic System Effects

3.2.2.1 Blur

Blur, that is, the spreading of the microscopic image structure, is a significant factor in determining the information in an image and therefore its quality. In the input scanner, blur is caused by the optical system, the size and properties of the light-sensing element, other electronic elements, and by mechanical and timing factors involved in motion. This blur determines whether the system is aliased. Roughly speaking, if the image of a point (the profile of which is called the point spread function) spreads over twice the sampling interval, the system is unaliased. The spreading also determines the contrast of fine details in the gray video image prior to processing. The cascading of these elements can be described conveniently by a series of spatial frequency responses [see later under MTFs for a detailed discussion] or other metrics that relate generally to the sharpness of optical images. It can be compensated for, in certain aspects, by subsequent electronic or computer image processing.

Blur in an output scanner is caused by the size of the writing spot, for example, the laser beam waist at focus, by modulation techniques and by the spreading of the image in any marking process such as xerography or photographic film. It is also affected by motion of the beam relative to the data rate and by the rate of motion of the light-sensitive receptor material. Output scanner blur more directly affects the appearance of sharpness in the final hard copy image that is presented to the HVS than does blur in the input scanner. Overall enhancement of the electronic input scanned image can, however, draw visual attention to details of the output image unaffected by blur limitations of either scanner.

Blur for the total system, from input scanner through various types of image processing to output scanner and then to marks on paper, is not easily cascaded, because the intervening processing of the image information is extremely nonlinear. This nonlinearity may give rise to such effects as a blurred input image looking very sharp on the edges of a binary output print because of the small spot size and low blur of the marking process. In such a case, however, the edges of square corners look rounded and fine detail such as serifs in text or textures in photographs may be lost. Conversely, a sharp input scan printed by a system with a large blurring spot would appear to have fuzzy edges, but the edge noise due to sampling would have been blurred together and would be less visible than in the first case. Moiré, from aliased images of periodic subjects caused by

low blur relative to the sample spacing, however, would still be present in spite of output blur. (Note, superposition of periodic patterns such as a halftoned document (see Section 3.2.2.3) and the sampling grid of a scanner results in new and often striking periodic patterns in the image commonly called Moiré patterns (see Bryngdahl⁵⁴). Once aliased, no amount of subsequent processing can remove this periodic aliasing effect from an image.) The popular technologies called “anti aliasing” deal with a different effect of undersampling, namely that binary line images exhibit strong visible staircase or jaggie effects on slanted lines when the output blur and sampling are insufficient for the visual system. These techniques nonlinearly “find” the stair steps and locally add gray pixels to reduce the visibility of the jaggie (see Section 3.7.2 and Figure 3.43).⁵⁵ Aliasing is also known as spurious response.¹³

It is apparent, then, that blur can have both positive and negative impacts on the overall image quality and requires a careful trade-off analysis when designing scanners.

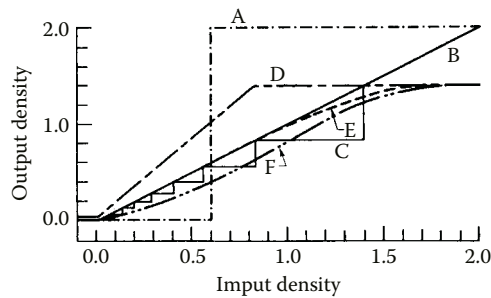
3.2.2.2 System Response

There are four ways in which electronic imaging systems display or print tonal information to the eye or transmit tonal information through the system:

1. By producing a signal of varying strength at each pixel, using either amplitude or pulse-width modulation.
2. By turning each pixel on or off (a two-level or binary system; see Section 3.5).
3. By use of a halftoning approach, which is a special case of binary imaging. Here, the threshold for the white–black decision is varied in some structured way over very small regions of the image, simulating continuous response. Many, often elaborate, methods exist for varying the structure; some involve multiple pixel interactions (such as error diffusion; see the end of Section 3.2.2.3) and others use subpixels (such as high addressability, extensions of the techniques mentioned in Section 3.7.2).
4. By hybrid halftoning combining the halftone concept in (3) with the variable gray pixels from (1) (e.g., see References 44 and 45).

From a hardware point of view, the systems are either designed to carry gray information on a pixel-by-pixel basis or to carry binary (two-level) information on a pixel-by-pixel basis. Because a two-level imaging system is not very satisfactory in many applications, some context is added to the information flow in order to obtain pseudogray using the halftoning approach.

Macroscopic tone reproduction is the fundamental characteristic used to describe all imaging systems’ responses, whether they are analog or digital. For an input scanner it is characterized by a plot of an appropriate, macroscopic output response, as a function of some representation of the input light level. The output may characteristically be volts or digital gray levels for a digital input scanner and intensity or perhaps darkness or density of the final marks-on-paper image for an output scanner. The correct choice of units depends upon the application for which the system response is being described. There are often debates as to whether such response curves should be in units of density or optical intensity, brightness, visual lightness or darkness, gray level, and so on. For purposes of illustration, see Figure 3.13.

**FIGURE 3.13**

Some representative input/output density relationships for (A) binary imaging response; (B) linear imaging response; (C) stepwise linear response; (D) saturation-limited linear response; (E) linear response with gradual roll-off to saturation; (F) idealized response curve for best overall acceptability.

Here we have chosen to use the conventional photographic characterization of output density plotted against input density using normalized densities. Curve A shows the case of a binary imaging system in which the output is white or zero density up to an input density of 0.6, at which point it becomes black or 2.0 output density. Curve B shows what happens when a system responds linearly in a continuous fashion to input density. As the input is equal to the output here, this system would be linear in reflectance, irradiance, or even Munsell value (visual lightness units).

Curve C shows a classic abridged gray (severely limited number of levels) system attempting to write linearly but with only eight levels of gray. This response becomes a series of small steps, but because of the choice of density units, which are logarithmic, the sizes of the steps are very different. Had we plotted output reflectance as a function of input reflectance, the sizes of the steps would have been equal. However, the visual system that usually looks at these tones operates in a more or less logarithmic or power fashion, hence the density plot is more representative of the visual effect for this image. Had we chosen to quantize in 256 gray levels, each step shown would have been broken down into 32 smaller substeps, thereby approximating very closely the continuous curve for B.

When designing the system tone reproduction, there are many choices available for the proper shape of this curve. The binary curve, as in A, is ideal for the case of reproducing high-contrast information because it allows the minimum and maximum input densities considerable variation without any change to the overall system response.

For reproducing continuous tone pictures, there are many different shapes for the relationship between input and output, two of which are shown in Figure 3.13. If, for example, the input document is relatively low contrast, ranging from 0 to 0.8 density, and the output process is capable of creating higher densities such as 1.4, then the curve represented by D would provide a satisfactory solution for many applications. However, it would create an increase in contrast represented by the increase in the slope of the curve relative to B, where B gives one-for-one tone reproductions at all densities. Curve D is clipped at an input density greater than 0.8. This means that any densities greater than that could not be distinguished and would all print at an output density of 1.4.

In many conventional imaging situations the input density range exceeds that of the output density. The system designer is confronted with the problem of dealing with this mismatch of dynamic ranges. One approach is to make the system respond linearly to density up to the output limit; for example, following curve B up to an output density

of 1.4 and then following curve D. This generally produces unsatisfactory results in the shadow regions for the reasons given earlier for curve D. One general rule is to follow the linear response curve in the highlight region and then to roll off gradually to the maximum density in the shadow regions starting perhaps at a 0.8 output density point for the nonlinear portion of the curve as shown by curve E. Curve F represents an idealized case approximating a very precisely specified version arrived at by Jorgenson.⁵⁶ He found the "S"-shaped curve resembling F to be a psychologically preferred curve among a large number of the curves he tried for lithographic applications. Note that it is lighter in the highlights and has a midtone region where the slope parallels that of the linear response. It then rolls off much as the previous case toward the maximum output density at a point where the input density reaches its upper limit.

3.2.2.3 Halftone System Response

One of the advantages of digital imaging systems is the ability to completely control the shape of these curves to allow the individual user to find the optimum relationship for a particular photograph in a particular application. This can be achieved through the mechanism of digital halftoning as described below. Historically important studies of tone reproduction, largely for photographic and graphic arts applications, include those of Jones and Nelson,⁵⁷ Jones,⁵⁸ Bartleson and Breneman,⁵⁹ and two excellent review articles, covering many others, by Nelson.^{60,61} Many recent advances in the technology of digital halftoning have been collected by Eschbach.⁷

The halftoning process can be understood by examination of Figure 3.14. In the top of this illustration two types of functions are plotted against distance x , which has been marked off into increments one pixel in width. The first functions are three uniform reflectance levels, R1, R2, and R3. The second function $T(x)$ is a plot of threshold versus distance, which looks like a series of up and down staircases, that produces the halftone pattern. Any pixels whose reflectance is equal to or above the threshold is turned on, and any that is below the threshold for that pixel is turned off.

Also sketched in Figure 3.14 are the results for the thresholding process for R1 on the second line and then for R2 and R3 on the third line. The last two are indistinguishable for this particular set of thresholding curves. It can be seen from this that the reflectance information is changed into width information and thus that the method of halftoning is a mechanism for creating dot growth or spatial pulse-width modulation over an area of several pixels. Typically, such threshold patterns (i.e., screens) are laid out two-dimensionally. An example is shown in Figure 3.15.

This thresholding scheme emulates the printer's 45° screen angle, which is considered to be favorable from a visual standpoint because the 45° screen is less visible (oblique effect⁴) than the same 90° screen. Other screen angles may also be conveniently generated by a single string of thresholds and a shift factor that varies from raster to raster.^{62,63} The numbers in each cell in the matrix represent the threshold required in a 32-gray-level system to turn the system on or off. The sequence of thresholds is referred to as the dot growth pattern. At the bottom, four thresholded halftone dots (Parts b–e) are shown for illustration. There are a total of 64 pixels in the array but only 32 unique levels. This screen can be represented by 32 values in a 4 × 8 pixel array plus a shift factor of four pixels for the lower set of 32, which enables the 45° screen appearance as illustrated. It may also be represented by 64 values in a single 8 × 8 pixel array, but this would be a 90° screen. It is also possible to alternate the thresholding sequence between the two 4 × 8 arrays, where the growth pattern in each array is most commonly in a spiral pattern, resulting in two unique sets of

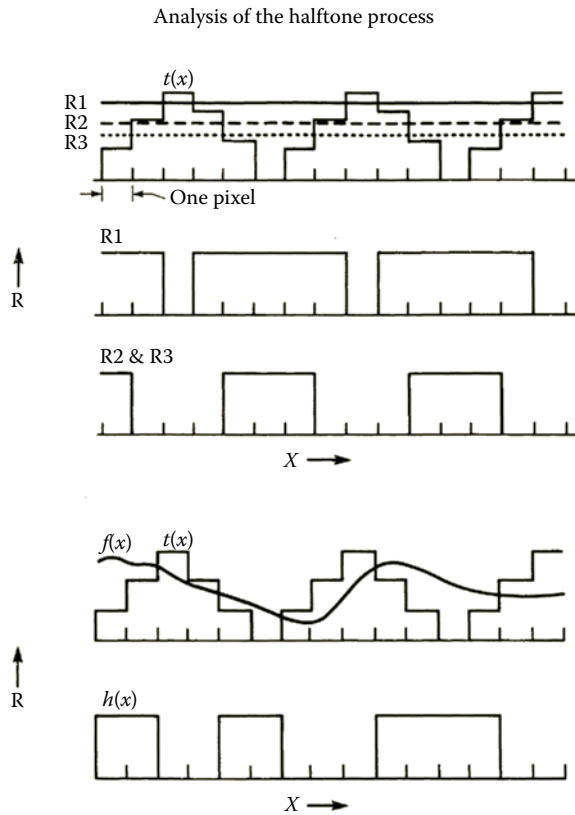
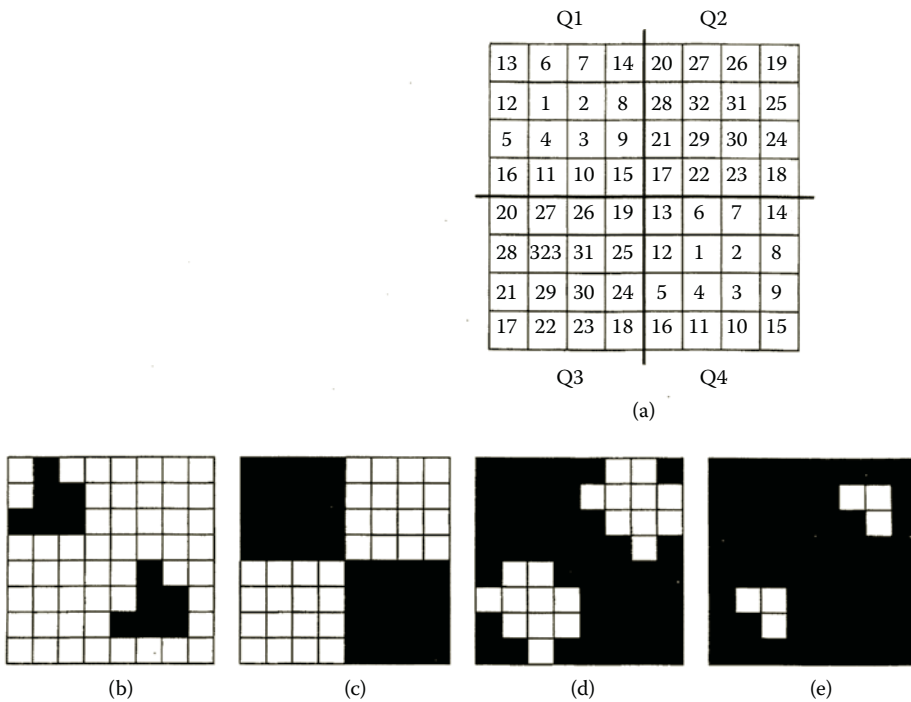


FIGURE 3.14

Illustration of halftoning process. Each graph is a plot of reflectance R versus distance X . $T(x)$ is the profile of one raster of the halftone threshold pattern, where image values above the pattern are turned on (creates black in system shown) by the halftone thresholding process. R1, R2, and R3 represent three uniform images of different average reflectances shown at the top as uniform input and in the middle of the chart as profiles of halftone dots after halftone thresholding. $f(x)$ represents an image of varying input reflectance and $t(x)$ is a different threshold pattern. $h(x)$ is the resulting halftone dot profile, with dots represented, here, as blocks of different width illustrating image variation.

32 thresholds for an equivalent of 64 different levels and preserving the screen frequency as shown. This screen is called a “double dot.” The concept is sometimes extended to four unique dot growth patterns and hence is named a “quad dot.” Certain percent area coverage dot patterns in these complex multicentered dot structures generate very visible and often objectionable patterns.

The halftone matrix described in Figure 3.15 represented 32 specific thresholds in a specific layout. There are many alternatives to the size and shape of the matrix, the levels chosen, the spatial sequence in which the thresholds occur, and arrangements of multiple, uniquely different matrices in a grouping called a super cell. Here there are many different cells (more than the four in a quad dot) varying slightly in shape and each may contain a slightly different number of pixels. This gives its designer even more gray levels since there are more cells and each may contain unique thresholds. There are also more available angles due to the size and shape differences of the individual cells giving the centers of the collection of all the supercells more precision to form a new screen angle. See Figure 3.16.

**FIGURE 3.15**

Example of two-dimensional quantized halftone pattern, with illustrations of resulting halftone dots at various density levels. (a) 8×8 spiral halftone matrix; (b) density = 0.10 or 20% fill 12/64 pixels; (c) density = 0.30 or 50% fill 32/64 pixels (d) density = 0.50 or 68% fill 44/64 pixels; (e) density = 1.00 or 90% fill 58/64 pixels.

The careful selection of these factors gives good control over the shape of the apparent tone reproduction curve, granularity, textures, and sharpness in an image. The halftone system's ability to resolve structures finer than the halftone screen array or cell size has been described as "partial dotting" by Roetling⁶⁴ and others and is an important and often misunderstood factor in image quality studies (Reference 53, p. 163; Reference 65, p. 403). It is the result of the high-resolution pixel-by-pixel comparison of the threshold matrix and the image detail which allows high-contrast image detail to pass through the halftone matrix, nearly unchanged.

There are also many other methods for converting binary images into pseudogray images using digital halftoning methods of a more complex form.^{66,67} These include alternative dot structures, that is, different patterns of sequences in alternating repeat patterns, random halftoning, and techniques known as error diffusion. In his book *Digital Halftoning*, Ulichney⁶⁸ describes five general categories of halftoning techniques:

1. Dithering with white noise (including mezzotint)
2. Clustered dot ordered dither
3. Dispersed dot ordered dither (including "Bayer's dither")
4. Ordered dither on asymmetric grids
5. Dithering with blue noise (actually error diffusion)

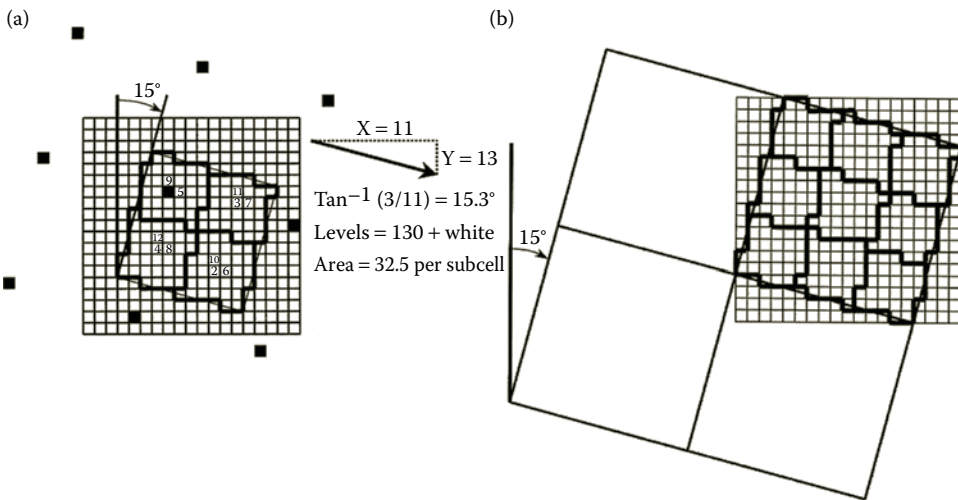


FIGURE 3.16

Examples of multicentered dots: (a) a classic quad dot showing the first three thresholds in each individual cell and the large black dots showing the repeat pattern centers at 15.255° and (b) a nine center "supercell" where the cell shape and size varies: from L to R 26, 27,27; 27,29,25; 27,27,26 pixels and the angle is 14.9°. (Reproduced with permission of the publisher, CRC Press, Boca Raton, FL, from Reference 16, p. 412 in Chapter 6 by Haines, Wang and Knox, 2003.)

He states that "spatial dithering is another name often given to the concept of digital halftoning. It is perfectly equivalent, and refers to any algorithmic process that creates the illusion of continuous tone images from the judicious arrangement of binary picture elements." The process described in Figures 3.15 and 3.16 falls into the category of a clustered dot ordered dither method (category 2) as a classical rectangular grid on a 45° base.

There is no universally best technique among these. Each has its own strengths and weaknesses in different applications. The reader is cautioned that there are many important aspects of the general halftoning process that could not be covered here. (See Reference 45 for a summary of digital halftoning technology and many references, and Reference 69 for many practical aspects of conventional halftoning for color reproduction.) For example, the densities described in Figure 3.15 only apply to the case of perfect reproduction of the illustrated pixel maps on nonlight-scattering material using perfect, totally black inks. In reality, each pattern of pixels must be individually calibrated for any given marking process. The spatial distribution interacts with various noise and blurring characteristics of output systems to render the mathematics of counting pixels to determine precise density relationships highly erroneous under most conditions. This is even true for the use of halftoning in conventional lithographic processes, due to the scattering of light in white paper and the optical interaction of ink and paper. These affect the way the input scanner "sees" a lithographic halftone original. Some of these relationships have been addressed in the literature, both in a correction factor sense^{69,70} and in a spatial frequency sense.⁷¹⁻⁷³ All of these methods involve various ways of calculating the effect that lateral light scattering through the paper has on the light reemerging from the paper between the dots.

The effects of blur from the writing and marking processes involved in generating the halftone, many of which may be asymmetric, require individual density calibrations for each of the dot patterns and each of the dithering methods that can be used to generate these halftone patterns. The control afforded through the digital halftoning process by the

careful selection of these patterns and methods enables the creation of any desired shape for the tone reproduction curve for a given picture, marking process, or application.

3.2.2.4 Noise

Noise can take on many forms in an electronic imaging system. First there is the noise inherent in the digital process. This is generally referred to as either sampling noise associated with the location of the pixels or quantization noise associated with the number of discrete levels. Examples of both have been considered in the earlier discussion. Next there is electronic noise associated with the electronic components from the sensor to the amplification and correction circuits. As we move through the system, the digital components are generally thought to be error-free and therefore there is usually no such noise associated with them.

Next, in a typical electronic system, we find the ROS itself, often a laser beam scanner. If the system is writing a binary file, then the noise associated with this subsystem is generally connected with pointing of the beam at the imaging material and is described as jitter, pixel placement error, or raster distortion of some form (see the next subsection). Under certain circumstances, exposure variation produces noise, even in a binary process. For systems with gray information, there is also the possibility that the signals driving the modulation of exposure may be in error, so that the ROS can also generate noise similar to that of granularity in photographs or streaks if the error occurs repeatedly in one orientation. Finally we come to the marking process, which converts the laser exposure from the ROS into a visible signal. Marking process noise, which generally occurs as a result of the discrete and random nature of the marking particles, generates granularity.

An electronic imaging system may enhance or attenuate the noise generated earlier in the process. Systems that tend to enhance detail with various types of filters or adaptive schemes are also likely to enhance noise. There are, however, processes (see Section 3.7.2) that search through the digital image identifying errors and substitute an error-free pattern for the one that shows a mistake.^{74,55} These are sometimes referred to as noise removal filters.

Noise may be characterized in many different ways, but in general it is some form of statistical distribution of the errors that occur when an error-free input signal is sent into the system. In the case of imaging systems, an error-free signal is one that is absolutely uniform, given a noise-free, uniform input. Examples would include a sheet of white microscopically uniform paper on the platen of an input scanner, or a uniform series of laser-on pulses to a laser beam scanner, or a uniform raster pattern out of a perfect laser beam scanner writing onto the light-sensitive material in a particular marking device. A typical way to measure noise for these systems would be to evaluate the standard deviation of the output signal in whatever units characterize it. A slightly more complete analysis would break this down into a spatial frequency or time-frequency distribution of fluctuations. For example, in a photographic film a uniform exposure would be used to generate images whose granularity was measured as the root-mean-square fluctuation of density. For a laser beam scanner it would be the root-mean-square fluctuation in radiance at the pixel level for all raster lines.

In general, certain factors that affect the signal aspect of an imaging system positively, affect the noise characteristics of that imaging system negatively. For example, in scanning photographic film, the larger the sampled area, as in the case of the microdensitometer aperture, the lower the granularity [Equation 3.1]. At the same time, the image information is more blurred, therefore producing a lower contrast and smaller signal level. In general the

signal level increases with aperture area and the noise level (as measured by the standard deviation of that signal level) decreases linearly with the square root of the aperture area or the linear dimension of a square aperture. It is therefore very important when designing a scanning system to understand whether the image information is being noise limited by some fundamentals associated with the input document or test object or by some other component in the overall system itself. An attempt to improve bandwidth, or otherwise refine the signal, by enhancing some parts of the system may, in general, do nothing to improve the overall image information, if it is noise in the input that is limiting and that is being equally "enhanced." Also, if the noise in the output writing material is limiting, then improvements upstream in the system may reach a point of diminishing returns.

In designing an overall electronic imaging system it should be kept in mind that noises add throughout the system, generally in the sense of an RSS (root of the sum of the squares) calculation. The signal attenuating and amplifying aspects, on the other hand, tend to multiply throughout the system. If the output of one subsystem becomes the input of another subsystem, the noise in the former is treated as if it were a signal in the latter. This means that noise in the individual elements must be appropriately mapped from one system to the other, taking into account various amplifications and nonlinearities. In a complex system this may not be easy; however, keeping an accurate accounting of noise can be a great advantage in diagnosing the final overall image quality. We expand on the quantitative characterization of these various forms of signal and noise in the subsequent parts of this chapter.

3.2.2.5 Color Imaging

Color imaging in general and especially digital color imaging have received considerable attention in the literature in recent years.^{5,6,14,29,30} An elementary treatment is given below covering a few major points important to scanning and image quality. See References 16 or 30 for a recent broad overview and literature survey of digital color imaging, and Reference 75 for a classic review of more traditional color reproduction systems and colorimetry.

3.2.2.5.1 Fundamentals

There are two basic methods of creating images, including digital images, in color, called additive and subtractive methods.

In an *additive color* system one creates the appropriate color image pixels by combining red (R), green (G), or blue (B) microsized lights, that is, pixels of varying intensities. Roughly equal amounts of each produce the sensation of "white" light on viewing. This applies to many self-luminous displays such as a CRT/TV or liquid crystal displays. The pixels must be small enough that the eye blurs them together. The eye detects these signals using sensors called "cones" in the retina. These are associated with the HVS sensations of red, green, and blue.

In the second method of color imaging, called *subtractive color*, light is removed from otherwise white light by filters that subtract the above components one at a time. Red is removed by a cyan (C) filter, green by a magenta (M) filter, and blue by a yellow (Y) filter. For an imaging system, these filters are created by an imagewise distribution of transparent colorants created (e.g., pixel by pixel) in varying amounts. They are laid down color layer by color layer. The "white" light may come from a projector as in the case of transparencies or from white room light reflected by a white sheet of paper with the imagewise distribution of transparent colorants bonded to it. Here the subtraction occurs once on the way to the paper and then a second time after reflection on the way to the eye. Color photographic reflection prints and color offset halftone printing both use this method.

A digital color imaging system, designed to capture the colors of an original object, breaks down light reflected (or transmitted) from the object into its R, G, and B components by a variety of possible methods. It uses separate red, green, and blue image capture systems and channels of image processing, which are eventually combined to form a full color image.

The visual response involves far more than just the absorption of light. It involves the human neurological system and many special processes in the brain. The complexity of this can be appreciated by observing the results of simple color matching experiments, in which an observer adjusts the intensities of three color primaries until their mixture appears to match a test color. Such experiments, using monochromatic test colors, lead to the development of a set of color matching functions for specific sets of colored light sources and specific observer conditions. Certain monochromatic colors require the subtraction of colored light (addition of the light to the color being matched) in order to create a match. Color matching experiments are described extensively in the literature^{4,10,14,29} and provide the foundation to the science of colorimetry.

Two such sets of color matching functions are shown in Figures 3.17a and b. The first set reports experimental results using narrow band monochromatic primaries. Note the large negative lobe on the third curve of "a," showing the region where "negative light" is needed, that is, where the light must be added to the color under test to produce a match. The second set has become a universally accepted representation defining the CIEs (Commission Internationale de l'Eclairage) 1931 2° Standard Colorimetric Observer. It is a linear transformation of standardized color matching data, carefully averaged over many observers and is representative of 92% of the human population having normal color vision. This set of functions provides the standardization for much of the science of the measurement of color, in other words, important colorimetry standards.

This overly simplistic description goes beyond the scope of this chapter to explain. Ideally the information recorded by a color scanner should be equivalent to that seen by an observer. In reality, the transparent colorant materials used to create images are not perfect. Significant failures stem from the nonideal shapes of the spectral sensitivities of the capturing device and the nonideal shapes of the spectral reflectance or transmittance of the colorants. Practical limitations in fabricating systems and noise also restrict the accuracy of color recording for most scanners. Ideal spectral shapes of sensitivities and filters would allow the system designer to better approximate the HVS color response. For example, an input original composed of conventional subtractive primaries such as real magenta (green absorbing) ink, not only absorbs green light, but also absorbs some blue light. Different magentas have different proportions of this unwanted absorption. Similar unwanted absorptions exist in most cyan and, to a lesser extent, in most yellow colorants. These unwanted characteristics limit the ability of complete input and output systems to reproduce the full range of natural colors accurately. Significant work has been carried out to define quality measures for evaluating the color quality of color recording instruments and scanning devices.^{32, 6 (ch. 5), 21 (ch. 19)}. See also Table 3.9: INCITS-WI, ANSI-IT8.

3.2.2.5.2 Colorimetry and Chromaticity Diagrams

This leads to two large problem areas in color image quality needing quantification, namely: (1) that the color gamuts of real imaging systems are limited; and (2) that colors which appear to match under one set of conditions appear different by some amount under another set of circumstances. This is conveniently described by a color analysis tool from the discipline of colorimetry (the science of color measurement) called a *chromaticity diagram*, shown in Figure 3.18. It describes color in a quantitative way. It can be seen, in this

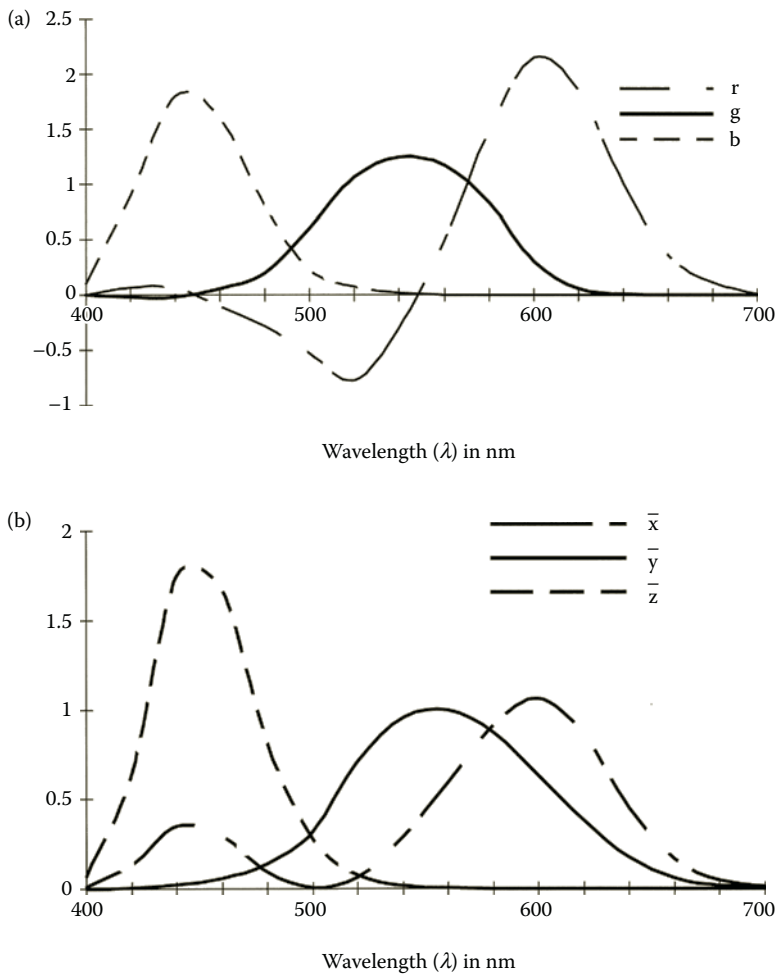


FIGURE 3.17

Color matching functions: (a) example of a directly measured result (adapted from Giorgianni, E.J.; Madden, T.E. *Digital Color Management Encoding Solutions*; Addison-Wesley: Reading, MA, 1998.); (b) a transformed result chosen as the CIE Standard Observer for 2° field of view.

illustration, that the monitor display is capable of showing different colors from a particular color printer. It is also possible, with this diagram, to show the color of an original. Note that a color gamut is the range of colors that can be produced by the device of interest as specified in some three- or more dimensional color space. It is important to note that a two-dimensional representation, like that shown here, while very helpful, is only a part of the whole three-dimensional color space. Variations derived from the chromaticity diagram, and the equations that define it, however, provide a basis for much of the literature that describes color image quality today. It is designed to facilitate description of small color differences, for example, between an original and a reproduced color or two different reproductions of the same color.

The reader must be warned, however, that the actual perception of colors involves many psychophysical and psychological factors beyond those depicted in this diagram.⁴ It is, however, a useful starting point. It describes any color in an image or a source and is

often the starting point in many of the thousands of publications on color imaging. There are also many different transformations of basic chromaticity diagram, a few primary examples of which we will describe here.

For the purposes of this chapter the basic equations used to derive the chromaticity diagram and to transform it provide an introduction to color image quality measurement. The outer, horseshoe-shaped curve, known as the “spectral locus,” represents the most saturated colors possible, those formed by monochromatic sources at different wavelengths. All other possible colors lie inside this locus. Whites or neutrals by definition are the least saturated colors, and lie nearer the center of the horseshoe-shaped area. The colors of selected broad-spectrum light sources A, B and C are shown later in Figure 3.47 (along with the equal energy white point E from Figure 3.18, plotted here for reference) using a more precise chromaticity diagram. Saturation (a perceptual attribute) of any color patch (transparent or reflection) can be estimated on this chart by a physical measure called excitation purity. It can be seen as the relative distance from the given illumination of the patch to the horseshoe limit curve along a vector. The dominant wavelength (approximate correlate with perceptual attribute of hue) is given by the intersection of that vector with the spectral locus. The lightness of the color is a third dimension, not shown, but is on an axis perpendicular to the plane of the diagram (coming out of the page). Use of dominant wavelength and purity to describe colors in the x, y version of the chromaticity diagram is shown in Figure 3.46 in Section 3.9. Different light sources may be used but standard source “C” (See Figure 3.47) was chosen here. These correlates are only approximate because lines of constant hue are slightly curved in these spaces.

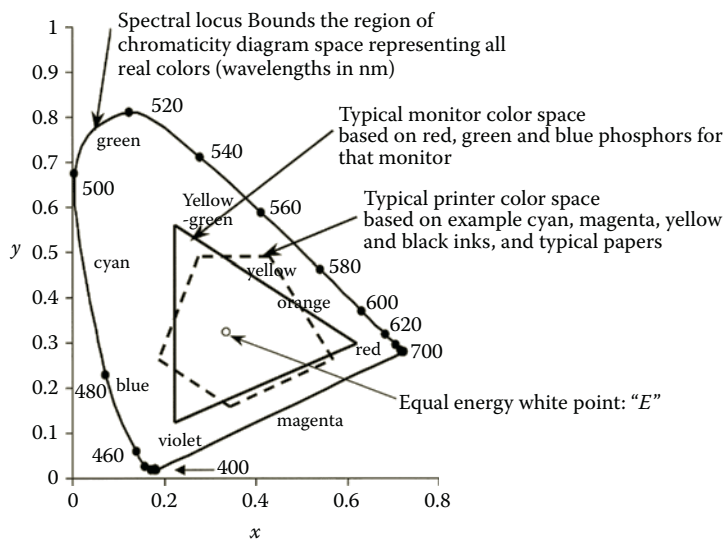


FIGURE 3.18

The x, y chromaticity diagram. Variations derived from it, and the equations that define it, provide a basis for much of the literature that describes color image quality today. It is designed to facilitate description of small color differences such as an original and a reproduced color or two different reproductions of the same color. Examples of the differences between possible colors at a given lightness formed in two different media, a printer and monitor, are shown (adapted from Adams, R.M.; Weisberg, J.B. *The GATF Practical Guide to Color Management*; GATF Press, Graphic Arts Technical Foundation: Pittsburgh, PA, 2000, which cites data from X-Rite Inc). A more precise chromaticity diagram is shown in Figure 3.46.

To understand the chromaticity coordinates, x and y , return to Figure 3.17b. From these curves for $\bar{x}, \bar{y}, \bar{z}$, the spectral power of the light source $S(\lambda)$, and the spectral reflectance (or transmittance) of the object $R(\lambda)$, one can calculate

$$X = k \sum_{\lambda=380}^{780} S(\lambda)R(\lambda)\bar{x}(\lambda) \quad (3.3a)$$

$$Y = k \sum_{\lambda=380}^{780} S(\lambda)R(\lambda)\bar{y}(\lambda) \quad (3.3b)$$

$$Z = k \sum_{\lambda=380}^{780} S(\lambda)R(\lambda)\bar{z}(\lambda) \quad (3.3c)$$

where k is normally selected to make $Y = 100$ when the object is a perfect white, that is, an ideal, nonfluorescent isotropic diffuser with a reflectance equal to unity throughout the visible spectrum. The spectral profile of several standard sources is given later in Figure 3.47.

These results are used to calculate the *chromaticity coordinates* in the above diagram as follows:

$$x = \frac{X}{(X + Y + Z)} \quad (3.4a)$$

$$y = \frac{Y}{(X + Y + Z)} \quad (3.4b)$$

$$z = \frac{Z}{(X + Y + Z)} \quad (3.4c)$$

One of the most popular transformations is the CIE $L^*a^*b^*$ version (called CIELAB for short) which is one of most widely accepted attempts to make distances in color space more uniform in a visual sensation sense^{76,203}. Here

$$L^* = 116(Y/Y_n)^{1/3} - 16 \quad (3.5)$$

which represents the achromatic lightness variable, and

$$a^* = 500 [(X/X_n)^{1/3} - (Y/Y_n)^{1/3}] \quad (3.6)$$

$$b^* = 500 [(Y/Y_n)^{1/3} - (Z/Z_n)^{1/3}] \quad (3.7)$$

represent the chromatic information, where X_n, Y_n, Z_n are the X, Y, Z tristimulus value of the reference white. Color differences are given as

$$\Delta E_{ab}^* = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2}. \tag{3.8}$$

In practical terms, results where $\Delta E_{ab}^* = 1$ represent approximately one just noticeable visual difference (see Section 3.8). However, the residual nonlinearity of the CIELAB chromaticity diagram, the remarkable adaptability of the human eye to many other visual factors, and the effect of experience require situation-specific experiments. Only such experiments can determine rigorous tolerance limits and specifications. Color appearance models that account for many such dependencies and nonlinearities have been developed.^{4,76} Attempts to standardize the methodology have been developed by CIE TC1-34 as CIECAM97s and proposed CIECAM02. (See Appendix A of Reference 4).

For readers not familiar with conventional graphic arts, printing and photographic analysis, densitometers are widely used to characterize those imaging systems. They measure transmission or reflection density, D ,

$$D_f = \text{Log}_{10}(1/R_f) \tag{3.9}$$

where $R_f = Y/Y_{\text{ref}}$ from the above equations. It is called the reflectance (or transmittance for films and filters) factor and may be expressed as a % or decimal. The subscript “ f ” indicates that there are many factors such as light source optical geometries and filters that need to be specified. The “ref” indicates the measurement of Y for a white reference, one of the many factors. Since many tests and test targets used to evaluate scanners and digital cameras are derived from these disciplines, it is useful to examine the relationship between density, L^* , and reflectance as in Figure 3.19. It is seen that to a first approximation (within $.05L^*$) a

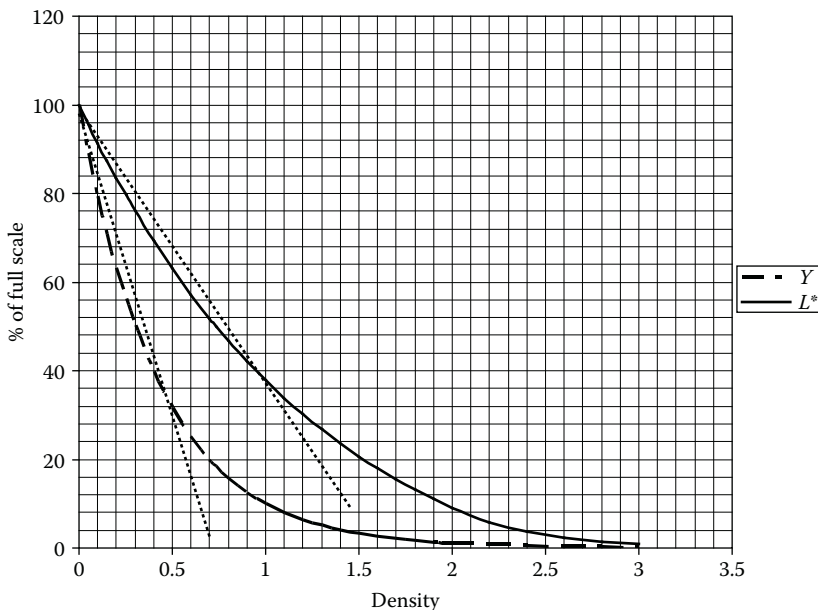


FIGURE 3.19 Reflectance (in the form of Y/Y_n as a %) and L^* as a function of density (See Table 3.8 in the appendix for values of all three) Dotted lines show linear approximations visually fit to each curve, anchored at the maximum value.

straight line approximates the density– L^* relationship to a density of about 1.2 which is a useful range for many image quality measurements, representing the eye's response better than the reflectance factor which is only approximately linear to density of 0.5. An abridged conversion table listing the corresponding values is given in the appendix as Table 3.8.

Another very important color description tool is the Munsell system in which painted paper chips of different colors have been arranged in a three-dimensional cylindrical coordinate system. The vertical axis represents *value* (akin to lightness), the radius represents *chroma*, and the angular position around the perimeter is called *hue*. These have been carefully standardized and are very popular as color references.⁷⁶

3.3 PRACTICAL CONSIDERATIONS

Several overall systems design issues are of some practical concern, including the choice of scan frequency as well as motion errors and other nonuniformities. They will be addressed here in fairly general terms.

3.3.1 Scan Frequency Effects

As digital imaging evolved in the previous decade, it had generally been thought that the spatial frequency, in raster lines or pixels per inch, which is used either to create the output print or to capture the input document, is a major determinant of image quality. Today there is a huge range of scan frequencies emanating from a huge range of products and applications from low-end digital cameras and fax machines, through office scanners and copiers, to high-end graphic arts scanners, all used with a plethora of software and hardware image processing systems that enlarge and reduce and interpolate the originally captured pixel spacings to something else. Then, other systems with yet additional processing and imaging affects are employed to render the image prior to the human reacting to the quality. It is only at this point in the process, where all the signal and noise effects roll up that the underlying principles from other parts of this chapter can be used to quantify overall image quality. Needless to say, scan frequency or pixel density is only one of these effects, and to assert it is *the* dominant effect is questionable in all but the most restrictive of circumstances. Yet it is an important factor and many type A shortcut experiments have attempted to address the connection between the technology variable of pixel density and various dimensions of overall image quality.

If a scanned imaging system is designed so that the input scanning is not aliased and the output reconstruction faithfully prints all of the information presented to it, then the scan frequency tends to determine the blur, which largely controls the overall image quality in the system. This is frequently not the case, and, as a result, scan frequency is not a unique determinant of image quality. In general, however, real systems have a spread function or blur that is roughly equivalent to the sample spacing, meaning they are somewhat aliased and that blur correlates with spacing. However, it is possible to have a large spot and much smaller spaces (i.e., unaliased), or vice versa (very aliased). The careful optimization of the other factors at a given scan frequency may have a great deal more influence on the information capacity of any electronic imaging system and therefore on the image-quality performance than does scan frequency itself. To a certain extent, gray information can be readily exchanged for scan frequency. We shall subsequently explore this further when dealing with the subject of information content of an imaging system.

In the spirit of taking a snapshot of this huge and complex subject, Figure 3.20 summarizes three types of practical findings, two about major applications of scanned or digital images, namely digital photography and graphic arts—digital reprographics, and one simplification of human perception. The curves in the lower graph (solid dots) show results of two customer acceptability experiments with digital photography, varying camera resolution and printing on 8 bpp contone printers (A1 from Reference 77, A2 from Reference 78). Experiments on digital reprographics are shown by the curves with the open symbols, which suggest acceptable enlargement factors for input documents scanned at various resolutions and printed at various output screen resolutions. Finally we can put this in perspective by noting, as triangles along the frequency axis, the resolution limitations of the HVS at normal and close inspection viewing distances using modest 6% and very sensitive 1% contrast detection thresholds. Returning to Figure 3.3, both applications are type A methods, while the HVS limits were inferred from visual algorithms.

A fairly general practice is to design aliased systems in order to achieve the least blur for a given scan frequency. Therefore, another major effect of scan frequency concerns the interaction between periodic structures in the input and the scanning frequency of the system that is recording the input information. These two interfere, producing beat patterns

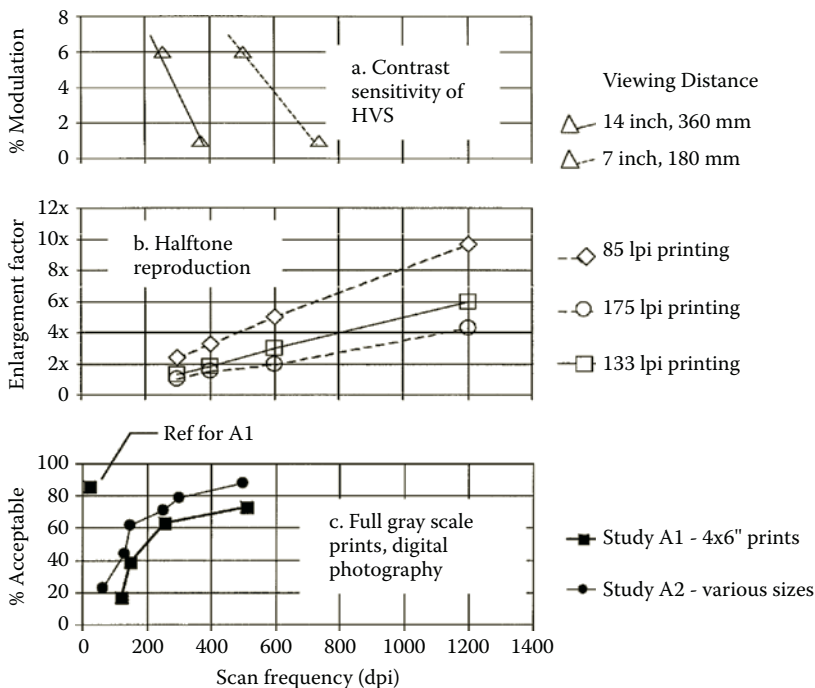


FIGURE 3.20

Summary of practical findings about sampling frequency in graphic arts (middle chart) for halftones (adapted from Cost, F. *Pocket Guide to Digital Printing*; Delmar Publishers: Albany, NY, 1997) in digital photography (bottom chart, adapted from Miller, M.; Segur, R. Perceived IQ and acceptability of photographic prints originating from different resolution digital capture devices. *Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Savannah, GA, 1999*; 131–137 and Daniels, C.M., Ptucha, R.W., Schaefer, L. The necessary resolution to zoom and crop hardcopy images, *Proceedings of IS & T Image Processing, Image Quality, Image Capture Systems (PICS) conference, Savannah, Georgia, 143, 1999*) and related HVS contrast sensitivity (top chart) reference values. (Adapted from Fairchild, M.D. *Color Appearance Models*; Addison-Wesley: Reading, MA, 1998.)

at sum and difference frequencies leading to the general subject of moiré phenomena. Hence, small changes in scan frequency can have a large effect on moiré.

One of the major considerations in selecting output scan frequency is the number of gray levels required from a given halftone screen. Recall the discussion of Figure 3.15 dot matrices from 4×4 to 12×12 are shown in Table 3.1 at a range of printing frequencies from 200 to 1200 raster lines per inch (7.87–47.2 raster lines/mm). For example, a 10×10 matrix of thresholds (on top row) can be used to generate a 51 gray level, 45° angle screen (two shifted 5×10 submatrices) as in Figure 3.15 but with different thresholds in each. Its screen frequencies are shown in the ninth column at the eight different printer scan frequencies and varies from 28 to 170 halftone dots/inch. Also indicated in the table between the bold lines is the approximate useful range for the visual system. The range starts at a lower limit of approx. 65 dots/in (2.56 dots/mm) halftone screen, formerly found in newspapers. This results in noticeably coarse halftones and has recently moved into the range of 85 dots/in (3.35 dots/mm) to 110 dots/in (4.33 dots/mm) in modern newspapers. The upper bound represents a

TABLE 3.1

Relationship among Halftone Matrix Size (Given in Pixels), Maximum Possible Number of Gray Levels in the Halftone, and Output Scan Frequency (in Pixels/Inch)

Matrix in pixels	4×4	3×3	4×4	6×6	5×5	8×8	6×6	10×10	12×12
Angle	45°	90°	90°	45°	90°	45°	90°	45°	45°
No of gray levels—type ^a :									
Conventional	9 - A	10 - B	17 - B	19 - A	26 - B	33 - A [§]	37 - B	51 - A	73 - A
Expanded ^b 2x	17 - C	19 - D	33 - D	37 - C	51 - D	65 - C	73 - D	101 - C	145 - C
Expanded ^b 4x	33	41 - E	65 - E	73	101 - E	129	145 - E	201	289
Scan freq. in pixels/in ↓									
1200	426	400	300	282	240	212	200	170	141
1000	352	333	250	236	200	176	167	142	118
800	284	267	200	188	160	142	133	114	94
600	212	200	150	141	120	106	100	85	71
500	176	166	125	118	100	88	84	71	59
400	142	133	100	94	80	71	67	57	47
300	106	100	75	71	60	53	50	43	36
200	71	67	50	47	40	35	33	28	24

Practical upper limit (~200 LPI)

Typical halftone (~100 LPI)

Halftone dot, practice lower limit (~65 LPI)

Entries are given in halftone dots/inch measured along the primary angle (row 2) of the halftone pattern. Dot types are given as (See quadrants of Figure 3.15): (A) the conventional 45° halftone where quadrants $Q1 = Q4$, $Q2 = Q3$; (B) conventional 90° halftone where $Q1 = Q2 = Q3 = Q4$. Expansions of the number of halftone gray levels show three new types: (C) = Type A except $Q3$ and $Q4$ thresholds are set at halfway between those in $Q1$ and $Q2$ (45° double dot), (D) where $Q1 = Q4$, but $Q2$ and $Q3$ thresholds are set halfway between those in $Q1$ (90° double dot); (E) where $Q1$ through $Q4$ thresholds are each set to generate intermediate levels among each other (90° quad dot). The number of gray values includes one level for white.

^a Type refers to specific halftone structures A–E (see caption) where appropriate.

^b Number of gray levels is increased by 2x or 4x over conventional by gray pixels or multicentered dots.

[§] Example shown in Figure 3.15.

materials limit of around 175 dots/in (6.89 dots/mm), which is a practical limit for many lithographic processes. The number of gray levels is shown in the third, fourth, and fifth rows. Conventional dots are single centered like Figure 3.15. Increased number of gray levels for double dots is shown as the “expanded by 2×” row and quad dots (four centers) by the row labeled as “expanded by 4×” are indicated by the row. This table assumes that the pixels are binary in nature. If a partially gray or high addressability output imaging system is employed then the number of levels in the table must be multiplied by the number of gray levels or subpixels per pixel appropriate to the technology. The use of these techniques and supercells to expand the gray level resolution has increased in recent years as real time microprocessing has enabled reasonable speed and memory for such approaches.

3.3.2 Placement Errors or Motion Defects

Since the basic mode of operation for most scanning systems is to move or scan rapidly in one direction and slowly in the other, there is always the possibility of an error in motion or other effect that results in locating pixels in places other than those intended. Figure 3.21 shows several examples of periodic raster separation errors, including both a sinusoidal and a sawtooth distribution of the error. These are illustrated at 300 raster lines/in (11.8 lines/mm) with ± 10 through ± 40 μm (± 0.4 through ± 1.6 mils) of spacing error, which refers to the local raster line spacing and not to the error in absolute placement accuracy. Error frequencies of 0.33 cycles/mm (8.4 cycles/in) and 0.1 cycles/mm (2.5 cycles/in) are illustrated.

For input scanners, which convert an analog signal to a digital one, the error takes the form of a change in the sampling of the analog document. Since sampling makes many mistakes,

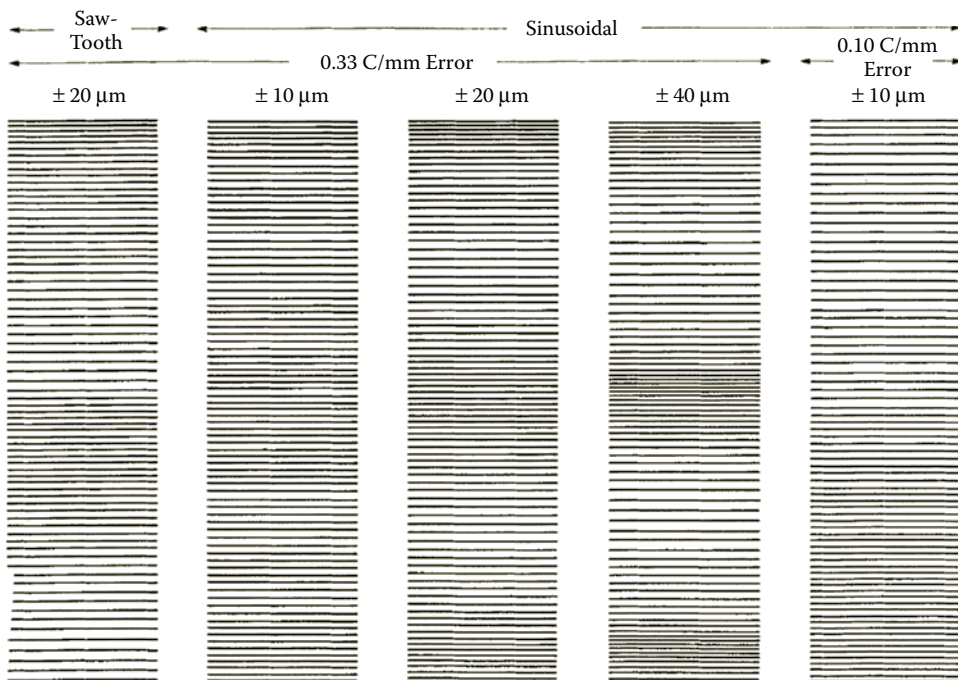


FIGURE 3.21
Enlarged examples of rasters with specified image motion variation at 300 dpi.

the sampling errors due to motion nonuniformity are most visible in situations where the intrinsic sampling error is made to appear repeatable or uniform, and the motion error, therefore, appears as an irregular change to an otherwise uniform pattern. Long angled lines that are parallel to each other provide such a condition because each line has a regular periodic phase error associated with it, and a motion error would appear as a change to this regular pattern. Halftones that produce moiré are another example, except that the moiré pattern is itself usually objectionable so that a change in it is not often significant.

In patterns with random phase errors such as text, the detection of motion errors is more difficult. Effects that are large enough to cause a two-pixel error would be perceived very easily; however, effects that produce less than one-pixel error on average would tend to increase the phase errors and noise in the image generally and would therefore be perceived on a statistical basis. Many identical patterns repeated throughout a document would provide the opportunity to see the smaller errors as being correlated along the length of the given raster line that has been erroneously displaced, and would therefore increase the probability of seeing the small errors.

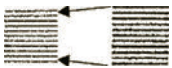







Motion errors in an output scanner that writes on some form of image-recording material can produce several kinds of defects. In Table 3.2, several attributes of the different types of raster distortion observables are shown. The first row in the matrix describes the general kind of error, that is, whether it is predominantly a pixel placement error or predominantly a developable exposure effect or some combination of the two. The second row is a brief word description or name of the effect that appears on the print. The third row describes the spatial frequency region in cycles/mm in which this type of error tends to occur. The next row indicates whether the effect is best described and modeled as one-dimensional or two-dimensional. Finally, a graphical representation of an image with the specific defect is shown in the top row, while the same image appears in the bottom row without the defect.

The first of the columns on the left is meant to show that if the frequency of the error is low enough then the effect is to change the local magnification. A pattern or some form of texture that should appear to have uniform spacings would appear to have nonuniform spacings and possibly the magnification of one part of the image would be different from that of another. The second column is the same type of effect except the frequency is much higher, being around 1 cycle/mm (25 cycle/in). This effect can then change the shape of a character, particularly one with angled lines in it, as demonstrated by the letter Y.

Moving to the three right-hand columns, which are labeled as developable exposure effects, we have three distinctly different frequency bands. The nature and severity of these effects depend in part on whether we are using a "write white" or "write black" recording system and on the contrast or gradient of the recording material. The first of these effects is labeled as structured background. When the separation between raster lines increases and decreases, the exposure in the region between the raster lines where the Gaussian profile writing beams overlap increases or decreases with the change. This gives an overall increase or decrease in exposure, with an extra large increase or decrease in the overlap region. Since many documents that are being created with a laser beam scanner have relatively uniform areas, this change in exposure in local areas gives rise to nonuniformities in the appearance in the output image.

In laser printers, for example, the text is generally presented against a uniform white background. In a positive "write white" electrophotographic process, such as is used in many large xerographic printers, this background is ideally composed of a distribution of uniformly spaced raster lines that expose the photoreceptor so that it discharges to a level where it is no longer developable. As the spacing between the raster lines increases, the exposure between them decreases to a point where it no longer adequately discharges the

TABLE 3.2
The Effects of Motion Irregularities, Defects, or Errors on the Appearance of Scanned Images

Error Category	Pixel Placement Error		Combination of Both		Developable Exposure Effects			
	Spacing Nonuniformity	Character Distortion/Fast Scan Jitter	Halftone Nonuniformity	Line Darkness Nonuniformity	Structured Background	Ragged/Structured Edges	Unsharp Line-text-edge Images	
Typical frequencies of motion error (c/mm)	<0.5	0.5-2	0.1-6	0.005-2	0.005-8	1-8	4-20+	
Dimensions of effect	1-D	2-D	2-D (1-D)	1-D (2-D)	1-D	2-D	2-D	
Scanner type	input/output	input/output	input/output	input/output	output	output	output	
Example with little or no defect needs space in "orno"								
Comment	Shows local reduction, magnification is also possible		Illustrates 2 levels, bottom case has very small defect		Shows "write white" system	Shows strong low & high freq.—approx 3 x 3 mm sample	Examples: higher magnification than at left eye blurs structures shown	

photoreceptor, thereby enabling some weak development fields to attract toner and produce faint lines on a page of output copy. For this reason among others, some laser printers use a reversal or negative "write black" form of electrophotography in which black (no light) output results in a white image. Therefore, white background does not show any variation due to exposure defects, but solid dark patches often do.

The allowable amplitude for these exposure variations can be derived from minimum visually perceivable modulation values and the gradient of the image-recording process.^{79,80} In the spatial frequency region near 0.5 cycles/mm (13 cycles/in), where the eye has its peak response at normal viewing distance, an exposure modulation of 0.004–0.001 $\Delta E/E$ has been shown to be a reasonable goal for a color photographic system with tonal reproduction density gradients of 1–4.⁸¹

If the frequency of the perturbation is of the order 1–8 cycles/mm (25–200 cycles/in), and especially if the edges of the characters are slightly blurred, it is possible for the nonuniform raster pattern to change the exposure in the partially exposed blurred region around the characters. As a result, nonuniform development appears on the edge and the raggedness increases as shown by the jagged appearance of the wavy lines in column 7. The effects are noticeable because of the excursions produced by the changes in exposure from the separated raster lines at the edges of even a single isolated character. The effect is all the more noticeable in this case because the darkened raster lines growing from each side of the white space finally merge in a few places. The illustration here, of course, is a highly magnified version of just a few dozen raster lines and the image contained within them.

In the last column we see small high-frequency perturbations on the edge, which would make the edge appear less sharp. Notice that structured background is largely a one-dimensional problem, just dealing with the separation of the raster lines, while character distortion, ragged or structured edges, and unsharp images are two-dimensional effects showing up dramatically on angled lines and fine detail. In many cases the latter require two dimensions to describe the size of the effect and its visual appearance.

Visually apparent darkness for lines in alphanumeric character printing can be approximately described as the product of the maximum density of the lines in the character times their widths. It is a well-known fact in many high-contrast imaging situations that exposure changes lead to line width changes. If the separation between two raster lines is increased, the average exposure in that region decreases and the overall density in a write white system increases. Thus, two main effects operate to change the line darkness. First, the raster information carrying the description of the width of the line separates, writing an actually wider pattern. Secondly, the exposure level decreases, causing a further growth in the line width and to some extent causing greater development, that is, more density. The inverse is true in regions where the raster lines become closer together. Exposure increases and linewidth decreases.

If these effects occur between different strokes within a character or between nearby characters, the overall effect is a change in the local darkness of text. The eye is generally very sensitive to differences of line darkness within a few characters of each other and even within several inches of each other. This means that the spatial frequency range over which this combination of stretching and exposure effect can create visual differences is very large, hence the range of 0.005–2 cycles/mm (0.127–50 cycles/in). Frequencies listed in Table 3.2 cover a wide range of effects, also including some variation of viewing distance. They are not intended as hard boundaries but rather to indicate approximate ranges.

Halftone nonuniformity follows from the same general description given for line darkness nonuniformity except that we are now dealing with dots. The basic effect, however, must occur in such a way as to affect the overall appearance of darkness of the small region of an otherwise uniform image. A halftone works on the principle of changing a

certain fractional area coverage of the halftone cell. If the spatial frequency range of this nonuniformity is sufficiently low, then the cell size changes at the same rate that the width of the dark dot within the cell changes. Therefore the overall effect is to have no change in the percent area coverage and only a very small change in the spacing between the dots. Hence, the region of a few tenths to several cycles/mm (several to tens of cycles/in) is the domain for this artifact. It appears as stripes in the halftone image.

The allowable levels for the effects of pixel placement errors on spacing nonuniformity and character distortion depend to a large extent upon the application. In addition to application sensitivity, the effects that are developable or partially developable are highly dependent upon the shape of the profile of the writing spot and upon amplification or attenuation in the marking system that is responding to the effects. Marking systems also tend to blur out the effects and add noise, masking them to a certain extent.

3.3.3 Other Nonuniformities

There are several other important sources of nonuniformity in a raster scanning system. First, there is a pixel-to-pixel or raster-to-raster line nonuniformity of either response in the case of an input scanner or output exposure in the case of an output scanner. These generally appear as streaks in an image when the recording or display medium is sensitive to exposure variations. These, for example, would be light or dark streaks in a printed halftone or darker and lighter streaks in a gray recorded image from an input scanner looking at a uniform area of an input document. A common example of this problem in a rotating polygon output scanner is the effect of facet-to-facet reflectivity variations in the polygonal mirror itself. The exposure tolerances described for motion errors above also apply here.

Another form of nonuniformity is sometimes referred to as *jitter* and occurs when the raster synchronization from one raster line to another tends to fail. In these cases a line drawn parallel to the slow scan direction appears to oscillate or jump in the direction of the fast scan. These effects, if large, are extremely objectionable. They will manifest themselves as raggedness effects or as unusual structural effects in the image, depending upon the document, the application, and the magnitude and spatial frequency of the effect.

3.3.3.1 Perception of Periodic Nonuniformities in Color Separation Images

Research on the visibility of periodic variations in the lightness of 30% halftone tints of cyan, magenta, yellow, and black color image separations printed on paper substrates has been translated into a series of guidelines for a specification for a high-quality color print engine.⁸² (Figure 3.22). They were chosen to be slightly above the onset of visibility. Specifically they are set at $\{[1/3] \times [(2 \times \text{“visible but subtle threshold”}) + (\text{“obvious threshold”})]\}$ and adjusted for a wider range of viewing distances and angles than during the experiments, which were at 38–45 cm. These guidelines are given in terms of colorimetric lightness units on the output prints. Visibility specifications must ultimately be translated into engineering parameters. We have selected the traditional CIE $L^*a^*b^*$ metrics version for illustration. These also tend to show the smallest, most demanding ΔE s. Guidelines developed in terms of ΔE for other color difference metrics (CMC 2:1 and CIE-94) have also been developed,⁸² and show different visual magnitudes, by as much as a factor of 2.

To translate these into a guidelines for the approximate optical scanner exposure variation, the ΔE values in this graph must be divided by the slope of the system response curve, in terms of $\Delta E/\Delta \text{exposure}$, for the color separation of interest. Exposure, H , is the general variable of interest since it is the integrated effect of intensity and time variations, both

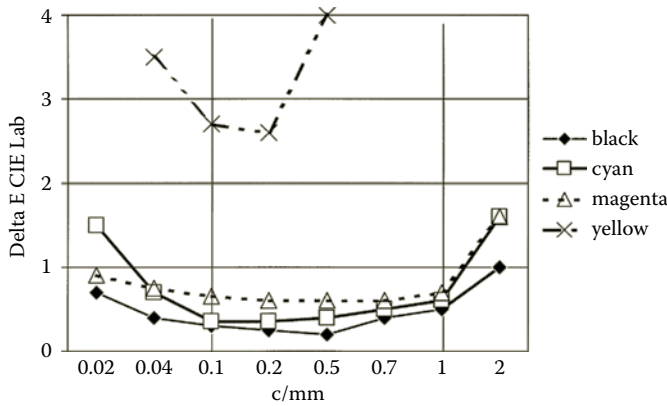


FIGURE 3.22

Guidelines for specification of periodic nonuniformities in black, cyan, magenta, and yellow color separations as indicated, in terms of ΔE derived from CIE $L^*a^*b^*$, plotted against the effective spatial frequency of the periodic disturbance. (Adapted from Goodman, N.B. Perception of spatial color variation caused by mass variations about single separations. *Proceedings of IS&T's NIP14: International Conference on Digital Printing Technologies, Toronto, Ontario, Canada, 1998; 556-559.*)

of which can result from the scanner errors discussed in the pages above. The system response would be approximated by the cascaded (multiplied) slopes of the responses of all the intermediate imaging systems between the scanner and the resulting imaging media assuming the small signal theory approximation to linearity of the cascaded systems. In the particular case where a system is positive working and linear such that, $\Delta R = \Delta H$, then taking the derivative of Equation 3.5 gives

$$\Delta E = \Delta L^*. \tag{3.10}$$

For $Y/Y_n = R = 0.70$. R and H are decimal output reflectance and normalized exposure values respectively (i.e., both full scale of 1 and a minimum of zero) R = 0.70 is the reflectance for a 30% halftone used above. For $\Delta E = 0.2$ this yields $\Delta R = 0.0041$ which also = ΔH , indeed a very small exposure value yielding visible errors.

Specific relationships for exposure and reflectance, and for any of the other colorimetric units described in this research should be developed for each real system. The linear gain = 1.0 assumption shown here should not be taken for granted. The reader is also reminded that these results are for purely sinusoidal errors of a single frequency and a single color and that actual nonuniformities occur in many complex spatial and color forms.

3.4 CHARACTERIZATION OF INPUT SCANNERS THAT GENERATE MULTILEVEL GRAY SIGNALS (INCLUDING DIGITAL CAMERAS)

In this section we will discuss the elementary theory of performance measurements and various algorithms or metrics to characterize them, the scanner factors that govern each, some practical considerations in the measurements, and visual effects where possible. Generally speaking, this is the subject of analyzing and evaluating systems that acquire

sampled images. Originally this was explored as analog sampled images in television, most notably by Schade^{83,84} in military and in early display technology.⁸⁵ As computer and digital electronics technology grew, this evolved into the general subject of evaluating digital sampled image acquisition systems, which include various cameras and input scanners. Modern scanners and cameras are different only in that a scanner moves the imaging element to create sampling in one direction while the camera imaging element is static, electronically sampling a two-dimensional array sensor in both dimensions. This topic can be divided into two areas. The first concerns scanners and cameras that generate output signals with a large number of levels (e.g., 256), where general imaging science using linear analysis applies.¹³ The second deals with those systems that generate binary output, where the signal is either on or off (i.e., is extremely nonlinear) and more specialized methods apply.⁸⁶ These are discussed in Section 3.5.

In recent years, the advent of digital cameras and the plethora of office, home, and professional scanners have promoted wide interest in the subject of characterizing devices and systems that produce digital images. Also, several commercially available image analysis packages have been developed for general image analysis, many using scanners or digital cameras, often attached to microscopes or other optical image magnification systems. Components of these packages and the associated technical literature specifically address scanner analysis or calibration.⁸⁷⁻⁸⁹ A variety of standards activities have evolved in this area.⁹⁰⁻⁹³ Additional related information is suggested by the literature on evaluating microdensitometers.^{94,95} These systems are a special form of scanners in which the sensor has a single aperture of variable shape. Much of this work relates to transmitted light scanners but reflection systems have also been studied.⁹⁶ Methods for evaluating digital cameras and commercially available scanners for specific applications have been described by many authors.^{77,93,97}

3.4.1 Tone Reproduction and Large Area Systems Response

Unlike many other imaging systems, where logarithmic response (e.g., optical density) is commonly used, the tonal rendition characteristics of input scanners are most often described by the relationship between the output signal (gray) level and the input reflectance or brightness. This is because most electronic imaging systems respond linearly to intensity and therefore to reflectance. Three such relationships are shown in Figure 3.23. In general these curves can be described by two parameters, the offset, O , against the output gray level axis and the gain of the system Γ , which is defined in the equation in Figure 3.23. Here g is the output gray level, and R is the relative reflectance factor. If there is any offset, then the system is not truly linear despite the fact that the relationship between reflectance and gray level may follow a straight-line relationship. This line must go through the origin to make the system linear.

Often the maximum reflectance of a document will be far less than the 1.0 (100%) shown here. Furthermore, the lowest signal may be significantly higher than 1% or 2% and may frequently reach as much as 10% reflectance. In order to have the maximum number of gray levels available for each image, some scanners offer an option of performing a histogram analysis of the reflectances of the input document on a pixel-by-pixel or less frequently sampled basis. The distribution is then examined to find its upper and lower limits. Some appropriate safety factor is provided, and new offset and gain factors are computed. These are applied to stretch out the response to cover as many of the total (256 here) output levels as possible with the information contained between the maximum and minimum reflectances of the document.

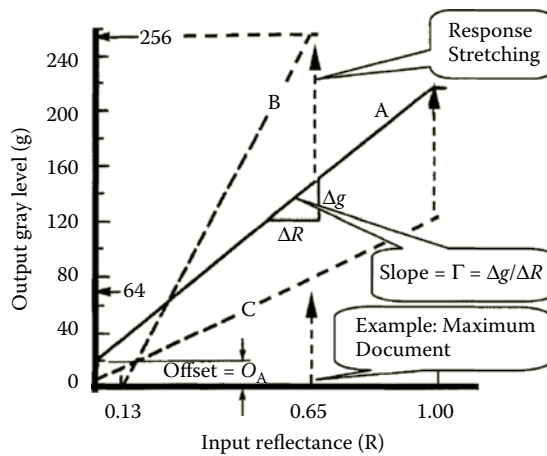


FIGURE 3.23

Typical types of scanner input responses, illustrating the definitions of “gain” (i.e., slope), “offset,” and “response stretching.”

Other scanners may have a full gray-scale capability from 4 to 12 bits (16–4096 levels). In Figure 3.23, curve C is linear, that is, no offset and a straight-line response up to a reflectance of 1.0 (100%), in this case yielding 128 gray levels. Curve A would represent a more typical general purpose gray response for a scanner while curve B represents a curve adjusted to handle a specific input document whose minimum reflectance was 0.13 and whose maximum reflectance was 0.65. Observe that neither of these curves is linear. This becomes very important for the subsequent forms of analysis in which the nonlinear response must be linearized before the other measurement methods can be applied properly. This is accomplished by converting the output units back to input units via the response function.

In a digital scanner the sensors themselves are fairly linear as can be seen in Figure 3.24 which plots exposure in linear units (lux-s) versus output in millivolts (mV). The response is strictly linear from 0 to 2.2 lux-s and then begins to roll over as it saturates. Notice the difference between the “linear saturation exposure” and the “saturation exposure” which is a graphical construct projecting the linear part of the curve to the maximum signal. It is often observed that digital sensors are linear but it can be seen from Figure 3.24 that this is only true for most but not all of the response curve. The scanner or camera designer is free to use as much or as little of the nonlinear high end of the curve as he desires. For digital cameras the indicated standard exposure differs by camera specifications but is usually in the linear region

It is also possible to arrange the electronics in the video processing circuit so that equal steps in exposure do not generate equal steps in electronic or digital response, but rather are appropriately spaced steps in some units that are more significant, either visually or in terms of materials properties. A logarithmic A/D converter is sometimes used to create a signal proportional to the logarithm of the reflectance or to the logarithm of the reciprocal reflectance (which is the same as “density”). Some scanners for graphic arts applications function in this manner. Another common conversion is making the signal proportional to L^* . Both of these require a larger number of levels to start with than what is output. These systems are highly nonlinear, but may work well with a limited number of gray levels, for example with 8 bits (256 levels) rather than the 10 or 12 bits as discussed earlier.

Many input scanners operate with a built-in calibration system that functions on a pixel-by-pixel basis. In such a system, for example, a particular sensor element that has greater

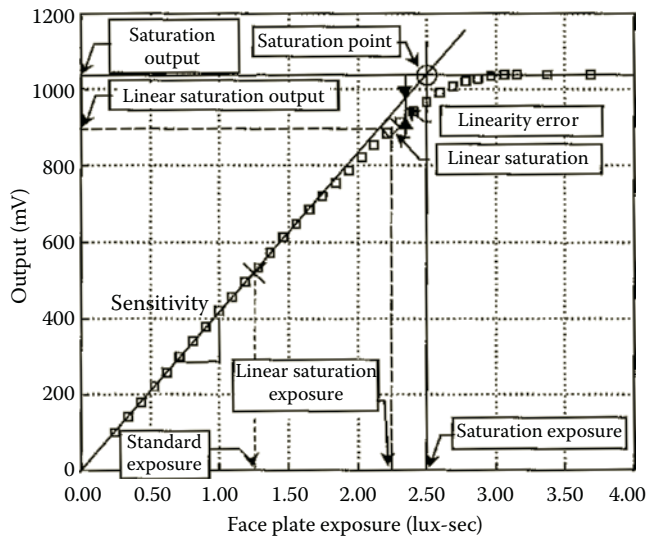


FIGURE 3.24

Fundamental electronic response to light of a sensor used in scanners and cameras showing the linear and nonlinear regions. (Reproduced with permission of the publisher from Nakamura, *J. Image Sensors and Signal Processing for Digital Still Cameras*; Taylor and Francis: Boca Raton, FL, 2006; Mizoguchi, T. Ch 6: Evaluation of image sensors, 179–203.

responsivity than others may be attenuated or amplified by adjusting either the gain or the offset of the system or both. This would ensure that all photosites (individual sensor elements) respond equally to some particular calibrated input, often, as is common with most light measuring devices such as photometers and densitometers, using both a light and dark reflectance reference (e.g., a white and black strip of paint).

It is possible in many systems for the sensor to be significantly lower or higher in responsivity in one place than another. As an example, a maximum responsivity sensor may perform as shown in curve A while a less sensitive photosite may have the response shown in curve C. If curve C was captured with the same A/D converter at the same settings (as is often the case in high-speed integrated circuits), the maximum signal range it contains has only 120 gray levels. A digital multiplier can operate upon this to effectively double each gray level, thereby increasing the magnitude of the scale to 220 or 240, depending upon how it handles the offset. Note that if some of the elements of a one-dimensional sensor responded as curve C, others as A, with the rest in between, then this system would exhibit a kind of one-dimensional granularity or nonuniformity, whose pattern depends upon the frequency of occurrence of each sensor type. This introduces a quantization error varying spatially in one-pixel-wide strips, and ranging, for this example, from strips with only 120 steps to others with 240 steps, yet covering the same distribution of output tones.

An ideal method for measuring tone reproduction is to scan an original whose reflectance varies smoothly and continuously from near 0% to near 100%, or at least to the lightest “white” that one expects the system to encounter. The reflectance is evaluated as a function of position, and the gray value from the scanner is measured at every position where it changes. Then the output of the system can be paired with the input reflectance at every location and a map drawn to relate each gray response value to its associated input reflectance. A curve like Figure 3.23 can then be drawn for each photosite and for various statistical distributions across many photosites.

The classic concepts of quality in tone reproduction generally extend to processes and devices beyond the capture device. Hence the idea of quality for a scanner involves how well it integrates into an overall system that would include a printer or display. This integration is facilitated by image processing, both hardwired in the scanner and through off-line software systems. The graphical construction of a multi quadrant “Jones Plot” has often been used in photography to characterize how a film integrates with camera/optics, film processing, an enlarger and printing paper and even the visual system.^{60,61} Similar systems plots can be constructed for the digital system starting either with the camera or the scanner. One such example, using representative system data is shown in Figure 3.25.

Starting at the axis labeled “original density” one creates four quadrants in a clockwise progression starting with Quadrant 1 (lower right) as a plot of digital output level (DOL) versus input Density (or equivalent Log Exposure) for the scanner or camera in question. This is a type of OECF (Optoelectronic Conversion Function) Curve.⁶⁵ In this illustration Density of the original target is plotted increasing to the left (Log exposure would increase to the right) and DOL (some call this value digital count or gray value) increases toward

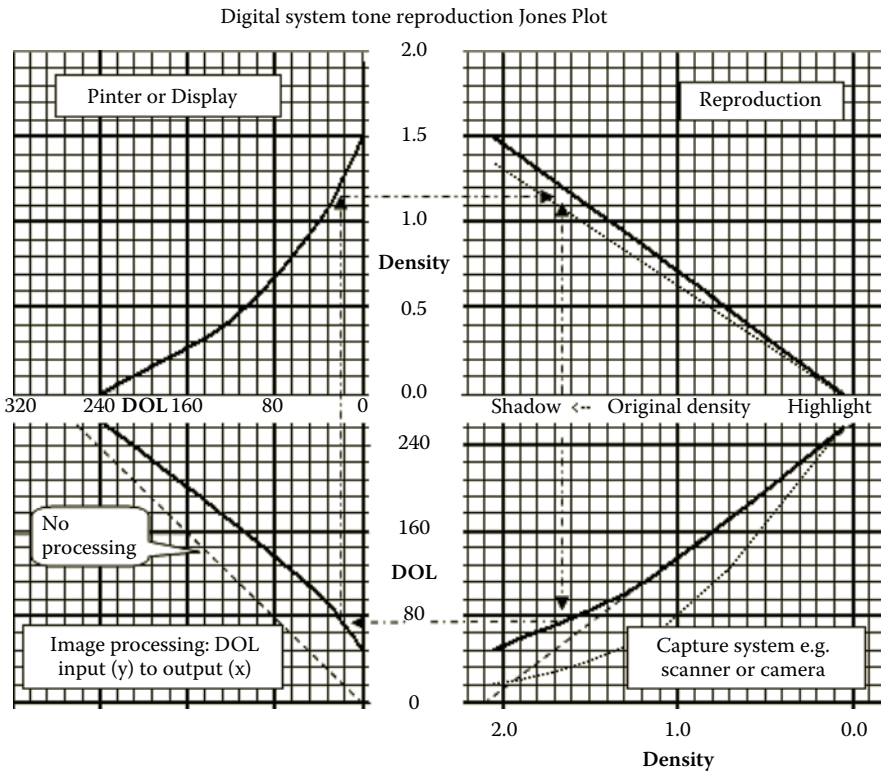


FIGURE 3.25

Jones plot for tonal response in a digital imaging system showing the cascade of components using four quadrants, Q1–Q4: Q1 (lower right) is for digital capture system (scanner or camera) showing density of the original test object (x-axis) mapped to digital output level (i.e., DOLs on y-axis), Q2 (lower left) is the image processing which maps the same DOLs on y-axis to image processing digital output levels (DOLs on x-axis) The latter are also digital input levels in Q3 (upper left) for either a printer or display. Here printer digital input levels map to output printed density (y-axis). In Q4 (upper right) the resulting solid curve (follow the dashed arrow through all four quadrants to see the cascading) gives printed density (y-axis) compared to original test object’s density (x-axis). This is the scanning (or photographic) system’s overall tone reproduction. See text for dotted/dashed curves.

the top. A dashed line indicates a linear response that follows the actual curve down to the dark region where it begins to “tail up”, due largely to flare light. The fact that the log values of density in the bolder solid curve agree so well with the linear values of electronic output (DOLs) suggests that the on-board image processing in this scanner is creating a nonlinear response (for the linear sensors as noted above) in order to better fit some output needs of printing or viewing. This would be typical of some digital cameras as well as some scanners where off-line image processing was expected. The lighter dotted line represents the output of a typical scanner integrated with the printer shown in Quadrant 3, a so-called all-in-one system or a digital copier.

In many such evaluations two of the other three quadrants are specified and the goal is to derive the missing curve. Consider that the rendering device (Quadrant 3 clockwise) is a printer with a fixed density response to a given array of input DOLs. Assume that the it is desired that reproduction (Quadrant 4) be a linear relationship between density of the original and that of the print, even though the maximum densities do not match. This leaves the image processing (Quadrant 2) to be determined. A linear, one for one, image processing between input from the first scanner and output DOLs (dashed curve) would result in a very light print with a somewhat curved density reproduction relationship. The solid curve in Quadrant 2 (Image Processing) results in the desired linear density relationship in Q4.

The second scanner curve (dotted) is less linear but includes on-board image processing which predistorts the output to compensate for the highly curved printer density response curve. This scanner response directly provides another linear final tone reproduction in Quadrant 4, although with slightly lower maximum density. In the Jones Plot this result uses the dashed “no image processing” curve in Quadrant 2 since off-line image processing is not possible in an all-in-one (copier) system. This scanner curve is the same one used in Figure 3.26.

Most scanners operate with sufficiently small detector sites or sensor areas that they respond to input granularity. Thus, a single pixel or single photosite measurement will not

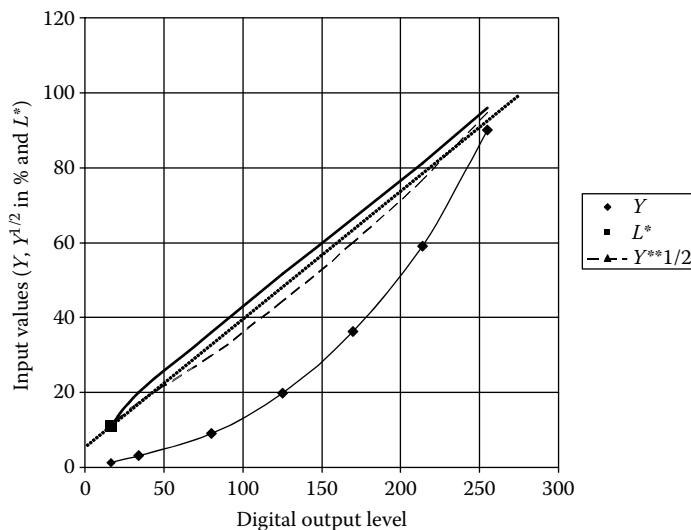


FIGURE 3.26

Scanner output digital levels (x-axis) as predicted by the input test target reflectance values or CIE Y (diamonds), L^* (large squares), $Y^{1/2}$ (small triangles), and a straight dotted line visually fit to the last two. The ordinate is the input value plotted on a relative scale of 0–100. Therefore Y (which \approx reflectance) and $Y^{1/2}$ are given in %.

suffice to get a solid area response to a so-called uniform input. Some degree of averaging across pixels is required, depending upon the granularity and noise levels of the input test document and the electronic system.

The use of a conventional step tablet or a collection of gray patches, where there are several discrete density levels, provides an approximation to this analysis but does not allow the study of every one of the discrete output gray levels. For a typical step tablet with approximately 20 steps 0.10 reflection density, half of the gray values are measured by only 4 steps, 0, 0.1, 0.2 and 0.3 density (or 50% reflectance). Thus a smoothly varying density wedge is more appropriate for the technical evaluation of an electronic input scanner.

However, suitable wedges are difficult to fabricate repeatably and the use of uniform patches of several discrete densities is common in many operations. See, for example, Reference 97 and the ISO standard targets in Figures 3.48b, 3.49 or the IEEE target in Figure 3.50b (see middle of pattern). Nonetheless, wedges are available (see for example the top of Figure 3.50b), and are essential to accurately evaluate binary scanning (see Section 3.5.2).

Returning to the large area tonal response of the scanner itself, it is tempting to describe it as the linear equation for the sensor itself but the fact is that most scanners today have some built-in image processing associated with them and it is more practical to use a curve. To compensate for some common printer and display response, scanner's tone response neglecting flare can often be mapped as

$$\text{DOL} = Hr^{1/\gamma} \quad (3.11)$$

where Hr is the relative exposure from the input and γ (gamma) is a constant designed to compensate for the exponential-shaped curves often found in output printers or displays. Values of 1.8, 2.2 are examples for Mac and PC monitors and 3 to emulate L^* but a general purpose scanner may desire to satisfy all these conditions with some hybrid and a few other terms. Results for a recent desktop scanner are shown in Figure 3.26—an x versus y inverted type of OECF²¹ curve—using the resulting digital output levels as the x-axis and various characterizations of the input as the y-axis to deduce the vendors image processing. The system is not linear in reflectance but is approximately linear in either L^* or $\gamma = 2$. Note $\gamma = 2$ is halfway between the Mac and PC standards.

Setting the maximum point equal to 100% input reflectance is often a waste of gray levels since there are no documents whose real reflectance is 100%. A value somewhere between 70% and 90% would be more representative of the upper end of the range of real documents. Some systems adjust automatically to the input target and are therefore difficult to evaluate. They are highly nonlinear in a way that is difficult to compensate. See Gonzalez and Wintz⁹⁸ for an early discussion of automatic threshold or gray scale adjustment and Hubel⁹³ for more recent comments on this subject as it relates to color image quality in digital cameras. Most amateur and some professional digital cameras fall into this automatic domain⁹³ as do many scanners. A system that finds this point automatically is optimized for each input differently and is therefore difficult to evaluate in a general sense.

An offset in the positive direction can be caused either by an electronic shift or by stray optical energy in the system (as shown in Q1 of Figure 3.25). If the electronic offset has been set equal to zero with all light blocked from the sensor, then any offset measured from an image can be attributed to optical energy. Typical values for flare light, the stray light coming through the lens, would range from just under 1% to 5% or more of full scale.⁹⁶ While offset from uniform stray light can be adjusted out electronically, signals from flare light are document dependent, showing up as errors in a dark region only when it is surrounded by a large field of white on the document. Therefore, correction for this

measured effect in the particular case of an analytical measurement with a gray wedge or a step tablet surrounded by a white field may produce a negative offset for black regions of the document that are surrounded by grays or dark colors. If, however, the source of stray light is from the illumination system, the optical cavity, or some other means that does not involve the document, then electronic correction is more appropriate. Methods for measuring the document-dependent contribution of flare have been suggested in the literature.^{96,97,99} Some involve procedures that vary the surround field from black to white while measuring targets of different widths,⁹⁶ others use white surround with different density patches.⁹⁷

A major point of confusion can occur in the testing of input scanners and many other optical systems that operate with a relatively confined space for the illumination system, document platen, and recording lens. This can be thought of as a type of integrating cavity effect. In this situation, the document itself becomes an integral part of the illumination system, redirecting light back into the lamp, reflectors, and other pieces of that system. The document's contribution to the energy in the illumination depends on its relative reflectance and on optical geometry effects relating to lamp placement, document scattering properties, and lens size and location. In effect the document acts like a position-dependent and nonlinear amplifier affecting the overall response of the system. One is likely to get different results if the size of the step tablet or gray wedge used to measure it changes or if the surround of the step tablet or gray wedge changes between two different measurements. It is best, therefore, to make a variety of measurements to find the range of responses for a given system. These effects can be anywhere from a few percent to perhaps as much as 20%, and the extent of the interacting distances on the document can be anywhere from a few millimeters to a few centimeters (fraction of an inch to somewhat over one inch). Relatively little has been published on this effect because it is so design specific, but it is a recognized practical matter for measurement and performance of input scanners. An electronic correction method exists.^{100,101}

3.4.2 MTF and Related Blur Metrics

We will now return to the subject of blur. Generally speaking, the factors that affect blur for any type of scanner include (Table 3.3): the blur from optical design of the system, motion

TABLE 3.3

Factors Affecting Input Scanner Blur and Pointers to Useful MTF Curves That Describe Selected Cases

Solid-state scanners

- Lens aberrations as functions of wavelength (see Figure 3.53 if diffraction limited, for example, some microscope optics), field position, orientation, focus distance (see $n = 3$ or 4 of Figure 3.54 for useful equation to fit system with various lens performance)
- Sensor: Aperture dimensions (see Figure 3.51), charge transfer efficiency (CCD), charge diffusion, leaks in aperture mask
- Motion of sensor during reading (Figure 3.51)
- Electronics rise time (measured frequency response)

Flying spot laser beam scanner

- Spot shape and size at document (Gaussian case see Figure 3.52)
 - Lens aberrations (as above)
 - Polygon aperture or equivalent
 - Motion during reading (Figure 3.51)
 - Sensor or detector circuit rise time (measured frequency response)
-

of the scanning element during one reading, electronic effects associated with the rise time of the circuit, the effective scanning aperture (sensor photo site) size, and various electro-optical effects in the detection or reading out of the signal. The circuits that handle both the analog and the digital signals, including the A/D converter, may have some restrictive rise times and other frequency response effects that produce a one-dimensional blur.

To explore the analysis of these effects, refer back to Figure 3.5. A primary concept begins with a practical definition of an ideally narrow line object and the image of it. Imagine that the narrow line object profile shown at the top right of Figure 3.5a was steadily reduced in width until the only further change seen in the resulting image Figure 3.5b is that the height of the image peak changes but not the width of its spreading. This is a practical definition of an ideally narrow line source. Under these conditions we would say that the peak of the image on the right of Figure 3.5b was a profile of the *line spread function* for the imaging system. [It is also seen at higher sampling resolution in Figure 3.7b.]

To be completely rigorous about this definition of the line spread function, we would actually use a narrow white line rather than a black line. If the input represented a very fine point in two-dimensional space we would refer to its full two-dimensional image as a *point spread function*. This spreading is a direct representation of the blur in any point in the image and can be convolved with the matrix of all the pixels in the sampled image to create a representation of the blurred image. The line spread function is a one-dimensional form of the spreading and is usually more practical from a measurement perspective. In the case illustrated, the line spread function after quantization would be shown in Figure 3.5c as the corresponding distribution of gray pixels.

There are several observations to be made about this illustration, which underscore some of the practical problems encountered in typical measurements. First, the quantized image in Figure 3.5c is highly asymmetric while the profile of the line shown in Figures 3.5a and b appears to be more symmetric. This results from sampling phase and requires that a measurement of the line spread function must be made, adjusting sampling phase in some manner (Figure 3.7b). This is especially important in the practical situation of evaluating a fixed sampling frequency scanner. Secondly, note the limited amount of information in any one phase. It can be seen that the smooth curve representing the narrow object in Figure 3.5b is only represented by three points in the sampled and quantized image.

The averaging of several phases would improve on this measurement, increasing both the intensity resolution and the spatial resolution of the measurement. One of the easiest ways to do this is to use a long narrow line and tip it slightly relative to the sampling grid so that different portions along its length represent different sampling phases. One can then collect a number of uniformly spaced sampling phases, each being on a different scan line, while being sure to cover an integer number of complete cycles of sampling phase. One cycle is equivalent to a shift of one complete pixel. The results are then combined in an interleaved fashion, and a better estimate of the line spread function is obtained. (This is tantamount to increasing the sampling resolution, taking advantage of the one-dimensional nature of the test pattern.) This is done by plotting the recorded intensity for each pixel located at its properly shifted absolute position relative to the location of the line. To visualize this consider the two-phase sampling shown in Figure 3.7b and c. There the resulting pixels from phase A could be interleaved with those from phase B to create a composite of twice the spatial resolution. Additional phases would further increase effective resolution.

In the absence of nonlinearities and nonuniformities, the individual line spread functions associated with each of the effects in Table 3.3 can be mathematically convolved with each other to come up with an overall system line spread function.

3.4.2.1 MTF Approaches

For engineering analysis, use of convolutions and measurements of spread functions are often found to be difficult and cumbersome. The use of an *optical transfer function* (OTF) is considered to have many practical advantages from both the testing and theoretical points of view. The OTF is the Fourier transform of the line spread function. This function consists of a modulus to describe normalized signal contrast attenuation (or amplification), and a phase to describe shift effects in location, both given as a function of spatial frequency. The signal is characterized as the modulation of the sinusoidal component at the indicated frequency. Therefore the contrast altering function is described as a *modulation transfer function* (MTF). The value of OTF analysis is that all of the components in a linear system can be described by their OTFs, and these are multiplied together to obtain the overall system response. The method and theory of this type of analysis has been covered in many journal articles and reference books.^{13,47,78,102}

Certain basic effects can be described in analytic form as MTFs and a few of these are indicated in Table 3.3 and illustrated in Section 3.9 in Figures 3.51–3.53, plotted in logarithmic form to facilitate graphical manipulation. Several photographic MTF curves are plotted in Figure 3.55 to provide a reference both as a range of input signals for film scanners or a range of output filters that transform optical signals to permanently readable form. One may also consider using these with an enlargement factor for understanding input of photographic prints to a desktop or graphic arts scanner. (For example, a spatial frequency of 10 cycles/mm on an 8× enlarged print is derived from the 80 cycles/mm pattern on the film. Therefore the film MTF at 80 cycles/mm is an upper limit input signal for an 8 × 10 in enlargement of a 35-mm film.) Other output MTFs would involve display devices such as monitors, projection systems, analog response ink systems, and xerographic systems. Obtaining the transform of the line spread function has many of the practical problems associated with measuring the line spread function itself plus the uncertainty of obtaining an accurate digital Fourier transform using a highly quantized input.

There are several commonly used methods for measuring the OTF. These include:

1. Measuring images of narrow lines using appropriate compensation for finite widths.
2. Directly measuring images of sinusoidal distributions of radiation.^{103,104}
3. Harmonic analysis of square-wave patterns.^{97,103,104}
4. Taking the derivative of the edge profile in the image of a very sharp input edge. This generates the line spread function, and then the Fourier transform is taken, taking care to normalize the results properly.^{13,92} (Table 3.9, ISO TC42, WG18 and Figures 3.48a and b)
5. Spectral analysis of random input (e.g., noise) targets with nearly flat spatial frequency spectrum.

It should also be mentioned at this point that for most characterizations of imaging systems the modulus, that is, the MTF, is more significant than the phase. The phase transfer function, however, may be important in some cases and can be tracked either by careful analysis of the relative location of target and image in a frequency-by-frequency method or by direct computation from the line spread function.

In general, these methods involve the use of input targets that are not perfect. They must have spatial frequency content that is very high. The frequency composition of the input target is characterized in terms of the modulus of the Fourier transform, $M_{in}(f)$, of

its spatial radiance profile. The frequency decomposition of the output image is similarly characterized, yielding $M_{\text{out}}(f)$. Dividing the output modulation by the input modulation yields the MTF as

$$\text{MTF}(f) = \frac{M_{\text{out}}(f)}{M_{\text{in}}(f)} \quad (3.12)$$

The success of this depends upon the ability to characterize both the input and the output accurately.

A straightforward method to perform this input and output analysis involves imaging a target of periodic intensity variations and measuring the modulation on a frequency-by-frequency basis. If the target is a set of pure sine waves of reflectance or transmittance, that is, each has no measurable harmonic content, and the input scanner is linear, then the frequency-by-frequency analysis is straightforward. Modulation of a sinusoidal distribution is defined as the difference between the maximum and minimum divided by their sum. The modulation is obtained directly, measuring the maximum and minimum output gray values g' , and the corresponding input reflectance (or transmittance or intensity) values, R , of Equation 3.17 for each frequency pattern. Expanding the numerator and denominator for Equation 3.16 and the case of sinusoidal patterns and linear systems yields

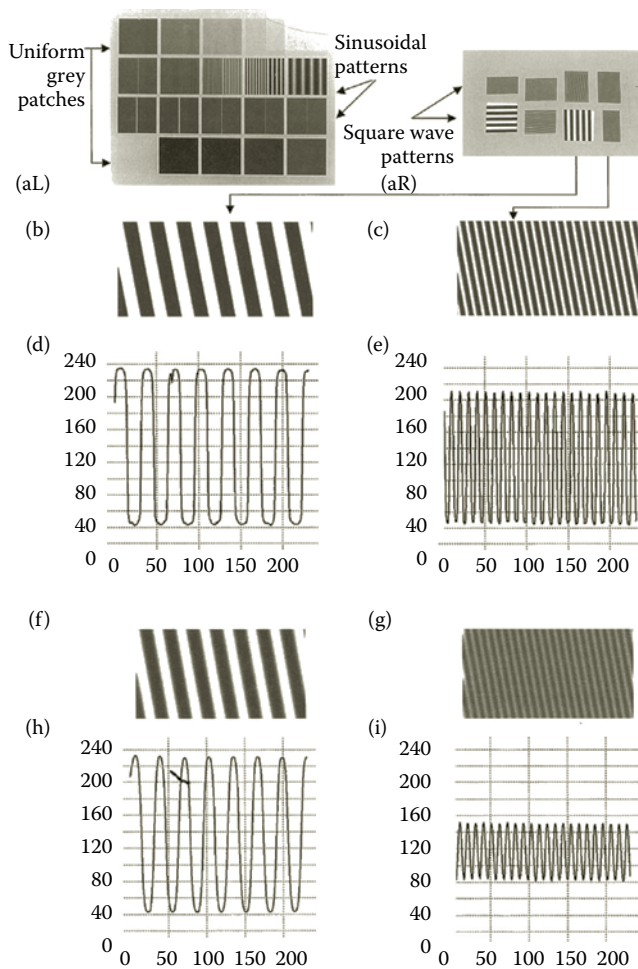
$$\text{MTF}(f) = \frac{-[g'_{\text{max}}(f) - g'_{\text{min}}(f)]/[g'_{\text{max}}(f) + g'_{\text{min}}(f)]}{[R_{\text{max}}(f) - R_{\text{min}}(f)]/[R_{\text{max}}(f) + R_{\text{min}}(f)]} \quad (3.13)$$

where the prime is used to denote gray response that has been corrected for any non-linearity as described below.

Figures 3.27 and 3.28 show an example of this process. In Figure 3.27 we see the layout of a representative periodic square-wave test target (aR) and a sinusoidal test target (aL) which exhibits features of well-known patterns¹⁰⁵ available today in a variety of forms (e.g., from Applied Image Reference 106). The periodic distributions of intensity (reflectance) are located in different blocks in the center of the pattern. Uniform reflectance patterns of various levels are placed in the top and bottom rows of the sinusoid to enable characterizing the tone response. A similar arrangement of uniform blocks is used with the square waves but not shown here. This enables correcting for its nonlinearities should there be any. Parts (b) and (c) show enlargements of parts of the square-wave pattern selecting a lower and a higher frequency. Parts (f) and (g) are enlargements of a gray image display of the electronically captured image of the same parts of the test target. Parts (d), (e), (h), and (i) show profiles of the patterns immediately above them.

To calculate a MTF, the modulation of each pattern is measured. For sinusoidal input patterns, one can use Equation 3.17 directly, finding the average maximum and minimum for many scan lines for each separate frequency. These modulation ratios, plotted on a frequency-by-frequency basis, describe the MTF. For square-wave input, the input and output signals must be Fourier transformed into their spatial frequency representations and only the amplitudes of the fundamental frequencies used in Equation 3.17. Schade¹⁰³ offers a method to compute the MTF by measuring the modulations of images of each square wave directly (i.e., the square wave response) and then unfolding for assumed perfect input square-waves without taking the transforms.

From a practical standpoint it is important to tip the periodic patterns slightly as seen in parts (d) and (e) to cover the phase distributions as described above under the spread

**FIGURE 3.27**

Example of images and profiles used in MTF analysis. Part (aL) shows a full pattern of gray patches and sinusoidal reflectance distributions at various frequencies. Part (aR) shows the frequency components of a square-wave test chart. Note that the bars are slanted slightly to facilitate measuring at different sampling phases. The figures on the left, (b), (d), (f), and (h), come from a low-frequency square-wave pattern as indicated by the arrow. The figures on the right, (c), (e), (g), and (i) are from a higher frequency square wave. Enlargements of the test patterns in (aR) are shown in parts (b) and (c). Slightly blurred images after scanning (as might be seen on a display of the scanner output) are shown in parts (f) and (g). Profiles of each of these images are displayed beneath them in parts (d), (e), (h), and (i), respectively. Because these are square-wave test patterns, special analysis of these patterns is required to compensate for effects of harmonics as described in the text. The reader should ignore small moiré effects caused by the reproduction process used to print this illustration.

function discussion. A new higher resolution image can be calculated by interleaving data points from the individual scan lines, each of which is phase shifted with respect to the sine or square wave.

Figure 3.28 shows several examples of linear and nonlinear response curves. It describes correcting the output of an MTF analysis (i.e., using g' in Equation 3.17) for the case of nonlinearity with offset. Here the maximum and minimum values for the sine waves are unfolded through the response curve to arrive at minimum and maximum input

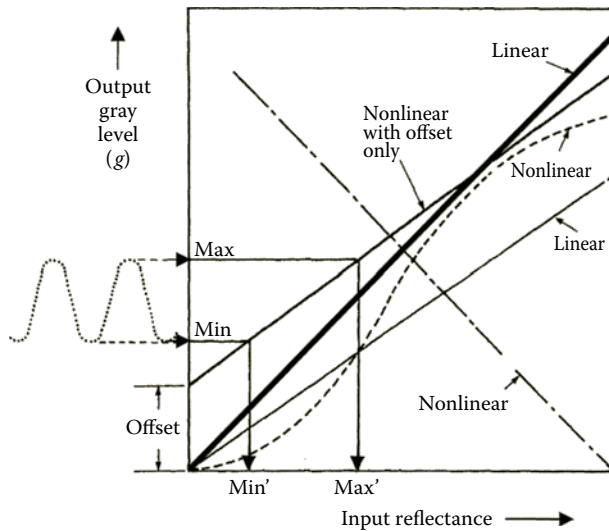


FIGURE 3.28

Examples of linear and nonlinear large-area response curves with an illustration of output modulation correction for offset using effective gray response at max' and min' .

reflectances, that is, the linear variables. For a full profile analysis, as needed for a Fourier transform method, and to obtain corrected modulation, each output gray level must be modified by such an operation.

If the response curve for the system was one of those indicated as linear in Figure 3.28, then no correction is required. It is important to remember that while the scanner system response may obey a straight-line relationship between output gray level and the reflectance, transmittance, or intensity of the input pattern, it may be offset due to either optical or electronic biases (e.g., flare light, electronic offset, etc.). This also represents a nonlinearity and must be compensated.

As the frequency of interest begins to approach the sampling frequency in an aliased input scanning system, the presence of sampling moiré becomes a problem. This produces interference effects between the sampling frequency and the frequency of the test pattern. If the pattern is a square wave, this may be from the higher harmonics (e.g., $3\times$, $5\times$ the fundamental). When modulation is computed from sampled image data using maxima and minima in Equation 3.17, errors may arise. There are no harmonics for the sinusoidal type of patterns, a distinct advantage of this approach.

See Figure 3.29 for an example of these phase effects on a representative MTF curve. It shows errors for test sine waves whose period is a submultiple of the sampling interval. Consider the case where the sinusoidal test pattern frequency is exactly one-half the sampling frequency, that is, the Nyquist frequency. In this case, when the sampling grid lines up exactly with the successive peaks and valleys of the sine wave, we get a strong signal indicating the maximum modulation of the sine wave (point A). When the sampling grid lines up at the midpoint between each peak and valley of the sinusoidal image (phase shifted 90° relative to the first position), each data point will be the same, and no modulation whatever results (point B). There is no right or wrong answer to the question of which phase represents the true sine-wave response, but the analog or highest value is often considered as the true MTF. Each phase may be considered as having its own sine-wave

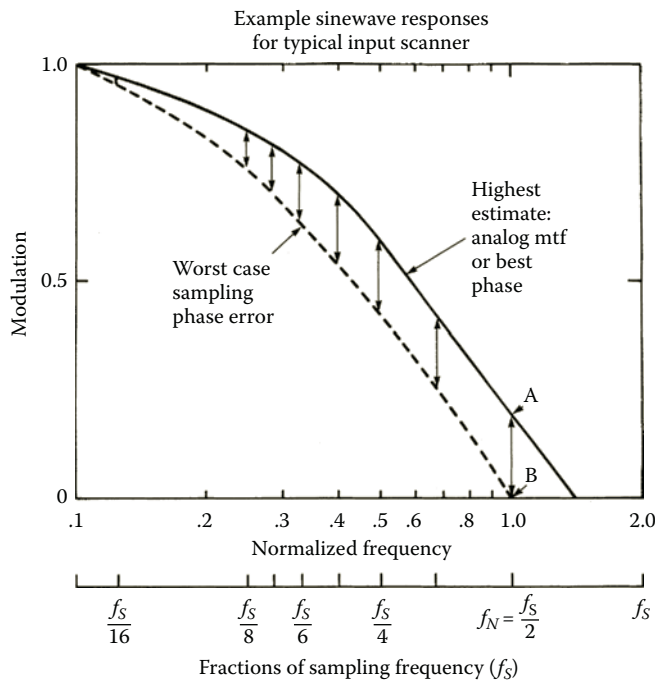


FIGURE 3.29

Example of the possible range of measured sine wave response values of an input scanner, showing the uncertainty resulting from possible phase variations in sampling.

response. Reporting the maximum and minimum frequency response and reporting some statistical average are both legitimate approaches, depending upon the intended use of the measurement. It is common practice to represent the average or maximum and the error range for the reported value.

The analog MTF, on the other hand, is only given as the maximum curve, representing the optical function before sampling. Therefore, the description of upper and lower phase boundaries for sine wave response shows the range of errors in the measurement of the MTF which one might get for a single measurement. This strongly suggests the need to use several phases to reduce error if the analog MTF is to be measured. Mathematically, phase errors may be thought of as a form of microscopic nonstationarity complicating the meaning of MTF for sampled images at a single phase. The use of information from several phases reduces this complication by enabling one to approximate the correct analog MTF that obeys the principle of stationarity.

In the case of a highly quantized system, meaning one having a relatively small number of gray levels, quantization effects become an important consideration in the design and testing of the input scanner. The graph in Figure 3.30 shows the limitation that quantization step size, E_q , imposes on the measurement of the MTF using sine waves. The number of gray levels used in an MTF calculation can be maximized by increasing the contrast of the sinusoidal signal that is on the input test pattern. It can also be increased by repeated measurements in which some analog shifts in signal level are introduced to cause the quantization levels to appear in steps between the previous discrete digital levels and therefore at different points on the sinusoidal distributions. The latter could be accomplished by changing the light level or electronic gain.

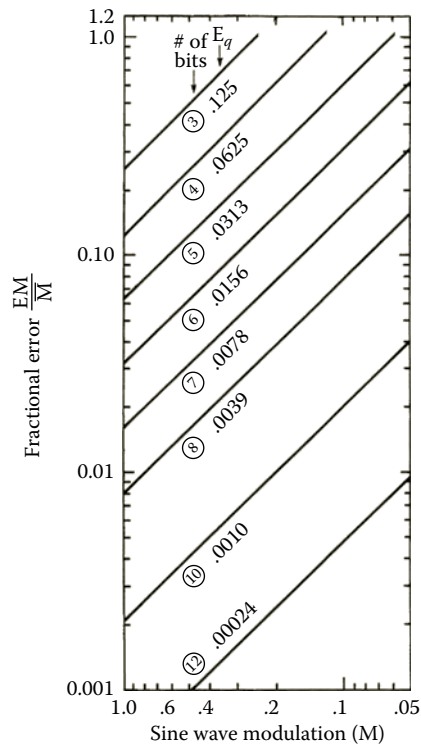


FIGURE 3.30

Errors in MTF measurements, showing the effects of modulation at various quantization errors. EM is in zero to peak units. M is average modulation. The numbers in the circles on each line indicate the system quantization in bits. E_q is the size of quantization step, where a full-scale signal = 1.0.

It is also important to note that because the actual MTF can vary over the field of view, a given measurement may only apply to a small local region over which the MTF is constant. (This is sometimes called an isoplanatic patch or stationary region within the image.) To further improve the accuracy of this approach, one can numerically fit sinusoidal distributions to the data points collected from a measurement, using the amplitude of the resultant sine wave to determine the average modulation. Taking the Fourier transform of the data in the video profile may be thought of as performing this fit automatically. The properly normalized amplitude of the Fourier transform at the spatial frequency of interest would, in fact, be the average modulation of the sine wave that fits the video data best.

The approach involving the application of square-wave test patterns (as opposed to sinusoidal ones, which have intrinsic simplicity as an advantage) has been shown by Newell and Triplett¹⁰⁴ to have significant practical advantages. They also show square-wave analysis has excellent accuracy when all-important details are carefully considered, especially the sampling nature of the analysis and the noise and phase effects. Square-wave test patterns are commonly found in resolving power test targets and are much easier to fabricate than sine waves, because the pattern exists as two states, foreground (e.g., black bars) and background (e.g., white bars). Two levels of gray bars may also be used depending on desired contrast. Fourier transform analysis and paying attention to the higher harmonics have been particularly effective.^{97,104} It has been shown that the general Discrete Fourier

Transform (DFT) algorithms where the length of the input can be altered is much better suited to MTF analysis than use of the Fast Fourier Transform (FFT) where the required power of two sampling points are a limitation.

One successful practice¹⁰⁴ included tipping the bar patterns to create a one pixel phase shift over eight scan lines and averaging over approximately 30 scan lines to reduce noise. The DFT was used and tuned to the precise frequency of the given bar pattern by changing the number of cycles of the square wave being sampled and using approximately 1000 data points. An improvement in MTF accuracy of several percent was demonstrated using the DCT over the more common FFT.

It is generally advisable to measure the target's actual harmonic content rather than to assume that it will display the theoretical harmonics of a perfect mathematical square wave. Likewise, other patterns of known spectral content can be calibrated and used. Edge analysis techniques are also popular.^{91,92,107} Using similar care such as a 5° slanted edge, a standardized algorithm, and specification on the edge quality, these achieve good accuracy too.

3.4.2.2 The Human Visual System's Spatial Frequency Response

As a matter of practical interest, several spatial frequency response measurements of the HVS are shown in Figure 3.31. These provide a reference to compare to system MTFs. The work of several authors is included.¹⁰⁸⁻¹¹⁴ The curves shown have all been normalized to 100% at their respective peaks to provide a clearer comparison. Except for the various normalizing factors, the ordinates are analogous to a modulation transfer factor of the type described by Equation 3.17. However, MTFs are applicable only to linear systems, which the human eye is not. The visual system is in fact thought to be composed of many independent, frequency-selective channels,^{115,116} which, under certain circumstances, combine to give an overall response as shown in these curves. It will be noted that the response of the visual system has a peak (i.e., modulation amplification relative to lower frequencies) in the neighborhood of 6 cycles/degree (0.34 cycles/milliradian) or 1 cycle/mm (25 cycles/in) at a standard viewing distance of 340 mm (13.4 inches). The variations among these curves reflect the experimental difficulties inherent in the measurement task and may also illustrate the fact that a nonlinear system such as human vision cannot be characterized by a unique MTF.¹¹⁷ For this reason such curves are called contrast sensitivity functions (CSFs) and not MTFs. For readers desiring a single curve, the luminance CSF reported by Fairchild⁴ is given in Section 3.9, Figure 3.56. While similar in shape to many curves in Figure 3.31, it displays a greater range of responses and also shows the red-green and blue-yellow chromatic CSFs.

3.4.2.3 Electronic Enhancement of MTFs: Sharpness Improvement

These visual frequency response curves suggest that the performance of an imaging system could be improved if its frequency response could be increased at certain frequencies. It is not possible with most passive imaging systems to create amplification at selected frequencies. The use of electronic enhancement, however, can impart such an amplified response to the output of an electronic scanner. Amplification here is meant to imply a high-frequency response that is greater than the very low-frequency response or greater than unity (which is the most common response at the lowest frequencies). This can be done by convolving the digital image with a finite-impulse response (FIR) electronic filter that has negative sidelobes on opposite sides of a strong central peak. The details of FIR

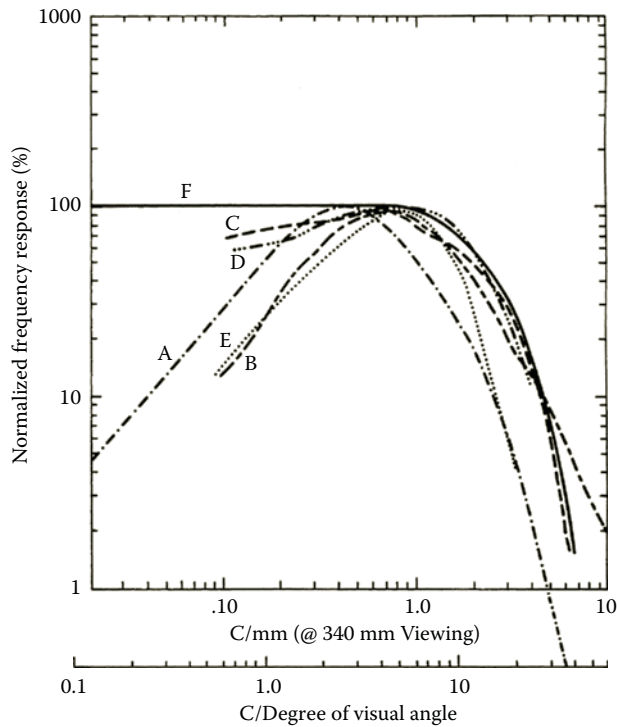


FIGURE 3.31

Measured spatial frequency response of the human visual system, showing the effects of experimental conditions on the range of possible results. Findings are presented according to (A) Campbell,¹⁰⁸ (B) Patterson¹¹⁴ (Glenn et al.¹¹²), (C) Watanabe et al.,¹¹¹ (D) Hufnagel (after Bryngdahl),¹¹⁷ (E) Gorog et al.,¹⁰⁹ (F) Dooley and Shaw.¹¹³ All measurements are normalized for 100% at peak and for 340-mm viewing distance. Note the universal visual angle scale at the bottom. See Figure 3.56 (from Fairchild⁴ for a seventh and more recent curve showing a larger response range and the two chromatic channels).

filter design are beyond the scope of this chapter, but the effects of two typical FIR filters on the system MTF are shown in Figure 3.32.

3.4.3 Noise Metrics

Noise in an input scanner, whether the scanner is binary or multilevel gray, comes in many forms (see Section 3.2.2). A brief outline of these can be found in Table 3.4. Various specialized methods are required in order to discriminate and optimize the measurement of each.

In this table we see that there are both fixed and time-varying types of noise. They may occur in either the fast or the slow scan direction and may either be additive noise sources or multiplicative noise sources. They may be either totally random or they may be structured. In terms of the spatial frequency content, the noise may be flat (white), that is, constant at all frequencies to a limit, or it may contain dominant frequencies, in which case the noise is said to be colored. These noise sources may be either random or deterministic; in the latter case, there may be some structure imparted to the noise.

The sources of the noise can be in many different components of the overall system, depending upon the design of the scanner. Instances of these may include the sensor of

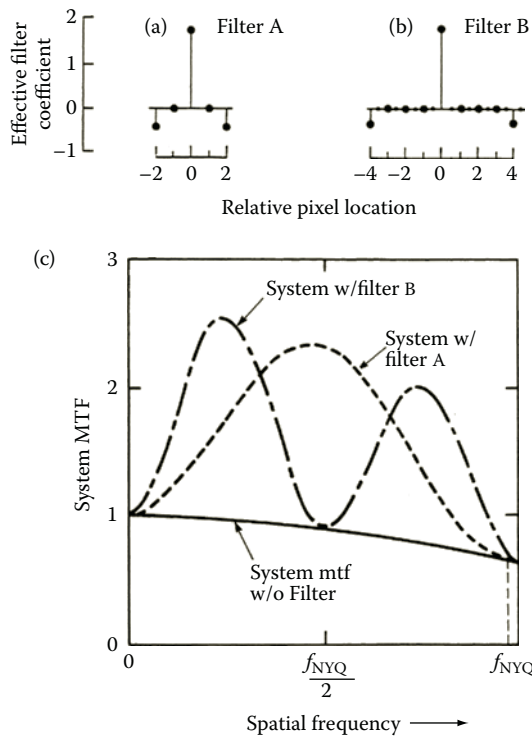


FIGURE 3.32

Two examples of enhancement of a scanning system MTF, using electronic finite-impulse response (FIR) filters in conjunction with an input scanner: (a) the one-dimensional line spread function for “filter A”; (b) line spread function for “filter B”, which is the same shown as in (a) but four pixels wider; (c) shows the effect of these filters on the MTF of an aliased high-quality scanner.

TABLE 3.4

Types and Sources of Noise in Input Scanners

Category	Type
Distribution	Fixed with platen, time varying
Type of operation	Multiplicative, additive
Spatial frequency	Flat (white), colored
Statistical distribution	Random, structured, image-dependent
Orientation	Fast scan, slow scan, none (two-dimensional)
Sources	Sensor, electronic system, motion error Calibration error, photons, lamp, laser Controls, optics

the radiation, or the electronics, which amplify and alter the electrical signal, including, for example, the A/D converter. Other noise sources may be motion errors, photon noise in low-light-level scanners, or noise from the illuminating lamp or laser. Sometimes the optical system, as in the case of a laser beam input scanner, may have instabilities that add noise. In many scanners there is a compensation mechanism to attempt to correct for fixed noise. This typically utilizes a uniformly reflecting or transmitting strip of one or more different densities parallel to the fast scan direction and located close to the input position

for the document. It is scanned, its reflectance(s) is memorized by the system, and it is then used to correct or calibrate either the amplifier gain or its offset or both.

Such a calibration system is, of course, subject to many forms of instabilities, quantization errors, and other kinds of noise. Since most scanners deal with a digital signal in one form or another, quantization noise must also be considered.

In order to characterize noise in a gray output scanner, one needs to record the signal from a uniform input target. The most challenging task is finding a uniform target with noise so low that the output signal does not contain a large component due to the input or document noise. In many of these scanners, the system is acting much like a microdensitometer, which reacts to such input noise as the paper fibers or granularity in photographic, lithographic, or other apparently uniform samples.

The basic measurement of noise involves understanding the distribution of the signal variation. This involves collecting several thousand pixels of data and examining the histogram of their variation or the spatial frequency content of that variation. Under the simplifying assumptions that we are dealing with noise sources that are linear, random, additive, and flat (white), a typical noise measurement procedure would be to evaluate the following expression:

$$s_s^2 = s_t^2 - s_o^2 - s_m^2 - s_q^2 \quad (3.14)$$

where s_s = the standard deviation of the noise for the scanner system (s); s_t = the total (t) standard deviation recorded during the analysis; s_o = the standard deviation of the noise in the input object (o) measured with an aperture that is identical to the pixel size; s_m = the standard deviation of the noise due to measurement (m) error; and s_q = the standard deviation of the noise associated with the quantization (q) error for those systems that digitize the signal. This equation assumes that all of the noise sources are independent. Removing quantization noise is an issue of whether one wants to characterize the scanner with or without the quantization effects, since they may in fact be an important characteristic of a given scanner design. The fundamental quantization error¹¹⁸ is

$$s_q^2 = \frac{2^{-2b}}{12} \quad (3.15)$$

where b = the number of bits to which the signal has been quantized.

The second and third terms in Equation 3.18 give the performance of the analog portion of the measurement. They would include the properties of the sensor amplification circuit and the A/D converter as well as any other component of the system that leads to the noise noted in the table above. The term s_q^2 characterizes the digital nature of the scanner and, of course, would be omitted for an analog scanning system.

Equation 3.18 is useful when the noise in the system is relatively flat with respect to spatial frequency or when the shape of the spatial frequency properties of all of the subsystems is similar. If, however, one or more of the subsystems involved in the scanner is contributing noise that is highly colored, that is, has a strong signature with respect to spatial frequency, then the analysis needs to be extended into frequency space. This approach uses Wiener or power spectral analysis.^{47,119} Systems with filters of the type shown in Figure 3.32 would exhibit colored (spatial frequency dependent) noise. A detailed development of Wiener spectra is beyond the scope of this chapter. However, it is important here to realize its basic

form. It is a particular normalization of the spatial frequency distribution of the *square* of the signal fluctuations. The signal is often in optical density (D), but may be in volts, current, reflectance, and so on. The normalization involves the area of the detection aperture responsible for recording the fluctuations. Hence units of the Wiener spectrum are often $[\mu\text{m } D]^2$ and can be $[\mu\text{m } R]^2$. (See Reference 119; the latter units are more appropriate for scanners because they respond linearly to reflectance, or, more generally, to irradiance.)

3.5 EVALUATING BINARY, THRESHOLDED, SCANNED IMAGING SYSTEMS

3.5.1 Importance of Evaluating Binary Scanning

Many output scanners accept only binary signals, that is, on or off signals for each pixel or subpixel. This translates to only black or white pixels on rendering. A binary thresholded image may be generated directly by the scanner, or reduced to this state through image processing just prior to delivery to an output scanner. It may also be the degenerate state of inappropriate gray or dithered image processing in which signals are overamplified in a variety of ways, to look like thresholded images. Irrespective of how they are generated, binary thresholded renderings remain an important class of images today and often produce image characteristics that are surprising to the uninitiated. Understanding and quantifying this type of imaging become an important part of the evaluation of the overall input scanner to output scanner-printer system.

As noted earlier, there are two types of binary digital images, either thresholded or dithered signals. To a first order, dithered systems (halftoned or error diffused) can be evaluated in a way similar to that used to evaluate full gray systems, with one simplifying assumption. The underlying concept is that, within the effective dither region over which a halftone dot is clustered or the error is diffused, the viewer does not notice the dither pattern. As a result, these systems are primarily evaluated using instruments and methods whose resolution is equal to or larger than the effective dither region and hence are confined largely to tone rendition and some forms of image noise. Extensive discussions of these measurement approaches are beyond the scope of this chapter. Many of the basic underlying principles are discussed in Section 3.5 of Chapter 3 of the earlier edition of this book and in 2.2.3 under halftone system response and detail rendition. Limited discussion follows here.

To understand and evaluate binary images, a few new concepts are explored and appropriate analytic methods developed.¹²⁰

3.5.1.1 Angled Lines and Line Arrays

To adequately describe performance over a range of sampling phases, it is important that the image structures must be measured at a large number of sampling phases. In other words the evaluation is repeated several times with respect to the input pattern at positions predetermined to create images at different sampling phases. These may be produced by shifting the components by various fractions of a pixel. Tilting lines or rectilinear patterns by a few degrees generates a continuum of phases along the edge of the designated structure. Without tilting for example, a fine line may be imaged in one test as two-pixel wide,

and on another random test and therefore at another sampling phase, it may be imaged as one-pixel wide.

3.5.2 General Principles of Threshold Imaging Tone Reproduction and Use of Gray Wedges

For binary system response, that is, binary tone reproduction, testing can best be accomplished by having smooth calibrated structures that allow finding the on-off binary transitions to a small fraction, say 1 part in 200, of an input characteristic like reflectance.

A calibrated gray wedge is useful. This device resembles a photographic step tablet except that it varies smoothly from a very low density to a very high density without steps. Ideally it would vary linearly in reflectance or transmittance as a function of distance, but the physical means for creating wedges often make them somewhat logarithmic. Accurate measurement of transmittance or reflectance versus distance from some reference mark on the wedge is used to calibrate the pattern, as shown in the graph in the top of Figure 3.33. Note that the picture at the bottom shows the image of a negative working system—that is, the maximum transmittance gives black and the minimum gives white.

The distance at which the wedge turns from black to white (or is 50% black and 50% white pixels for a noisy image) is measured for a given *gray threshold* and converted to a *transmittance or reflectance threshold*.

3.5.2.1 Underlying Characteristic Curve and Noise

If one is trying to determine the underlying characteristic curve of the scanner, a series of specified transmittances or reflectances can be determined along the wedge. The digital

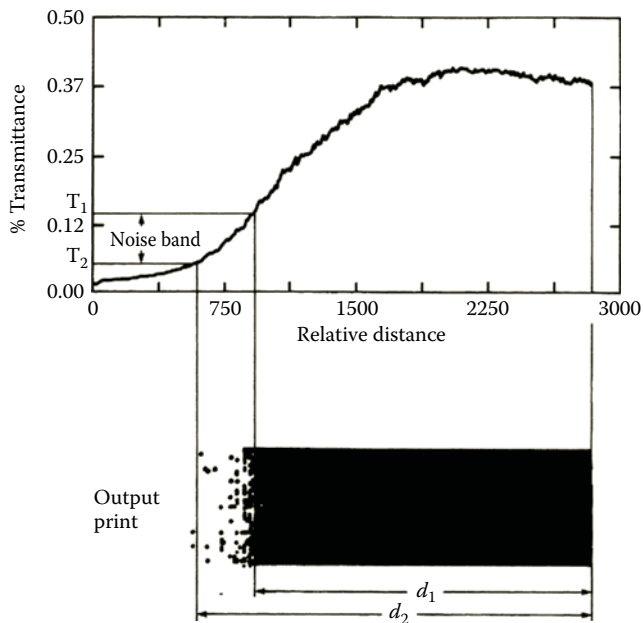


FIGURE 3.33

Transmittance profile of a gray wedge and a corresponding output print (binary image), both as function of distance in arbitrary units. Smallest dots at left are individual pixels.

output gray level of the threshold setting that creates the black to white transition at each specified transmittance or reflectance is then plotted. This describes the underlying characteristic curve of the binary system in output threshold gray level versus input transmittance or reflectance and is a type of OECF (optoelectronic conversion function) curve (Reference 65).

Because of noise in the typical system, including noise on the input document, the location along the length of the wedge where the image changes from white to black will not be a sharp straight line. Rather it will be a region of noise as shown in Figure 3.33. Typically the middle of this transition region $(T_2 + T_1)/2$ is identified as the transmittance or reflectance at which an average threshold is set. A unit called Gray Wedge Noise = GWN = $(T_1 - T_2)$ can be set here where T_2 and T_1 the transmittances at a fixed probability of finding a minimal or maximum response (such as 95% white or black).

3.5.3 Binary Imaging Metrics Relating to MTF and Blur

Given the on-off nature of a binary thresholded image, a linear approach such as MTF analysis does not work. To deal with this nonlinearity we can pose three specific types of questions about imaging performance: (1) *Detectability*: what is the smallest isolated detail that the system can detect? (2) *Discriminability of fine detail*: what is the finest, most complex small structure or fine texture that the system can handle (legibility, resolving power)? (3) *Fidelity of reproduction*: for the larger details and structures, how do the images compare to the original input such as some reasonable width line? To create a specific metric in each of these categories, one defines a specific test object or test pattern that relates to the imaging application. One then defines a set of rules or criteria by which to judge performance against that pattern. These include rules for determining threshold variation/selection and phase probabilities for decision criteria. These are more completely described in the earlier edition (Reference 53) including line width detectability and fidelity.

3.5.3.1 Resolving Power (A Measure for Discrimination of Fine Detail)

Resolving power is a commonly used descriptor of image quality for nearly every kind of imaging system. Its application to binary electronic imaging and scanning systems is therefore appealing. However, because of its extreme sensitivity to threshold and test pattern design, it must be applied with great care to prevent misleading results. Its primary value is in understanding performance for fine structures. The metrics noted above apply to isolated detail, while resolving power tends to emphasize the ability to distinguish many closely spaced details. In general, it can often be considered as an attempt to measure the cutoff frequency, that is, the maximum frequency for the MTF of a system. Binary systems are so nonlinear that even an approximate frequency-by-frequency MTF analysis cannot be considered.

The basic concept of a resolving power measurement is to attempt (through somewhat subjective visual evaluations) to detect a pattern in the thresholded video, which, to some level of confidence, resembles the pattern presented in the test target. For example, one may establish a criterion of 75% confidence that the image represents five black bars and four white spaces at the appropriate spacing. Values of 50%, 95%, 100% or any other confidence could also be used. As in all the metrics above, each judgment must be measured over a wide range of sampling phases for the bar pattern, resulting in an appropriate average confidence over all phases. Tipping the bar target so the length of the bars intercepts an integer number of sampling phases is again convenient. More detail on this test is given

in the earlier edition (Reference 53) including unusual and non-intuitive sampling patterns that may arise, pseudo- or spurious resolution (Reference 121) and gray in white spaces between bars in well-resolved image that is caused by light scattering in test target substrates⁷¹⁻⁷³ and may change expected thresholding results.

Binary imaging systems have extremely powerful contrast enhancement properties under the right circumstances. Selecting exactly the right threshold, one between the light and dark part of a resolving power image, amplifies a 1% or 2% modulation of the optical or gray electronic image to an on-off pattern (i.e., 100% modulation) that can be easily resolved in the video bit map.

Because it is possible to detect these low contrast patterns, it is also common to detect the situation known as pseudo- or spurious resolution. Here the blurring due to the input scanner is in a particular form that causes the light bars of the pattern to turn dark, and the dark bars to turn light.¹²¹ There are many strategies available for using resolving power. These would include:

1. Varying the threshold and noting the pattern that is resolved (see Figure 3.34).
2. Fixing the bar spacing of interest, and looking for the threshold at which it is detected.

It should be noted in any digital system (binary or gray) that there is a strong dependence on angular orientation. Unlike resolving power in a conventional optical system, a nonzero or non-90° orientation may in fact perform better because of the independent MTFs in the x - and y -directions and the rectangular sampling grid.

Resolving power test targets come in many forms and substrates and these make a significant difference in the results, as noted earlier. Some of these are illustrated in Figure 3.35.

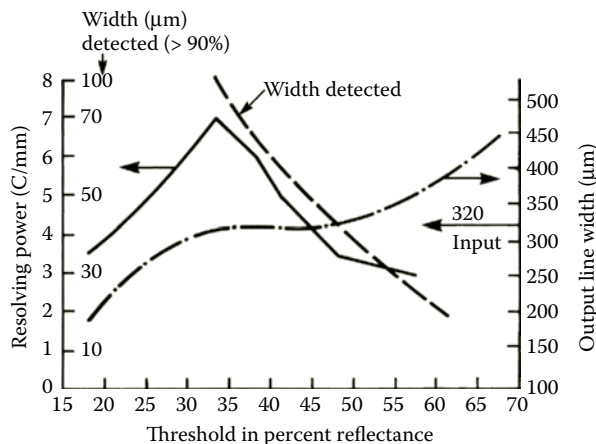
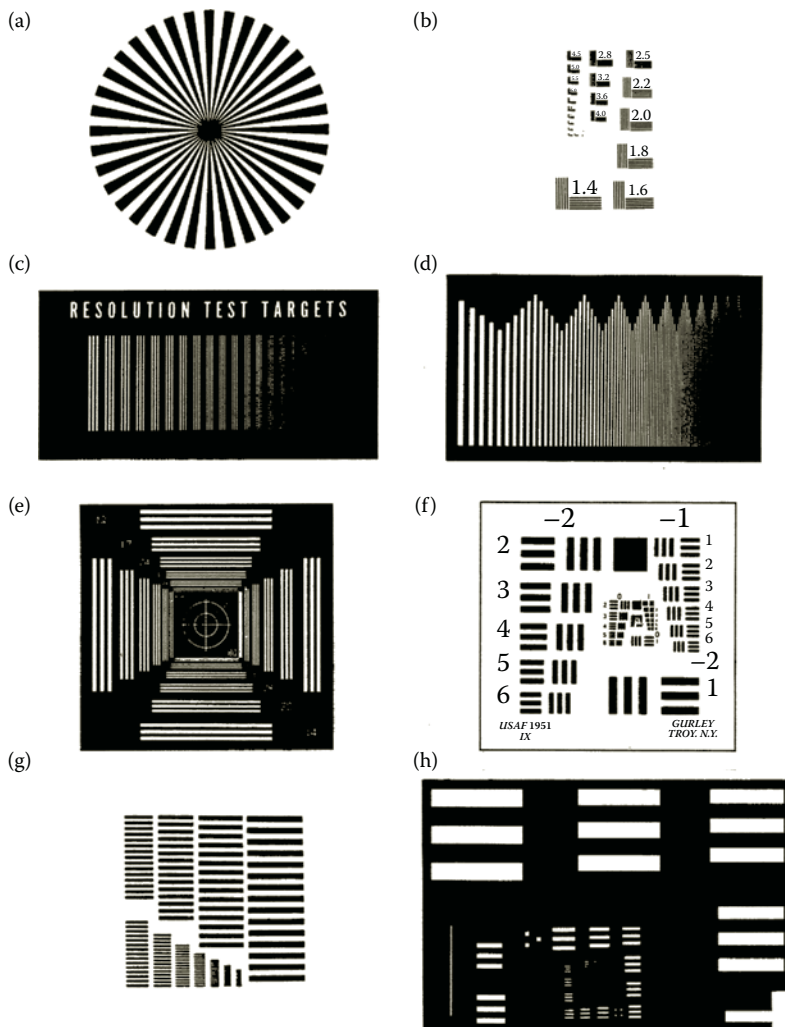


FIGURE 3.34

Plots of line width detectability, fidelity, and resolving power as functions of threshold setting in a binary imaging system. (Adapted from Lehmbek, D.R. *Imaging Performance Measurement Methods for Scanners that Generate Binary Output*. 43rd Annual Conference of SPSE, Rochester, NY, 1990; 202–203.). Arrows on each curve indicate which axis represents the ordinate for that curve. “Output line width” in μm is for images of 320 μm “input line width” as noted by arrow at right. The “Width Detected” curve refers to the left inside axis and is given in μm of the input line width, which is detected at the designated threshold for >90% of the sampling phases.

**FIGURE 3.35**

Images made with various bar pattern test targets in common use for measurement of resolving power and related metrics. See text for identification and description of each type: (a) radial graded frequency chart, (b) NBS Microcopy test chart, (c) a machine readable chart, (d) Sayce chart – a linear graded frequency pattern, (e) NBS lens testing chart, (f) USAF test chart, (g) Ealing test pattern, and (h) portion of ANSI resolving power chart.

Only the coarser patterns are imaged in this illustration and no attempt should be made by the reader to use these images for testing. They are illustrations only. Two general forms exist: those with discrete changes in the bar spacing and those with continuous variation in the bar spacing. In the former category there are several fairly commonly encountered types, designed for visual testing, namely the NBS lens testing type (e), NBS microcopy test chart type (b), the US Air Force type (f), the Cobb chart type (2 bars, not shown) and finally, the ANSI Resolving Power test patterns (h). This form also includes the extended square-wave types as represented by either Ronchi rulings (not shown) or ladder charts (not shown), which are simply larger arrays of the Ealing test pattern (g) which shows 15 bars of each square wave. Machine-readable forms are also useful where the modules are arranged in a pattern that can be scanned in a single straight line, as in (c). The

differences between these can be seen in the aspect ratio of the bars in the various patterns, the number of bars per frequency, the layout of the pattern itself, whether it is a spiral or a rectangle displaying progression of spacing, and the actual numerical progression of different frequency patterns within the target. In many cases low contrast versions or reverse polarity (white and black parts are switched) are also available.

The second major class, the continuously varying frequency pattern, is exemplified by the Sayce chart (d) and the radial graded frequency chart (a). The Sayce patterns are particularly useful for automated readout, provided the appropriate phase information is obtained (coordinates of each black bar) to prevent the pseudoresolution phenomena described above.

3.5.3.2 Line Imaging Interactions

A strategy for evaluating line and text imaging against all of the above metrics is to establish a fixed threshold that optimizes system performance for one of the major categories, such as line width fidelity, and then to report performance for the other variables, such as detectability and resolving power, at that threshold. One may also choose to plot detectability, fidelity, and resolving power as a function of threshold on a single plot in order to observe the relationships among the three and find an optimum threshold, trading off one against the other. This is illustrated in Figure 3.34 for a particular scanner. Such a plot provides several useful perspectives relating to the effects of blur on a binary system. It is clear in this example that the maximum fidelity occurs between 35% and 45% threshold while the maximum fine-line detectability keeps growing as the threshold drops below 30%. The resolving power has a distinct maximum at about 33%. Such a plot is different for each system and is governed by the shape of the underlying MTF curves and the various nonlinear interactions produced by image processing and the electronics.

3.5.4 Binary Metrics Relating to Noise Characteristics

Conventional approaches to measuring the amplitude of the noise fluctuations using various statistical measures of the distribution are not appropriate for binary systems. In these systems the noise shows up as pixels that are of the wrong polarity; that is, a black pixel that should have been white or a white pixel that should have been black. In general it is the distribution and location of these errors that need to be characterized. The practical approach to this problem is to examine the noise in a context equivalent to the main applications of interest for the binary imaging system. The resulting metrics include:

1. The range of uncertainty associated with determining the threshold using a gray wedge as described above in Section 3.5.2, which has led to the *gray wedge metric for noise*.
2. The noise seen on edges of lines and characters, which has led to the *line edge range metric for noise*.
3. The characterization of noise in a halftone image, that is, *halftone granularity*.

These are all described below.

3.5.4.1 Gray Wedge Noise

Figure 3.33 shows the transmittance profile of a gray wedge as a function of distance. The thresholded image of that gray wedge is shown below this profile with the x-axis lined

up to correspond to the position in the profile plot. Here, as noted earlier The GWN (Gray Wedge Noise) = $T_2 - T_1$.

To make a statistically satisfactory measure of noise, a probability distribution is used with the criteria for determining the positions d_1 and d_2 . As illustrated, these are the point where the signal is 95% black and the point where the signal is 95% white. Under the assumption of normally distributed noise this would represent plus or minus approximately two standard deviation limits on the noise distribution.

To fully characterize a binary system with this metric, one plots the width of the noise band in effective transmittance as a function of the independent variable, threshold (converted to transmittance).

3.5.4.2 Line Edge Noise Range Metric

The noise associated with the edges of lines is an important type of noise to be directly evaluated for many practical reasons. The image of every line has a microscopic gray region associated with it where the intensity falls off gradually from the white surround field into the black line. For the image of the edge of a line oriented at a very small angle to the sampling matrix there is a distribution of gray varying along the edge of the line. It gradually increases from white to black. In a binary system this scan line appears white until it reaches a fractional coverage along the tilted edge required by the threshold, and then changes to black. Much like the case associated with the wedge in the previous metric as the edge approaches the transition point where the threshold causes a change in the binary signal, the probability for an error resulting from noise increases. Thus the binary signal along the length of this slightly tipped line acts much like the signal for the wedge in the previous example, oscillating between black and white. This provides the basis for a second metric, which we refer to as the line edge noise range metric.

In Figure 3.36, a slightly tilted input line is shown relative to several scan lines. The binary video bit map for this line is shown in the lower part of Figure 3.36. Vertical lines mark the location at which the edge of the line makes a transition from the center of one raster line to the next. In the video bit map this transition is noisy and the two ranges in which this uncertainty of the black to white transition exists are indicated as N_1 and N_2 . The centers

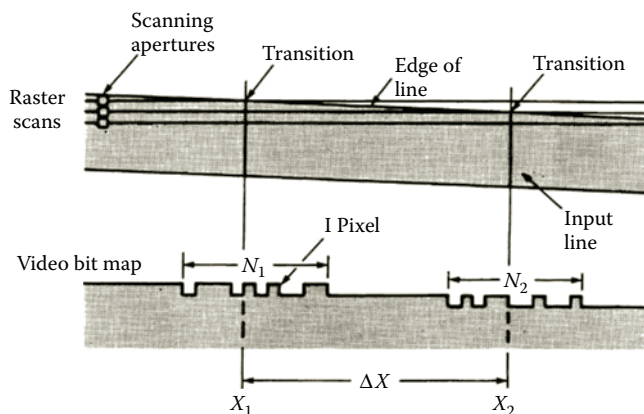


FIGURE 3.36

Scanning of a slightly tilted line, with the corresponding binary video bit map image, showing noise effects, which define edge noise range (ENR).

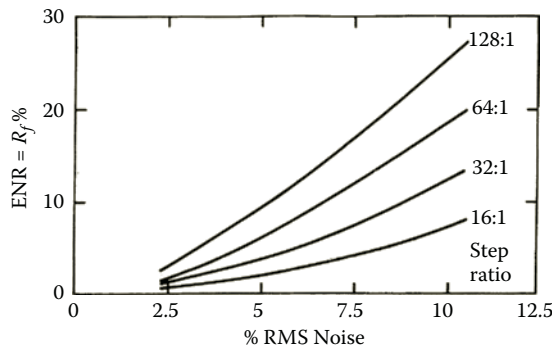


FIGURE 3.37
Relationship between LENR and RMS noise for various step ratios.

of these noisy transition regions are marked by the transition lines at X_1 and X_2 and are separated by the distance ΔX . The metric can be applied to edges of lines, or edges of solids, or any straight edge and is therefore generically referred to as “edge noise range” or ENR and simply defined as

$$ENR = \frac{\sum_{i=1}^n N_i}{\sum_{i=1}^{n-1} \Delta X_i} \tag{3.16}$$

The numerator and the denominator are averages over a large number of transitions along one or more constantly sloping straight lines. ΔX is the number of pixels per “step.” The range N is determined by subtracting the pixel number of the first white pixel in the black region from that of the last black pixel in the white region along the length of the line in each transition region.

Figure 3.37 shows the relationships among ENR, the step ratio ΔX , and the percentage of RMS noise in the imaging system, assuming additive white Gaussian noise distributions (private communication, J.C. Dainty, 1984) These are not intuitive relationships. For example, it should be noted from Figure 3.37 that an increase of a factor of 2 in RMS noise for a given angle line produces a line edge noise range increase of anywhere from $2\frac{1}{2}$ to $4\times$ depending upon the slope of the line and the exact noise level. It is noted that the noise is highly dependent on the angle of the line. Gradually sloping lines not only produce a larger absolute range but also a larger fractional range.

Here the MTF of the imaging system was considered to be perfect. The effect of blur, that is, decreases in MTF, is to increase the magnitude of ENR above those values shown.

It must also be noted that document noise will create extra fluctuations along the edge of the line and also increase the length of the range.

3.5.4.3 Noise in Halftoned or Screened Digital Images

This is the binary situation where a gray signal is created by a gray scanner and then converted to a binary halftone signal via processing in order to print on a binary rendering device (i.e., it is not a characteristic of a binary scanner per se). Scanning a typical

photograph and applying a halftone screen of the type described earlier (2.2.3) results in a bit map in which some of the arrangements of pixels in the halftone cells do not follow the prescribed growth pattern for the screen (see Figure 3.38). The noise in the scanning system itself can produce pixel-by-pixel changes in the effect of threshold at each one of the sites within the halftone cell. Some of these errors are introduced by the partial dotting mechanism, described earlier, when the granularity of the otherwise uniform input document, which was scanned to create the image, is of sufficient contrast to change the structure inside areas formed by individual halftone cell's threshold matrix. See Section 3.2.2.3 and Figures 3.14 and 3.15 for a review of these mechanisms.

One way of evaluating this type of noise is to create images of a series of perfectly uniform patches of differing density and process them through the halftoning method of choice. One then measures the RMS fluctuation in the percent area coverage for the resulting halftone cells, one patch at a time. To the extent that the output system is insensitive to the orientation of the bit map inside the halftone cell, this fluctuation becomes a reasonable measure of the granularity of the digital halftone pattern. For the electronic image, it can be calculated with a simple computer program that searches out each halftone cell and calculates its area coverage, collecting the statistics over a large number of halftone cells.

There are many image analysis packages on the market that will find particles in a digital image and evaluate their statistical distribution. They are found in biology, medicine, or metallurgy software applications as well as in image analysis packages.⁸⁷⁻⁸⁹ In this case a "particle" is a halftone dot whose area corresponds to the number of pixels.

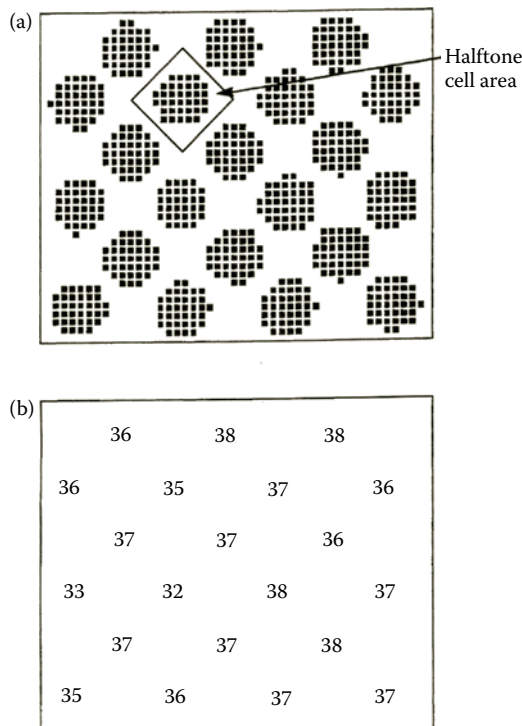


FIGURE 3.38

Noise (granularity) in a binary halftone image, part (a) Bitmap of the halftone rendering of a scanned image of a uniform area on an original. Part (b) Number of black pixels in each cell where the average number of black pixels per cell is 36.4 and the estimate of the standard deviation is 1.56 pixels.

If the image has been printed, a microdensitometer can be set up with an aperture that exactly covers one halftone cell, then scanned along rows of halftone dots. The RMS fluctuations or the low-frequency components of the noise spectra can then be evaluated. This is sometimes referred to as *aperture filtered granularity* measurement.

3.6 SUMMARY MEASURES OF IMAGING PERFORMANCE

Many attempts have been made to take the general information on image quality measurement and reduce it to a single measure of imaging performance. These often take the form of shortcut "D" in Figure 3.3. While none of the resulting measures provides a single universal figure of merit for overall subjective image quality, each brings additional insight to the design and analysis of particular imaging systems and its quality. Each has achieved some level of success in a limited range of applications. Preferred image quality, however, is a psychological reaction to a complex set of trade-offs and visual stimuli. There is a very subjective, application-oriented aspect to this reaction that does not readily lend itself to analytical description.

Instead, in an attempt to help the engineer control or design his systems, we shall describe a number of metrics.

The metrics are described here in their general form, many were developed for analog imaging systems such as cameras and film, others for digital. To the extent that the scanning systems in question are unaliased and have a large number of gray levels associated with them, the direct application of analog metrics is valid. In general, it should be remembered that digital imaging systems are not symmetric in slow and fast scan orientations in either noise or spatial frequency response (MTF). Therefore, what is given below in one-dimensional units must be applied in both dimensions for successful analysis of a digital input scanner. These concepts can be extended to an entire imaging system with little modification if the subsequent imaging modules, such as a laser beam scanner, provide gray output writing capability and generate no significant sampling or image conversion defects of their own (i.e., they are fairly linear). Since full gray scale input scanners are usually linear, most of these concepts can be applied to them, with the qualification that some display or analysis technique is required to convert the otherwise invisible electronic image to a visual or numerical form.

Most of these summary measures can be described by curves like the illustrative ones shown in Figure 3.39. Here we show some common measures, generically symbolized by F for both signal and noise, such as intensity, modulation or (modulation)² plotted as a function of spatial frequency f . A signal $S(f)$ is shown generally decreasing from its value at 0 spatial frequency to the frequency f_{\max} . A limiting or noise function $N(f)$ is plotted on the same graph starting at a point below the signal; it too varies in some fashion as spatial frequency increases. The various unifying constructs (metrics) involve very carefully considered approaches to the relationship between S and N , to their respective definitions, and to the frequency range over which the relationship is to be considered, along with the frequency weighting of that relationship

3.6.1 Basic Signal-to-Noise Ratio

The simplest of all signal-to-noise measures is the ratio of the mean signal level $S(0)$ to the standard deviation $N(0)$ of the fluctuations at that mean. If the system is linear and the

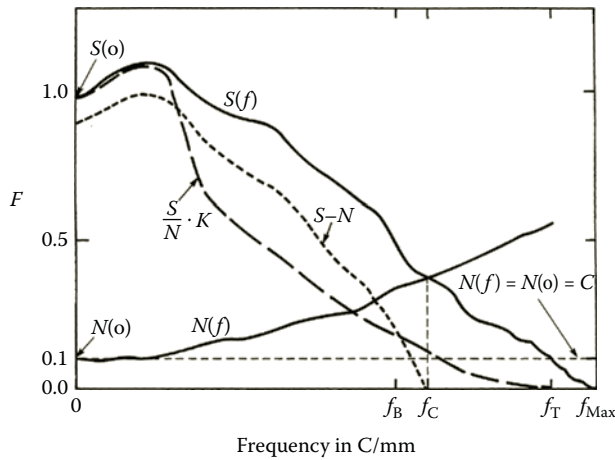


FIGURE 3.39 Signal, $S(f)$, noise, $N(f)$, and various measures of the relationships between them, plotted as functions of spatial frequency. Various critical frequencies are noted as points on the frequency axis.

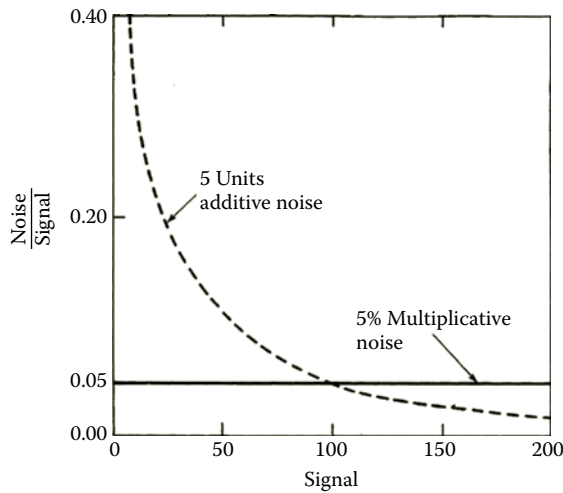


FIGURE 3.40 Relative noise (noise/signal) as a function of signal level for additive and multiplicative noise.

noise is multiplicative, this is a useful single number metric. If the noise varies with signal level, then this ratio is plotted as a function of the mean signal level to get a clearer picture of performance. Hypothetical elementary examples of this are shown in Figure 3.40, in which are plotted both the multiplicative type of noise at 5% of mean signal level (here represented as 100) and additive noise of 5% with respect to the mean signal level. It can be seen why such a distinction is important in evaluating a real system. It should be noted that in some cases, multiplicative or additive noise might vary as a function of signal level for some important design reason.

In comparing signals to noise, one must also be careful, to ensure that the detector area over which the fluctuations are collected is appropriate for the application to which the signal-to-noise calculation pertains. This could be the size of the input or output pixel, of

the halftone cell, or of the projected human visual spread function. The data must also be collected in the orientation of interest. In general, for scanned imaging systems there will be a different signal-to-noise ratio in the fast scanning direction than in the slow scanning direction.

3.6.2 Detective Quantum Efficiency and Noise Equivalent Quanta

When low light levels or highly noise limited situations occur, it is desirable to apply the concepts of detective quantum efficiency (DQE) and noise equivalent quanta (NEQ). These fundamental measurements have been extensively discussed in the literature.⁴⁷

If we set gain of a system response to a constant r in arbitrary output units, and assume the distribution of fluctuations obeys normal statistics, then we can write

$$DQE = \frac{r^2 q}{\sigma_0^2} \quad (3.17)$$

$$NEQ = \frac{r^2 q^2}{\sigma_0^2} \quad (3.18)$$

where σ_0 represents an estimate of the standard deviation of the distribution of the output fluctuations. Here for DQE the average signal level q is divided by the square of the standard deviation (i.e., the variance) and contains a modifier that is related to the characteristic amplification factors associated with a particular detection system, namely r , which also enters as a square. It should also be pointed out that DQE is an absolute measure of performance, since q is an absolute number of exposure events, that is, number of photons or quanta.

For illustration purposes, Figure 3.39 shows all of the above constructs.

3.6.3 Application-Specific Context

The above descriptions are frequently derived from the fundamental physical characteristics of various imaging systems, but the search for the summary measure of image quality usually includes an attempt to arrive at some application-oriented subjective evaluation, correlating subjective with objective descriptions. Applications that have been investigated extensively include two major categories: those involving detection and recognition of specific types of detail and those involved in presenting aesthetically pleasing renderings of a wide variety of subject matter. These have centered on a number of imaging constraints, which can usually be grouped into the categories of display technologies and hard-copy generation. Many studies of MTF have been applied to each.^{85,97,122–126} All of these studies are of some interest here. Note that modern laser beam scanning tends to focus on the generation of hard copy where the raster density is hundreds of lines per inch and thousands of lines per image compared with the hundreds of lines per image for early CRT technology used in the classical studies of soft display quality.

3.6.4 Modulation Requirement Measures

One general approach characterizes $N(f)$ in Figure 3.39 as a “demand function” of one of several different kinds. Such a function is defined as the amount of modulation or signal required for a given imaging and viewing situation and a given target type. In one class

of applications, the curve $N(f)$ is called the threshold detectability curve and is obtained experimentally. Targets of a given format but varying in spatial frequency and modulation are imaged by the system under test. The images are evaluated visually under conditions and criteria required by the application. Results are stated as the input target modulation required (i.e., “demanded”) for being “just resolved” or “just detected” at each frequency. It is assumed that the viewing conditions for the experiment are optimum and that the threshold for detection of any target in the image is a function of the target image modulation, the noise in the observer’s visual system, and the noise in the imaging system preceding the observer. At low spatial frequencies this curve is limited mostly by the HVS, while at higher frequencies imaging system noise as well as blur may determine the limit.

One such type of experiment involves measuring the object modulation required to resolve a three-bar resolving power target. For purposes of electronic imaging, it must be recalled that the output video of an input scanner cannot be viewed directly, and therefore any application of this method must be in the systems context, including some form of output writing or display. This would introduce additional noise restrictions. The output could be a CRT display of some type, such as a video monitor with gray-scale (analog) response. Another likely output would be a laser beam scanner writing on xerography or on silver halide film or paper. The details for measuring and using the demand function can be found in work by Scott¹²⁷ for the example of photographic film and in Biberman,⁸⁵ especially Chapter 3 for application to soft displays.

3.6.5 Area under the MTF Curve (MTFA) and Square Root Integral (SQRI)

Modulation detectability, while useful for characterizing systems in task-oriented applications, is not always useful in predicting overall image quality performance for a broad range of imaging tasks and subject matters. It has been extended to a more general form through the concepts of the threshold quality factor¹²⁸ and area under the MTF curve (MTFA).^{129,130} These concepts were originally developed for conventional photographic systems used in military photo-interpretation tasks.¹²⁸ They have been generalized to electro-optical systems applications for various forms of recognition and image-quality evaluation tasks, mostly involving soft displays.⁸⁵ The concept is quite simple in terms of Figure 3.39. It is the integrated area between the curves $S(f)$ and $N(f)$ or, equivalently, the area under the curve labeled $S-N$. In two dimensions, this is

$$\text{MTFA} = \int_0^{f_{cx}} \int_0^{f_{cy}} [S(f_x, f_y) - N(f_x, f_y)] df_x df_y \quad (3.19)$$

where S is the MTF of the system and N is the modulation detectability or demand function as defined above, and f_{cx} and f_{cy} are the two-dimensional “crossover” frequencies equivalent to f_c shown in Figure 3.39.

This metric attempts to include the cumulative effects of various stages of the scanner, films, development, the observation process, the noise introduced into the perceived image by the imaging system, and the limitations imposed by psychological and physiological aspects of the observer by building all these effects into the demand function $N(f)$. Extensive psychophysical evaluation and correlation has confirmed the usefulness of this approach¹³⁰ for recognition of military reconnaissance targets, pictorial recognition in general, and for some alphanumeric recognition.

Related approaches using a visual MTF weighting have been successfully applied to a number of display evaluation tasks, showing good correlation with subjective quality.¹²³

Many studies examine differences in quality where noise factors are relatively constant. One of these is the square root integral (SQRI) model of Bartend.^{124,131} Here, the demand function is specified by a general contrast sensitivity of the HVS and the comparison of the quality of two images of interest is specified in JND units (see Section 3.8 for a definition of JND, a just noticeable difference).

$$J = \frac{1}{\ln 2} \int_0^{W_{\max}} \sqrt{\frac{M(w) d(w)}{M_t(w)w}} \quad (3.20)$$

where $M(w)$ is the cascaded MTF of the image components, including that of the display and $M_t(w)$ is threshold MTF of the HVS, both in units of angular spatial frequency w . Results are to be interpreted with the understanding that 1 JND is “practically insignificant.” It is equal to a 75% correct response in a paired comparison experiment. Note that 3 JND is “significant,” and 10 JND is “substantial.”^{124,132} The $M_t(w)$ term describes the HVS as the threshold contrast for detecting a grating of angular frequency w as follows:

$$1/M_t(w) = aw \exp(-bw) \sqrt{1 + c \exp(bw)} \quad (3.21)$$

where

$$a = \frac{540(1 + 0.7/L)^{-0.2}}{1 + 12/[s_w(1 + w/3)^2]} \quad (3.22a)$$

$$b = 0.3(1 + 100/L)^{0.15} \quad (3.22b)$$

$$c = 0.6 \quad (3.22c)$$

L is the display luminance in cd/m^2 and s_w is the display size or width in degrees. These equations have been shown to have high correlation with perceived quality over a wide range of display experiments,¹²⁴ one of which is shown in Figure 3.41. Here the resolution, size, and subject matter of projected slides were varied and the equation was fit to the data.

It is noted by Barten that noise should be taken into account in the modulation threshold function, which is done by using a root sum of squares method of a weighted noise modulation factor to $M_t(w)$.¹²¹ Other authors have expanded on these concepts, extending them to include more fundamentals of visual mechanisms.^{125,126,133}

3.6.6 Measures of Subjective Quality

Several authors have explored the broader connection between objective measures of image quality and overall aesthetic pictorial quality for a variety of subject matters encountered in amateur and professional photography.^{134–140} The experiments to support these studies are difficult to perform, requiring extremely large numbers of observers to obtain good statistical measures of subjective quality. The task of assessing overall quality is less well defined than the task of recognizing a particular pattern correctly, as evaluated in most of the studies cited above. It would appear that no single measurement criterion has become

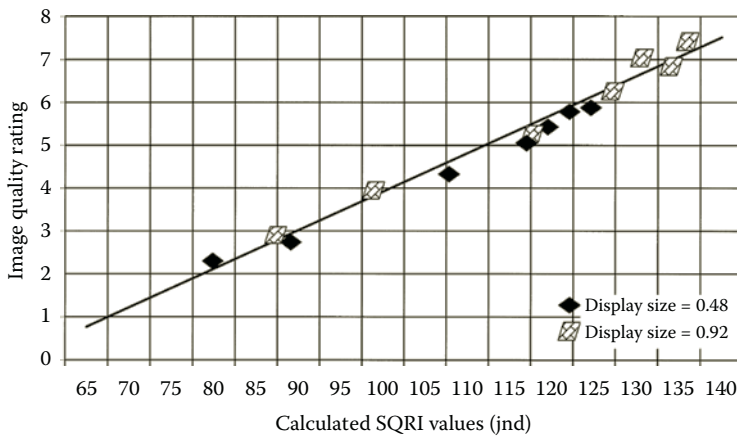


FIGURE 3.41

Linear regression between measured subjective quality and calculated SQRI values for projected slides of two different sizes, as indicated, illustrating the good fit. (Adapted from Barten, P.G.J. The Square Root Integral (SQRI): A new metric to describe the effect of various display parameters on perceived image quality. *Proceedings of SPIE conference on Human Vision, Visual Processing, and Digital Display, Los Angeles, CA, 1989; Vol. 1077, 73–82.*)

universally accepted by individuals or organizations working in this area. Below we shall discuss a few of the key descriptors, but we do not attempt to list them all.

Many of the earlier studies tended to focus on the signal or MTF-related variable only. In one such series of studies,^{134,135} S in Figure 3.39 is defined as the modulation of reflectance on the output print (for square waves) divided by the modulation on the input document (approximately 0.6 for these experiments). The quality metric is defined as the spatial frequency at which this ratio falls to 0.5. This is indicated in Figure 3.39 by the frequency f_b for the curve S as drawn. In these studies, a landscape without foreground was rated good if this characteristic or critical frequency was 4–5 cycles/mm (100–125 cycles/in), but for a portrait 2–3 cycles/mm (50–75 cycles/in) proved adequate. Viewing distance was not a controlled variable. By using modulation on the print and not simply MTF, the study has included the effects of tone reproduction as well as MTF. Granularity was also shown to have an effect, but was not explicitly taken into account in the determination of critical frequency.

Several studies have shown that the visual response curve discussed earlier can be connected with a measure of $S(f)$ to arrive at an overall quality factor. See, for example, system modulation transfer acutance (SMT acutance) by Crane¹³⁹ and an improvement by Gendron¹⁴⁰ known as cascaded area modulation transfer (CMT) acutance. One metric, known as the subjective quality factor (SQF),¹³⁶ defines an equivalent passband based on the visual MTF having a lower (initial) cutoff frequency at f_i and an upper (limiting) frequency of f_l . Here, f_i is chosen to be just below the peak of the visual MTF, and f_l is chosen to be four times f_i (two octaves above it). For prints that are to be viewed at normal viewing distance [i.e., about 340 mm (13.4 in)], this range is usually chosen to be approximately 0.5–2.0 cycles/mm (13–50 cycles/in).

The MTF of the system is integrated as follows.

$$\text{SQF} = \int_{f_x=0.5}^2 \int_{f_y=0.5}^2 S(f_x, f_y) d(\log_{10} f_x) d(\log_{10} f_y). \quad (3.23)$$

This function has been shown to have a high degree of correlation with pictorial image quality over a wide range of picture types and MTFs. It is possible that a demand function similar to that described in the MTF concepts above could be applied to further improve the performance. The SQF metric is applied to the final print as it is to be viewed and may be scaled to the imaging system, when reduction or enlargement is involved, by applying the appropriate scaling factor to the spatial frequency axis.

It should be noted that there is a significant difference between the upper band limit of this metric at 2 cycles/mm (50 cycles/in) and the critical frequency described above in Biedermann's work for landscapes, which is in the 4–5 cycles/mm (100–125 cycles/in) region. But there is good agreement for the portrait conclusions of the earlier work, which cites an upper critical frequency of 2–3 cycles/mm (50–75 cycles/in). Authors of both metrics acknowledge the importance of granularity or noise without directly incorporating granularity into their algorithms. Granger¹⁴¹ discusses some effects of granularity and digital structure in the context of the SQF model, but calls for more extensive study of these topics before incorporating them into the model.

It is clear that when the gray content and resolution of the digital system are high enough to be indistinguishable from an analog imaging system, then these techniques, which are general in nature, should be applicable. The quantization levels at which this equivalence occurs vary broadly. Usually 32 to 512 levels of gray suffice, depending on noise (higher noise requires fewer levels), while resolution values typically range from 100 to 1000 pixels/in (4–40 pixels/mm), depending on noise, subject matter, and viewing distance.

Another fairly typical approach to quantifying overall subjective image quality involves measuring the important attributes of a set of images made under a range of technology variables of interest and then surveying a large number of observers, usually customers for the products using the technology. They are asked for their overall subjective reaction to each image. A statistical regression is then performed between the measured attributes and the average subjective score for each image. This is the "type D shortcut" illustrated in Figure 3.3. An equation describing quality is derived using only the most important terms in the regression, that is, those that describe most of the variance. The "measures" may also be visual perceptions, that is, the "nesses," in which case the result is Engeldrum's "image quality models,"^{12,41} but must include all the factors that could have any reasonable bearing on quality. Sometimes the technology variables themselves are used (type A shortcut, Figure 3.3). This makes the resulting equation less general in its applicability but gives immediate answers to product questions.

Below is an example of an *image quality model*¹⁴² selecting five visual perception attributes from a list of 10 general image quality attributes¹⁴³ to describe a series of 48 printed color images from lithography, electrophotography, inkjet, silver halide, and dye diffusion, under a wide variety of conditions. A linear regression against overall preferences of 61 observers yielded the following equation.

$$\begin{aligned} \text{Avg. Preference} = & 8.8 \text{ Color Rendition} + 5.5 \text{ MicroUniformity} \\ & + 4.4 \text{ Effective Resolution} + 3.5 \text{ MacroUniformity} \\ & + 1.9 \text{ Gloss Uniformity.} \end{aligned} \quad (3.24)$$

A plot showing the Preference versus a fitted three-dimensional surface for the top two correlates is given in Figure 3.42.

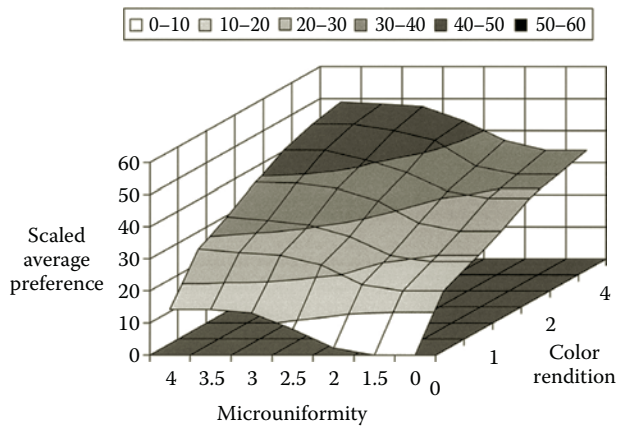
**FIGURE 3.42**

Illustration of the multivariate nature of a typical image quality model showing relationship between scaled image quality preference and two of several variables: color rendition and microuniformity. (Adapted from Natale-Hoffman, K.; Dalai, E.; Rasmussen, R.; Sato, M. *Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Savannah, GA, 1999; 266–273.*)

Other regression equations between physical or measured image parameters and customer preference have been developed in many different imaging environments some tackling huge lists of variables.¹⁴⁴

$$\begin{aligned}
 \text{Portrait IQ rating} = & 0.393 \times 10^{-8} (C^* \text{ of red } 100\%)^{5.2} \\
 & + 69.51 \times \exp(-0.125 \times \text{graininess of cyan } 60\%) \\
 & - 0.000173 \times (H^0 \text{ of } 1.0 \text{ Neutral Solid} - 305.0)^2 \\
 & - 0.409 \times (C^* \text{ of blue } 10\% - 4.90)^2 \\
 & + 47.7 \times \exp(-0.0766 \times \text{graininess of skin color}) \\
 & - 0.0197 \times (C^* \text{ of blue } 40\% - 23.5)^2 \\
 & - 0.0452 \times (C^* \text{ of Cyan } 70\% - 36.8)^2 \\
 & - 15.22
 \end{aligned}$$

A study comparing perceived image quality and acceptability of photographic prints of images from different resolution digital capture devices⁷⁷ directly compared perceived quality by a range of individuals with varying experience in photography and computers. Photographic prints (4 × 6 in) showing optimum tone and color rendering were used as output for viewing. Their results and others were given earlier in Figure 3.20.

3.6.7 Information Content and Information Capacity

There are numerous articles in the imaging science literature that analyze imaging systems in terms of information capacities and describe their images as having various information contents.

Using basic statistics of noise and spread function concepts from Section 3.2.1.3 a simple description of image information is given by Equation 3.33.^{46,50,51} It defines image

information, H , as

$$H = a^{-1} \log_2 \left[\frac{\text{probability of "density message" being correct}}{\text{probability of a specific density as input}} \right] \quad (3.25)$$

where a is the area of the smallest resolvable unit in the image (i.e., 2×2 pixels based on unaliased sampling from the sampling theorem) and the log factor is from the classic definition of information in any message,¹⁴⁵ here being messages about density (any other signal units can be used if done so consistently and they constitute a meaningful message in some context). To convert this into more useful terms let

$$H = a^{-1} \log_2 \left[\frac{p}{1-M} \right] \underset{p \rightarrow 1}{\cong} \left[\frac{\log_2 M}{a} \right] \quad (3.26)$$

where the numerator is set equal to p , the probability that a detected level within a set of levels is actually the correct one (i.e., the reliability), and M is the number of equally probable distinguishable levels (i.e., the quantization) from Equation 3.1 in Section 3.2.1.3. Assuming a high reliability such that p approaches unity, the simplification on the right results. The standard deviation of density in Equation 3.1 must be measured with a measuring tool whose aperture area is equal to a .

An approximation useful in comparing different photographic materials uses the standard deviation of density σ_a at a mean density of approximately 1 to 1.5 and Equation 3.30 results:

$$H = a^{-1} \log_2 \left[\frac{L}{6\sigma_a} \right] \quad (3.27)$$

where k was set to ± 3 ($= 6$), leading to $p = 0.997$ (~ 1); L is the density range of the imaging material.

Since the standard deviation of density is strongly dependent on the mean density level, it is more accurate and also common practice to measure the standard deviation at several average densities and segment the density scale into adjacent, empirically determined, unequal distinguishable density levels. These levels are separated by k standard deviations of density as measured for each specific level.^{46,50,51} If the input scanner itself is very noisy, then the σ_a term must represent the combined effects of both input noise and scanner noise. This was covered in Section 3.4.3 (see References 46, 50, and 51 for further information).

Another approach uses all of the spatial frequency based concepts developed above for the MTF and the Wiener spectrum and can incorporate the HVS as well. It produces results in bits/area that are directly related to the task of moving electronic image data from an input scanner to an output scanner or other display. Much of the research in this area began on photographic processes, but has also been applied to electronic scanned imaging. Both are addressed here. The basic equation for the spatial frequency based information content of an image is given¹⁴⁶ by

$$H_i = \frac{1}{2} \int_{-\infty}^{\infty} \log_2 \left[1 + \frac{\Phi_S(f)}{\Phi_N(f)} \right] df \quad (3.28)$$

where H_i is the information content of the image, Φ_s is the Wiener spectrum of the signal, Φ_N is the Wiener spectrum of the system noise, and f is the spatial frequency, usually given in cycles per millimeter. This equation is in one-dimensional form for simplicity, in order to develop the basic concepts. For images, these concepts must, as usual, be extended to two dimensions. Unlike the work dealing with the photographic image, the assumption of uniform isotropic performance cannot be used to simplify the notation to radial units. For digital images the separation of the orthogonal x - and y -dimensions of the image must be preserved.

Alternative methods for calculating information capacity do not include explicit spatial frequency dependence but do explicitly handle probabilities.^{46,50,51,147} They served as the basis for our discussion of quantization and Equation 3.29, is rewritten as

$$H_i = N \log_2(pM) \quad (3.29)$$

where N is the number of independent information storage cells per unit area. It may be set equal to the reciprocal of the smallest effective cell area of the image, for example, a number of pixels or the spread function. Here, p is the reliability with which one can distinguish the separate messages within an information cell, and M is the number of messages per cell. M is determined by the number of statistically different gray levels that can be distinguished in the presence of system noise at the reliability p , using noise measurements made with the above cell area.

Generalizing Equations 3.29 and 3.1 to the "generic" units of Figure 3.39, L is set equal to S_0 and σ_a is set equal to σ_s for the maximum signal and its standard deviation, respectively. We select a spread function for an *unaliased* system equal to 2×2 pixels and translate this to frequency space using the reciprocal of the sampling frequencies f_{sx} and f_{sy} in the x - and y -directions. This gives a generalized, sampling-oriented version of Equations 3.29 and 3.32 as

$$H_i = \frac{f_{sx} f_{sy}}{4} \log_2 \left[\frac{S_0}{k \sigma_s} \right] \quad (3.30)$$

where S and σ_s are measured in the same units. k can be set to determine the reliability for a given application. Values from 2^{45} to 20^{43} have been proposed for k for different applications; 6 is suggested here, making $p = 0.997$. This assumes that σ_s is a constant (i.e., additive noise) at all signal levels. If not, then the specific functional dependence of σ_s on S must be accounted for in determining the quantity in the brackets, measuring the desired number of standard deviations of the signal at each signal level over the entire range.^{46,51} While this approach predicts text quality and resolving power¹⁴⁷ and deals with the statistical nature of information, it does not (as noted above) permit the strong influence of spatial frequency to be handled explicitly.

Equation 3.31 may be expanded to illustrate the impact of the MTF on information content, giving

$$H_i = \frac{1}{2} \int_{-\infty}^{\infty} \log_2 \left[1 + \frac{K^2 \Phi_i(f) |\text{MTF}(f)|^2}{\Phi_N(f)} \right] df \quad (3.31)$$

where $\Phi_i(f)$ is the Wiener spectrum of the input scene or document and $MTF(f)$ is the MTF of the imaging system (assumed linear) with all its components cascaded. At this point we need to begin making some assumptions in order to carry the argument further. The constant K in the equation is actually the gain of the imaging system. It converts the units of the input spectrum into the same units as the spectral content of the noise in the denominator. For example, a reflectance spectrum for a document may be converted into gray levels by a K factor of 256 when a reflectance of unity (white level) corresponds to the 256th level of the digitized (8-bit) signal from a particular scanner; the noise spectrum is in units of gray levels squared.

Various authors have gained further insight into the use of these general equations. Some of those investigating photographic applications have extended their analysis to allow for the effect of the visual system,¹³⁷ others attempted to apply some rigor to the terms in the equation that are appropriate for digital imaging.^{148,149} Others have worked on image quality metrics for digitally derived images,¹³⁸ but some have tended to focus on the relationship to photointerpreter performance.¹⁵⁰

Several of these authors have suggested that properly executed digital imagery does not appear to be greatly different from standard analog imagery in terms of subjective quality or interpretability. One almost always sees these images using some analog reconstruction process to which many analog metrics apply. It therefore seems reasonable to combine some of this work into a single equation for image information and to hypothesize that it has some direct connection with overall image quality when applied to a scanner whose output is viewed or printed by an approximately linear display system. It must also be assumed that the display system noise and MTF are not significant factors or can be incorporated into the MTF and noise spectra by a single cascading process. A generic form of such an equation is given below as Equation 3.35 without the explicit functional dependencies on frequency in order to show and explain the principles that follow (expanding on the analysis in Reference 137).

$$H_i = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \log_2 \left[1 + \frac{K^2 \Phi_i |MTF|^2 R_1^2}{\left\{ 1 + 12(f_x^2 + f_y^2) \right\} \left\{ [\Phi_a + \Phi_n + \Phi_q] R_2^2 + \Phi_E \right\}} \right] df_x df_y. \quad (3.32)$$

Let us begin by examining the numerator. Several authors have attempted to multiply the MTF of the imaging system by a spatial frequency response function for the HVS to arrive at an appropriate weighting for the signal part of Equation 3.34. Kriss and his coworkers¹³⁷ observed that a substantial increase in the enhancement beyond the eye's peak response produced larger improvements in overall picture quality than did equivalent increases in enhancement at the peak of the eye's response. The pictures with large enhancement at the eye's peak response were "sharper," but were also judged to be too harsh. These results indicate that the HVS does not act as a passive filter and that it may weigh the spatial frequencies beyond the peak in the eye's response function more than those at the peak.

Lacking a good model for the visual system's adaptation to higher frequencies as described above, Kriss et al. proposed the use of the reciprocal of the eye frequency response curve as a weighting function, $R_1(f)$, that could be applied to the numerator. The conventional eye response $R_2(f)$ should be applied in the denominator to account for the perception of the noise, since the eye is not assumed to enhance noise but merely to filter it. The noise term, Φ_N , in the earlier equation has been replaced by the expression in the square brackets and multiplied by $R_2^2(f)$. The reciprocal response, R_1 , is set equal to 0.0 at 8 cycles/mm (200 cycles/in) in order to limit this function.

Next let us examine the noise effects themselves. A major observation is that noise in the visual system, within one octave of the signal's frequency, tends to affect that signal. It can be shown that the sum in the first curly brackets in the denominator of Equation 3.35 provides a weighting of noise frequencies appropriate to this one-octave frequency-selective model for the visual system.¹³⁷ Several authors^{116,126,133,151} describe frequency-selective models of the visual system. The present construct for noise perception was first described by Stromeyer and Julesz.¹⁵² A term for the Wiener spectrum of the noise in the visual process, Φ_{E_v} , has been added to the second factor in the denominator to account for yet one more source of noise. It is not multiplied by the frequency response of the eye, since it is generated after the frequency-dependent stage of the visual process.

The factor in square brackets in the denominator contains three terms unique to the digital imaging system.¹⁴⁹ These are Φ_a , the Wiener spectrum of the aliased information in the passband of interest; Φ_n , the Wiener spectrum of the noise in the electronic system, nominally considered to be fluctuations in the fast scan direction; and Φ_q the quantization noise determined by the number of bits used in the scanning process. We have thus combined in Equation 3.35 important information from photographic image quality studies, including vision models and psychophysical evaluation, with scanning parameters pertinent to electronic imaging.

The study of information capacity, information content, and related measures as a perceptual correlate to image quality for digital images is an ongoing activity. By necessity it is focused on specific types of imaging applications and observer types. For example, an excellent database of images and related experiments on quality metrics was built for aerial photography as used by photointerpreters.¹⁵⁰

Experiments correlating subjective quality scores with the logarithm of the basic information capacity, taking the log of H_i as defined in Equations 3.1 and 3.29 showed correlation of 0.87 and greater for subjective quality of pictorial images.¹⁵³ Specific MTF and quantization errors were studied. The results were normalized by the information content of the original.

By use of various new combinations of the same factors discussed above, it was possible to obtain even higher correlations. A digital quality factor was defined¹⁵³ as

$$\text{DQF} = \left[\frac{\int_0^{f_n} \text{MTF}_s(f) \text{MTF}_v(f) d(\log f)}{\int_0^{10} \text{MTF}_v(f) d(\log f)} \right] \times \log_2 \left[\frac{L}{L/M + 2\sigma} \right] \quad (3.33)$$

where we retain the one-dimensional frequency description used for simplicity by the original authors and the subscripts "s" and "v" refer to the system under test and the visual process, respectively. L is the density range of the output imaging process, f_n is the Nyquist frequency, M is the number of quantization levels, and σ is the RMS granularity of the digital image using a $10 \times 1000 \mu\text{m}$ microdensitometer slit. The first factor is related to the SQF described in Equation 3.27, and the second factor is related to the fundamental definition of image information capacity in Equation 3.29. A correlation coefficient of 0.971 was obtained for these experiments, using student observers and pictures showing a portrait together with various test patterns. It must be noted, however, that information capacity or any of these information-related metrics cannot be accepted, without psychophysical verification, as a general measure of image quality when different imaging systems or circumstances are to be compared.^{15,154} Since systems models are used to determine MTFs and information capacities and hence arrive at useful descriptions of technology variables,

these are good examples of the type A shortcut regression models described in Figure 3.3, but are restricted to the limitations of such regression shortcuts.

In conclusion, this brief overview of specific quality metrics should give the reader some perspectives on which ones may be best suited to his or her needs. The variety of these metrics, and the considerable differences among them, are evidence of the inherent diversity of imaging applications and requirements. Given this diversity, together with the large and rapidly expanding range of imaging technologies, it is hardly surprising that no single universal measure of quality has been found.

3.7 SPECIALIZED IMAGE PROCESSING

Most scanned images either begin or end in a digital form that needs to be efficiently managed in the larger context of a computer system, often in a network with other devices. This brings other dimensions to scanned image quality, namely the need to control the size of the files and the quality of the scanned images beyond the devices themselves. Controlling the file size is the subject of image compression.^{11,155-157} Compression is an image quality issue because several methods do so at the expense of image quality, with lossy compression being one example and reduced sampling versus increased gray resolution, that is, resolution enhancement,⁵⁵ being another. Finally there is the color management challenge: finding a method to ensure that a color scanned image created by any of a number of scanners will look well when printed on any of a number of differently designed or maintained color printing devices.^{6,14}

3.7.1 Lossy Compression

Image compression is a technology of finding efficient representation^{155,156} for digital images to:

1. Reduce the size and cost of on-board or off-line computer memory and disk drive space required for their storage;
2. Reduce the bandwidth and or time needed to manipulate, send, or receive images in a communication channel; and
3. Improve effective access time when reading from storage systems.

The need to improve storage is easily seen in the graphic arts business, where an 8.5 × 11 in, 600 × 600 in, 32 bit color image is approximately 10⁹ (or a billion) bits/image. Even the good quality portable amateur still cameras require 6+ megabytes (1 byte = 8 bits) per color image. Needless to say, transmitting such large files or accessing them takes tremendous amounts of time or bandwidth. Many standards groups are actively trying to create order out of the plethora of possible compression methods in order to reduce the number and types of tools needed to work in our highly interactive world of communications and networks.

There are generally two types of compression: lossless and lossy. Lossless takes advantage of better ways to encode highly redundant spatial or spectral information in the image, such as many contiguous white pixels in a text document. Results vary from compression ratios well over 100:1 on some text to 1.5:1 or less on many pictures. Group 3 and

Group 4 facsimile (“fax”) standards, established by the CCITT (Consultative Committee of the International Telephone and Telegraph), now ITU-T, the Telecommunication Standardization sector of the International Telecommunication Union) are perhaps the best known and apply only to binary images.¹⁵⁷ Other standards include JBIG,^{11,158} which is especially important for black and white halftones where it achieves about 8:1 compression while best CCITT methods actually expand file size by almost 20% over the uncompressed version.^{11,159} See Table 3.9 for links to these standards groups and the latest upgrades.

All compression involves several different operations from transformation of the data to allow for efficient coding (e.g., discrete cosine transform (DCT)) to the actual symbol-encoding step where many technologies have developed. The latter include Huffman,¹⁶⁰ LZ,^{161–163} and LZW¹⁶⁴ encoding, which are often cited as important parts of complete compression schemes.

Lossy compression is important from an image quality perspective since it removes information contained in the original image and therefore potentially causes a reduction in image quality to gain a compression advantage. Sometimes lossy compressions are said to be “visually lossless” in that they only give up information about the original that they claim cannot be detected by the HVS (recall the limits discussed in Section 3.2.1.3). Simply invoking binary imaging, for example, is an excellent method of compression, which is visually lossless when scanning ordinary black text on a white substrate at high resolution. It reduces a gray image from 8 bits to one and preserves all the edge information if it is high enough in resolution while throwing away all the useless gray levels in between. It does not work well on a photograph, where the primary information is in the tones that are all lost! Most lossy compression methods are very complex, involving advanced signal processing and information theory^{6,17,157} beyond the scope of this chapter.

The best-known lossy compression technique is called JPEG (after the Joint Photographic Experts Group formed under the joint direction of ISO-IEC/JTC1/SC2/WG10—see Table 3.9—and CCITT SGVIII NIC in 1986).¹¹ It is aimed at still-frame, continuous tone, monochrome, and color images. In the case of JPEG, the underlying algorithm is a DCT of the image one 8×8 pixel cell at a time. It then makes use of the frequency-dependent quantization sensitivity of the eye (Figure 3.11) to alter the quantization of the signal on a frequency-by-frequency basis within each cell.

Many lossy compression methods are adjustable depending on the users’ needs, so that the amount of compression is proportional to the amount of loss. They can be adjusted to a visually lossless state or to some acceptable state of degradation for a given user or design intent. The JPEG technique is adjustable by programming a table of coefficients in frequency space, called a Q table, which specifies the quantization at each of several spatial frequency bands. It can also be adjusted using a scaling factor applied to the Q table. Psychometric experiments (see Section 3.8) should be employed to determine acceptable performance in making such changes, using the exact scanning and marking methods and objects of interest. There are many other features of the JPEG approach that cannot be covered here. It has routinely been able to show an order of magnitude better compression over raw continuous tone pictures¹¹ with very little to no apparent visual loss of quality.

As noted earlier, compression is often aimed at improving communication of data and as such it is closely linked with file formats. In recent years a heavy focus on both the Internet and fax^{159,165} has led to significant progress. JPEG and GIF have become widely used in the Internet,¹⁶⁵ where, in a greatly simplified view, it is seen that the former is lossy in spatial terms while the latter is lossy in color terms.

Color fax standards^{159,166} have recently been developed in which the color, gray, and bitonal information is encoded into multiple layers for efficient transmission and compression.

These are formally known as TIFF-FX formats and generally fall into the broad category of mixed raster content or MRC.¹⁵⁹

A new standard, JPEG 2000¹⁶⁷ has been developed, which, in addition to several other improvements, utilizes wavelets as an underlying technology and includes several optional file formats called the JP family of file formats. One, the JPM file format¹⁶⁸ with extension .jpm is aimed at compression of compound images, those having multiple regions each with differing requirements for spatial resolution and tonality. It employs this multiple layer approach. MRC formats allow the optimization of image quality, color quality via good color management, and best compression, all in one package. The base mode of MRC decomposes a mixed content image into three layers: a bitonal (binary) Mask layer, and color Foreground, and Background layers. The wavelet approach in JPEG 2000 causes less objectionable artifacts than the DCT-based baseline JPEG.¹⁶⁸

3.7.2 Nonlinear Enhancement and Restoration of Digital Images

The characteristics of a scanned image may be altered in nonlinear ways to enable its portability between output devices of different resolutions while maintaining image quality and consistency of appearance. This may also be done to improve quality by reducing sampling effects or otherwise enhancing image appearance when compared to a straightforward display or print of the bit map. These are the general goals of digital image enhancement and restoration, topics that have been covered extensively in the literature and pursued by many imaging and printer corporations. They have been summarized by Loce and Dougherty.⁵⁵ Many of the techniques fall in the domain of morphological image processing,^{3,169} which treats images as collections of well-defined shapes and operates on them with other well-defined shapes. It is most often used with binary images where template matching, that is, finding an image shape that matches the filter shape and then changing some aspect of the image shape, is a good general example. Two particular examples illustrate some of the underlying concepts.

“Anti-aliasing” is a class of operations in which “jaggies” or staircases (i.e., sampling artifacts or “aliased” digital images of tilted lines) in binary images are reduced to a less objectionable visual form. In Figure 3.43 the staircased image of a narrow line is analyzed by a filter programmed to find the jaggies (template matching) and then operated on, pixel by pixel, to replace certain all-white or all-black edge pixels with new pixels, each at an appropriate level of gray, in this case one of three levels. The gray pixels may be printed using

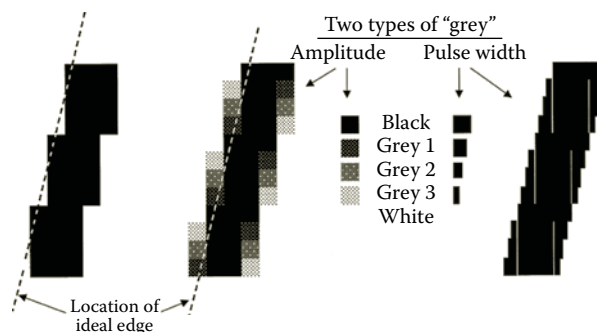


FIGURE 3.43

Anti-aliasing by the amplitude and pulse-width methods. Pixels narrowed by pulse-width changes are shown separated from the full pixels by a narrow white line only to illustrate where each is located.

conventional means of gray writing such as varying exposure on a continuous tone printing medium at the output stage. This may be thought of as amplitude modulation. A similar but often more satisfactory effect, producing sharper edges, can be achieved by using high addressability or pulse-width modulation in conjunction with printing processes having an inherently sharp exposure threshold rather than continuous tone response.

Methods to evaluate the prints to determine the reduction of the appearance of the jaggies involve scanning along the edge of a line containing the effects of interest with a long microdensitometer slit whose length covers the space from the middle of the black line to the clear white surround. The resulting reflection profile is proportional to the excursions of the edge. It indicates the additional effects of the printing and measurement processes on decreasing or increasing the jaggies and can be analyzed for its visually significant components against an appropriate CSF.¹⁷⁰ Some of these components are random based on the marking process, others are periodic based on the angle of the line and the resulting frequency of the staircase effect.

Figure 3.44 shows an example of several practical effects of such enhancement and restoration on an italic letter “b.” The upper figure is a representation of a conventional bit map of the original computer generated letter. Note the jaggies or staircase on the straight but tilted stroke at the left and a variety of undesirable effects throughout the character. Using the observation window employed by Hewlett Packard’s RET (Resolution Enhancement Technology)^{55,171,172} as shown on the left, roughly 200 pixel-based templates are compared to the surrounding pixels for each individual pixel in the original “b,” a part of which is shown here. A decision is made regarding how large a mark, if any, should replace that pixel, based on a series of rules developed for a particular enhancement scheme, in this case the RET algorithm. The mark in this case is created by modifying the width of the pulse in the horizontal dimension as illustrated. The resulting map of full and width-modulated pixels is shown in the lower part of the letter “b.” Note that some of the narrow pixels can be positioned left or right. This is called pulse-width position modulation, PWPM.

When the individual pulses are blurred and developed by the marking process, they will tend to merge into the body of the letter, both physically and visually, to produce even

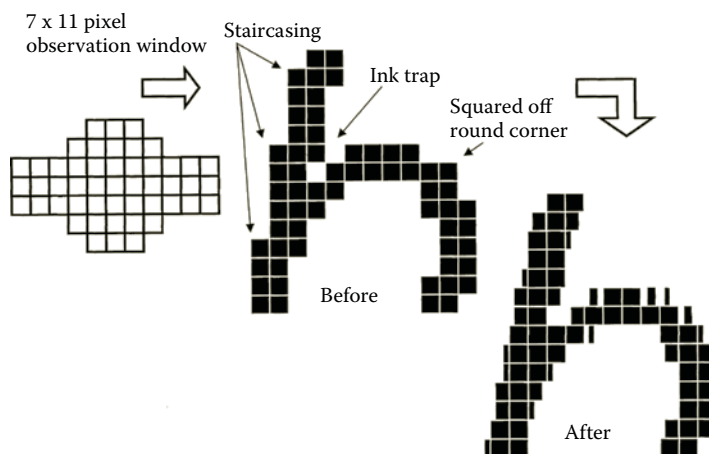


FIGURE 3.44

An example of resolution enhancement on a portion of an italic letter “b.” (Adapted from Tong, C., *Resolution enhancement in laser printers, Proceedings of SPIE conferences on color Imaging: Device-Independent color, Color Hardcopy and Graphic Arts II*, San Jose, CA, 1997.

smoother edges than shown as a bit map image here. There are many similar techniques patented prior to and following the above and sold by other companies such as Xerox,¹⁷³ IBM,¹⁷⁴ Destiny,¹⁷⁵ and DP-Tek,¹⁷⁸ now owned by Hewlett Packard, to name only a few, each with its own special features. They all create the effect on the HVS equivalent to that provided by a higher resolution print and many enhance the images in other ways as well, such as removing ink traps or sharpening the ends of tapered serifs.⁵⁵

3.7.3 Color Management

Color measurement systems, as discussed earlier in Section 3.2.2.5, are the key to managing color reproduction in any situation. The advances in scanned color imaging systems that separated input and output scanning devices and inserted networks, electronic image archives, monitors and preprinting (prepress) software in between have made it desirable to automate the management of accurate or pleasing (not the same) color reproduction. This in turn has meant automating or at least standardizing and carefully controlling the objective measurement of the color performance of the many input, output, and image manipulation devices and a variety of methods for insuring consistency.^{6,14}

The basic concept is to encode, transmit, store, and manipulate images in a device-independent form, carrying along additional information to enable decoding the files at the step just before rendering to an output device, that is, just before making it device specific. CIELAB (i.e., $L^*a^*b^*$) based reference color space, above, is commonly used to relate characteristics of both of these types of devices to an objective standard. Standardized operating system software, operating with standardized tools and files accomplished this, but is beyond the scope of this chapter. See References 29 and 30 as well as 6-ch 4, and the papers cited in them for more details and Reference 23 for a practical guide to using the tools that are available at this time.

Today, a common ANSI standard target known as IT-8.7 (see Figure 3.49) is manufactured by Kodak (shown), Fuji, and Agfa, each using their own photographic dyes. It is scanned by the scanner of interest into a file of red, green, and blue (RGB) pixels. It is also measured with a spectrophotometer to determine the CIE $L^*a^*b^*$ values for all 264 patches. Color management software then compares both results and constructs a *source profile* of the scanner color performance.

A well-known example of a color management system is the approach organized by the International Color Consortium (ICC). See Table 3.9. It has created a standard attempting to serve as a cross-platform device profile format to be used to characterize color devices. It enables the device-independent encoding and decoding primarily developed for the printing and prepress industries, but allows for many solutions providers.

This profile, often called an "ICC profile" if it follows the Consortium's proforma, is a lookup table that is carried with all RGB files made with that scanner. It is useful for correction as long as nothing changes in the scanner performance or setup. Similarly, a destination profile is created, typically for a printer or a monitor. Here, known computer-generated patterns of color patches are displayed or printed, and measured with a spectrophotometer in $L^*a^*b^*$. Again, a comparison between the known input and the output is performed by the color management software, which creates a lookup table as a *destination profile*.

The color management architecture incorporates two parts. The first part is the *profiles* as described above. They contain signal processing transforms plus other material and data concerning the transforms and the device. Profiles provide the information necessary to convert device color values to and from values expressed in a color space known as a *profile connection space* ($L^*a^*b^*$ in the ICC example). The second basic part is the *color management module* (CMM), which does the signal processing of the image data using the profiles.

Progress in color management and the ICC in particular have pulled together an important set of structures and guidelines.^{6,14} These enable an open color management architecture that has made major improvements. Of course, gamut differences like those in Figure 3.18, are not a problem that color management, *per se*, can ever solve. It is also important to note that drift in the device characteristics between profile calibrations cannot be removed. It is reported¹⁷⁷ that (averaging over a wide range of colors) rotogravure images in a long run show $\Delta E_{ab}^* = 3.0$ and for offset $\Delta E_{ab}^* = 5.5$, (i.e., the range for 90% of images) while they report for *input scanners* $\Delta E_{ab}^* = 0.4$. They also report that the use of color management and ICC profiles improved system results from $\Delta E_{ab}^* = 9$ down to 5, and suggest in general, with good processes, that this is inherently as good as one can achieve. Similarly, Chung and Kuo¹⁸² found they could achieve an $\Delta E_{ab}^* = 6.5$ as the average for the best scenario in color matching experiments using ICC profiles for a graphic arts application. Control over specific limited sets of system components, colors or small color ranges as well as newer measurement technologies can show much tighter tolerances than these. There is still a great deal of analysis and work that must be carried out to make color management more universal, easier, and more successful.^{176,179–181}

3.8 PSYCHOMETRIC MEASUREMENT METHODS USED TO EVALUATE IMAGE QUALITY

3.8.1 Relationships between Psychophysics, Customer Research, and Psychometric Scaling

As one attempts to develop a scanned imaging system, there are usually some image quality questions that cannot be answered by previous experience or by reference to the literature. Often this reduces to a question of determining quantitatively how “something” new *looks visually* for “some task.” It is a problem because no one else has ever evaluated the “something” or never used it for “some task” or both. We give the reader at least some pointers to the basic visual scaling discipline and tools to attack his own specialized problems.

As the Image Quality Circle^{12,40} and the full framework in Figure 3.3 indicates, there are many places where one needs to quantify the human visual responses. Sometimes this is in the short-cut paths connecting technology variables (the “something”) directly to customer quality preferences for “some tasks” through *customer research*. Sometimes, it is in creating a more thorough understanding by developing visual algorithms, which connect the physical image parameters, that is, attributes (other types of “something”), with the fundamental human perceptions of these attributes. The science of developing these latter connections is referred to as *psychophysics*. The underlying discipline for doing both engineering-oriented customer research and psychophysics is psychometric *psychometric scaling*. Hundreds of good technical papers, chapters, and whole books have been written on these subjects, but are often overlooked in imaging science and engineering for a variety of reasons. Many of the papers cited in this chapter draw on the rich resources of psychometric scaling disciplines in certain large corporations, government agencies, and universities to develop their algorithms. Engeldrum¹² has recently distilled many of the basic disciplines and compiled many of the classic references into a useful book and software toolkit for imaging systems development.

3.8.2 Psychometric Methods

There are many classes of psychometric evaluation methods, the selection of which depends on the nature of the imaging variable and the purpose of the evaluation. We can only describe them at a high level in this section. Figure 3.45 describes a framework for considering psychometric experiments, starting with two fundamental purposes, at the left, each of which breaks down into three basic approaches and six types of data.

The way in which the sample preparation is done, observer (called “respondent” in market research) quantity and selection methods, and the numbers of images shown can all be very different, depending on the purpose. In general the customer-user experiments require significantly more care in all areas, are restricted to user-like displays of relatively few images, and require several dozen to hundreds of respondents. They tend to focus on quantifying the “Customer Quality Preference” block in Figure 3.3.

Visual sciences experiments on psychophysics and perception are useful for developing the image quality models and especially the visual algorithms of Figure 3.3 and the comparisons between the HVS and measurements indicated in Figure 3.2. Here smaller numbers of observers, from a few to several dozen, are often deemed adequate. These observers are often experts or technical personnel and can be told to overlook certain defects in samples and concentrate on the visual characteristic of interest. Such observers can be asked to try more fatiguing experiments. These are often broken into several visits to the laboratory, something not possible with customer research.

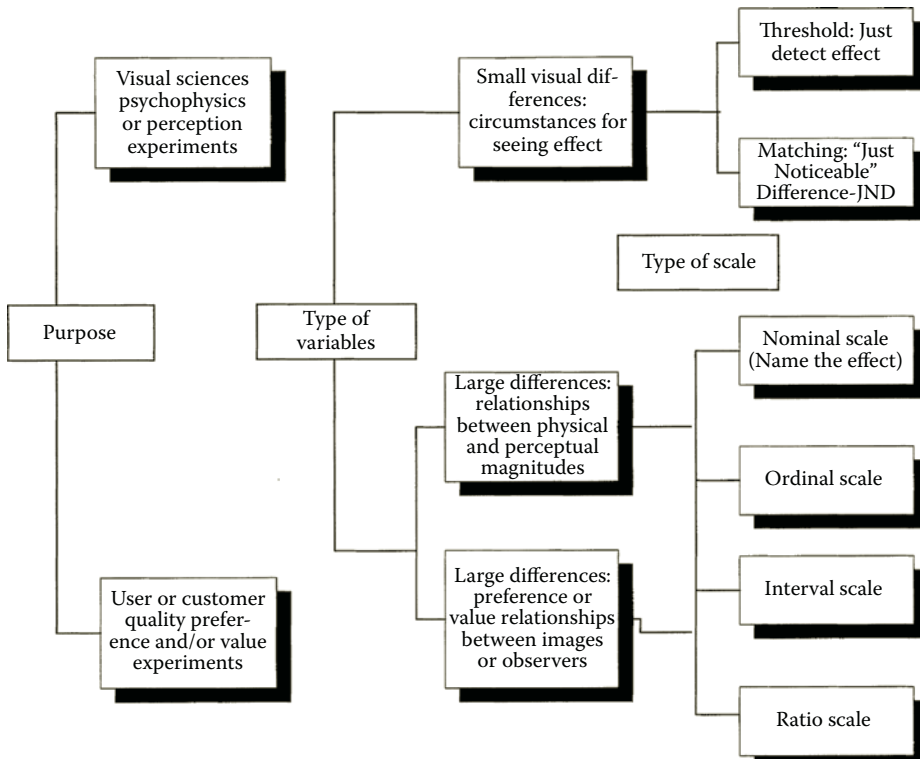


FIGURE 3.45 Psychometric experiments for diverse purposes, grouped in two classes here, can be further classified into three types of variables, which in turn lead to a few basic but significantly different types of scales.

In general, experiments with good statistical design should be used, in which a targeted confidence level is established. It is common practice in many customer and general experiments to seek 95% confidence intervals (any basic statistics book¹⁸³ will provide equations and tables to enable this, provided the scaling method is properly classified as shown below). This requires estimating the size of the standard deviation between observers and using it along with the confidence interval equation to determine the number of independent observations that translate to the number of observers. The experimenters in visual sciences can use fewer observers than a customer researcher because the visual sciences use variables and trained observers that have much better agreement, that is, smaller standard deviations. Also, these experimenters may require less statistical confidence because they are often more willing to use other technical judgment factors such as models and inferences from other work.

For either purpose, the decisions regarding basic approach must be determined by looking at the types of variables and the types of scales to be built. If the goal is to determine when some small signal or defect (such as a faint streak) is just visually detected, *threshold* scales are developed. They show the probability of detection compared to the physical attribute(s) of the image samples or the observation variables. One may wish to compare readily visible signals, such as images of well-resolved lines, trying to distinguish when one is just visually darker than another. This involves determining the probability, in *matching* experiments, of what levels of a variable(s) cause two images to be seen as *just noticeably different (JND)*, that is, just do not match each other. Dvorak and Hamerly¹⁸⁴ and Hamerly¹⁸⁵ give examples of JND scaling for text and solid area image qualities.

These experiments often explore fundamental mechanisms of vision and can draw on a relatively small number of observers in well-controlled experimental situations using electronic displays with side-by-side image comparisons. The temptation to substitute an easily controlled electronic display experiment for one in which the imaging media is identical to the actual images of interest (e.g., photographic transparencies viewed by projection, or xerographic prints viewed in office light) must be carefully weighed in each situation. Various forms of image noise, display factors affecting human vision (especially adaptation), as well as visual and psychological reference cues picked up from the surround, are often important enough to outweigh the ease of electronic display methods.

When the magnitude of visible variables is large and the goal is to compare quality attributes over a large range, as in the bottom two “variables” boxes in Figure 3.45, then a decision about the mathematical nature of the desired scale and the general nature and difficulty of the experimental procedure becomes important. The four basic types of scales¹⁸⁶ shown here were developed by Stevens¹⁸⁷⁻¹⁸⁹ and are shown in increasing order of “mathematical power” in Figure 3.45 and with very short descriptions in Table 3.5.

There is an abundance of literature on the theory and application of scaling methods,¹² some of which are indicated in the table as column headings. Additional general references include References 190–199. Below is a very brief summary of the methods to assist the reader in beginning to sort through these choices. Here we assume the samples are “images,” but they could just as well be patches of colored chips, displays on a monitor, pages of text, or any other sensory stimulus.

3.8.3 Scaling Techniques

The methods by which the various types of scales (Table 3.5) may be constructed are listed with very brief descriptions and the type of scale they may be used to construct in (). Where only complicated procedures enable a type of scale to be derived it is unbolded.

TABLE 3.5

Types of Scales

Type of Scale	Description and Analysis Operations
Nominal	Names of categories/classes
Ordinal	Ordered along variables, determines "greater than" or "less than," (gives arbitrary/unknown distances on the variable scale)
Interval	Ordinal scale + magnitude of differences are quantified. $y = ax + b$ (Equality of intervals may be determined and any linear transformation is OK. Mean, standard deviation, coefficient of correlation are valid. Basis for much image quality analysis)
Ratio	Interval scale where "none" of an attribute is assigned 0 response. $y = ax$ (i.e., $b = 0$) (Interval scale operations are OK and coefficient of variation and equality of ratios are valid—many lightness scales are an example with an absolute zero)

Engeldrum¹² gives a good discussion of all these methods. References using these methods are also listed for each.

3.8.3.1 Identification (Nominal)

In this simple scaling method, observers group images by identifying names for some attributes and collecting images with those attributes. The resulting nominal scales are useful in organizing collections of images into manageable categories.

3.8.3.2 Rank Order (Ordinal)

Observers arrange a set of images according to decreasing or increasing amount of the perceived attribute.^{12,188,190,194} A median score for the group is frequently used to select the rank for each sample. Agreement between observers can be tested to understand the nature of the data by calculating the coefficient of concordance or the rank order coefficient.¹⁹⁷

3.8.3.3 Category (Nominal, Ordinal, Interval)

Observers simply separate the images into various categories of the attribute of interest, often by sorting into labeled piles. This is useful for a large number of images, many of which are fairly close in attributes, so that there are some differences of opinion over observers or over time as to which category is selected. Interval scales can be obtained if the samples can be assumed to be normally distributed on the perceived attributed.¹⁹⁹

3.8.3.4 Graphical Rating (Interval)

The observers score the magnitude of the image attribute of interest by placing an indicator on a short line scale that has defined endpoints for that attribute. The mean of the positions on the scale for all observers is used to get a score for each image.¹²

3.8.3.5 Paired Comparison (Ordinal, Interval, Ratio)

All images are presented to all observers in all possible pairwise combinations, usually one pair at a time, sometimes with a reference. The observer selects one of the pair as having

more of the attribute of interest. If there are N different images then there are $N(N - 1)/2$ pairs. The proportion of observers for which each particular image is selected over each other image is arrayed in a matrix. The average score for each image (i.e., any column in the matrix) is then computed to determine an ordinal scale.^{12,191,197,206} If it is assumed that the perceived attributes are normally distributed, then, as with the category method, an interval scale can be determined. This is done using Thurstone's Law of Comparative Judgement¹⁹⁹ in which six types of conditions for standard deviations describing the datasets are used to construct tables of Z-Deviates,^{12,197} from which interval scales are directly obtained.

3.8.3.6 Partition Scaling (Interval)

The observer is given two samples, say S1 and S9, and asked to pick a third sample from the set, whose magnitude of the appearance variable under test is halfway between the two samples; call it S5 in this case. Next he finds a sample halfway between S1 and S5, call it S3, then he finds one between S5 and S9, calling it S7, and so on, until he has built a complete interval scale using as many samples and as fine a scale as desired.⁴

3.8.3.7 Magnitude Estimation (Interval, Ratio)

The observer is asked to directly score each sample for the magnitude of the attribute of interest.^{12,187,189,194,195} Often, the observer is given a reference image at the beginning of his scoring process, called an anchor, whose attribute of interest is identified with a moderately high, easy to remember score, such as 100. His scores are based on the reference and he is coached in various ways to use values that reflect ratios. This process implies that a zero attribute gets a zero response and hence generates a ratio scale. However actual observations sometimes are more in line with an interval scale, and this needs to be checked after the test.

3.8.3.8 Ratio Estimation (Ratio)

This test may be done by selecting samples that bear specific ratios to a reference image. The experimenter does *not* assign a value to the reference. Alternatively, the observers may be shown two or more specific images at a time and asked to state the apparent ratios between them for the attribute of interest.^{4,195}

3.8.3.9 Semantic Differential (Ordinal, Interval)

Typically used for customer research.¹⁹⁰ The image attributes of interest are selected and a set of bipolar adjectives is developed for the attributes. For example, if the attribute class were tone reproduction, the adjectives could be such pairs as darker–lighter, high contrast–low contrast, good shadow detail–poor shadow detail. Each image in the experiment is then rated on a several point scale between each of the pairs. Each scale is treated as an interval scale and the respondents' scores for each image and each adjective pair are averaged. A profile is then displayed.

3.8.3.10 Likert Method (Ordinal)

Typically used for customer research and attitude surveys. A series of statements about the image quality attributes of a set of images is provided (e.g., "The overall tones are perfect in

this images," "the details in the dark parts of this image are very clear"). The respondents are then asked to rate each statement on the basis of the strength of their personal feelings about it: strongly agree (+2); agree (+1); indifferent (0); disagree (-1); strongly disagree (-2). Note signs on numbers reverse for negative statements. The statements used in the survey are often selected from a larger list of customer statements. A previous set of judges maybe was used to determine those statements that produce the greatest agreement in terms of scores assigned to this set of images.

3.8.3.11 Hybrids (Ordinal, Interval, Ratio)

There are many approaches that combine the better features of these different methods to enable handling different experimental constraints and obtaining more accurate or more precise results. A few are noted here:

1. *Paired Comparison for Ratio*: Paired comparisons reduced to an interval scale that is fairly precise and transformed to accommodate a separate ratio technique (accurate but less precise) to set a zero. This gives a highly precise ratio scale.¹⁹⁷
2. *Paired Comparison Plus Category*: The quality of each paired comparison is evaluated by the observer, using something like a Likert scale below. A seven-level scale from strongly prefer "left" (e.g., +3) to strongly prefer "right" (e.g., -3) is used.²⁰⁶
3. *Paired Comparison Plus Distance* using distance (e.g., linear scale on a piece of paper) to rate the magnitude of the difference between each pair, giving the same information as the graphical rating methods discussed earlier, but with the added precision of paired comparison.
4. *Likert and Special Categories*: A variety of nine-point symmetrical (about a center point) word scales can provide categories of preferences that are thought to be of equal intervals. One scale attributed to Bartleson¹⁹⁴ goes from: Least imaginable "... ness" → very little "... ness" → mild "... ness" → moderate "... ness" → average "... ness" → moderate high "... ness" → high "... ness" → very high "... ness" → highest imaginable "... ness." Another similar scale is 1 = Bad, 2 = Poor, 3 = Fair, 4 = Good, 5 = Excellent. Many other such scales are found in the literature.

3.8.4 Practical Experimental Matters Including Statistics

Each of these techniques has been used in many imaging studies, each with special mathematical and procedural variations well beyond the scope of this chapter. A short list of common procedural concerns is given in Table 3.6 (from literature^{4,10,14,29} plus a few from the authors' experience).

The * items represent a dozen practical factors that must *always* be considered in designing nearly any major experiment on image quality or attributes of images.

The statistical significance of the results are often overlooked but cannot be stressed enough. For an interval scaling experiment that samples a continuous variable like darkness, standard deviations and means and subsequent confidence intervals on the responses can be calculated in straightforward ways to determine if the appearances of two samples are statistically different or to determine the quality of a curve fit. (See any statistics book on the confidence interval for two means given an estimate of the standard deviations for each sample's score, or to determine the confidence for a regression.)

TABLE 3.6

Factors that should be Considered in Designing Nearly Any Major Experiment on Image Quality or on Attributes of Images

Most important	Important
Complexity of observer task ^a	State of adaptation
Duration of observation sessions ^a	Background conditions
Illumination level ^a	Cognitive factors (many)
Image content ^a	Context
Instructions ^a	Control and history of eye movements
Not leading the observer in preference experiments ^a	Controls
Number of images ^a	Feedback (positive and negative effects)
Number of observers ^a	Illumination color
Observer experience ^a	Illumination geometry
Rewards ^a	Number of observation sessions
Sample mounting/presentation/identification methods ^a	Observer acuity
Statistical significance of results ^a	Observer age
	Observer motivation
	Range effects
	Regression effects
	Repetition rate
	Screening for color vision deficiencies
	Surround conditions
	Unwanted learning during the experiment

^a See text.

In detection experiments it is often desired to know if two scanned images, which gave two different percentages of observers who saw a defect or an attribute, are significantly different from each other (market researchers call such experiments sampling for attributes). This involves computing confidence intervals for proportions and therefore estimating standard errors for proportions, a procedure less commonly encountered in engineering. If p = the fraction of observers detecting an attribute, q = the fraction not detecting an attribute (note $p + q = 1.0$) and n = number of observers, assuming n is a very small fraction of the population being sampled, then the standard error for proportions is

$$S_p = [(p \times q)/n]^{0.5} \quad (3.34)$$

and the confidence interval around p is

$$CI = Z \times S_p \quad (3.35)$$

where, for example, $Z = 1.96$ for 95% confidence and 1.28 for 80%. A few cases are illustrated in Table 3.7 to give the reader perspective on the precision of such experiments and the number of observers required. The first column shows the value of p , the fraction of observers finding the attribute of interest. The second column gives the confidence desired in % where 95 is common in many experiments, and 80 is about the lowest confidence cited

TABLE 3.7

Confidence Intervals Around p for Attribute Data from Statistics of Proportions

<i>P</i>	% Confidence	<i>n</i> = 4	<i>n</i> = 8	<i>n</i> = 20	<i>n</i> = 100	<i>n</i> = 500
0.99	95	<i>0.10</i>	<i>0.07</i>	<i>0.04</i>	0.02	0.01
	80	<i>0.06</i>	<i>0.05</i>	<i>0.03</i>	0.01	0.005
0.95	95	<i>0.21</i>	<i>0.15</i>	0.10	0.04	0.02
	80	<i>0.13</i>	<i>0.09</i>	0.06	0.03	0.01
0.90	95	<i>0.29</i>	0.21	0.13	0.06	0.03
	80	<i>0.17</i>	0.12	0.08	0.04	0.02
0.80	95	0.39	0.28	0.18	0.08	0.03
	80	0.23	0.17	0.11	0.05	0.02
0.60	95	0.48	0.34	0.22	0.10	0.04
	80	0.31	0.21	0.13	0.06	0.02
0.50	95	0.49	0.35	0.22	0.10	0.04
	80	0.32	0.23	0.14	0.06	0.03

Statistical uncertainties in experimental results for proportion data (e.g., percentages of “yes” or “no” answers). Table entries gives 80% and 95% confidence as one-sided confidence intervals, that is, positive or negative deviation from the p value in column 1, at a few percentages of positive responses “ p ” (row headings) and a few numbers of respondents “ n ” (i.e., sizes of groups interviewed) as column headings. Italic unbolded entries are for p values that cannot be realized or closely approximated with the associated n values.

in many texts and statistical tables. The numbers reported in the table are the deviations about the fraction in column one that constitute the confidence interval for the population of all observers that would detect the attribute. The unbolded italic numbers correspond to values of p that cannot be realized by straightforward means for an observer population as small as indicated (e.g., a “ p ” value of 0.99 could not be observed with only four people—it would take 100!). As an example, for a sample with an attribute that was seen 90% of the time by a sample of 20 observers, one can be 80% confident that 82% to 98% (0.90 ± 0.08) of all observers would see this attribute. One would also be 95% confident that between 77% and 100% (numerically 103%, which here is equivalent to 100%) would see it.

3.9 REFERENCE DATA AND CHARTS

The following pages are a collection of charts, graphs, nomograms, and reference tables, which, along with several earlier ones, the authors find useful in applying first-order analyses to many image quality engineering problems. Needless to say, a small library of computer tools covering the same material would provide a useful package. In addition to those in this section, there are a few graphs, charts, and tables of value to engineering projects included in the text where their tutorial value was considered more important.

These include Figure 3.17 on CIE standard observer color matching function, Figure 3.20 on scan frequency effects, Figure 3.22 on nonuniformity guidelines, Figures 3.30 and 3.31 on MTF, and, finally, Figure 3.37 on edge noise calculations. The tables include Table 3.1 on halftone calculations, Table 3.3, which serves as a directory to Figures 3.50 to 3.53 in this section, and Table 3.7, giving confidence intervals for proportions.

In this section additional graphs on basic colorimetry are provided as Figure 3.46 for a more precise x, y chromaticity diagram with a dominant wavelength example and some

standard light sources, and Figure 3.47 for spectral characteristic of four standard light sources. Table 3.8 gives useful conversions between imaging oriented variables of density or reflectance and colorimetry units of L^* (see also Figure 3.19).

Figures 3.48–3.50 show, through annotations, the important structures in useful industry standard test patterns, two each for monochrome in Figure 3.48 and 3.50 and one for color in Figure 3.49 (reproduced here in black and white). Figure 3.48a is the monochrome ISO 12233 reflection test chart for testing digital still cameras which is also well suited to testing flat platen scanners for detail rendition and resolving power. Figure 3.48b is a monochrome test pattern suited for measuring pictorial monochrome scanners using methods referred to in ISO 16067 (see Table 3.9). Unlike Figure 3.48a which is two-level pattern, it contains a range of gray information to allow mapping tonal response and MTF analysis. Notice that both have many tilted edges and lines to show a wide range of sampling phases as well as untilted ones. Bar pattern and edge patterns are suited to software analysis of spatial frequency response using gray response while resolving power targets can be measured directly and visually with a monitor under magnification.

Next are some useful MTF equations and their corresponding graphs (plotted in log–log form for easy graphical cascading).

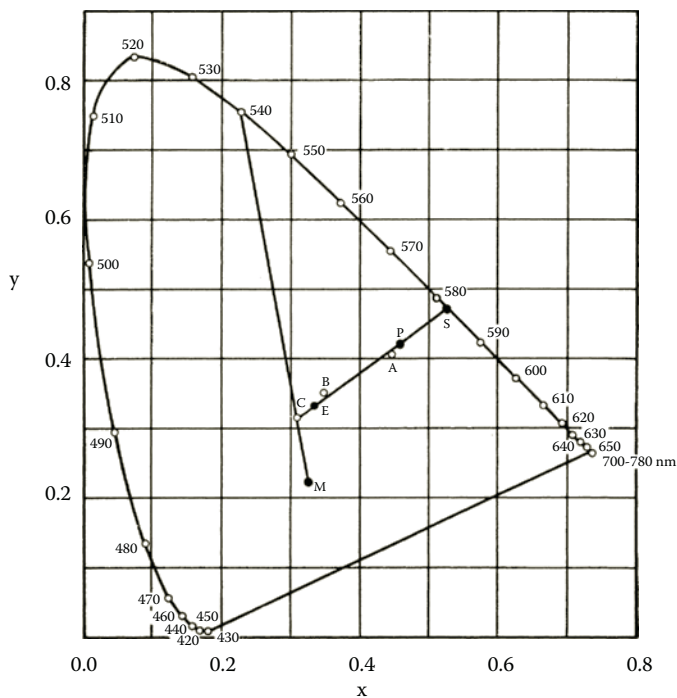


FIGURE 3.46

Dominant wavelength and purity plotted on the CIE x, y chromaticity diagram. The dominant wavelength for point P under illuminant C is found by drawing a straight line from the illuminant C point through P to the spectrum locus, where it intersects at 582 nm, the dominant wavelength. Excitation purity is the percentage defined by CP/CS , the percentage the distance from illuminant C to P is of the total distance from illuminant C to spectrum locus. Standard illuminants A, B, and E are also shown. See Figure 3.47 for the relative spectral power distributions of A, B, and C. E has equal amounts of radiation in equal intervals of wavelength throughout the spectrum. (From Hunter, R.S.; Harold, R.W. *The Measurement of Appearance*, 2nd Ed.; John Wiley and Sons: New York, 1987; 191; reproduced with permission of John Wiley & Sons, Inc.).

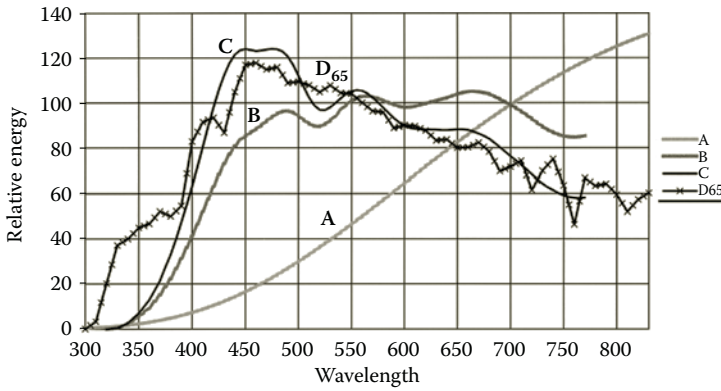


FIGURE 3.47
Standard illuminants A, B, C, and D₆₅ showing relative spectral energy distribution. Wavelength is in nm.

TABLE 3.8

An Abbreviated Conversion Table for Density, % Reflectance (%Y/
Y_m × 100) and L* Covering Densities to 4

Density	%Y	L*	Density	Y	L*
0	100	100	1.4	3.98	23.61
0.05	89.1	95.62	1.5	3.16	20.67
0.1	79.4	91.41	1.6	2.51	17.96
0.15	70.8	87.39	1.7	2	15.49
0.2	63.1	83.49	1.8	1.58	13.11
0.25	56.2	79.73	1.9	1.26	10.99
0.3	50.1	76.13	2	1	8.99
0.4	39.8	69.33	2.2	0.631	5.7
0.5	31.6	63.01	2.4	0.398	3.59
0.6	25.1	57.17	2.6	0.251	2.27
0.7	20	51.84	2.8	0.158	1.43
0.735	18.4	50	3	0.1	0.9
0.8	15.8	46.71	3.2	0.063	0.57
0.9	12.6	42.15	3.4	0.04	0.36
1	10	37.84	3.6	0.025	0.23
1.1	7.94	33.86	3.8	0.016	0.14
1.2	6.31	30.18	4	0.01	0.09
1.3	5.01	26.76			

Figure 3.51 is the MTF of two uniform, sharply bounded spread functions. The MTF of a uniform disc point spread function is defined as

$$T(N) = \frac{2J_1(Z)}{Z} \tag{3.36}$$

where $Z = \pi DN$, $N = \text{cycles/mm}$, and $D = \text{diameter of disk in mm}$.

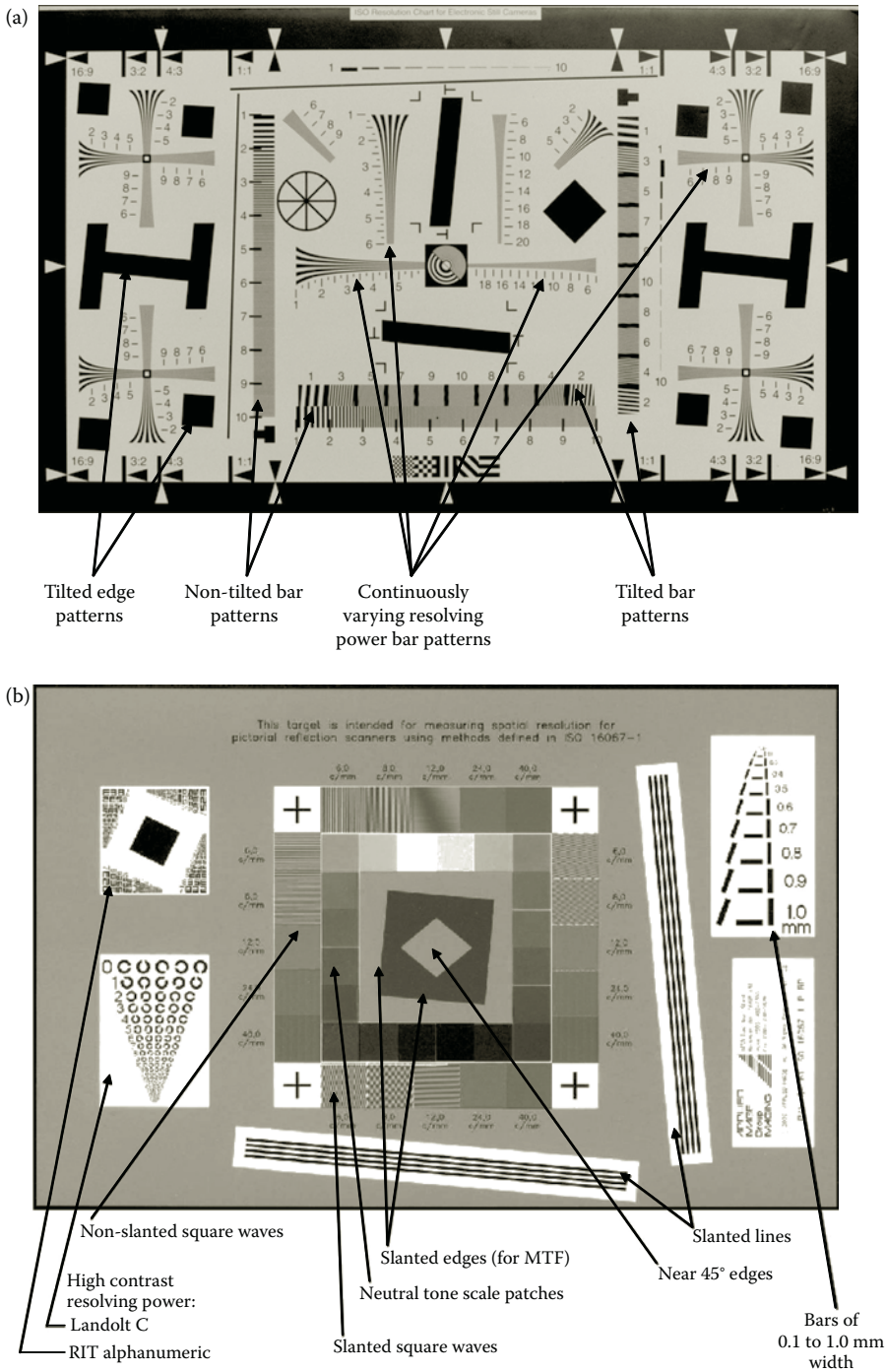


FIGURE 3.48

(a) Digital resolution target specified for digital cameras by ISO12233 which is also useful for testing scanners. (b) Applied Image¹⁰⁶ version of the ISO target for measuring spatial resolution for pictorial reflection scanners using methods defined in ISO16067-1. See Table 3.9 for pointers to standards. (Do not use this printed reproduction for testing, it is considerably degraded.)

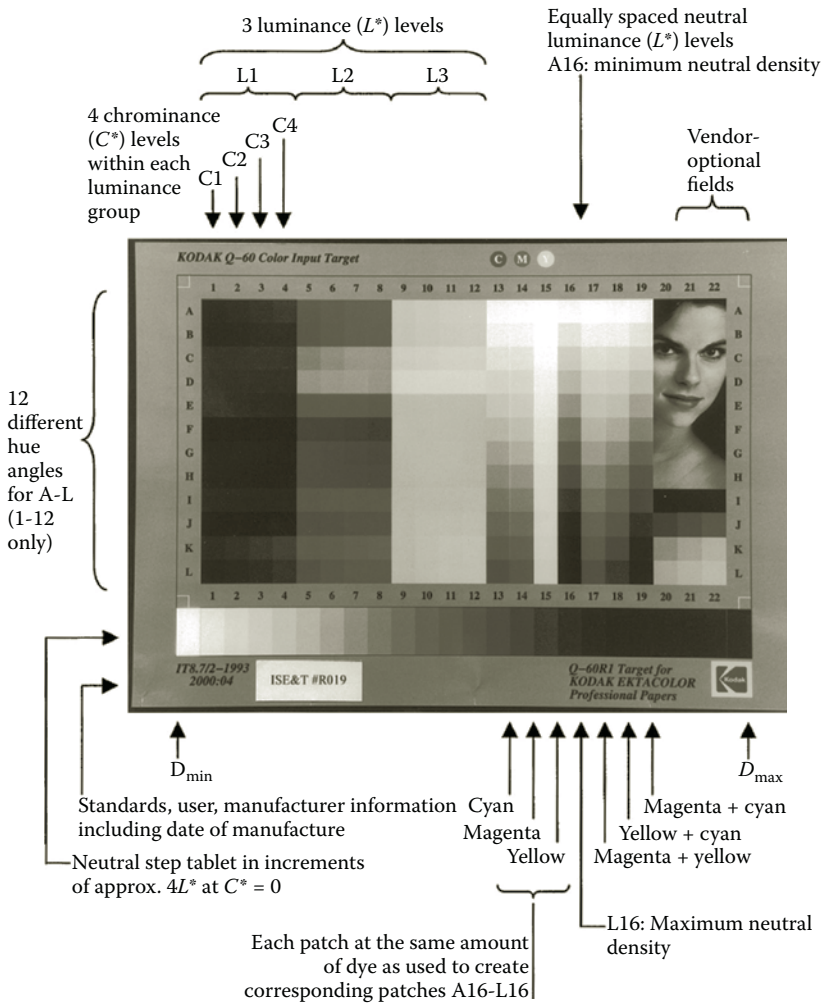


FIGURE 3.49

Layout of the IT8.7/1 (transmissive) and IT8.7/2 (reflective) scanner characterization targets. Details of colors are described in Table 5-1, 5-2, and 5-3 of Reference 23, or ISO IT8.7/1 and 2-1993 (from Reference 23). (Note: Do not attempt to use this reproduction as a test pattern.)

The MTF of a uniform slit or uniform image motion is defined as

$$T_{\text{slit}}(N) = \frac{\sin pDN}{pDN} \tag{3.37}$$

where D = width of slit in mm (or width of rectangular aperture or length of motion during image time) and N = cycles/mm.

Figure 3.52 is the MTF of a Gaussian spread function $S(r)$

$$T(N) = e^{-a^2N^2} \tag{3.38}$$

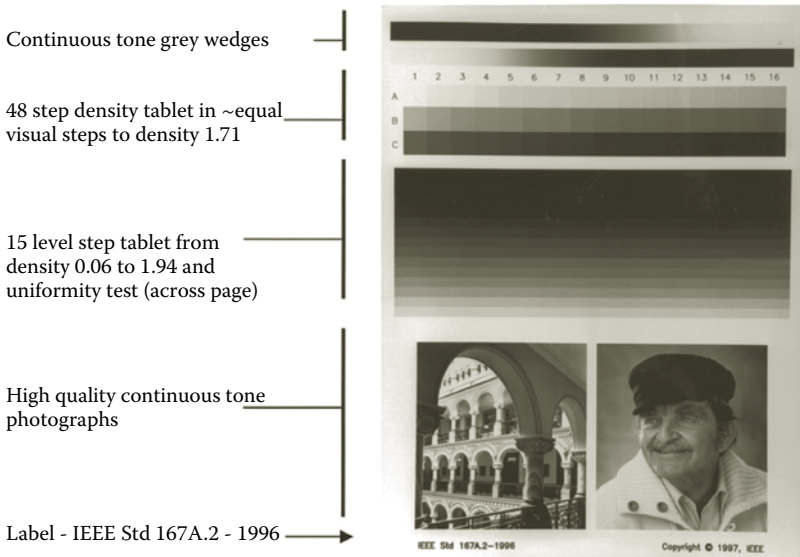
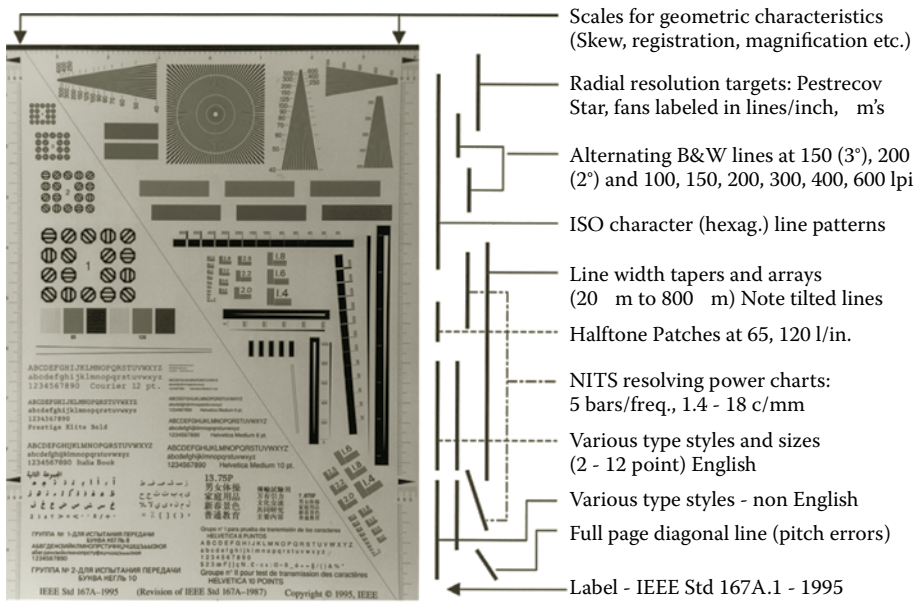


FIGURE 3.50

Images of two IEEE Standard Facsimile Test Charts (now “withdrawn” but still in use) which contain many elements valuable in assessing performance of scanning systems; (a) (top) IEEE Std. 167A. 1-1995-Bi-Level (black and white) chart, (b) (lower) IEEE Std. 167A.2-1996, High Contrast (gray scale) chart printed on glossy photographic paper. To identify what test pattern element each annotation refers to, project the relative vertical position of the bar in the specific annotation horizontally across the image of the test pattern. The bars are arranged from left to right in sequence. A composite using many parts of both (a) and (b) plus other elements is available today as the Eastman Kodak/Digital Science Imaging Test Chart (TL. 5003) from Applied Image Corp¹⁰⁶ under keyword Q4.60. See Figure 3.34 for other resolving power targets and Table 3.9 for pointers to other standard test patterns. (Note: Do not attempt to use these reproductions as test patterns.)

TABLE 3.9
Standards of Interest in Scanned Imaging

Principal Standards Group	Subgroup	Sampling of Areas of Work or Example Standards	Web site(s) and/or Address(es)
ANSI—American National Standards Institute (those assigned to NPES)	CGATS: Committee for Graphic Arts Technologies Standards	CGATS-4-1993 Graphic Technology—Graphic Arts Reflection Densitometry Measurements—Terminology, Equations, Image Elements and Procedures	http://www.npes.org/standards/cgats.html NPES Assoc. for Suppliers of Printing and Publishing Technologies, 1899 Preston White Dr., Reston, VA 22091
	IT8: Committee on Digital Data Exchange and Color Definition (assigned to CGATS in '94)	IT8.1 Exchange of Color Picture Data IT8.7/2-1993-Reaffirmed 2008 Color Reflection Target for Input Scanner Calibration	http://www.ansi.org American National Standards Inst, 1819 L St. NW #230 Washington, DC
Graphic Communications Association	GRACoL <i>Regrouped significantly in 2004</i>	General Requirements for Applications in Commercial Offset Lithography—specifications, best practices (<i>e.g.</i> SWOP procedures)	http://www.gracol.org/index.html IDEAlliance, 1421 Prince St. NW #230 Alexandria, VA
ICC—International Color Consortium	NA	ICC Color Management Profile Specification, <i>e.g.</i> ICC. 1:2003-09 ver 4.1.0	http://www.color.org/index.xalter see ANSI-NPES above
ISO/IEC	JTC 1/SC 28 (joint technical committee on office equipment)	ISO 13660-2001 image quality measurement for hard copy output: ISO 12653-2:2000 Test target for black and white scanning	http://www.iso.ch www.iec.ch ISO Secretariat International Organization for Standardization, 1, CH, de la Voie-Creuse Case postale 56, CH-1211 Geneva 20, Switzerland IEC Central Office 3, Rue de Varembeé, PO Box 131 CH-1211 Geneva 20, Switzerland
	JTC 1/SC 29. (coding of multimedia information)	Coding of audio, picture, multimedia and hypermedia information includes bilevel and limited bits-per-pixel still pictures	

<p>ISO-TC42-WG18 (electronic still picture imaging covers scanners and digital still cameras) see Section 3.4.2</p>	<p>ISO 16067-1&2, 2003 &4: Electronic scanners for photographic images—Spatial Resolution Measurements; ISO 21550-2004 Dynamic range measurement; ISO 20462-3-2005, Psychophysical experimental methods for estimating image quality; ISO 12233 Photography—Electronic still picture cameras—resolution measurements</p>	<p>ISO 16067-1&2, 2003 &4: Electronic scanners for photographic images—Spatial Resolution Measurements; ISO 21550-2004 Dynamic range measurement; ISO 20462-3-2005, Psychophysical experimental methods for estimating image quality; ISO 12233 Photography—Electronic still picture cameras—resolution measurements</p>
<p>INCITS (International Committee for Information Technology Standards); formerly NCITS '97-01 X3 '61-96</p>	<p>W1 (office equipment</p>	<p>Includes copiers, multifunction, fax machines, page printers, scanners and other office equipment.⁹⁰ Collaborates, w JTC1/SC 28—for example on ISO 13660, above.</p>
<p>ITU-T (International Telecommunication Standardization Sector)</p>	<p>Joint Photographic Expert Group</p>	<p>ITU-T Rec. T.800 and ISO/IEC 15444, Information Technology—Digital compression and coding of continuous tone still images (JPEG 2000) [see References 93,167,168,180 D. Lee pp. 248-267]</p>
<p>IEEE</p>	<p>Joint with ISO/IEC/JTC1 SC29 WG1</p>	<p>JBIG2 ITU-T Rec.T.88</p>
<p>Transmission Systems Committee of IEEE Communications Society</p>	<p>Facsimile imaging, which is of value to many general scanning areas of interest. See Figure 3.50 for example test patterns</p>	<p>http://www.ieee.com The Institute of Electrical and Electronics Engineers, Inc., Corporate Offices 3 Park Ave, 17th Floor New York, NY 10016</p>

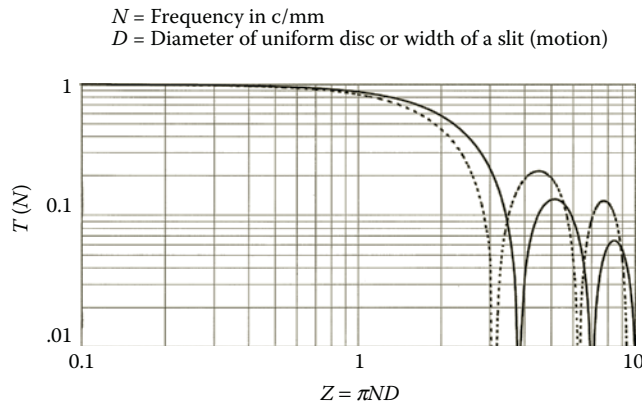


FIGURE 3.51 MTF of a uniform disk (solid line) and slit (dashed line) spread functions where N = frequency in cycles/mm and D = diameter of uniform disk, width of slit or rectangular aperture, or length of motion.

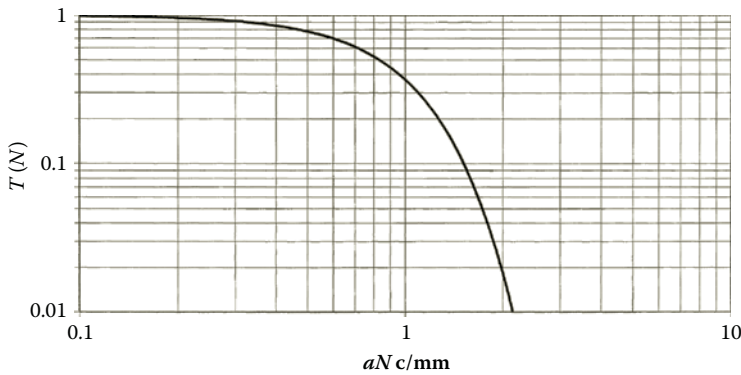


FIGURE 3.52 MTF for imaging system with Gaussian spread function.

where $a = \pi/c$, and c = width of Gaussian spread function $S(r)$ of the form

$$S(r) = 2c^2 e^{-c^2 r^2} = 2c^2 e^{-c^2(x^2 + y^2)} \tag{3.39}$$

where r = radius such that $r^2 = x^2 + y^2$; all are in mm^2 .

Figure 3.53 is the MTF of a diffraction-limited lens, where

$$T(N) = \frac{2}{p} [\cos^{-1} g - g\sqrt{1 - g^2}] \tag{3.40}$$

where $\gamma = N\lambda f$ (object at ∞), N = cycles/mm, λ = wavelength of light in mm, and f = aperture ratio = [focal length]/[aperture diameter].

It has been suggested (Reference 200) that terms in Equation 3.37 raised to the powers of 2, 3, and 4 (where D = the spacing between sensor elements) is useful approximations

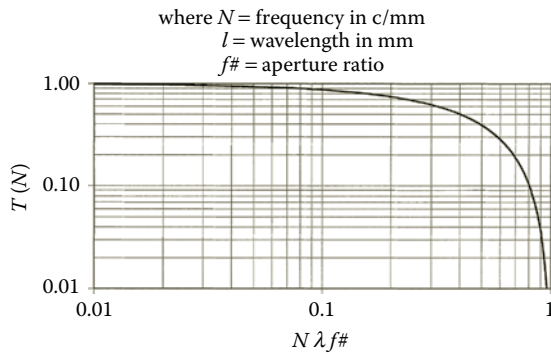


FIGURE 3.53

MTF of a diffraction-limited lens where N = frequency in cycles/mm, λ = wavelength in mm, and $f\#$ = aperture ratio.

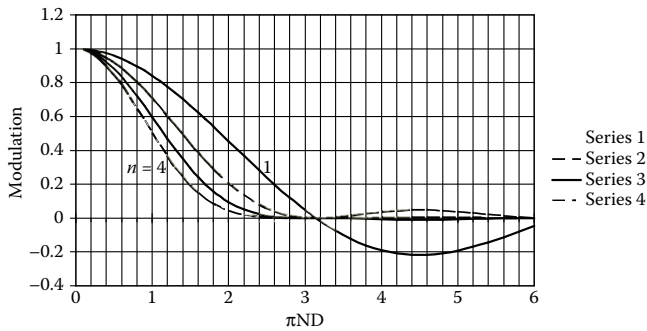


FIGURE 3.54

The general MTF family represented by $[\sin \pi DN/\pi DN]^n$ showing curves for $n = 1-4$. $n = 1$ case is shown previously in Figure 3.51 where the terms are explained. $N = 3$ and 4 approximate many real scanners.

to certain cases of actual scanner MTF performance. These are shown in Figure 3.54 where n = the power of the $[\sin \pi DN/\pi DN]$ term. In one case averaging over all sampling phases with an ideal sensor (where the sensor width = array spacing, i.e. 100% fill), the $n = 2$ case was a good approximation. In the case of several real film scanners (Reference 200) where other degradations from optics enter in, the $n = 3$ case was shown to be a good fit. Finally it appears that some inexpensive flatbed scanners which have even more degradation fit the $n = 4$ case.

Figure 3.55 presents data on four representative modem films, plotted here to provide perspective on the range of practical photographic characteristics. They are shown here to set scanning performance in perspective. These are not intended to be performance specifications of specific films.

Lastly we finish the reference curves with visual performance relationships. Figure 3.56 illustrates recently developed visual contrast sensitivity curves (related to MTF of linear systems) including color components of vision, after Fairchild,⁴ drawn with scales relating to the earlier published visual frequency response characteristics shown in Figure 3.31. Figure 3.57 shows the line luminance visibility threshold as a function of line width, originally described as display “seam visibility” from display experiments after Alphonse and

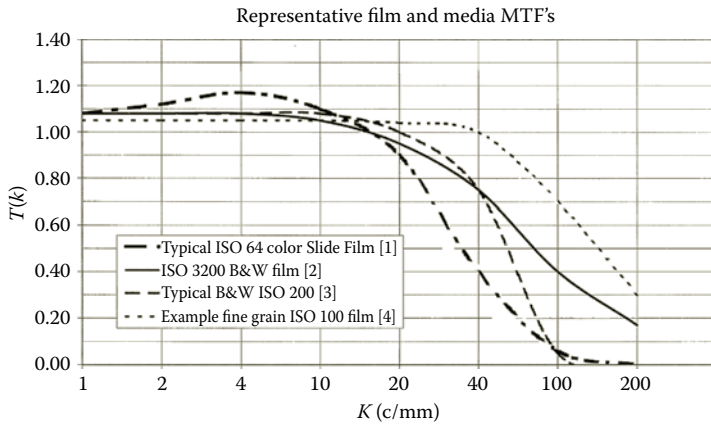


FIGURE 3.55
Data on four representative modern films.

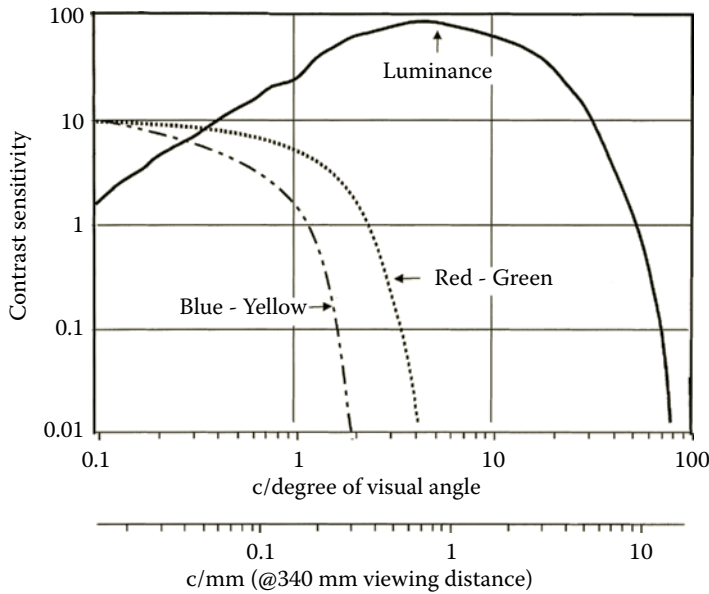


FIGURE 3.56
Typical visual spatial contrast sensitivity functions for luminance and indicated chromatic contrasts at constant luminance. (Adapted from Fairchild, M.D. *Color Appearance Models*; Addison-Wesley: Reading, MA, 1998.)

Lubin.¹²⁵ Figure 3.58 shows the edge contrast threshold visibility from display experiments of Lubin and Pica.¹²⁵

The closing reference, Table 3.9, is a chart showing a *very* sparse cross section of the standards that intercept the digital and scanning image quality technical world. These enable an engineer to get some orientation and pointers to important standards organizations.

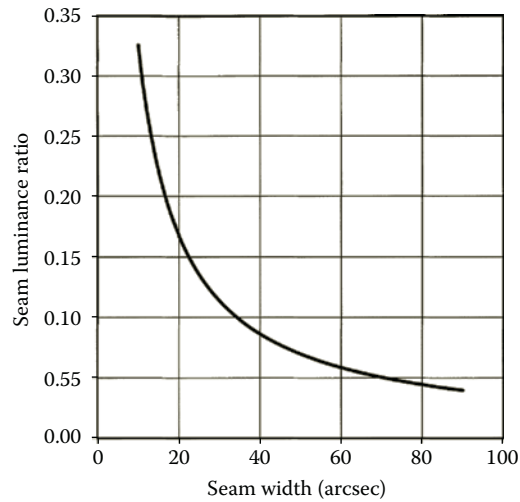


FIGURE 3.57

Line luminance visibility threshold as a function of line width for a black or white line on the opposite background derived from seam visibility for CRT displays. (Adapted from Lubin, J. The use of psychophysical data and models in the analysis of display system performance. In *Digital Images and Human Vision*; Watson, A.B., Ed.; MIT Press: Cambridge, MA, 1993; 163–178.)

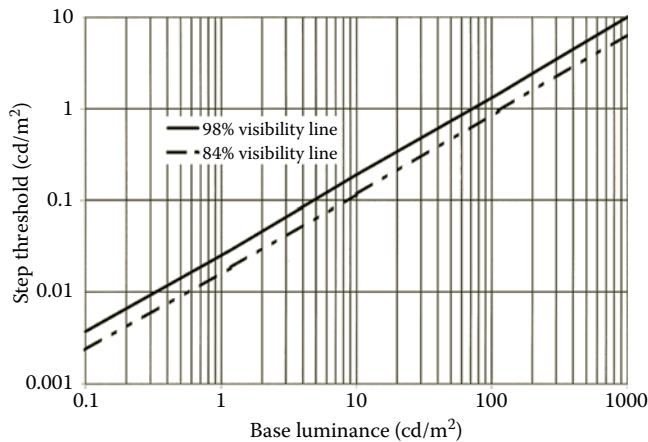


FIGURE 3.58

Thresholds for the visibility of a luminance difference at a step edge in a 17 by 5.25° CRT display where 84% detection = $dL_{85} = 0.01667L^{0.8502}$. (Adapted from Lubin, J. The use of psychophysical data and models in the analysis of display system performance. In *Digital Images and Human Vision*; Watson, A.B., Ed.; MIT Press: Cambridge, MA, 1993; 163–178.)

ACKNOWLEDGMENTS

Memorial Remarks This chapter is dedicated to the memory of Dr. John C. Urbach who died in his home, in Portola Valley, California, in February 2002, after several months of illness. He was a brilliant, dedicated, and prolific contributor to the field of optics and

scanning. His outstanding contributions to Xerox research and the careers and ideas of many he worked with at Xerox, as a consultant and in the general technical community, are widely regarded with the highest esteem. The previous edition of this chapter would not have been completed without his efforts, as he continued to help in its editing, even during his last days. We all miss his learned advice, special humor, and profound insights.

Contributions We are indebted to Cherie Wright, John Moore, David Lieberman, Roger Triplett, and others on the staff of the Imaging Sciences Engineering and Technology Center of the Strategic Programs Development Unit at Xerox Corporation for their helpful participation and support in the preparation of enduring parts of the earlier version of this chapter, and to Xerox Corporation for the use of their resources. I'd also like to thank my colleagues in the Imaging and Photographic Technology Department²⁰¹ at RIT for their support as I reworked the current edition. All illustrations, except as noted, are original drawings created for purposes of this chapter. However, those inspired by another author's way of illustrating a complex topic or providing a collection of useful data, reference his or her contribution with the note "(From Reference ____)." We are grateful for these authors' ideas or data, which we could build on here. We also wish to thank our reviewers Martin Banton, Guarav Sharma, Robert Loce, and Keith Knox for their time and many valuable suggestions, and our wives Jane Lehmebeck and Mary Urbach for their support and encouragement during work on the earlier edition and Jane during the current revision.

REFERENCES

1. Mertz, P.; Gray, F.A. A theory of scanning and its relation to the characteristics of the transmitted signal in telephotography and television. *Bell System Tech. J.* 1934, 13, 464–515.
2. Hornack, J.P. *Visual Encyclopedia of Imaging Science and Technology*; J. Wiley: New York, 2002—In depth treatment of many topics including section on "Imaging Systems."
3. Dougherty, E.R. *Digital Image Processing Methods*; Marcel Dekker: New York, 1994.
4. Fairchild, M.D. *Color Appearance Models*; Addison-Wesley: Reading, MA, 1998.
5. Eschbach, R.; Braun, K. Eds. *Recent Progress in Color Science*; Society for Imaging Science & Technology: Springfield, VA, 1997.
6. Sharma, G. Ed. *Digital Color Imaging Handbook*; CRC Press: Boca Raton, FL, 2003.
7. Eschbach, R. Ed. *Recent Progress in Digital Halftoning I and II*; Society for Imaging Science & Technology: Springfield, VA, 1994, 1999.
8. Kang, H. *Color Technology for Electronic Imaging Devices*; SPIE Press: Bellingham, WA, 1997.
9. Dougherty, E.R. Ed. *Electronic Imaging Technology*; SPIE Press: Bellingham, WA, 1999.
10. MacAdam, D.L. Ed. *Selected Papers on Colorimetry—Fundamentals, MS 77*; SPIE Press: Bellingham, WA, 1993.
11. Pennebaker, W.; Mitchell, J. *JPEG Still Image Data Compression Standard*; Van Nostrand Reinhold: New York, 1993.
12. Engeldrum, P.G. *Psychometric Scaling: A Toolkit for Imaging Systems Development*; Imcotek Press: Winchester, MA, 2000.
13. Vollmerhausen, R.H.; Driggers, R.G. *Analysis of Sampled Imaging Systems Vol. TT39*; SPIE Press: Bellingham, WA, 2000.
14. Giorgianni, E.J.; Madden, T.E. *Digital Color Management Encoding Solutions*; Addison-Wesley: Reading, MA, 1998.
15. Watson, A.B. Ed. *Digital Images & Human Vision*; MIT Press: Cambridge, MA, 1993.

16. Sharma, G. *Digital Color Imaging Handbook*; CRC Press, Boca Raton, FL, 2003. An excellent in depth review of many topics including: fundamentals, psychophysics, color management, digital color halftones, compression and camera image processing and more.
17. Russ, J.C. *The Image Processing Handbook*, 5th Ed.; Taylor and Francis-CRC: Boca Raton, FL, 2007—In depth discussion of numerous image processing topics.
18. Graham, R. *The Digital Image*; CRC Press-Whittles Publishing: Boca Raton, FL 2005—Excellent tutorial on fundamentals of digital imaging and especially photography.
19. Cost, F. *Pocket Guide to Digital Printing*; Delmar Publishers: Albany, NY, 1997.
20. Ohta, N., Rosen, M. *Color Desktop Printing Technology*; Taylor & Francis-CRC Div: Boca Raton, FL, 2006—Comprehensive overview with useful detail.
21. Gann, R.G. *Desktop Scanners Image quality Evaluation*; (Prentice Hall PTR, Upper Saddle River, NJ, 1999) overall practical serious-user oriented with especially useful practical tests.
22. Matteson, R. *Scanning for the SOHO Small Office and Home Office*, Virtualbookworm.com Publishing PO Box 9949, College Station TX 2004—Excellent very basic tutorial on all aspects of scanning.
23. Adams, R.M.; Weisberg, J.B. *The GATF Practical Guide to Color Management*; GATF Press: Pittsburgh, PA, 2000.
24. Sharma, G.; Wang, S.; Sidavanahalli, D.; Knox, K. "The impact of UCR on scanner calibration." Proceedings of IS&T Image Processing, Image Quality, Image Capture, Systems Conference. Portland, OR, 1998; 121–124.
25. Knox, K.T. "Integrating cavity effect in scanners." Proceedings of IS&T/OSA Optics and Imaging in the Information Age, Rochester, NY, 1996; 156–158.
26. Sharma, G.; Knox, K.T. "Influence of resolution on scanner noise perceptibility." Proceedings of IS&T 54th Annual and Image Processing, Image Quality, Image Capture, Systems Conference, Montreal, Quebec, Canada, 2001; 137–141.
27. Loce, R.; Roetling, P.; Lin, Y. Digital halftoning for display and printing of electronic images. In *Electronic Imaging Technology*; Dougherty, E.R., Ed.; SPIE Press: Bellingham, WA, 1999.
28. Lieberman, D.J.; Allebach, J.P. "On the relation between DBS and void and cluster." Proceedings of IS&T's NIP 14: International Conference on Digital Printing Technologies, Toronto, Ontario, Canada, 1998; 290–293.
29. Sharma, G.; Trussell, H.J. Digital color imaging. *IEEE Trans. on Image Proc.* 1997, 6, 901–932.
30. Sharma, G.; Vrhel, M.; Trussell, H.J. Color imaging for multimedia. *Proc. IEEE* 1998, 86, 1088–12108.
31. Jin, E.W.; Feng, X.F.; Newell, J. "The development of a color visual difference model (CVDM)." Proceedings of IS&T Image Processing, Image Quality, Image Capture, Systems Conference, Portland, OR, 1998; 154–158.
32. Sharma, G.; Trussell, H.J. Figures of merit for color scanners. *IEEE Trans. on Image Proc.* 1997, 6, 990–1001.
33. Shaw, R. "Quantum efficiency considerations in the comparison of analog and digital photography." Proceedings of IS&T Image Processing, Image Quality, Image Capture, Systems Conference, Portland, OR, 1998; 165–168.
34. Loce, R.; Lama, W.; Maltz, M. Vibration/banding. In *Electronic Imaging Technology*; Dougherty, E.R. Ed.; SPIE Press: Bellingham, WA, 1999.
35. Dalal, E.N.; Rasmussen, D.R.; Nakaya, F.; Crean, P.; Sato, M. "Evaluating the overall image quality of hardcopy output." Proceedings of IS&T Image Processing, Image Quality, Image Capture, Systems Conference, Portland, OR, 1998; 169–173.
36. Rasmussen, D.R.; Crean, P.; Nakaya, F.; Sato, M.; Dalai, E.N. "Image quality metrics: Applications and requirements." Proceedings of IS&T Image Processing, Image Quality, Image Capture, Systems Conference, Portland, OR, 1998; 174–178.
37. Loce, R.; Dougherty, E. Enhancement of digital documents. In *Electronic Imaging Technology*; Dougherty, E.R., Ed.; SPIE Press: Bellingham, WA, 1999.
38. Lieberman, D.J.; Allebach, J.P. "Image sharpening with reduced sensitivity to noise: A perceptually based approach." Proceedings of IS&T's NIP 14: International Conference on Digital Printing Technologies, Toronto, Ontario, Canada, 1998; 294–297.

39. Keelan, B.W. *Handbook of Image Quality Characterization and Prediction*; Marcell Dekker: New York, 2002—Comprehensive and detailed.
40. Engeldrum, P.G. "A new approach to image quality." Proceedings of the 42nd Annual Meeting of IS&T, 1989; 461–464.
41. Engeldrum, P.G. A framework for image quality models. *Imaging Sci. Technol.* 1995, 39, 312–323.
42. Engeldrum, P.G. *Psychometric Scaling: A Toolkit for Imaging Systems Development*; IMCOTEK Press: Winchester, MA, 2000; chapter 2, 5–17.
43. Shannon, C.E. A mathematical theory of communication. *Bell System Tech. J.* 1948, 27, 379, 623.
44. Roetling, "P.G. Visual performance and image coding." Proceedings of the Society of Photo-Optical Instrumentation Engineers on Image Processing, Vol. 74, 1976; 195–199.
45. Roetling, P.G.; Loce, R.P. Digital halftoning. In *Digital Image Processing Methods*; Dougherty, E.R., Ed.; Marcel Dekker: New York, 1994; 363–413.
46. Eyer, J.A. The influence of emulsion granularity on quantitative photographic radiometry. *Photog. Sci. Eng.* 1962, 6, 71–74.
47. Dainty, J.C.; Shaw, R. *Image Science: Principles, Analysis and Evaluation of Photographic-Type Imaging Processes*; Academic Press: New York, 1974.
48. Selwyn, E.W.H. A theory of graininess. *Photog. J.* 1935, 75, 571–589.
49. Siedentopf, H. Concerning granularity, resolution, and the enlargement of photographic negatives. *Physik Zeit.* 1937, 38, 454.
50. Altman, J.H.; Zweig, H.J. Effect of spread function on the storage of information on photographic emulsions. *Photog. Sci. Eng.* 1963, 7, 173–177.
51. Lehmbeck, D.R. Experimental study of the information storing properties of extended range film. *Photog. Sci. Eng.* 1967, 11, 270–278.
52. Vaysman, A.; Fairchild, M.D. "Degree of quantization and spatial addressability tradeoffs in the perceived quality of color images." *Proc SPIE on Color Imaging III* 1998, 3300, 250.
53. Marshall, G. *Handbook of Optical and Laser Scanning*, chapter 3; Marcell Dekker, NY, 2004 (previous edition this book & chapter).
54. Bryngdahl, O.J. *Opt. Soc. Am.* 1976, 66, 87–98.
55. Loce, R.P.; Dougherty, E.R. *Enhancement and Restoration of Digital Documents*; SPIE Optical Engineering Press: Bellingham, WA, 1997.
56. Jorgensen, G.W. Preferred tone reproduction for black and white halftones. In *Advances in Printing Science and Technology*; Banks, W.H., Ed.; Pentech Press: London, 1977; 109–142.
57. Jones, L.A.; Nelson, C.N. The control of photographic printing by measured characteristics of the negative. *J. Opt. Soc. Am.* 1942, 32, 558–619.
58. Jones, L.A. Recent developments in the theory and practice of tone reproduction. *Photogr. J. Sect. B* 1949, 89B, 126–151.
59. Bartleson, C.J.; Breneman, E.J. Brightness perception in complex fields. *J. Opt. Soc. Am.* 1967, 57, 953–957.
60. Nelson, C.N. Tone reproduction. In *The Theory of Photographic Process*, 4th Ed.; James, T.H., Ed.; Macmillan: New York, 1977; 536–560.
61. Nelson, C.N. The reproduction of tone. In *Neblette's Handbook of Photography and Reprography: Materials, Processes and Systems*, 7th Ed.; Sturge, J.M., Ed.; Van Nostrand Reinhold: New York, 1977; 234–246.
62. Holladay, T.M. An optimum algorithm for halftone generation for displays and hard copies. *Proceedings of the SID* 1980, 21, 185–192.
63. Roetling, P.G.; Loce, R.P. Digital halftoning. In *Digital Image Processing Methods*; Dougherty, E.R., Ed.; Marcel Dekker: New York, 1994; 392–395.
64. Roetling, P.G. Analysis of detail and spurious signals in halftone images. *J. Appl. Phot. Eng.* 1977, 3, 12–17.
65. ISO-TC-42, ISO 14524-1999 and 12232:2006(E), International Standards Organization, Geneva, Switzerland, 2006, See Table 10, this chapter. OECF stands for Opto-electronic Conversion Function—As applied to cameras in ISO 14524 which is conceptually the same for scanners. 12232 deals with speed metrics derived from OECF's.

66. Stoffel, J.C. *Graphical and Binary Image Processing and Applications*; Artech House: Norwood, MA, 1982; 285–350.
67. Stoffel, J.C.; Moreland, J.F. A survey of electronic techniques for pictorial image reproduction. *IEEE Trans. Comm.* 1981, 29, 1898–1925.
68. Ulichney, R. *Digital Halftoning*; The MIT Press: Cambridge, MA, 1987.
69. Clapper, R.; Yule, J.A.C. The effect of multiple internal reflections on the densities of halftone prints on paper. *J. Opt. Soc. Am.*, 43, 600–603, 1953, as explained in Yule, J.A.C. *Principles of Color Reproduction*; John Wiley and Sons: New York, 1967; 214.
70. Yule, J.A.C.; Nielson, W.J. The penetration of light into paper and its effect on halftone reproduction. In *Research Laboratories Communication No. 416*; Kodak Research Laboratories: Rochester, NY, 1951 and in TAGA Proceedings, 1951, 3, 65–76.
71. Lehmbeck, D.R. "Light scattering model for predicting density relationships in reflection images." Proceedings of 28th Annual Conference of SPSE, Denver, CO, 1975; 155–156.
72. Maltz, M. Light-scattering in xerographic images. *J. Appl. Phot. Eng.* 1983, 9, 83–89.
73. Kofender, J.L. "The Optical Spread Functions and Noise Characteristics of Selected Paper Substrates Measured in Typical Reflection Optical System Configurations," MS thesis, Rochester Institute of Technology: Rochester, NY, 1987.
74. Klees, K.J.; Holmes, J. "Subjective evaluation of noise filters applied to bi-level images." 25th Fall Symposia of Imaging (papers in summary form only). Springfield, VA, *Soc. Phot. Sci. & Eng.*, 1985.
75. Hunt, R.W.G. *Reproduction of Colour in Photography, Printing & Television*, 5th Ed.; The Fountain Press: Tolworth, England, 1995.
76. Hunt, R.W.G. *Measuring Colour*; Ellis Horwood Limited, Halstead Press, John Wiley & Sons: NY, 1987.
77. Miller, M.; Segur, R. "Perceived IQ and acceptability of photographic prints originating from different resolution digital capture devices." Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Savannah, GA, 1999; 131–137.
78. Smith, W.J. *Modern Optical Engineering*; McGraw Hill: New York, 1966; 308–324.
79. Bestenreiner, F.; Greis, U.; Helmberger, J.; Stadler, K. Visibility and correction of periodic interference structures in line-by-line recorded images. *J. Appl. Phot. Eng.* 1976, 2, 86–92.
80. Sonnenberg, H. Laser-scanning parameters and latitudes in laser xerography. *Appl. Opt.* 1982, 21, 1745–1751.
81. Firth, R.R.; Kessler, D.; Muka, E.; Naor, K.; Owens, J.C. A continuous-tone laser color printer. *J. Imaging Technol.* 1988, 14, 78–89.
82. Goodman, N.B. "Perception of spatial color variation caused by mass variations about single separations." Proceedings of IS&T's NIP14: International Conference on Digital Printing Technologies, Toronto, Ontario, Canada, 1998; 556–559.
83. Shade, O. Image reproduction by a line raster process. In *Perception of Displayed Information*; Biberman, L.M., Ed.; Plenum Press: New York, 1976; 233–277.
84. Shade, O. Image gradation, graininess and sharpness in TV and motion picture systems. *J. SMPTE* 1953, 67, 97–164.
85. Biberman, L.M. Ed. *Perception of Displayed Information*; Plenum Press: New York, 1976.
86. Lehmbeck, D.R.; Urbach, J.C. "Scanned Image Quality," Xerox Internal Report X8800370; Xerox Corporation: Webster, NY, 1988.
87. Kipman, Y. Imagexpert Home Page, <http://www.imagexpert.com>; Nashua NH, 2003 (describes several scanning-based image quality tools).
88. Wolin, D.; Johnson, K.; Kipman, Y. "Importance of objective analysis in IQ evaluation." IS&T's NIP14: International Conference on Digital Print Technologies, Toronto, Ontario, Canada, 1998; 603.
89. Briggs, J.C.; Tse, M.K. "Beyond density and color: Print quality measurement using a new hand-held instrument." Proceedings of ICIS 02: International Congress of Imaging Science, Tokyo, Japan, May 13–17, 2002, and describes other scanning-based image quality tools at QEA Inc., <http://www.qea.com> (accessed 2003).

90. Yuasa, M.; Spencer, P. NCITS-W1: "Developing standards for copiers and printers." Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Savannah GA, 1999; 270.
91. Williams, D. "Debunking of specsmanship: Progress on ISO/TC42 standards for digital capture imaging performance." Proceedings of IS&T Processing Images, Image Quality, Capturing Images Systems Conference (PICS), Rochester, NY, 2003; 77–81.
92. Williams, D. "Benchmarking of the ISO 12233 slanted edge spatial frequency response plug-in." Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Portland, OR, 1998; 133–136.
93. Hubel, P.M. "Color IQ in digital cameras." Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, 1999; 153.
94. Swing, R.E. *Selected Papers on Microdensitometry*; SPIE Optical Eng Press: Bellingham, WA, 1995.
95. Swing, R.E. *An Introduction to Microdensitometry*; SPIE Optical Eng Press: Bellingham, WA, 1997.
96. Lehmbek, D.R.; Jakubowski, J.J. Optical-principles and practical considerations for reflection microdensitometry. *J. Appl. Phot. Eng.* 1979, 5, 63–77.
97. Ptucha, R. "IQ assessment of digital scanners and electronic still cameras." Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Savannah, GA, 1999; 125.
98. Gonzalez, R.C.; Wintz, P. *Digital Image Processing*; Addison, Wesley: Reading, MA, 1977; 36–114.
99. Jakubowski, J.J. Methodology for quantifying flare in a microdensitometer. *Opt. Eng.* 1980, 19, 122–131.
100. Knox, K.T. "Integrating cavity effect in scanners." Proceedings of IS&T/OSA Optics and Imaging in the Information Age, Rochester, NY, 1996; 156–158.
101. Knox, K.T. US Patent #5,790,281, August 4, 1998.
102. Perrin, F.H. Methods of appraising photographic systems. *J. SMPTE* 1960, 69, 151–156, 239–249.
103. Shade, O. *Image Quality, a Comparison of Photographic and Television Systems*; RCA Laboratories: Princeton, NJ, 1975.
104. Newell, J.T.; Triplett, R.L. An MTF analysis metric for digital scanners. *Proceedings of IS&T 47th Annual Conference/ICPS, Rochester, NY, 1994*; 451–455.
105. Lamberts, R.L. The prediction and use of variable transmittance sinusoidal test objects. *Appl. Opt.* 1963, 2, 273–276.
106. Applied Image, on line catalog, (Applied Image Inc., 1653 E. Main Street, Rochester NY USA, 2009, <http://www.aig-imaging.com/>). Nearly all test patterns referred to in this chapter and various standards are available from this source along with detailed descriptions in their on-line catalog. See also Reference 200.
107. Scott, F.; Scott, R.M.; Shack, R.V. The use of edge gradients in determining modulation transfer functions. *Photog. Sci. Eng.* 1963, 7, 345–356.
108. Campbell, F.W. *Proc. Australian Physiol. Soc.* 1979, 10, 1.
109. Gorog, I.; Carlson, C.R.; Cohen, R.W. "Luminance perception—Some new results." In Proceedings, SPSE Conference on Image Analysis and Evaluation; Shaw, R., Ed.; Toronto, Ontario, Canada, 1976; 382–388.
110. Bryngdahl, O. Characteristics of the visual system: Psychophysical measurements of the response to spatial sine-wave stimuli in the photopic region. *J. Opt. Soc. Am.* 1966, 56, 811–821.
111. Watanabe, H.A.; Mori, T.; Nagata, S.; Hiwatoshi, K. *Vision Res.* 1968, 8, 1245–1254.
112. Glenn, W.E.; Glenn, G.; Bastian, C.J. "Imaging system design based on psychophysical data." In Proceedings of the SID 1985, 26, 71–78.
113. Dooley, R.P.; Shaw, R. A statistical model of image noise perception. In *Image Science Mathematics Symposium*; Wilde, C. O., Barrett, E., Eds.; Western Periodicals: Hollywood, CA, 1977; 10–14.
114. Patterson, M. In Proceedings of the SID 1986, 27, 4.
115. Blakemore, C.; Campbell, F.W. *J. Physiol.* 1969, 203, 237–260.
116. Rogowitz, B.E. *Proceedings of the SID* 1983, 24, 235–252.

117. Hufnagel, R. In *Perception of Displayed Information*; Biberman, L., Ed.; Plenum Press: New York, 1973; 48.
118. Oppenheim, A.V.; Schafer, R. *Digital Signal Processing*; Prentice-Hall: Englewood Cliffs, NJ, 1975; 413–418.
119. Jones, R.C. New method of describing and measuring the granularity of photographic materials. *J. Opt. Soc. Am.* 1955, 45, 799–808.
120. Lehmbeck, D.R. *Imaging Performance Measurement Methods for Scanners that Generate Binary Output*. 43rd Annual Conference of SPSE, Rochester, NY, 1990; 202–203.
121. Vollmerhausen, R.H.; Driggers, R.G. *Analysis of Sampled Imaging Systems Vol. TT39*; SPIE Press: Bellingham, WA, 2000; 50–72.
122. Kriss, M. Image structure. In *The Theory of Photographic Process*, 4th Ed.; James, T.H., Ed.; Plenum Press: New York, 1977; Chap. 21, 592–635.
123. Carlson, C.R.; Cohen, R.W. A simple psychophysical model for predicting the visibility of displayed information. *Proc. of SID* 1980, 21, 229–246.
124. Barten, P.G.J. "The square root integral (SQRI): A new metric to describe the effect of various display parameters on perceived image quality." Proceedings of SPIE conference on Human Vision, Visual Processing, and Digital Display, Los Angeles, CA, 1989; Vol. 1077, 73–82.
125. Lubin, J. The use of psychophysical data and models in the analysis of display system performance. In *Digital Images and Human Vision*; Watson, A.B., Ed.; MIT Press: Cambridge, MA, 1993; 163–178.
126. Daly, S. The visible differences predictor: an algorithm for the assessment of image fidelity. In *Digital Images and Human Vision*; Watson, A.B., Ed.; MIT Press: Cambridge, MA, 1993; 179–206.
127. Scott, F. Three-bar target modulation detectability. *J. Photog. Sci. Eng.* 1966, 10, 49–52.
128. Charman, W.N.; Olin, A. Image quality criteria for aerial camera systems. *J. Photogr. Sci. Eng.* 1965, 9, 385–397.
129. Burroughs, H.C.; Fallis, R.F.; Warnock, T.H.; Brit, J.H. *Quantitative Determination of Image Quality*, Boeing Corporation Report D2: 114058–1, 1967.
130. Snyder, H.L. "Display image quality and the eye of the beholder." Proceedings of SPSE Conference on Image Analysis and Evaluation, Shaw, R., Ed.; Toronto, Ontario, Canada, 1976; 341–352.
131. Barten, P.G.J. The SQRI method: A new method for the evaluation of visible resolution on a display. *Proc. SID* 1987, 28, 253–262.
132. Barten, P.G.J. "Physical model for the contrast sensitivity of the human eye." Proceedings of the SPIE on Human Vision, Visual Processing, and Digital Display III, San Jose, CA, 1992; Vol. 1666, 57–72.
133. Daly, S. "The visible differences predictor: an algorithm for the assessment of image fidelity." Proceedings of the SPIE on Human Vision, Visual Processing, and Digital Display III, San Jose, CA, 1992; Vol. 1666, 2–15.
134. Frieser, H.; Biederman, K. Experiments on image quality in relation to modulation transfer function and graininess of photographs. *J. Phot. Sci. Eng.* 1963, 7, 28–46.
135. Biederman, K. *J. Photog. Korresp.* 1967, 103, 41–49.
136. Granger, E.M.; Cupery, K.N. An optical merit function (SQF) which correlates with subjective image judgements. *J. Phot. Sci. Eng.* 1972, 16, 221–230.
137. Kriss, M.; O'Toole, J.; Kinard, J. "Information capacity as a measure of image structure quality of the photographic image." Proceedings of SPSE Conference on Image Analysis and Evaluation, Toronto, Ontario, Canada, 1976; 122–133.
138. Miyake, Y.; Seidel, K.; Tomamichel, F. Color and tone corrections of digitized color pictures. *J. Photogr. Sci.* 1981, 29, 111–118.
139. Crane, E.M. *J. SMPTE* 1964, 73, 643.
140. Gendron, R.G. *J. SMPTE* 1973, 82, 1009.
141. Granger, E.M. Visual limits to image quality. *J. Proc. Soc. Photo-Opt. Instr. Engrs* 1985, 528, 95–102.

142. Natale-Hoffman, K.; Dalai, E.; Rasmussen, R.; Sato, M. Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Savannah, GA, 1999; 266–273.
143. Dalal, E.; Rasmussen, R.; Nakaya, F.; Crean, P.; Sato, M. "Evaluating the overall image quality of hardcopy output." Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Portland, OR, 1998; 169–173.
144. Inagaki, T.; Miyagi, T.; Sasahara, S.; Matsuzaki, T.; Gotoh, T. Color image quality prediction models for color hard copy. *Proceedings of SPIE* 1997, 2171, 253–257.
145. Goldman, S. *Information Theory*; Prentice Hall: New York, 1953; 1–63.
146. Felgett, P.B.; Linfoot, E.H.J. *Philos. Trans. R. Soc. London* 1955, 247, 369–387.
147. McCamy, C.S. On the information in a photomicrograph. *J. Appl. Opt.* 1965, 4, 405–411.
148. Huck, F.O.; Park, S.K. Optical–mechanical line-scan image process—Its information capacity and efficiency. *J. Appl. Opt.* 1975, 14, 2508–2520.
149. Huck, F.O.; Park, S.K.; Speray, D.E.; Halyo, N. Information density and efficiency of 2-dimensional (2-D) sampled imagery. *Proc. Soc. Photo-Optical Instrum. Engrs* 1981, 310, 36–42.
150. Burke, J.J.; Snyder, H.L. Quality metrics of digitally derived imagery and their relation to interpreter performance. *SPIE* 1981, 310, 16–23.
151. Sachs, M.B.; Nachmias, J.; Robson, J.G. *J. Opt. Soc. Am.* 1971, 61, 1176.
152. Stromeyer, C.F.; Julesz, B. Spatial frequency masking in vision: critical bands and spread of masking. *J. Opt. Soc. Am.* 1972, 62, 1221.
153. Miyake, Y.; Inoue, S.; Inui, M.; Kubo, S. An evaluation of image quality for quantized continuous tone image. *J. Imag. Technol.* 1986, 12, 25–34.
154. Metz, J.H.; Ruchti, S.; Seidel, K. Comparison of image quality and information capacity for different model imaging systems. *J. Photogr. Sci.* 1978, 26, 229.
155. Hunter, R.; Robinson, A.H. International digital facsimile coding standards. *Proc. IEEE* 1980, 68(7), 854–867.
156. Rabbani, M. *Image Compression. Fundamentals and International Standards, Short Course Notes*; SPIE: Bellingham, WA, 1995.
157. Rabbani, M.; Jones, P.W. *Digital Image Compression Techniques, TT7*; SPIE Optical Engineering Press: Bellingham, WA, 1991.
158. Joint BiLevel Working Group. *ITU-T Rec. T.82 and T.85*; Telecommunication Standardization Sector of the International Telecommunication Union, March 1995, August 1995.
159. Buckley, R.; Venable, D.; McIntyre, L. "New developments in color facsimile and internet fax." Proceedings of IS&T 5th Annual Color Imaging Conference, Scottsdale, AZ, 1997; 296–300.
160. Huffman, D. A method for the construction of minimum redundancy codes. *Proc. IRE* 1962, 40, 1098–1101.
161. Lempel, A.; Ziv, J. Compression of 2 dimensional data. *IEEE Trans Info. Theory* 1986, IT-32 (1), 8–19.
162. Lempel, A.; Ziv, J. Compression of 2 dimensional data. *IEEE Trans Info. Theory* 1977, IT-23, 337–343.
163. Lempel, A.; Ziv, J. Compression of 2 dimensional data. *IEEE Trans Info. Theory* 1978, 1T-24, 530–536.
164. Welch, T. A technique for high performance data compression. *IEEE Trans Comput.* 1984, 17(6), 8–19.
165. Beretta, G. Compressing images for the internet. *Proc. SPIE, Color Imaging, III*, 1998, 3300, 405–409.
166. Lee, D.T. Intro to color facsimile: Hardware, software, standards. *Proc. SPIE* 1996, 2658, 8–19.
167. Marcellin, M.W.; Gornish, M.J.; Bilgin, A.; "Boliek, M.P. An overview of JPEG 2000." SPIE Proceedings of 2000 Data Compression Conference, Snowbird, Utah, 2000, 2658, 8–30.
168. Sharpe II, L.H.; Buckley, R. "JPEG 2000.jpm file format: a layered imaging architecture for document imaging and basic animation on the web." Proceedings SPIE 45th Annual Meeting, San Diego, CA, 2001; 4115, 47.
169. Dougherty, E.R. Ed. *An Introduction to Morphological Image Processing*; SPIE Optical Engineering Press: Bellingham, WA, 1992.
170. Hamerly, J.R. An analysis of edge raggedness and blur. *J. Appl. Phot. Eng.* 1981, 7, 148–151.

171. Tung, C. "Resolution enhancement in laser printers." Proceedings of SPIE Conference on Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts II, San Jose, CA, 1997.
172. Tung, C. Piece Wise Print Enhancement. US Patent 4,847,641, July 11, 1989, US Patent 5,005,139, April 2, 1991.
173. Walsh, B.F.; Halpert, D.E. Low Resolution Raster Images, US Patent 4,437,122, March 13, 1984.
174. Bassetti, L.W. Fine Line Enhancement, US Patent 4,544,264 October 1, 1985, Interacting Print Enhancement, US Patent 4,625,222, November 25, 1986.
175. Lung, C.Y. Edge Enhancement Method and Apparatus for Dot Matrix Devices. US Patent 5,029,108, July 2, 1991.
176. Tuijn, W.; Cliquet, C. "Today's image capturing needs: going beyond color management." Proceedings IS&T/SID 5th Color Imaging Conference, Scottsdale, AZ, 1997; 203.
177. Gonzalez, G.; Hecht, T.; Ritzer, A.; Paul, A.; LeNest, J.F.; "Has, M. Color management—How accurate need it be." Proceedings IS&T/SID 5th Color Imaging Conference, Scottsdale, AZ, 1997; 270.
178. Frazier, A.L.; Pierson, J.S. Resolution transforming raster based imaging system, US Patent 5,134,495, July 28, 1992, Interleaving vertical pixels in raster-based laser printers, US Patent 5,193,008, March 9, 1993.
179. Has, M. Color management—Current approaches, standards and future perspectives. IS&T, 11NIP Proceedings, Hilton Head, SC, 1995; 441.
180. Buckley, R. *Recent Progress in Color Management and Communication*; Society for Imaging Science and Technology (IS&T): Springfield, VA, 1998.
181. Newman, T. "Making color plug and play." Proceedings IS&T/SID 5th Color Imaging Conference, Scottsdale, AZ, 1997; 284.
182. Chung, R.; Kuo, S. "Colormatching with ICC Profiles—Take one." Proc. IS&T/SID 4th Color Imaging Conference, Scottsdale, AZ, 1996; 10.
183. Rickmers, A.D.; Todd, H.N. *Statistics, an Introduction*; McGraw Hill: New York, 1967.
184. Dvorak, C.; Hamerly, J. Just noticeable differences for text quality components. *J. Appl. Phot. Eng.* 1983, 9, 97–100.
185. Hamerly, J. Just noticeable differences for solid area. *J. Appl. Phot. Eng.* 1983, 9, 14–17.
186. Bartleson, C.J.; Woodbury, W.W. Psychophysical methods for evaluating the quality of color transparencies III. Effect of number of categories, anchors and types of instructions on quality ratings. *J. Photo. Sci. Eng.* 1965, 9, 323–338.
187. Stevens, S.S. *Psychophysics: Introduction to Its Perceptual, Neural and Social Prospects*; John Wiley and Sons: New York, 1975; Reprinted: Transactions Inc.: New Brunswick, NJ, 1986.
188. Thurstone, L.L. Rank order as a psychophysical method. *J. Exper. Psychol.* 1931, 14, 187–195.
189. Stevens, S.S. On the theory of scales of measurement. *J. Sci.* 1946, 103, 677–687.
190. Kress, G. *Marketing Research*, 2nd Ed.; Reston Publishing Co. Inc.: a Prentice Hall Co.: Reston, VA, 1982.
191. Morrissey, J.H. New method for the assignment of psychometric scale values from incomplete paired comparisons. *JOSA* 1955, 45, 373–389.
192. Bartleson, C.J.; Breneman, E.J. Brightness perception in complex fields. *JOSA* 1967, 57, 953–960.
193. Bartleson, C.J. The combined influence of sharpness and graininess on the quality of color prints. *J. Photogr. Sci.* 1982, 30, 33–45.
194. Bartleson, C.J.; Grum, F. Eds. Visual measurements. In *Optical Radiation Measurements*, Academic Press: Orlando, FL, Vol. 5, 1984.
195. Gescheider, G.A. *Psychophysics: The Fundamentals*, 3rd Ed.; Lawrence Erlbaum Assoc. Inc.: Mahwah, NJ, 1997.
196. Guilford, J.P. *Psychometric Methods*; McGraw Hill Book Co.: New York, 1954.
197. Malone, D. Psychometric methods. In *SPSE Handbook of Photographic Science and Engineering*, Chapter 19.4; A Wiley Interscience Publication, John Wiley & Sons: New York, 1973; 1113–1128.
198. Nunnally, J.C.; Bernstein, I.R. *Psychometric Theory*, 3rd Ed.; McGraw Hill Book Co.: New York, 1994.

199. Torgerson, W.S. *Theory and Methods of Scaling*; J. Wiley & Sons: New York, 1958.
200. Koren, N. Making fine prints in your digital darkroom, Understanding image sharpness and MTF; <http://www.normankoren.com>, validated. 2/11/2009—Good practical and theoretical insights with pointers to a software product “Imatest” used to measure these and many other characteristics of cameras and scanners.
201. A. Davidhazy, Ed. <http://www.rit.edu/cias/photo/ipt-faculty/> (Rochester Institute of Technology, Rochester, NY, 2009) which lists faculty and related courses that influenced updates herein.
202. Hunter, R.S.; Harold, R.W. *The Measurement of Appearance*, 2nd Ed.; John Wiley and Sons: New York, 1987; 191.
203. Scheff'e, H. An analysis of variance for paired comparisons. *J. Am. Statist. Assoc.* 1952, 47, 381–395.
204. Daniels, C.M.; Ptucha, R.W.; Schaefer, L. “The necessary resolution to zoom and crop hardcopy images.” Proceedings of IS&T Image Processing, Image Quality, Image Capture Systems (PICS) Conference, Savannah, Georgia, 1999; 143.
205. Nakamura, J. *Image Sensors and Signal Processing for Digital Still Cameras*; Taylor and Francis: Boca Raton, FL, 2006; Mizoguchi, T. Ch 6: Evaluation of image sensors, 179–203.