The Print Quality Toolkit: An Integrated Print Quality Assessment Tool†

Jim Grice*,‡ and Jan P. Allebach*,‡

Electronic Imaging Systems Laboratory, School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana

The Print Quality Toolkit is a system for printing, scanning, and evaluating a test target on a workstation or personal computer using image processing techniques for the purpose of high-speed, objective, and mostly automated print quality assessment. The metric set used for this assessment is specified in the ISO/IEC document 13660, that, in turn, was primarily derived from the literature on print quality. As examples of potential applications for the system, we investigate the dependence of these metrics on scan resolution, line width, paper type, and printer technology.

Journal of Imaging Science and Technology 43: 187–199 (1999)

Introduction

Shortly after the appearance of electrophotographic copiers, research in the imaging industry began focusing on the quality of output produced by these devices. The evolution of the laser printer from this technology, and the subsequent development of thermal ink jet devices contributed to the need in industry for objective as well as subjective evaluative means for providing feedback to printer and printer software engineers. While some similarities exist between digital imaging and conventional photography, and some metrics, such as granularity/graininess, were used practically verbatim, much of this analysis was unique to the emerging field.

In the literature on image quality, one typically finds articles on individual metrics (or small sets of metrics), or on print quality evaluation *systems*, which usually consist of a printed target, imaging system, and analysis software. Along with these categories, there is also a marked contrast between purely objective metrics or measurement systems, and those that attempt to correlate a metric or set of metrics to subjective impressions of quality. Our project belongs to the print quality system category, and calculates purely objective metrics, with the exception of the granularity/graininess metric that incorporates subjective weighting.

Some examples of recent literature on print quality assessment are Edinger and Newell,[1] who correlate the width of a lower-case "L" with perceived text quality; Lee, Winslow, and Bares,[2] who test three image quality evaluation schemes on 41 paper samples; Engledrum,[3] who develops a model of image quality and uses it to find the correlation between subjective image quality and a weighted sum of objective metrics; and Wang[4] and Kipman,[5] both of whom develop print quality evaluation systems similar in concept to ours.

The Print Quality Toolkit calculates the print quality metrics specified in the ISO/IEC guidelines on hardcopy print quality assessment.[6] These guidelines include metrics for four distinct categories of printed areas: line and character metrics, solid-fill (colorant blocks) metrics, tint solid (halftoned blocks) metrics, and background field (non-printed areas) metrics. The document completely specifies the measurement and calculation techniques for the metrics, with the exception of *granularity*, that we developed via the work of others in the area.[7,8,9] We have developed a system to implement these metrics.

The step-by-step procedure for using the Print Quality Toolkit (PQTK) is illustrated in Fig. 1. The file containing the test target is printed. The printed test target is then scanned at the desired resolution, and the resulting scanned test target is transferred to a workstation or PC.

We describe the ISO/IEC metrics and their calculation in our Description of Metrics section. In the section, Description of Test Target, we describe the test target developed for the procedure. The calibration and confirmation of high frequency performance of our scanner is discussed in Scanner Calibration and MTF Testing section. The PQTK program, developed for calculating these metrics, is outlined in the section, Description of the PQTK Tool. Some experiments we performed and the results obtained are discussed in the Experimental Results section and finally, we present our Conclusions.

Description of Metrics

The set of print quality metrics calculated in the PQTK program are those defined in the ISO/IEC document 13660.[6] Some necessary definitions, descriptions of these metrics, the methodology for calculating them, and their background in the print quality literature (where applicable) follows.

Original manuscript received April 13, 1998

* IS&T Member

† This research was supported by the Hewlett-Packard Company

‡ e-mail: (grice + allebach)@ecn.purdue.edu

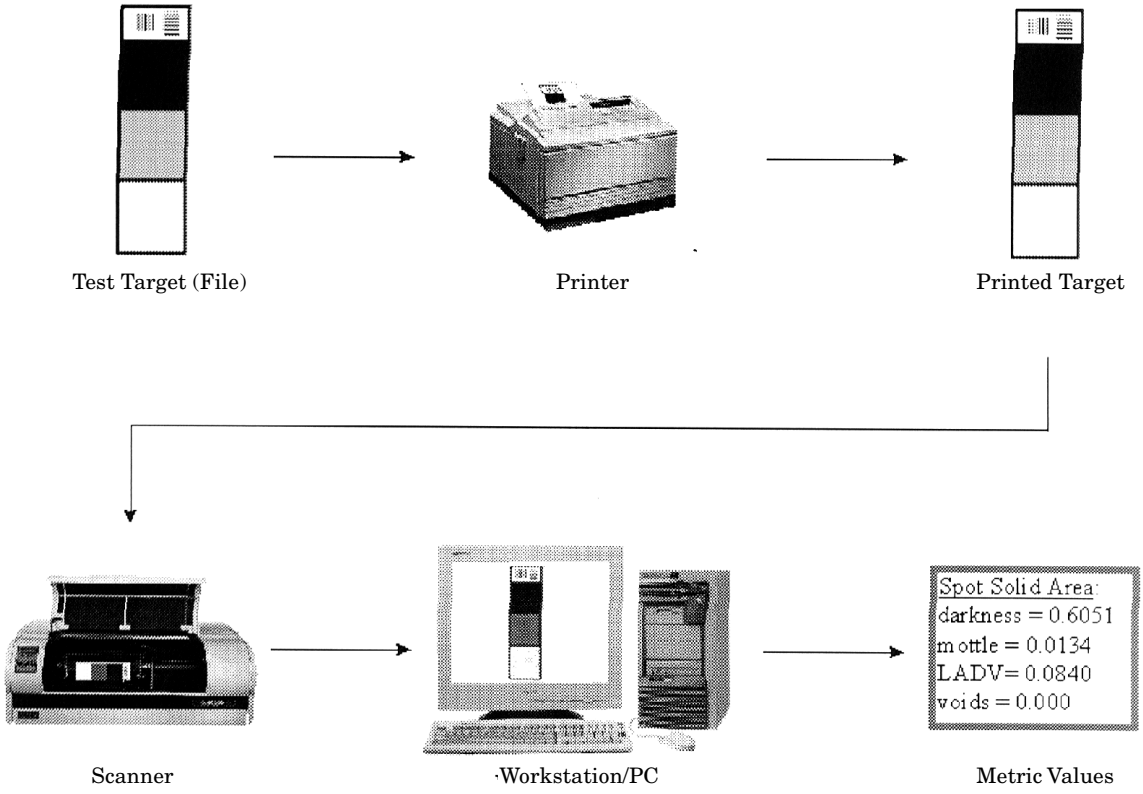


Figure 1. Overview of print quality analysis with the PQTK.

Definitions

character field the region within 500 μm of the T_{10} thresholds of the line of interest, excluding the line itself (i.e., outside of the T_{10} thresholds).

R_{max} in line metrics, it refers to the maximum reflectance value produced by a 63.5×254 mm^2 averaging window within the character field. This window is oriented with its long dimension along the length of the line. In area metrics, it refers to the maximum reflectance value of the substrate.

R_{min} in line metrics, it refers to the minimum reflectance value produced by a 63.5×254 mm^2 averaging window within the character field. This window is oriented with its long dimension along the length of the line. In area metrics, it refers to the minimum reflectance value of a solid colorant area.

T_{pp} **threshold** a number, calculated as

$$T_{pp} = R_{max} - \frac{pp}{100}(R_{max} - R_{min}),$$

representing a specific reflectance value.

T_{pp} **boundary** the physical location, interpolated if necessary, where the reflectance of the edge in question equals the T_{pp} threshold.

R_{line} the mean reflectance of a line within its T_{90} thresholds.

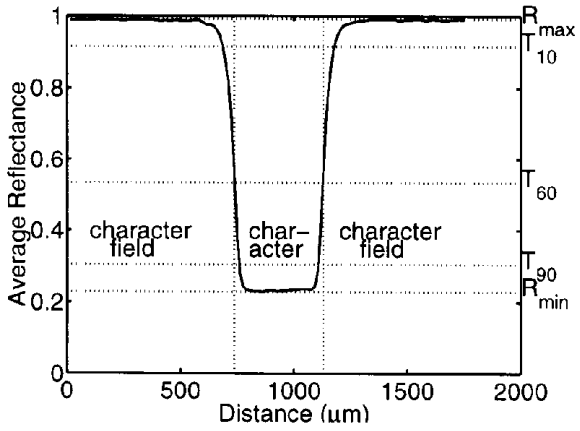


Figure 2. Visual representation of quantities used in the ISO/IEC metric definitions.

Figure 2 illustrates these quantities. It represents the reflectance profile perpendicular to a line segment. The character field regions, R_{max}, R_{min}, and several important threshold values T_{pp} are indicated on the plot.

Line Metrics

Line metrics refer to the attributes of a printed line segment (or linear character segment) and its neighborhood. These include:

Blurriness the sharpness of the transition from substrate to colorant, and vice-versa, at a line edge. It is measured as the mean

physical distance between the T_{10} and T_{90} boundaries for the edge profile, measured on both sides of the line.

Much of the literature concerning print quality assessment is in agreement on the measurement technique to be used for blurriness. However, the label placed on this metric varies. In Dvorak,[10] as well as much other literature, the term "NEP" or "Normal Edge Profile" is used. This refers to the reflectance profile along a path normal to the line edge. This profile's spatial property is, in effect, the metric. In Hamerly and Dvorak,[11] the term "edge transition width" is used, that is perhaps the most descriptive label relative to the actual measurement procedure. In an article by Jansson,[12] the descriptor is "sharpness", which bears an inverse relationship to blurriness.

Contrast the relationship between the reflectance of a line and the substrate. It is calculated by the standard formula:

$$\text{Contrast} = (R_{\max} - R_{\text{line}}) / R_{\max}.$$

Fill the homogeneity of the optical density of a line. It is measured as the ratio of the number of pixels within the T_{90} boundaries of the line with reflectance greater than the T_{75} threshold to the total number of pixels within the T_{90} boundaries.

This metric appears to be a direct adaptation of the technique used by Tanaka and Abe[13] in their print quality assessment tool to measure what they deem "void rate". The threshold used for void rate is 0.6 optical density units, that corresponds to a reflectance threshold of T_{75} in an image with a white point of 100% reflectance and a dark point of 0% reflectance. In effect, the ISO/IEC definition determines the threshold for voids from the darkness and background haze metrics.

Darkness the mean optical density within the T_{90} boundaries of a line. This metric requires calibration of the scanning device for absolute density values.

Raggedness the geometric distortion of an edge from ideal. A straight line is fitted in a least-squares sense to the T_{60} boundaries of the printed line. The raggedness metric is calculated as the standard deviation of the residuals of the actual T_{60} contour to the fitted line.

In the literature, there are many variations on the raggedness-like metrics, under names such as "raggedness",[13–16] "tangential edge roughness",[17] and "tangential edge profile" or "TEP".[18] Closely related to raggedness is "stair-stepping", the effect caused by approximating a diagonal line with printer dots on a discrete matrix. Most of the referenced literature includes the stair-stepping effect in the raggedness-like metric, although Edinger[19] treats it separately by measuring the instantaneous slope change of an "O" at several fixed angles.

* Note that this RMS definition agrees with that of the ISO/IEC document if one normalizes the standard deviation by N, the number of samples. However, the accepted form of the standard deviation, and the one we use in the PQTK program, is the unbiased estimate, i.e. normalized by $N - 1$ (the square root of the sample variance.)

Nearly as varied as the names given to the metrics are the details of their measurement: Schein and Beardsley,[14] Crawford, Elzinga and Yudico,[17] and Dvorak and Hamerly[18] use the RMS* deviation of an edge from a line fitted to it. Edinger[15] uses the standard deviation of the edge from the intended edge, not from a fitted edge. Lastly, Tanaka and Abe[13] parameterize raggedness into three components: "SD" that, as in the ISO/IEC document, is the standard deviation of the edge from a fitted edge, "PP", the peak-to-peak amplitude of the residuals to the fitted line, and "linearity", the ratio of the regression line length to the edge boundary length.

Stroke Width the mean width of a line. The line width is the physical distance between the T_{60} boundaries on either side of a line.

This definition appears to be taken from Crawford, Elzinga, and Yudico[17] with no revision. Edinger[15] refers to stroke width as "line broadening", the ratio of line width to nominal line width (evidently defining the line boundary at an unspecified reflectance threshold).

Extraneous Marks the presence of stray marks near a line. It is measured as the ratio of the number of pixels in the character field with reflectance less than the T_{20} threshold to the total number of pixels in the character field.

A similar measurement technique is employed by Edinger,[15] whose metric uses a ratio definition to calculate a "graininess" measure. No reflectance (or optical density) threshold to distinguish marks from background is mentioned. The test target used is the area between the vertical bars of "*N*". Here, the marks are called "satellites".

Background Haze the presence of haze near a line, not discernible as marks. Haze is defined as the ratio of the mean reflectance of the character field, excluding extraneous marks, to the mean reflectance of the background (substrate) outside of the character field. Edinger[15] uses the same technique to derive a graininess measure for background haze (or simply "background" in the paper) as for extraneous marks, but with two differences: the test target is a large area of substrate, and only colorant particles with diameters less than 40 μm are considered. Compare this with the ISO/IEC metric, that considers background elements based on their reflectance, not size.

Solid-Fill Metrics Solid-fill metrics refer to the attributes of an area of solid colorant.

Overall Darkness the mean optical density within the T_{90} boundaries of a solid-fill area. As with line darkness, this metric requires scanner calibration to derive absolute optical density units.

Mottle the standard deviation of optical density of small (2×2 mm^2) regions within the T_{90} boundaries of a solid-fill area. The region of interest is converted into opti-

cal density units, then convolved with a 2×2 mm^2 averaging window. The standard deviation of the resulting array entries is the metric value.

Tobias, Ricks, and Chadwick[20] developed a mottle tester that uses an approach similar to the ISO/IEC document. However, their window size was unspecified and their metric was calculated by taking the average deviation from the mean, as opposed to the standard deviation of the samples as in the ISO/IEC document.

Large Area Density Variation (LADV) the largest difference in density values of 4 mm square regions of a solid-fill area. The region of interest is converted into optical density units, then convolved with a 4 mm square averaging window. The difference between the largest and smallest array entries after this operation is the LADV.

This appears to be an adaptation of *mottle* by the ISO/IEC committee, as no references are given and the two metrics are very similar.

Voids percentage of a solid-fill area devoid of colorant. It is calculated as the ratio of the number of pixels within the T_{90} boundaries of the area to the number of pixels in this region with reflectance greater than the T_{40} threshold.

The area *voids* metric has the same origin and measurement technique as the fill metric for lines. Please see that subsection for the specifics.

Tint Solid Tint solid metrics refer to the attributes of a tint solid (halftone) area. Included in the tint solid metrics are *overall darkness, mottle,* and *LADV*. These are defined and referenced in the literature as in the solid-fill metrics. Please see that subsection for details.

Granularity the only metric whose measurement procedure is not completely specified in the ISO/IEC document. We derived our approach from that proposed in Refs. 7 and 9. The measurement involves the Weiner noise spectrum of the reflectance, i.e., fluctuation in the tint solids region, weighted by a human visual system transfer function and the mean optical density of the region. The measurement technique and the quantities involved in the measurement formulae will be explained in depth in the subsection on granularity.

Background Field Metrics Background field metrics refer to the attributes of an area of the substrate (paper).

Extraneous Marks proportion of the background field containing tint material. The metric is calculated as the ratio of the number of pixels with reflectance less than the T_{20} threshold to the total number of pixels in the background field region.

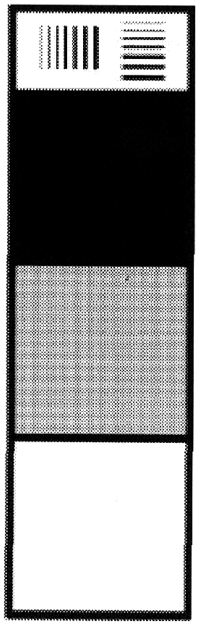


Figure 3. Test target developed for the PQTK.

This metric has identical measurement procedures as the extraneous marks method for lines, except that the measurement region is not near a printed line in this case. See the line metric equivalent for a literature discussion.

Background the standard deviation of the reflectance of the pixels in a non-printed area.

Description of Test Target

The test target developed for calculating the metrics in the ISO/IEC document is shown in Fig. 3. The different areas of the target are as follows:

Border The dark border surrounding each area of the target simplifies segmentation of the scanned target into its constituent parts, as well as providing reference lines for checking the skew of the scanned test target relative to the normal orientation.

Line Area The uppermost target area as shown in Fig. 3 is the line area. It consists of seven horizontal and seven vertical lines, one quarter inch long and ranging in width from one to seven printer dots in one dot intervals at 300 dpi printing resolution (equivalent to 85 to 600 μm). Each line is separated from its nearest neighbor by about 1000 μm or 12 printer dots at 300 dpi. This arrangement prevents the adjacent character fields (see definition in "Description of Metrics") from interfering with each other.

Solid Fill Area Under the line area is the solid fill area, simply a one inch square area of solid colorant.

1	0	1	0	1	0	1	0
0	0	0	0	0	0	0	0
1	0	1	0	1	0	1	0
0	0	0	0	0	0	0	0
1	0	1	0	1	0	1	0
0	0	0	0	0	0	0	0

Figure 4. Pattern of the tint solid (halftoned) area of test target.

Tint Solid Area Below the solid fill area is the tint solid area. This area consists of a 1 in. square area of a standard quarter-fill pattern (every other dot is printed in each spatial direction at the full resolution of the printer). Figure 4 shows a portion of the bitmap printed in this area.

Background Area Beneath the tint solid area is the background area. This is a non-printed region of the substrate.

Scanner Calibration and MTF Testing

Although the ISO/IEC document makes no explicit reference to calibration of the target acquisition device to some absolute standard, most of the values of the metrics are dependent on such a calibration. For example, all of the "darkness" metrics are directly affected by calibration. Also, those line metrics that use thresholds, T_{pp}, (see the section on Description of Metrics) to determine whether a particular pixel belongs to a line segment are indirectly affected by scanner calibration, as differing calibrations lead to different line boundaries. In the Scanner Calibration section, we discuss the methodology followed for calibrating the scanner used in our experiments.

Also of concern when doing high-resolution imaging is the high-frequency response of the imaging device. The standard metric for evaluating such performance from imagers is the Modulation Transfer Function (MTF). Our procedure for evaluating the MTF of our scanner is covered in the section on MTF Evaluation.

Scanner Calibration. The scanner calibration was performed as follows:

1. Twenty-five three-eighths-inch square test patches with gray-values ranging from 0 to 1 were generated in Matlab. These patches were set together in a matrix with ample (≈ 1/4 in.) border between each patch and saved in TIFF file format.
2. The TIFF file was printed on a laser printer using the printer driver's halftoning technique at 600 dpi.
3. The patches were measured with a (calibrated) Gretag SPM-50 spectrophotometer for optical density values (ANSI standard). Ten measurements were taken for each patch, and the results were averaged. The resulting density was converted to reflectance (0-1) values, then rescaled to fall in the range 1-256 for consistency with Matlab's image format.
4. The patches were then scanned using a Howtek D4000 scanner at 1000 dpi. The resulting patch images were cropped to avoid edge effects and spatially averaged over the remaining area. This resulted in a scaled reflectance value (1-256) for each patch.

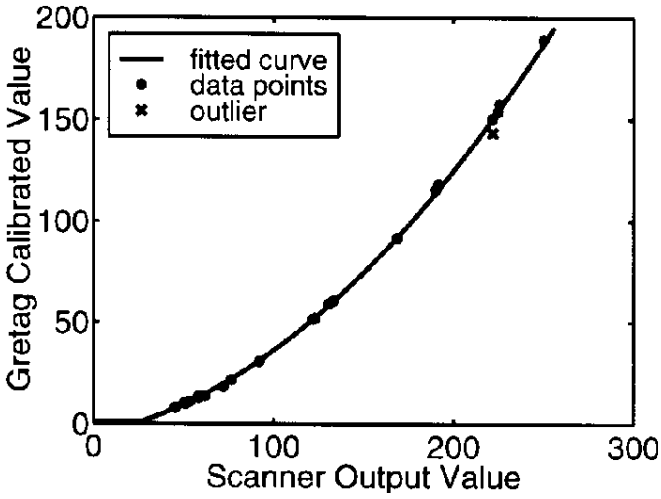


Figure 5. Raw data and fitted curve for the Howtek D4000 drum scanner calibration. Encircled data point was omitted from polynomial fit calculations, and is presumed due to measurement error.

5. The Howtek data was fitted to the Gretag data using a second-order (quadratic) polynomial via the Matlab "polyfit" function. This yielded the coefficients $c_0 = 2.650 \bullet 10^{-3}$, $c_1 = 1.242 \bullet 10^{-1}$, and $c_2 = -2.526$ that relate the Gretag G and Howtek H data by the following equation:

$$G = c_2 \times H^2 + c_1 \times H + c_0.$$

6. Whenever an image file is read by the PQTK program, it is calibrated using these coefficients to generate a look-up table that is then used to map the raw image to a calibrated image.

The raw data and the fitted polynomial curve are shown in Fig. 5. Note that the scanner calibration is dependent on both the media and the colorant. Ideally, it should be redone when either of these components are changed.

MTF Evaluation. The scanner MTF was determined using a Sine Patterns M-6-80 target.[21] This is a transmissive film imprinted with sine patterns of various frequencies (0.375 to 80 cycles/mm) at 80% modulation,

$$M = \frac{A - B}{B},$$

where A is the peak transmittance of a sine pattern, and B is the mean transmittance of the pattern.

One could calculate the MTF directly from scanned data, but according to the engineering notes supplied by the target manufacturer,[22] this approach is prone to system noise and any "wedging" effect, or slow drift of the mean transmittance of the target. Rather, we used the suggested Fourier approach as follows:

1. The test target includes fourteen solid gray areas with known optical densities, that we used to calibrate the Howtek D4000 scanner for transparencies, exactly in the same manner where the reflectance calibra-

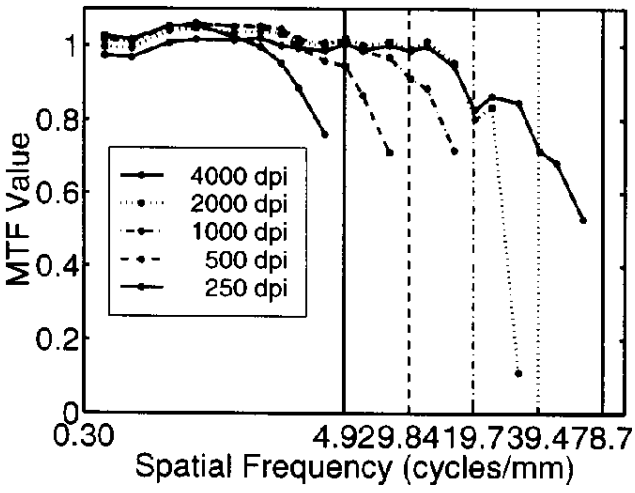


Figure 6. Results of MTF measurements on the Howtek D4000 drum scanner. The vertical lines represent the Nyquist frequencies for the respective scanning resolutions.

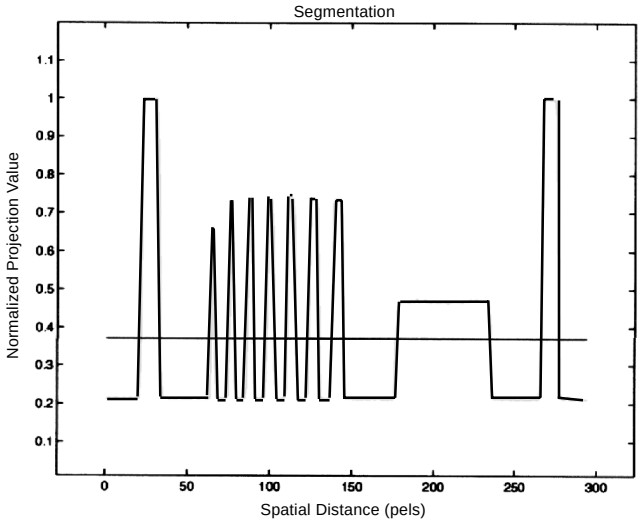


Figure 7. Screen capture of the segmentation routine in demo mode, showing the test target segmentation in progress.

in the same manner where the reflectance calibration was performed in the subsection on scanner calibration.
2. For scanning resolutions of 4000, 2000, 1000, 500, and 250 dpi, each of the test patterns with spatial frequencies up to the Nyquist limit was scanned.
3. The width of each scan was adjusted to begin and end at, as closely as possible, the same point in the sinusoidal cycle of the test targets, so that each scan contained an integer number of cycles.
4. When the Fast Fourier Transform (FFT) of a sample with an integer number of cycles is taken, the modulation is simply $2F(n) = F(0)$, where F is the FFT of the sample, and n is the number of cycles in the sample.
5. The MTF is simply the measured modulation divided by the actual modulation, that is provided by the target manufacturer.
6. We then have the MTF as a function of spatial frequency for each scanned resolution.

The results of these calculations are shown in Fig. 6. In our judgement, the scanner exhibited sufficiently flat high-frequency response to perform accurate scans at all resolutions, up to the Nyquist limit. However, if a flatter overall frequency response were desired, the PQTK program could be modified to correct the frequency characteristics of the scans via digital filtering. We have not done this for the results reported in this paper. It is possible that the scanner MTF might differ between the transmissive and reflective modes. However, with our scanner, the position of the illuminating source is the only thing that changes between reflective and transmissive modes. Any change in MTF due to scattering in the reflective mode would be attributed to the target being measured and not the instrument itself.

Description of the PQTK Tool

Overview of Procedure. Analysis of print quality with the PQTK involves the following steps:

1. The test target is stored in compressed TIFF file format, one bit per pixel. It is then converted to Postscript language and sent to a printer, or imported into a drawing/graphics program and printed directly, depending on the computer platform. The resolution is chosen in the software so that the target is printed at the correct size, nominally 1×3.5 in^2.
2. The test target is converted to hardcopy by the printer.
3. The target hardcopy is then mounted on a scanning device, and scanned at the desired resolution. The scan is then saved to disk in uncompressed TIFF format.
4. Once scanned and saved to disk, the reproduced test target is read and analyzed by the PQTK software.
5. After analysis by the software, each metric is assigned a numerical value, that can be stored in a computer file or viewed interactively through the software.

Program Methodology. Within the PQTK program are several substages, some of which employ a unique methodology to achieve an aspect of the general goal: to map portions of the scanned data to image elements, and apply the ISO/IEC metrics to these elements. A description of the substages that use such unique methodologies follows.

The inverse relation between aperture and depth of field is well-known in photography and imaging. Thus, at high (≈ 2000 dpi or more) scanning resolution, the out-of-roundness of the drum on our scanner necessitated scanning smaller parts of the image separately. Rather than trying to recombine them into a contiguous image, we allow within the program the ability to read the separate sections of the test target from disk.

Another complication at higher scanning resolutions is the memory required to load the scanned test target data into the program—at 4000 dpi, a 1×3.5 in^2 scan with boundary requires upwards of 448 megabytes of memory in the Matlab 4.x environment. We developed a functionality called *Partial Image Read* to address this difficulty by allowing segments of the test target to be loaded into the program and analyzed separately.

The role of the segmentation portion of the PQTK program is to determine the indices of different areas of the image: all of the vertical and horizontal lines and the tint solid, solid fill, and background regions. This function is made much easier due to the fact that the structure and dimensions of the test page are known *a priori*. The segmentation subroutine uses projections,

$$\frac{1}{16}\begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix} = \frac{1}{4}\begin{bmatrix}1&1\\1&1\end{bmatrix} ** \frac{1}{4}\begin{bmatrix}1&0&1&0\\0&0&0&0\\1&0&1&0\\0&0&0&0\end{bmatrix}$$

Figure 8. Realization of 4×4 unweighted moving average as a composition of a 2×2 unweighted moving average with a 4×4 sparse kernel. The double asterisk denotes 2D convolution.

target in one of the dimensions, to determine the location of region limits. A capture of the action window during the segmentation routine is reproduced in Fig. 7. Here we see the second stage of segmentation—the vertical extent of the various regions of the test page have been determined, and summing vertically down the line region produces the plot shown. From here, it is a simple matter to extract the horizontal boundaries of the individual vertical lines and the horizontal boundaries of the horizontal line area. The process is then repeated in different directions and regions to completely segment the image.

A secondary role of the segmentation routine is to check the skew of the scanned test page, that is accomplished by extracting the bottom-most boundary segment from the image and computing its slope. The ISO/IEC specify that the test-page should have no more than a ±1:5 degree skew. We did not have difficulty aligning the target to be scanned on the drum within these tolerances. No skew correction was provided for the metrics, as ±1:5 degrees of skew causes less than 0.035% error in distance measurements.

Special Image Processing Techniques. There are instances where particular image processing techniques are employed to minimize processing time or conserve memory, or both. This section describes these techniques. All Fourier transform techniques used are found in Ref. 23.

Separable Filtering and Efficient Data Usage. The *mottle* and *LADV* metrics are calculated by area averaging operations, that are computationally intensive on large images. We optimized these computations in two steps:

1. Using separable averaging filters. By convolving the area first with an $N \times 1$ averaging window and then convolving the resulting array with a $1 \times N$ averaging window, the same result is obtained as if the original image is convolved with an $N \times N$ averaging window, but with substantial computational savings.
2. Using the results of the mottle convolution as data for the LADV convolution. Fortunately, the LADV averaging window (4×4 mm^2) is exactly twice the size of the averaging window for mottle in each dimension (2×2 mm^2). We can realize an unweighted moving average over a $2N \times 2N$ pixel window (the LADV calculation) by convolving the results of an unweighted moving average over an N × N pixel window (the mottle calculation) with a special $2N \times 2N$ pixel sparse kernel containing ones at locations (0; 0), (0; N), (N; 0), and (N; N), with zeroes elsewhere. Figure 8 illustrates this for the case of $N = 2$. In the actual implementation, only the computations involving nonzero terms in the sparse kernel are performed.

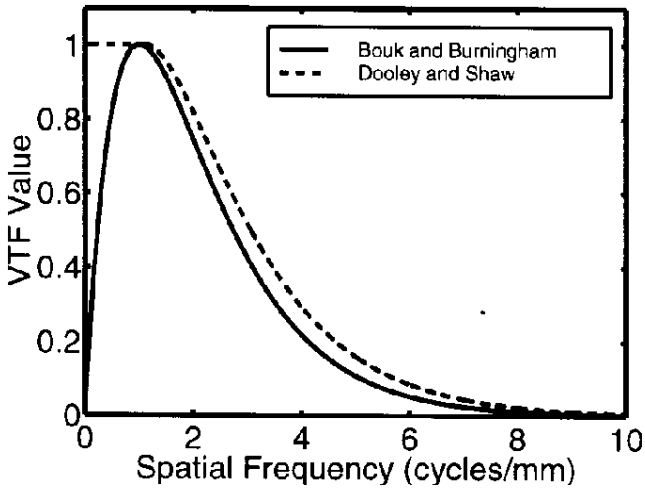


Figure 9. Visual transfer functions for the human visual system found in the literature on print quality assessment.

Granularity Calculation and Optimization. Granularity is taken in the literature to be the objective measure of graininess, which is a subjective quality, both concepts being derived directly from the field of photography. As mentioned in the Metrics section, we derived our approach to granularity primarily from Dooley and Shaw.[7] In addition, following Bouk and Burningham,[9] we added a notch filter to the equation to remove the frequency components of the underlying halftone pattern. Both references report high correlations between measured granularity and perceived graininess.

The PQTK program analyzes a 0.5×0.5 in.^{2} segment of the tint solids area for the graininess metric, which is an area small enough to avoid any boundary effects in the 1×1 in.^{2} tint solids area, while being large enough to include significant low frequency content.

The resulting equation implemented in the PQTK is:

$$G = e^{-1.8D}\sqrt{\sum_{u,v}\left[W(u,v)\left(1-\hat{W}_{ht}(u,v)\right)V(u,v)^2\right]}, \qquad (1)$$

where

D = is the mean optical density of the area,
u and v = the discrete spatial frequencies in the area,
$V(u, v)$ = the visual transfer function of the human visual system,
$W(u, v)$ = the Weiner power spectrum of the area,
$\hat{W}_{ht}(u,v)$ = is the normalized Weiner power spectrum of the halftone present in the area.

A discussion of the various quantities follows:

$e^{-1.8D}$ This quantity, a function of the mean optical density of the area, compensates for the effect that darker areas appear less grainy than lighter ones.

$V(u, v)$ The visual transfer function (VTF) we used was from Dooley and Shaw,[7] which incorporates flat low-frequency (below 1 cycle/mm) response. A plot of this VTF, as well as that used by Bouk and Burningham,[9] appears in Fig. 9. These VTFs assume a viewing distance of 350 mm. Subjective

testing by Hino et al.[8] revealed high correlation between graininess and granularity for either choice of VTF in Fig. 9. As a result of weighting the metric by the human visual system's response, we can limit the frequency extent of the three quantities under the summation in Eq. 1, because $V(u, v)$ above 10 cycles/mm is negligible (≤ 0.0026). Thus, at scanning resolutions above 508 dpi, we begin limiting the size of the two-dimensional quantities $V(u, v)$, $W(u, v)$, and $\hat{W}_{ht}(u,v)$. This data reduction allows savings in storage space and computational complexity.

$W(u, v)$ By Fourier identities, we have:

$$W(u, v) = \mathcal{F}\{C_{\Delta R}(x, y)\} \tag{2}$$

where
$C_{\Delta R}(x, y)$ is the autocorrelation function of $\Delta R(x, y)$,
$\Delta R(x, y) = I(x, y) - \overline{I(x, y)}$, the reflectance fluctuation,
$I(x, y)$ is the reflectance image, and $\overline{I(x, y)}$ is its mean.

We also know by Fourier identities that

$$C_{\Delta R}(x, y) = \mathcal{F}^{-1}\{\mathcal{F}\{\Delta R(x, y)\}\mathcal{F}^*\{\Delta R(x, y)\}\} \tag{3}$$

Therefore, by combining Eqs. 2 and 3, we arrive at the simplified equation for calculating $W(u, v)$:

$$W(u, v) = \mathcal{F}\{\Delta R(x, y)\}\mathcal{F}^*\{\Delta R(x, y)\} \tag{4}$$

The calculation of $W(u, v)$ therefore involves only the calculation of $\Delta R(x, y)$, which is an averaging and subtraction operation, $\mathcal{F}\{\Delta R(x, y)\}$, which is a Discrete Fourier Transform, and $\mathcal{F}^*\{\Delta R(x, y)\}$, a complex conjugation operation. Using the simplifications inherent in Eq. 4 rather than the direct method of Eq. 2 to compute $W(u, v)$ leads to a large savings in computational complexity, as Eq. 2 contains an implicit convolution of the reflection fluctuation image with itself.

$\hat{W}_{ht}(u,v)$. Because the halftone pattern used in the tint solids area is known (see Fig. 4), it is straight forward to calculate its Weiner power spectrum. We simulate the halftone pattern as a 2-D sinusoidal pattern, that contains only the fundamental frequencies of the halftone. Then Eq. 4 is applied to this pattern to calculate the Weiner power spectrum of the halftone.

The granularity metric requires us to take the FFT of a potentially large 2-D data set, whose size ranges from 125×125 to 2000×2000, depending on the scanning resolution. Thus, there are severe computational and memory allocation issues to consider. To this end, we have enacted the following steps to minimize such requirements:

- Because of the reduction in complex multiplies and additions when using a radix-two FFT routine and the corresponding large reduction in computation time, we use only power-of-two processing for the FFT operations.
- We calculate granularity last among the metrics that use the tint solids data for their calculation. Thus, we simply modify the array already allocated for storage to hold the FFT calculation. The separability of the FFT operation makes this possible.

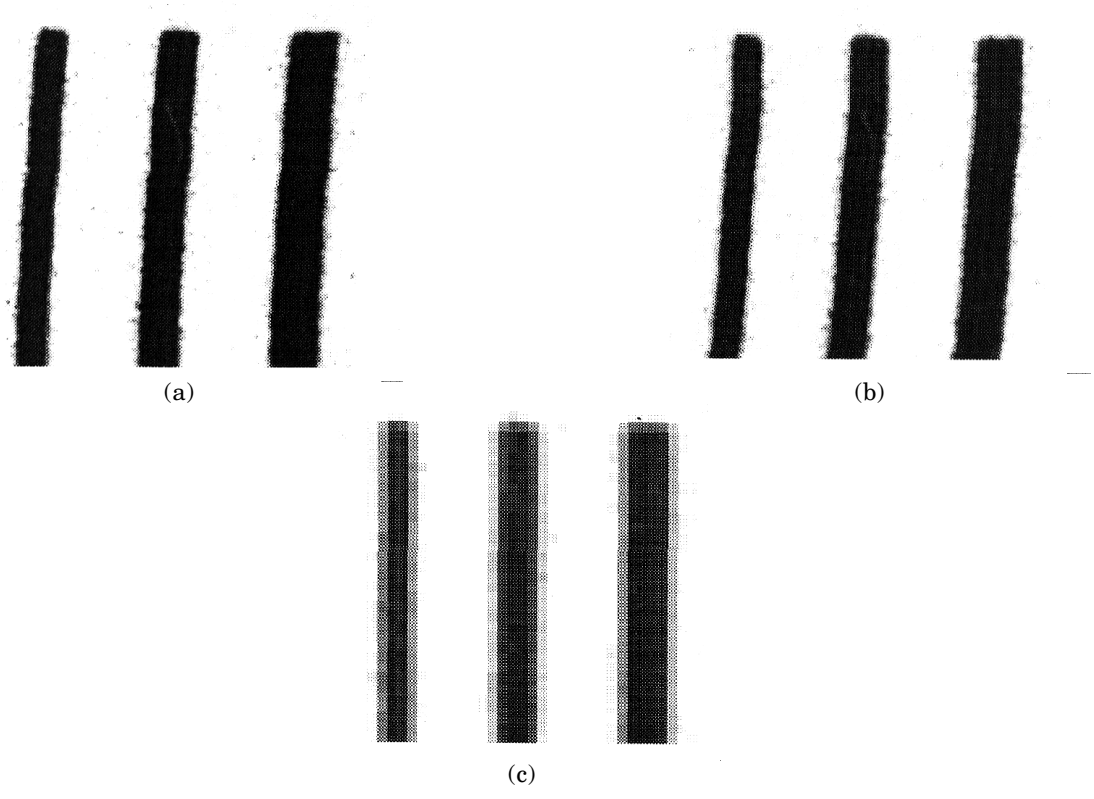
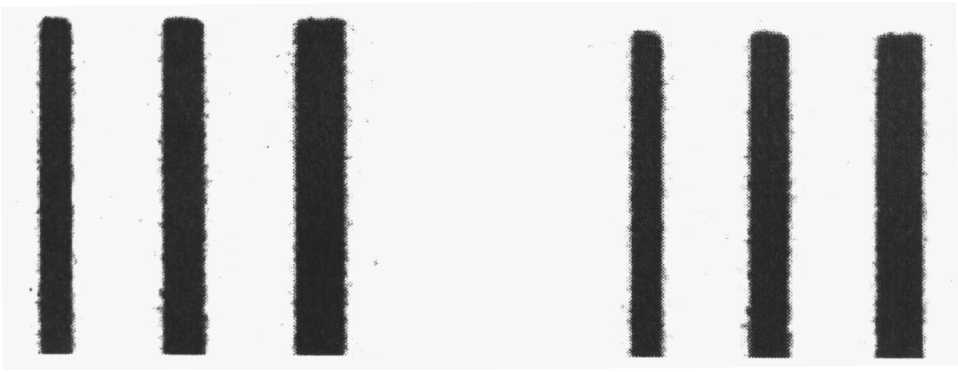
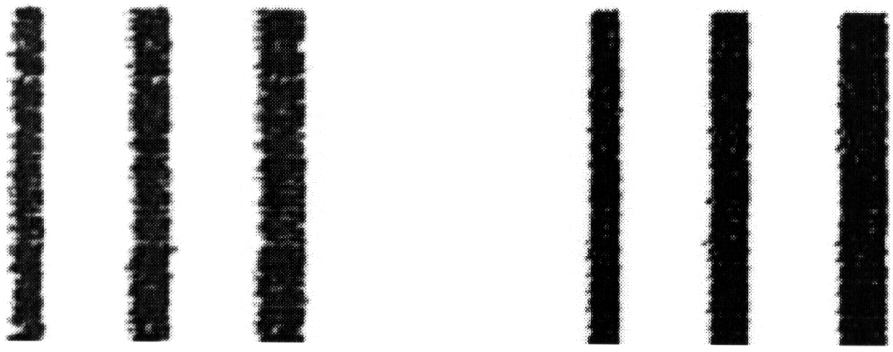


Figure 10. Vertical line segments from scan of target on standard paper and laser printer at scanning resolutions of (a) 4000 dpi, (b) 1000 dpi, and (c) 250 dpi. Effective magnification is approximately 12.



(a) (b)

Figure 11. Vertical line segment from 4000 dpi scan of target on a laser printer on (a) coated paper and (b) cotton bond. Effective magnification is approximately 12.



(a) (b)

Figure 12. Vertical line segment from 4000 dpi scan of target on an inkjet printer on (a) standard paper and (b) special inkjet paper. Effective magnification is approximately 12.

- As mentioned in the subsection on granularity, the human visual transfer function can be truncated at 10 cycles/mm, thus the FFT can also be truncated, reducing the size of the data set.
- Because the reflection fluctuation data is real, the resulting 2-D FFT is symmetrical about the zero-lag point (or DC). Therefore, we only need to calculate two of the four quadrants surrounding the zero-lag, and can then reconstruct the other two quadrants based on this symmetry.
- In those instances where we must use the partial image read functionality, the entire image data for the tint solids area are not available. To remedy this, we read the individual columns of data from the image file, convert them to intensity fluctuations, and take their FFT. We then take the FFT of the rows, and are left with the 2-D FFT.

Experimental Results

We performed three experiments to demonstrate potential applications for the PQTK. In each experiment, we printed the test target specified in the Description of Test Target section on the printer and paper pertinent to that experiment, and scanned each printed target at 4000, 2000, 1000, 500, and 250 dpi with the Howtek D4000 drum scanner. The resulting scans were calibrated for absolute reflectance using the correction curve obtained with the Gretag SPM-50 Spectrophotometer, then analyzed by the PQTK program. We only did a single calibration for the entire set of scans, using output from the laser printer on "standard" paper. The three experiments differ as follows:

1. In the first experiment, we printed ten samples of the test target on a laser printer on "standard" printer paper. This paper was chosen to represent printer paper likely to be used in an office or personal printer. To illustrate the differences caused by changing the scan resolution, Fig. 10 shows segments of the vertical line section from this experiment scanned at 4000, 1000, and 250 dpi.
2. In the second experiment, we printed ten samples of the test target on the laser printer on coated paper with a very smooth, non-porous surface, and on paper with a rough, porous surface (cotton bond). Figure 11 shows vertical line segments from this experiment scanned at 4000 dpi.
3. In the third experiment, we printed ten samples of the test target on an inkjet printer on the aforementioned standard paper, and on a paper formulated for inkjet printers. Figure 12 shows a segment of the vertical line section from this experiment, scanned at 4000 dpi.

We then analyzed the behavior of a subset of the line metrics, namely darkness, fill, raggedness, and stroke width. These metrics were studied for three types of dependencies: resolution dependence, line width dependence, and printer technology and paper type dependence. These dependencies are roughly categorized as weak (< 10% change), moderate (< 50% change), and

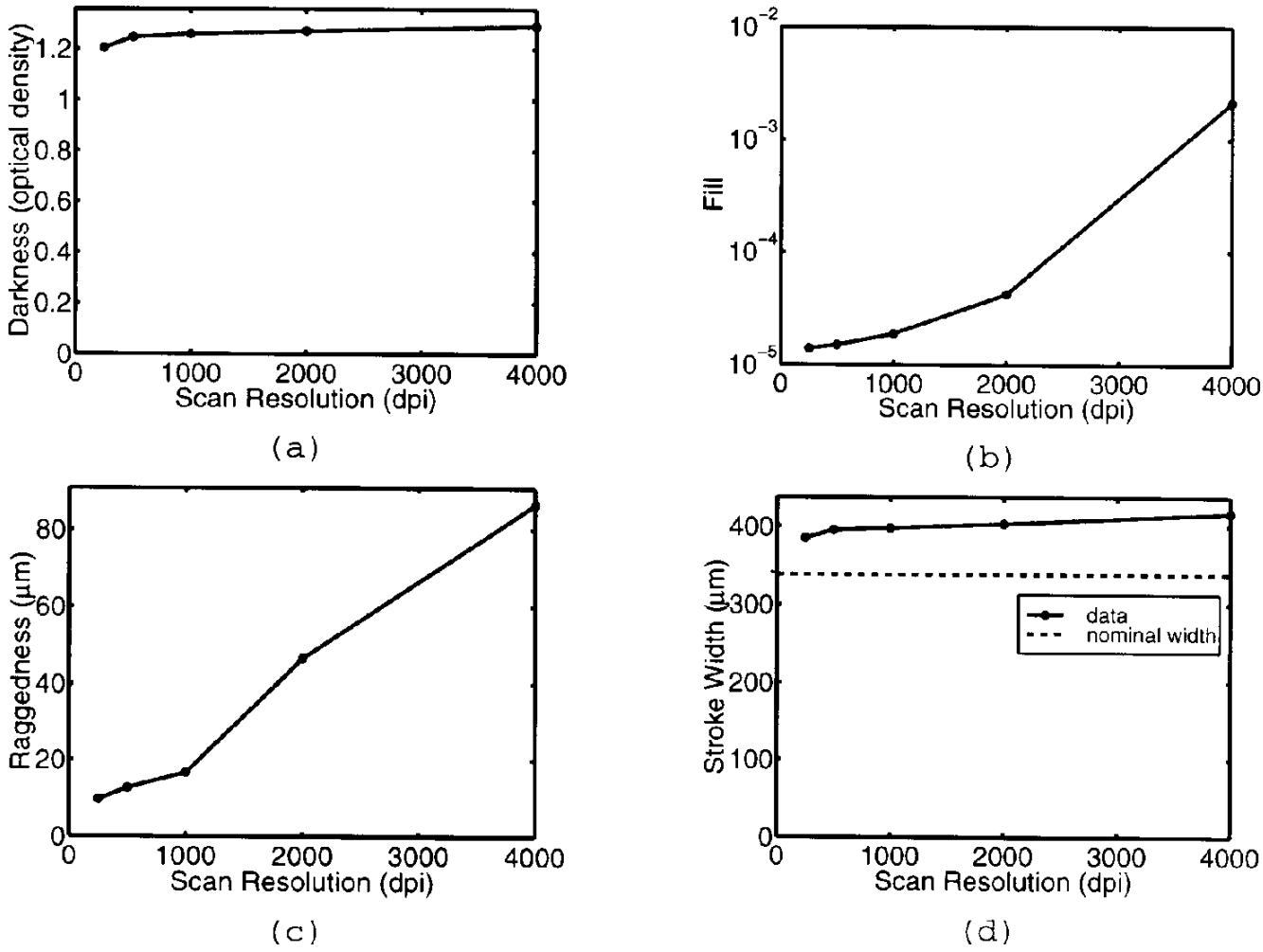


Figure 13. Effect of scanning resolution on the (a) darkness, (b) fill, (c) raggedness, and (d) stroke width metrics.

strong (≥ 50% change). Finally, Table I lists all the metrics and their dependencies.

Dependence of Metrics on Scanning Resolution

This section will investigate the dependencies of the chosen metrics on specific scanner resolutions. Figure 13 shows the plots of these metrics vs. scanning resolution. The data used for this analysis are those from the standard paper and a nominal printed line width of 4/300 in (the average line width of the target). We then look at each metric as a function of resolution only.

Darkness versus Resolution. A plot of darkness versus scanning resolution is shown in Fig. 13(a). From the figure, it is evident that the darkness metric is only weakly dependent on scanner resolution. This result is logical, because darkness is an area-averaging metric, averaging pixels over a relatively large area, in this case, the interior and exterior of a line segment to arrive at the metric value. Therefore, increasing the scanning resolution has a negligible effect on the metric.

Fill versus Resolution. A plot of fill versus scanning resolution is shown in Fig. 13(b). The value of the fill metric is strongly dependent on scanning resolution, increasing in an exponential fashion above 500 dpi (the fill metric is 0 at and below 500 dpi). This effect is easily explained considering the measurement technique for the fill metric—finding individual pixels with reflectances above a certain threshold. With higher scan resolutions, and thus less area averaged per data point, the probability of finding such voids increases dramatically.

Raggedness versus Resolution. A plot of raggedness versus scanning resolution is shown in Fig. 13(c). It can be seen that the metric value increases nearly linearly with the scanning resolution, thus this metric is strongly dependent on scanning resolution. At lower scanning resolutions, the blurring effect of the scanner aperture reduces edge raggedness, thus, variations in the scanning resolution have a profound effect on this metric.

Stroke Width versus Resolution. A plot of stroke width versus scanning resolution is shown in Fig. 13(d). From the figure, it can be seen that stroke width is only weakly dependent on resolution. Thus, blurring of the stroke edge does not significantly shift the location of the T_{60} threshold used to define this edge.

Note that the nominal line width is 339 μm, as shown in the figure.

Dependence of Metrics on Line Width

This section investigates the dependencies of the chosen metrics on the width of the line being analyzed. Fig.

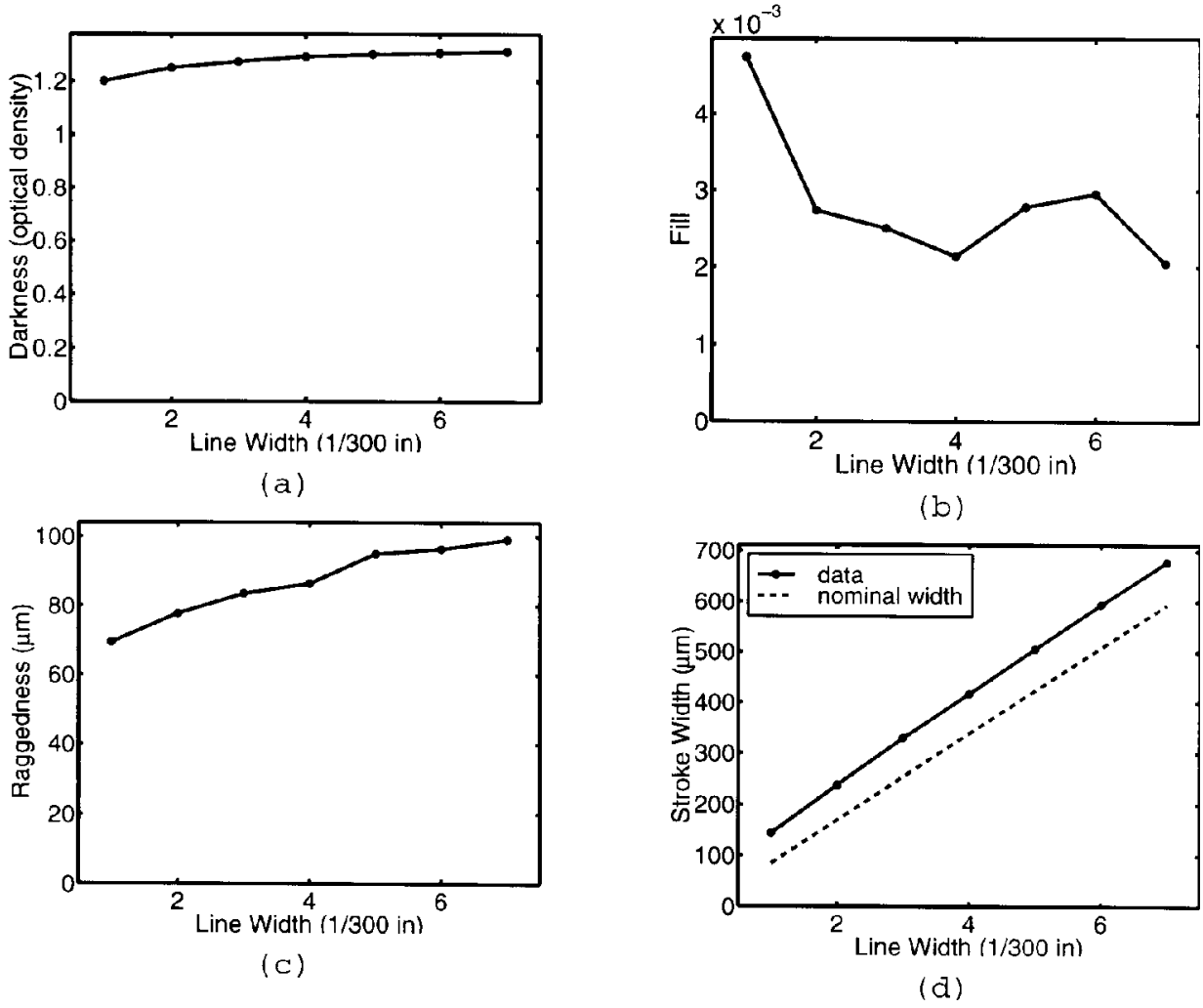


Figure 14. Effect of line width on the (a) darkness, (b) fill, (c) raggedness, and (d) stroke width metrics.

14 shows the plots of these metrics versus line width. Note the distinction here between line width, the theoretical width of the line when it is printed, and stroke width, a metric for the actual width of the line as rendered by a printer and analyzed and measured by a scanner and the PQTK tool. The data used for this analysis are the standard paper and a scanning resolution of 4000 dpi. We then look at each metric as a function of line width only.

Darkness versus Line Width. A plot of darkness versus line width is shown in Fig. 14(a). As was the case with darkness versus scanning resolution, darkness is seen to have only a weak dependence on line width. Again, we conclude that the area-averaging nature of the metric is responsible for this independence.

Fill versus Line Width. A plot of fill versus line width is shown in Fig. 14(b). We see that the metric exhibits a moderate to strong dependence on line width, depending on whether the 1/300 in. wide line is included in the analysis. The line width dependence, however, is nonmonotonic, and hence, unpredictable.

Raggedness versus Line Width. A plot of raggedness versus line width is shown in Fig. 14(c). The figure shows distinct, monotonically increasing raggedness with respect to line width, that falls into the category of moderately dependent.

Stroke Width versus Line Width. The plot of stroke width versus line width is shown in Fig. 14(d). As one would expect, there is an almost perfectly linear relationship (strong dependence) between stroke width and line width. Stroke width therefore has a strong dependence on line width.

The nominal line width is plotted in the figure. Note that there is a nearly constant offset (or error) between the measured and theoretical line widths. This effect is presumed to be caused by dot gain.

Dependence of Metrics on Printer Technology and Paper Type

This section will investigate the dependencies of the chosen metrics on (a) the type of paper used, keeping the printer used constant, and (b) the printer technology used, keeping the paper type constant. The data used for (a) are those obtained from the laser printer with standard, coated, and cotton bond paper and the inkjet printer with standard and special paper. The data used for (b) are those obtained from the laser and inkjet printer with standard paper. Both experiments used

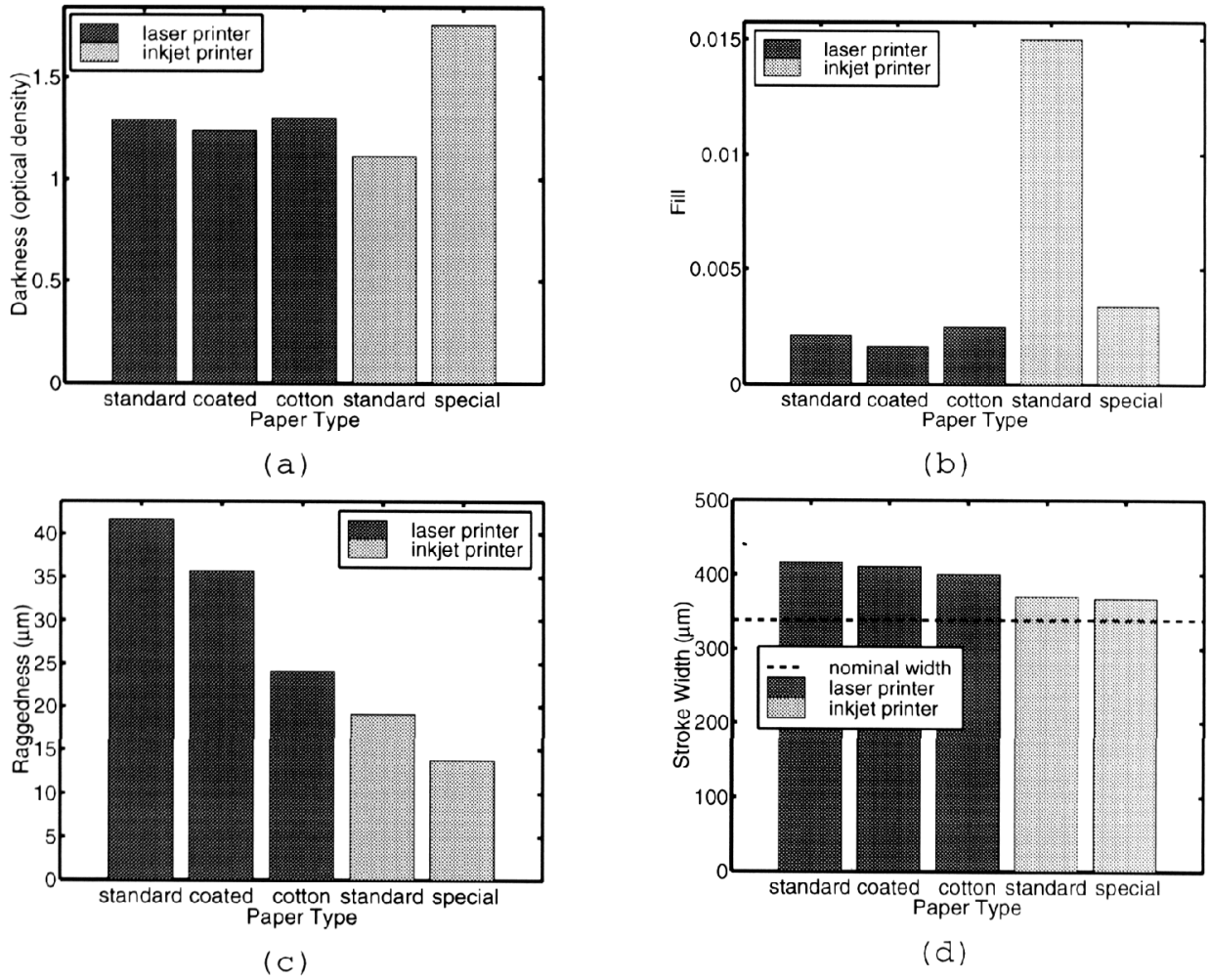


Figure 15. Effect of paper type and printer technology on the (a) darkness, (b) fill, (c) raggedness, and (d) stroke width metrics.

TABLE I. Summary of Dependencies of Metrics

	Type of Dependence			
Metric	Scan Res.	Line Wid.	Paper	Print Tech.
Line Metrics				
Blurriness	moderate	moderate	moderate	moderate
Contrast	weak	weak	weak	weak
Darkness	weak	weak	weak/moderate	moderate
Fill	strong	strong	moderate	strong
Haze	weak	weak	weak	weak
Marks	weak	weak	weak	weak
Raggedness	strong	moderate	moderate	moderate
Stroke Width	weak	strong	weak	weak
Solid-Fill Metrics				
Darkness	weak	—	weak/strong	moderate
LADV	moderate	—	moderate/strong	strong
Mottle	moderate	—	strong	strong
Voids	strong	—	moderate/strong	strong
Tint Solid Metrics				
Darkness	weak	—	weak	moderate
Granularity	moderate	—	—	—
LADV	moderate	—	moderate	moderate
Mottle	weak	—	moderate/strong	moderate
Background Field Metrics				
Marks	strong	—	strong	strong
Uniformity	strong	—	strong/moderate	weak

4000 dpi scan resolution and a 339 μm nominal line width. We then look at each metric as a function of printer technology and paper type only. The plots of the metric subset versus paper type and printer technology are shown in Fig. 15.

Darkness versus Technology/Paper. The plot of darkness versus technology/paper is shown in Fig. 15(a). It is apparent that darkness is only weakly dependent on paper type in the laser printer, but is moderately dependent on paper type in the inkjet printer. Any conclusions must be viewed with caution, however, because the paper types were not the same on the respective printers. The overwhelmingly darker lines of the inkjet and special paper combination attests to the special paper's ability to produce solid colorant areas well. The metric is also moderately dependent on printer technology.

Fill versus Technology/Paper. The plot of fill versus technology/paper is shown in Fig. 15(b). This figure shows that fill has a moderate dependence on paper type in the laser printer, and a large dependence in the inkjet printer. Fill also exhibits a large dependence on print technology, owing to the poor fill characteristics of the inkjet printer on the standard paper. Even with the special paper, the inkjet technology has more voids (the inverse of fill) than any of the laser/paper combinations.

Raggedness versus Technology/Paper. The plot of raggedness versus technology/paper is shown in Fig. 15(c). The raggedness metric exhibits a strong dependence on paper type in both printers. In the laser printer, the ordering of the papers is nearly counter-intuitive: the roughest paper had the best raggedness values. One plausible explanation for this phenomenon is that the toner particles adhere to the cotton bond better, reducing toner displacement after transfer as the paper moves through the printer. Raggedness also showed a strong dependence on printer technology, as the inkjet sample measured much less ragged than the laser sample.

Stroke Width versus Technology/Paper. The plot of stroke width versus technology/paper is shown in Fig. 15(d). The nominal line width is plotted on the figure as a dotted line. The effects of paper type on stroke width are weak. Stroke width is clearly dependent on printer technology, based on the figure. It is interesting to note that the inkjet technology reproduced the theoretical line width much more accurately than the laser printer. This effect could be due to the more numerous voids on the inkjet/standard paper combination that artificially narrow the printed line, and due to the superior ability of the special paper to prevent the ink from wicking away from the desired area.

Table of Metric Dependencies. Table I shows the dependencies of the metrics on scanner resolution, line width, printer technology and paper as determined by the experiments outlined in this section.

Conclusions

We have described a system for objective print quality assessment that implements the ISO/IEC metrics specified in document 13660. This system uses a drum scanner to capture an image of a printed test target, that was designed via the ISO/IEC specifications. The program is written in Matlab, for platform portability and ease of modification, and includes many optimizations to conserve memory usage and speed up calculations.

We have also shown that the metrics in the ISO/IEC document exhibit varying degrees of dependence on scanning resolution, line width, printer technology and paper type. These investigations are good examples of applications where the PQTK may be useful. Based on these results, we offer the following recommendations:

- The ISO/IEC document specifies a scanning resolution of 400, or optionally 600 dpi. In regards to our experimental results, we find that resolution to be too low for some of the metrics. Interestingly, some metrics may be better measured at lower scanning resolutions, for example voids. A scanning resolution that is too high may identify such voids in the printed image that cannot be discerned by a viewer at normal viewing distances.
- Line width is not specified in the ISO/IEC document. Because we found that many metrics were dependent on line width, we recommend choosing line width based on the specific application. For example, an investigation in printed character quality should use line widths approximately the width of the characters.

References

1. J. R. Edinger and C. Newell, Subjective Impression of the Weight of Text Correlates with Measured L-Width, *J. Imag. Sci. Technol.* **41,** 174–177 (1997).
2. D. L. Lee, A. T. Winslow and S. Bares, Performance of Three Image Quality Merit Factors Applied to Inkjet Printing on Plain Papers, *Proc. Eighth International Congress on Advances in Non-Impact Printing Technologies*, IS&T, Springfield, VA, 1992.
3. P. G. Engeldrum, A Framework for Image Quality Models, *J. Imag. Sci. Technol.* **39,** 312–318 (1995).
4. P. Wang, Print Image Quality Evaluation System, *Proc. Tenth International Congress on Advances in Non-Impact Printing Technologies*, IS&T, Springfield, VA, 1994.
5. Y. Kipman, Image Quality Tests for Printers, *Proc. IS&T's 50th Annual Conference*, IS&T, Springfield, VA, 1997.
6. Document B123: NP 13660—Office equipment—Measurement of image quality attributes for hardcopy output {Binary monochrome text and graphic images, 7th Working Draft, ISO/IEC, Geneva, Switzerland, 1995.
7. R. P. Dooley, and R. Shaw, Noise Perception in Electrophotography, *J. Applied Photo. Eng.* **5,** 190 (1979).
8. M. Hino, K. Kagitani and M. Kojima, A New Method for the Graininess Evaluation, *Proc. Eleventh International Congress on Advances in Non-Impact Printing Technologies*, IS&T, Springfield, VA, 1995.
9. T. Bouk and N. Burningham, Measurement of Graininess for Halftone Electrophotography, *Proc. Eighth International Congress on Advances in Non-Impact Printing Technologies*, IS&T, Springfield, VA, 1992.
10. C. A. Dvorak, Text Sharpness, Its Components and Text Quality, *J. Applied Photo. Eng.* **9,** 109–111 (1983).
11. J. R. Hamerly and C. A. Dvorak, Detection and Discrimination of Blur in Edges and Lines, *J. Opt. Soc. Am.* **71,** 448–452 (1981).
12. Lars Jansson, Print Quality, *Byte,* 199–207 (September 1987).
13. Y. Tanaka and T. Abe, Quantitative Analysis of Print Quality Features, *J. Imag. Tech.* **13,** 202–207 (1987).
14. L. B. Schein and G. Beardsley, Offset Quality Electrophotography, *J. Imag. Sci. Tech.* **37,** 451–461 (1993).
15. J. R. Edinger, The Image Analyzer—A Tool for the Evaluation of Electrophotographic Text Quality, *J. Imag. Sci.* **31,** 177–183 (1987).
16. J. R. Hamerly and R. M. Springer, Raggedness of Edges, *J. Opt. Soc. Amer.* **71,** 285–288 (1981).
17. J. L. Crawford, C. D. Elzinga and R. Yudico, Print Quality Measurements for High Speed Electrophotographic Printers, *IBM J. Res. Dev.* **28,** 276 (1984).
18. C. A. Dvorak and J. R. Hamerly, Just-Noticeable Differences for Text Quality Components, *J. Applied Photo. Eng.* **9,** 97–100 (1983).
19. J. R. Edinger, Jr., A Measure for Stairstepping in Digitized Text that Correlates with the Subjective Impression of Quality, *J. Imag. Sci. Tech.* **39** (1985).
20. P. E. Tobias, J. Ricks and M. Chadwick, Objective, Reproducible Measurement of Printing Mottle with a Mottle Tester, *Tappi J.* **72,** 109–112 (1989).
21. Sine Patterns, 236 Henderson Drive, Penfield, NY, 14526.
22. R. L. Lamberts, *Use of Sinusoidal Test Patterns for MTF Evaluation*, Engineering Notes provided by Sine Patterns.
23. J. G. Proakis and D. G. Manolakis, 2nd ed., *Digital Signal Processing: Principals, Algorithms, and Applications*, Macmillan, New York, 1992.