Application of Digital Image Analyses to Measure Print Quality

William Lim and Suresh Mani—Flint Ink Research Center*

An important aspect of poor print quality in high-speed printing is color density variation in solid print areas, which is described as mottle. Armel and Wise provide a more detailed description and the various types of mottle.[1] Traditionally, these variations were assessed by visual side-by-side comparisons. These assessments were influenced by subjective factors of human vision, such as fatigue and variations in color interpretation between different people. Therefore, an objective method of mottle measurement was needed to identify the sources of mottle and plan corrective action. In recent years, personnel at Heidelberg Web Press developed the concept of an absolute scale by selecting representative samples at various levels of mottle. *Figure* 1 shows the samples along with the Heidelberg Web Press numerical ranking. Although this helped establish a scale, it was still limited to the need of a side-by-side comparison for a proper rating, which would require the analyst to have the standard available. This is not always the case, nor would it be practical to continually send it from one location to another. Reproduction of the standard would also be impractical because, in addition to error from differences in human perception, artifacts caused by the reproduction process would introduce errors.

Because the technique of choice needs to be quantitative, objective, and reproducible, digital image analysis was chosen as suggested by the study of dirt specks by Rosenberger[2] and ink jet print quality measurements by Kowalczyk and Trksak.[3] In this paper, we describe an image analyses system that satisfies the previously mentioned criteria and costs less than $3,000. The resolution and related features of the components is also described. A procedure to quantify print mottle using this system is highlighted, and the results from this method are compared with visual rankings. To address the issue of a universal measurement, tests on scanner calibration were performed. This method and the merits of calibration are discussed. The technique was also tested on prints on both paper and plastic film substrates. Finally, a modification of the technique was used to measure the mottle on halftone prints, and the results are compared with those from the Tobias[4] mottle tester.

Mottle Measurement

At any given illumination, research has suggested that humans can discern about 60 to 150 levels of gray.[5] Therefore, a scanned image with 256 gray levels should be sufficient. The lightness channel from a 24-bit CIELab image was chosen as the measured parameter first because it provides the information to which our eyes are most sensitive—i.e., to changes of neutral or grays,[5] and second because it provides the 8-bit resolution (or 256 discernable levels) needed. Most of today's scanners and scanning software scan at 8 bits per channel or higher, and resample down to 8 bits per channel if necessary. Other scanning parameters are recommended to ensure that a minimal number of artifacts are introduced by scanners and their software.

(1) Use the optical resolution of the scanner. A higher resolution does not provide more information. It simply interpolates to a higher resolution, requiring more memory. Scanning at a lower resolution than the optical resolution throws out information.

(2) Use the default color, tone correction, and lighting settings of the scanner software. If there is no default preset, make sure there are no color filters or sharpening. If a tone correction must be applied, use the liner correction. Using the default settings allows multiple users to scan using the same settings. A color management system should be able to overcome the variations in the set-

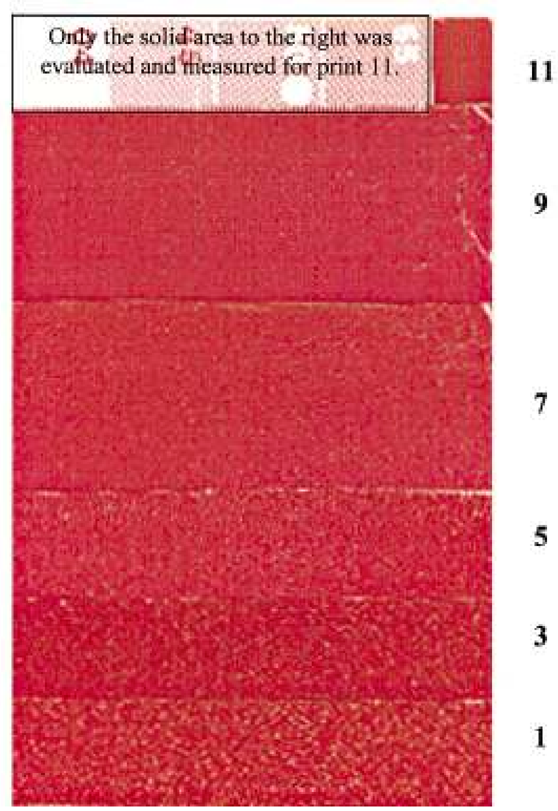


Figure 1—Heidelberg print standards for mottle.

*4600 Arrowhead Drive, Ann Arbor, MI 48105.

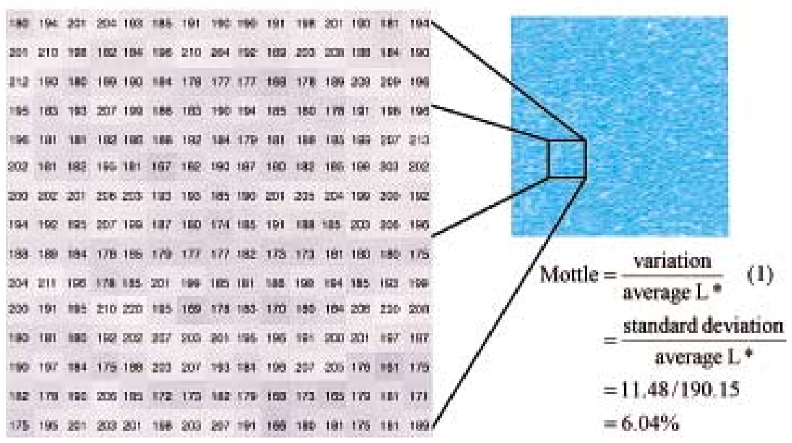


Figure 2—The lightness component of the cyan print is shown magnified to show individual pixel lightness values.

tings as long as a scan of the calibration target is scanned under the same settings as your samples.

(3) Only one scan of the target is necessary per scanning session. The scanning condition does not change significantly, and multiple calibrations may actually introduce more error.

Mottle, using the scanner method, is defined as the lightness "variation" divided by the average lightness [see equation (1) in *Figure* 2]. The average lightness is calculated from the lightness values of each individual pixel. The variation is taken as the standard deviation in lightness values for all the pixels in the image. The lightness scale, 0 = black, 255 = white, was chosen because the paper is generally white. For this reason, lighter values of the same color will have less of a contrast between the ink and paper. *Figure* 2 explains the details of the measurement.

EXPERIMENTAL

Calibration Tests for Universal Measurements

In order to obtain universal measurements, a calibration step is necessary to ensure that different scanners are detecting the same color and lightness values. This study shows that mottle measurements are more consistent when calibration steps are used.

Sample Selection: In order to cover mottle values under a full range of color and densities, a set of 12 gravure acrylic prints was used to evaluate the effects of scanner calibration on mottle measurement. This set of prints included the four process colors—cyan, magenta, yellow, and black (CMYK)—printed at three different densities (light = 1, medium = 2, and dark = 3). *Figure* 3 shows an example of the color and density ranges measured.

Image Analysis Software: The software used for this study was NIH image. Measurements of images can be done in batches, which can be saved to a data file for calculation and analysis. Adobe Photoshop was also used for image cropping and converting color space.

Color Management System: In order to improve the consistency of the measurements between scanners with different specifications and ages, a color calibration system was used. The calibration software for the scanners used in this study is ColorSynergy. This compares a scan of a Kodak Color Target (IT8) against the reference levels for each color square in the target. A transformation profile is made based on the difference in the scanned color and the reference values that can be used for batch processing of images.

Scanners Used in the Evaluation: Five scanners were used to simulate the variety of scanners that different locations might use for mottle evaluation. The selection of scanners provides the worst case scenario by including scanners with different optical resolutions (400-600 pixels per inch, ppi), and color bit depths (24–36 bits). To meet the 256 grayscale level requirement, 8 bits ($2^8 = 256$) are need in the lightness channel of a calibrated CIELab image. This requires the calibrated image to contain 24 bits. Therefore, the original image must also have 24 bits. This is why the minimum number of bits chosen for the scanners is 24. The details of the scanners are given in *Table* 1.

Scanning: A mask was placed over each print set with square openings slightly larger than 12.7 mm (1/2 in.) over each color and density (12 openings on the mask). This ensures that the same area is scanned each time on all scanners, reducing any variability effects due to sampling.

Each scanning software was set to the default settings (full density range, no sharpening, etc.). In addition, a resolution of 500 ppi was used (The Epson Scan software did not have a 500 ppi setting. Therefore, the closest resolution, 480 ppi, was used.) This was chosen to capture the detail of the print mottle, yet keep the TIFF file sizes at a reasonable size, 228 KB per color square.

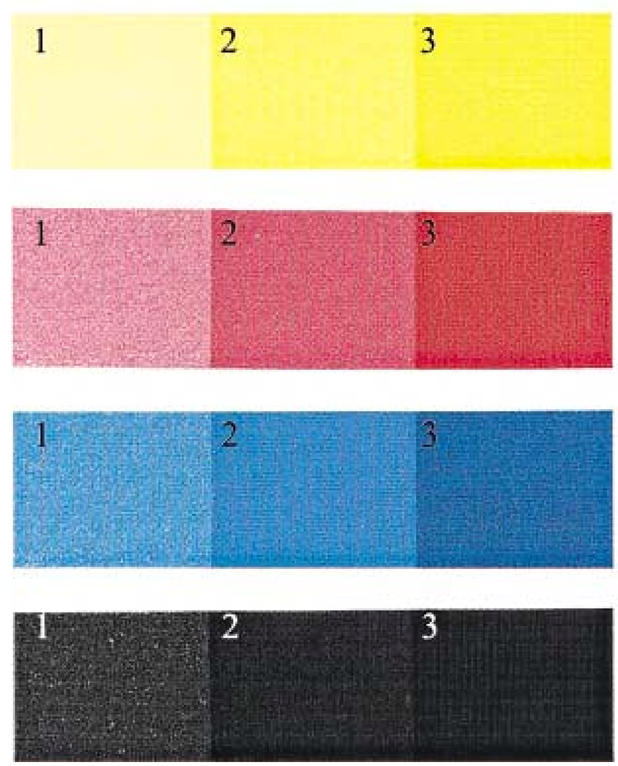


Figure 3—Patches from the gravure acrylic prints.

Table 1—Summary of Scanner Specifications

Manufacturer and Scanner	Optical Resolution (ppi)	Color Bit Depth	Software
AGFA StudioScan II	400	30	FotoLook
Hewlett Packard ScanJet 3c	400	24	DeskScan
Microtek[a] ScanMaker III	600	30	Scan Wizard
Microtek[a] ScanMaker III	600	30	Scan Wizard
Epson Expression 636	600	36	Epson Scan!

(a)Two Microtek scanners were used at different locations, The University of Michigan and the Flint Ink Research Center.

The calibration target and each set of prints were scanned, each in a single pass. This process was repeated five times for each scanner. Each scan was saved as a RGB (red, green, blue) image using the TIFF format. In Photoshop, 250 by 250 ppi squares were cut from each color square on the prints and stored as individual images (using the same RGB-TIFF format). Data from the five scans was used to evaluate the consistency of a single scanner, in addition to providing data for average values used to compare the different scanners.

Calibration: The scan of the target (IT8) was imported into ColorSynergy. The image was compared with the reference file, and a profile was created. This profile was used to make a transformation for the ColorCircuit application (packaged with ColorSynergy), for batch processing of images. The ColorCircuit application transformed the RGB images into device independent L*a*b* images. These new L*a*b* images were stored as separate TIFF files.

Measurement: The images were imported into NIH image. The average lightness (L* component in L*a*b*) was measured for each image. In addition, the standard deviation in L* of all the pixels in the image was measured. The mottle value, expressed as a percentage, was calculated from the variation and the average value as in equation (1) (see *Figure* 2).

To maintain the lightness measurement for the images that were not transformed, the RGB images were first converted to L*a*b* images in Photoshop. Note that these images are still device dependent because they have not been transformed with a calibration profile.

Two transformation methods were tested using the five scans of each sample and five scans of the (IT8) color target. The first method transformed each scan using its corresponding transformation profile. The second method transformed each scan using only the third transformation profile. This is to determine if a single calibration is necessary per scanning session.

RESULTS

Scanner-to-Scanner Consistency

The calibration techniques improve the scanner-to-scanner consistency in mottle measurement (as measured by the standard deviation in measurements) by improving the consistency in lightness and "variation" measurements as shown in *Figures* 4 and 5. The standard deviations in lightness measurements for the acrylic prints ranged from 9.0 to 32.2 levels of lightness (out of 256 levels) for the uncalibrated scans. Once the scans were transformed, the standard deviations ranged from 1.6 to 8.1 levels of lightness. The standard deviations in the variation measurements ranged from 1.2 to 4.3 for the untransformed, acrylic samples. Once transformed, the standard deviation in variation ranges from 0.8 to 4.0 levels. These calibration improvements change the standard deviation ranges in mottle from 0.5-9.7% to 0.3-3.4%. *Figures* 6 and 7 show the improved consistency in mottle measurements between scanners before and after the calibration transformation. *Table* 2 summarizes these results and *Figure* 8 visually shows the effects of calibration on a cyan print.

The effects of calibration were most dramatic in the black samples. The standard deviations in mottle measurements for the black prints are reduced to about a quarter of the uncalibrated values. The standard deviations in mottle of the cyan prints were reduced to about half the uncalibrated values. The standard deviations in mottle measured for the magenta and yellow prints were reduced to about three-quarters of the uncalibrated values.

Although the consistency of mottle measurement for yellow is improved, one needs to be aware of the mottle values, to determine whether a comparison is valid. This is because the mottle values for yellow are lower than the error associated with multiple scanners. If the mottle value is less than about one percent in yellow, the measurement may be valid only when compared with measurements from the same scanner. This is not considered to be a problem because yellows are rarely measured for mottle. For black, cyan, and magenta, minimal values are 3, 2, and 1.5%, respectively. At these low values, mottle is typically not visually apparent.

Effect of the Calibration on Consistency of a Single Scanner

When the images were transformed using all five profiles (as in transforma-

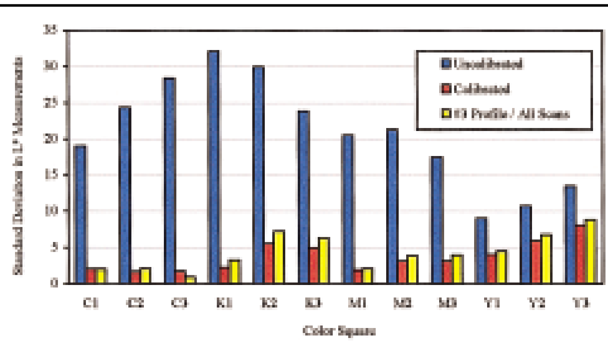


Figure 4—Standard deviations in L measurements for the acrylic prints.*

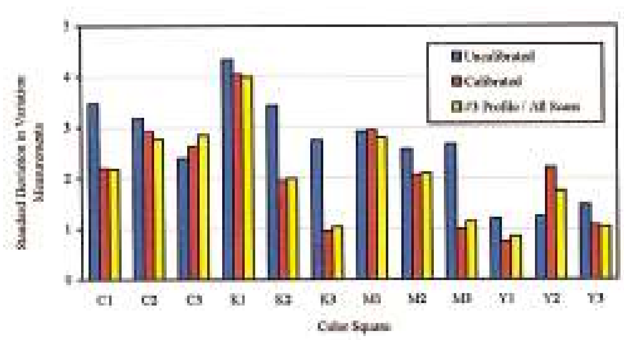


Figure 5—Standard deviations in variation measurements for the acrylic prints.

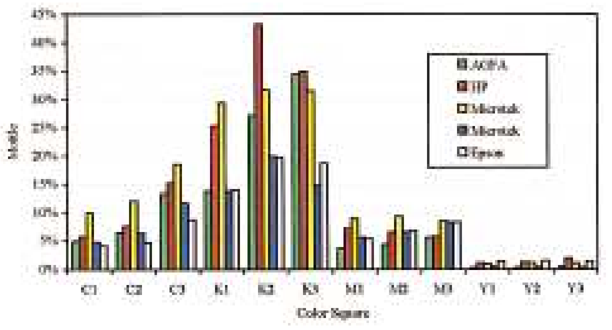


Figure 6—Comparison of mottle values for the images from five uncalibrated scanners.

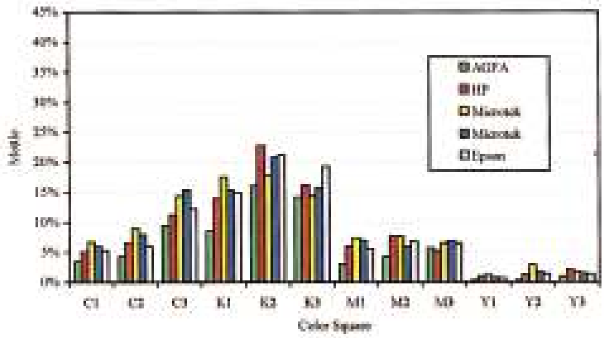


Figure 7—Comparison of mottle values for the images from five calibrated scanners.

tion method 1), 27 out of 60 L* measurements improved in consistency. However, only three of the measurements that had more inconsistencies introduced an error of more than three lightness levels. This is much smaller that the minimum error introduced by using different scanners without calibration (as indicated in *Table* 2, nine levels of lightness). In reference to the variation measurements, only about a quarter of the measurements improved. However, only three of the 60 measurements resulted in inconsistencies above one lightness value. Results from the transformation method, three were similar. This is likely because all profiles were used for this method also.

When only one calibration profile was used (as in transformation method 2), 42 out of 60 L* measurements improved in consistency. In addition, only two of the remaining 18 measurements had an increase in inconsistency of more than one lightness level. Consistency in "variation" measurements improved in half the number of measurements. Only three of the remaining measurements had inconsistencies above 0.5 lightness level. Again, an important factor is that if there is any greater inconsistency, it typically changes by less than 0.5 lightness levels. This is small compared to inconsistencies introduced by uncalibrated multiple scanners (up to 4.3 levels).

DISCUSSION

Heidelberg Standards Measured

Using the scanning method, we have seen a promising correlation with visual assessment of the Heidelberg[6] series of magenta standard prints (refer to *Figure* 1). These samples were tested, applying the scanner and calibration method (described previously in the experimental section), to verify the correlation between digital analysis and visual assessment. The magenta prints ranked from 1 to 11 (odd numbers only) in the order from highly mottled to no mottle. A Microtek scanner was used to scan the samples. The data from the standard shows a linear correlation with the visual ranking (see *Figure* 9).

Cryovac Poultry Films Measured

Measurements of print mottle on poultry films[7] provided by Cryovac show the mottle testing procedure's versatility on the type of substrate that can be tested. Using the cling properties of the wrap, the print was smoothed onto the scanner glass using a soft cloth. This removed any bubbles that may be trapped under the wrap and smoothed out any creases, both of which can cause artifacts in the mottle measurement. These samples were scanned at 600 ppi using the Epson Expression 636. The lid of the scanner provided a reflective white background for the cling wrap.

The two samples with observable mottling have significantly higher measured mottle values compared to the control sample. The control sample, which has an acceptable level of mottle, has a measured mottle value of 4.5%. The unacceptable "return" sample has a mottle of 18.4%, about four times the control value. The sample labeled 200 has a mottle of 27.2%, a value six times that of the control sample. This sample con-

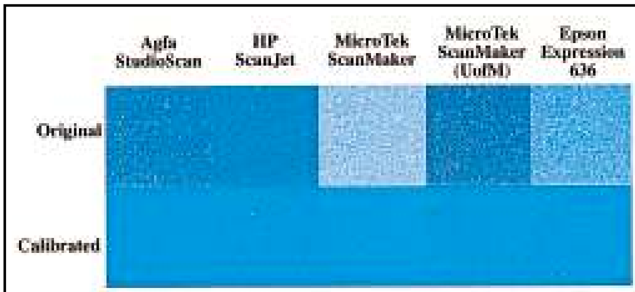


Figure 8—Comparison of five scans by different scanners of a cyan print. The top row shows uncalibrated images; the bottom row shows the calibrated images.

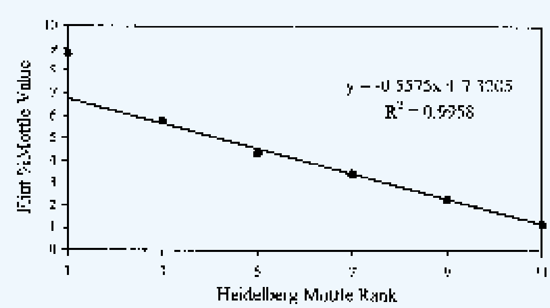


Figure 9—Comparison of mottle values measured digitally (Flint Ink method) and visually (Heidelberg Ranking). Prints with a Heidelberg rank of one would be extremely mottled, and not typically measured for print quality. Therefore, it is not included in the trending analysis.

tained 200% of an additive that was intended to induce mottling.

Rohm and Haas Halftone Prints

Fourteen prints of halftone cyan were tested using the scanning method and a modified image analysis. At a normal viewing distance, the eye does not detect the halftone pattern. Instead, the dots are blended to see a solid area. However, the scanner can detect these differences because the optical resolution is typically higher than that of the screen. Using a modified method for screened prints, the obstacle of screen artifacts can be overcome. This modified analysis attempts to simulate what the eye sees when it makes a measurement by "blending" the screen into a solid color. Specifically, the image is broken into "tiles." The length scale of the tile was chosen based on a definition of mottle presented in a tutorial[7] by J.R. Edinger. The definition describes mottle as an "aperiodic fluctuation of density at frequencies less than one cycle per millimeter in all directions." Therefore, the image was broken into tiles at intervals of every 1 mm. Hence, variations at length scales less than 1 mm (as is the case for a screen) will be ignored by the use of averaging. Variations at length scales greater than 1 mm will be measured using a standard deviation. An AGFA StudioScan II scanner was used to scan the samples. Each sample was scanned at 400 dpi, the optical resolution of the scanner, and the images were calibrated through ColorSynergy. The analyses of the images were performed in NIH image and Excel. The following provides a more detailed description of the method used for these samples:

(1) A 50.8 mm × 50.8 mm (2 in. × 2 in.) area from each of the samples was scanned at the optical resolution of the scanner, 400 ppi.

(2) Transform the image using the calibration file for the scanner.

(3) Break the image into tiles at 1 mm intervals (approximately every 15th pixel for a 400 ppi image).

(4) To remove the screen effect, the average lightness, L*, of each tile is measured.

(5) Next, the standard deviation in the average L* is calculated for each tile and its nearest neighbors (groups of nine tiles) as shown in *Figure* 10.

(6) The "nearest-neighbor" standard deviations are averaged to provide a typical variation.

(7) This variation is divided by the average L* to provide the final mottle value.

Figure 11 shows the results of the modified Flint Ink scanner method and the Tobias mottle tester. Note the similarities in the ranking of the prints, labeled A through N.

Table 2—Comparison of the Uncalibrated and Calibrated Consistencies[a]

Acrylic Samples	Uncalibrated	Calibrated
Standard deviation range in lightness	9.0 - 32.2	1.6 - 8.1
Standard deviation range in variation	1.2 - 4.3	0.8 - 4.0
Standard deviation range in mottle	0.5% - 9.7%	0.3% - 3.4%

(a) Averaged data from the five scanners was used to calculate the standard deviation of the mottle parameters. The ranges are broad because they include all four colors and three levels of print densities.

CONCLUSIONS

The digital image analyses method is capable of objective and reproducible measurements of print mottle. The results of the scanner calibration test show improved consistency in the mottle values between different scanners. This would allow many locations to have their own analysis station without compromising a "universal" measurement. The mottle measurements of the Heidelberg standards correspond to the visual ranking. The robustness of the scanner method was shown by tests of prints on different substrates. Finally, a modified version of the analysis shows that halftone prints can also be tested. The method described here shows good agreement with the Tobias mottle tester.

List of Definitions and Terms

Variation: The standard deviation of pixel values in an image. This is used to calculate mottle values.

L*: The lightness component in the L*a*b* color space.

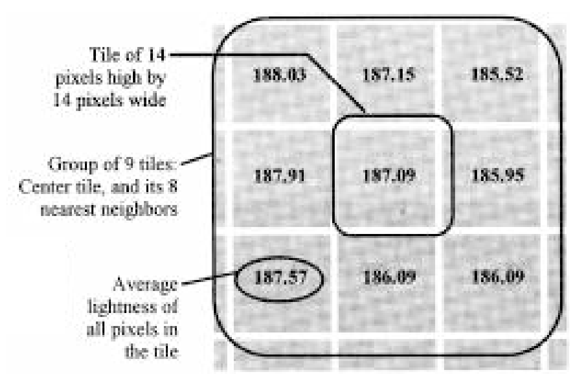
Standard Deviation: The standard deviation is used in two ways:

(1) To quantify the range of lightness values of a scanned print (see Variation, above)

(2) To measure the error associated with a measurement due to inconsistencies between scans and/or scanners.

ACKNOWLEDGMENTS

We thank James Watson, of Rohm and Haas, for providing samples of the half-



Example: Standard deviation of tile group = 0.93

Figure 10—Magnified view of the tiles cut from the image of a screened print.

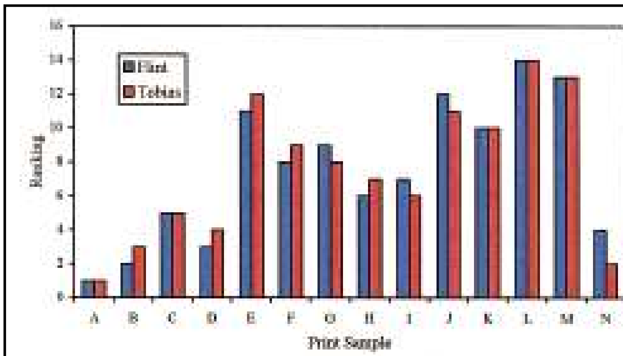


Figure 11—Comparison of digital analyses performed by Flint Ink and Tobias on Rohm and Haas prints A-N.

tone prints and their Tobias mottle values, and Cryovac for providing poultry films for the mottle measurement of prints on clear films.

We acknowledge useful discussions with Jim Vrotacoe, of Heidelberg Web Press, Dr. Bruce Blom, of Mead Paper, and James Watson, of Rohm and Haas.

The authors wish to thank Gary McBee, of Flint Ink, for doing some of the initial work in this project, and Dan Lawrence, also of Flint Ink, for discussions on instrumentation.

References

(1) Armel, D. and Wise, J., "An Analytical Method for Quantifying Mottle," *Flexo* (December 1998)

(2) Rosenberger, R.R., *Prog. Paper Recycling*, 4, No. 3, 39 (1995).

(3) Kowalczyk, G.E. and Trksak, R.E., "Image Analysis of Ink-Jet Quality for Multi-Use Office Paper," *Tappi J.*, 81, No. 10, 181 (1998).

(4) Tobias, P.E. et al., *Tappi J.*, 72, No. 5, (May 1989).

(5) Green, P., *Understanding Digital Color*, Graphic Arts Technical Foundation, 1995.

(6) Vrotacoe, J. from Heidelberg Web Press, has performed similar studies on the scanning method, confirming our results.

(7) Edinger, J.R., "Fundamentals of Black and White Image Quality," IS&T's NIP13: International Conference on Digital Printing Technologies, 1997.

(8) Portions of this paper have been published in the IS&T NIP14 Conference Proceedings. ©1998, IS&T-The Society for Imaging Science and Technology.

(9) Portions of this paper have been presented at the NPIRI Technical Conference, October 1998.