The Image Analyzer - A True Dot Area Meter?

David J. Romano*

Keywords: Video, Measurement, Halftone, Images

Abstract: Measuring halftone dot areas on aluminum printing plates with an video image analyzer is a simple procedure but tends to be rather subjective. The major variables in such a system are image capture, aperture selection and thresholding. Some of the theoretical and practical considerations surrounding these variables are discussed and some proposals are made regarding procedure. It is shown that a very simple method of measurement can be both accurate and fairly insensitive to user subjectivity.

Introduction

Halftone dot area measurements of printing plates have traditionally been done in laboratories using an instrument called a planimeter. It is is the same as an image analyzer, whose basic components include:

- CCD or other video camera
- Microscope
- Frame grabber (image digitizer)
- Image analysis software

---

*Bayer Corporation, Agfa Division

Image analysis systems have never gained much popularity with printers. The main reason is that there is not any need for printers to measure the halftones on conventional (analog) plates. With the platemaking process standardized, dot gain is characterized by measuring the film and the press sheet. Computer To Plate (CTP) technology does not use film though, and the printer has no dot value information except for press sheets. The plate has essentially replaced the film and there is suddenly a need to measure it. This has led to a recent interest over measuring plates with densitometers, sometimes using the Yule-Nielsen n-factor in order to "correct" the measurements. However, as this paper will remind us, there are no absolute answers, only consistent ones.

The original purpose of the n-factor was to account for light scattering and optical dot gain on translucent surfaces.[1] Since a plate's surface is opaque, the n-factor loses its meaning. It becomes a fudge factor. In itself, this is fine, but the measurements themselves are meaningless unless they can be repeated. Repeatability is the real issue. Densitometers are normally repeatable to ±.01 density with an inter-instrument agreement of ±.02. Dot area measurements are based on 3 density measurements: the paper, the solid and the tint; which compounds the error of the three. This is not a problem when measuring prints, but the density range on a plate is insufficient to allow for this degree of variation. Silver halide CTP plates have black images on gray backgrounds, and the density range is typically about .70 D. Often, there is a variation in the solid density in different areas of the plate, and it may mottled or un-uniform. Photopolymer CTP plates achieve slightly more contrast with their colored emulsions, about .80 D, but it depends on how well the reflection spectrum of the emulsion complements the spectral characteristics of the densitometer filters. In some cases of *analog* plates, this can result in a contrast of up to 2.0 D, making the measurements fairly reliable. By calibrating the densitometer, as described in the next section, it can be made more accurate, but it cannot be made more repeatable. In this paper we will see that the image analyzer is capable of delivering accurate as well as repeatable measurements. This paper is not a formal comparison between image analysis and densitometry.

Calibration and threshold

If we are to measure the relative areas of halftone dots, whether they be on plate, film or print one important question must be asked: What qualifies as a dot? At what point does the dot start and the background end? Somewhere in the processing of raw data an assumption must be made as to what a dot is. It doesn't matter at which point we make it, as long as it is acceptable to people, most of the time and is theoretically sound. Even the well known methods of conventional reflection and transmission dot area measurements are based on some assumptions, whether the user is aware of them or not. For example, the Murray-Davies equation used in reflection readings, assumes that the Reflectance (R) of the ink is 0% and the R of the paper is 100%. It also assumes that the amount of ink in the solid tone indicates the amount on the individual dots. We know that these assumptions are not completely true, yet the method provides fairly precise measurements and we use it anyway. A transmission densitometer assumes that the Dmin, even between the dots, that the dots have sharp edges and both transmit and reflect 0% of the light reaching them. Once again, these assumptions are not completely true, but in most cases, it doesn't matter. In fact, both types of "dot area meters" are not measuring dot area at all, they are only inferring it from an integral density measurement. In this sense, the image analyzer *is* a true dot area meter because it does not need to measure relative to a *black* and a *white*. This does not make the image analyzer any more correct, but it does make it more versatile. Instead of measuring density, it uses an image histogram and a thresholding to disciminate between the image and non-image areas. From there, dot areas, line widths, perimeters and a variety of other image features can be measured directly. Different applications may demand different strategies for thresholding. A histogram of a 50% tint is found below in figure 1. It has a bimodal distribution of gray levels, with the lower levels representing the image areas and the higher levels representing the surrounding plate areas. The threshold will define how dark a pixel in the image has to be for it to be counted as part of a dot. All that is needed is to have a method for determining this threshold.

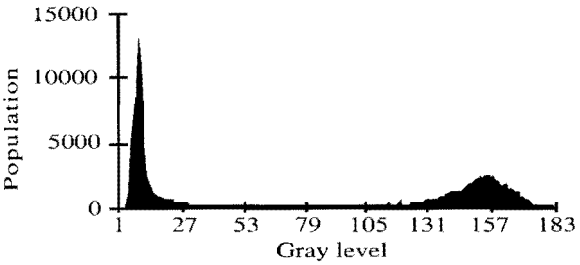

Figure 1 Image histogram with two symmetrical peaks, at gray levels 7 and 155

One logical solution to this is to choose a threshold representing 0.5 R, or at "half-contrast". This threshold correlates well with transmission dot area measurements.[2] This point would be found by averaging the mean level of the dots with the mean level of the background. As long as the two histogram modes are reasonably symmetrical, these values are represented by their peaks. A good way to check this is to measure the edge profile of a dot and check the levels of the Dmin and Dmax areas, as in figure 2.

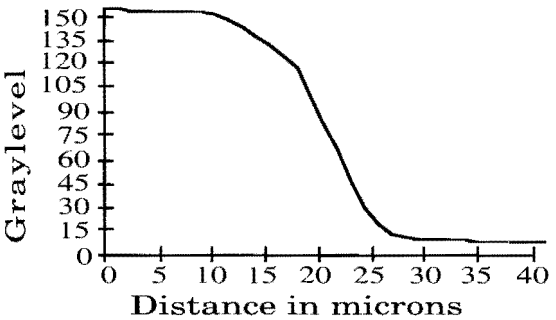

Figure 2 Edge profile of a dot showing Dmin and Dmax levels equal the histogram's peaks

Another way of determining threshold is to calibrate the image analyzer to a target with known dots areas. By making an exposure to a plate which reproduces positive and negative microlines to the same line width, it can be assumed that the halftone dots will reproduce without gain or loss.[3] At this exposure, the dot area read by a transmission dot area meter will be the same on the plate. A good film to use is an UGRA Wedge because a dot for dot exposure is easier to achieve when the dots are very hard dot and the density between the dots is very low and consistent. The measurement of the film should be done on a transmission densitometer which reads to 0.1% area. Densitometers which read to 1% are repeatable to ±1% and ones that read to 0.1% can be off by ±0.4%.[4] This effectively creates a calibration target with "known" dot areas, and simply devising a thresholding procedure which makes the measured values match the known values, calibrates the image analyzer. This method could also be used to set the n-factor on the densitometer. Remember that this is only addressing the issue of accuracy, not repeatability, which is the densitometer's downfall. But how accurate is this calibration target? If the film's dot areas can first be read to a precision of only ±0.4% and a then contact exposure is made, adding its own error, it is easy to have an uncertainty of nearly ±1%, even without considering a possibly uneven Dmin on the plate. This paper makes some simple suggestions for threshold selection, but the best method must be determined by the user.

The Image

Before making a measurement, the system must be set up properly to obtain a high quality image. The output of the camera should be linear, and all automatic functions should be switched off. Start by plotting a histogram of the dark frame, (a picture with the lens covered or the shutter closed) and noting its gray level. No part of an image can be darker than this level, and all image histograms will be clipped there, regardless. Next, place a solid tone area under the microscope and adjust the incident light intensity until the

image's gray level is as close to the dark level as possible without being clipped. Check this by plotting an image histogram. Once the black point has been found, place the lightest halftone tint under the microscope. Use the camera's gain function, in balance with further adjustment of the incident light intensity, until the best image quality is obtained and there is no clipping of gray levels at either end of the scale. This means no pixels should be level 255 or equal to the dark level.

Image sharpness and contrast are essential to making precise measurements. The edges of the halftone dots should be well defined and stand out strongly from the background. One way to achieve this is to illuminate with the color of light which complements the hue of the plate's emulsion. Color cameras have the advantage of offering red, green and blue channels, much like a densitometer, and one of these can be chosen instead. Unless it is green though, the same color filter must also be used in the light path. Otherwise, it will be impossible to perform any critical focusing due to chromatic aberrations in the objective.

Image quality should be judged using two criteria: (1) The distance between the histogram's peaks (contrast) and (2), the depth of the histogram valley (sharpness) at the 50% tint. In fact, the depth of the valley is the only important characteristic of the histogram, but it is generally improved by increasing image contrast. The reason it is checked at the 50% dot is, interestingly, the same reason we measure dot gain at this point. It has the most perimeter, and therefore, the most unsharpness.

Even illumination is another very important factor in image quality. If the illumination is not even across the frame, the respective areas of the dots will be inherently different. Kohler illumination is the best method for correcting this.[5] The illumination should then also be corrected mathematically using the equation:

$$R = \left[\frac{s - d}{r - d}\right] \times \overline{r - d} \qquad (1)$$

With,

R: result

s: sample

r: reference

d: dark level

Reference image *r* is an image which is illuminated exactly the same as sample *s*, yet without image structure. It is merely a record of the illumination pattern, and therefore can be divided out. The reference image must contain as little visible structure as possible, because when the division is carried out, the noise in the referenece will add to the noise in the resulting image. The best way of accomplishing this is to use the bare plate surface and fit a high degree polynomial surface to the image. This will provide the smoothest surface possible while retaining the form of the unven illumination pattern.

Another important consideration is sample noise. Since halftone dots differ slightly in size from one another, it is important to include as many as possible in each measurement. A densitometer might include 100-200 dots in a single measurement, averaging out many types of errors. Generally, the precision of a measurement made with an image analyzer is dependent on the amount of edge area within the field of view. When the threshold is being chosen, a deviation of 1 gray level will not affect the measurement of a sharp image as much as as an unsharp image. Also, what is the distribution of gray levels within the dots and the background areas like? Are the peaks symmetrical? Usually, the darker gray levels of the image areas have a fairly normal distribution with a lower standard deviation than the gray levels of the background areas. Refer to figure 1.

The Aperture

The aperture size, or field of view, is critical to making accurate measurements and deserves special discussion. Since the field of view seen with a microscope is generally small, possibly enclosing only a handful of dots, a great deal of error can result when the aperture is randomly placed. Apertures in standard transmission and reflection dot area meters use a constant aperture size for all screen rulings, a round one usually 2 or 3 millimeters in diameter. By randomly placing an aperture of arbitrary size over halftones of varying pitch, the inclusion of partial dots or just missing dots leads to measurement error.[6] Franz Sigg of RIT has calculated, that the greatest amount of error occurs when the center of the aperture lies on the center of a dot or the center of a hole, and the amount of error is inversely

proportional to the product of screen frequency and aperture size.[7] Conversely, zero error occurs when the center is placed on the edge of a dot and hole.[8] (More accurately, on the edge of a *cycle*. This is most easily visualized with a checkerboard.) However, this is assuming that the aperture is not designed to sample full unit areas, which is possible to do with image analysis. Not only can the aperture dimensions be accurately adjusted, but also its placement. A unit area is defined as the area of a dot and a hole together. In figure 3, the center of the aperture is placed on the center of a dot, but because it has been adjusted to enclose exactly two unit areas, the error is still zero. So, this arrangement provides complete lateral symmetry. The aperture can be placed anywhere in the frame.

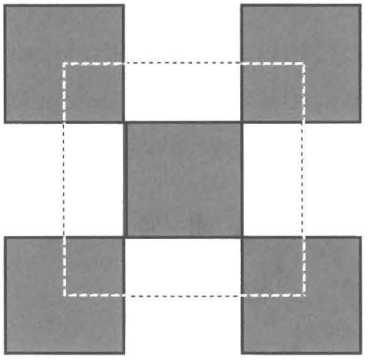

Figure 3 Screen angle 45°

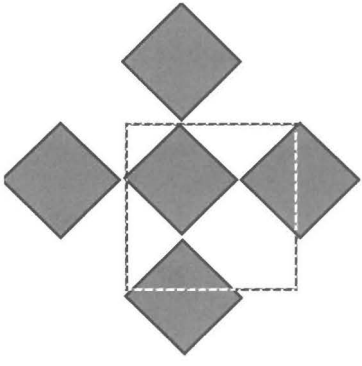

Figure 4 Screen angle 90°

In figure 4, the same aperture is used, but it is no longer aligned with the screen angle. (Note that figure 4 has been reduced in size to fit on the page) The result of this is that it now encloses more than 2 unit areas, but not quite 3 either. It is only because its center has been placed on the edge of a dot and a hole, does this arrangement result in zero error. The arrangement in figure 4, then, provides rotational symmetry and some lateral symmetry. Provided the aperture is placed in such a way, and has a symmetrical shape, it can be of any size and the error will be zero. Either arrangement can be used, but it is usually easier to define the size of the aperture rather than the location of its center. Note that the opposite would be true if the aperture size had been adjusted for the 90° screen instead as in figure 5 and figure 6.

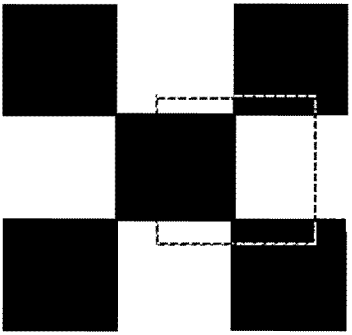

Figure 5

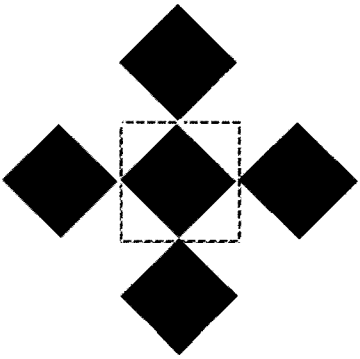

Figure 6

Densitometers use circular apertures for two reasons. First, these apertures are large enough to drown out this error. Second, since it cannot be guaranteed that the aperture is aligned with the screen correctly, the circle will average out this error while the square would have exaggerated it.

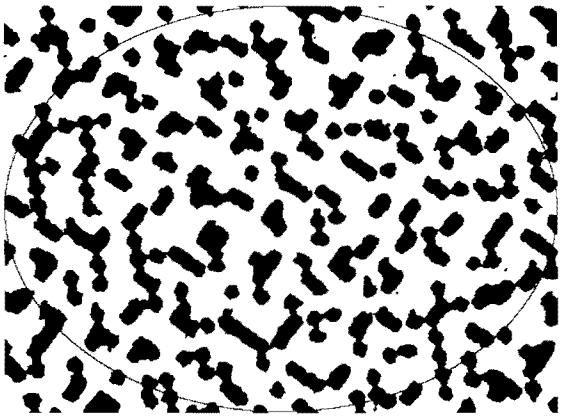

Fig. 7 FM halftone with oval aperture to round out errors

With FM halftones, there are some additional considerations when drawing and placing the aperture. Since the dots are not all the same shape and are not evenly spaced, the best aperture shape to use is an oval rather than a square. See figure 7. Square edges are more likely to interfere with the grid of the addressability than round edges. Also, the errors caused by aperture size and placement will almost always be different in the x and y directions of the oval. Averaging them will then result in less error. A circular or a square aperture

would have the same amount of error in both directions.[9] If it were possible to generate an aperture with a truly random shape, then the error would theoretically be zero.[10] A star shape might be a good approximation of this. In an FM halftone, a unit area is not easily defined. One would have to know exactly where the pseudo-random pattern begins to repeat itself, and it is unlikely that this area is symmetrical in shape or easily drawn and placed correctly. It is also likely to be different for every dot area. Though it is difficult to adjust the size of the aperture correctly, the sheer number of unit areas sampled will be great enough to reduce errors to a negligible degree.

Flare

Flare is caused by a number of factors. For the purposes of this paper, we will use this term for any phenomenon which reduces edge sharpness and image contrast. Flare can be greatly reduced by limiting the illumination to the field of view. This can be done mechanically with a black mask, or optically with the microscope's field aperture. Both will decrease the amount of stray light striking the objective and increase overall contrast; but will not have much effect on the sharpness of the dot edges. Just as light outside the microscope's field of view scatters in all directions, the light within the image is also scattered by the rough surface of the plate. Optical factors that influence edge sharpness include the graininess of the plate surface, the numerical aperture of the microscope, the wavelength of the length used and the quality of the optics. Microscopes with zoom lenses have complex optics and will increase light scatter. Though offering convenience, they cannot provide the image quality of a prime lens. Even if the image is perfect when it reaches the CCD, or other photoreceptor, light can bounce off of the chip and the surroundings before striking the chip again, in a different place, causing some unsharpness. Current CCD technology requires that a fair percentage of the chip's surface is covered by the "irrigation ditches" which carry the electrical charges to the amplifier. Light can reflect off of these surfaces as

well as off of the photoreceptors themselves. In addition, an effect known as *blooming* occurs when electrons spread from one pixel site to another.[11]

Further edge degradation can also be caused by the digitizer board and from the electronics of the computer itself. Regardless of the number of the photoreceptors in the camera, the vertical and horizontal spatial resolution of the image is limited by the format of the frame grabber.[12] In an RS170 system, the vertical sampling is limited to about 480 lines, and to about 580 in the CCIR system. For example, the camera may use an array of 786x572 pixel sites, but this signal will be re-sampled by the digitizer board as 640x480 pixels.

In reflection microscopy, flare commonly causes dots of different size to also differ in brightness. A 10% dot is surrounded by a sea of light. If it is assumed that the amount of flare, is constant, a 10% dot will be more affected than a 50% dot, even though it has the same density. With increasing coverage, both image area and non image areas will appear darker. The less reflective image areas deprive the non image areas of the scattered light they would have received otherwise, and the image areas appear darker for the same reason.

Precision

Image analyzers are not quite like densitometers. Densitometers ask little of the user, and are not as complicated as image analyzers. How do we control the variables? Better yet, how much control is possible and which are the most important variables? One could argue that image sharpness is the most important variable. If an image histogram were ideal, (and the image were binary), there would only be 2 brightness levels; one for the dots and one for the background. Assume for now that of 100 pixels sampled, 50 are completely black and 50 are completely white. The tint would always measure 50%, regardless of where the threshold was set. In a real image, there is a distribution of gray levels between these two modes, called the valley. If the valley is low, a small change in the threshold will not change the dot area by

much. For example, with an aperture size of 160000 pixels, the threshold happens to be reduced by 1 gray level and changes 320 pixels from white to black (pixel error = 320). This would decrease the dot area by only 0.2%. That is a high degree of precision, more than is normally needed. If the screen ruling of this image were doubled, and the aperture remained the same, the pixel error would theoretically increase to 620 and the dot area would increase by 0.4%. Although there is not a formula which defines how much precision can be obtained from a certain screen ruling, consider another argument about precision.

Dot area 60 l/cm (152 lpi)	*Change in % for 1µm diameter change*
1% (19µm) round	.109%
5% (42µm)	.241%
10% (60µm)	.339%
40% (105µm) square	.763%
50% (118µm) square	.852%

Figure 8

For an equal change in diameter, larger dots will produce more error than small dots, because they have a longer perimeter. As figure 8 shows, for a square 50% dot a 60 l/cm (150 lpi), a change in diameter of 1µm will result in nearly 1% error, while this is not the case for the 10% dot. For an equal change in diameter, a 50% dot at 120 l/cm (300 lpi), will change twice as much as one at 60 l/cm. That means that perimeter is proportional to screen ruling. It also means that it is inversely proportional to magnification, because doubling the screen ruling has the same effect on image perimeter as zooming out by a factor of two. This ties both magnification and screen

ruling to precision. Figure 9 in the appendix calculates the change in diameter and the resulting change in dot area when the threshold is changed.

Conclusions

Image analysis vary widely in cost and complexity and can be purchased as an integrated package or be customized by the user. It is often time consuming to make measurements on them and one is limited to a sample size which will fit under the microscope. Presently, only manufacturers, research organizations and a handful of large printers need, and can afford the time and cost to run one. Possibly the drive toward filmless printing systems will create enough of a market for a simple and affordable system. The author hopes that this has been an informative overview of image analysis and dot area measurements.

Caveats

The author would also like to extend some advice to those considering an image analysis system. In the near future, the results provided by image analysis systems may be considered to be the final word, providing a reference measurement that all others are compared to and that further research is based on. Image analysis is a very versatile tool, capable of providing high accuracy and repeatability, but there is no standardization as of now. Non standardized practices such as image capture, preprocessing, aperture placement and thresholding can result in large discrepancies. It has recently been shown that densitometers do not provide repeatable or accurate measurements of lithographic plates. In fact, a dot area measurement made with an image analyzer is prone to even more error than this because of aperture selection alone. So, unless all of the information regarding procedure is provided along with the results, it may not be possible to compare them with those made by another.

References and further reading

[1] John A. C. Yule and W. J. Nielsen, *The Penetration of Light Into Paper and its Effects on Halftone Reproduction,* TAGA Proceedings 1951, p.p. 65-76.

[2] R. Van Esch, *Dot Quality in PTS* Recorders, SPSE 43rd Annual Conference Proceedings, p.p. 195-198.

[3] Franz Sigg, *A Few Things About Microlines That Most People Do Not Know*, TAGA Proceedings 1988, p.p. 428-449.

[4] Franz Sigg and David Romano, *How to Calibrate and Linearize an Imagesetter Using the Digital UGRA/FOGRA Wedge*, IS&T 4th Technical Symposium on Prepress, Proofing and Printing, 1995, p.p. 88-92.

[5] Eastman Kodak Company, *Photography Through The Microscope*, 9th Ed, 1988, p.p. 28-37.

[6,7,8,9,10] Franz Sigg, *Errors in Measuring Halftone Dot Areas*, Journal of Applied Photographic Engineering 1983, p.p. 27-32.

[11] Johansson, Yu. A. Lebendinsky, and K.G. Predko, *Experimental MTF Measurements of CCD Using an Interfereometrically Generated Test Pattern*, Journal of Imaging Science and Technology 1991, p.p. 320-325.

[12] G.W. Awcock and R. Thomas, *Applied Image Processing*, McGraw-Hill, Inc., 1996, p.p. 84-86.

H. Brent Archer, *Photometric Measurement of Dot Area* , TAGA Proceedings 1966, p.p. 35-44.

A.W. Pearson, *Instrumented Characterization of Lithographic Plate Surfaces and Halftone Images*, TAGA Proceedings 1981, p.p. 77-85.

R.S. Fisch and R.D. Cavin, *A Report on the Percent Area Measurements Taken From Plates Used in the GCA 1985 Print Properties Test*, TAGA Proceedings 1986, p.p. 267-268.

Appendix

Diameter measurements

When measuring an object with the microscope digitally, the image is broken into pixels whose size is determined by the magnification. It would seem at first that the amount of error in dot area measurements is directly related to the size of the pixels and therefore to the magnification. What is the relationship between image sharpness, pixel error and change in dot diameter? What is the repeatability of a diameter measurement? An experimental way of doing this would be to sample the a large population of dots of the same size, as in an FM tint, and measure their areas and diameters. With the experiment repeated many times, turning the light source and of and on and allowing different warm up times, refocusing and so forth (but without moving the sample), the variation of the diameter measurements and dot areas is found.

There are some additional considerations with this method. For example, the diameter can be calculated easily enough by assuming round spots and dividing by π. However, π cannot ever equal 3.1416 for a digital system because it assumes square pixels. Mathematically, π is defined as the relationship between a circle's diameter and its perimeter but for a square, this

relationship is not 1:3.1416, it is 1:4. We should not simply assume a circle or a square in order to calculate diameter. The shape is defined by its circularity, the relationship between area and perimeter. By first redefining π using the measured perimeter and area values, and replacing π with shape factor πf it can be determined what its average diameter is. See equation (2). This will provide a general idea of the amount of precision a system is capable of and suggests a logical method of measuring dot diameters.

$$d \equiv 2\sqrt{\frac{a}{\pi f}} \qquad (2)$$

where

$$\pi f \equiv \frac{p}{2\sqrt{\frac{a}{\pi}}}$$

With,

d: diameter

a: area

p: perimeter

πf: shape factor

In figure 9, the change in dot area, as a function of the change in diameter, is calculated for the range of 10%-50%. The dots begin as round in the highlights and continuously progress to a square, 50% checkerboard. Note that shape factor πf changes linearly from 3.29 at 10% to 4.0 at 50%.

1	2	3	4	5	6	7	8	9
Dot %	Aperture in pixels	# of Dots Sampled	Dot size in pixels	**Pixel Error**	Diameter Change µm	πf	**% Area Change**	**% Area Change = Pixel Error ÷ Aperture**
10%	140.625	9	1563	**51**	.10 µm	3.29	**.16%**	**.16%**
20%	140.625	9	3175	**316**	.45 µm	3.45	**.22%**	**.22%**
30%	140.625	9	4688	**410**	.46 µm	3.62	**.29%**	**.29%**
40%	140.625	9	6250	**479**	.46 µm	3.80	**.34%**	**.34%**
50%	140.625	9	7813	**500**	.42 µm	4.00	**.36%**	**.36%**

Figure 9

In the above figure, the last column calculates dot percent error by taking the number of pixels found at the threshold (the depth of the valley of the

histogram) and dividing it by the total number of pixels sampled (aperture size). This calculates the amount of error caused by varying the threshold by 1 gray level. Column 8 calculates the same thing, but does so by determining the change in diameter, (caused by a change in area), which would then lead to a difference in dot area. The data in columns 8 and 9 are the same, proving that the precision of a dot area measurement can be measured with a simple histogram analysis.