

Towards Instrumental Analysis of Perceptual Image and Print Quality

Ming-Kai Tse and David Forrest; Quality Engineering Associates (QEA), Inc.; Burlington, MA, USA

Abstract

Quantitative image quality analysis technologies have advanced significantly in the past twenty years. Practical instruments for objective analysis of engineering quality attributes are now used widely and successfully in the digital printing industry. However, in the design and optimization of an imaging system, measurement of engineering quality alone is not sufficient. It is equally if not more important to understand how image quality is perceived by the end user. Unfortunately, unlike engineering attributes, perceptual image quality is much more difficult to measure and quantify. Recognizing the technical challenges and the potential gain, we have initiated a study on the feasibility of instrumental analysis of perceptual image quality. Given the complexity of the problem, we submit that even incremental advances make a valuable contribution to the imaging community. In this paper, we discuss our methodology and our initial results on perceived noise, and present an assessment of the efficacy of our approach.

Introduction

Image quality involves many different attributes that affect viewer preference, including color accuracy, effective resolution, uniformity, line quality, gloss, etc… The subject of this paper is image uniformity and its counterpart, image noise. The objective is to devise a single objective metric that correlates with human perception of image noise.

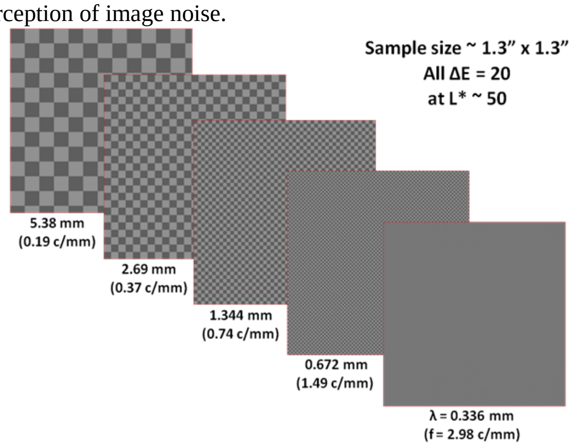

***Figure 1**. A series of images with different size scale non-uniformity.*

For a digitized image (i.e. represented as a bitmap) there are many tools available to quantify image noise [1]. The standard deviation of all pixel values provides a simple indication of image uniformity. Note that in the ideal case (i.e. a perfectly uniform image) all pixel values are identical and the standard deviation is zero. But a simple standard deviation metric does not generally correlate with human perception of noise, primarily because the standard deviation is insensitive to the spatial scale or periodicity of variations. It is well know that human contrast sensitivity is highly dependent on spatial frequency [2] and therefore very small features and high frequency noise are not perceived. Perception of noise in color images is further complicated by the tri-chromatic nature of human vision; for example, noise in the luminance direction is more objectionable than noise in the hue direction. Finally, beyond human sensitivity, human preference also plays an important role. Viewers tend to object more to large scale noise versus small scale noise (of equal magnitude) and to high contrast variations versus low contrast variations. It is therefore very difficult to create an objective metric to correlate with human perception.

Graininess and Mottle

Figure 1 shows a series of images comprised of squares with exactly two density levels, arranged in a checkerboard pattern. The same density levels are used in each image, the only difference between the images is the size of the component squares.

Note that the standard deviation computed for each image in Figure 1 is identical. Yet under normal viewing conditions the noise in the right-most image, at 2.98 cycles/mm, is not perceived. This is explained by the frequency dependence of human contrast sensitivity. Viewers also generally perceive the image with 1.49 cycles/mm to be a "grainy" image, whereas the images at lower frequencies are judged to have a different class of non-uniformity called "mottle." Finally, for images with mottle, viewers tend to object more to lower frequency variations, although this is contrary to the human contrast sensitivity function at the corresponding frequencies.

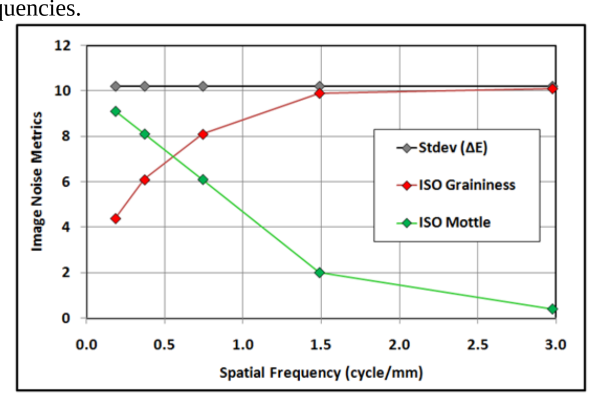

***Figure 2**. Noise metrics for the images in Figure 1.*

ISO-13660 is an international standard incorporating a wide range of print quality attributes, including the uniformity metrics graininess and mottle. Graininess is high frequency noise (> 0.4 cycles per mm) and mottle is low frequency noise (< 0.4

cycles/mm). ISO-13660 is intended for monochrome images and the units for variation computations are specified as optical density. In this study, since we are interested in color images and especially human perception of noise, we use the CIE-L*a*b* space for all computations. Figure 2 shows standard deviation, graininess, and mottle (in L*) computed for the images in Figure 1. Clearly the ISO metrics include some frequency sensitivity, and therefore are more relevant than the standard deviation which is constant for all the images. The mottle metric seems to correlate well with perception: at high frequencies the mottle is very small. But even in this simple example the graininess metric has a clear weakness: at the highest frequency, no graininess is perceived in the image but it is at a maximum according to the ISO graininess metric.

Note that graininess and mottle are closely related to the simple standard deviation statistic: according to ISO 13660, noise is classified as either graininess or mottle, but the total noise is still approximated by the standard deviation with no consideration for size scale. To agree with perception, we need to develop a noise metric that recognizes the variation at 2.98 cycles/mm in Figure 1 is negligible.

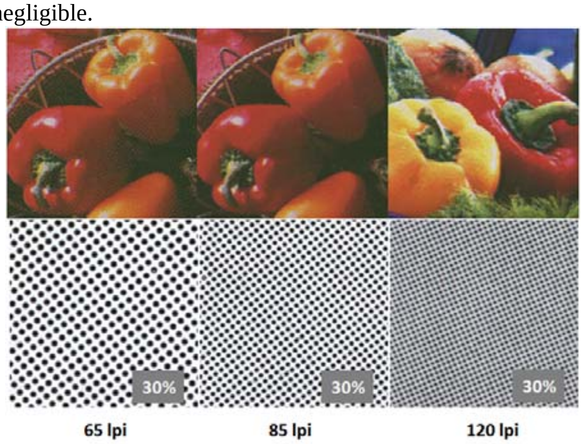

Figure 3*. Halftone images at various frequencies. Perception of non-uniformity decreases with increasing half-tone frequency.*

Spatial Filtering

Figure 3 shows three images with different halftone frequencies. The perceivable noise in the images is inversely related to the halftone frequency (higher frequency corresponds to less perceived noise).

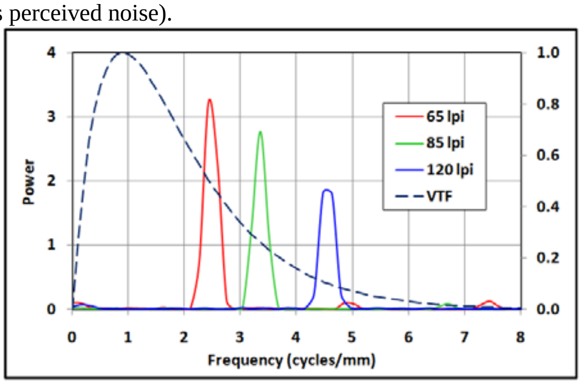

Figure 4*. The human visual transfer function (VTF) showing sensitivity as a function of frequency, and the power spectra from the halftones in Figure 3.*

Since the standard deviation metric is insensitive to variation in spatial frequency, the standard deviation for all three images in Figure 3 is similar. However, we can apply a spatial filter to the image to transform the bitmap into a better representation of what is perceived by a human viewer. The filtering is done in the frequency domain using a 2-dimensional version of the human Visual Transform Function (VTF), shown in one dimension as a dotted line in Figure 4. Figure 4 also shows the power spectra for the three halftoned images.

Note that the VTF in Figure 4 correctly indicates the low perceivability of the high frequency halftone pattern at 120 lpi. Applying the 2-dimensional VTF filter to the image before computing the standard deviation yields a noise metric in better agreement with perception, as listed in Table 1:

Halftone Line Screen	L^*_{mean} @30% Gray Patch	Visibility of Halftone	Std Dev No Filtering	Std Dev VTF Filtered
65 lpi	67.4	High	22.8	14.1
85 lpi	66.0	Medium	21.0	8.5
120 lpi	61.9	Low	18.5	1.0

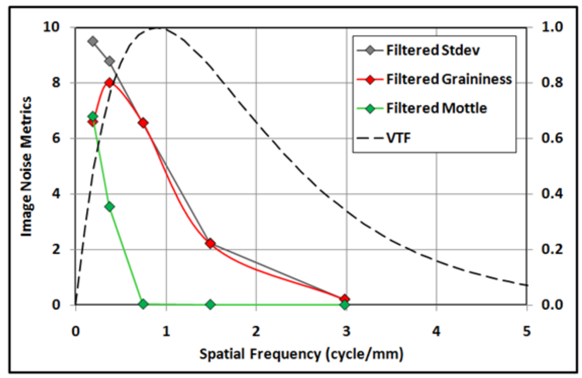

Figure 5*. Noise metrics for the images in Figure 1 after the image has been filtered by the VTF (in two dimensions).*

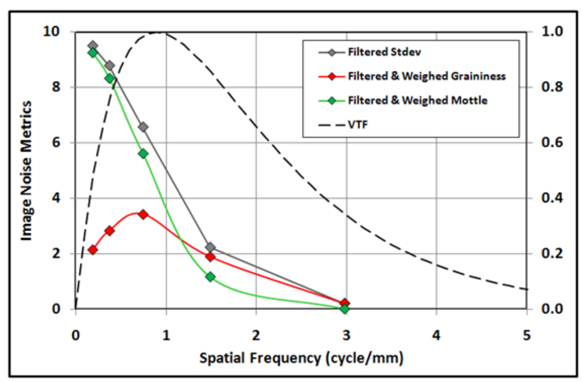

Figure 6*. Noise metrics for the images in Figure 1 after the image has been filtered by the VTF (in two dimensions) and using a tile size of 0.35mm for mottle (versus 1.27mm specified in ISO-13660).*

The same filtering when applied to the images from Figure 1 improves the correlation between the subjective graininess and the ISO graininess metric as shown in Figure 5. Note that at high frequencies the graininess becomes negligible in agreement with perception. However, the graininess metric is now greater in value than the mottle metric for all of the images except the lowest frequency. This does not agree with perception: according to most viewers' perception the images <1.49 cycles/mm exhibit "mottle" rather than "graininess." We can improve the correlation by adjusting the tile size used in the mottle calculation from 1.27mm to 0.35mm. This essentially changes the size scale threshold between graininess and mottle. Results are shown in Figure 6.

Contrast

Figure 7 shows three images with the same size scale and variation magnitude (in the L* dimension), but with different average luminance. The objective noise metrics (standard deviation, graininess, and mottle) make no distinctions among these three images. Yet human observers perceive the noise to be worse in the darker images. Generally, this preference is related to the contrast in the image, where contrast is the variation divided by the mean value. Simply scaling the noise metrics (standard deviation, graininess, and mottle) by the contrast provides better agreement with perception.

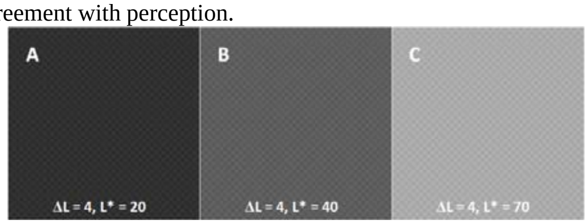

***Figure 7**. Images with the same spatial size and magnitude variations, but at different mean levels of luminance. Viewers generally judge image A to be most objectionable.*

Figure 8 shows a density tone sweep corresponding noise metrics before and after scaling by contrast.

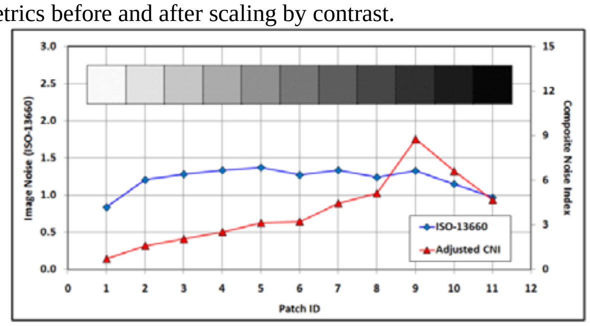

***Figure 8**. Noise metrics for a half-tone sweep, before scaling by contrast (ISO-13660) and after scaling by contrast (Adjusted CNI).*

Note that the ISO-13660 noise metric shows little variation throughout the tone sweep (blue curve with diamond markers in Figure 8). But viewers generally perceive more objectionable non-uniformity near patch #9. After scaling by contrast, the noise metric agrees with observed viewer preference (red curve with triangle markers in Figure 8).

Noise in Color Dimensions

Since color perception is tri-chromatic in nature, color images include two more dimensions where noise may be present. In the CIE XYZ color space, the tristimulus values X, Y, and Z are roughly red, green and blue, respectively.

The sensitivity of human perception to variations in each of these channels is different, so an objective metric for image noise must weigh noise in each of these dimensions accordingly. Figure 9 is an example illustrating that variations in the blue channel are less noticeable than variations of the same spatial scale and magnitude in the green and red channels.

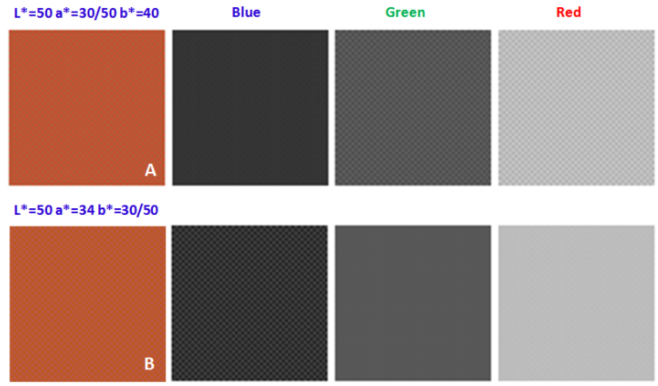

***Figure 9**. Images A and B have the the same spatial size and magnitude variations, but the variations in A are in the a* channel and in B they are in the b* channel. Color separations are shown to the right of each image. Note that most of the variation in image B is in the blue channel. Viewers generally judge image A to be less uniform.*

The factors used to weigh the color noise, α, β, and γ, are determined empirically and their use is described in the next section.

Composite Noise Index

To develop an objective metric that correlates with subjective perception of image noise, it is important to understand the factors that influence perception and give them proper weight. As described above, the key factors include:

1. Variation magnitude
2. Spatial size and frequency of variations
3. Image contrast
4. Relative importance of variations in different color dimensions

The first factor, variation magnitude, is quantified by the standard deviation, which is the starting point for all the basic noise metrics (standard deviation, mottle, and graininess). The second factor, spatial size, is covered by filtering the image according to the VTF. Also, proper selection of the mottle tile size provides a meaningful distinction between graininess and mottle. Adding weights according to image contrast and color dimension leads to the composite noise index as follows:

$$N = \sqrt{{G_{RMS}}^2 + {M_{RMS}}^2}$$

$$G_{RMS} = \sqrt{G_L^2 + G_a^2 + G_b^2} \qquad M_{RMS} = \sqrt{M_L^2 + M_a^2 + M_b^2}$$

$$G_L = K\frac{\sigma_L}{\alpha^2 \mu_L}\hat{G}_L \qquad G_a = K\frac{\sigma_a}{\beta^2 \mu_L}\hat{G}_a \qquad G_b = K\frac{\sigma_b}{\gamma^2 \mu_L}\hat{G}_b$$

$$M_L = K\frac{\sigma_L}{\alpha^2 \mu_L}\widehat{M}_L \qquad M_a = K\frac{\sigma_a}{\beta^2 \mu_L}\widehat{M}_a \qquad M_b = K\frac{\sigma_b}{\gamma^2 \mu_L}\widehat{M}_b$$

where $\boldsymbol{N}$ is the composite noise index, α, β, and γ are weights representing the influence of the L*, a*, and b* variations respectively, μ_L is the mean in the L* channel, σ_x is the standard deviation in the x channel (where x is L, a, or b), and $\hat{G}_x$ and $\hat{M}_x$ are the graininess and mottle metrics from the spatially filtered image respectively in the x channel (where x is L, a, or b). Note K is a scalar selected for convenience just so that the final noise metrics are in a familiar range for "real world" images (roughly 0 to 100).

Values for the weights, α, β, and γ, are determined empirically so that the composite noise index is well correlated with subjective perception of image noise in color images. Note that since the composite noise index, N, is a combination of grain and mottle, it is closely related to the standard deviation which could be computed for the spatially filtered image and adjusted by the color channel weights, α, β, and γ.

Application to Printed Images

Figure 10 shows images scanned from a wide range of samples printed with both laser and inkjet printers. The images shown are medium gray patches, intended to be uniform. The images are arranged in order of decreasing composite noise index.

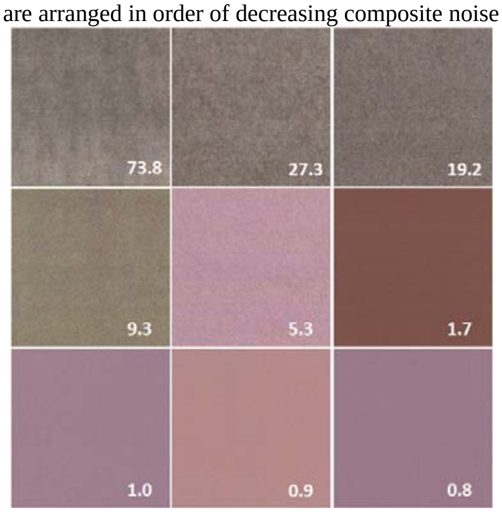

Figure 10. *Scanned images of medium gray patches printed with a range of technologies. Composite noise index values are displayed.*

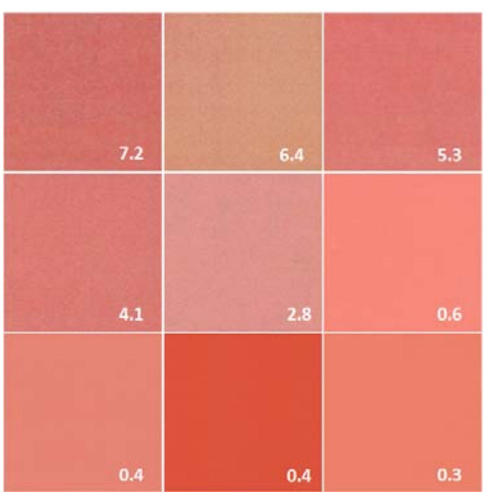

Figure 11. *Scanned images of medium red patches printed with a range of technologies. Composite noise index values are displayed.*

The ranking by composite noise index (CNI) correlates well with the subjective ranking determined by a panel of judges. The method works well despite the variation in printing technologies and the obvious variations in lightness and hue.

A similar correlation is observed for medium density red patches, as illustrated in Figure 11. This lends credence to the approach even for a very different hue regime. Note that the range of CNI values in Figure 11 overlap the values in Figure 10, and when all 18 images were sorted by CNI, the ranking is reasonable despite the mixture of grey and red images.

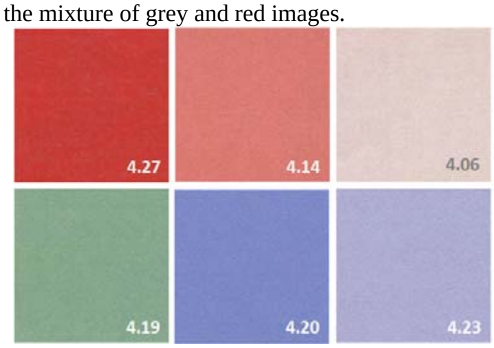

Figure 12. *Scanned images of various colored patches all with approximately equal composite noise index values.*

Various color images with approximately equal CNI values are shown in Figure 12. Indeed these images are subjectively assessed to be very similar in uniformity, despite the color differences. This indicates a good potential for using the CNI metric to quantify image noise for a broad range of images.

Summary

A single metric, composite noise index, CNI, has been proposed as an objective measurement of image noise which correlates well with subjective perception. The CNI involves 2-dimensional image filtering using the human visual transform function (VTF) to account for the spatial frequency dependence of variation perception. The CNI is also scaled according to image contrast and different empirically determined weighting factors are used for each color dimension.

The CNI metric has been shown to correlate well with subjective assessment of noise in a variety of images, including a range of printing technologies, image uniformity, colors, lightness, and hues.

References

[1] Brian W. Keelan, "Handbook of Image Quality, Characterization and Prediction," Marcel Dekker (2002).

[2] Peter G. J. Barten, "Contrast Sensitivity of the Human Eye and its Effects on Image Quality," SPIE Press (1999).

Author Biography

Dr. Ming-Kai Tse is the founder and president of Quality Engineering Associates (QEA), Inc. in Burlington, MA, USA. Prior to founding QEA, Dr. Tse was an Associate Professor of Mechanical Engineering at the Massachusetts Institute of Technology (MIT). His research interest is in measurement technology, with special application in digital printing and image quality analysis. He earned his BS at Cornell University and his SM and PhD at MIT. He is a Senior Member of IS&T. He can be reached at mingkaitse@qea.com or 1-781-221-0080.