

Measuring camera Shannon Information Capacity with a Siemens Star Image

Norman L. Koren, Imatest LLC, Boulder, Colorado, USA

Abstract

Shannon information capacity, which can be expressed as bits per pixel or megabits per image, is an excellent figure of merit for predicting camera performance for a variety of machine vision applications, including medical and automotive imaging systems. Its strength is that it combines the effects of sharpness (MTF) and noise, but it has not been widely adopted because it has been difficult to measure and has never been standardized.

We have developed a method for conveniently measuring information capacity from images of the familiar sinusoidal Siemens Star chart. The key is that noise is measured in the presence of the image signal, rather than in a separate location where image processing may be different—a commonplace occurrence with bilateral filters. The method also enables measurement of SNRI, which is a key performance metric for object detection.

Information capacity is strongly affected by sensor noise, lens quality, ISO speed (Exposure Index), and the demosaicing algorithm, which affects aliasing. Information capacity of in-camera JPEG images differs from corresponding TIFF images from raw files because of different demosaicing algorithms and nonuniform sharpening and noise reduction.

Introduction

In electronic communications systems, Shannon information capacity C defines the maximum rate in bits per second that data can be transmitted through a channel without error. For additive white gaussian noise, it is given by the deceptively simple Shannon-Hartley equation.

$$C = W \log_2 \left(1 + \frac{S}{N} \right) = W \log_2 \left(\frac{S+N}{N} \right) \quad (1)$$

While it is quite logical to extend this definition to imaging systems, where C has units of bits/pixel, a number of questions arise. How should bandwidth W (closely related to sharpness) be defined? What to use as signal S ? Can we trust measurements from consumer camera JPEG images, which often have nonuniform image processing (bilateral filtering) [1] that reduces noise in smooth areas (lowpass filtering) but sharpens the image near contrasty features such as edges (high frequency boost)?

Because nonuniform image processing is so commonly applied, it is highly desirable to measure signal and noise at the same location in the image, i.e., to measure noise in the presence of signal. We have developed a method to accomplish this with the well-known Siemens Star chart, which is a part of the ISO 12233:2014/2017 standard [2].

Measurement background

To measure signal and noise at the same location, we use an image of a sinusoidal Siemens-star test chart consisting of n_{cycles} total cycles, which we analyze by dividing the star into k radial segments (32 or 64), each of which is subdivided into m angular segments (8, 16, or 24) of length P_{seg} . The number sine wave cycles in each angular segment is $n = n_{cycles}/m$.

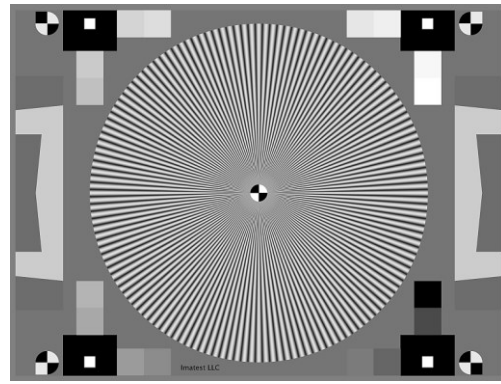


Figure 1. 144-cycle Siemens star pattern.

For radius r in pixels, the spatial frequency in Cycles/Pixel is $f = n_{cycles}/2\pi r$. This means that the Nyquist frequency, $f_{nyq} = 0.5$ C/P is located at $r = 46$ pixels for a 144-cycle star and 23 pixels for a 72-cycle star [3].

We calculate the ideal (pure sine wave + second harmonic) signal $S_{ideal}(\varphi)$ for each segment from the actual (noisy) input signal $S_{input}(\varphi)$, using Fourier series coefficients a_j and b_j [4].

$$S_{ideal}(\varphi) = \sum_{j=1}^2 a_j \cos\left(\frac{2\pi j n \varphi}{P_{seg}}\right) + b_k \sin\left(\frac{2\pi j n \varphi}{P_{seg}}\right) \quad (2)$$

where

$$a_j = \frac{2}{P} \int_P S_{input}(x) \cos\left(\frac{2\pi j n x}{P_{seg}}\right) dx$$

$$b_j = \frac{2}{P} \int_P S_{input}(x) \sin\left(\frac{2\pi j n x}{P_{seg}}\right) dx \quad (3)$$

The second harmonic term ($j=2$) term improves results for JPEG images from cameras, which frequently have “shoulders” (regions of reduced contrast) in their tonal response curves, making them difficult to linearize with a simple gamma formula.

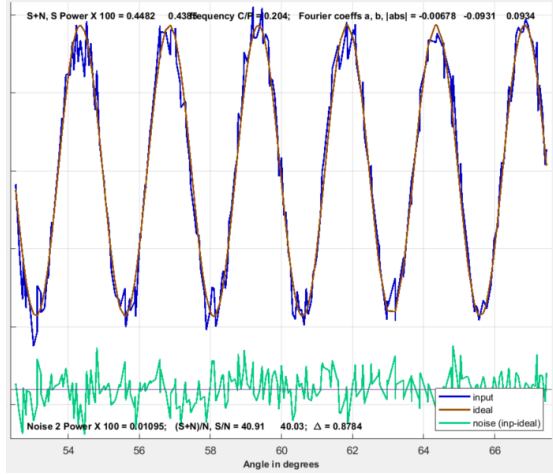


Figure 2. Illustration of $s_{input}(\varphi)$ (blue), $s_{ideal}(\varphi)$ (brown), and noise $N(\varphi) = s_{ideal}(\varphi) - s_{input}(\varphi)$ (green).

The spatial domain noise in each segment is

$$N(\varphi) = s_{ideal}(\varphi) - s_{input}(\varphi) \quad (4)$$

The roughness of s_{input} and $N(\varphi)$ in Figure 2, which is derived from all points in the segment (sorted), does not represent the actual frequency distribution of the noise, and does not affect the calculation of frequency domain noise power $N(f)$. The noise spectrum is more closely related to the noise image (with signal removed), illustrated in Figure 8.

The mean signal power used to calculate C is $S(f) = \text{mean}(a_n^2 + b_n^2)/2$ for all angular segments at radius r (where $f = n_{cycles}/2\pi r$). Frequency domain noise power $N(f)$ is the mean of the variance (σ^2) of the middle 80% of $N(\varphi)$ for each angular segment (there can be irregularities near the ends).

We started the analysis the integral form of the Shannon information capacity equation [5], which is appropriate for one-dimensional systems:

$$C = \int_0^B \log_2 \left(1 + \frac{S(f)}{N(f)} \right) df = \int_0^B \log_2 \left(\frac{S(f)+N(f)}{N(f)} \right) df \quad (5)$$

Where C is the channel capacity in bits/pixel; B = Nyquist frequency f_{Nyq} ; $S(f)$ is the signal power spectrum, $N(f)$ is the noise power spectrum (which we interpret as broadband noise in the presence of signal $S(f)$).

The difficulty with this equation is that Information capacity applies to two-dimensional images, whereas all sharpness-related measurements (MTF and $S(f)$) refer to *one* dimension, for example, MTF in units of cycle/pixel actually means cycle per pixel *length*. To overcome this limitation, we express C as a double integral.

$$C = \iint_0^B \log_2 \left(\frac{S(f_x, f_y) + N(f_x, f_y)}{N(f_x, f_y)} \right) df_x df_y \quad (6)$$

where f_x and f_y are frequencies in the x and y -directions, respectively. In order to evaluate this integral, we transform x and y into polar coordinates r and θ .

$$C = \int_0^{2\pi} \int_0^B \log_2 \left(\frac{S(f_r, f_\theta) + N(f_r, f_\theta)}{N(f_r, f_\theta)} \right) f_r df_r df_\theta \quad (7)$$

Since $S(f_r, f_\theta)$ and $N(f_r, f_\theta)$ are only weakly dependent on θ , (7) can be rewritten in one-dimension.

$$C = 2\pi \int_0^B \log_2 \left(\frac{S(f) + N(f)}{N(f)} \right) f df \quad (8)$$

The minimum star frequency is $f_{min} = n_{cycles}/2\pi r_{star}$ where r_{star} is the radius of the star in pixels. The contribution of frequencies below f_{min} is calculated by extrapolation— by assuming that $S(f) = S(f_{min})$ and $N(f) = N(f_{min})$ for $f < f_{min}$. This assumption is reasonable for $N(f)$ but somewhat pessimistic for $S(f)$. However, the effect on C is minimal because of the presence of f_r in the integrand of Equation (8).

C is strongly dependent on noise $N(\varphi)$, which is calculated from the difference between the input signal (which includes noise) and the perfect sine wave signal. This causes C to *decrease* in the presence of artifacts such as aliasing and saturation. Aliasing has little effect on standard MTF measurements, and saturation (or clipping) actually improves measurements even though it degrades system performance. We will discuss aliasing in more detail.

Measurement method

A key challenge in measuring information capacity is how to define mean signal power S . Ideally, the definition should be based on a widely-used test chart. For convenience, the chart should be scale-invariant (so precise chart magnification does not need to be measured). And, as we indicated, signal and noise should be measured at the same location. For different observers to obtain the same result the chart design and contrast should be standardized.

To that end we recommend a sinusoidal Siemens star chart similar to the chart specified in ISO 12233:2014/2017, Annex E. Contrast should be as close as possible to 50:1 (the minimum specified in the standard; close to the maximum achievable with matte media). Higher contrast can make the star image difficult to linearize. Lower contrast is acceptable, but should be reported with the results. The chart should have 144 cycles for high resolution systems, but 72 cycles is sufficient for low resolution systems. The center marker (quadrant pattern), used to center the image for analysis, should be 1/20 of the star diameter.

Acquire a well-exposed image of the Siemens star in even, glare-free light. Exposures should be reasonably consistent when multiple cameras are tested. The mean pixel level of the linearized image inside the star should be in the range of 0.16 to 0.36. (The optimum has yet to be determined.)

The center of the star should be located as close as possible to the center of the image to minimize measurement errors caused by optical distortion (if present).

The size of the star in the image should be set so the maximum spatial frequency, corresponding to the minimum radius r_{min} , is larger than the Nyquist frequency f_{Nyq} , and, if possible, no larger than $1.3 f_{Nyq}$, so sufficient lower frequencies are available for the channel capacity calculation. This means that a 144-cycle star with a 1/20 inner marker should have a diameter of 1400-1750 pixels and a 72-cycle star should have a diameter of 700-875 pixels. For high-quality inkjet printers, the physical

diameter of the star should be at least 9 (preferably 12) inches (23 to 30 cm).

Other features may surround the chart, but the average background should be close to neutral gray (18% reflectance) to ensure a good exposure (it is OK to apply exposure compensation if needed). Figure 3 shows a typical star image in a 24-megapixel camera.



Figure 3. Typical image of Siemens star, scaled to the maximum frequency is slightly above f_{Nyq} .

Information capacity results

We tested three cameras with excellent lenses that produced both raw and JPEG output for information capacity as a function of Exposure Index (ISO speed setting).

1. Panasonic Lumix LX5. An older (2010) compact 10.1-megapixel camera with a $2.14 \mu\text{m}$ pixel pitch and a Leica $f/2$ zoom lens set to $f/4$.
2. Sony A6000. A 24-megapixel micro four-thirds camera with a $3.88 \mu\text{m}$ pixel pitch and a 60mm Canon macro lens set to $f/8$.
3. Sony A7Rii. A 42-megapixel full-frame camera with a backside-illuminated (BSI) sensor with $4.5 \mu\text{m}$ pixel pitch and a 90mm Sony macro lens set to $f/8$.

We captured both JPEG and raw images, which were converted to 24-bit sRGB TIFF images with no sharpening or noise reduction (designated as raw/TIFF) and $\gamma \cong 2.2$. Results with 48-bit Adobe sRGB conversion were nearly identical.

Figure 4 illustrates typical results for the camera 2 raw/TIFF image at ISO 100. The noise spike near the Nyquist frequency (0.5 C/P) is primarily caused by aliasing from the raw converter (the highest quality algorithm in dcraw). It disappears when the lens is stopped down.

The signal $S(f)$ (which is proportional to MTF), noise $N(f)$ (scaled the same as signal), and $(\text{Signal}+\text{Noise})/\text{Noise}$ ratio in Figure 4 are all calculated from the same sinusoidal star pattern. Noise measurements are more accurate than for flat patches (found in standard test charts such as ISO 14524, ISO 15739, and others) in the presence of widely-used bilateral filtering, which reduces noise more aggressively in smooth areas (see the upper-left of Figure 8).

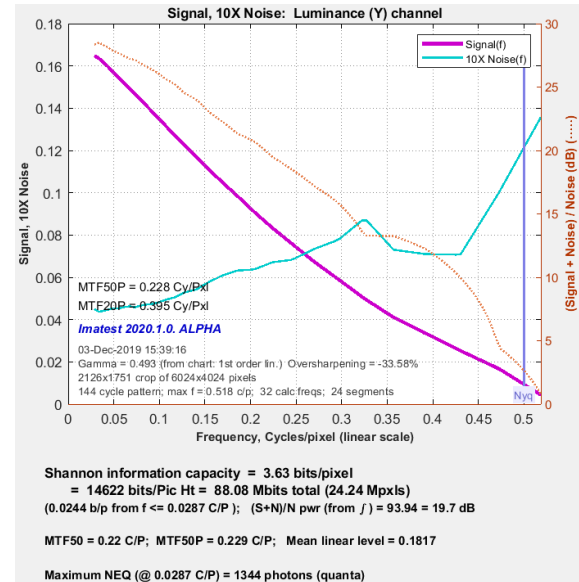


Figure 4. Signal, 10x noise, and $(S+N)/N$ plot for Camera 2, TIFF from raw, ISO 100, 32 frequencies, 32 radial segments. Summary results are shown below the plot. Note that signal $S(f)$ is proportional to $\text{MTF}(f)$.

Figure 5 shows results from raw/TIFF images (solid lines) and JPEG images (dotted lines) as a function of ISO speed. Raw/TIFF results are consistent for the three cameras. Results for camera 2 are similar to camera 1 at $2.5\times$ the ISO speed (2000 vs. 800), and results for camera 3 are similar to camera 2 at $3\times$ the ISO speed (6000 vs. 2000). The improvement between cameras 1 and 2 is somewhat lower than the ratio of the pixel areas, but the improvement between cameras 2 and 3 is much greater because camera 3 has a BSI sensor (a greatly improved technology).

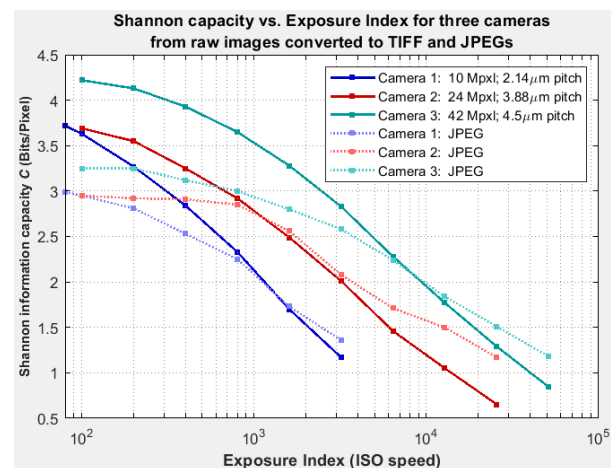


Figure 5. Information capacity for the three cameras as a function of Exposure Index: solid lines for TIFFs derived from raw images; dotted lines for JPEGs.

Using the results in Figure 5, we can compare images from two different cameras with similar information capacities (1.7 bits/pixel): camera 1 ($2.14 \mu\text{m}$ pixel pitch) at ISO 1600 and camera 3 ($4.5 \mu\text{m}$ BSI pixel pitch) at ISO 12800.

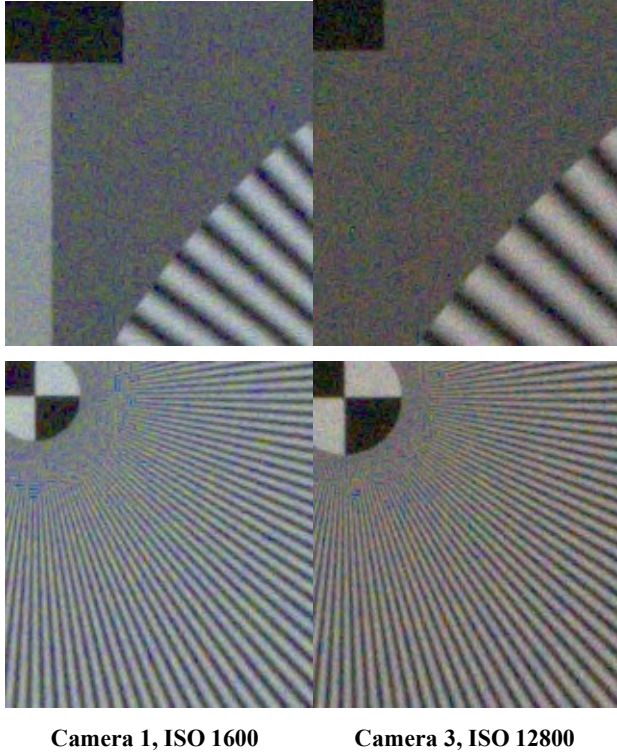


Figure 6. Comparison of raw/TIFF images from two different cameras with similar information capacity (≈ 1.7), significantly degraded from ISO 100.

Noise appears very similar in the upper images, and sharpness and artifacts are similar in the lower images. Note that since the star chart is scale invariant, sharpness at the same position of the two lower images can be directly compared, even though the center pattern differs in size.

Sharpening and noise reduction

Unsurprisingly, we found no combination of linear sharpening and noise reduction (lowpass filtering) that improved information capacity C .

In Table 1, the baseline image (Camera 2, raw/TIFF, ISO 100) has no sharpening or noise reduction. USM is Unsharp Mask; R_n = sharpening radius; A_n = sharpening amount. Gaussian g is Gaussian filter parameter g .

Table 1. Capacity losses for sharpening & noise reduction

Image	C	MTF50 (c/p)	MTF50P(c/p)
Baseline	3.69	0.22	0.229
USM R2 A1	3.65	0.345	0.323
USM R1 A2	3.63	0.407	0.397
Gaussian 0.7	2.99	0.162	0.168
Gaussian 1.0	2.25	0.138	0.143
USM R2A1, Gaussian 0.7	3.06	0.241	0.239

Sharpening increases both bandwidth and noise. The noise increase with sharpening (RMS noise voltage = 0.01 for the USM R1A2 image vs. 0.004 for the baseline: 8 dB difference) was consistent with measurements in gray regions just outside the

star circle. The USM R1A2 (sharpened) image had $\text{SNR} = 30.0 = 29.6 \text{ dB}$. The baseline image had $\text{SNR} = 70.2 = 36.9 \text{ dB}$.

JPEG behavior

Calculated information capacity (Figure 5) is lower for camera JPEG images than for raw/TIFF images at low ISO speeds, but it decreases more slowly as ISO speed increases, and can be higher at high ISO speeds.

The reasons are twofold. (1) JPEG images tend to have strong sharpening at low ISO speeds, which boosts noise, but less sharpening at high ISO speeds, as illustrated by MTF_{50P} —the spatial frequency where MTF drops to 50% of its peak value—in Figure 7. By comparison, raw/TIFF images maintain relatively constant MTF_{50P} , except at the highest ISO speeds.

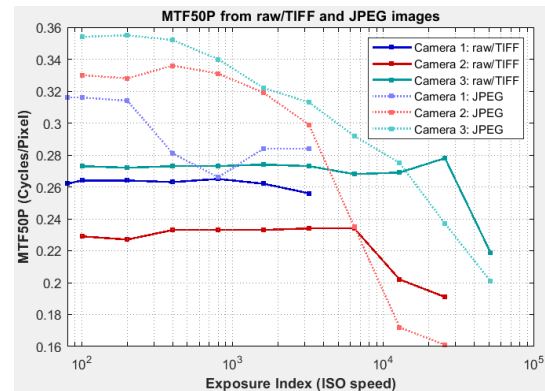


Figure 7. MTF_{50P} as a function of Exposure Index (ISO speed). Relatively constant for raw/TIFF images, except at the highest ISO speeds.

(2) Additional noise reduction (bilateral filtering) [1] is often applied to JPEG images, especially at high ISO speeds. Information capacity measurements are not reliable in the presence of bilateral filtering because low contrast detail is smoothed (lowpass filtered). This increases measured values of information capacity, while decreasing the actual capacity. For this reason, measurements of C from JPEG images should always be treated with caution. Bilateral filtering is not an issue with raw images.

Using Equation (4) we created complete noise images (with signal removed) and examined them for appearance with the intent of detecting and compensating bilateral filtering. Figure 8 shows a portion of the noise images for raw/TIFF and JPEG images from camera 2 at ISO 25600.

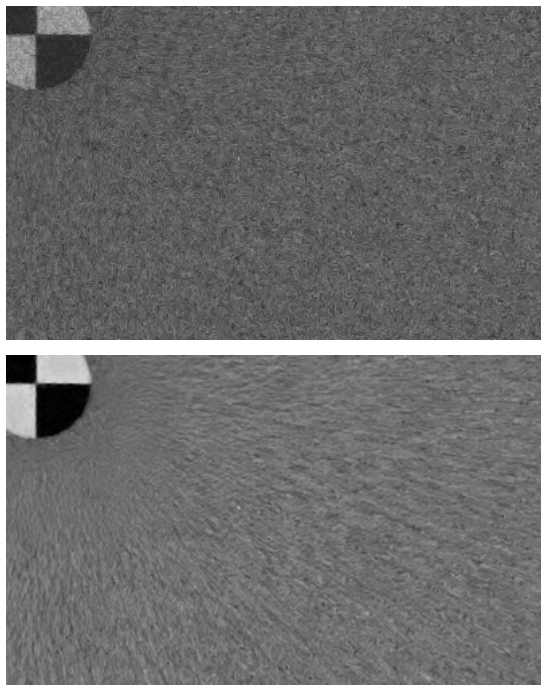


Figure 8. Noise-only images near center of star for camera 2 at ISO 25600. Raw/TIFF (above) and JPEG (below).

The appearance is strikingly different, especially in the center marker, where noise has been effectively removed in the JPEG image and is greatly reduced just outside the center marker. The spectra for the two images are also different (as expected), but this difference was not seen consistently when comparing RAW/tiff and JPEG images from other cameras.

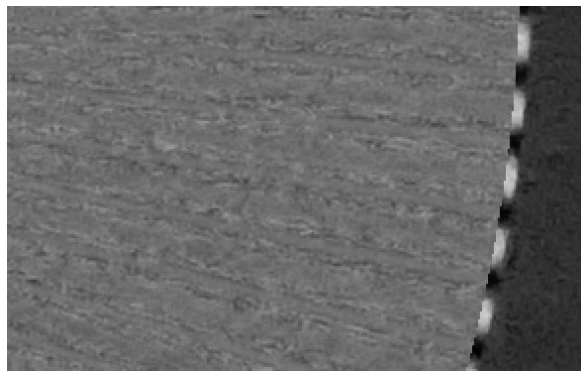


Figure 9. Noise-only image near the edge of star for camera 2 JPEG at ISO 25600.

Careful examination of Figure 9 shows that noise is largest at angles where the brightness gradient is largest. We tried measuring information capacity using only these regions, but results were inconsistent, apparently because the noise variation did not continue in the interior of the star.

At this time, we have found no reliable means of detecting bilinear filtering. On a positive note, the errors caused by bilateral filtering are not extremely large. For cameras 1-3 set to their highest ISO speeds (for extremely poor lighting), the discrepancies for C are 0.2, 0.5, and 0.3 bits/pixel, respectively.

Ideal image

The Information capacity of an “ideal” image is of more than passing interest. We used a small version of the image used to print test charts (which would have been about 2 cm high if printed).

The “ideal” image has not passed through a Bayer color filter array, has not been demosaiced, and has not been processed in any way other than gamma encoding.

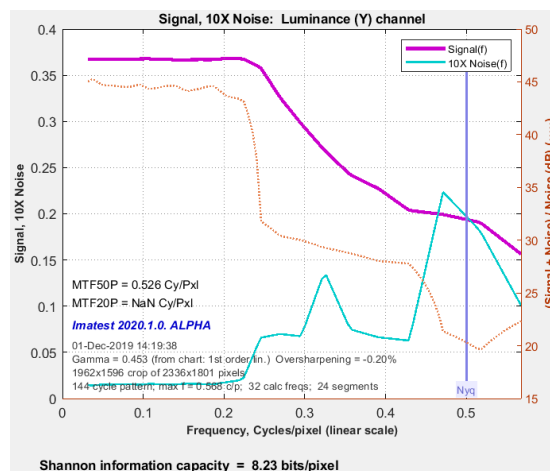


Figure 10. Signal, 10X noise, and SNR for “ideal” image.

$C = 8.23$ bits/pixel is slightly above the expected maximum of 8 bits/pixel due to undetermined numerical causes. It’s twice the highest value we’ve seen from the best camera images, all of which have been passed through Bayer color filter arrays, then demosaiced. We were not able to test a Foveon sensor (where all colors are at the same location) in time for the paper. The signal rolloff above 0.25 cycles/pixel is caused by anti-aliasing filtering intended to smooth jagged features.

RGB channels

Most of the results in this paper are from the luminance (Y) channel, where $Y = 0.2125 \cdot R + 0.7154 \cdot G + 0.0721 \cdot B$. Many authors recommend measuring C separately for the R, G, and B channels. For the Camera 2 raw/TIFF image, which has $C = 3.69$ bits/pixel for the Y-channel, $C_R = 2.90$, $C_G = 3.53$, and $C_B = 3.07$ for the R, G, and B channels, respectively. C is higher for the Y-channel because combining the uncorrelated noise from the R, G, and B channels reduces the Y-channel noise.

Total $C = C_R + C_G + C_B$ is nearly triple when the R, G, and B color channels are analyzed separately. But we can now think of a pixel as having 24 bits instead of 8.

The undemosaiced version of this image has information capacities of 5.62, 5.95, 5.14, and 5.81 bits/pixel for the R , G_R , B , and G_B channels, respectively. The mean of the four undemosaiced channels, 5.63, is well above the mean of the R, G, and B channels, 3.17. Likely causes include (1) the strong correlation between demosaiced channels, where the contents of each color channel is strongly influenced by the contents of neighboring channels, and (2) the lack of aliasing from demosaicing. Further study is required to fully characterize the effects of demosaicing on information capacity.

In the future we plan to analyze chroma channels for information capacity. The most likely candidates are B-Y and R-

Y, which are used for chroma noise. C_B and C_R (from $Y C_B C_R$) are derived from B-Y and R-Y, and should have similar SNR and information capacity C . We will need special charts (most likely stars with R-G and B-G patterns).

Aliasing and demosaicing

There has recently been increased interest in aliasing—low frequency artifacts such as moiré fringing that can appear when the signal reaching the image sensor has significant energy above the Nyquist frequency, f_{Nyq} (0.5 Cycles/Pixel) [3]. Images from Bayer sensors can have significant aliasing because the Nyquist frequency of the red and blue channels is half that of the sensor as a whole.

The amount of aliasing is strongly dependent on the demosaicing algorithm used by the raw converter. Simple algorithms such as bilinear demosaicing have severe aliasing. Most modern cameras use sophisticated algorithms that use detail from all color channels to construct the image in each individual channel.

We examined the effects of demosaicing using the four algorithms available in dcrw [6]: bilinear (well-known for poor quality), VNG, PPG, and AHD (in order of increasing quality) and the recommended AMaZE algorithm in RawTherapee [7], which offers many additional algorithms.

Table 2. Comparison of demosaicing algorithms

Demosaicing algorithm	C (bits/pixel)	$MTF50P$ (C/P)
dcrw bilinear	1.8	0.191
dcrw VNG	2.51	0.219
dcrw PPG	3.14	0.229
dcrw AHD	3.69	0.229
RawTherapee AMaZE	3.95	0.236

Information capacity C is a far more sensitive measure of demosaicing quality than $MTF50P$.

The enlarged images below compare three of the algorithms. The orange circle is f_{Nyq} . The two visible cyan circles are at $0.75f_{Nyq}$ and $0.5f_{Nyq}$.

Severe aliasing in bilinear demosaicing (Figure 13; a very simple form of linear interpolation) is plainly visible.

To summarize, these results show that information capacity C is an excellent measurement for evaluating demosaicing quality, which strongly affects aliasing. Unfortunately, we presently have no easy way of separating the effects of aliasing from noise and other artifacts.

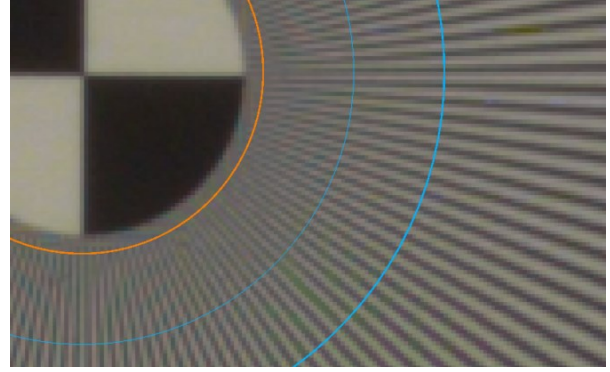


Figure 11. RawTherapee AMaZE demosaicing.

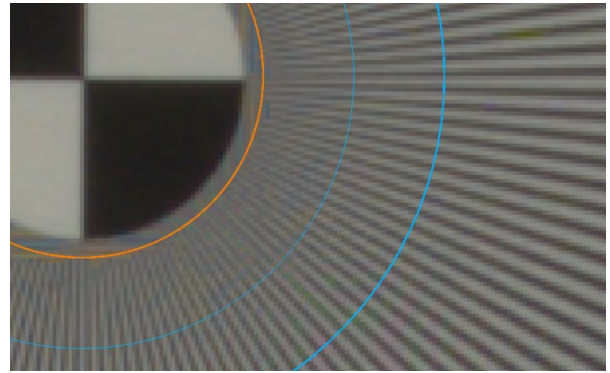


Figure 12. dcrw Adaptive Homogeneity-Directed (AHD) demosaicing.

Both AMaZE and AHD demosaicing are high quality, but AMaZE has subtly lower aliasing and fewer artifacts near f_{Nyq} . The differences are difficult to see.

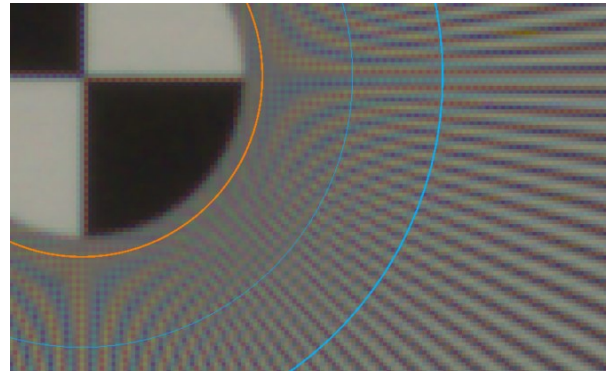


Figure 13. Bilinear demosaicing (famously low quality)

Data compression

We used Information capacity measurements to study data compression by taking a high quality raw/TIFF image acquired at ISO 100 from camera 2, loading it into Irfanview [8], then saving it as JPEG and JPEG 2000 (JP2) files at quality levels from 10 to 100. Information capacity is a strong function of quality level for both file types. We treat both types of compression as “black box” image processing. We have not examined their inner workings.



Figure 14. Information capacity C vs. quality level.

[9] lists several metrics used to evaluate JPEG data compression, then states, "Of all measures applied to the compression system, the JPEG 'quality factor', used to specify the compression level, was found to have the best performance when predicting the results of subjective tests."

More significant is information capacity as a function of file size. JPEG 2000 seems to be the clear winner here.

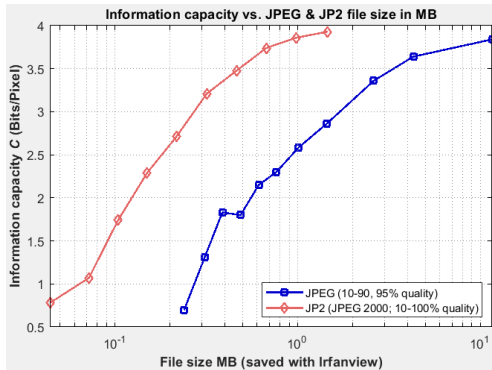


Figure 15. Information capacity C vs. file size in MB. Original file size = 72 MB ($4000 \times 6000 \times 3$).

Figure 15 may be compared with results in [9].

MTF_{50P} , the spatial frequency in Cycles/Pixel where MTF drops to half its peak value, muddies the comparison somewhat. It is calculated for both from the star and a slanted-edge adjacent to the star. MTF_{50P} numbers for the star may be biased slightly high because of the limited low frequencies for normalization.

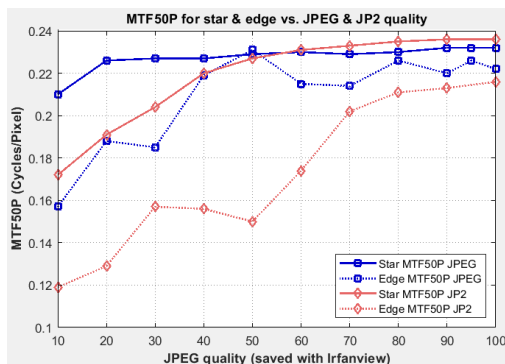


Figure 16. MTF_{50P} vs. quality level. Biased slightly high for the star. JPEG 2000 has greater MTF_{50P} loss at quality levels below 70%.

These results illustrate the potential of using information capacity for evaluating image compression algorithms. Low contrast and colored stars may be needed to complete the analysis, and information capacity needs to be correlated with perceptual image quality.

Figures 17 and 18 show JPEG image degradation at quality levels of 70 and 50%. Note that these images have a similar appearance to the original in Figure 11, in part because the spatial frequencies in the star are inversely proportional to the radius, hence half the information capacity C comes from the small region between $f_{Nyq}/2$ (the strong cyan circle) and f_{Nyq} (the orange circle) near the center of the star.

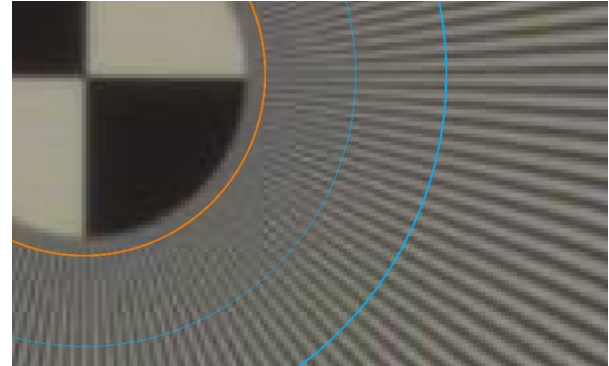


Figure 17. 70% JPEG quality level. Compare with Fig. 9. Some contrast loss is observed near the Nyquist frequency.

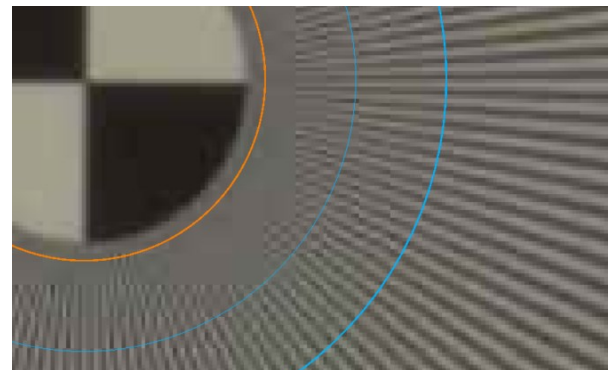


Figure 18. 50% JPEG quality level. Significant contrast loss is observed above 0.75 Nyquist (inside the thin cyan circle).

Another reason Figures 17 and 18 have a similar appearance is that information capacity C is an *informational* rather than a *perceptual* metric. SSIM (structural similarity) [10] may be more appropriate when perceptual differences need to be emphasized. The SSIM plot below clearly emphasizes where image appearance differs. (SSIM = 1 for no difference.)

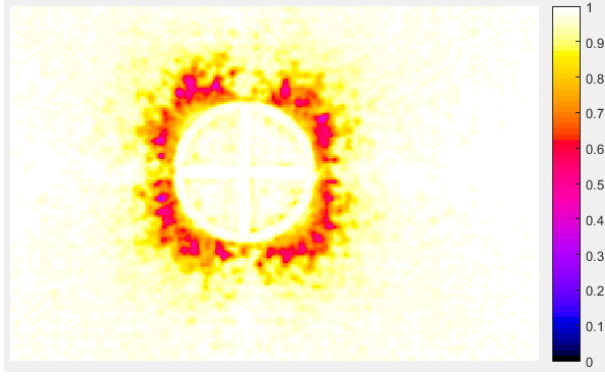


Figure 19. SSIM results, comparing 100% and 50% images. Regions where the images differ visually are strongly emphasized.

SNRI and Object Detection

The measurements described above enable the calculation of *SNRI*, which is a key figure of merit for the task of detecting small objects in an image. Reference [11], Equation (3) defines *SNRI* for detecting a “difference object” whose Fourier transform is $G(v)$, where v is spatial frequency.

$$SNRI^2 = K^2 \int \frac{|G(v)|^2 MTF_{sys}^2(v)}{NPS(v)} dv \quad (9)$$

Noting that $K^2 MTF_{sys}^2(v)$ is equivalent to $S(f)$ and $NPS(v)$ is equivalent to $N(f)$ in Equation (5),

$$SNRI^2 = \int \frac{|G(f)|^2 S(f)}{N(f)} df \quad (10)$$

For simplicity we assume that the difference object is a one-dimensional rectangle of amplitude h and width w , with Fourier transform,

$$G(f) = h \frac{\sin(\pi w f)}{\pi w f} \quad (11)$$

If necessary, we could change this to a two-dimensional rectangle used in [11], but for now one dimension should be sufficient because object detection is limited by the smaller rectangle dimension. We could also calculate $G(f)$ for different types of objects.

[11] has an excellent discussion of the uses of *SNRI*, which will not be discussed further here.

Noise Equivalent Quanta (NEQ)

Equations (6) and (7) lead directly to the definition of Noise Equivalent Quanta [12-14].

$$NEQ(f) = \frac{S(f)}{N(f)} \quad (12)$$

NEQ is the number of quanta that would result in measured Signal-to-Noise Ratio $SNR(f) = \sqrt{NEQ(f)}$ where *SNR* is a *vol-tage* ratio. In the future we plan to use *NEQ* to calculate Detective Quantum Efficiency $DQE(f) = NEQ(f)/\bar{q}$ where \bar{q} is the average number of input quanta per unit area. This will require careful calibration involving the illumination level and spectrum. *DQE*

is of particular interest for applications where available light is limited.

Reference [12] discusses several uses of *NEQ*.

Total camera capacity from slanted-edges

The next step after finding the information capacity of a pixel is to find the total capacity, C_{total} , for the camera. Unfortunately, it can't be reliably obtained by multiplying C by the number of megapixels because lens sharpness (*MTF* response) tends to be nonuniform, typically decreasing with distance from the image center.

Although the preferred method of measuring C_{total} is to use an array of Siemens stars, we have developed a fast and convenient alternative method using an array of slanted-edges. The calculation is called “star-equivalent information capacity,” C_{starEq} — the approximate information capacity that would have been calculated from a Siemens star. The procedure sounds complex, but it is fast and easy to perform, and it is quite effective because C_{starEq} tracks the star measurement C_{star} (designated as C , above).

C_{starEq} is only recommended for minimally-processed images converted from raw capture because it is strongly affected by sharpening and noise reduction (problematic because sharpness and noise are measured at separate locations for slanted edges). It is not recommended for in-camera JPEG images, and it should not be used as a primary information capacity measurement.

To measure C_{starEq} , begin by linearizing the image (if required), then measure $MTF(f)$ from low contrast slanted-edges. (4:1 contrast ratio, specified in ISO 12233, is recommended.) Be sure the edge is exposed well, i.e., not saturated. Note that the edge contrast doesn't directly affect the results.

Measure the pixel levels of the light and dark parts of the slanted-edge region (away from the edge), Px_{lt} and Px_{dk} , and the noise voltages in the two regions, $N_{lt} = \sigma(Px_{lt})$ and $N_{dk} = \sigma(Px_{dk})$. The mean values of the pixel levels and noise are Px_{mean} and N_{mean} , respectively. Px_{mean} should be close to the mean pixel level of the star measurement.

Because C_{starEq} is extremely sensitive to saturation, which reduces the measured values of N_{dk} or N_{lt} , it should not be calculated if $Px_{lt} > 0.95$ or $Px_{dk} < 0.03$.

For assumed equivalent star contrast $Contrast_{star}$, the modulation (Michelson contrast) is

$$modulation = Px_{mean} \frac{Contrast_{star}-1}{\sqrt{2} (Contrast_{star}+1)} \quad (13)$$

The $\sqrt{2}$ factor causes *modulation* to have units of RMS voltage: useful because *modulation*² has units of mean power. The signal power for calculating C_{starEq} is

$$S(f) = (modulation \times MTF(f))^2 \quad (14)$$

The noise power of the assumed star is

$$N(f) = N_{mean}^2 \quad (15)$$

C_{starEq} is calculated by substituting (14) and (15) into (8). The total camera information capacity is

$$C_{total} = \frac{C_{star}(\text{center}) \times \text{mean}(C_{starEq})}{C_{starEq}(\text{center})} \times \text{megapixels} \quad (16)$$

$\text{mean}(C_{\text{starEq}})$ is the mean of C_{starEq} taken over a large number of slanted-edges distributed over the image. This formula also applies to arrays of Siemens stars, where it simplifies to $C_{\text{total}} = \text{mean}(C_{\text{star}}) \times \text{megapixels}$.

Summary and future work

Information capacity is of great interest as a figure of merit for evaluating camera image quality because it combines sharpness with noise. But until now it hasn't been easy to measure. We have presented a method for conveniently measuring information capacity, signal $S(f)$ (proportional to MTF), and noise $N(f)$ from images of Siemens star test charts.

We have shown that information capacity behaves as expected for cameras with different sensors and ISO speed settings, that visual comparisons of images with similar information capacities are similar in appearance, and that it is sensitive to losses from aliasing and data compression.

Although information capacity is still unfamiliar to most engineers in the imaging industry, its units— information bits per pixel (or total image)— are intuitive and easy to understand. In short, we believe that it is a better indicator of a camera's potential performance (after tuning) than plain megapixels.

In the future we would like to see information capacity — either megabits per pixel or megabits total at specified ISO speeds (exposure indices) or light (lux) levels — become an integral part of a standard camera specifications.

For information capacity to become widespread, the measurement procedure must be standardized. That means that chart contrast (which determines signal level), exposure, chart size in the image, and other factors that affect the results need to be specified in a recognized standard.

The basic signal and noise parameters measured with the Siemens star, $S(f)$ and $N(f)$, can be used to calculate additional measurements used in detection theory, such as $SNRI$ and Noise Equivalent Quanta (NEQ), which are discussed elsewhere [12,13].

In the near future we plan to calculate color channel capacity using U and V (from YUV) (chroma) channels with special color star charts.

Future work includes

- additional correlation of information capacity C with visual appearance for a variety of images, without and with additional image processing (sharpening, noise reduction, color correction, etc.),
- correlating C with performance of machine vision and Artificial Intelligence systems,
- evaluating the quality of cameras with non-traditional color filter arrays (RCCC and RCCB, where C designates "Clear"),
- extending the model of C to include viewing conditions and the human visual system (contrast sensitivity function) to obtain a "visual information capacity" analogous to visual noise or acutance.
- verifying the suitability of $SNRI$ for performing specific tasks (determining the detectability of features of various sizes),
- using NEQ as the basis for Detective Quantum Efficiency (DQE) calculations.

Acknowledgments

The author wishes to offer special thanks to Robin Jenkin for his generous assistance and Paul Romanczyk and Ty Cumby for their help in developing the key equations.

References

- [1] Tomasi, C., and R. Manduchi. "Bilateral Filtering for Gray and Color Images". Proceedings of the 1998 IEEE International Conference on Computer Vision. Bombay, India. Jan 1998, pp. 836–846.
- [2] ISO 12233:2017 Photography – Electronic still picture imaging – Resolution and spatial frequency response [Online], www.iso.org/standard/71696.html.
- [3] U. Artmann, "Quantify Aliasing a new approach to make resolution measurement more robust," Electronic Imaging 2019, Society for Imaging Science and Technology.
- [4] en.wikipedia.org/wiki/Fourier_series, (Eq.1) [Online]
- [5] C. Shannon, "Communication in the Presence of Noise", Proceedings of the I.R.E., January 1949, pp. 10-21.
- [6] www.dechifro.org/dcrow/dcrow.1.html [Online] dcrow offers four demosaicing algorithms, listed under Interpolation Options.
- [7] rawtherapee.com/ [Online] RawTherapee is a free raw converter program that has many demosaicing algorithms.
- [8] www.irfanview.com/ [Online] A popular free image viewer/editor.
- [9] R. Jenkin, R. Jacobson, M. Richardson, "Use of the first order Wiener kernel transform in the evaluation of SQRIn and PIC quality metrics for JPEG compression," Proc. SPIE 5294, Image Quality and System Performance, 2003, Electronic Imaging 2004.
- [10] Z. Wang, A. Bovik, and H. Sheikh, "Image Quality Assessment: From Error Visibility to Structural Similarity", IEEE Transactions on Image Processing, Vol. 13, No. 4, April 2004.
- [11] P. Kane, "Signal Detection Theory and Automotive Imaging," Electronic Imaging 2019, Society for Imaging Science and Technology.
- [12] B. Keelan, "Imaging Applications of Noise Equivalent Quanta," Electronic Imaging 2016, Society for Imaging Science and Technology.
- [13] I. Cunningham and R. Shaw, "Signal-to-noise optimization of medical imaging systems," Journal of the Optical Society of America, Vol. 16, No. 3, March 1999, pp. 621-632.
- [14] "Medical Imaging - The Assessment of Image Quality", ICRU Report 54 (International Commission of Radiation Units and Measurements, Bethesda, 1995).

Author Biography

Norman Koren became interested in photography while growing up near the George Eastman House photographic museum in Rochester, NY. He received his BA in physics from Brown University (1965) and his Masters in physics from Wayne State University (1969). He worked in the computer storage industry simulating digital magnetic recording systems and channels for disk and tape drives from 1967-2001. He founded Imatest LLC in 2003 to develop software and test charts to measure the quality of digital imaging systems.