Measuring Camera Information Capacity with Slanted Edges

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

Abstract

As Machine Vision (MV) and Artificial Intelligence (AI) are incorporated in an ever-increasing range of imaging applications, there is a corresponding need for camera measurements that accurately predict the performance of these systems. At the present time, the standard practice is to separately measure the two major factors—sharpness and noise (or Signal-to-Noise Ratio)— along with several others, then to estimate system performance based on a combination of these factors. This estimate is usually based on experience, and is often more of an art than a science.

Camera information capacity, based on Claude Shannon's groundbreaking work on information theory [1],[2] holds great promise as a figure of merit for a variety of imaging systems, but it has traditionally been difficult to measure [3],[4],[5],[6].

We describe a new method for measuring camera information capacity that uses the popular slanted-edge test pattern, specified by the ISO 12233:2014/2017 standard [7]. Measuring information capacity requires no extra effort; it essentially comes for free with slanted-edge MTF measurements. Information capacity has units of bits per pixel or bits per image for a specified ISO speed and chart contrast, making it easy to compare very different cameras. The new measurement can be used to solve some important problems, such as finding a camera that meets information capacity requirements with a minimum number of pixels— important because fewer pixels mean faster processing as well as lower cost.

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.

$$C = W\ \log_2\left(1+\frac{S}{N}\right) = \int_0^W \log_2\left(1+\frac{S(f)}{N(f)}\right) \tag{1}$$

While it is quite logical to extend this definition to imaging systems, where *C* has units of bits/pixel, bandwidth, *W*, signal power, *S,* and noise power, *N,* have to be measured and applied with great care. Consumer camera JPEG images often have nonuniform image processing (bilateral filtering) [6] that sharpens images near contrasty features such as edges (boosting high frequencies) but reduces noise elsewhere (lowpass filtering), complicating measurements.

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 slanted edge pattern, which is a part of the ISO 12233:2014/2017 standard [7].

The slanted edge measurement

For context, we briefly review the slanted-edge algorithm.

1. **The image should be well-exposed**, avoiding the dark "toe" and light "shoulder" regions, where the image deviates from standard log-linear behavior, $\log(pixel\ level) = gamma_{encoding} \times \log(exposure)$
2. **Linearize the image** by applying the inverse of the measured encoding gamma curve or using the edge itself to obtain an approximate linearization if the chart contrast is known.

3. **Find the center of the transition** between the light and dark regions for each horizontal scan line, $y_l(x)$.
4. Fit a polynomial curve to the center locations.
5. Depending on the location of the curve on the scan line, add each appropriately shifted scan line to one of four bins.
6. **Combine** the mean signal in each bin to obtain the 4× oversampled averaged edge for L scan lines, $\mu_s(x)$, illustrated in the upper plot of Figure 1.

$$\mu_s(\mathrm{x}) = \frac{1}{L}\sum_{l=0}^{L-1} y_l(x-\delta) \tag{2}$$

7. **Modulation Transfer Function, *MTF(f),*** is calculated by differentiating the averaged edge, windowing it, then taking the magnitude of the Fourier transform, normalized to 1 (100%) at zero frequency. *MTF(f)* is illustrated in the lower plot of Figure 1.

The overlooked capability of the ISO 12233 e-SFR algorithm

A simple addition to the ISO 12233 binning algorithm described above allows the ***variance*** of the signal, σ_s^2 (the noise power), to be calculated in addition to the mean, μ_s.

In addition to $\sum y_l(x)$, **calculate the sum of the squares of each scan line**, $\sum y_l^2(x)$. Then,

$$\sigma_s^2(\mathrm{x}) = \frac{1}{L}\sum_{l=0}^{L-1}(y_l(x)-\mu_s(x))^2 = \frac{1}{L}\sum_{l=0}^{L-1} y_l^2(\mathrm{x}) - \left(\frac{1}{L}\sum_{l=0}^{L-1} y_l(\mathrm{x})\right)^2 \tag{3}$$

$\sigma_s^2(x)$ and $\sigma_s(x)$ are the noise power, $N(x)$, and noise voltage, $\sqrt{N(x)}$, *not the noise itself,* at each position on the oversampled array— including the edge transition, where noise was traditionally difficult to measure.

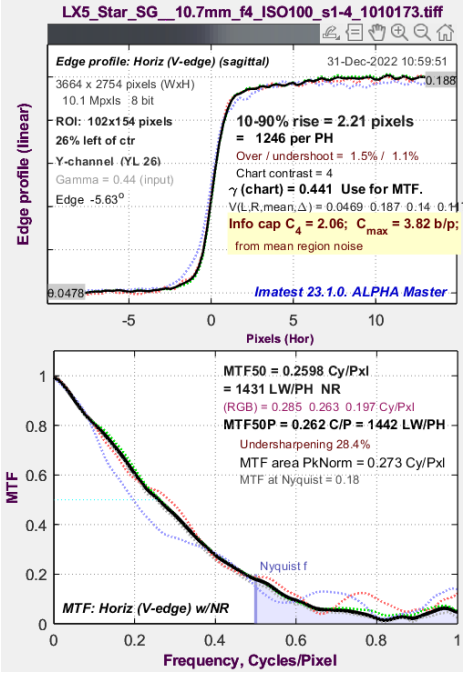


***Figure 1**. Edge and MTF plot for compact digital camera for Unsharpened TIFF from raw. Upper: Mean edge μ_s ; Lower: MTF(f). **C_4** is the Shannon information capacity for a 4:1 contrast ratio edge.*

Noise measured in the presence of a signal can be more accurate than noise measured in flat areas, and can be used in the Shannon-Hartley equation for channel capacity, C.

Binning noise

$\sigma_s^2(x)$ and $\sigma_s(x)$ must be corrected for binning noise— a recently-discovered artifact of the ISO 12233 binning algorithm, which has identical statistics to quantization noise. It is largest near the image transition— where the Line Spread Function, $LSF(x) = d\mu_s(x)/dx$ (Figure 3) is maximum, and it can affect information capacity measurements. It appears because the individual scan lines are added to one of four bins, based on a polynomial fit to the center locations of the scan lines.

Assume that n identical signals, $\mu_s(x_k)$, are binned over an interval [-Δ/2, Δ/2], where Δ = 1 in the 4× oversampled output of the binning algorithm (noting that Δ = (original pixel spacing)/4). The values of $\mu_s(x_k)$ are summed at uniformly-distributed locations x_k over the interval Δ, so they take on values

$$\mu_k = \mu_s(x_k) = \mu_s(x_0 + \delta) = \mu_s(x_0) + \delta\frac{d\mu(x)}{dx} = \mu_s(x_0) + \delta\ LSF(x) \tag{4}$$

Noting that δ is uniformly distributed over [-½, ½], and applying the equation for quantization noise,

$$\sigma_{Bnoise}^2 = LSF^2\sigma_{Qnoise}^2 = \frac{LSF^2}{12} \quad \text{or} \quad \sigma_{Bnoise} = LSF/\sqrt{12} \tag{5}$$

Although this equation is approximate, we have had good success calculating the corrected noise,

$$\sigma_s^2(\text{corrected}) = \sigma_s^2 - \sigma_{Bnoise}^2. \tag{6}$$

Results are shown in Figure 2. Binning noise has no effect on standard MTF calculations.

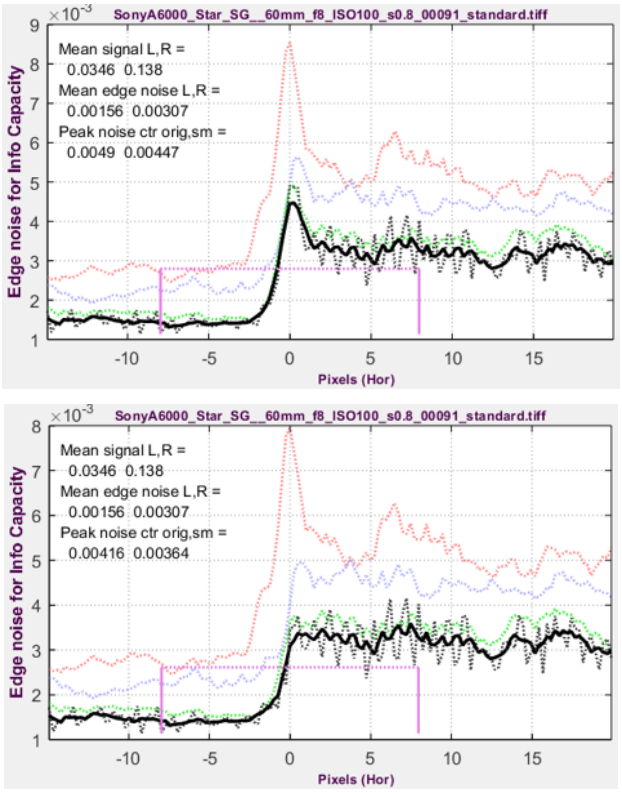


***Figure 2**. Edge noise for a Micro Four-Thirds digital camera, Y-channel from raw image converted to TIFF with minimal processing (raw→TIFF). Top: with binning noise (σ_s^2(uncorrected)) Bottom: binning noise removed (σ_s^2(corrected)) The **bold black** curve is the smoothed Y-channel.*

Signal power S

The peak-to-peak signal voltage, $V_{p\text{-}p}$, at low spatial frequencies is the measured difference between the means of the light and dark regions of the linearized slanted edge, $\mu_s(x)$.

$$V_{p-p} = \Delta\mu_s = \mu_{sLight} - \mu_{sDark} \tag{7}$$

If we assume a uniform distribution between the limits μ_{sLight} and μ_{sDark}, which maximizes information capacity, the

variance, which is the average signal power at low spatial frequencies, is

$$\sigma_V^2 = S_{avg}(0) = (\mu\_sLight - \mu\_sDark\ )^2/12 = V_{p-p}^2/12 \tag{8}$$

The Shannon-Hartley equation uses the *average* frequency-dependent signal power, $S_{avg}(f)$.

$$S_{avg}(f) = \left(V_{p-p}\ \times MTF(f)\right)^2/12 \tag{9}$$

Signal power, S, is proportional to the square of the chart contrast if the image has been properly linearized, which is easy to accomplish if the camera is operating in its log-linear region, i.e., is not approaching saturation at low or high pixel levels. $S_{max} \leq 1$ for linearized images normalized to 1.

Noise power, N

Noise power, N, has the same units as signal power, S; hence S/N is dimensionless.

Noise near the edge transition— not noise measured in flat patches— dominates system performance. The calculation of N depends on the detected image type. Two distinct image types cover most cases of interest.

(1) **Uniformly or minimally-processed images**, often TIFFs converted from raw files (raw→TIFF). Most cameras intended for Machine Vision/Artificial Intelligence are in this category.
Since noise can be a very rough function of x (Figure 4), a moderately large region size is used for measuring N.
We average noise over a region defined as the edge center $\pm$ 1.5×*PW20*, where *PW20* is the width of the region where the Line Spread function $LSF(x) = d\mu_s(x)/dx \geq 0.20\ LSF_{max}$.

$$N_{method-1} = \text{mean}\left(\sigma_s^2(x)\right)$$

$$\text{for}\quad x = \text{edge center} \pm 1.5 \times PW20. \tag{10}$$

(2) **Bilateral-filtered images** [8] include most JPEG images from consumer cameras. Bilateral filters sharpen images near contrasty features such as edges, but blur them (to reduce noise) elsewhere. This causes a noise peak close to the edge transition, which can dominate camera performance. We have long known about the noise peak, but previously had no convenient way to measure or detect it.
Noise power, $N_{method-2}$, is the noise at the peak, smoothed slightly (with a rectangular kernel of length $PW20/2$) to remove jaggedness. This is a somewhat arbitrary choice, but it produces reasonably consistent results. Method (2) also works with minimally-processed images, but results are less consistent than method (1).

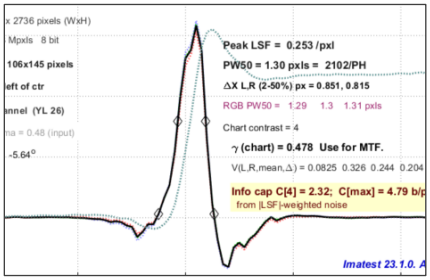
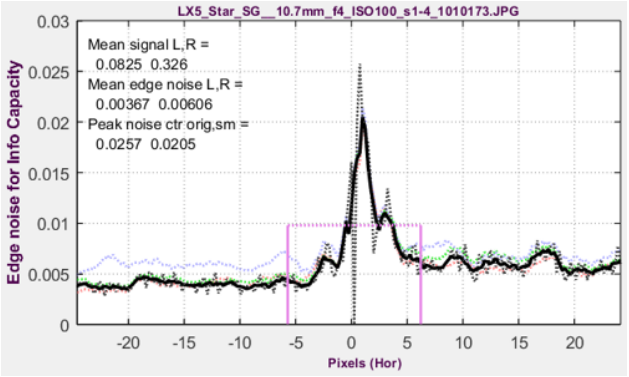


***Figure 3**. Line Spread Function (LSF) for a sharpened JPEG. The x-axis is distance in pixels.*

The noise calculation method may be selected manually or automatically, based on whether or not a peak is detected near the transition. Some additional considerations:

- Noise is not exactly white, but is close enough to yield good results. This assumption is strongly supported by experimental results in [9].
- Noise power is larger on the lighter side of the edge due to photon shot noise, which increases with the number of photons reaching the sensor pixels. The mean includes both sides.
- More generally, noise power increases with exposure. For linear sensors it follows the function $N(V) = k_0 + k_1 V$, where k_1 is the coefficient for photon shot noise.

Figure 4 illustrates the noise voltage $\sigma_s(x) = \sqrt{N(x)}$ for a bilateral-filtered JPEG file (top) and for minimally-processed TIFF files (bottom). The JPEG has a large, distinct peak not present with the TIFF. The solid dark curves are for the luminance (Y) channel smoothed with a 5 pixel-wide rectangular function (1.25 pixels before 4× oversampling) to improve plot appearance.



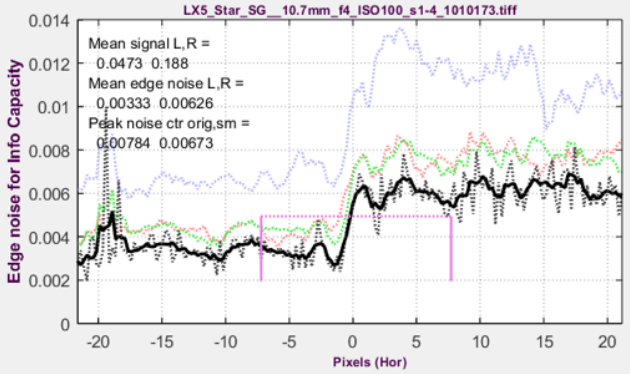


Figure 4. *Edge noise voltage @ ISO 100. The* ***bold black*** *curve is the smoothed Y-channel.*
Top*: Bilateral-filtered in-camera JPEG;*
Bottom*: Unsharpened TIFF from raw.*
The x-axis is the original pixel location of the 4× oversampled signal. Note that the spike around x = -19 of the lower plot is a noise outlier likely caused by a speck of dust on the chart.

Bandwidth, W

Bandwidth, W, is 0.5 cycles/pixel (the Nyquist frequency). Signals above Nyquist do not contribute to the information content; they can actually reduce it by causing aliasing— spurious low frequency signals like Moiré that can interfere with the actual image. Frequency dependence comes from MTF(f), which is a component of $S_{avg}(f)$.

Combining S, N, and W to obtain information capacity, C

$$C = \int_0^{0.5} \log_2\left(1 + \frac{S_{avg}(f)}{N}\right) df \cong \sum_{i=0}^{0.5/\Delta f} \log_2\left(1 + \frac{S_{avg}(i\Delta f)}{N}\right) \Delta f \quad (11)$$

MTF(f) can take a large bite out of C, especially since it is squared in the above equation. Since it has a strong frequency-dependence, it is sometimes confused with bandwidth. For the raw-converted image in Figure 1, bottom, it drops to zero around 0.6 cycles/pixel — typical of a well-focused high quality camera with no sharpening that makes good use of the sensor pixels. Since it is a nearly straight line, $MTF(f) \cong 1 - f/0.6$ for $f \leq 0.6$. The integral of $MTF^2(f)$ for $0 \leq f \leq 0.5$ is approximately 0.204: a significant loss from the value of 0.5 for a perfect (no rolloff) response.

Would increasing MTF help? The relationship between MTF and signal spread (or extent) is explored for diffraction-limited systems in [10] and summarized on an ***Imatest*** web page, Diffraction, Optimum Aperture, and Defocus [11]. If all the energy of a point of light were inside one pixel, there would be little MTF loss. This corresponds to MTF = 0.69 at the Nyquist frequency (0.5 C/P), dropping to 0.4 at twice Nyquist (1.0 C/P). Such a system would have extreme aliasing: low frequency artifacts such as Moiré that degrade its performance. The camera in Figure 1 has only a little energy above Nyquist, so aliasing is reasonably well-controlled. But the pulse is spread over two pixels, with the 10-90% rise distance = 2.21 pixels in Figure 1, leading to a significant loss. This appears to be an unavoidable tradeoff.

Technique

Test chart edge contrast should be between 2:1 and 10:1, with 4:1 (specified in the ISO 12233 e-SFR standard) recommended. Edge contrast greater than 10:1 increases the likelihood of nonlinear operation (saturation or clipping), which compromises results.

Images should be well-exposed because saturation or clipping (both deviations from log-linearity) can give misleading good results.

The camera should be well-focused (unless you're testing misfocus). Sturdy camera support should be employed.

Although results are relatively insensitive to ROI selection, some care must be taken to obtain good consistency. ROIs should be reasonably large; at least 30x60 pixels is recommended. The edge should be centered in the selected region, and there should a reasonable amount of "breathing room" on the sides. A good initial "rule of thumb": The ratio of light to dark space at the ends of the edges should be no larger than 35/65.

Additional assumptions

A key assumption is that the camera's dynamic range (the range of tones that can be reproduced with good contrast and Signal-to-Noise Ratio (SNR)) is sufficient for the intended task. Most modern image sensors have dynamic ranges greater than 60dB (1000:1); high dynamic range (HDR) sensors have 120 dB or more. The majority of scenes in pictorial, medical, or robotic imaging have tonal ranges under 60 dB. Lens flare (stray light) typically limits practical camera dynamic range to 100dB or less, which can impact automotive night driving by fogging the important dark to middle tones. If there are concerns about dynamic range, we strongly recommend measuring it with a transmissive chart.

Other assumptions: sensor nonuniformities (fixed-pattern noise, also called PRNU (Photo Response Nonuniformity) are included in noise measurements. Tonal response is well-behaved (typically following a gamma curve, except for the extreme highlights and shadows).

Because the value of C is closely tied to the n:1 chart contrast ratio, where $n \leq 10$ to minimize saturation or clipping, we specify n when C is reported, e.g., C_4 for charts with a 4:1 contrast ratio.

Sensitivity to exposure

4:1 edges may appear to have relatively low contrast, but they can occupy a substantial portion of the available linearized and normalized signal voltage, V, where $0 \leq V \leq 1$. The portion is strongly dependent on exposure. For a standard "good" exposure, where $V_{mean} \approx 0.20$, the voltage from the 4:1 edge occupies 24% of the total range. However, it can occupy as much as 75%, as shown in Table 1.

V_{mean}	V_{min} ($0.4\ V_{mean}$)	V_{max} $1.6\ V_{mean}$)	Range = ΔV = $V_{p\text{-}p}$
0.12	0.048	0.192	0.144
0.20	0.08	0.32	0.24
0.40	0.16	0.64	0.48
0.60	0.24	0.96	0.72

Table 1: V_{mean}, V_{min}, $V_{P\text{-}P}$, and normalized signal voltage range $\Delta V = V_{max}\text{-}V_{min} = V_{p\text{-}p}$ for 4:1 contrast ratio edges.

Because both noise power, N, and voltage range, ΔV, increase with exposure, C_4 **is a strong function of exposure.**

Consistent exposure can be difficult to maintain with autoexposure consumer cameras because their JPEG output files often have "shoulders" in their tonal response (regions of reduced highlight contrast intended to improve pictorial quality by minimizing saturated ("burnt out") highlights).

Implementing a shoulder requires extra headroom, i.e., a degree of underexposure, which can vary for different camera models. Since autoexposure is optimized for JPEG output, minimally processed files, typically TIFFs converted from raw with simple gamma curves, often appear to be underexposed.

Maximum information capacity C_{max} — a more stable metric

Because the strong exposure-dependence of C_4 reduces its value as a performance metric, we have developed a new metric for maximum information capacity, C_{max}, that is nearly independent of exposure. It is obtained in two steps.

Step1: Replace the measured peak-to-peak voltage range, $V_{p\text{-}p}$, with the maximum allowable value, $V_{p-p\_max} = 1$. This may seem like a simplification, but it works well for most cameras. Referring to the section on Noise Power, N,

$$S_{avg}(f) = \left(V_{p-p\_max}\, MTF(f)\right)^2/12 = MTF(f)^2/12 \tag{12}$$

Step 2: Replace the measured noise power, N, with N_{mean}, the mean of N over the range $0 \le V \le 1$ (where 1 is the maximum allowable normalized signal voltage V). The general equation for N for linear image sensors is

$$N(V) = k_0 + k_1 V \tag{13}$$

k_0 is the coefficient for constant noise (dark current noise, Johnson (electronic) noise, etc.). k_1 is the coefficient for photon shot noise. Noise powers $N_1 = \sigma_1^2$ and $N_2 = \sigma_2^2$ are measured along with signal voltages V_1 and V_2 on either side of the edge transition.

Assuming $N_1 = k_0 + k_1 V_1$ and $N_2 = k_0 + k_1 V_2$, we can solve two equations in two unknowns to obtain k_0 and k_1.

$$k_0 = \frac{N_1 V_2 - N_2 V_1}{V_2 - V_1}\ ; \quad k_1 = \frac{N_2 - N_1}{V_2 - V_1} \tag{14}$$

For bilateral-filtered images (most JPEGs from consumer cameras), where N_1 and N_2 are reduced by lowpass filtering, we make an approximate adjustment to N to compensate for the filtering.

$$N \rightarrow k_N N, \text{ where } k_N = N_{method\_2}/N_{method\_1} \tag{15}$$

C_{max} as less accurate for bilateral-filtered images than for uniformly-processed images.

The mean noise power N_{mean} over the range $0 \le V \le 1$ for calculating C_{max} is

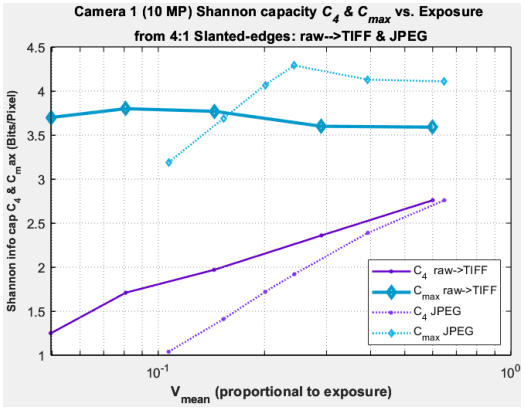
$$N_{mean} = \int_0^1 N(V)\, dv \Big/ \int_0^1 dv = \int_0^1 (k_0 + k_1 V) dv$$
$$= k_0 + k_1/2 \tag{16}$$

Using $N = N_{mean}$, $V_{p-p\_max} = 1$, and $S_{avg}(f) = MTF(f)^2/12$, Equation (11) becomes

$$C_{max} = \int_0^{0.5} \log_2\left(1 + \frac{MTF(f)^2}{12\ N_{mean}}\right) \mathrm{df}$$
$$\cong \sum_{i=0}^{0.5/\Delta f} \log_2\left(1 + \frac{MTF(i\Delta f)^2}{12\ N_{mean}}\right) \Delta f \tag{17}$$

C_{max} (Figure 5) is nearly independent of exposure for minimally or uniformly-processed images with linear sensors, where noise power, N, is a known function of signal voltage, V.

C_{max} is approximate for imaging systems with bilateral filtering or HDR (nonlinear) sensors, where noise power N is not a simple function of V (Equation (13)).



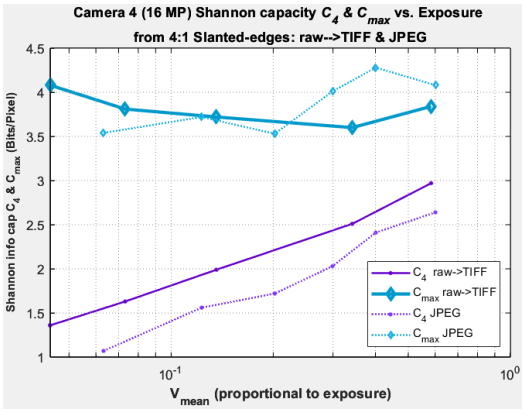


Figure 5. *C_4 and C_{max} for minimally processed raw→TIFF and JPEG images for two cameras: Top. 10 MP compact, Bottom. 16 MP micro 4/3.*
C_{max} is consistent, especially for the raw→TIFF image, except for the lowest (severely underexposed) exposures.

Information capacity results

Table 2 shows three cameras that produced both raw and JPEG output that we tested for information capacity as a function of Exposure Index (ISO speed setting).

1.	Panasonic Lumix LX5	2.14 µm pixel pitch. Compact 10.1-megapixel camera with a Leica f/2 zoom lens set to f/4.
2.	Sony A6000	3.88 µm pixel pitch. 24-megapixel micro four-thirds camera
3.	Sony A7Rii	4.5 µm pixel pitch. A 42-megapixel full-frame camera with a Backside-Illuminated (BSI) sensor

Table 2. Cameras used in the tests

The image in Figure 6, which was analyzed in "Measuring camera Shannon Information Capacity with a Siemens Star Image" [12], contains a 50:1 contrast Siemens star and four 4:1 contrast slanted edges. We used the upper-left slanted edge for most tests. The average background of the chart should be close to neutral gray (18% reflectance) to ensure a good exposure.

Figure 6. *Typical image (cropped) including Siemens star and slanted-edges to the left and right of the star.*

We captured both JPEG images and raw images, converted by LibRaw to 24-bit sRGB TIFF (designated as raw→TIFF) with minimal processing (no sharpening, no noise reduction, and simple gamma-encoding). The luminance channel (Y = 0.2125·R + 0.7154·G + 0.0721·B) was analyzed. Results with 48-bit Adobe RGB conversion were similar.

Results for JPEG and minimally-processed raw→TIFF images

Figures 7-9 show results for the as a function of ISO speed (Exposure Index, which is proportional to analog gain and should not be confused with exposure) for the raw→TIFF images (solid lines) and JPEG images (dotted lines). For the raw→TIFF images, the relationship between ISO speed and C is similar for all three cameras.

$N_{method\text{-}1}$ is used for the raw→TIFF images; $N_{method\text{-}2}$ is used for the bilateral-filtered JPEGs.

C_4: 4:1 slanted edge

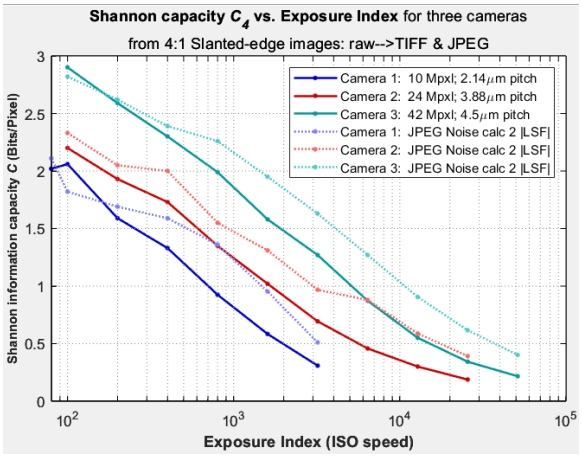


Figure 7*. Information capacity, C_4, from 4:1 slanted-edge images. Solid lines for raw→TIFF images; Dotted lines for JPEGs.*

The information capacity for 4:1 contrast edges, C_4, shows similar trends to C_{max} and C_{star} (shown below), but since the relatively low 4:1 contrast uses only a fraction of the available signal level, C_4 is lower than either measurement.

C_{max}: maximum information capacity

C_{max} is the maximum information capacity of the camera, derived from measurements of 4:1 edges. It is relatively accurate for minimally or uniformly-processed (often raw→TIFF) images, and is much less sensitive to exposure than C_4. C_{max} is a robust measurement, well-suited for comparing the performance of different cameras.

Comparisons of information capacity between different cameras are similar regardless of the measurement method.

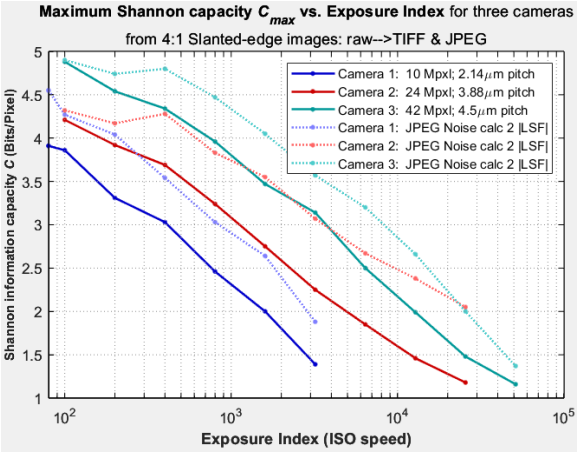


Figure 8. *Information capacity, C_{max}, from slanted-edge results. Solid lines are for TIFFs derived from raw images; dotted lines for JPEGs.*

C_{star}: Siemens star (50:1 contrast)

Information capacity, C_{star}, for the star is generally higher than C_4, but only slightly lower than C_{max} for slanted edges because the star images don't use the entire available tonal range.

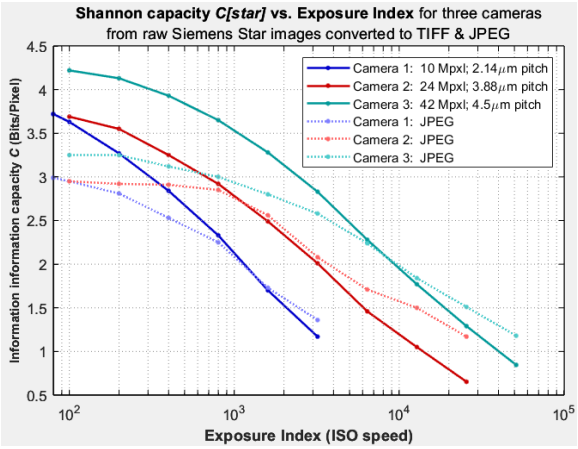


Figure 9. *Information capacity, C_{star}, from Siemens star images*

Color channels

The separate R, G, and B channels tend to have slightly lower C_4 than the Y-channel because the uncorrelated noise from the separate channels are combined in the Y-channel. Example: for Camera 2 (24 Megapixels, Micro Four-Thirds) at ISO 400, $C_{4Y} = 1.96$, $C_{4R} = 1.16$, $C_{4G} = 1.81$. and $C_{4B} = 1.36$ bits/pixel. The noise for each channel is shown in Figure 10 from best (lowest noise, highest C) to worst: Y, G, B, and R. The green channel has the best SNR because the image sensor is most sensitive to green, and hence the green channel had the least boost in the white balance process.

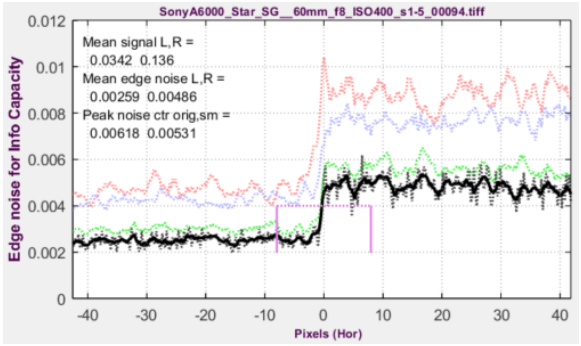


Figure 10*. Noise for camera 2 @ ISO 400, showing the different color channels*

$C_R + C_G + C_B$ is nearly triple C_Y. But this is to be expected because the three color channels occupy 24 bits instead of 8 (for a single channel).

Even though we've focused on demosaiced images, the slanted-edge method can be applied to raw (undemosaiced) images. For this camera, $C_{4Ru} = 2.09$, $C_{4GRu} = 2.43$, $C_{4Bu} = 1.6$. and $C_{4GBu} = 2.47$ bits/pixel, where each undemosaiced (u) channel has a quarter as many pixels as the demosaiced channels. We haven't worked on interpreting these results.

Effects of sharpening

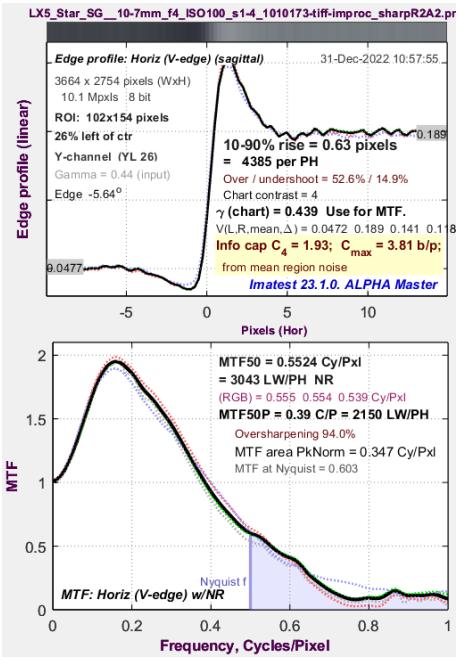


Figure 11*. Edge/MTF plots derived from the same image as Figure 1, where C_4 = 2.06 b/p and C_{max} = 3.82 b/p, raw→TIFF, ISO 100 Sharpening Radius = 2; Amount = 2. C_4 = 1.93 b/p; C_{max} = 3.81 b/p.*

The example in Figure 11 (one of several we ran) shows that sharpening has little effect on slanted-edge information capacity, as expected for a valid measurement. The image (initially a minimally-processed TIFF) has been strongly Unsharp Mask sharpened with Radius = 2 and Amount = 2. It can be compared to Figure 1, where $C_4 = 2.06$ and $C_{max} = 3.82$ b/p.

We observed a similar insensitivity of C to sharpening with Siemens stars.

The insensitivity of C to sharpening is an indicator of the calculation's validity.

Total information capacity

We have focused on information capacity, C, in bits per pixel. The total information capacity, C_{total}, for the entire image must take variations in C over the image into account. In *Imatest*, the mean value of C for the image can be displayed in the 3D plots for multi-region slanted-edge modules (Figure 12). This mean is always unweighted for information capacity displays.

$$C_{total} = \text{mean}(C) \times \text{megapixels} \tag{18}$$

The mean information capacity, C_{max}, is 2.847 bits/pixel. Since the camera has 16 Megapixels, total capacity, $C_{maxTotal}$,for the Luminance (Y) channel = 45.44 MB.

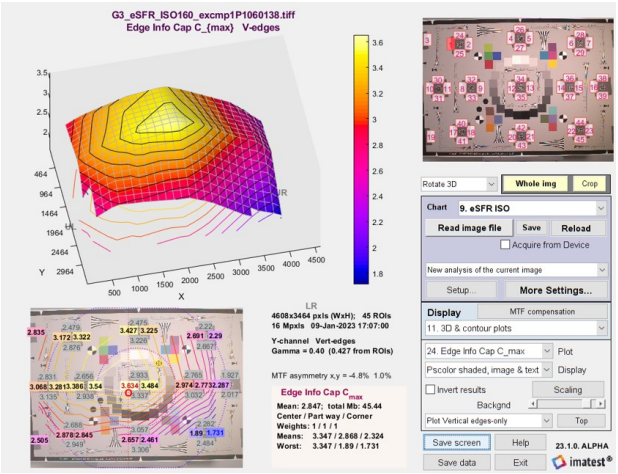


***Figure 12**. 3D contour eSFR ISO plot of Cmax for the Luminance (Y) channel, ISO 100*

Comparisons of the slanted-edge and Siemens star methods

Slanted-edge method

- Any slanted-edge test image with printed contrast ≤ 10:1 (4:1 recommended) can be used to obtain C. Most older images can be analyzed.
- For multi-region images, C can be mapped over the entire image to find total information capacity.
- For bilateral-filtered images (most in-camera JPEGs), results are useful, but less accurate than for minimally or uniformly processed images.
- C does poorly for measuring the effects of artifacts. Clipped images may show improved performance due to sharp corners on the edge transition and reduced noise in clipped areas.

Siemens star method

- Sensitive to optical distortion. Best if the star is in the center of the image.
- Computationally slower than slanted edges.
- Appropriate response to bilateral filtering and image processing artifacts: enables comparison of demosaicing techniques, image compression, aliasing, etc. C is correctly reduced for clipped images.

Both methods assume that the camera dynamic range is sufficient for the intended task, which may not hold for automotive night driving, where it may be limited by stray light.

When comparing cameras, the same measurement method (chart type, contrast, etc.) should be used.

Future work

- Collaborate with partners in industry and academia to correlate information capacity, C, with performance of Machine Vision and Artificial Intelligence systems.
- Work to include camera information capacity in several standards, especially ISO TC42.
- Explore the correlation between C with the subjective visual appearance of a variety of images, without and with additional image processing.
- Develop predictions of the effects of image processing (sharpening, noise reduction) on MV/AI performance. Our initial approach will be to measure SNRI, based on the work of Paul Kane and collaborators [4], [9], [13].

Summary

We have focused on the newest of two methods for measuring camera information capacity— the slanted edge method, which is fast, convenient, and requires no special effort. The key concepts are

1. Information capacity, which combines sharpness, noise, contrast loss, and (for the Siemens star) the effects of several types of artifact, is a fundamental figure merit for imaging systems.
2. Spatially varying noise power $N(x)$ can be extracted from slanted-edge regions.
3. The noise peak in bilateral-filtered images allows them to be distinguished from uniformly-processed images, so that the optimum noise calculation can be selected.
4. Information capacity, C_n, measured from n:1 contrast slanted edges (typically 4:1), is sensitive to chart contrast and exposure, but it can be extended to calculate a maximum information capacity, C_{max}, that is insensitive to these factors.

The Siemens Star and Slanted Edge methods have similar *relative* trends (for comparing cameras).

Camera information capacity is still a novel measurement. Significant effort will be required to make it better known. But the units— information bits per pixel (or total image) for a specified ISO speed— are intuitive and easy to understand.

We would like to see information capacity at specified ISO speeds (exposure indices) or light (lux) levels— become an integral part of a standard camera specifications, especially for

machine vision applications. We are optimistic that this can lead to improved performance and reduced energy use [14].

References

[1] C. E. Shannon, “A mathematical theory of communication,” Bell Syst. Tech. J., vol. 27, pp. 379–423, July 1948; vol. 27, pp. 623–656, Oct. 1948.

[2] C. Shannon, “Communication in the Presence of Noise”, Proceedings of the I.R.E., January 1949, pp. 10-21.

[3] R. Shaw, “The Application of Fourier Techniques and Information Theory to the Assessment of Photographic Image Quality,” Photographic Science and Engineering, Vol. 6, No. 5, Sept.-Oct. 1962, pp.281-286. Available for download.

[4] Robin Jenkin, Paul Kane, "Fundamental Imaging System Analysis for Autonomous Vehicles", *Proc. IS&T Int'l. Symp. on Electronic Imaging: Autonomous Vehicles and Machines*, 2018, pp 105-1 – 105-10, https://doi.org/10.2352/ISSN.2470-1173.2018.17.AVM-105

[5] Frédéric Cao, Frédéric Guichard, Hervé Hornung, “Information capacity: a measure of potential image quality of a digital camera”, DxO Labs, 3 rue Nationale, 92100 Boulogne Billancourt, FRANCE, 2010

[6] F.-X. Thomas, T. Corbier, Y. Li, E. Baudin, L. Chanas, F. Guichard, "RAW Image Quality Evaluation Using Information Capacity", Proc. IS&T Int'l. Symp. on Electronic Imaging: Image Quality and System Performance XVIII, 2021, pp 218-1 - 218-7, https://doi.org/10.2352/ISSN.2470-1173.2021.9.IQSP-218

[7] ISO 12233:2017 Photography – Electronic still picture imaging – Resolution and spatial frequency response [Online], www.iso.org/standard/71696.html.

[8] 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.

[9] Orit Skorka, Paul J. Kane, "Object Detection Using an Ideal Observer Model", *Proc. IS&T Int'l. Symp. on Electronic Imaging: Autonomous Vehicles and Machines*, 2020, pp 41-1 - 41-7, https://doi.org/10.2352/ISSN.2470-1173.2020.16.AVM-041

[10] Robert D. Feite, “Modeling the Imaging Chain of Digital Cameras”, SPIE Digital Library, 2010, Chapter 8, “The Story of Q”.

[11] N. Koren, “Diffraction, Optimum Aperture, and Defocus”, on internet, https://www.imatest.com/docs/diffraction-and-optimum-aperture/

[12] Norman L. Koren, "Measuring camera Shannon Information Capacity with a Siemens Star Image", *Proc. IS&T Int'l. Symp. on Electronic Imaging: Image Quality and System Performance XVII*, 2020, pp 347-1 - 347-10, https://doi.org/10.2352/ISSN.2470-1173.2020.9.IQSP-347

[13] Paul J. Kane, "Signal detection theory and automotive imaging", *Proc. IS&T Int'l. Symp. on Electronic Imaging: Autonomous Vehicles and Machines Conference*, 2019, pp 27-1 - 27-8, https://doi.org/10.2352/ISSN.2470-1173.2019.15.AVM-027

[14] Soumya Sudhakar, “Data centers on wheels: the carbon footprint of self-driving cars”, TEDxBoston

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.