

# Measuring Information Capacity with Imatest

Norman Koren

May 2023

revised May 29, 2023

Information capacity is a key performance indicator (KPI) for a large variety of imaging systems, including

- Machine vision
- Automotive (Driver assistance and autonomous vehicles)
- Artificial Intelligence
- Robotics
- Security

It is important because Machine vision and AI (MV/AI) algorithms operate on *information*, not *pixels*, making information capacity is a far better predictor of system performance than either sharpness (MTF) or noise.

**Imatest** has developed a highly convenient method for measuring information capacity and related KPIs from the most widely used ISO standard resolution test pattern — the slanted edge.

## Introduction

Traditional image quality measurements are based on several image quality factors, the best-known of which are sharpness, noise, dynamic range, optical distortion, tonal and color response, and spatial uniformity.

These measurements have proven useful for human vision, where tradeoffs are often required, for example, sharpening an image may increase noise, but makes fine features more visible to the human eye. Choices are often based on experience; they come down to what looks best, i.e., what has the most pleasing appearance for the application.

Machine Vision/Artificial Intelligence (MV/AI) systems are different. System performance is not dependent on image appearance. A more objective metric is required.

## Information

The information is a metric of how much is learned from a measurement. For example, an individual pixel in a blurred image is highly correlated with its neighbors, so little can be learned from its contents. But if the image is sharp, it is weakly correlated, and much more can be learned from individual pixels.

The concept of information was quantified in 1948 and 49 in two celebrated papers by [Claude Shannon](#) [1],[2]. We present a concise introduction to basic information concepts in [Appendix I](#), below. Earlier work on measuring information capacity from Siemens Star images [3] will only be touched on in this document.

## The slanted edge

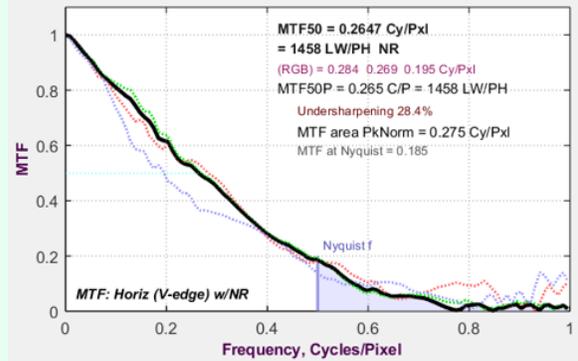
The slanted edge, which is a key part of the ISO 12233 standard, “Photography — Electronic still picture imaging — Resolution and spatial frequency responses” [4], is the most convenient and widely-used resolution test pattern. It is highly efficient in its use of space (with multiple edges, sharpness can be mapped over the image surface), and calculations are very fast.



Information capacity is calculated from an overlooked capability of slanted-edge regions that was quite literally hidden in plain sight. To understand it, we present a brief summary of the standard ISO 12233 Edge SFR (e-SFR) algorithm.

1. **The image should be well-exposed**, avoiding the dark “toe” and light “shoulder” regions.
2. **Linearize the image** by applying the inverse of the encoding gamma curve or using the edge itself if the chart contrast is known.
3. **Find the center of the transition** between the light and dark regions for each horizontal scan line.
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 the scan lines,  $\mu_s(x) = \frac{1}{L} \sum_{l=0}^{L-1} y_l(x - \delta)$ .

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 displayed in the lower plots of the Edge/MTF figure. Example on the right.



## The overlooked capability of the ISO 12233 binning algorithm: the Edge variance method

By adding the sum of the squares of each scan line,  $\sum y_l^2(x)$ , we can calculate the edge variance (noise power)  $\sigma_s^2(x)$  and noise amplitude  $\sigma_s(x)$  in addition to the mean,  $\mu_s$ . This sum is,

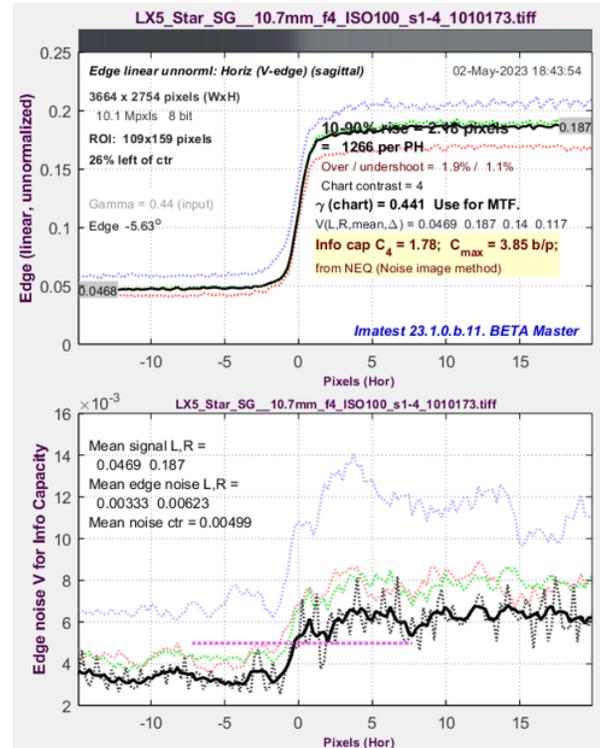
$$\rho_s(x) = \frac{1}{L} \sum_{l=0}^{L-1} y_l^2(x - \delta)$$

This allows edge variance  $\sigma_s^2(x)$  and noise amplitude  $\sigma_s(x)$  to be calculated from  $\sum y_l(x)$  and  $\sum y_l^2(x)$ .

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

The figure on the right shows

- Upper plot: the average 4x oversampled edge,  $\mu_s(x)$ . The thick black line is the luminance channel.
- Lower plot: the noise amplitude (voltage),  $\sigma_s(x)$ . The thick black line is the smoothed luminance channel. The plot of  $\sigma_s(x)$  is new: spatially-dependent noise was previously difficult to observe.



### Calculating information capacity from $\mu_s(x)$ and $\sigma_s(x)$

The next step is to calculate the information capacity,  $C$ , typically in units of bits per pixel, from the signal and noise power.  $C$  is calculated by substituting the correct values of signal and noise power into the Shannon Hartley equation.

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

Where  $S(f)$  and  $N(f)$  are frequency-dependent signal and noise power, and  $W$  is the bandwidth, which is always equal to 0.5 cycles/pixel (the Nyquist frequency). Frequency-dependence is entered into  $S(f)$  using  $MTF(f)$  (described below).

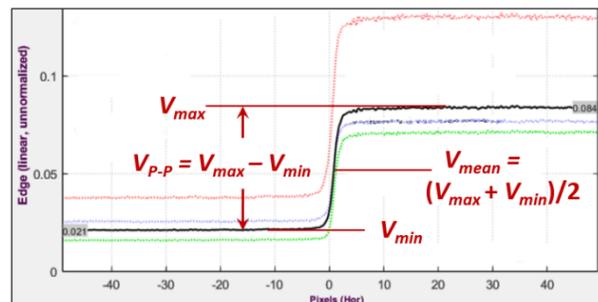
Note that this method, which is called the *edge variance method*, is the first of two methods for calculating  $C$ . The second method, called the *noise image method*, may be slightly more accurate, but it only works for uniformly or minimally-processed images; not for bilateral-filtered images (mostly in-camera JPEGs), to be described below.

### Signal power $S$

The peak-to-peak signal amplitude at low spatial frequencies is the measured difference between the means of the light and dark regions of the linearized slanted edge  $V(x) = \mu_s(x)$ .

$$V_{p-p} = \Delta\mu_s = \mu_{sLight} - \mu_{sDark} = V_{max} - V_{min}$$

The signal power is the *variance* of this signal. If we assume a uniform distribution between the limits  $V_{max}$  and  $V_{min}$ , which maximizes information capacity, we note that the [variance of the uniform distribution](#), which is the average signal power at low spatial frequencies, is



$$\sigma_V^2 = S_{avg}(0) = (V_{max} - V_{min})^2/12 = V_{p-p}^2/12$$

The [Shannon-Hartley equation](#) uses the *average* frequency-dependent signal power,  $S(f)$ .

$$S_{avg}(f) = (V_{p-p} MTF(f))^2 / 12$$

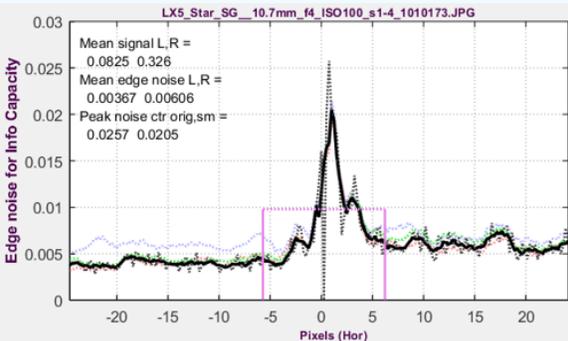
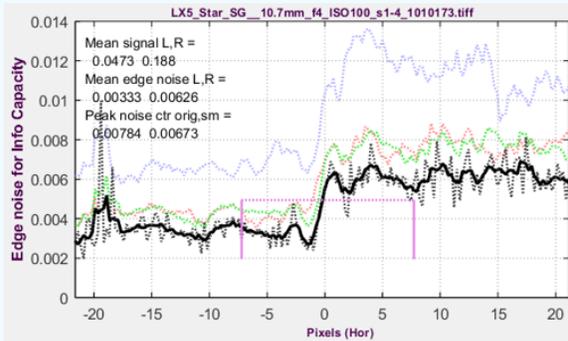
Signal power  $S$  is proportional to the square of the chart contrast if the image has been properly linearized.  $S_{max} \leq 1$  for linearized images normalized to 1.

### Noise power $N$

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

In examining a great many images, we observe two broad classes of images with very different noise properties, visible in  $\sigma_s(x)$ . We call them (1) uniformly/minimally processed and (2) bilateral filtered images. The value of noise power,  $N$ , used to calculate  $C$ , is different for the two image types.

For “black box” cameras with unknown image processing, the table below shows how to distinguish the two image types. If the image processing pipeline is known and understood, the table may not be necessary. For most measurements, uniformly/minimally processed images are preferred.

The two image types: Plots of $\sigma_s(x)$ (4× oversampled)	
<p><b>Bilateral-filtered image:</b> sharpened near the edge; noise-reduced elsewhere. Nearly universal in consumer camera JPEG images. Image processing appears to increase information capacity <math>C</math>, even though information is actually removed. For this reason, it is important to use the peak noise <math>\sigma_s^2(x)</math> (as described below) to calculate <math>C</math>.</p>	<p><b>Minimally (i.e., uniformly) processed image:</b> converted from raw with an external raw converter, with no sharpening or noise reduction.</p>
<p><b>A strong <math>\sigma_s(x)</math> peak is visible</b> near the edge transition. (This peak below is stronger than usual.)</p>	<p><b>Little or no peak is visible</b> in <math>\sigma_s(x)</math>. Noise increases on the right because noise power is proportional to signal power (the mean number of photons striking each pixel) for linear sensors.</p>
<p>For calculating <math>C</math>, <math>N</math> is the square of the peak noise, smoothed with a rectangular kernel of length <math>PW20/2</math>.</p>	<p>For calculating <math>C</math>, <math>N = \text{mean}(\sigma_s^2(x))</math> for all values of <math>x</math> in the ROI.</p>
	

For both images, the solid line is the smoothed noise amplitude, $\sigma_s(x)$	
Avoid for evaluating cameras for MV/AI systems.	Recommended, where available.
<b>Texture:</b> is reduced in low contrast portions of the image. Bilateral filtering is the reason texture is measured with charts such as Spilled Coins and Log F-Contrast: measurements can be very different from slanted edges.	<b>Texture:</b> relatively uniform; affected very little by image contrast. Charts such as Spilled Coins and Log F-Contrast should have MTF similar to the slanted-edge.

Bilateral-filtered images are of interest because we often measure “black box” cameras, where we don’t know whether bilateral filtering is present (it’s important to know), but we want to obtain a reasonable approximation of  $C$ .

Uniformly/minimally-processed images should be used for evaluating cameras for use in MV/AI systems,.

Binning noise is a type of quantization noise that affects the Line Spread Function, but has no effect on conventional MTF measurements. It is described in [Appendix 2, below](#).

### Bandwidth $W$

Bandwidth  $W$  is always 0.5 cycles/pixel (the Nyquist frequency). Signals above Nyquist do not contribute to the information content; they can reduce it by causing aliasing — spurious low frequency signals like Moiré that can interfere with the true image. Frequency-dependence comes from  $MTF(f)$ .

### Combining $S_{avg}(f)$ , $N$ , and $W$ to obtain information capacity $C$

$S_{avg}(f)$ ,  $N$ , and  $W$  are entered into the Shannon-Hartley equation.

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

$MTF(f)$  can take a large bite out of  $C$ , especially since it is squared in the above equation. Because of its frequency-dependence, it is sometimes confused with bandwidth.

$C$  is measured with relatively low contrast test charts to ensure the camera is operating in its linear region and to minimize errors from saturation. For most of our work, we use charts with a 4:1 contrast ratio (Michelson contrast =  $\frac{V_{max}-V_{min}}{V_{max}+V_{min}} = 0.6$ ), following the ISO 12233 standard [4].

Since  $V_{p,p}$  is directly proportional to chart contrast, we label  $C$  according to the contrast ratio:  $C_n$  for n:1 contrast ratio. We use  $C_4$  throughout this document.

By measuring  $C_4$  from a variety of exposures, we quickly learned that (a)  $C_4$  is highly dependent on the exposure level, and (b)  $C_4$  does **not** represent the maximum information capacity of the camera.

### Maximum information capacity $C_{max}$ — a more consistent metric

The strong dependence of  $C_4$  on exposure reduces its value as a performance metric. The reasons for this dependence are (1) voltage range  $\Delta V = V_{p,p}$  is a strong function of exposure, and (2) noise power  $N$  is also a function of exposure.

**We have developed a new metric for maximum information capacity:  $C_{max}$ , that is nearly independent of exposure.** It is obtained in two steps, shown inside a “green for geeks” box below, which can be skipped by most readers.

**Step 1:** Replace the measured peak-to-peak voltage range  $V_{p-p}$  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 [Signal Power  \$S\$](#) ,

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

$$N(V) = k_0 + k_1V$$

$k_0$  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  $N_1 = k_0 + k_1V_1$  and  $N_2 = k_0 + k_1V_2$  we can use two equations in two unknowns to solve for  $k_0$  and  $k_1$ .

$$k_0 = \frac{N_1V_2 - N_2V_1}{V_2 - V_1}; \quad k_1 = \frac{N_2 - N_1}{V_2 - V_1}$$

$N$  closely approximates the noise used in noise calculation method (1) (used for minimally-processed images that don't have bilateral filtering). But if method (2) (the smoothed peak noise) is used (recommended for in-camera JPEGs with bilateral filtering),  $N$  is generally larger, and must be modified.

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

In bilateral-filtered images (most JPEGs from consumer cameras), lowpass filtering (for noise reduction) may affect  $N_1$  and  $N_2$  strongly enough so the equation  $N(V) = k_0 + k_1V$  does not reliably hold. This can adversely affect the accuracy of  $C_{max}$ .

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

$$N_{mean} = \int_0^1 N(V) dv / \int_0^1 dv = \int_0^1 (k_0 + k_1V) dv = k_0 + k_1/2$$

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

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

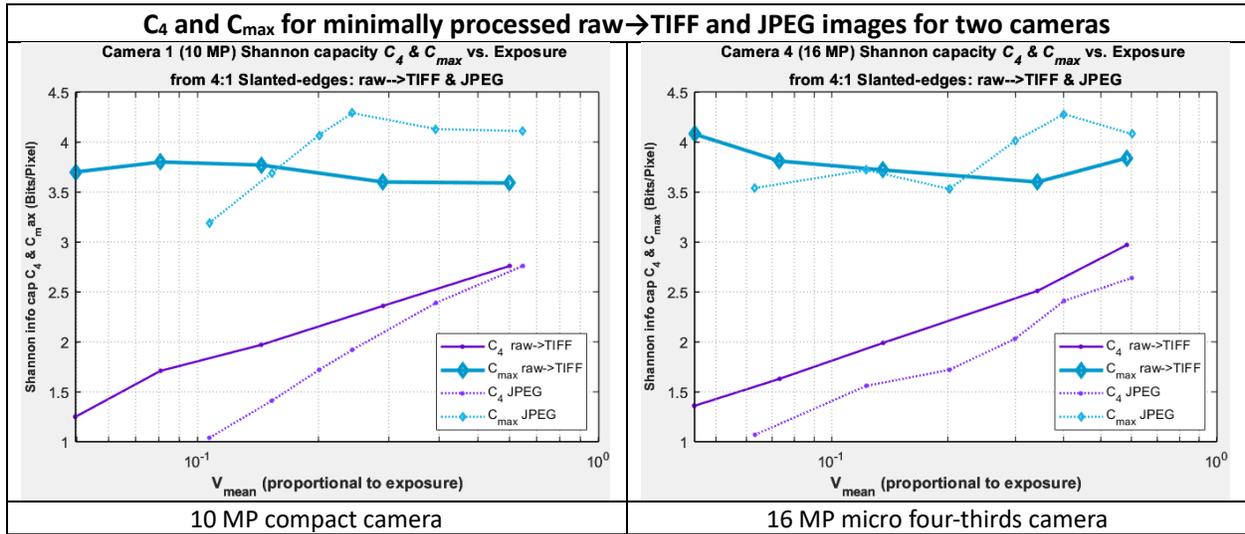
Because noise in High Dynamic Range (HDR) sensors does not follow the simple equation for linear sensors, we recommend giving the image sufficient exposure so the brighter side of the edge is close to (but definitely below) saturation, then leaving  $N$  unchanged ( $N_{mean} = N$ ).

$C_{max}$  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  $V$ .

### Consistency of $C_{max}$

We performed a set of analysis on two cameras with a range of exposures (indicated by  $V_{mean}$ ). The results showed that  $C_{max}$  was highly consistent with exposure for the raw→TIFF images (which were not bilateral-filtered), but less consistent with the bilateral-filtered (JPEG) images.  $C_4$  varied as expected.

Because of the inconsistency, we don't recommend using bilateral-filtered images where accurate information capacity measurements are required.



C<sub>max</sub> may be inaccurate if the image is incapable of spanning the entire range of Digital Numbers (DNs), for example, 0-255 for images with bit depth = 8. None of the information capacity measurements work if local tone mapping is applied.

## Obtaining Results

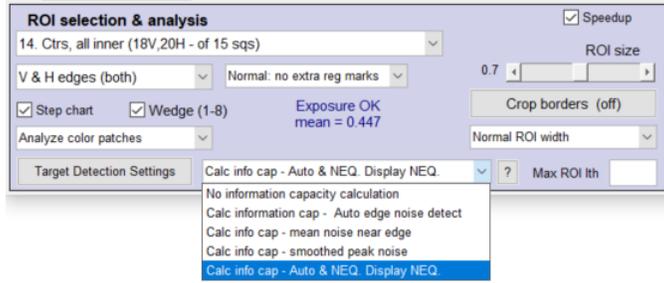
The settings below apply to both the Edge Variance and Noise Image method. As long as the image is known to be minimally or uniformly-processed, the user doesn't need be concerned about the method.

Now that we have described the two major results from the Edge Variance method: Information capacity (C<sub>4</sub> and C<sub>max</sub>,) and spatially-dependent noise (power  $\sigma_s^2(x)$  or amplitude  $\sigma_s(x)$ ), we show how to obtain them. They can be calculated in any Imatest slanted-edge module, including [SFR](#) (manual ROI detection), and [SFRplus](#), [eSFR ISO](#), [Checkerboard](#), or [SFRreg](#) (auto ROI detection). We focus on the settings in the Auto detection modules (the locations for SFR are slightly different).

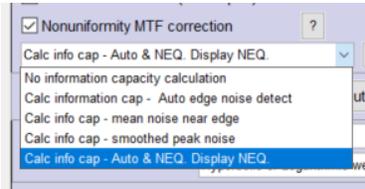
Test chart edge contrast should be between 2:1 and 10:1, with 4:1 (the ISO 12233 e-SFR standard [4]) strongly recommended. General good technique is recommended for acquiring the image: lighting should be uniform and glare-free; the image should be well-exposed; sturdy camera support should be used; ROIs should be reasonably large: at least 30x60 pixels is recommended.

### To turn on the calculations (this works with both methods),

Make the selection in the **Setup** window,

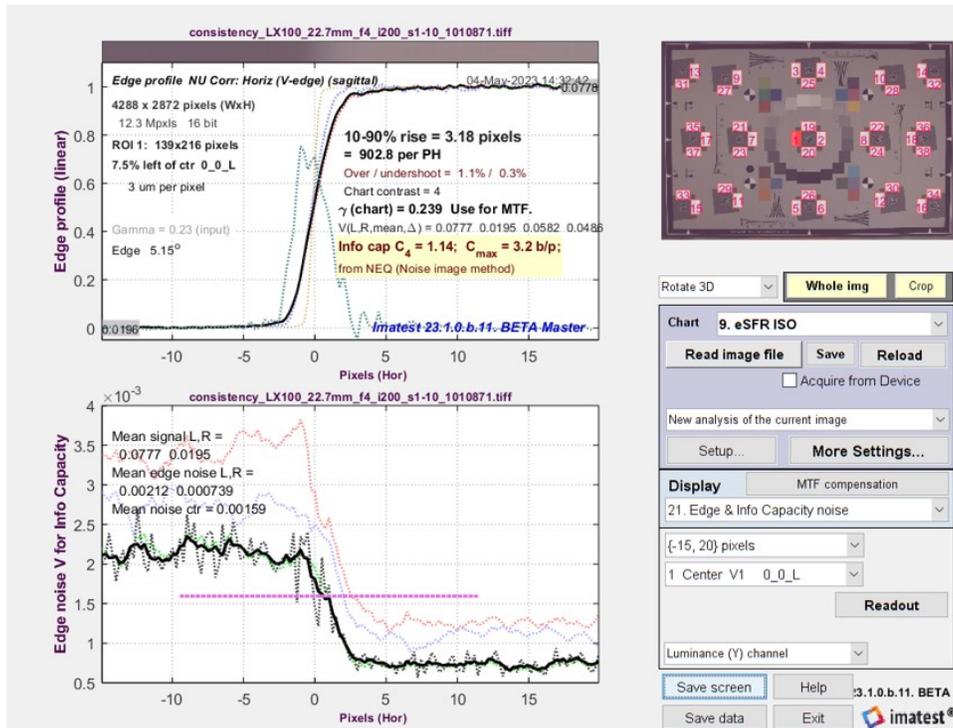


— or — in the **More settings** window.



- The first selection turns off all information capacity calculations. This is the default at the time of the 23.1 release. We may change it.
- The remaining selection determines what gets displayed in the Edge and MTF and Edge & Info capacity noise plots.
- The second selection is reasonable when you don't know whether your image is bilateral-filtered.
- The fifth (last) selection displays the *NEQ* information capacity (described below) in the Edge/MTF figure, which is slightly more accurate than the Edge Variance *C*. It's the best selection for minimally/uniformly-processed images.

For the two plots, Information capacity is displayed next to the Edge (upper) plot.



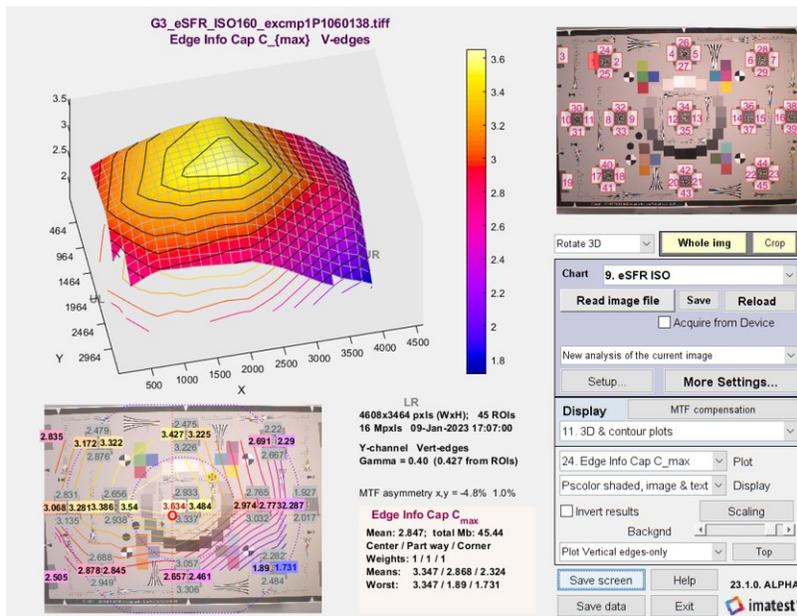
**Edge & information capacity noise plot**

## Total information capacity

The results we have presented thus far are for information capacity  $C$  in bits per pixel. The total information capacity,  $C_{total}$ , for the entire image takes variations in  $C$  over the image into account.

To obtain  $C_{total}$  for auto-detected slanted-edge modules, [SFRplus](#), [eSFR ISO](#), and [Checkerboard](#), select **3D & contour plots**, then select **Edge info Cap C\_max** (on the right of the Rescharts window, below). The mean value of  $C_{max}$  for the image will also be displayed. For the information capacity plots ( $C_4$  and  $C_{max}$ ), the zone weights are always [1, 1, 1].

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



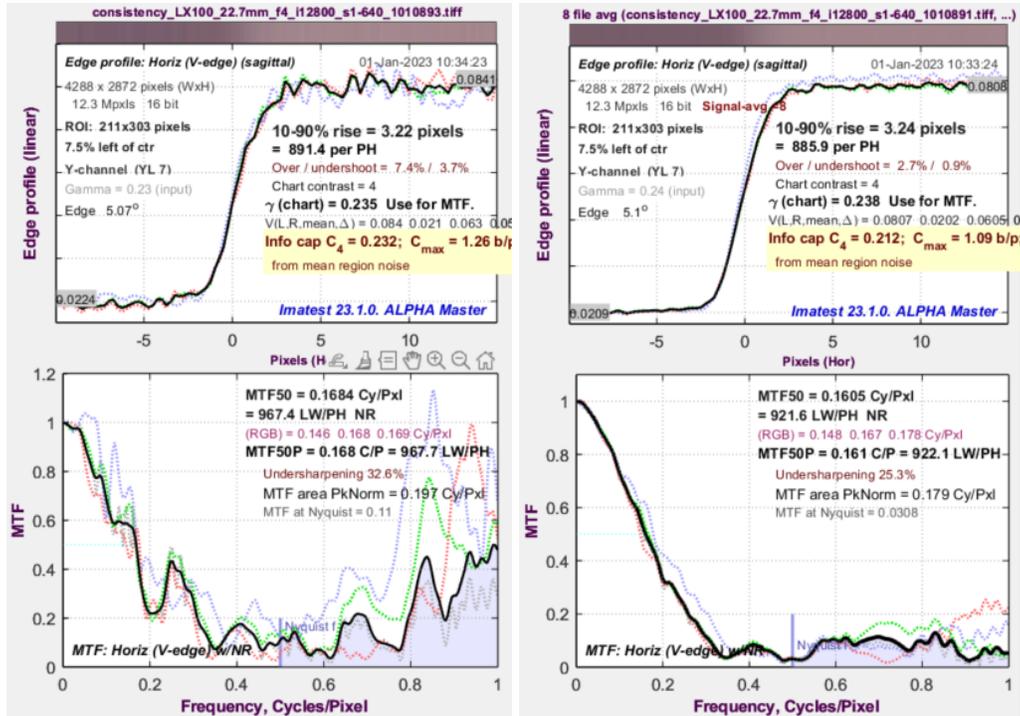
The mean information capacity  $C_{max}$  (unweighted for information capacity calculations) is 2.847 bits/pixel. Since the camera has 16 Megapixels, total capacity  $C_{maxTotal}$  for the Luminance (Y) channel = 45.44 MB.

## Signal averaging

This well-known technique can be used to improve the accuracy and consistency for measurements of noisy images for both the Edge Variance and Noise image methods.

Extremely noisy images, typically acquired in dim light or at high Exposure Indices, may result in inaccurate measurements of  $MTF$  and  $C$ . Signal averaging, where  $n$  identical captures of the same image are averaged, is a classic technique for reducing the effect of noise and obtaining better measurements from noisy images. When  $n$  images are averaged, the sum of the signal voltage and the sum of the noise power (noise voltage<sup>2</sup>), which is uncorrelated, are both proportional to  $n$ . This causes noise voltage to be proportional to  $\sqrt{n}$ , so that SNR increases by  $\sqrt{n}$ : by 3dB whenever  $n$  is doubled. To obtain correct information capacity measurements when the signal is averaged, the noise power is multiplied by  $n$  in the information capacity calculation.

This effect is illustrated below for a camera with a one-inch sensor, which was imperfectly focused, at ISO 12800. A single image is shown on the left. Note that MTF is rough and has significant high frequency noise bumps. For the average of 8 images is shown on the right, information capacity  $C$  is slightly reduced because MTF is better behaved, i.e., there is less spurious high frequency response.



### Some key results of the Edge Variance method

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

#### Cameras used in the tests

1.	Panasonic Lumix LX5	<b>2.14 <math>\mu\text{m}</math> pixel pitch.</b> An older (2010) compact 10.1-megapixel camera with a Leica f/4 zoom lens set to f/4.
2.	Sony A6000	<b>3.88 <math>\mu\text{m}</math> pixel pitch.</b> A 24-megapixel micro four-thirds camera with a 60mm Canon macro lens set to f/8
3.	Sony A7Rii	<b>4.5 <math>\mu\text{m}</math> pixel pitch.</b> A 42-megapixel full-frame camera with a Backside-Illuminated (BSI) sensor and a 90mm f/2.8 Sony macro lens set to f/8

We captured both JPEG and raw images, which were converted to 24-bit sRGB (encoding gamma  $\cong 1/2.2$ ) TIFF images (designated as raw→TIFF) with [LibRaw](#), with minimal processing (defined as no sharpening, no noise reduction, and a simple gamma-encoding function). Results for 48-bit Adobe sRGB conversion were nearly identical.



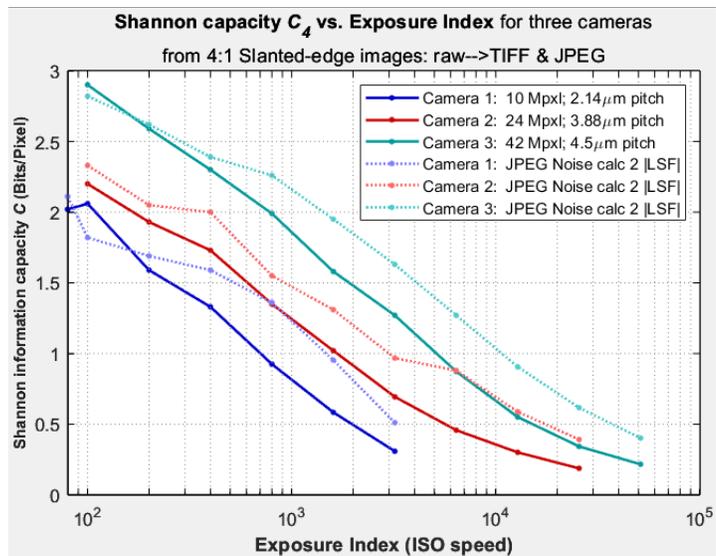
The image on the right, which was analyzed in “[Camera Information Capacity: A Key Performance Indicator for Machine Vision and Artificial Intelligence Systems](#)” [3], contains a 50:1 contrast Siemens star and four 4:1 contrast slanted edges on the sides. We used the upper-left slanted edge for most tests. The average background of the chart is close to neutral gray (18% reflectance) to ensure a good exposure (exposure compensation may be applied if needed and available).

### Results for JPEG and minimally-processed raw→TIFF images

The two figures below show results for the luminance ( $Y = 0.2125 \cdot R + 0.7154 \cdot G + 0.0721 \cdot B$ ) channel as a function of ISO speed (Exposure Index) 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.

#### $C_4$ 4:1 slanted edge

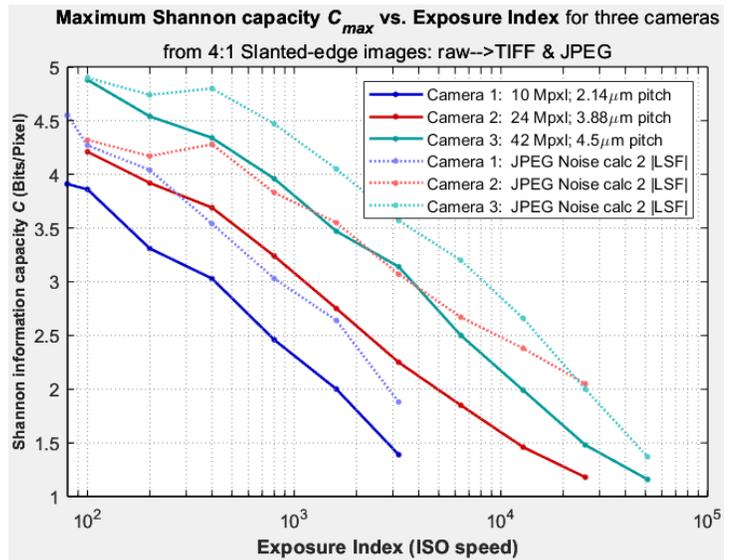
The information capacity for 4:1 contrast edges,  $C_4$ , shows similar trends to  $C_{max}$ , but since the relatively low 4:1 contrast uses only a fraction of the available signal level,  $C_4$  is lower than either measurement. It is also highly sensitive to exposure.



## $C_{max}$ 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  $C_4$ , making it robust and well-suited for comparing the performance of different cameras.

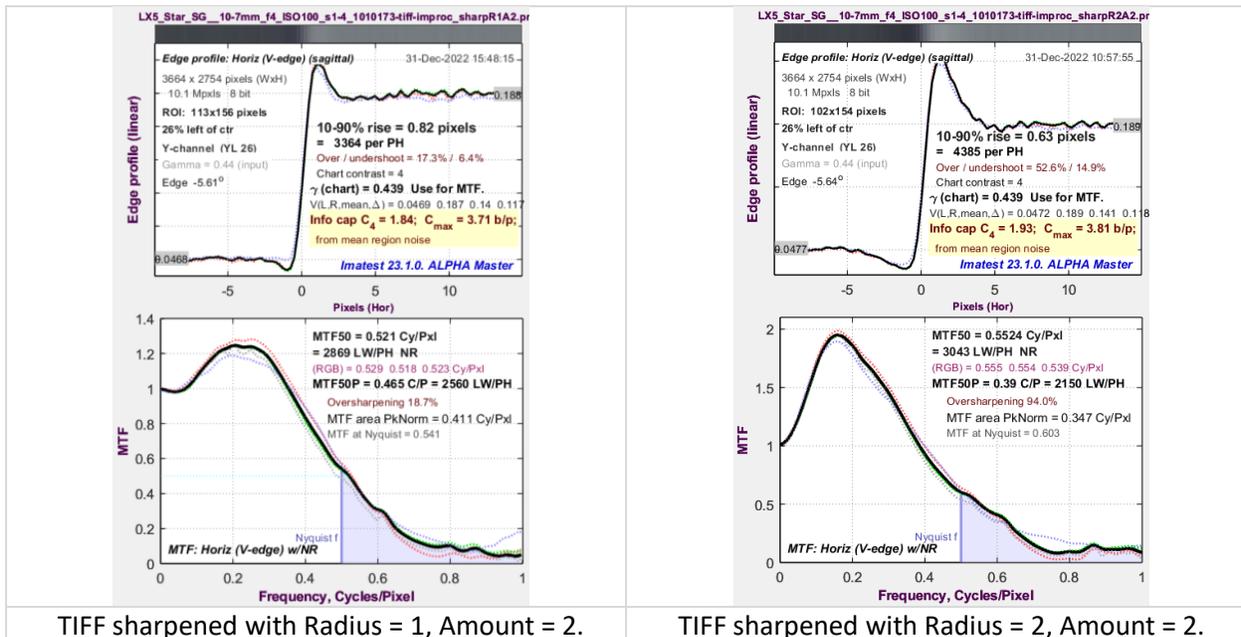
Both  $C_4$  and  $C_{max}$  give the expected results:  $C$  is higher for the higher quality (larger pixel) sensor, and decreases for increased Exposure Index (less exposure and more analog gain, resulting in poorer SNR).



## Sharpening

Simple sharpening, which has the same effect on the signal and noise response, would not be expected to have a strong effect on  $C$ . This is indeed the case.

The two examples below show that sharpening has little effect on slanted-edge information capacity, as expected for a valid measurement. The two images (originally a minimally-processed TIFF) have been strongly USM sharpened in the [Imatest Image Processing](#) module with Radii = 1 and 2 and Amount = 2. The original unsharpened TIFF has  $C_4 = 2.06$  and  $C_{max} = 3.82$  b/p.



This highlights another benefit of information capacity measurements. Unlike MTF50, they do not reward excessive sharpening, which creates “halos” near edges. These halos improve the human perception of sharpness when applied in moderation, but create artifacts that degrade image appearance when applied in excess [9]. They also have a bad reputation for machine vision applications.

### Summary of the Edge Variance method

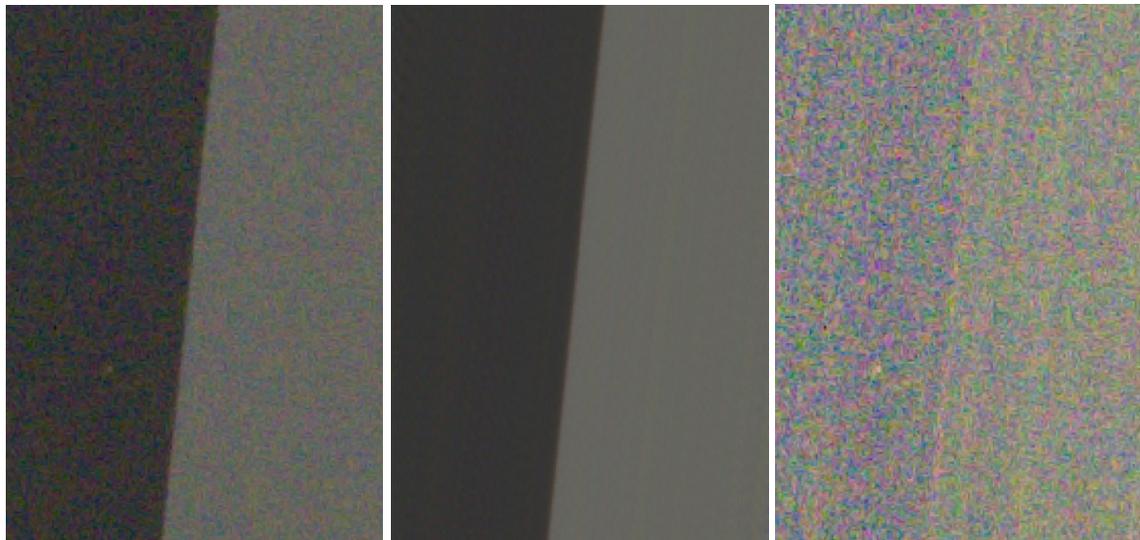
- The Edge Variance method is the first of two methods for calculating information capacity,  $C$ , from slanted edges.
- It has a limited set of results.
  - Information capacities  $C_4$  and  $C_{max}$ ,
  - A plot of spatially-dependent noise power  $\sigma_s^2(x)$  or amplitude  $\sigma_s(x)$ , which can be useful for determining if the image has been bilateral-filtered.
- Produces an interesting approximate measurement of  $C$  for bilateral-filtered images, but uniformly/minimally processed images give the most accurate results, and should *always* be used when a camera is being evaluated for use in MV/AI systems.
- Results are simple and convenient to obtain, even though the algorithms behind them can be complex. For the most part, the *Imatest* user doesn't need to be concerned about the calculation method.

### The Noise Image method of calculating information capacity-related metrics

The Noise Image method is the second of two methods for calculating information capacity and related figures of merit. It was developed shortly after the Edge Variance method. It offers a particularly rich set of measurements.

This method involves inverting the ISO 12233 binning procedure. Noting that the 4× oversampled edge was created by interleaving the contents of 4 bins, each of which contains an averaged (noise-reduced) signal derived from the original image, we apply an inverse of the binning algorithm to set the contents of each scan line to its corresponding interleave (**Inverse binned... ROI, below**). Since the inverse-binned image is a nearly noiseless replica of the original image, we can create a noise image by subtracting the inverse-binned image from the original image (which must be corrected for illumination nonuniformity in the direction of the edge). This image is shown, adjusted to make the mean (zero) value middle gray, as the **Noise image ROI**, on the right below.

The three images are shown below. The noise image (below-right), which has a mean value of 0, has been lightened and contrast-boosted for display. The other images are displayed with gamma-correction.



Original ROI

(2) Inverse-binned /  
de-interleaved / reverse-projected ROI

(3) Noise image ROI

These images allow several key image quality parameters to be calculated, including Noise Power Spectrum and Noise Equivalent Quanta, well-known in medical imaging systems, and described in [an excellent review paper by Ian Cunningham and Rodney Shaw \[10\]](#). These measurements are not well-known outside of medical imaging, in part because they have been difficult to measure.

### Displaying the noise image results

The key Noise image results are in the **Noise Spectrum, NEQ, SNRi** plot, which has numerous display options.

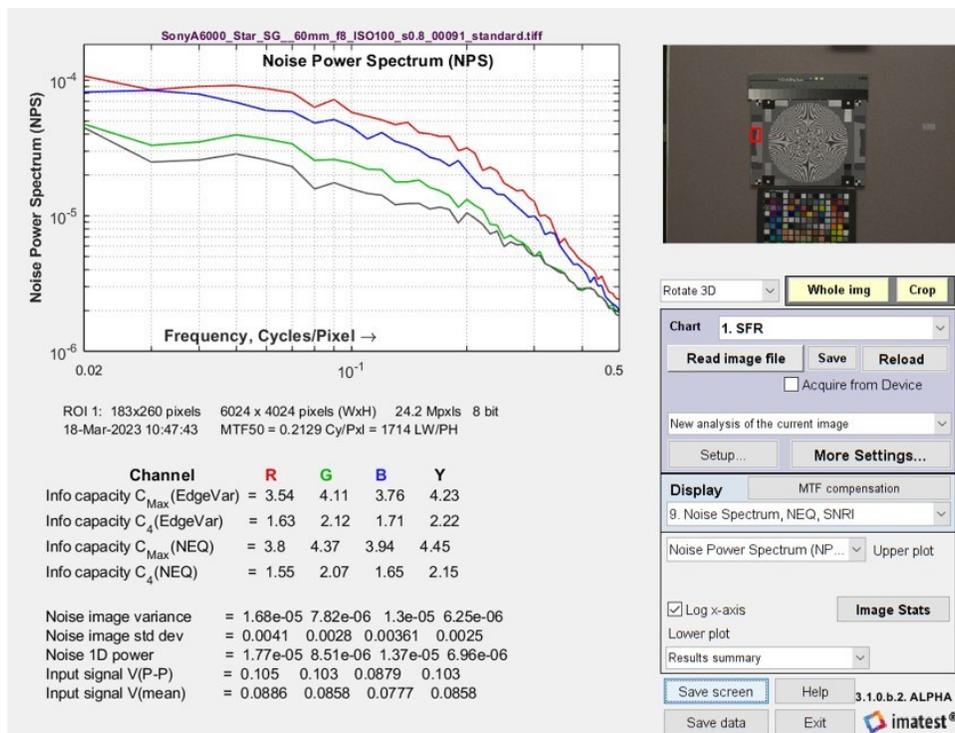
This plot displays two results: one at the top and one at the bottom. The contents of the upper and lower plots are selected in the Display area on the right of the Rescharts window, shown in the middle of the table below.

<u>Upper or Lower plot</u>	Display settings	<u>Lower plot-only</u>
<p>(All these results available in either the upper or lower plot.)</p> <p><b><u>Noise Voltage Spectrum</u></b></p> <p><b><u>Noise Power Spectrum (NPS)*</u></b> (shown above)</p> <p><b><u>Noise Equivalent Quanta (NEQ)</u></b></p> <p><u>MTF*</u></p> <p><u>Edge linearized unnormalized*</u></p> <p><b><u>Noise autocorrelation</u></b></p>		<p><b><u>Original image crop</u></b> (1)</p> <p><b><u>Unbinned image crop</u></b> (1) (Reverse-projected; low noise)</p> <p><b><u>Noise image crop</u></b> (1) (Original – Noise)</p> <p><b><u>Results summary</u></b> (1) (Shown above)</p> <p><b><u>SNRi 2D square w x w</u></b></p> <p><b><u>SNRi 2D rectangle w x 4w</u></b></p> <p><b><u>Square visibility image</u></b> (1)</p> <p><b><u>Square visibility – LARGE</u></b> (1)</p> <p><b><u>Noise Voltage Spectrum</u></b></p>

<u>Upper or Lower plot</u>	<u>Display settings</u>	<u>Lower plot-only</u>
<u>SNRi 2D square <math>w \times w</math></u> <u>SNRi 2D rectangle <math>w \times 4w</math></u> *for checking inputs to NEQ calculation (These plots are available elsewhere.)		<u>Noise Power Spectrum (NPS)</u> <u>Noise Equiv. Quanta (NEQ)</u> <u>MTF*</u> <u>Edge linearized unnormalized*</u> <u>Noise autocorrelation</u>

(1) Plots displayed in dark red available in upper plot-only.

Here is an example, with Noise Power Spectrum (NPS) displayed on the top and Results summary displayed on the bottom.



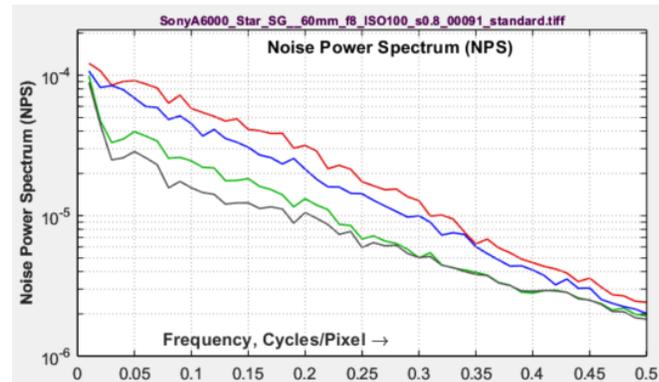
**Noise Spectrum, NEQ, SNRi plot with Noise Power Spectrum (NPS) displayed on the top and Results summary (showing the two different information capacity calculations) on the bottom.**

We now list the measurement (Figures of Merit) available with the Noise image method. Some are unfamiliar, and some are experimental.

## Noise Voltage or Power Spectrum (NPS)

The *NPS* (upper plot above) is displayed above with a logarithmic x-axis and on the right with a linear x-axis (selectable by a checkbox). The Noise Power and Voltage Spectrum plots have the same shape: only the y-axis labels are different.

The 1D Noise Power or Voltage spectrum is derived from a 2D Fourier transform (FFT) of the noise image. The initial 2D FFT has zero frequency at the image center. The image is divided into several annular regions, and the average noise power is found for each region. *NPS* is used in the *NEQ* and *SNR<sub>i</sub>* calculations.



## Noise Equivalent Quanta (NEQ)

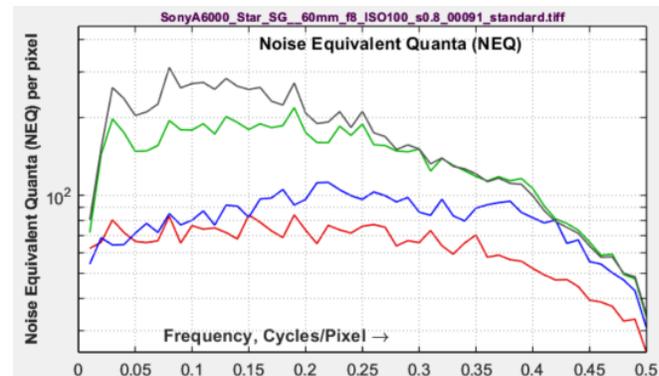
*NEQ* is a figure of merit used in medical imaging, but is unfamiliar in general imaging. It is described in a 2016 paper by Brian Keelan [5] and in an earlier paper by Cunningham and Shaw [10]. Essentially, it is a frequency-dependent Signal-to-Noise (power) Ratio, in contrast to *MTF*, which is signal amplitude response-only.

Units are the equivalent number of detected quanta that would generate the measured SNR when photon shot noise is dominant.

$$NEQ(f) = \frac{\mu^2 MTF(f)^2}{NPS(f)}$$

where the mean linear signal,  $\mu$ , can be defined in either of two ways, depending on how *NEQ* is to be interpreted.

In the standard definition of *NEQ*, where *NPS* is dominated by photon shot noise,  $\mu^2 = V_{mean}^2 = \bar{q}^2$ , where  $\bar{q}$  is the mean count of the detected quanta. But because noise is uncorrelated,  $NPS = \mu = \bar{q}$ . Therefore, *NEQ* is proportional to the count of detected quanta,  $\bar{q}$ . For example,  $NEQ = 200$  corresponds to a mean of  $\bar{q} = 200$  detected quanta detected at each pixel (assuming that the noise is dominated by photon shot noise).



The above equation,  $\mu = V_{mean} = \bar{q}$ , is appropriate if *NEQ* is to be used for calculating *DQE* (Detective Quantum Efficiency), where  $DQE(f) = NEQ(f)/\bar{q}_i$ , where  $\bar{q}_i$  is the mean number of quanta incident on each pixel. Measuring *DQE* requires a separate (and very exacting) measurement of  $\bar{q}_i$ , which we may add in the future.

Getting familiar with the meaning and use of *NEQ* may take some time. [Characterization of imaging performance in differential phase contrast CT compared with the conventional CT: Spectrum of noise](#)

[equivalent quanta NEQ\(k\)](#) [16] by Tang et. al. is an excellent example of how NEQ is used in medical imaging.

The NEQ plot is rough because of the relatively small size of the slanted-edge ROIs (Regions of interest). It can be improved using [Signal Averaging](#).

### Information capacity from NEQ:

A special form of  $NEQ$ ,  $NEQ_{info}(f)$  (not plotted), calculated using  $\mu = V_{p-p}/\sqrt{12}$  (to be consistent with the Edge Variance calculation), is used to calculate information capacity,  $C_{NEQ}$ .

$$C_{NEQ} = \int_0^W \log_2(1 + NEQ_{info}(f)) df$$

where bandwidth  $W$  is the camera's Nyquist frequency,  $W = f_{Nyq} = 0.5$  Cycles/Pixel. [Author's note: I thought I'd discovered this connection, but it's in papers on PET scanners and Digital Mammography by Christos Michail et. al. [6],[7]. Not papers anybody outside medical imaging is like encounter.]

The key results,  $C_4(NEQ)$  and  $C_{max}(NEQ)$ , are included in the Results summary. They are slightly different from the Edge Variance results, most likely because the calculated Noise Power Spectrum,  $NPS(f)$ , is used. (The Edge Variance calculation assumes constant  $NPS$ ).

Channel	R	G	B	Y
Info capacity $C_{Max}$ (EdgeVar)	= 3.54	4.11	3.76	4.23
Info capacity $C_4$ (EdgeVar)	= 1.63	2.12	1.71	2.22
Info capacity $C_{Max}$ (NEQ)	= 3.8	4.37	3.94	4.45
Info capacity $C_4$ (NEQ)	= 1.55	2.07	1.65	2.15
Noise image variance	= 1.68e-05	7.82e-06	1.3e-05	6.25e-06
Noise image std dev	= 0.0041	0.0028	0.00361	0.0025
Noise 1D power	= 1.77e-05	8.51e-06	1.37e-05	6.96e-06
Input signal V(P-P)	= 0.105	0.103	0.0879	0.103
Input signal V(mean)	= 0.0886	0.0858	0.0777	0.0858

### Ideal Observer SNR (SNRi)

is a measure of the detectability of small objects. It is described in papers by Paul Kane [14] and Orit Skorka and Paul Kane [15]. The two-dimensional equation in [15] gives the correct results.

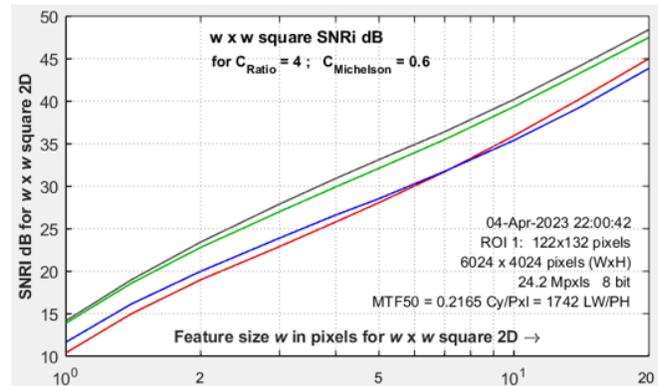
$$SNRi^2 = \iint \left( \frac{\mu^2 \Delta S^2(v_x, v_y) MTF^2(v_x, v_y)}{NPS(v_x, v_y)} \right) dv_x dv_y$$

$\mu \Delta S(v_x, v_y) = G(v_x, v_y)$  is the Fourier transform of the object to be detected; defined below.

$MTF(v)$  and  $NPS(v)$  are defined in one dimension, where spatial frequency  $v =$

$\sqrt{v_x^2 + v_y^2}$  has units of Cycles/Pixel, and the linearized signal is normalized to a maximum value of 1.

The object to be detected is typically a rectangle of dimensions  $w \times kw$ , where  $k = 1$  (for a square) or 4 for a 1x4 aspect ratio rectangle. Its amplitude (for the initial analysis) is the peak-to-peak voltage of the slanted edge,  $\Delta Q = V_{p-p}$  which is typically obtained from a chart with a 4:1 contrast ratio.



SNRi curves, Micro 4/3 camera, ISO 100

$$\Delta g(x, y) = \Delta Q \cdot \text{rect}\left(\frac{x}{w}\right) \cdot \text{rect}\left(\frac{y}{kw}\right)$$

where  $\text{rect}(x) = 1$  for  $-1/2 < x < 1/2$ ; 0 otherwise.

$G(v_x, v_y)$  is the Fourier transform of the object to be detected,  $\Delta g(x, y)$ . It is expressed in two dimensions.

$$G(v_x, v_y) = kw^2 \Delta Q \frac{\sin(\pi w v_x)}{\pi w v_x} \frac{\sin(\pi k w v_y)}{\pi k w v_y}$$

$SNRi^2$  is calculated numerically by creating a two-dimensional array of frequencies (0 to 0.5 c/p in 51 steps) that has  $v_x$  on the x-axis  $v_y$  on the y-axis, and is filled with  $v = \sqrt{v_x^2 + v_y^2}$ . These frequencies are used to create a 2D array that can be numerically summed [15].

$$SNRi^2 = \Delta v_x \Delta v_y \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{MTF^2(i, j) V_{P-P}^2}{NPS(i, j)} \Delta S^2(i, j)$$

$SNRi$  is displayed for each color channel for  $w$  from 1 to 40 in increments of approximately the square root of  $w$  (1, 1.4, 2, ...), typically,  $w = 1, 2, 3, 4, 7, 10, 14, 20$ .

Unlike  $C$ ,  $SNRi$  is affected by signal processing (sharpening, etc.), but we have yet to find a sharpening setting that consistently improves  $SNRi$ . More work is needed.

### SNRi information capacity: $C_{SNRi}$ (new and experimental)

$C_{SNRi}$  is an experimental measurement that takes advantage of the fact that the argument of the  $SNRi^2$  integral,

$$\frac{\mu^2 \Delta S^2(v_x, v_y) MTF^2(v_x, v_y)}{NPS(v_x, v_y)},$$

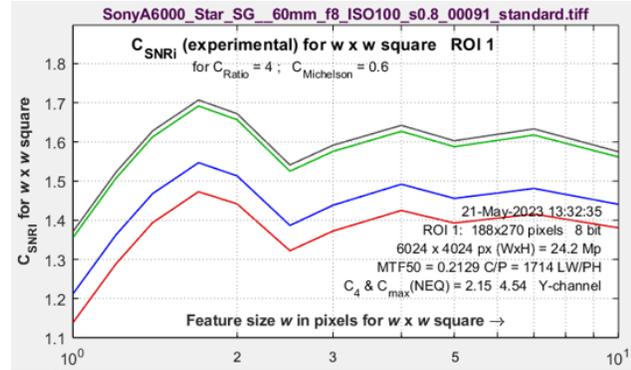
has units of Signal-to-Noise Ratio power, and hence can be used in the Shannon-Hartley equation,

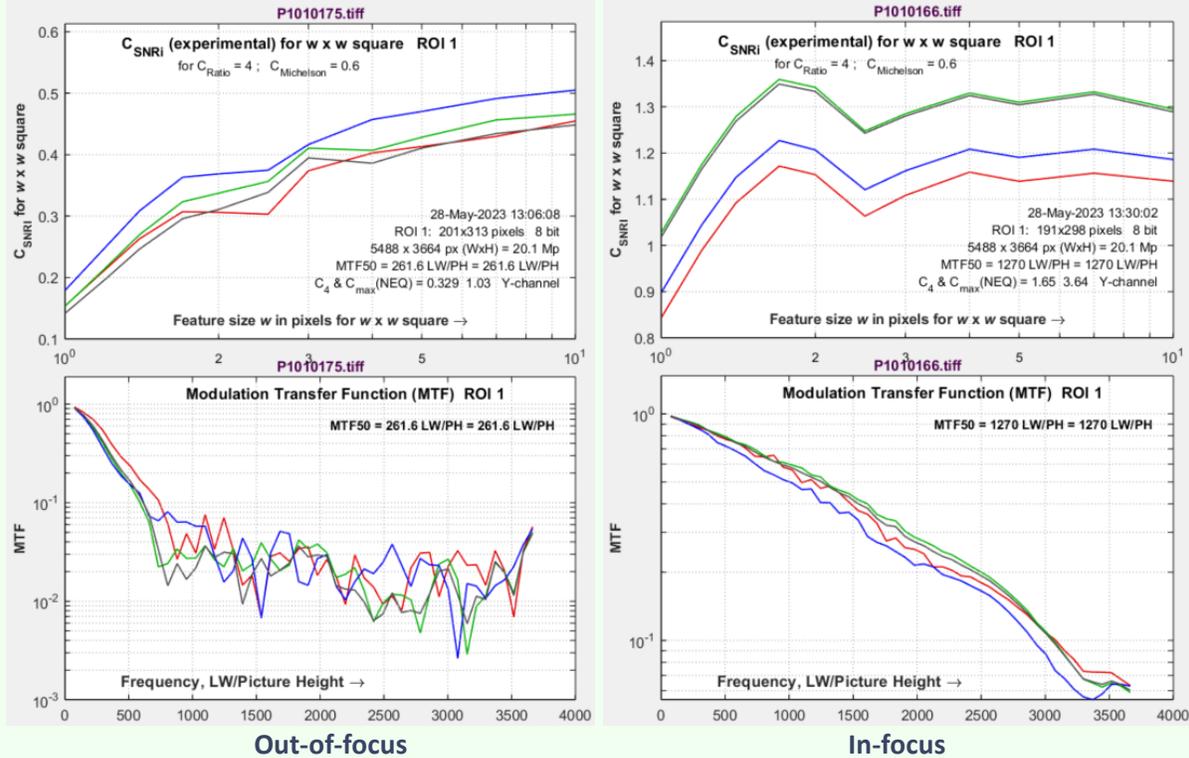
$$C_{SNRi} = \int \int \log_2 \left( 1 + \frac{\mu^2 \Delta S^2(v_x, v_y) MTF^2(v_x, v_y)}{NPS(v_x, v_y)} \right) dv_x dv_y$$

The numerical evaluation of  $C_{SNRi}$  follows the same steps as  $SNRi^2$ , above.

$C_{SNRi}$  is not yet fully validated, but it is promising (we are working on it). Plots are generally more readable and easier to interpret than  $SNRi$ . For images with poor MTF,  $C_{SNRi}$  increases more slowly for small widths and reaches a lower asymptotic value. The (rough) asymptotic value is close (though not identical) to  $C_4(\text{NEQ})$ .  $C_{SNRi}$  should be available in the pilot program by June 2023. (Not sure of the release schedule).

Here is an example from two images from a camera with a one-inch sensor described in [FocusField](#).





Note that  $C_{SNRi}$  increases more rapidly with  $w$  for the in-focus image and approaches a higher asymptotic value ( $\approx 1.2$  vs.  $0.47$ ).

## Object visibility

The goal of  $SNRi$  measurements is to predict object visibility for small, low contrast squares or 4:1 rectangles. The  $SNRi$  prediction begs for visual confirmation. A simulated image that can do this is shown in Figure 3 of a classic  $SNRi$  paper [8].

We have developed a display for Imatest that does this with a real slanted-edge image. Despite the trickery, the data is directly from the acquired image.

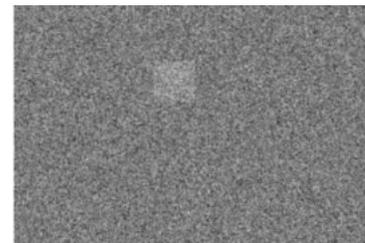
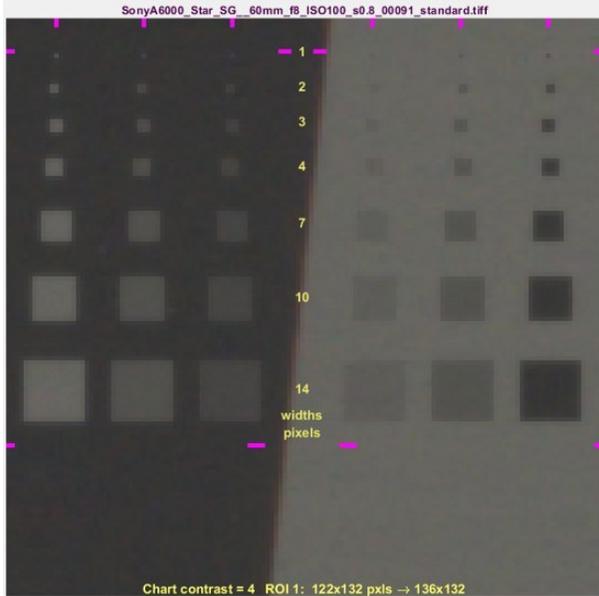


Figure 3. Simulated images of square objects

We show two sets of results: one for a relatively low noise image and one high noise image (both from a camera with Micro Four-Thirds sensors, at ISO 100 and 12800, respectively). The sides of the squares are  $w = 1, 2, 3, 4, 7, 10, 14,$  and  $20$  pixels. The original chart has a 4:1 contrast ratio (light/dark = 4), equivalent to a Michelson contrast  $C_{Mich} = (\text{light-dark})/(\text{light+dark}) = 0.6$ . The outer squares have  $C_{Mich} = 0.6$ . The middle and inner squares have  $C_{Mich} = 0.3$  and  $0.15$ , respectively.

**How to use these images** — Inconspicuous magenta bars are designed to help finding the small squares, which are hard to see. The yellow numbers are the square widths in pixels. The  $SNRi$  curves (initially, at least) represent the chart contrast — with 4:1 (the ISO 12233 standard [4]) strongly recommended. The outer patches correspond to the  $SNRi$  curves, where, according to the [Rose model](#) [10],  $SNRi$  of 5 (14 dB) should correspond to the threshold of visibility.

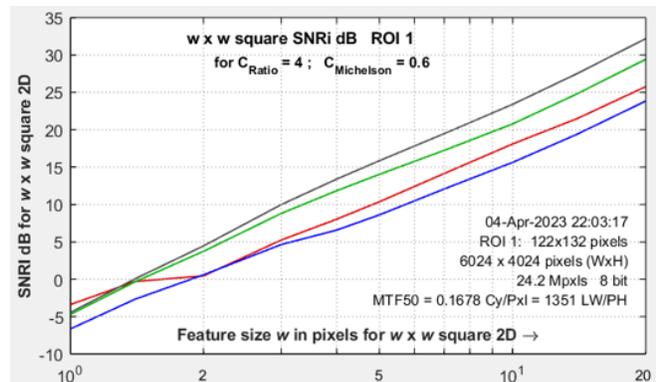


Low noise image, ISO 100



Noisy image, ISO 12800

The  $SNR_i$  curve on the right is for the noisy ISO 12800 image on the right, above. The  $w = 1$  squares are invisible; the  $w = 2$  and 3 squares are only marginally visible, and  $w = 4$  squares are clearly visible. In the plot,  $SNR_i$  at  $w = 2$  is 0-5 dB; it reaches 5-10 dB for  $w = 3$ ; close to the expectation that the threshold of visibility is around 14 dB.



Only original pixels were used in these two images, but we used a little smoke and mirrors to make the squares that have the same blur as the device under test.

#### How the squares were made

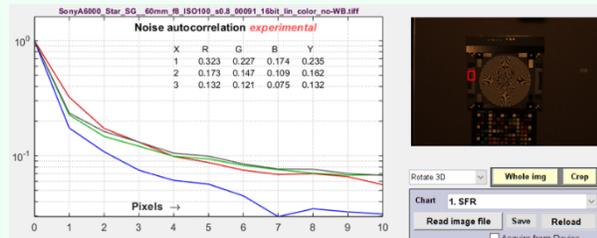
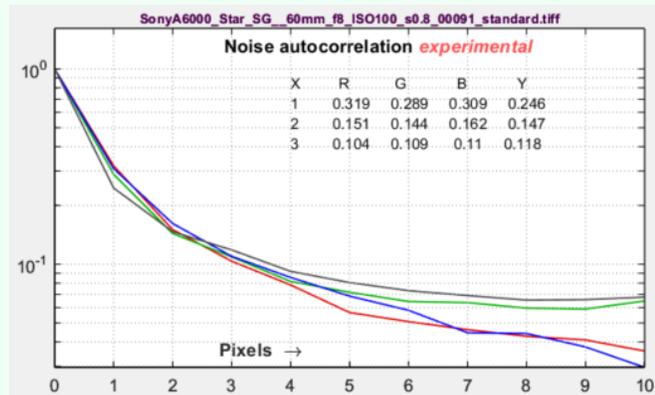
1. Expand the image if needed (if the original is less than 170 pixels wide) to make room for all the squares by adding mirrored versions of image to the left and right to the sides of the image. If needed, add a cropped vertical mirrored image to the bottom.
2. Create a (horizontal) mirror of the full image. This is the “mirror” part.
3. Create a mask consisting of ideal  $w \times w$  squares. The background is 0 and the squares are 1. The sides are sharp.

4. Blur the squares with the MATLAB filter2 function. This is the “smoke” part. Determining the blur kernel was challenging. We found that we couldn’t get good results by just using the 1D Line Spread function (LSF) in 2D. A more complex transformation was required.
5. Linearize the two images (remove the gamma encoding).
6. Combine them using the mask, keeping the original image where the mask = 0, using the mirrored image where the mask = 1, and blending them elsewhere.
7. Reapply the gamma encoding.

## Noise autocorrelation

This plot has been added to examine the hypothesis that the noise power spectrum (and autocorrelation) indicate the amount of electrical crosstalk of image sensor when the effects of demosaicing and fixed-pattern noise are removed and the primary noise source is photon shot noise. The idea behind the hypothesis is that light incident on the sensor is entirely uncorrelated, so that if there were no crosstalk, the noise would be white.

This image used for the first plot was white-balanced. The curve is the  $|\text{inverse Fourier transform}|$  of the noise spectrum, based on the author’s understanding of the [Wiener-Khinchin](#) theorem.



The image on the (lower) right was not White-Balanced. The red channel has a larger autocorrelation distance than the other channels, as we would expect. Click on the image to enlarge it.

A similar autocorrelation plot can also be obtained from a flat field image in the [Image Statics](#) module. Illumination nonuniformity has been corrected to decrease the (spurious) autocorrelation at large distances.

$MTF(f)$  and Edge Voltage  $V$  are now shown here because they are included in the standard calculations.

## Key measurements from the Noise image method

$NEQ(f)$  is relatively unfamiliar outside of medical radiology.

Measurement	Description
-------------	-------------

<a href="#">Noise Power Spectrum, <math>NPS(f)</math> (or Noise Voltage Spectrum)</a>	Used in $NEQ$ and $SNR_i$ calculations. $NPS$ was implicitly assumed to be constant (white noise) in the Edge Variance method.
<a href="#">Noise Equivalent Quanta, <math>NEQ(f)</math> and <math>NEQ_{info}(f)</math></a>	A measure of frequency-dependent signal-to-noise ratio ( $SNR$ ). $NEQ(f) = \mu^2 MTF(f)^2 / NPS(f)$ , where $\mu = V_{mean}$ has been used for quantifying medical image quality, but is much less familiar in general imaging. $NEQ(f)$ is equivalent to the number of quanta detected by the sensor when photon shot noise is dominant. It is appropriate for calculating Digital Quantum Efficiency (DQE), when the density of quanta reaching the image sensor is known. $NEQ_{info}(f)$ is derived from $\mu = VP - P/12 - \nu$ , making it well-suited for calculating information capacity CNEQ.
<a href="#">Information capacities <math>C_4(NEQ)</math> and <math>C_{max}(NEQ)</math></a>	correspond to $C_4$ and $C_{max}$ from the Edge variance method. Derived from $NEQ_{info}(f)$ . They are close, but not identical.
<a href="#">Ideal observer Signal-to-Noise Ratio, <math>SNR_i</math></a>	From Kane [14] and Skorka and Kane [15], "The Ideal Observer is a <a href="#">Bayesian decision maker</a> that maximizes the statistical precision of a hypothesis test with two possible outcomes." $SNR_i$ is a metric of the detectability of small objects (squares or rectangles), typically of low contrast.
<a href="#">Object visibility</a>	Images of low contrast squares of various sizes: a visual indicator of object visibility. Correlates with $SNR_i$ .
<a href="#">Noise autocorrelation</a>	The inverse Fourier transform of the Noise Voltage Spectrum. May be related to sensor electrical crosstalk.

### Summary of the noise image method

- The Noise Image method is the second of two methods for calculating information capacity,  $C$ , from slanted edges.
- It only gives reliable results with uniformly or minimally processed images, which can be distinguished from bilateral-filtered images by the absence of a peak in  $\sigma_s^2(x)$  or  $\sigma_s(x)$  displays.
- It produces a rich set of related results, including Noise Power Spectrum ( $NPS$ ), Ideal observer SNR ( $SNR_i$ ), Noise Equivalent Quanta ( $NEQ$ ), and a second set of information capacity measurements, derived from  $NEQ$ , that can be compared with the Edge variance results (they are slightly more accurate because  $NPS$  is not assumed to be constant).
- Some of the results are new and unfamiliar. It may take some time before we have a full understanding of their value.

### Overall summary

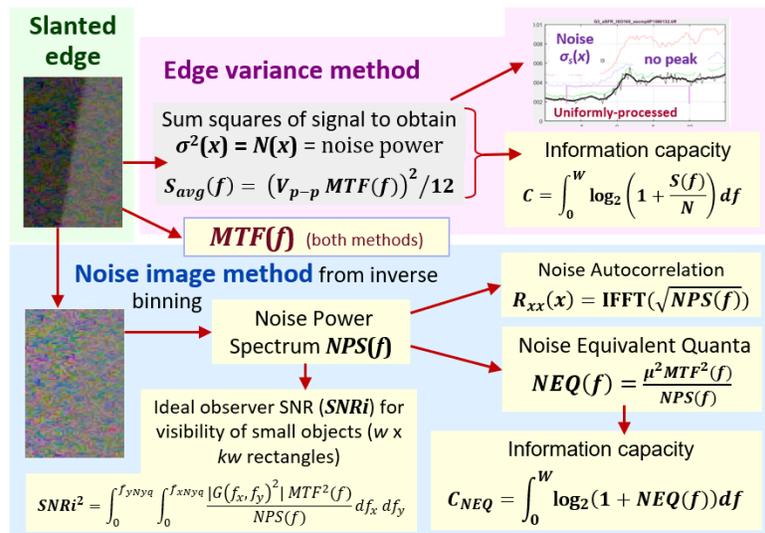
We have presented several new Figures of Merit for imaging systems that are especially applicable to Machine Vision/Artificial Intelligence systems. The most important of these is **information capacity**, which combines sharpness and noise, and is calculated in two different ways (with similar results).

Information capacity can be used to optimize camera selection for machine vision applications. This involves specifying the required information capacity, then finding the camera that meets this spec with the minimum number of pixels, which will result in the fastest calculations, lowest power consumption, and—perhaps most importantly—minimum cost.

The new measurements are extremely easy to obtain from any of *Imatest's* slanted-edge analyses. All you need to do is [turn them on](#), and they appear in the Edge/MTF plot. But the math and algorithms behind the measurements contain concepts that are unfamiliar to most imaging scientists. They are worth learning.

Compared to the earlier Siemens star information capacity method [3], the slanted-edge method is faster, more convenient, better for mapping results over the entire image, and better for calculating the total information capacity. For reliable measurements, Siemens stars should be well-centered, especially if there is significant optical distortion. Siemens stars are better for quantifying the effects of demosaicing methods, image compression, and image saturation.

The diagram below contains a summary of the two slanted-edge methods, illustrating the rich interconnections between the new KPIs. For the most part, the *Imatest* user does not need to be concerned about which of the two methods is used.



As of May, 2023, there is still much work to be done.

- Work with partners in industry and academia to correlate information capacity  $C$  with performance of Machine Vision and Artificial Intelligence systems (accuracy, speed, and power consumption).
- Add camera information capacity in several standards, especially ISO TC42.
- Search for metrics that can be used to predict the effects of image processing (sharpening, noise reduction) on MV/AI performance.
- Find better ways of modeling High Dynamic Range (HDR) sensors, where noise is not a simple monotonic function of signal.

## 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.  
[people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf](http://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf)
- [2] C. Shannon, "Communication in the Presence of Noise", Proceedings of the I.R.E., January 1949, pp. 10-21.  
[fab.cba.mit.edu/classes/S62.12/docs/Shannon\\_noise.pdf](http://fab.cba.mit.edu/classes/S62.12/docs/Shannon_noise.pdf)
- [3] N. L. Koren, "[Measuring camera Shannon Information Capacity with a Siemens Star Image](#)", Electronic Imaging conference, 2020. The [Camera Information Capacity](#) (Siemens Star) white paper is much more readable.
- [4] ISO 12233:2017: <https://www.iso.org/standard/71696.html>
- [5] Brian W. Keelan, "Imaging Applications of Noise Equivalent Quanta" in *Proc. IS&T Int'l. Symp. on Electronic Imaging: Image Quality and System Performance XIII*, 2016, <https://doi.org/10.2352/ISSN.2470-1173.2016.13.IQSP-213>.
- [6] Michail C, Karpetas G, Kalyvas N, Valais I, Kandarakis I, Agavanakis K, Panayiotakis G, Fountos G., [Information Capacity of Positron Emission Tomography Scanners](#). *Crystals*. 2018; 8(12):459. <https://doi.org/10.3390/cryst8120459>.
- [7] Christos M. Michail, Nektarios E. Kalyvas, Ioannis G. Valais, Ioannis P. Fudos, George P. Fountos, Nikos Dimitropoulos, Grigorios Koulouras, Dionisis Kandris, Maria Samarakou, Ioannis S. Kandarakis, "Figure of Image Quality and Information Capacity in Digital Mammography", *BioMed Research International*, vol. 2014, Article ID 634856, 11 pages, 2014. <https://doi.org/10.1155/2014/634856>.
- [8] Robert D. Feite, "Modeling the Imaging Chain of Digital Cameras", [SPIE Digital Library, 2010](#), Chapter 8, "The Story of Q".
- [9] N. L. Koren, [Correcting Misleading Image Quality Measurements](#), Electronic Imaging Conference, 2020
- [10] I.A. Cunningham and R. Shaw, "Signal-to-noise optimization of medical imaging systems", Vol. 16, No. 3/March 1999/pp 621-632/*J. Opt. Soc. Am. A*
- [11] 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. Reprinted in "Selected Readings in Image Evaluation," edited by Rodney Shaw, SPSE (now SPIE), 1976. [Available for download](#).
- [12] R. Jenkin and P. Kane, [Fundamental Imaging System Analysis for Autonomous Vehicles](#), Electronic Imaging, 2018
- [13] N. L. Koren, "[Using images of noise to estimate image processing behavior for image quality evaluation](#)", Electronic Imaging conference , 2021
- [14] 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>.
- [15] Orit Skorka and Paul J. Kane, "Object Detection Using an Ideal Observer Model", Electronic Imaging conference, 2020.
- [16] X. Tang, Y. Yang, S. Tang, Characterization of imaging performance in differential phase contrast CT compared with the conventional CT: Spectrum of noise equivalent quanta NEQ(k), *Med Phys*. 2012 Jul; 39(7): 4467–4482. Published online 2012 Jun 29. [doi: 10.1118/1.4730287](https://doi.org/10.1118/1.4730287).

## Appendix I. Information theory background

Because concepts of information theory are unfamiliar to most imaging engineers, we present a brief introduction. To learn more, we recommend a text such as "[Information Theory— A Tutorial Introduction](#)" by James V Stone, available on [Amazon](#). Shannon's 1948 and 1949 papers [1],[2] are highly readable.

## What is information?

Information is a measure of surprise or the resolution of uncertainty. The classic example is a coin flip. For a “fair” coin, which has a probability of 0.5 for either a head or tail outcome (which we can designate 1 or 0), the result of such a flip contains one bit of information. Note that two coin flips have four possible outcomes (00, 01, 10, 11); three coin flips have eight possible outcomes, etc. The number of information bits is  $\log_2$ (the number of outcomes), which is the number of flips.

Now, suppose you have a weirdly warped coin that has a probability of 0.99 for a head (1) and 0.01 for a tail (0). Little information is gained from the results of a flip. The equation for the information in a trial with  $m$  outcomes, where  $p(x_i)$  is the probability of outcome  $i$  and  $\sum_{i=1}^m p(x_i) = 1$ , is

$$H = \sum_{i=1}^m p(x_i) \log_2 \frac{1}{p(x_i)}$$

$H$  is called “entropy”, and is often used interchangeably with “information”. It has units of bits (binary digits). Note that this definition is subtly different from the physical memory element called a “bit.”

For the fair coin, where  $p(x_1) = p(x_2) = 0.5$ ,  $H = 1$  bit. But for the warped coin, where  $p(x_1) = 0.99$  and  $p(x_2) = 0.01$ ,  $H = 0.0808$  bits. If the results of the warped coin toss were transmitted without coding, each channel bit would contain 0.0808 information bits. That would be extremely inefficient.

Claude Shannon was one of the genuine geniuses of the twentieth century— renowned among electronics engineers, but little known to the general public. The medium.com article, [11 Life Lessons From History’s Most Underrated Genius](#), is a great read. (Perhaps Shannon is considered “underrated” because history’s most famous genius lived in the same town.) There are also nice articles in [The New Yorker](#) and [Scientific American](#). And IEEE has an [article connecting Shannon with the development of Machine Learning and AI](#). The 29-minute video “[Claude Shannon – Father of the Information Age](#)” is of particular interest to the author of this report because it was produced by the [UCSD Center for Memory and Recording Research](#), which he visited frequently in his previous career.



*Claude Shannon*

## Channel capacity

Shannon and his colleagues developed two theorems that form the basis of information theory.

The first, Shannon’s source coding theorem, states that for any message there exists an encoding of symbols such that each channel input of  $D$  binary digits can convey, on average, close to  $D$  bits of information. For the above example, it implies that a code can be devised that can convey close to 1 information bit for each channel bit—a huge improvement over the uncoded value of 0.0808.

The second, known as the Shannon-Hartley theorem, states that the [channel capacity](#),  $C$ , i.e., the theoretical upper bound on the [information rate](#) of data that can be communicated at an arbitrarily low [error rate](#) through an analog communication channel with bandwidth  $W$ , average received signal power,  $S$ , and [additive Gaussian noise](#) power,  $N$ , is

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

This equation is challenging to use because bandwidth  $W$  is not well-defined, noise is not white, and it applies to one-dimensional systems, whereas imaging systems have *two* dimensions, at least for Siemens stars. Slanted-edge analysis is one-dimensional. We have developed methods for calculating  $C$  for both the Siemens star and slanted edge test patterns.

At this point we can hazard a guess as to why camera information capacity has been ignored for cameras. For most of its history the hot topic in information theory was the development of efficient codes, which didn't approach the Shannon limit until the 1990s—nearly fifty years after Shannon's original publication. But channel coding is not a part of image capture (though coding is important for image and video compression). Also, camera information capacity was not critically important when the primary consumers of digital images were humans (though it is related to perceived image quality), but that is changing rapidly with the development of new AI and machine vision systems. And finally, convenient methods of measuring it didn't exist. (Rodney Shaw's heroic efforts with film in the early 1960s are very impressive [11].)

## Appendix 2. Binning noise

This “green for geeks” box can be skipped by most readers.

Binning noise, which has identical statistics to [quantization](#) noise, is a recently-discovered artifact of the ISO 12233 binning algorithm. It is largest near the image transition — where the Line Spread Function  $LSF(x) = d\mu_s(x)/dx$  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, which is a continuous function.

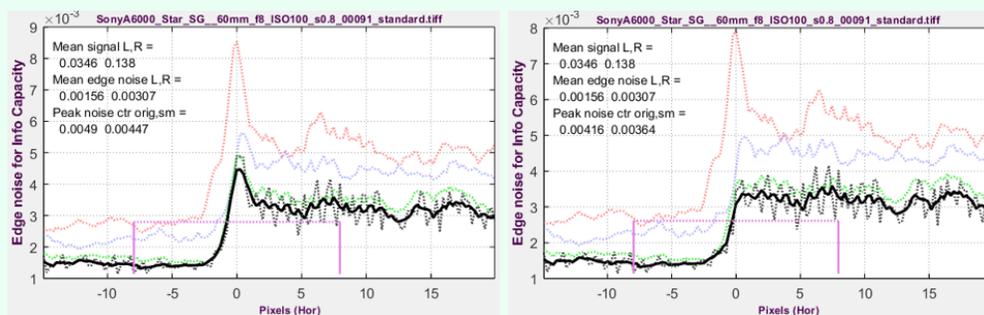
Assume that  $n$  identical signals  $\mu_s(x)$  are binned over an interval  $\{-\Delta/2, \Delta/2\}$ , where  $\Delta = 1$  in the  $4\times$  oversampled output of the binning algorithm (noting that  $\Delta = (\text{original pixel spacing})/4$ ). If there were no binning noise, we would expect the binning noise power  $\sigma_{Bnoise}^2$  to be zero. However, the values of  $\mu_s(x_k)$  are summed at uniformly-distributed locations  $x_k$  over the interval  $\Delta$ , so they take on values

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

for Line Spread Function  $LSF$ . Noting that  $\delta$  is uniformly distributed over  $\{-1/2, 1/2\}$  we apply the equation for the [variance of a uniform distribution](#) (similar to [quantization noise](#)) to get

$$\sigma_{Bnoise}^2(x) = LSF^2(x) \sigma_{Uniform}^2 = LSF^2(x)/12 \quad \text{or} \quad \sigma_{Bnoise} = LSF(x)/\sqrt{12}.$$

Although this equation involves some approximations, we have had good success calculating the corrected noise,  $\sigma_s^2(\text{corrected}) = \sigma_s^2 - \sigma_{Bnoise}^2$ . Binning noise has no effect on conventional MTF calculations.



*Edge noise for a Micro Four-Thirds digital camera, ISO 100, Y (Luminance) channel from raw image converted to TIFF with minimal processing.*

*Left: with binning noise*

*Right: binning noise removed*

Binning noise also affects JPEG files with bilateral filtering (nonuniform sharpening). Removing it improves the robustness of Edge Variance calculations.

At the time of the 23.1 release (May 2023), the Slanted edge calculation setting, on the lower-left of the **More settings** window, must be set to **Imatest 22.1 (recommended)**.