Image Information Metrics and Applications: Reference

Norman Koren, Imatest, LLC

December 5, 2023

The market for cameras that produce images for Machine vision (MV) and Artificial Intelligence (AI), in contrast to pictorial images for human vision, is steadily growing. Applications include automotive (driver assistance and autonomous vehicles), robotics, security, and medical imaging systems.

Two questions arise when designing camera systems for such applications.

1. How best to select (or qualify) cameras for MV/AI applications?
2. What image processing (ISP or filtering) is optimal?

To answer these questions, we must go beyond standard measurements of sharpness (MTF) and noise and apply metrics derived from **information theory**, including information capacity and related metrics for object and edge detection.

These metrics are important because Object Recognition (OR), MV, and AI algorithms operate on *information*, not *pixels*, making them far better predictors of system performance than MTF or noise.

Imatest has developed a highly convenient method for measuring information capacity and related metrics from the most widely used ISO standard resolution test pattern — the slanted edge. We describe how the new metrics can be used to select (or qualify) cameras and determine the optimum Image Signal Processing (ISP) for Object Recognition, which is likely to improve the performance of MV and AI algorithms.

A shorter, easier-to-read version of this white paper with fewer equations, **"Image Information Metrics in Imatest,"** is linked from **www.imatest.com/solutions/image-information-metrics**.

This document describes features of ***Imatest 24.1***, which will be available in the ***Imatest 24.1 Pilot program*** until it the spring 2024 release.

Table of Contents

Image Information Metrics and Applications: Reference 1

Table of Contents 2

Outline 3

Introduction 4

Information 4

The slanted edge 4

The Edge Variance method 5

Signal and noise results 6

Calculating information capacity from $\mu_s(x)$ and $\sigma_s(x)$ 6

Signal power S 6

Noise power N 7

Bandwidth W 8

Combining $S_{avg}(f)$, N, and W to obtain C 8

Maximum information capacity C_{max} — a more consistent metric 9

Consistency of C_{max} 10

Total information capacity 10

Signal averaging 11

Some key results of the Edge Variance method 12

C_4 4:1 slanted edge 13

C_{max} maximum information capacity 13

Sharpening 13

Edge location variance (or standard deviation) 14

Summary of the Edge Variance method 15

The Noise Image method of calculating information capacity-related metrics 16

Displaying the results 16

Noise Voltage or Power Spectrum (NPS) 18

Noise autocorrelation ... 19
Noise Equivalent Quanta (NEQ) ... 19
Information capacity from NEQ, C_{NEQ} ... 20
Ideal Observer SNR ($SNRi$) ... 20
SNRi displayed in dB per pixel^2 ... 22
Object visibility ... 22
Edge Signal-to-Noise Ratio ($Edge\ SNRi$) ... 24
Line Spread Function (LSF) doublet results ... 24
Edge SNRi in frequency and spatial domain ... 25
Table of key measurements from the Noise Image method ... 26
Summary of the Noise Image method ... 27
Image Signal Processing (ISP) ... 27
Pre-filtering: effects of sharpening and lowpass filtering ... 28
Matched filter ... 30
Units and Exposure sensitivity ... 31
Summary ... 32
References ... 33
Appendix I. Information theory background ... 34
What is information? ... 35
Channel capacity ... 35
Appendix 2. Obtaining Results with *Imatest* ... 36
Setting Channel capacity calculations, ... 37
1. Edge & Info Capacity noise ... 38
2. Information-related: NEQ, SNRi, ... 38
Appendix 3. Filtering images with the *Imatest* Image Processing module ... 39
Appendix 4. Correlation between information capacity and object/edge detection metrics ... 39
Appendix 5. Binning noise ... 40

Outline

This white paper begins with an Introduction, then it describes two mathematical approaches to calculating information capacity and related metrics from the slanted edge: the edge variance method and the noise image method.

We describe several information capacity-related metrics (measurements that combine sharpness and noise). This leads to the key object and edge detection metrics, $SNRi$ and $Edge\ SNRi$, that can be used to design a filter that significantly improves edge and object detection prior to sending the image to the MV/AI algorithm.

We will not cover important metrics not directly related to information, including dynamic range and susceptibility to stray light.

Finally, five appendices introduce information theory, describe how to filter images and obtain results from ***Imatest***, show how the key detection metrics correlate with information capacity, and explain binning noise (obscure but important).

Introduction

Traditional image quality measurements are based on several image quality factors, including 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 makes fine features more visible to the human eye, but it increases noise. Choices are often based on experience; they come down to what looks best, i.e., what has the most pleasing appearance for the application.

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

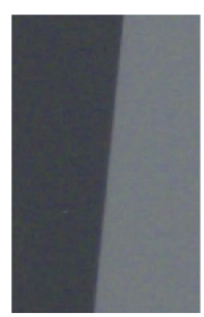
Information

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

The concept of information dates from 1948 and 49 in two celebrated papers by Claude Shannon [1],[2]. Appendix I contains a brief introduction to information theory. Earlier work on measuring information capacity from Siemens Star images [3] will only be briefly referenced in this document.

In electronic communications, information capacity is the maximum rate that information can be transmitted through a channel without error. In images, it is the maximum amount of information that a pixel or image can hold.

The slanted edge

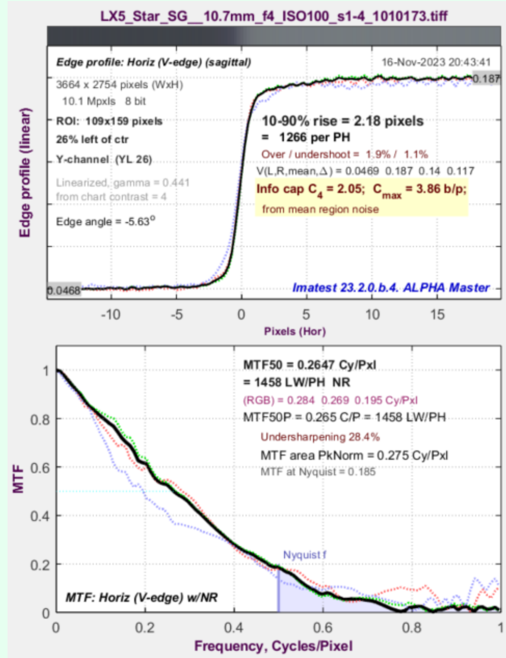


The slanted edge, which is a key part of the ISO 12233 standard, “Photography — Electronic still picture imaging — Resolution and spatial frequency responses” [4], is the most convenient and widely used resolution test pattern. It is highly efficient in its use of space (with multiple edges, sharpness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

Imatest offers several charts with multiple edges that can be automatically detected and rapidly analyzed. Some of the charts offer additional color, tone, noise, and distortion analysis.

Information capacity can be calculated from an overlooked capability of slanted-edge regions that was quite literally hidden in plain sight. To understand it, we start with a 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 signals in each bin to obtain the 4× oversampled averaged edge for the scan lines, $\mu_s(\mathrm{x}) = \frac{1}{L}\sum_{l=0}^{L-1} y_l(x)$.
7. **Modulation Transfer Function** $MTF(f)$ can be 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.



Combined mean $\mu_x(x)$ (upper); MTF(f) (lower)

The Edge Variance method

The Edge Variance method uses an overlooked capability of the ISO 12233 binning algorithm to calculate information capacity.

By summing the squares of each scan line, $\rho_s(x) = \frac{1}{L}\sum_{l=0}^{L-1} y_l^2(x)$, we calculate the edge variance (the spatially dependent noise power) $\sigma_s^2(x) = N(x)$ and noise amplitude $\sigma_s(x)$ in addition to the mean, $\mu_s(x)$.

Edge variance $\sigma_s^2(x)$ and noise amplitude $\sigma_s(x)$ are calculated from $\sum y_i(x)$ and $\sum y_i^2(x)$.

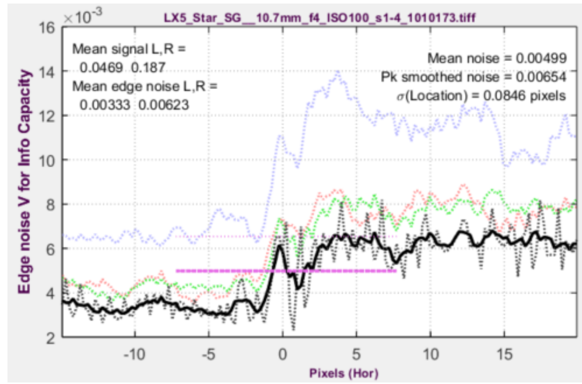
$$N(x) = \sigma_s^2(x) = \frac{1}{L}\sum_{l=0}^{L-1}\left(y_l(x) - \mu_s(x)\right)^2 = \frac{1}{L}\sum_{l=0}^{L-1} y_l^2(x) - \left(\frac{1}{L}\sum_{l=0}^{L-1} y_l(x)\right)^2 = \rho_s(x) - \mu_s^2(x)$$

Signal and noise results

The 4× oversampled average edge, $\mu_s(x)$, is shown in the upper plot (in the green box, above). Information capacities are shown with a yellow background.

The noise amplitude (voltage), $\sigma_s(x)$, is shown on the right. The thick black line is the smoothed luminance channel.

$\sigma_s(x)$ plot is a new measurement: spatially dependent noise was previously difficult to observe.



Spatially dependent noise, calculated by the Edge Variance method

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, by entering appropriate values of the signal and noise power, $S(f)$ and $N(f)$, into the **Shannon-Hartley equation**.

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

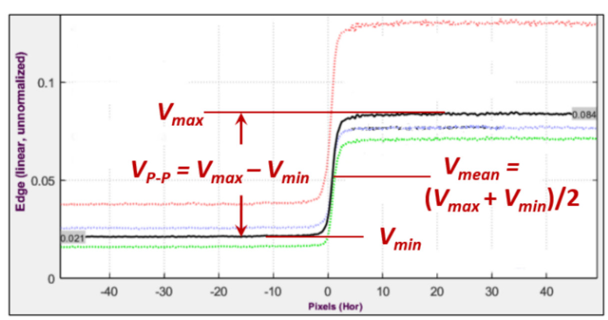
$S(f)$ and $N(f)$ are frequency-dependent signal and noise power, and W is the bandwidth, which is always equal to 0.5 cycles/pixel (the Nyquist frequency). Frequency-dependence is entered into $S(f)$ using $MTF(f)$ (described below).

> 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 only suitable for uniformly or minimally processed images; it should not be used 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 $V(x) = \mu_s(x)$.

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



Signal amplitude from slanted edge

The signal power is the *variance* of this signal. For calculating C, we assume a uniform distribution between the limits V_{max} and V_{min}, which maximizes information capacity, noting 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, $S_{avg}(f)$.

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

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. (It may be less in systems that limit signal levels.)

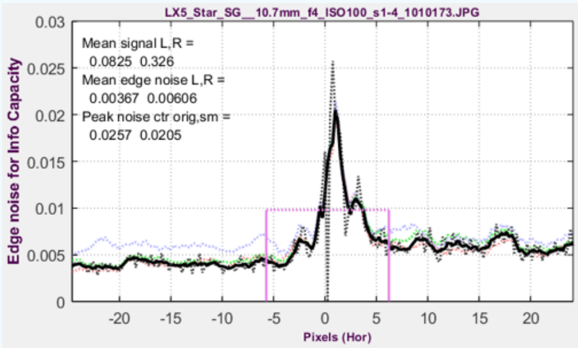
Noise power N

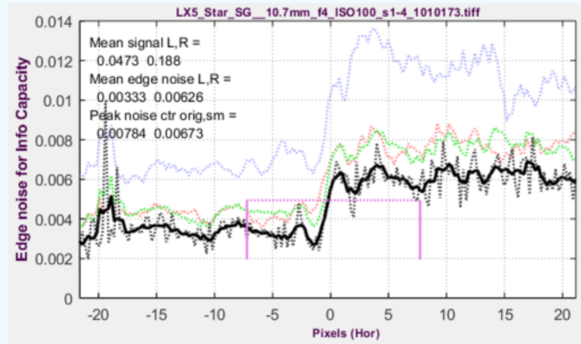
Noise power N has the same units as signal power S; hence S/N is dimensionless.

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 applications, uniformly/minimally processed images are preferred.

<table>
<tr><th colspan="2">The two image types: Plots of $\sigma_s(x)$ (4× oversampled)</th></tr>
<tr><th>Bilateral-filtered image</th><th>Minimally (i.e., uniformly) processed image</th></tr>
<tr><td>Sharpened near the edge; usually noise-reduced elsewhere. Nearly universal in consumer camera JPEG images. Image processing appears to increase information capacity C, even though information is removed. For this reason, it is important to use the peak noise $\sigma_s^2(x)$ (as described below) to calculate C.</td><td>Converted from raw with an external raw converter, with no sharpening or noise reduction.</td></tr>
<tr><td rowspan="2">A strong $\sigma_s(x)$ peak is visible near the edge transition. (This peak below is stronger than usual.)</td><td>Little or no peak is visible in $\sigma_s(x)$. Noise increases on the right because noise power is proportional to signal power (the mean number of photons striking each pixel) for linear sensors.</td></tr>
<tr><td>Strongly sharpened images can have a moderate amplitude peak, probably because the binning algorithm, which bins scan lines based on the polynomial fit to the centers (but not the centers themselves) is designed to optimize the MTF calculation but not the edge noise.</td></tr>
<tr><td>For calculating C,
N is the peak noise power, smoothed with a rectangular kernel of length PW20/2.</td><td>For calculating C,
$N = \text{mean}\left(\sigma_s^2(x)\right)$ for all values of x in the ROI.</td></tr>
</table>

<table>
<tr><td>

Noise σ_s(x) for bilateral-filtered image</td>
<td>

Noise σ_s(x) for uniformly processed image</td></tr>
<tr><td colspan="2">For both images, the solid line is the smoothed noise amplitude, $\sigma_s(x)$</td></tr>
<tr><td>Avoid for evaluating cameras for MV/AI systems.</td><td>Recommended, if available.</td></tr>
<tr><td>Texture: 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.</td><td>Texture: 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.</td></tr>
</table>

Bilateral-filtered images are of interest because we often measure "black box" cameras, where we don't know whether bilateral filtering is present, but we want to obtain a reasonable estimate 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 standard MTF measurements. It is described in Appendix 5, 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 C

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

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

$MTF(f)$ 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 that the camera is operating in its linear region. 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 $V_{P\text{-}P}$ 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.

Measurements of C_4 from a variety of exposures make it clear 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

C_4 is strongly dependent on exposure because (1) voltage range $\Delta V = V_{p\text{-}p}$ is a strong function of exposure, and (2) noise power N is also a function of exposure (derived from image sensor properties).

We have developed a 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.

Calculating maximum information capacity, C_{max}

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 Signal Power S,

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). For linear (non-HDR) image sensors, the general equation for noise power N as a function of V is

$$N(V) = k_0 + k_1 V$$

k_0 is the coefficient for constant noise (dark current noise, Johnson (electronic) noise, etc.). k_1 is the coefficient for photon shot noise. They are calculated from noise powers $N_1 = \sigma_1^2$ and $N_2 = \sigma_2^2$, which are measured along with signal voltages on the darker and lighter sides of the edge transition.

Assuming $N_1 = k_0 + k_1 V_{min}$ and $N_2 = k_0 + k_1 V_{max}$ where $N_1 < N_2$ for linear image sensors, we can solve two equations in two unknowns for k_0 and k_1.

$$k_0 = \frac{N_1 V_{max} - N_2 V_{min}}{V_{max} - V_{1min}} \;;\quad k_1 = \frac{N_2 - N_1}{V_{max} - V_{min}}$$

N closely approximates the noise used in noise calculation method (1) (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 be affect N_1 and N_2 strongly enough so the equation $N(V) = k_0 + k_1 V$ does not reliably hold. This can adversely affect the accuracy of C_{max}.

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

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

To handle cases where noise is not lighter on the light side of the edge, which can happen with HDR image sensors or with weird image processing, use

$$N_{mean} = \max(N_{mean}, N_{min}, N_{max})$$

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

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

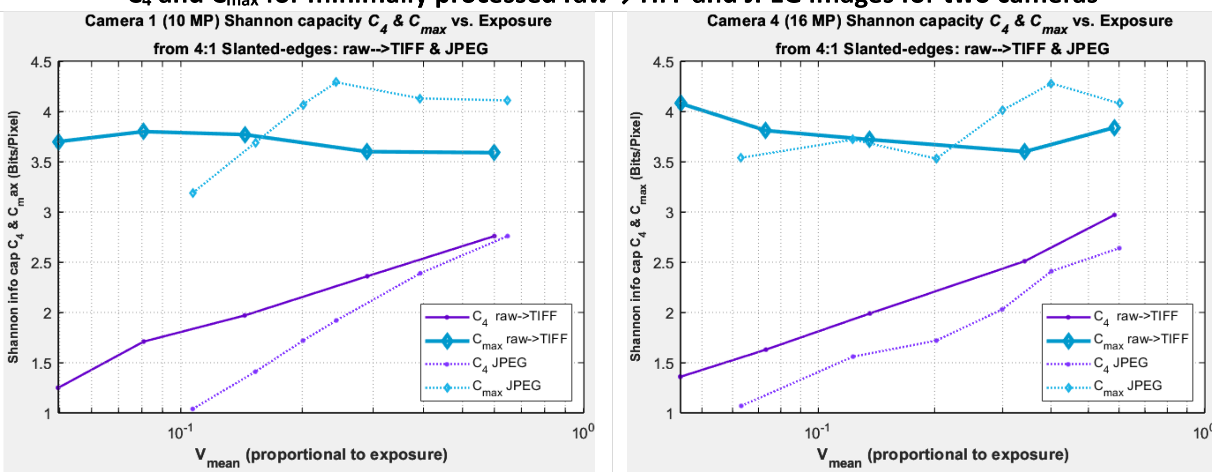
Because noise in High Dynamic Range (HDR) sensors does not follow the simple equation for linear sensors, we recommend giving the image sufficient exposure so the brighter side of the edge is near (but definitely below) saturation, then, if the noise does not increase with exposure, use N_{mean}, as indicated above.

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— especially when cameras are being evaluated for use in MV/AI systems.

C_4 and C_{max} for minimally processed raw→TIFF and JPEG images for two cameras



10 MP compact camera	***16 MP Micro Four-Thirds camera***

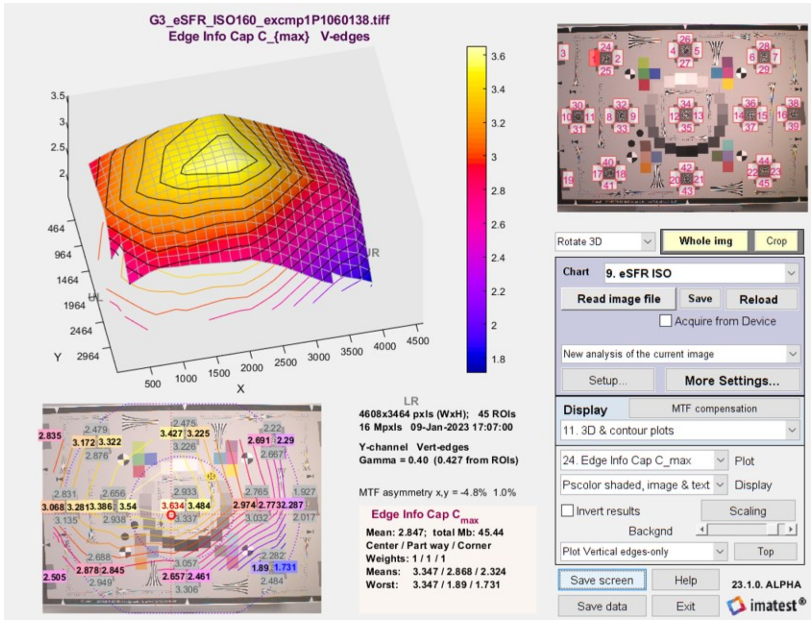
C_{max} may be need to be adjusted if the image is incapable of spanning the entire range of Digital Numbers (DNs), for example, 0-255 for images with bit depth = 8. Information capacity measurements fail if local tone mapping has been applied.

Total information capacity

Thus far, we have presented 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, or 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].

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



3D plot of information Capacity C_{max} over the image surface

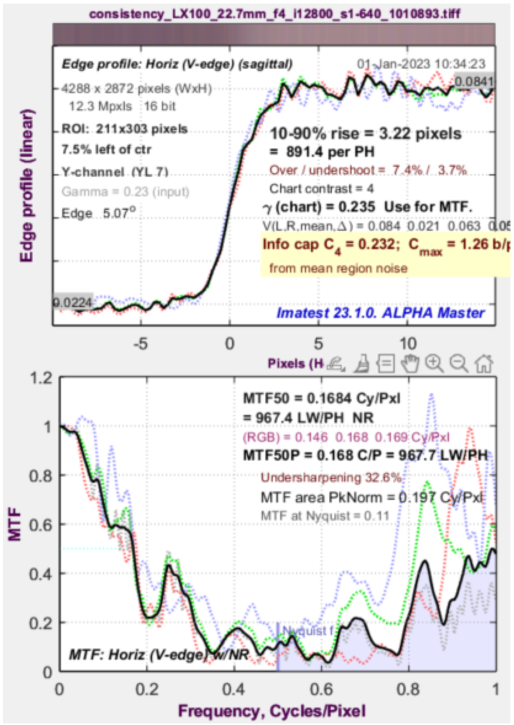
The mean information capacity C_{max} is 2.847 bits/pixel. Since the camera has 16 Megapixels, total capacity $C_{maxTotal}$ for the Luminance (Y) channel = 45.44 MB.

Signal averaging

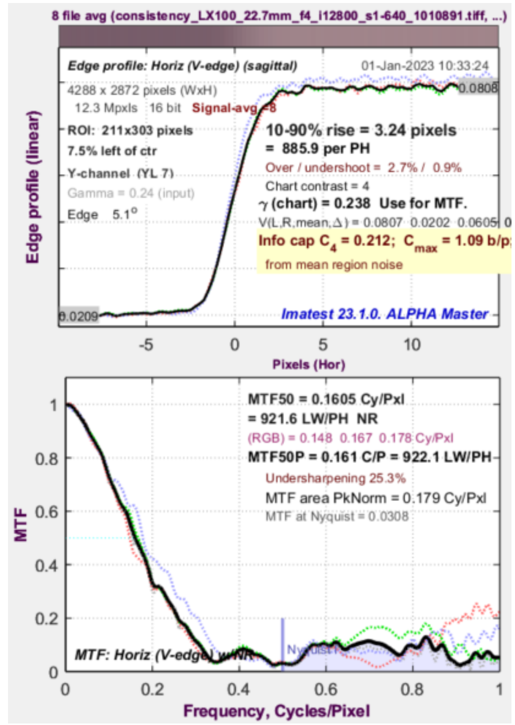
Signal averaging is a well-known technique that can improve the accuracy and consistency for measurements of noisy images for both the Edge Variance and Noise Image methods.

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 obtaining more consistent measurements by reducing the effect of uncorrelated noise. When n images are averaged, the sum of the signal voltage and the sum of the noise *power* (noise voltage^2), which is uncorrelated, are both proportional to n. This causes noise amplitude 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.

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.



Single image



n = 8 averaged

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	**2.14 µm pixel pitch.** An older (2010) compact 10.1-megapixel camera with a Leica f/2 zoom lens set to f/4.
2.	Sony A6000	**3.88 µm pixel pitch.** A 24-megapixel micro four-thirds camera with a 60mm Canon macro lens set to f/8
3.	Sony A7Rii	**4.5 µm pixel pitch.** 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

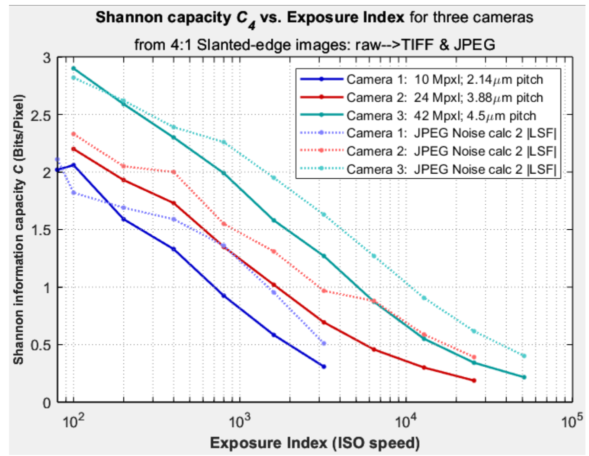
We captured both JPEG and raw images, which were converted to 24-bit sRGB (encoding gamma ≅ 1/2.2) TIFF images (designated as raw→TIFF) with LibRaw, with minimal processing (defined as no sharpening, no noise reduction, and a simple gamma-encoding function). Results for 48-bit Adobe sRGB conversion were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

The 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 C_{max} or C measured on Siemens stars. It is also highly sensitive to exposure.

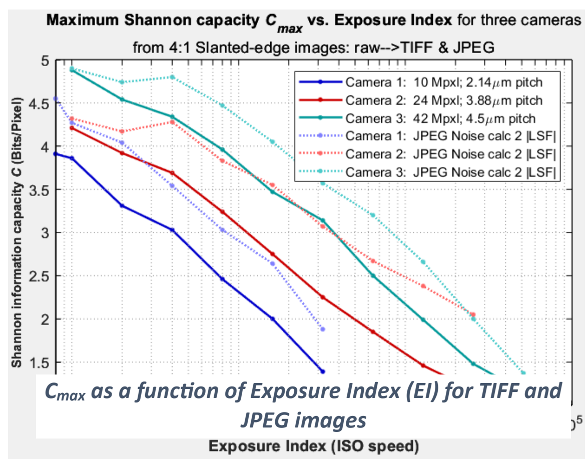


C_4 as a function of Exposure Index (EI) for TIFF and JPEG images

C_{max} maximum information capacity

C_{max} is 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).

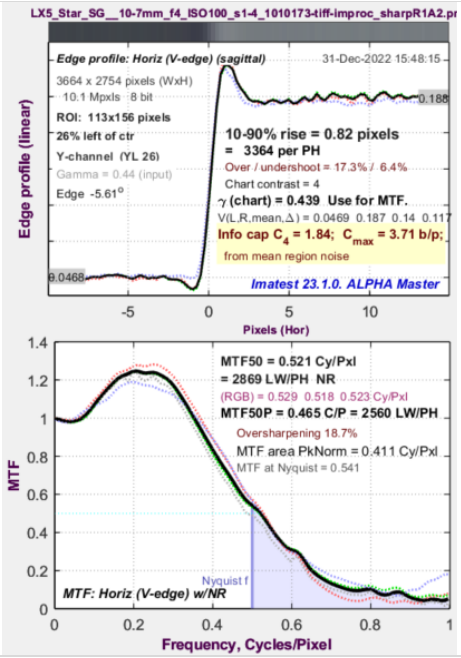


C_{max} as a function of Exposure Index (EI) for TIFF and JPEG images

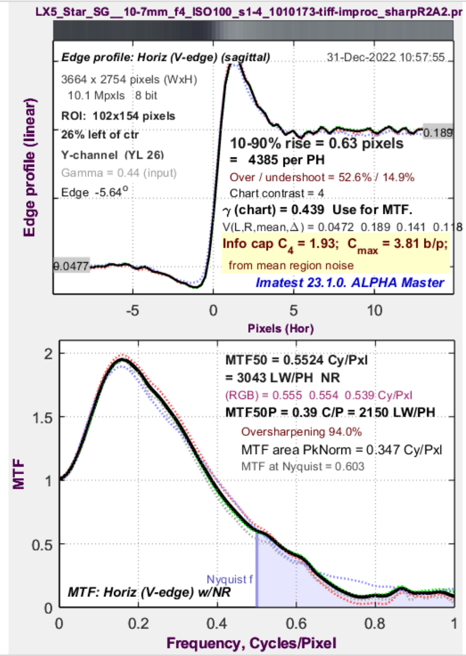
Sharpening

Simple sharpening, which has the same effect on the signal and noise response, and therefore does not change $S(f)/N(f)$, would not be expected to have much effect on C. This is indeed the case.

The two examples below show that (uniform) USM sharpening has little effect on slanted-edge information capacity. 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.



TIFF sharpened with Radius = 1, Amount = 2.



TIFF sharpened with Radius = 2, Amount = 2.

For R=1, A=2, $\sigma(\text{Location})$ = 0.11 pixels. For R=2, A=2 (stronger sharpening), $\sigma(\text{Location})$ = 0.108 pixels. This is a relatively small increase over the unsharpened $\sigma(\text{Location})$ = 0.0846 pixels.

This highlights another benefit of information capacity measurements. Unlike MTF50, they do not reward excessive sharpening, which creates "halos" near edges, making the image look sharp in small displays, but creating artifacts that degrade image appearance on large displays [9]. They also have a bad reputation for machine vision applications.

Edge location variance (or standard deviation)

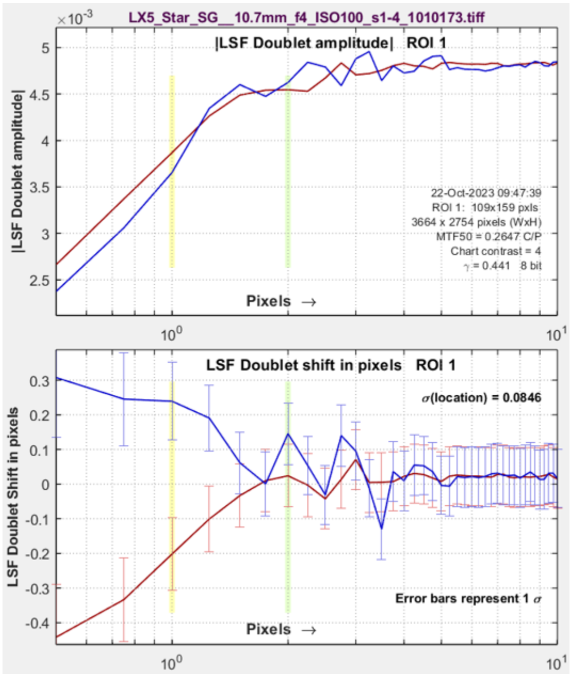
An additional result can be derived from the Edge Variance method: ***The edge location variance (or standard deviation),*** $\sigma^2(\text{Location})$ or $\sigma(\text{Location})$.

Edges are important because they are often required to distinguish an object. For example, athe only way to distinguish a gray vehicle from a gray concrete background is with the edges.

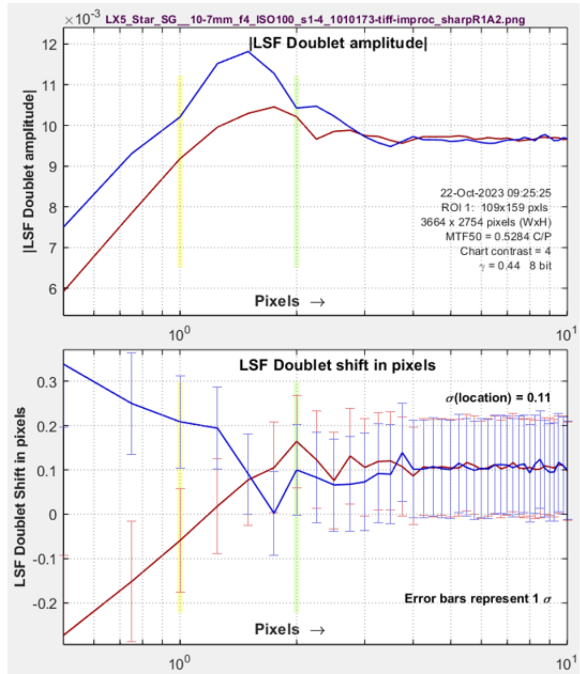
For signal voltage $V(x)$, the edge is defined as the location x where the Line Spread Function $LSF(x)$ = $dV(x)/dx$ (in units of 1/pixels) has its peak value. The standard deviation of the edge location is

$$\sigma(\text{Location}) = \frac{\text{maximum } \sigma_x(x)}{\text{maximum } dV(x)/dx} \cong \frac{\sigma_x(x) \text{ at peak } LSF(x)}{\text{peak } LSF(x)} \qquad \text{in units of pixels}$$

The actual location of an edge is affected by interference from neighboring edges (mostly the closest edge) as well as $\sigma(\text{Location})$. When edges are close together (small w), interference causes edge amplitudes decrease, which increases sigma, and it also causes edge locations to shift. Displays of LSF amplitude and shift versus spacing are shown below for the image used for Signal and noise result (above), and the same image moderately sharpened with Radius = 1 and Amount = 2, in Sharpening (above). $\sigma(\text{Location})$ is not a major metric. *SNRi* and *Edge SNRi* are more useful.



TIFF unsharpened



Sharpened with Radius = 1, Amount = 2.

The difference between the two results is not large. At a pixel spacing of 1 (labeled 10^0 on the x-axis), the difference between $\Delta LSF_{maximum}$ + $\sigma(\text{Location})$ and $\Delta LSF_{minimum}$ - $\sigma(\text{Location})$ is 0.36+0.31 = 0.67 for the TIFF unsharpened image and 0.32+0.17 = 0.49 for the R1A2 sharpened image: a modest improvement. For a strongly oversharpened (R2A5) image, $\sigma(\text{Location})$ increases to 0.125, but ΔLSF_{total} @ pixel spacing 1 = 0.07+0.34 = 0.43. This improvement was surprising since the edge has a large peak. At higher ISO speeds, $\sigma(\mathit{Location})$ would have been much larger and there would be less improvement with sharpening. We will present better performance metrics, *SNRi* and *Edge SNRi*, below.

Summary of the Edge Variance method

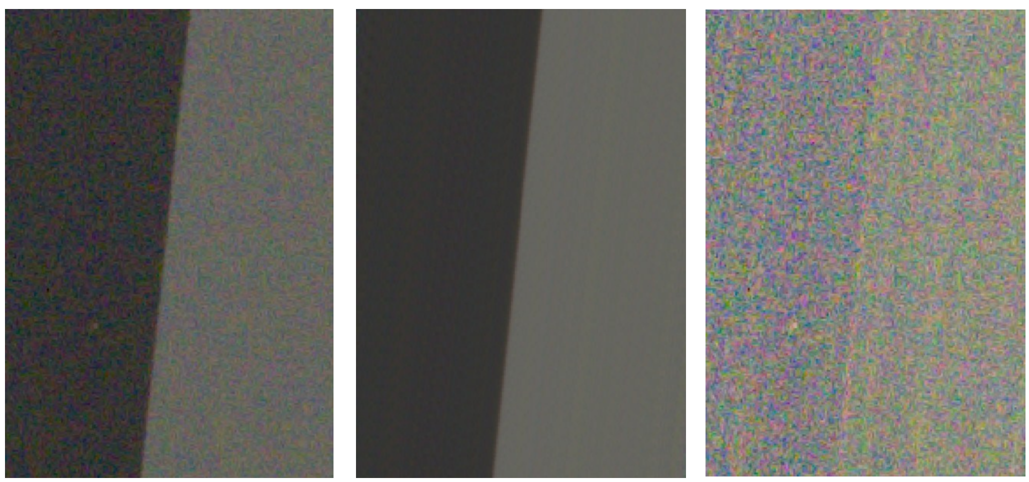
- The Edge Variance method is the first of two methods for calculating information capacity, C, from slanted edges.
- It has a limited set of results. There are many more in the Noise Image method.
  - Information capacities C_4 and C_{max} and $\sigma(\text{Location})$,
  - A plot of spatially dependent 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 a useful approximate measurement of C for bilateral-filtered images, but more accurate results are obtained from uniformly/minimally processed images, which should *always* be used when a camera is being evaluated for use in MV/AI systems.
- Results are easy to obtain, even though the algorithms behind them can be complex. For the most part, ***Imatest*** users don'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 metrics. It was developed shortly after the Edge Variance method, and 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).

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.



(1) Original ROI ***(2) Inverse-binned / de-interleaved / reverse-projected*** ***(3) Noise image ROI***

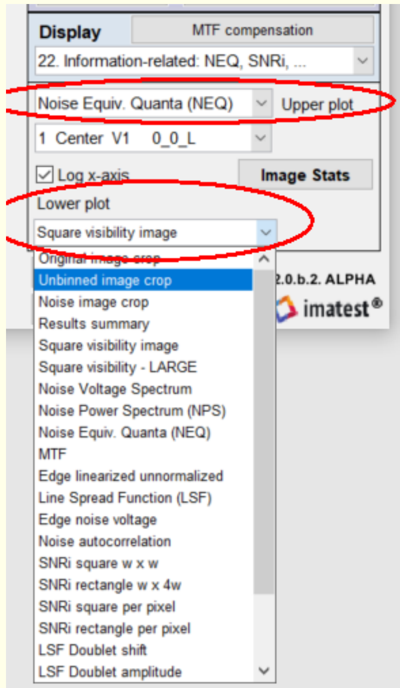
These images allow several additional image quality parameters to be calculated, including Noise Power Spectrum (*NPS*) and Noise Equivalent Quanta (*NEQ*), well-known in medical imaging systems, and described in an excellent review paper by Ian Cunningham and Rodney Shaw [10]. These measurements are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

An alternative information capacity measurement, C_{NEQ}, derived from NEQ, is described below.

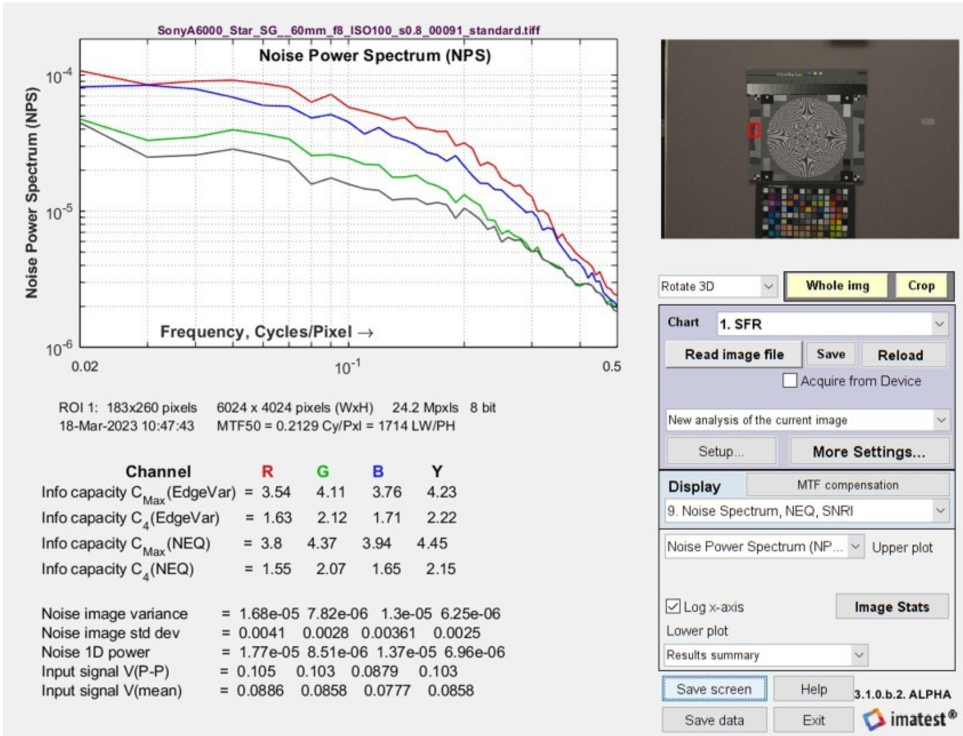
Displaying the results

The key results are in the **Information-related, NEQ, SNRi,…** plot of the ***Imatest*** **Rescharts** window (used to run SFRplus, eSFR ISO, SFRreg, and Checkerboard interactively).

This plot has numerous display options. It 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 column of the table below.

Upper or Lower plot	Display settings	Lower plot-only
Noise Voltage Spectrum Noise Power Spectrum (NPS) Noise Equivalent Quanta (NEQ) MTF Edge linearized unnormalized Line Spread Function (LSF) Edge noise voltage Noise autocorrelation SNRi square *w* x *w* SNRi rectangle *w* x 4*w* SNRi square per pixel[2] SNRi rectangle per pixel[2] LSF Doublet shift LSF Doublet amplitude LSF doublet S/N energy Edge SNRi square w x w Edge SNRi rectangle w x 4w Edge SNRi 1D doublet Object matched filter Edge matched filter		Original image crop Unbinned image crop (Reverse-projected; low noise) Noise image crop (Original – Noise) Results summary (Shown above) Square visibility image Square visibility – LARGE

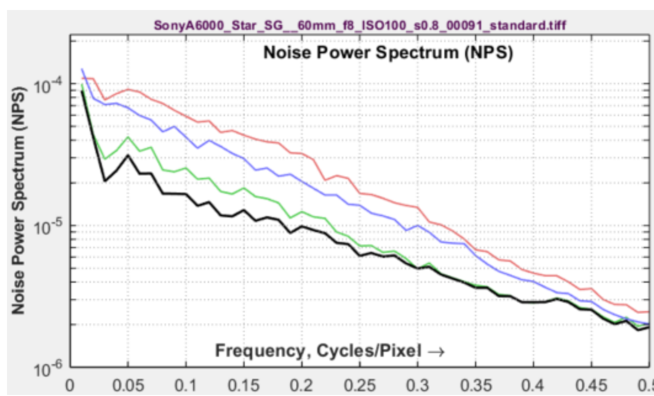
Here is an example, with Noise Power Spectrum (NPS) displayed on the top and Results summary displayed on the bottom.



Noise Power Spectrum (NPS) displayed on the top and Results summary (showing the two different information capacity calculations) on the bottom.

Noise Voltage or Power Spectrum (NPS)

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



Noise Power Spectrum (NPS)

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 for the NEQ and $SNRi$ calculations.

Because this procedure does not maintain the invariance in energy between the spatial and frequency domain implied by Parseval's theorem, $\iint \sigma^2(x,y)\, dx\, dy = \iint NPS(v_x, v_x)\, dv_x\, dv_x$, where v is frequency, we must apply a correction to the NPS.

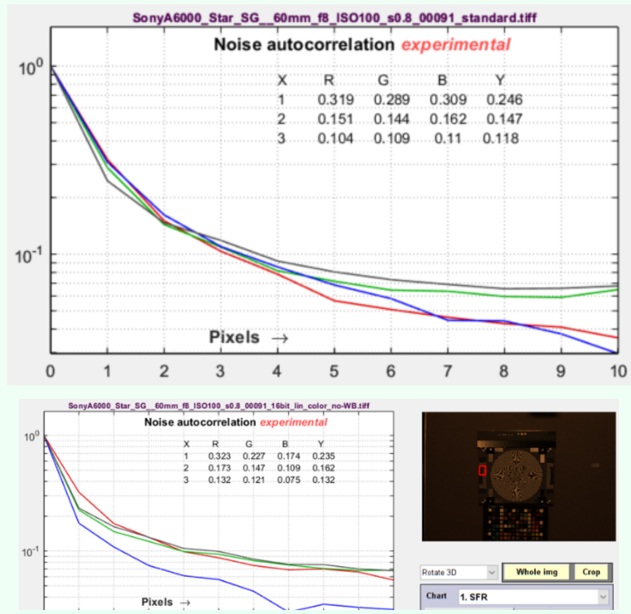
Noise autocorrelation

This plot is used to examine the hypothesis that noise autocorrelation (the inverse Fourier transform (IFT) of the NPS) indicates 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 uncorrelated, so that if there were no crosstalk, the noise would be white.

This image used for the upper plot was white-balanced. The curve of |IFT(NPS)| is based on the author's understanding of the Wiener-Khinchin theorem.

The image in the lower right was not white-balanced. This increases the red channel autocorrelation distance, as expected.



Noise autocorrelation

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

Noise Equivalent Quanta (NEQ)

NEQ is a figure of merit used in medical imaging [5], but is unfamiliar in general imaging. It is described in a 2016 paper by Brian Keelan 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.

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

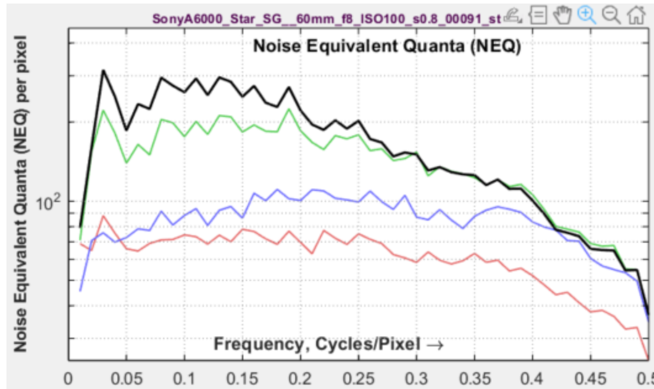
where the mean linear signal, μ, can be defined in either of two ways, depending on how NEQ is to be interpreted.

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 per pixel (assuming photon shot noise is dominant).

The above equation, $\mu = V_{mean} = \bar{q}$, is appropriate if NEQ is to be used for calculating Detective Quantum Efficiency), $DQE(f) = NEQ(f)/\bar{q}_i$, where $\bar{q}_i$ is the mean number of quanta *incident* on each pixel. Measuring *DQE* requires a separate (and very exacting) measurement of $\bar{q}_i$. It is not yet in ***Imatest***.

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) [17] by Tang et. al. is an excellent example of how NEQ is used in medical imaging.

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



Noise Equivalent Quanta (NEQ)

Information capacity from NEQ, C_{NEQ}

A variant 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}.

$$C_{NEQ} = \int_0^W \log_2(1 + NEQ_{info}(f))\, df = \int_0^{0.5} \log_2\left(1 + \frac{\mu^2 MTF^2(f)}{NPS(f)}\right) df$$

where bandwidth $W = f_{Nyq} = 0.5$ Cycles/Pixel, is the camera's Nyquist frequency. [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, i.e., white noise.

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

Results summary

Ideal Observer SNR ($SNRi$)

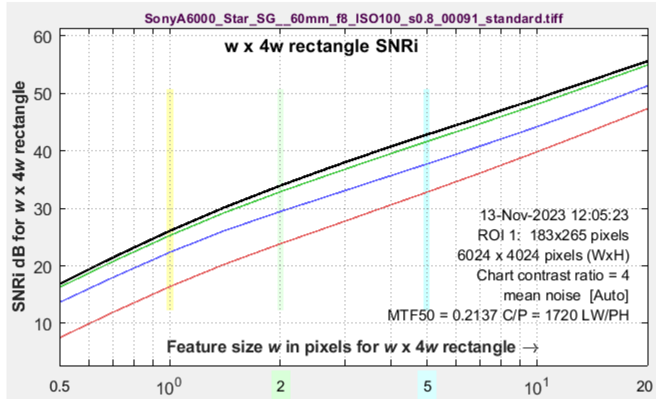
$SNRi$ is a measure of the detectability of objects, described in ICRU Report 54 [16] and in papers by Paul Kane [14] and Orit Skorka and Paul Kane [15]. The two-dimensional equation in [15] gives the best results.

$$SNRi^2 = \iint \left(\frac{\left|G(\nu_x, \nu_y)\right|^2 MTF^2(\nu_x, \nu_y)}{NPS(\nu_x, \nu_y)}\right) d\nu_x\, d\nu_y$$

$G(\nu_x, \nu_y)$ is the Fourier transform of the rectangular object to be detected, defined below.

$MTF(\nu)$ and $NPS(\nu)$ are defined in one dimension. Spatial frequency $\nu = \sqrt{\nu_x^2 + \nu_y^2}$ has units of Cycles/Pixel.

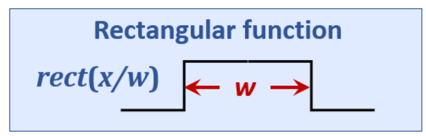
Objects to be detected are typically rectangles of dimensions $w \times kw$, where k = 1 for a square or 4 for a 1:4 aspect ratio rectangle.Amplitude, $V_{P\text{–}P}$, is typically obtained from a chart with a 4:1 contrast ratio.



SNRi for sharp, low-noise (ISO 100) image

$$g(x,y) = V_{P-P} \cdot \text{rect}\left(\frac{x}{w}\right) \cdot \text{rect}\left(\frac{y}{kw}\right)$$

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



$G(\nu_x,\nu_y)$ is the Fourier transform of the object, $g(x,y)$, expressed in two dimensions.

$$G(\nu_x, \nu_y) = kw^2\, V_{P-P} \frac{\sin(\pi w \nu_x)}{\pi w \nu_x} \frac{\sin(\pi k w \nu_y)}{\pi k w \nu_y} = V_{P-P}\, G_{rect}(\nu_x, \nu_y)$$

where $G_{rect} = w\,\text{sinc}(\omega w/2) = w\,\text{sinc}(\pi w \nu)$ is the Fourier transform of $rect(x/w)$ for frequency ν. Note that G has units of $1/\nu^2$, and since ν has units of cycles/pixel, G has units of pixels^2.

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

$$SNRi^2 = \Delta\nu_x\, \Delta\nu_y \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{MTF^2(i,j)\, G^2(i,j)}{NPS(i,j)}$$

$SNRi$ is displayed for each color channel for w = 0.5, 0.7, 1, 1.4, 2, 3, 4, 7, 10, 14, 20.

Note that like C_4, $SNRi$ is strongly affected by exposure and chart contrast. But unlike C, $SNRi$ is affected by image signal processing (sharpening, etc.).

Although $SNRi$ is a powerful measurement, we currently prefer a closely related measurement, ***Edge SNRi***, for determining the performance of pre-filtering (Image Signal Processing performed before sending the image to the Object Recognition/Machine Vision/AI block).

***SNRI²* is equivalent to the total (integrated) Signal/Noise energy of the object in the spatial domain.** This is best illustrated in one dimension, using Parseval's theorem.

$$\int_{-\infty}^{\infty} |r(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |R(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |R(2\pi f)|^2 df$$

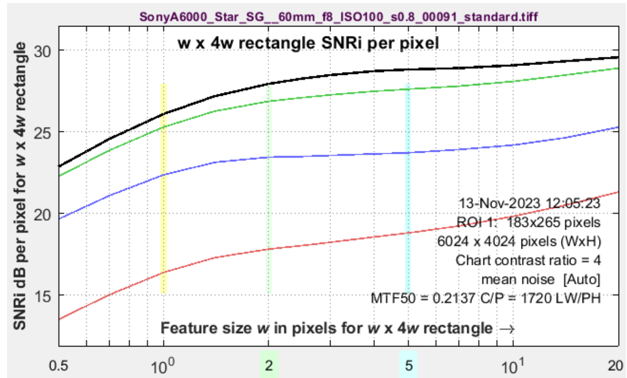
where

$$r(x) = (V(x) - V(x-w))/\sigma \quad \text{and} \quad R(f) = \frac{G(f)\, MTF(f)}{\sqrt{NPS(f)}}$$

SNRi displayed in dB per pixel²

Because standard *SNRi* plots can be difficult to read (in part because *SNRi* has units of pixels²), we have added a plot of *SNRi* in dB per pixel², shown on the right. It is somewhat easier to read than the standard *SNRi* plot, but it is more of a *relative* measurement—for evaluating changes from image processing.

Tip— Click on Data cursor in the dropdown below the thumbnail on the upper right to get a reading of the actual value.

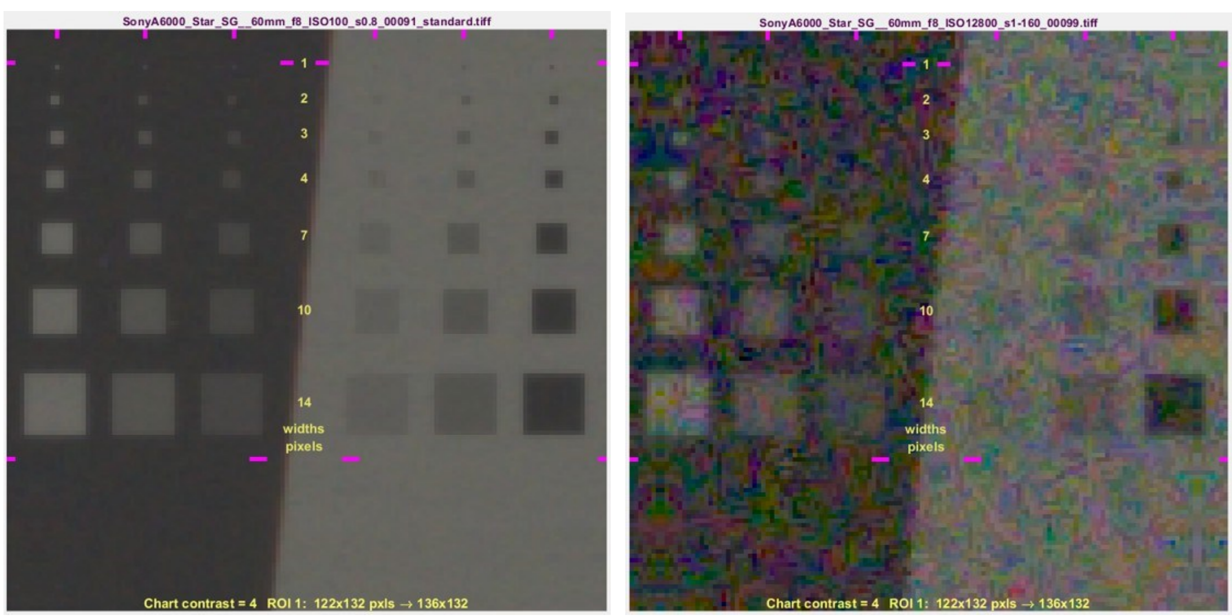


SNRI in dB per pixel² for low-noise (ISO 100) image

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.

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



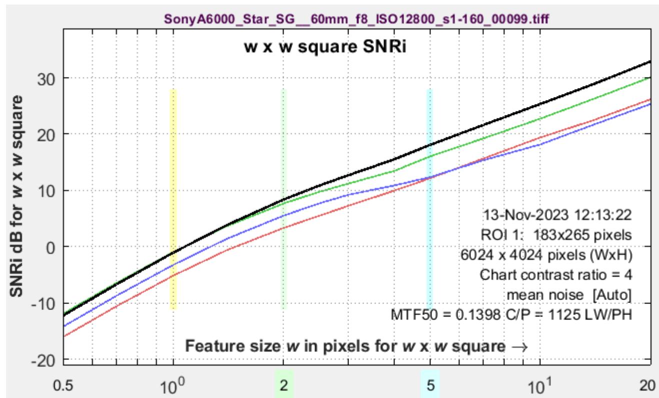
Low noise ISO 100 (left) **Noisy ISO 12800 (right)**

***MTF50* = 0.214 c/p; C_{max} = 4.24 b/p;** ***MTF50* = 0.140 c/p; C_{max} = 1.37 b/p.**

We show two images, above: one for a relatively low noise image and one for a noisy image (both from a camera with a Micro Four-Thirds sensor, 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, $C_{Mich} = \frac{\text{light-dark}}{\text{light+dark}} = 0.6$. 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.

The $SNRi$ 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, the Y (luminance) channel $SNRi$ at w = 2 is 9 dB; it reaches 11 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 of squares, but we used a little smoke and mirrors to make the squares that have the same blur as the device under test. Feel free to skip this explanation.

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×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.

Edge Signal-to-Noise Ratio (*Edge SNRi*)

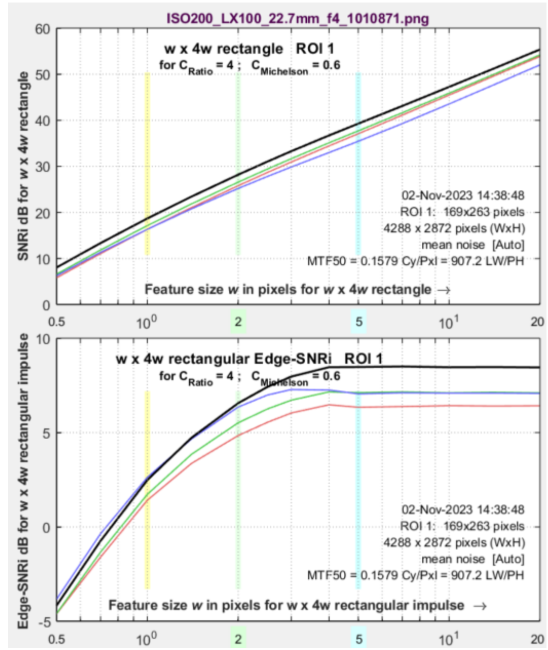
Edge SNRi is an edge-based measure of the detectability of the edges of small objects, similar to *SNRi*, described above and in papers by Paul Kane [14] and Orit Skorka and Paul Kane [15].

$$Edge\ SNRi^2 = \iint \left(\frac{\left|H(\nu_x,\nu_y)\right|^2 MTF^2(\nu_x,\nu_y)}{NPS(\nu_x,\nu_y)} \right) d\nu_x\, d\nu_y$$

$H(\nu_x,\nu_y)$ is the Fourier transform of the *edges* (the gradient) of the object to be detected; defined below.

For a rectangle of dimensions $w \times kw$, the function is the derivative, $h(x,y)$, of the rectangle, $g(x,y)$, that describes the object.

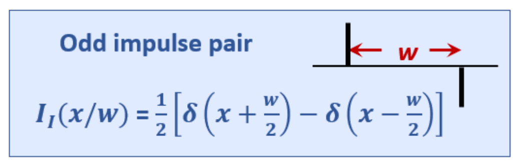
V_{P-P} is typically obtained from a chart with a 4:1 contrast ratio.



SNRI curve (Upper), Edge SNRi curve (lower), 1" sensor raw-converted with minimal processing, ISO 200

$$h(x,y) = V_{P-P} \cdot d\left[\text{rect}\left(\frac{x}{w}\right)\right]/dx \cdot d\left[\text{rect}\left(\frac{y}{kw}\right)\right]/dy \;=\; V_{P-P} \cdot I_I\left(\frac{x}{w}\right) \cdot I_I\left(\frac{y}{kw}\right)$$

where $I_I(x/w) = d(\text{rect}(x/w)/dx$ is called the "odd impulse pair," consisting of a pair of Dirac delta functions of opposite polarity separated by the object width w. It is shown on the right.



$H(\nu_x,\nu_y)$ is the Fourier transform of the edges of the object to be detected, equivalent to $2\pi\nu\, G(\nu_x,\nu_y)$ for frequency ν. Expressed in two dimensions,

$$H(\nu_x,\nu_y) = 2\, V_{P-P} \sin(\pi w \nu_x) \sin(\pi k w \nu_y)$$

Edge SNRi^2 is calculated using a similar equation to standard *SNRi*^{2}.

$$Edge\ SNRi^2 = \Delta\nu_x\, \Delta\nu_y \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{MTF^2(i,j)\, H^2(i,j)}{NPS(i,j)}$$

Edge SNRi is displayed for each color channel for w = 0.5, 0.7, 1, 1.4, 2, 3, 4, 7, 10, 14, 20.

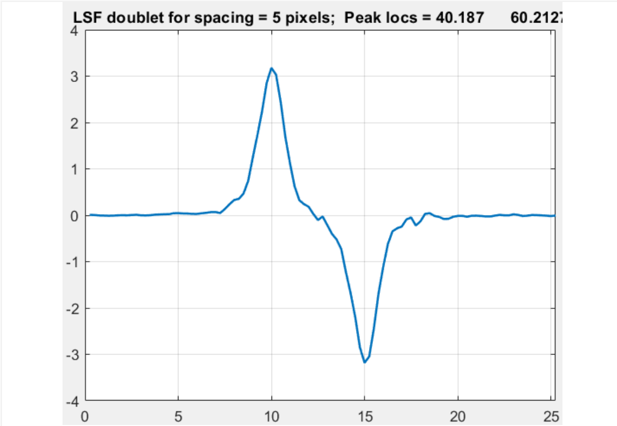
Unlike *C*, *Edge SNRi* is affected by signal processing (sharpening, etc.), making it useful for evaluating pre-filtering (ISP filtering applied prior to the object recognition/machine learning/AI blocks).

Line Spread Function (LSF) doublet results

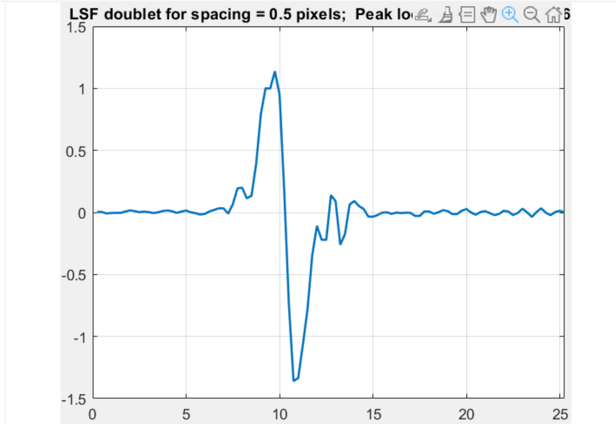
Edge SNRi is based on pairs of Line Spread Functions of opposite polarity called *LSF doublets*, $r(x)$, which are also used in several key calculations.

$$r(x) = (LSF(x) - LSF(x-w))/\sigma \quad \text{and} \quad R(\nu) = \frac{H(\nu)\, MTF(\nu)}{\sqrt{NPS(\nu)}}$$

LSF Doublets, $r(x)$, are illustrated below for w = 5.0 and 0.5 pixels.



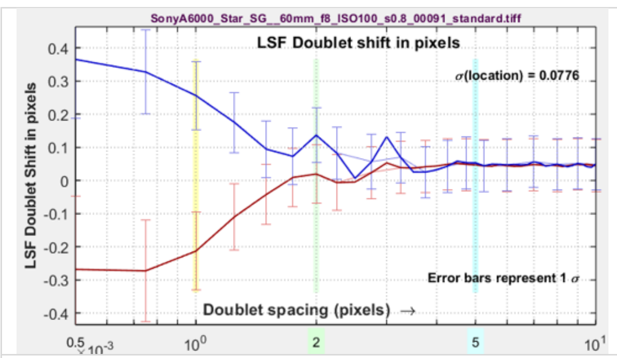
LSF Doublet. w = 5.0 pixels.



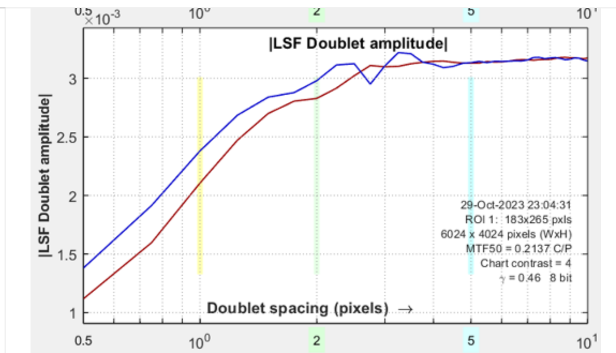
LSF Doublet. w = 0.5 pixels. Amplitude is 1/3 as large as for w = 5.0 pixels.

As spacing w decreases,

- the peaks are closer (but shifted more from their original locations), and
- amplitude decreases. These are plotted below.



LSF Doublet shift as a function of spacing w



LSF Doublet amplitudes as a function of spacing w

Edge SNRi in frequency and spatial domain

As a result of Parseval's theorem, ***Edge SNRI***[2]**, which is defined in frequency domain, is equivalent to the total (integrated) Line Spread Function doublet divided by Noise energy in the spatial domain.**

This is best illustrated in one dimension. For LSF doublets in both domains,

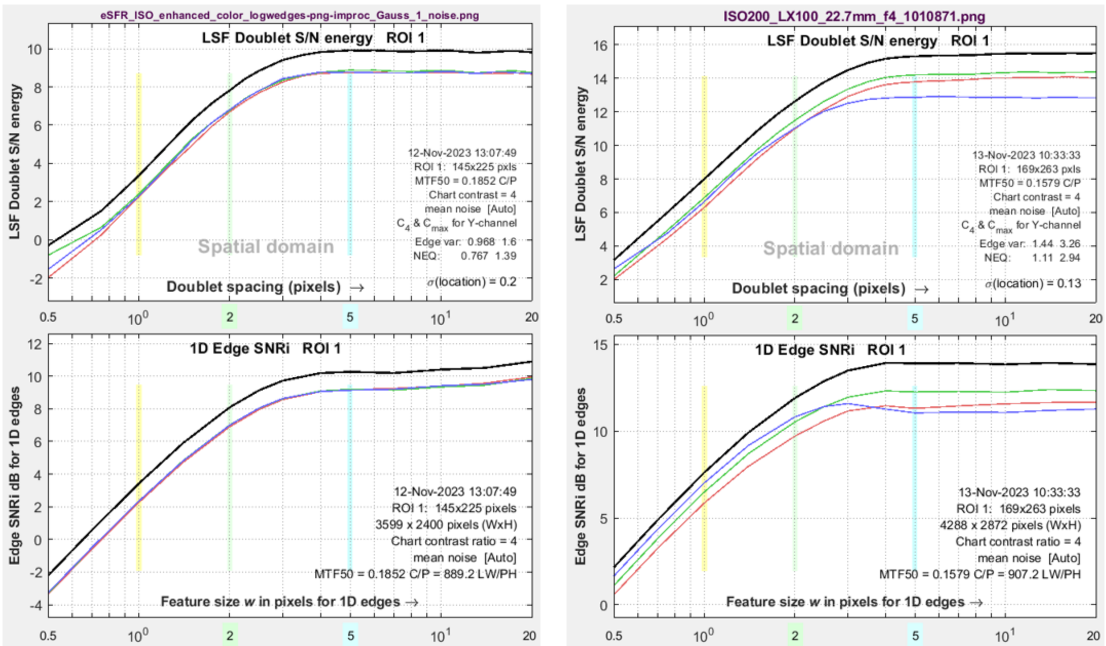
$$\int_{-\infty}^{\infty} |r(x)|^2 dx \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} |R(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |R(2\pi v)|^2 dv \cong 2\int_{0}^{fNyq} |R(2\pi v)|^2 dv$$

Because the integrands are energy, dB = 10 $\log_{10}$.

The upper plots below are the spatial domain result, $\int_{-\infty}^{\infty} |r(x)|^2 dx = \int_{-\infty}^{\infty} |LSF(x)/noise|^2 dx$ as a function of doublet spacing (feature size w).

The lower plots are the frequency domain results, $Edge\ SNRi = 2\int_0^{fNyq} |B(2\pi f)|^2 df$ as a function of w.

There is slightly more discrepancy between the lower and upper plots for the real image (on the right) because noise is not white, as assumed by the spatial domain calculation. In the real image on the right, the Noise Spectrum (Power or Voltage) falls off with frequency.



Simulated image with white noise *Real image with spectral noise*

Comparison of spatial and frequency-domain Edge SNRi measurements

Table of key measurements from the Noise Image method

Measurement	Description
Noise Power Spectrum, *NPS(f)* (or Noise Voltage Spectrum)	Used in NEQ and $SNRi$ calculations. NPS was implicitly assumed to be constant (white noise) in the Edge Variance method.
Noise autocorrelation	The inverse Fourier transform of the Noise Voltage Spectrum. May be related to sensor electrical crosstalk. An experimental measurement.
Noise Equivalent Quanta, $NEQ(f)$ and $NEQ_{info}(f)$	A measure of the frequency-dependent signal-to-noise ratio (SNR). $NEQ(f) = \mu^2\ MTF(f)^2/NPS(f)$, where $\mu = V_{mean}$ is relatively unfamiliar outside of medical radiology. $NEQ(f)$ is equivalent to the number of quanta detected by the sensor when photon shot noise is dominant. It can be used for calculating Digital Quantum Efficiency (DQE), when the density of quanta reaching the image sensor is known. $NEQ_{info}(f)$, derived from $\mu = V_{P-P}/\sqrt{12}$, is used to calculate information capacity C_{NEQ}.

Information capacities $C_4(NEQ)$ and $C_{\max}(NEQ)$	correspond to C_4 and $C_{\max}$ from the Edge Variance method. Derived from $NEQ_{info}(f)$. They are close, but not identical.
Ideal observer Signal-to-Noise Ratio, *SNRi*	From Kane [14] and Skorka and Kane [15], "The Ideal Observer is a Bayesian decision maker that maximizes the statistical precision of a hypothesis test with two possible outcomes." *SNRi* is a metric of the detectability of small objects (squares or rectangles), typically of low contrast. $SNRI^2$ is equivalent to the total (integrated) Signal/Noise energy of the object in the spatial domain.
Object visibility	Images of low contrast squares of various sizes: a visual indicator of object visibility. Correlates with *SNRi*.
Edge SNRi	Similar to *SNRi*, except that it is derived from the object edges, i.e., Line Spread Function doublets (pairs of LSFs representing the edges of the object). *Edge* $SNRI^2$ is equivalent to the total (integrated) Signal/Noise energy of the LSF doublet Signal/Noise energy in the spatial domain.

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 uses a 2D image of the noise to calculate several image quality metrics.
- 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 ($SNRi$), $Edge\ SNRi$, Noise Equivalent Quanta (NEQ), and a second information capacity measurement, derived from NEQ, that can be compared with the Edge Variance results (they are slightly more accurate because $NPS(f)$ is not assumed to be constant).

Image Signal Processing (ISP)

Several recent papers [18],[19],[20] state that appropriate image processing prior to Object Recognition, Machine Vision or AI algorithms may improve system performance (accuracy, speed, and power consumption). Because information capacity changes relatively little with Image Signal Processing— at least with ISP that does not remove information, such as Unsharp Mask (USM) sharpening— it provides little guidance about how to design optimal image processing.

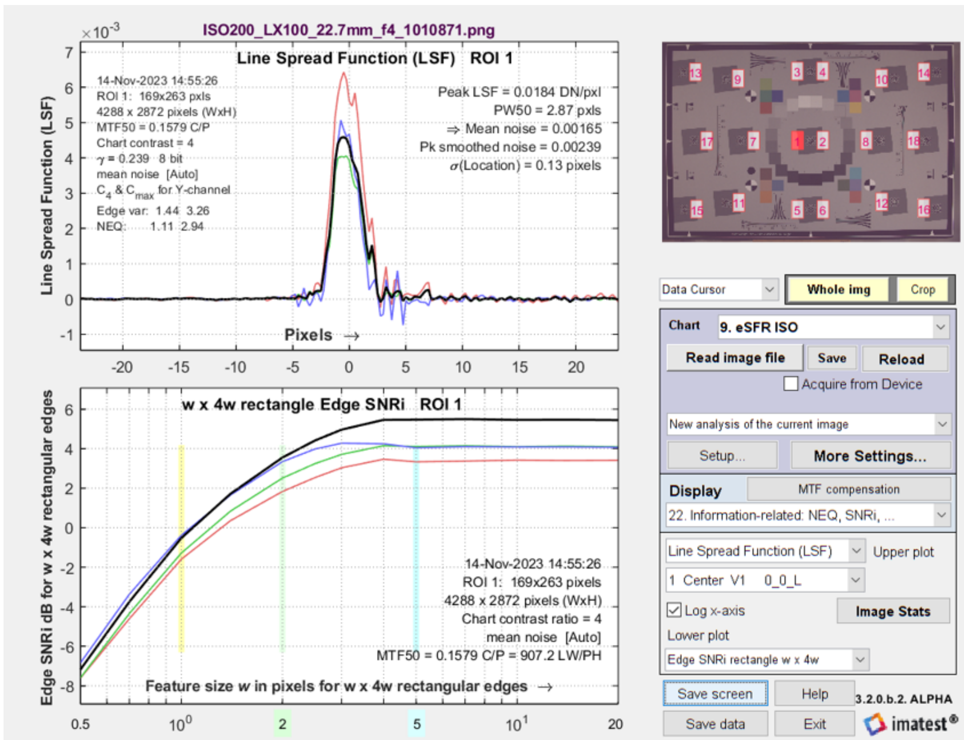
Image signal processing algorithms can be designed to optimize a specific task, for example, the detection of an object of a specific size, often a small rectangle, or its edges. In practice, ISP needs to perform well over a range of tasks: detecting objects of edges greater than a minimum size and limiting interference from neighboring objects.

SNRi has some drawbacks as an object detection metric. Plots of *SNRi* are challenging to interpret because *SNRi* increases with feature size. And there is the problem of object color. What If the object has the same color as the background (e.g., gray cars in front of gray concrete)? In such cases it is the *edge* that matters. Because of these shortcomings, we prefer ***Edge SNRi***.

Pre-filtering: effects of sharpening and lowpass filtering

Starting with an unsharpened image, we applied sharpening and/or lowpass filtering (blurring) using the [***Imatest*** **Image Processing**] module.

We show the entire ***Imatest*** **Rescharts** interactive window displaying the Line Spread Function and $Edge\ SNRi$ for a $w{\times}4w$ rectangle (as a function of the narrower edge, w). We selected smoothed peak noise for the calculations, which gives lower performance than the mean noise, but should be more representative of Line Spread Function edge performance. *Edge SNRi* is 5.3 dB for large *w*; -1 at *w* = 1.



Results with no filtering. LSF (top), Edge SNRi (bottom)

The first filter is USM (Unsharp Mask) sharpened with Radius = 2 and Amount = 3. This is moderately strong sharpening. R = 2 was chosen because the original image is not extremely sharp: R = 1 might be better for sharper image (with some MTF remaining at the Nyquist frequency).

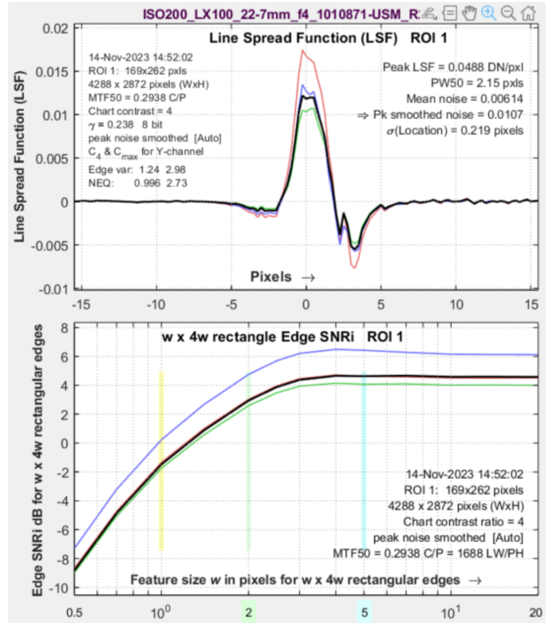
***Edge SNRi* (4.7 dB for large w; -1.5 at w = 1) is worse than for the unsharpened image at all spacings.** The reason is that sharpening increases noise, σ(location) is significantly larger (0.219 pixels) than for the unsharpened image (0.13 pixels)

As expected for sharpening, *PW50* (full width half maximum) is reduced and *MTF50* is increased. Information capacity C_4 and C_{max} is slightly lower (due to numerical calculations).

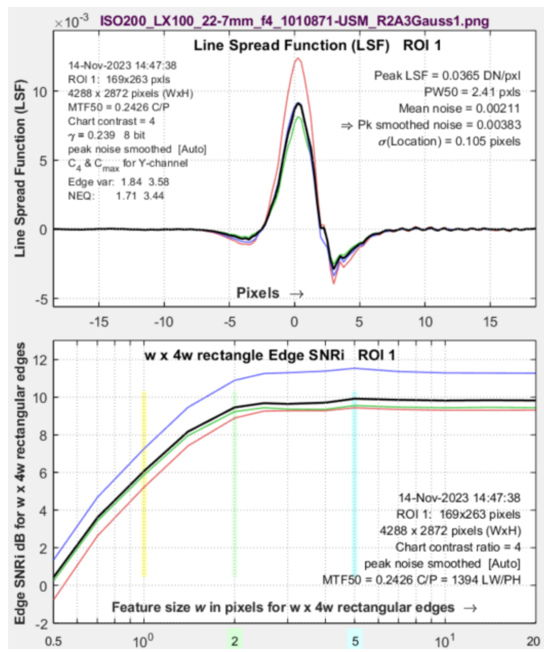
This result agrees with opinions I've heard (alas, I don't have good references) — that sharpening does not improve performance of machine vision systems. But it is not the end of the story.

The second filter was USM sharpened (with the same values: Radius = 2 and Amount = 3) with an added Gaussian lowpass filter with σ = 1 pixel (determined by old-fashioned trial and error). The filter was created with the ***Imatest*** Image Processing module, but an external program could have been used.

***Good news! Edge SNRi* (9.8 dB for large *w*; 6.1 at *w* = 1) was better than either the unfiltered or USM-only filtered image.** This is an extremely significant result. It shows that correctly chosen filtering can improve the performance of a key task (edge detection) before the image is sent to the object recognition/machine vision/AI processing block.



Results with R2A3 USM sharpening. LSF (top), Edge SNRi (bottom)



Results with R2A3 + Gaussian USM sharpening. LSF (top), Edge SNRi (bottom)

The key results (*Edge SNRI* and *SNRi* in dB per pixel²) for a *w*×4*w* object are shown in the Table.

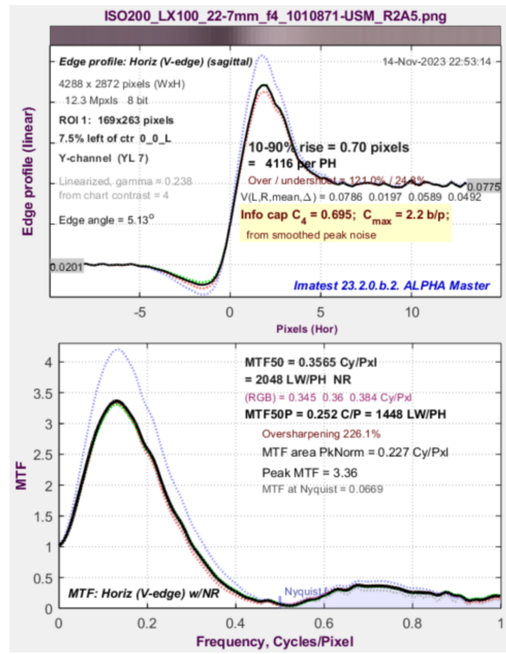
Filter	***MTF50* *C/P***	***Edge SNRi* *w* = 1**	***Edge SNRi* large *w***	***SNRi* dB/pxl² *w* = 1**	***SNRi* dB/pxl² *w* = 5**	***C_{max}* *(NEQ)***	***σ(loc.)* *pixels***
None	0.158	-0.52	5.49	21.7	28.3	2.94	0.13
USM R2A3	0.294	-1.5	4.7	20.7	26.5	2.73	0.219
USM R2A3 + σ = 1 Gaussian LPF	0.243	6.1	9.8	24.7	30.0	3.44	0.105

σ = 1 Gaussian LPF	0.122	1.56	8.5	24.9	33.7	2.72	0.89
USM R2A5 (*extreme oversharpening*)	0.357	-6.8	-1.1	14.7	20.1	2.02	0.26

This important result shows that filtering can improve object detection, indicating that it may be able to improve Object Recognition, Machine Vision, and Artificial Intelligence system performance.

Edge SNRi appears to be slightly more sensitive than *SNRi* dB per pixel^2 (showing greater differences for different filtering). Sharpening + lowpass filtering gives the best result. Results are well-correlated with edge location noise, σ(location).

The excessively oversharpened USM R2A5 image, plotted on the right is illustrated because it's all too common, and we do our best to discourage it: it is a cheap way of improving *MTF50* measurements and image appearance on tiny displays (phones), but it creates "halos" (peaks near edges) that degrade appearance in large displays. The poor *Edge SNRi* and other results are additional reasons to avoid this type of image processing, as have described in [9].



Edge/MTF plot for extremely oversharpened image. MTF50 correlates poorly with performance.

Matched filter

In the above section, we discussed a applying a filter $\mathcal{F}(f)$ to optimize either *SNRi* or *Edge SNRi*.

An optimum filter can be determined if a task (for example, detecting an edge of a certain size) is defined. Such a filter is called a *matched filter*, $\mathcal{F}_{matched}(f)$, which has the same frequency spectrum as the measurement used to derive it. For our edge and object detection tasks,

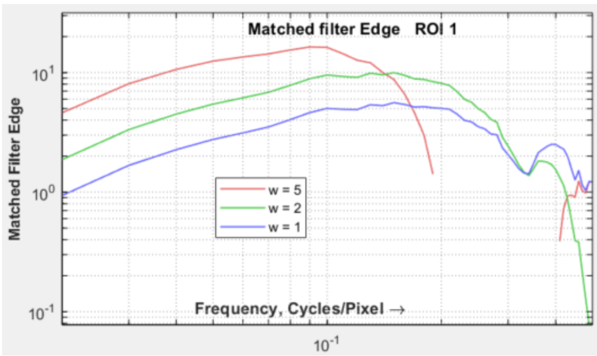
$$\mathcal{F}_{matched}(f) = \frac{|P(f)|\, MTF(f)}{\sqrt{NPS(f)}}$$

$P(f)$ is either $G(f)$ (for *SNRi*) or $H(f)$ (for *Edge SNRi*). Precise matched filters (for the above equation) are difficult to implement exactly, and rarely needed for at least two reasons.

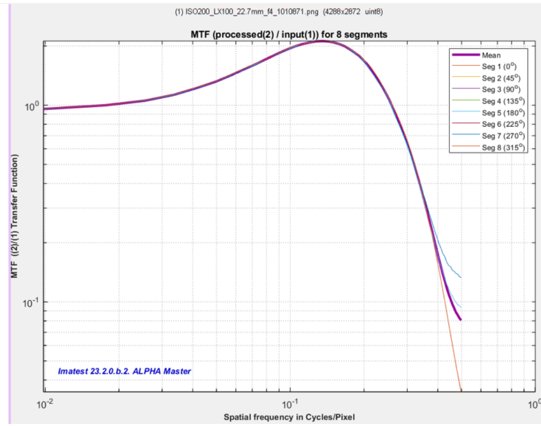
- They don't have to be exact to perform well.
- Real world imaging systems perform a multitude of tasks: detecting objects and edges of varying sizes and colors. Since large objects are usually detected well, it makes sense to design ISP to perform well with small objects (or edges).

In the example below, there is some resemblance between the matched filter for edge detection with *w* = 2 pixels and the sharpening + lowpass filter used above, also shown in Appendix 3. Both the matched filter and the USM+LPF Image Processing filter have response peaks around 0.15 Cycles/Pixel.

Fortunately, filters don't have to be exact to perform well. The matched filter response can be used for guide for designing the filter.



Matched filter for Edge Detection



USM + LPF filter from Image Processing

We have gone far enough down this deep and potentially fruitful rabbit hole. We were fortunate that the filter parameters we found by trial-and-error were reasonably close to the matched filter calculated from the image properties (sharpness and noise) and that they enhanced edge and object detection.

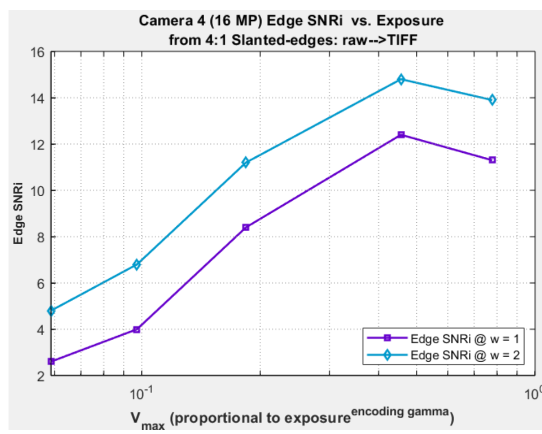
Units and Exposure sensitivity

$NEQ(f)$ is dimensionless because $\mu^2 MTF(f)^2$ has the same units as $NPS(f)$.

$SNRi$ has units of pixels^2 because G has units of pixels^2 = $1/v^2$ (for frequency v), where v has units of cycles/pixel. This is why $SNRi$ increases by about 6 dB (4× energy) for every doubling of feature size w.

$Edge\ SNRi$ and $SNRi$ displayed in dB per pixel^2 are both dimensionless, making them somewhat easier to work with than $SNRi$.

Like C_4, $Edge\ SNRi$ and other noise image metrics vary with exposure. The plot on the right shows $Edge\ SNRi$ vs. exposure for the same camera data used to plot C_4 and C_{max} vs. exposure, above.



Edge SNRi vs. maximum ROI pixel level V_{max} for w = 1 and 2.

Standard exposure — Ultimately, a standard exposure will be needed for comparing cameras (and will need to be in the nascent **ISO 23654** standard). For images encoded with $gamma \approx 0.454 = 1/2.2$ (sRGB, etc.), $V_{max} \approx 0.5$ is appropriate. For linear ($gamma = 1$) images, the equivalent exposure results in $V_{max} = 0.5^{2.2} = 0.22$ (where V_{max} is normalized to a maximum of 1).

Summary

We have developed a powerful toolkit of new measurements — Figures of Merit for imaging systems that combine sharpness and noise — that are especially applicable to Object Recognition, Machine Vision, and Artificial Intelligence systems. The key measurement is ***information capacity***, which can be used to predict camera performance for MV/AI systems. We also have metrics related to specific tasks, most importantly object and edge detection, and are potentially useful for designing ISP filters that optimize OR/MV/AI system performance.

Using ***Edge SNRi***, which is closely related to the more traditional object-based $SNRi$, we have shown an example of image processing (sharpening + lowpass filtering) that improves object detection and is likely to improve MV/AI system performance. This needs to be tested.

> In Appendix 4 we show that **Information capacity C has a monotonic relationship with key metrics for object and edge detection, $SNRi$ and $Edge\ SNRi$, i.e., increasing C_{NEQ} increases $SNRi$ and $Edge\ SNRi$.** This does not hold for standard sharpness metrics based on MTF-only.
>
> This relationship holds because $G_{rect}(f)$ and $H_{impulse}(f)$ (the Fourier transforms of the objects to be detected) are independent of $K(f)$, and hence C_{NEQ}.
>
> ***In other words, object and edge detection performance are functions of information capacity.***

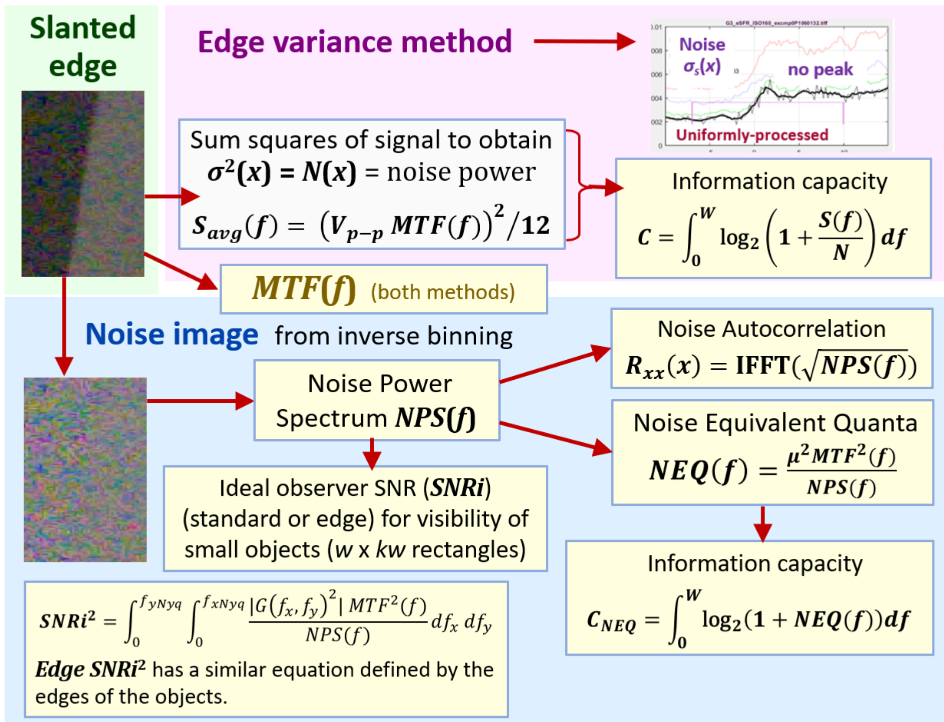
As we become more familiar with information capacity and determine the requirements for effectively performing tasks, we should be able to select cameras with the minimum number of pixels to meet the spec, resulting in faster calculations, lower power consumption, and reduced cost.

The new measurements are extremely easy to obtain from any of ***Imatest's*** slanted-edge analyses. By default, they are included in the Edge/MTF plot and other outputs.

> The key takeaway of this document is that
>
> **Information capacity C is a key measurement for predicting Object Recognition/Machine Vision/Artificial Intelligence system performance.**
>
> **Several additional metrics based on C, most importantly $Edge\ SNRi$ (for edge detection), measure specific task performance and can be used to design filters to optimize OR/MV/AI system performance.**
>
> **Standard MTF measurements are insufficient for this purpose.**

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 total information capacity. For reliable measurements, Siemens stars need to 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 illustrates the two slanted-edge methods, showing the rich interconnections between the new KPIs. For the most part, the ***Imatest*** user does not need to be concerned about details of the two methods.



Summary of calculations for both methods

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

- Better understand the numeric results for $SNRi$ and $Edge\ SNRi$
- Partner with researchers in industry and academia to correlate information capacity C end *Edge SNRi* with performance of Object Recognition, Machine Vision, and Artificial Intelligence systems (accuracy, speed, and power consumption).
- Continue working on a new standard for measuring camera information capacity, assigned **ISO 23654** by the ISO TC42 committee.
- Find better ways of characterizing information capacity in 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

[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

[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] ICRU Report 54, Medical Imaging – The Assessment of Image Quality, 1996, https://www.icru.org/report/medical-imaging-the-assessment-of-image-quality-report-54/

[17] 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.

[18] Chyuan-Tyng Wu et. al. (Intel Corp. Santa Clara), "VISIONISP: Repurposing the Image Signal Processor for Computer Vision Applications," IEEE International Conference on Image Processing (ICIP), 2019, pp. 4624-4628, https://doi.org/10.1109/ICIP.2019.8803607

[19] P. Hansen, et. al., ISP4ML: "Understanding the Role of Image Signal Processing in Efficient Deep Learning Vision Systems," *25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 2438-2445, https://doi.org/10.48550/arXiv.1911.07954

[20] K. Park, M. Chae, and J.H. Cho, "Image Pre-Processing Method of Machine Learning for Edge Detection with Image Signal Processor Enhancement," *Micromachines* **2021**, *12*, 73. https://doi.org/10.3390/

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 classic 1948 and 1949 papers [1],[2] are highly readable.

What is information?

Information is a measure of 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. 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.

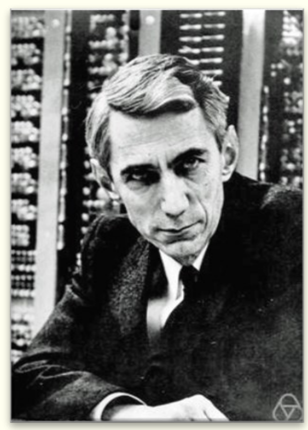
Now, suppose you have a weirdly warped coin that has a probability of 0.99 for a head (1) and 0.01 for a tail (0). Little information is gained from the results of a flip. The equation for the information in a trial with m outcomes, where $p(x_i)$ is the probability of outcome i and $\sum_{i=1}^{m} p(x_i) = 1$, 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 $p(x_1) = 0.95$ and $p(x_2) = 0.05$, $H = 0.286$ bits. If the results of the warped coin toss were transmitted without coding, each symbol would contain 0.286 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 I visited frequently in my 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 without error. 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 0.286.

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. 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, there were no convenient methods of measuring it. (Rodney Shaw's heroic efforts with film in the early 1960s are very impressive [11].)

Appendix 2. Obtaining Results with *Imatest*

Information capacity (C_4 and C_{max}) and related measurements can be calculated from any of ***Imatest's*** ISO 12233-based slanted-edge modules. If you are a beginner with Imatest, we recommend **Using *Imatest* – Getting started**.

Some of the newer methods in this white paper are available (as of November 2023) in the ***Imatest 24.1*** **Pilot Program**. ***Imatest 24.1*** will be released in spring, 2024.

Here are some recommended links for slanted-edge modules (from the documentation page, **www.imatest.com/docs**).

SFR (manual ROIs), SFRplus, eSFR ISO, SFRreg, Checkerboard (auto ROI detection)

Detailed instructions for information capacity and related calculations are on

Image information metrics from Slanted edges: Equations and Algorithms – figures of merit that combine sharpness and noise, conveniently measured from any slanted edge, including NPS, NEQ, and SNRi.

Image information metrics from Slanted edges: Instructions – instructions on the new calculations.

We focus on the settings in the Auto detection modules (settings for SFR, which uses manually selected ROIs, are similar).

The test chart edge contrast should be between 2:1 and 10:1, with 4:1 (the ISO 12233 e-SFR standard [4]) strongly recommended.

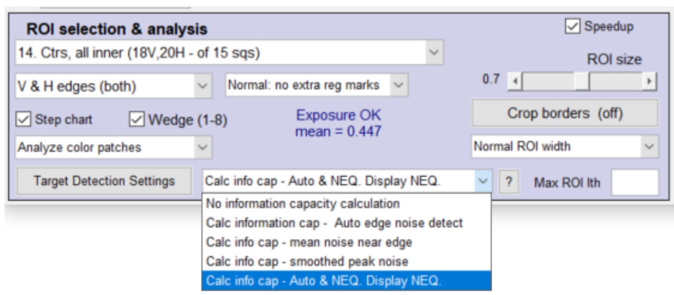
General good technique is recommended for acquiring images:

- Lighting should be uniform and glare-free;

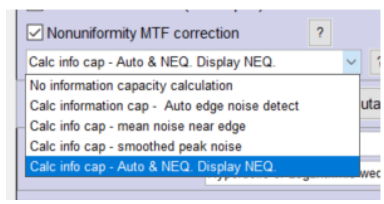
- The image should be well-exposed. Avoid saturation (clipping or operating in response regions with strong nonlinearities— either highlights or shadows). For consistency in comparing cameras, standard exposure is recommended.
- Use sturdy camera support,
- ROIs should be reasonably large: at least 30x60 pixels is recommended. More are better.
- For evaluating cameras for use in OR/MV/AI systems, we recommend minimally or uniformly processed images: avoid bilinear filtering (commonly found in JPEGs from consumer cameras) if possible. This can be done by starting with raw files, then converting them with LibRaw (for commercial files) or Read Raw (for custom binary files). Tone mapping (locally adaptive image processing) should also be avoided.

Setting Channel capacity calculations,

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



—or— in the **More settings** window, which can be opened at any time from interactive (Rescharts) modules.

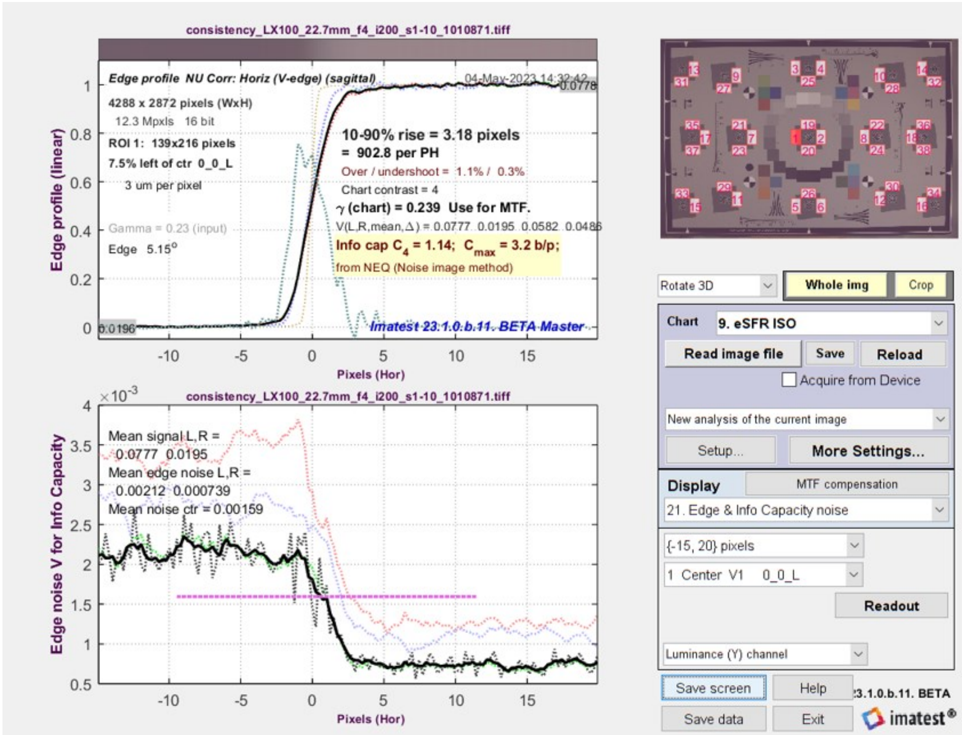


- The first selection turns off all information capacity calculations. This is the default at the time of the 23.1 release. We may change it.
- The remaining selections determine what gets displayed in the Edge and MTF and Edge & Info capacity noise plots.
- The second selection (Auto…) is reasonable when you don't know whether your image is bilateral-filtered.
- The fourth selection (smoothed peak noise) is the current recommendation for edge detection calculations ($Edge\ SNRi$).
- 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 is the best selection for minimally/uniformly processed images.

Information capacity is displayed on the upper (edge) plot of the standard Edge and MTF display. Two selections have been added to the **Display** dropdown menu for displaying Information capacity and related results:

1. Edge & Info Capacity noise

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



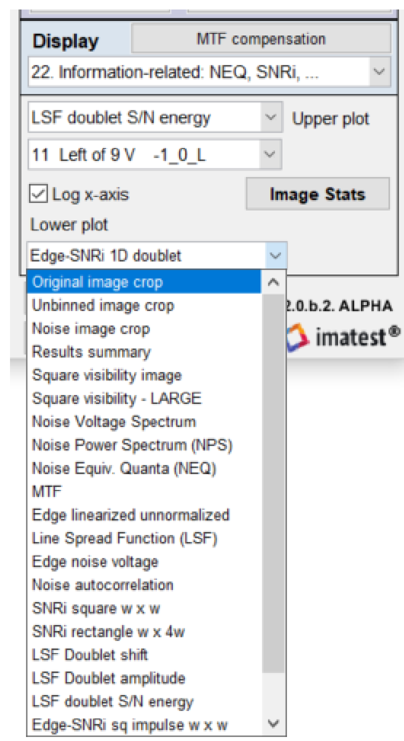
Edge & information capacity noise plot

2. Information-related: NEQ, SNRi, …

A large dropdown menu allows any two of a large number of selections to be displayed: one on top and one on the bottom.

Here are the selections. Click on the links for examples.

Selections in the bottom plot-only		Selections in both plots
Original image crop		Noise Voltage Spectrum
Unbinned image crop		Noise Power Spectrum (NPS)
Noise image crop		Noise Equiv. Quanta (NEQ)
Results summary		MTF
Square visibility image		Edge linearized unnormalized
Square visibility - LARGE		Line Spread Function (LSF)
		Edge noise voltage
		Noise autocorrelation
		SNRi square w x w
		SNRi rectangle w x 4w
		SNRi square per pixel^2
		SNRi rectangle per pixel^2
		LSF Doublet shift
		LSF Doublet amplitude
		LSF Doublet S/N energy
		Edge SNRi sq impulse w x w
		Edge SNRi rect impulse w x 4w
		Edge SNRi 1D doublet



		Object matched filter [**Edge matched filter**]

Appendix 3. Filtering images with the *Imatest* Image Processing module

The ***Imatest*** Image Processing module (instructions on www.imatest.com/ docs/image-processing/) includes typical camera degradation and enhancement functions. We used just two of the functions for the filtering in this document: Gaussian filtering and USM (Unsharp Mask) sharpening. Here is an example of the Image Processing window.

Basic instructions are

1. Press **Read input file (1)** to open the image to be processed.
2. Select the settings. In this case, select **Filter 1** to be Gaussian with sigma = 1 pixel, and select **Sharpen (USM)** with Radius = 2 and Amount = 3.
3. Press **Update calculations** (lower-left).
4. To save the filtered image, press **Save image file (2)**. Make sure the file goes to the location you want, and its name makes sense.
5. To see the effects of the filtering, Check **Side-by-side view** and crop the image. Until this is done, the whole image (input or output) is displayed.

To see the MTF of the filter (below right), uncheck **Side-by-side view** and press **MTF**.

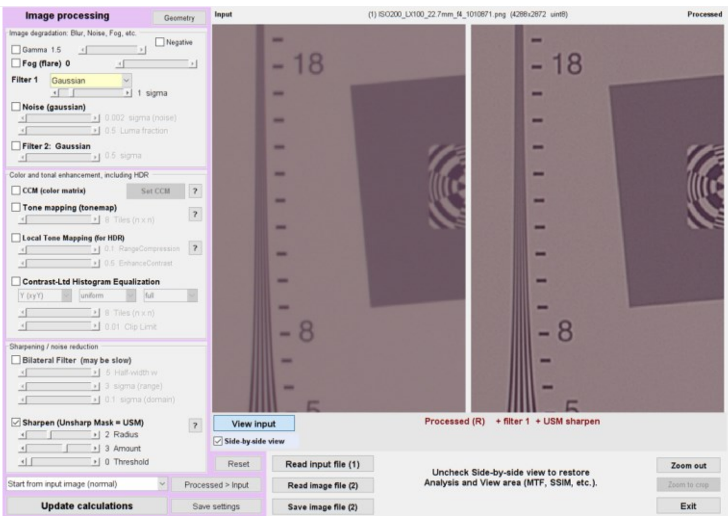
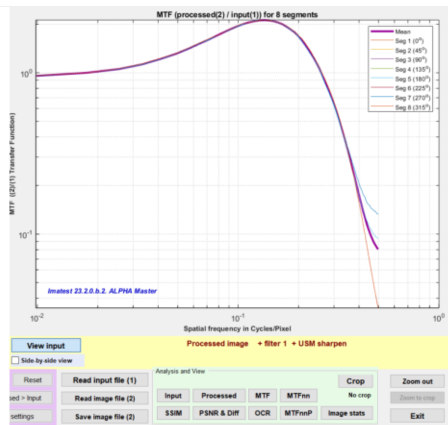


Image Processing module, showing side-by-side view after processing



MTF display for Gaussian Lowpass filtering (σ = 1) and USM sharpening (Radius = 2; Amount = 3) from the Image Processing module

Appendix 4. Correlation between information capacity and object/edge detection metrics

In this section we show how information capacity correlates with the key metrics for object and edge detection, $SNRi$ and $Edge\ SNRi$, which should be predictors of MV/AI system performance.

We start with the integral form of the Shannon-Hartley equation from Wikipedia, derived in Shannon's second paper [2].

> **Frequency-dependent (colored noise) case** [edit]
>
> In the simple version above, the signal and noise are fully uncorrelated, in which case $S+N$ is the total power of the received signal and noise together. A generalization of the above equation for the case where the additive noise is not white (or that the S/N is not constant with frequency over the bandwidth) is obtained by treating the channel as many narrow, independent Gaussian channels in parallel:
>
> $$C = \int_0^B \log_2\left(1+\frac{S(f)}{N(f)}\right) df$$

We define $K(f) = S(f)/N(f)$ as the ***kernel*** of the information capacity equation.

Relating Wikipedia's nomenclature to ours, $N(f) = NPS(f)$ is the Noise Power Spectrum and $S(f) = S_{avg}(f) = \ (\boldsymbol{k\ MTF(f)})^2 = \ \left(\boldsymbol{V_{p-p}\ MTF(f)}\right)^2/\boldsymbol{12}$ is the signal power for calculating C.

To clarify the correlation between the metrics, it is useful to express $SNRi$ and $Edge\ SNRi$, in one dimension, $SNRi^2$ or $Edge\ SNRi^2 = \int \left(\frac{|P_{obj}(f)|^2 V_{P-P}^2\ MTF^2(f)}{NPS(f)}\right) df = \int \left|P_{obj}(f)\right|^2 K(f)\, df$

where $P_{obj}(f) = G_{rect}(f) = kw\ \frac{\sin(\pi w f)}{\pi w f}$ for $SNRi^2$, or

$P_{obj}(f) = H_{impulse}(f) = 2\pi f\, G_{rect}(f) = 2\sin(\pi w f)$ for $Edge\ SNRi^2$.

Grouping the equations for NEQ, C_{NEQ}, $SNRi$, and $Edge\ SNRi$, expressed as functions of $K(f)$, reveals something important.

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

$$C_{NEQ} = \int_0^W \log_2(1 + NEQ_{info}(f))\, df = \int_0^{0.5} \log_2(1 + \mu^2\, K(f))\, df$$

$$SNRi^2 = \int |G_{rect}(f)|^2\, K(f)\, df\ ; \quad Edge\ SNRi^2 = \int \left|H_{impulse}(f)\right|^2 K(f)\, df$$

$NEQ(f)$, C_{NEQ}, and detection metrics $SNRi$ and $Edge\ SNRi$ have a monotonic relationship with each other, based on $K(f)$, i.e., they all increase or decrease with $K(f)$.

Effects of filtering — Because uniform processing — sharpening or lowpass filtering — does not affect the $MTF^2(f)/NPS(f)$ ratio or $K(f)$, it does not affect $NEQ(f)$ or C_{NEQ}, as expected from the data processing inequality. It does, however, affect $SNRi^2$ and $Edge\ SNRi^2$, which have an additional $\left|P_{obj}(f)\right|^2$ term inside the integral, and can be improved with appropriate filtering.

Appendix 5. Binning noise

This "**green for geeks**" box can be skipped by most readers.

November 10, 2023. We are working on an improved binning calculation that we hope will improve the distinction between uniformly sharpened and bilateral-filtered images.

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.

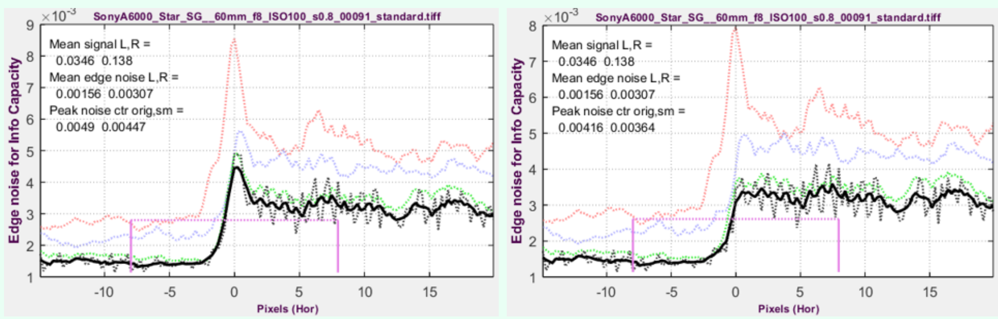
Assume that n identical signals $\mu_s(x)$ are binned over an interval $\{-\Delta/2, \Delta/2\}$, where $\Delta = 1$ in the 4× oversampled output of the binning algorithm (noting that Δ = (original pixel spacing)/4). If there were no binning noise, we would expect the binning noise power $\sigma_{Bnoise}{}^2$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

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

for Line Spread Function LSF. Noting that δ 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

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

Although this equation involves some approximations, we have had good success calculating the corrected noise, $\sigma_s^2(\text{corrected}) = \sigma_s^2 - \sigma^2_{Bnoise}$. 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.

For now, the Slanted edge calculation setting, on the lower-left of the **More settings** window, must be set to ***Imatest* 22.1 (recommended)**.