Image information metrics from slanted edges

a toolkit of metrics to aid object recognition, machine vision, and artificial intelligence systems

Norman L. Koren, Imatest LLC, Boulder, Colorado

Revised October 29, 2026

This document is an updated and revised version of the paper of the same name, presented at Electronic Imaging 2004. It is intended to be the primary technical introduction to *Imatest's* image information metrics — a new set of measurements for quantifying Machine Vision/Artificial Intelligence (MV/AI) system performance. It contains all the important equations and algorithms.

The original paper is available on library.imaging.org/ei/articles/36/9/IQSP-256. Although this document contains numerous changes and additions, the key calculations are unchanged. Several Appendices have been added for material that isn't appropriate for the main body of text because it's too technical (the equations for binning noise) or experimental (noise autocorrelation as an indicator of sensor *SFR*).

Table of Contents

Abstract 3

Introduction 3

- Related documents 4
- Information capacity 4
- Test charts 4
- The slanted-edge measurement 5

The Edge Variance method for spatially dependent noise power, $N(x)$ 6

- Signal power, S 6
- Noise power, N, and image processing 7
- Bandwidth, W 10
- Combining S, N, and W to obtain information capacity, C 10
- Edge Variance results: Edge, *SFR*, C 10

Measurement technique 10
Additional assumptions 11
Sensitivity to exposure 11
Maximum information capacity C_{max} — a more stable metric than C_4 12
Information capacity results 13
Effects of sharpening 15
Signal averaging 16
Total information capacity 16
Summary of the Edge variance method 17
The Noise Image Method 17
The noise image 17
Noise Power Spectrum, *NPS* 18
Noise Equivalent Quanta, *NEQ* 20
Information capacity from *NEQ* 20
Ideal Observer SNR (*SNRi*) 21
Object Detection Error Probability, *ODEP* 23
Task: Detection, Classification, Identification (DRI) 23
Object visibility and *SNRi* 24
Edge SNRi and Object Location Standard Deviation (*Object Location* σ) 26
Line Spread Function (LSF) doublet results 27
Effects of Image Signal Processing (ISP) 28
Matched filters 30
InfoDR chart, exposure, and pixel level 32
Heuristic and experimental metrics 32
Summary 33
Future work 34
References 35
Author Biography 36
Appendix I. Information theory background 36
What is information? 36
Channel capacity 37
Appendix 2. Binning noise 38
Appendix 3. Edge location variance (or standard deviation) 39

Appendix 4. Noise autocorrelation 40

Abstract

We present a comprehensive framework for conveniently measuring camera information capacity and related performance metrics from the widely used slanted-edge (e-SFR) test patterns.

The goal is to develop a set of image quality metrics that can predict the performance of Machine Vision (MV) and Artificial Intelligence (AI) systems, assist with camera selection, and be used for designing electronic filters to optimize system performance. The new methods go far beyond the standard approach of estimating system performance based on sharpness and noise (or Signal-to-Noise Ratio) — which often involves more art than science.

Metrics include Noise Power Spectrum (NPS), Noise Equivalent Quanta (NEQ), and for quantifying the detectability of objects and edges, Ideal Observer Signal-to-Noise Ratio (SNRi), Object Location standard deviation (σ_{objLoc}), and Object Detection Error Probability (ODEP). We show how to use these metrics to design electronic filters that optimize object and edge detection performance.

The new measurements can be used to set camera specifications and finding a camera that meets performance requirements with a minimum number of pixels— important because fewer pixels mean faster processing and lower energy consumption and cost.

Introduction

We introduce the concept of information capacity, which is calculated from signal power, noise power, bandwidth, and image contrast. We show how noise power (and hence information capacity) is calculated *from the same location* using the familiar slanted-edge (e-SFR) test pattern, specified by the ISO 12233 standard [1]. This algorithm (with some refinements) is in process of being incorporated into ISO 23654, with release expected in 2028.

Noise is calculated in two steps.

1. For any image, regardless of how it was processed, the **spatially dependent noise power, $N(x) = \sigma_s^2(x)$**, is calculated by the **Edge variance method**, which allows uniform image processing to be distinguished from nonuniform processing. It can be used to obtain a reasonable approximation to the information capacity of images with bilinear (nonuniformly) filtered images, found in JPEG image files from most consumer cameras.

2. For uniformly processed images, the **noise power spectrum, $NPS(f)$** is calculated from a **noise image**, obtained by inverse binning. This method fully characterizes the noise, enabling a more accurate calculation of information capacity as well as several additional performance metrics, including Noise Equivalent Quanta, *NEQ*, Ideal observer Signal-to-Noise Ratio, *SNRi*, *Object location σ*, and Object Detection Error Probability (*ODEP*).

We describe how to use these metrics to evaluate camera performance and design electronic filters, called *matched filters*, for optimizing object detection and edge location performance.

Related documents

Solutions - Image Information Metrics contains a concise summary of information capacity, along with news and links to pages, white papers, and videos. It's the go-to page for finding documentation on information metrics.

Image quality testing based on information metrics is a tutorial/mini-textbook/crash course on image quality testing, emphasizing the new information metrics. It is the most important companion document, containing broad information on image quality testing, but fewer equations that this document, though there is some overlap.

Information capacity

Camera information capacity, based on Claude Shannon's ground-breaking work on information theory [2-3] has long held promise as a figure of merit for a variety of imaging systems. It has been discussed in several technical papers [4-7] and in two textbooks [8-9], but it has failed to gain widespread adoption in the imaging industry, primarily because it was difficult to measure.

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

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

It is quite logical to apply this definition to imaging systems, where C has units of bits/pixel. Signal power, $S(f)$, and noise power, $N(f)$, must be measured with care, which was traditionally difficult and error-prone because $S(f)$ and $N(f)$ had to be measured at separate locations.

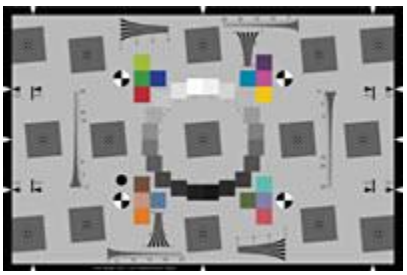
To make matters worse, most JPEG images from consumer cameras usually have nonuniform image processing (bilateral filtering) [10] that sharpens images near contrasty features such as edges (boosting high frequencies) but reduces noise elsewhere (lowpass filtering). This can increase the measured information capacity while removing information.

We resolved this problem by measuring noise from the same image used to calculate MTF (sharpness) and calculating noise power N at the edge, where MTF is measured. This approach is highly convenient and reduces chances of errors in scaling.

Test charts

Several types of test chart can be used for measuring information metrics with slanted edges. We recommend the following charts.

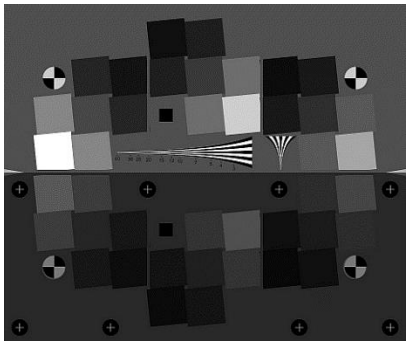
1. Auto-detection slanted-edge charts, which include SFRplus, eSFR ISO, Checkerboard, and SFRreg. These chart types are described and compared on the Documentation page (Table of contents). They operate in interactive and fixed (batch-capable) modes. The eSFR ISO chart is shown on the right. These charts are excellent for mapping information (as well as standard) metrics over the image. Figure 14 (below) is an



eSFR ISO chart

example of the results.

2. The InfoDR (Information Dynamic Range) chart, released in early 2026, consisting of patches with a wide range of tones bounded by 4:1 contrast (0.6 Michelson contrast) slanted edges. The most useful version, consisting of two layers of photographic film, has 42 grayscale patches and 48 slanted edges, with a patch density range of 5.20 (104 dB) and a mean edge density range of 4.90 (92 dB). Its use for measuring performance as a function of illumination is described in Image quality testing based on information metrics. More detailed documentation can be found in Part 1: InfoDR Introduction — Part 2: Using InfoDR — Part 3: InfoDR Results.



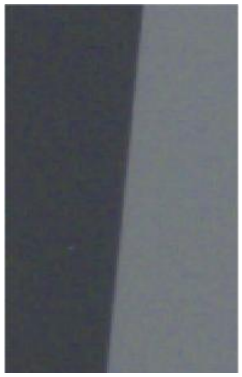
InfoDR chart

The companion document, Image quality testing based on information metrics, discusses how the two chart types complement each other, with

- spatial performance over the image measured by the auto-detect slanted edge charts, and
- tonal performance over a wide range of exposure measured by the InfoDR chart,

to fully characterize camera information metric performance.

The slanted-edge measurement



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

1. **The image should be well-exposed**, avoiding the dark "toe" and light "shoulder" response regions.
2. **Linearize the image** by applying the inverse of the encoding gamma curve. Linearization does not need to be precise (see below).
3. **Find the center of the transition** between the light and dark regions for each horizontal scan line, $y_l(x)$.
4. **Fit a polynomial curve** to the center locations.
5. **Add each appropriately shifted scan line to one of four bins**, depending on the location of the curve relative to the scan line.
6. **Combine the mean signal in each bin** to obtain the 4× oversampled averaged edge for L scan lines, $\mu_s(x)$, illustrated in the upper plot of Figure 8.

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

7. **Calculate Spatial Frequency Response, *SFR*(*f*) (synonymous with *MTF*(*f*)),** by differentiating the averaged edge, windowing it, then taking the magnitude of the Fourier transform, normalized to 1 (100%) at zero frequency, illustrated in Figure 1.

Nomenclature — Spatial Frequency Response (*SFR*) and Modulation Transfer Function (*MTF*) are often used synonymously, but *SFR* is generally preferred in recent literature [1]. We use *SFR* (mostly) here, although some plots are labeled *MTF* (which is more familiar) and we keep summary metrics such as *MTF50*, the spatial frequency where *SFR* drops to 50% of its zero-frequency value.

Linearization — A relatively simple method for linearizing the image based on the mean pixel levels, Digital Numbers DN_L and DN_D on the light and dark sides of a slanted edge, provides excellent accuracy for calculating *SFR* (*MTF*) and information metrics. For an edge with an n:1 contrast ratio in a camera whose response follows a standard gamma (γ) curve, $DN = E^{\gamma}$, for illumination E, $DN_L = E_L^{\gamma}$, $DN_D = E_D^{\gamma}$, where $E_L/E_D = n$. Noting that $DN_L/DN_D = E_L^{\gamma}/E_D^{\gamma} = (E_L/E_D)^{\gamma} = n^{\gamma}$, and solving for γ by taking the logarithm,

$$\gamma = \frac{\log\,(DN_L/DN_D)}{\log\,(n)} \qquad (2a)$$

The linearized Digital Number for calculating $SFR(f)$ is $DN_{linearized} = DN^{(1/\gamma)}$.

The Edge Variance method for spatially dependent noise power, *N*(*x*)

For the Edge Variance method, which was introduced in an earlier paper [13], a simple addition to the ISO 12233 binning algorithm described above allows the ***variance*** of the signal, $\sigma_s^2(x) = N(x)$ (the noise power), to be calculated in addition to the mean, $\mu_s(x)$.

In addition to $\mu_s(x) = \sum y_l(x)$, calculate the sum of the squares of each scan line, $\sum y_l^2(x)$. Then, the noise power is

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

$\sigma_s(x) = \sqrt{N(x)}$ is the noise amplitude. This equation holds for the entire oversampled array—including at the edge transition, where noise was traditionally difficult to measure.

[13] and Appendix 2 describes a form of quantization noise called *binning noise* that is largest near the image transition— where the Line Spread Function, $LSF(x) = d\mu_s(x)/dx$ is maximum. Subtracting it from $\sigma_s^2(x)$ modestly improves calculation accuracy.

Noise amplitude results ($\sigma_s(x)$) are shown in the section on Noise Power, below.

Signal power, *S*

The peak-to-peak signal amplitude, $V_{P\text{-}P}$, (Figure 4) at low spatial frequencies is the measured difference between the means of the light and dark regions of the linearized slanted edge, $\mu_s(x)$.

$$V_{p-p} = \Delta\mu_s = \mu_{sLight} - \mu_{sDark} = V_{max} - V_{min} \qquad (4)$$

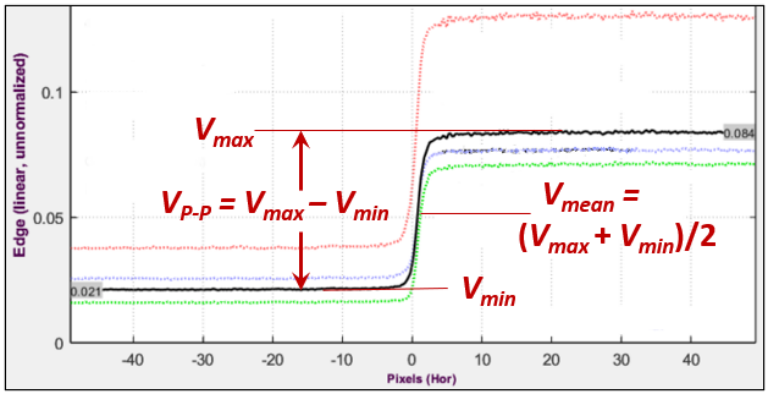


Figure 4. Slanted-edge amplitudes (voltages, V)

Since our intent is to calculate the information (or channel) *capacity*, which is the *maximum* information for the $V_{P\text{-}P}$ signal, we assume a uniform signal distribution that maximizes the information capacity. Its variance, which is the average signal power at low spatial frequencies, is

$$\sigma_V^2 = S_{avg}(0) = (\mu_{sLight} - \mu_{sDark})^2/12 = V_{p-p}^2/12 \qquad (5)$$

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

$$S_{avg}(f) = \left(V_{p-p}\, SFR(f)\right)^2/12 \qquad (6)$$

Signal power, S, is proportional to the square of the chart's Michelson contrast, $C_{Mich} = (\mu_{sLight} - \mu_{sDark})/(\mu_{sLight} + \mu_{sDark})$, for a properly linearized image, which is easy to obtain if the image does not approach saturation at low or high pixel levels. Note that $S_{max} \leq 1$ for linearized images normalized to 1. (Some cameras restrict S_{max} to values < 1)

Noise power, *N*, and image processing

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

Depending on how the image has been processed, system performance may be dominated by the mean of the noise power, mean($N(x)$) or by the noise power in the vicinity of the edge transition region, which is defined by the Line Spread Function, $LSF(x) = d\mu_s(x)/dx$, shown in Figure 5.

For an *LSF* peak located at X_{LSF}, the edge transition region is defined by $X_{LSF} - PW20 < x < X_{LSF} + PW20$, where *PW20* is the length of the region where $LSF(x) = d\mu_s(x)/dx \geq 0.20\ LSF_{max}$.

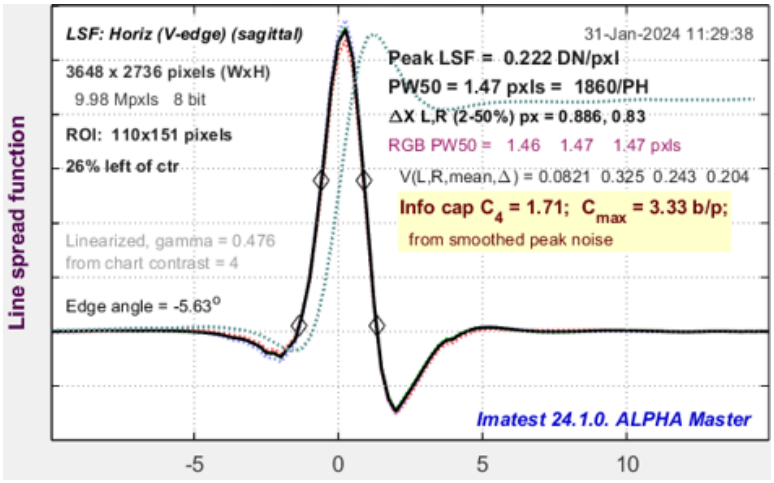


Figure 5. Line Spread Function, LSF = $d\mu_s(x)/dx$ for the strongly sharpened bilateral-filtered image in the example below.

Two image processing types cover most cases of interest. A third type is extremely rare.

1. Minimally and uniformly -processed (but not strongly sharpened) images, typically TIFFs converted from raw files (raw→TIFF). Most cameras to be evaluated for Machine Vision/Artificial Intelligence are in this category. Minimally processed images can be identified by the lack of a strong noise peak near the transition.

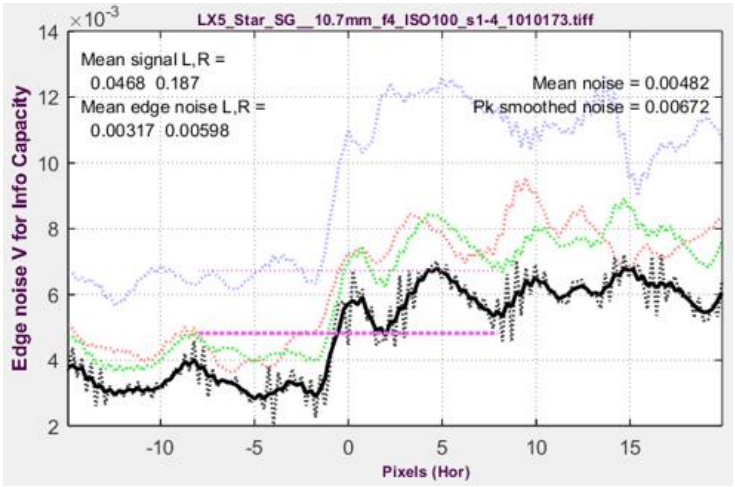


Figure 6. Noise amplitude $\sigma_s(x)$ for uniformly processed image (TIFF from raw; no sharpening or noise reduction). ISO 100. The bold black curve is the Y-channel, smoothed with a 1.25-pixel kernel before 4× oversampling.

Since noise amplitude, $\sigma_s(x) = \sqrt{N(x)}$, can be quite rough (Figure 6), the mean of the entire acquired edge is recommended for calculating N for the Shannon-Hartley equation.

$$N_{uniform} = \text{mean}\left(\sigma_s^2(x)\right) \qquad (7)$$

2. Bilateral-filtered images [10] include most JPEG images from consumer cameras. Bilateral filters sharpen images near contrasty features such as edges, but blur (lowpass-filter) them to reduce noise elsewhere. This causes a distinct noise peak, shown in Figure 7, close to the edge transition, which can dominate camera performance because SFR is also measured at the transition. The noise-reduced (lowpass-filtered) region should not be used for calculating *NPS*(*f*) or information metrics. We have long known about the noise peak, but there was previously no easy way to observe it.

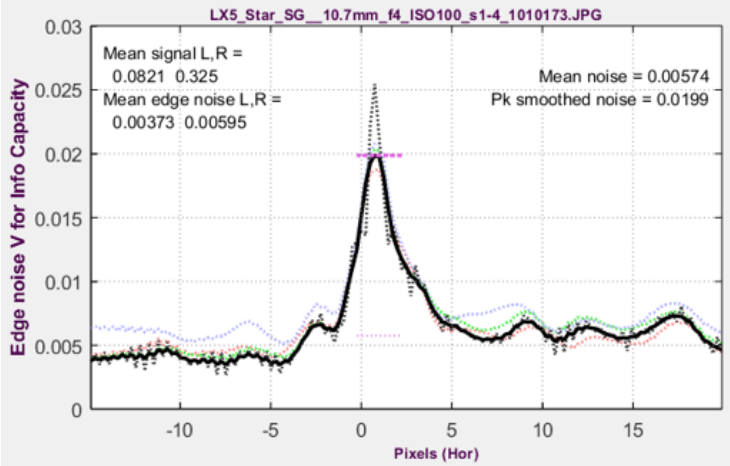


Figure 7. Noise amplitude $\sigma_s(x)$ for bilateral-filtered image (sharpened near edges; noise-reduced elsewhere) from a camera JPEG. ISO 100..

For bilateral-filtered images, use Noise power, N_{peak}, the square of the peak noise amplitude, smoothed with a rectangular kernel of length $PW20/2$. This is an arbitrary choice, but it removes most of the jaggedness and produces reasonably consistent results.

3. A very few **strongly but uniformly-sharpened images** may have strong noise peaks caused by defects such as dust specks. These defects have little effect on standard SFR measurements. In the rare event that this occurs, noise $N_{uniform\text{-}sh}$ can be calculated from $\text{mean}\left(\sigma_s^2(x)\right)$ *excluding* the edge transition region, X_{LSF} – PW20 < x < X_{LSF} + PW20.

The selection of noise ($N_{uniform}$, N_{peak}, or $N_{uniform\text{-}sh}$) to enter into the Shannon-Hartley equation is based the presence of a detected peak near the transition. Some additional considerations:

- Noise is close enough to white to yield reasonably good results. This assumption is supported by experimental results in [22]. For minimally or uniformly-processed images, the Noise Image calculation, below, which uses the noise power spectrum, *NPS*(*f*), is more accurate.
- Noise power is larger on the lighter side of the edge due to photon shot noise, which increases with the number of photons reaching the sensor pixels (for linear— but not HDR— sensors). The mean, $N_{uniform}$, includes both sides.

- For linear sensors, noise power increases with exposure, following the function $N(V) = k_0 + k_1 V$, where k_1 is the coefficient for photon shot noise, derived in [13]. But Signal-to-Noise Ratio (SNR) improves.
- Because N_{peak} is measured at the same location as the signal used to measure *SFR*, it provides a reasonable approximation to the information capacity for bilateral-filtered images, which includes almost all JPEGs from cameras. It can be useful for estimating the performance of "black box" cameras, which have unknown image processing.

Bandwidth, W

Bandwidth, W, is always 0.5 cycles/pixel (the Nyquist frequency, f_{Nyq}). 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 actual image. Frequency dependence comes from $SFR(f)$, which is a component of $S_{avg}(f)$.

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

Once signal power, S, and selected noise power, N ($N_{uniform}$, $N_{uniform\text{-}sh}$, or N_{peak}, as appropriate), have been obtained, information capacity for the Edge Variance method, C_{EV}, can be calculated.

$$C_{EV} = \int_0^{0.5} \log_2\left(1 + \frac{S_{avg}(f)}{N}\right) df \quad \cong \quad \sum_{i=0}^{0.5/\Delta f} \log_2\left(1 + \frac{S_{avg}(i\Delta f)}{N}\right) \Delta f \qquad (8)$$

Edge Variance results: Edge, *SFR*, *C*

Figure 8 shows the Edge and SFR response as well as calculated information capacity values (C_4 and C_{max}, to be introduced below). C_4 is the information capacity for a 4:1 contrast ratio edge.

$SFR(f)$, which is sometimes confused with bandwidth, can take a large bite out of C, especially since it is squared in the equation for signal power, $S_{avg}(f)$. [13] contains an explanation of how increasing SFR can lead to significant aliasing-related artifacts, such as Moiré, that degrade performance.

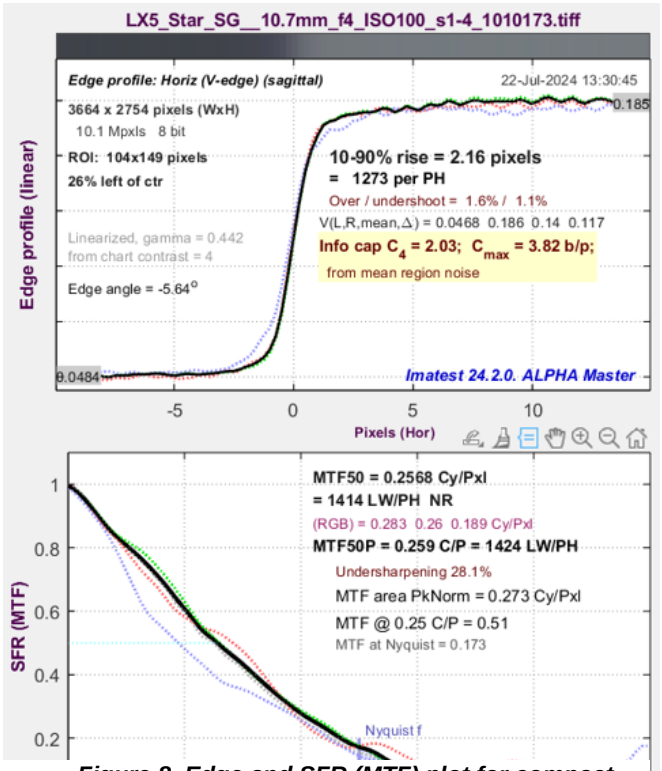


Figure 8. Edge and SFR (MTF) plot for compact digital camera for an unsharpened TIFF from raw. Upper: Mean edge $\mu_s(x)$; Lower: SFR(f).

Measurement technique

- Test chart edge contrast should be between 2:1 and 10:1, with ≤ 4:1 (specified in the ISO 12233 e-SFR standard) strongly recommended. Edge contrast greater than 10:1 increases the

likelihood of nonlinearities (saturation or clipping) that will compromise the results.

- Images should be well-exposed because saturation or clipping can cause misleading results. Exposure should be consistent if different cameras are being compared. It may be of interest, however, to measure *C* as a function of exposure, especially for dim light. The InfoDR chart is particularly convenient for this.
- The camera should be well-focused. Sturdy camera support should be employed.
- Although results are relatively insensitive to ROI selection, some care must be taken to obtain good consistency. ROIs should be reasonably large; at least 30x60 pixels where possible. If possible, the edge should be centered in the selected region, and there should a reasonable amount of "breathing room" on the sides.

Additional assumptions

A key assumption is that the camera's dynamic range (the range of tones that can be reproduced with good contrast and Signal-to-Noise Ratio (SNR)) is sufficient for the intended task. Most modern image sensors have dynamic ranges greater than 60dB (1000:1); high dynamic range (HDR) sensors have 120 dB or more. The majority of scenes in pictorial, medical, or robotic (but *not* automotive) imaging have tonal ranges under 60 dB. Lens flare (stray light) typically limits practical camera dynamic range to under 100 dB, which can impact automotive night driving by fogging important dark to middle tones. If there are concerns about dynamic range, we strongly recommend measuring it with a transmissive grayscale chart, preferably InfoDR .

Other assumptions: sensor nonuniformities (fixed-pattern noise, also called PRNU (Photo Response Nonuniformity) are included in noise measurements. Tonal response is well-behaved (typically following a gamma curve, except for the extreme highlights and shadows). Stray (flare) light is not severe in regions where *SFR* is measured.

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

Sensitivity to exposure

Because both noise power, N, and amplitude range, ΔV, increase with exposure, **C_4 is a strong function of exposure,** as illustrated in Figure 9**.** This makes it useful for measuring low light performance, especially with the **InfoDR** chart, as illustrated in **Figure 31**.

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

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

Maximum information capacity C_{max} — a more stable metric than C_4

In addition to using C_4 to measure information capacity as a function of exposure (Figure 9), we are also interested in measuring a camera's *maximum* information capacity, C_{max}, which we expect to be independent of exposure. It is obtained in two steps.

Step1: In the equation for C, replace the measured peak-to-peak amplitude range, $V_{P\text{-}P}$, with the maximum allowable value, $V_{P-P_max} = 1$ (for systems normalized to a maximum amplitude of 1). This may seem like a simplification, but it works well for most cameras. Referring to the section on Noise Power, N,

$$S_{avg}(f) = \left(V_{p-p_max}\, SFR(f)\right)^2/12 = \frac{SFR(f)^2}{12} \qquad (9)$$

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

$$N(V) = k_0 + k_1 V \qquad (10)$$

Equations for k_0 and k_1 and an adjustment to C_{max} for bilateral-filtered images (which are less accurate than for minimally processed images) are derived in [13].

C_{max} (Figure 9) is nearly independent of exposure for minimally or uniformly-processed images with linear sensors, where noise power, N, is a known function of signal amplitude, V, but is only approximately independent for imaging systems with bilateral filtering or HDR (nonlinear) sensors, where noise power N is not a simple function of V.

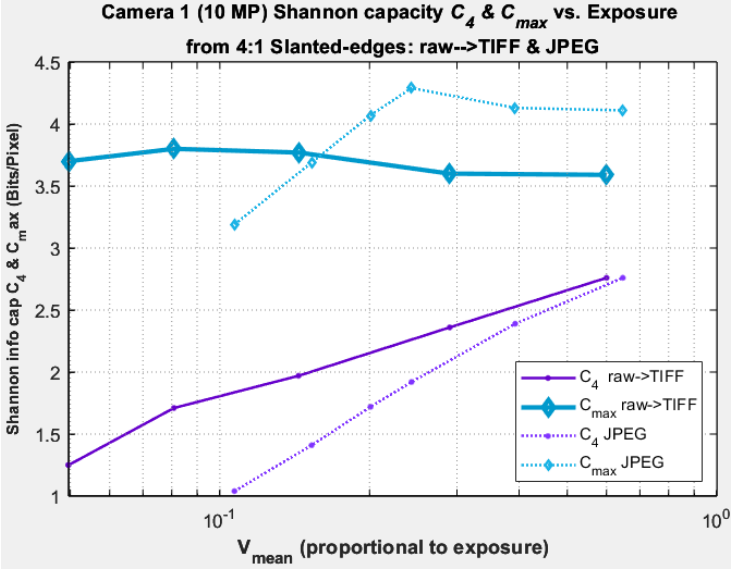


Figure 9. C_4 and C_{max} for minimally processed raw→TIFF and JPEG images for a 10 MP compact camera. C_{max} is consistent for the raw→TIFF image. C_4 is useful for quantifying low light performance.

Limitations of C_{max}

– Results are not valid when the maximum pixel level is below the maximum allowable level for the bit depth (2N-1 for bit depth N, e.g., 12-bit data in a 16-bit container). This is not uncommon.

– Results are not reliable for High Dynamic Range (HDR) sensors, as described below.

High Dynamic Range (HDR) images

C_{max} is challenging to calculate for HDR sensors, which have several cycles of SNR and noise as exposure increases [14].

Noise $N(V)$ increases monotonically in dark regions, but jumps at the discontinuities of the SNR plot (Figure 10), around $\text{Log}_{10}(DN) = 3.6$ and 4.8. The noise measurement depends on the location on the sawtooth curve where the measurement is made. Because each HDR sensor is different, there is no simple equation, comparable to (10) for calculating C_{max}. A rough approximation requires an assumption about the maximum SNR, perhaps limiting it to the mean value in the sawtooth region (above $\text{Log}_{10}(DN) = 3$ in Figure 10).

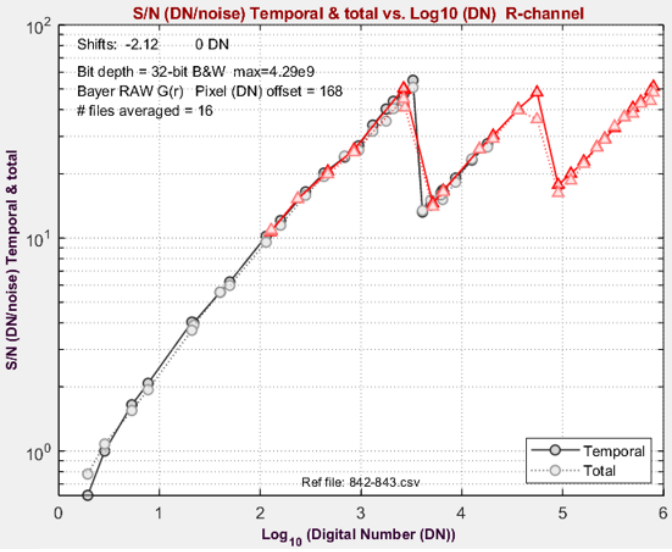


Figure 10. Cyclic response of Signal-to-Noise Ratio for HDR sensor

For now, C_{max} is not recommended for characterizing the performance of HDR sensors. But measurements of C_4 as a function of exposure, which can be performed conveniently with the InfoDR chart, are of definite value. The **heuristic $C_4\Delta$ metric**, may prove to be a useful substitute for C_{max}.

Information capacity results

Table 1 shows three cameras with both raw and JPEG output that we tested for information capacity as a function of Exposure Index (ISO speed setting).

Table 1. Cameras used in the tests

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

The image in Figure 11, which was analyzed in [15], contains a 50:1 contrast Siemens star and four 4:1 contrast slanted edges. We used the upper-left slanted edge for most tests. The average background of the chart is close to neutral gray (18% reflectance) to ensure a good exposure.

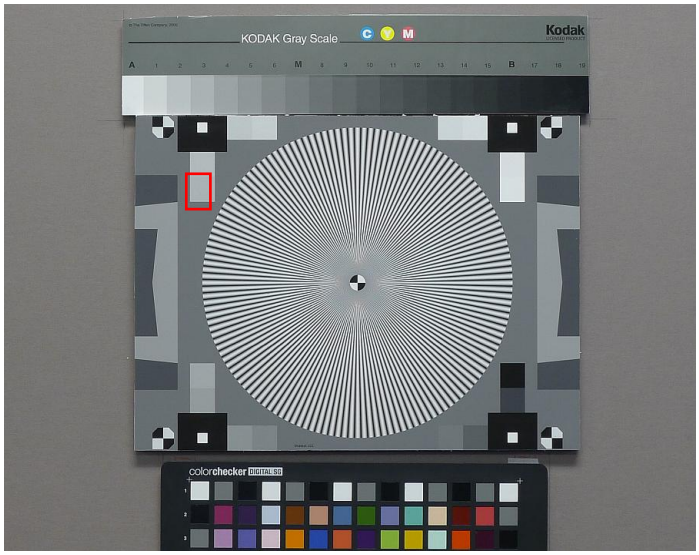


Figure 11. Typical image (cropped) including Siemens star and slanted-edges to the left and right of the star.

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

Figure 12 shows C_4 as a function of ISO speed (Exposure Index, which is proportional to analog gain) for 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.

Note that increasing the Exposure Index *decreases* the exposure (the total light reaching the image sensor).

$N_{uniform}$ was used for the raw→TIFF images; N_{peak} was used for the bilateral-filtered JPEGs.

C_{max} has a similar trend to C_4, but is higher about 1.6. [13] contains more detail on its behavior.

Color channels

The separate R, G, and B channels tend to have slightly lower C_4 than the Y-channel because the noise in the separate channels is uncorrelated. Color is discussed in more detail in [13].

Although this document has focused on demosaiced images,

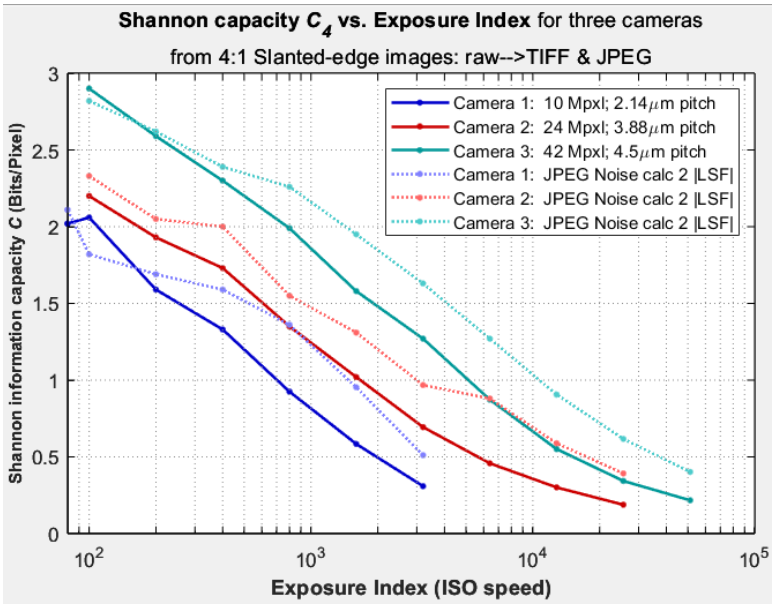


Figure 12. Information capacity, C_4, from 4:1 slanted-edge images. Solid lines for raw→TIFF images; Dotted lines for JPEGs. Note that increasing the Exposure Index decreases the exposure (total light reaching the image sensor).

the slanted-edge method can also be applied to raw (undemosaiced) images.

Effects of sharpening

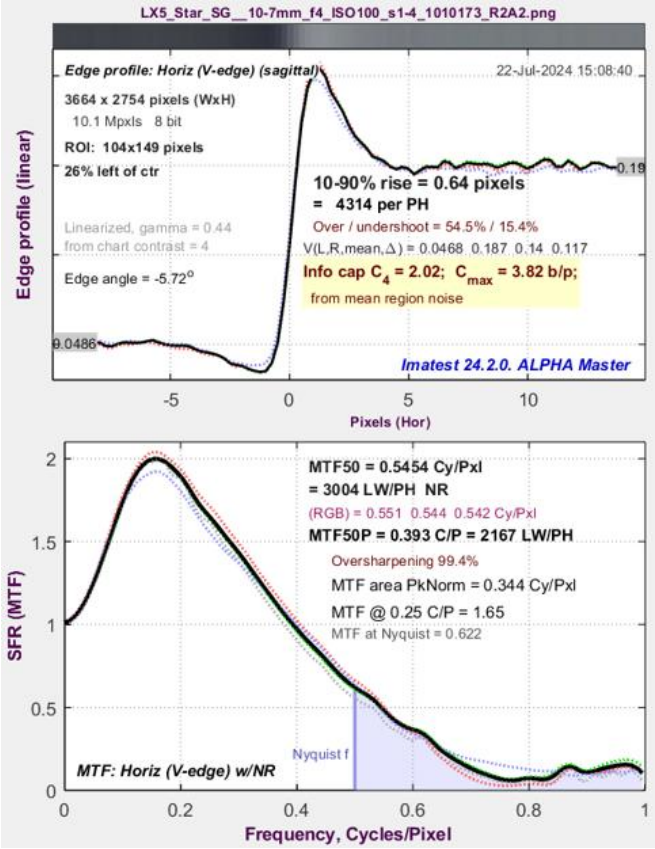


Figure 13. Edge/SFR (MTF) plots derived from the same image as Figure 8, where C_4 = 2.03 b/p and C_{max} = 3.82 b/p, raw→TIFF, ISO 100 Sharpening Radius = 2; Amount = 2. C_4 = 2.02 b/p; C_{max} = 3.82 b/p. (The exact agreement is a coincidence: they're typically very close.)

As expected, the examples in Figures 8 and 13 (and many others we ran) show that sharpening (and more generally, reversable image processing) has little effect on slanted-edge information capacity. The image used for Figure 13, which has been strongly Unsharp Mask (USM) sharpened with Radius = 2 and Amount = 2 (R2A3), was initially the minimally-processed TIFF used for Figure 8, where $C_4 = 2.03$ and $C_{max} = 3.82$ b/p. We observed a similar insensitivity of C to sharpening with Siemens stars [15].

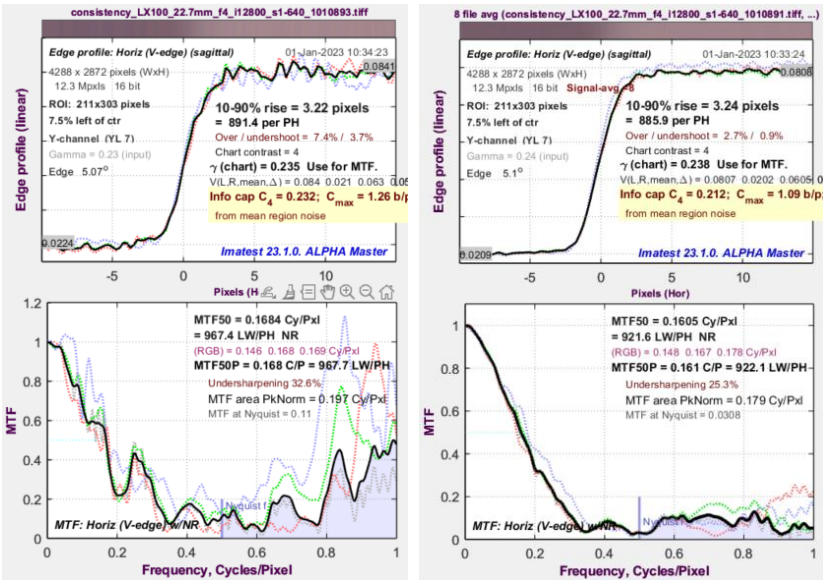
One of our original goals in developing the information capacity measurements was to be able to measure camera quality independently of sharpening and image processing, which strongly varies from camera to camera. The information capacities for the in-camera JPEG version of the image in Figures 8 and 13, which is strongly sharpened (with peak $MTF \cong 1.5$) is $C_4 = 1.91$ and $C_{max} = 3.61$ b/p, which is reasonably close.

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***

Total information capacity

The total information capacity, C_{total}, for the entire image is calculated from

$$C_{total} = \text{mean}(C) \times \text{megapixels} \qquad (11)$$

From Figure 14, the mean value of C_{max} is 2.847 bits/pixel. Since this camera has 16 Megapixels, the total capacity, $C_{maxTotal}$, for the Luminance (Y) channel = 45.55 MB.

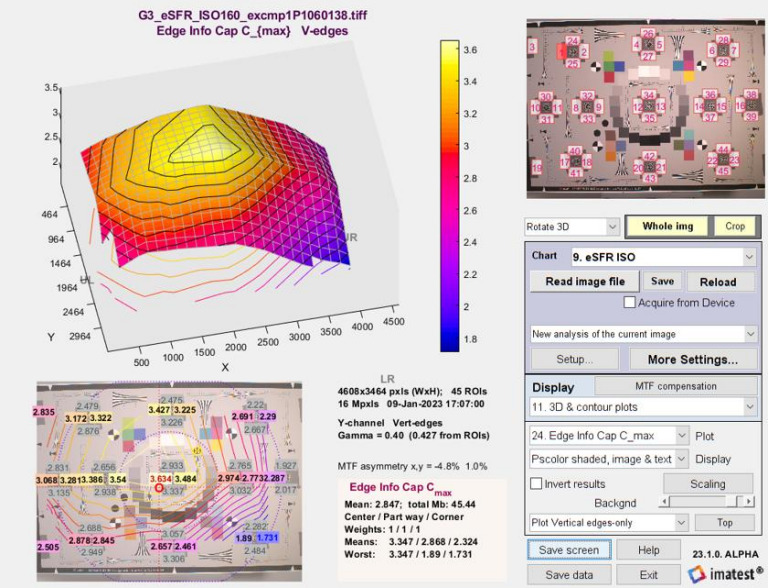


Figure 14. 3D contour eSFR ISO plot of Cmax for the Luminance (Y) channel, ISO 100

Summary of the Edge variance method

The edge variance calculation of the spatially dependent noise power, $N(x)$, does not fully characterize the noise. It is, however, good enough for a reasonable approximation to information capacity C that works with nonuniform as well as uniform image processing. The mean (or peak value) of the noise is substituted into the Shannon-Hartley equation to calculate information capacity, C_n (measured directly from an n:1 contrast chart) or C_{max} (the maximum information capacity, which is independent of exposure and chart contrast).

The Noise Image method for calculating the Noise Power Spectrum, $NPS(f)$, described below, is required to derive key metrics for edge and object detection. But it not reliable for most nonuniformly processed (especially bilateral filtered) images, where the noise is lowpass filtered in regions not immediately adjacent to the edge, which is where *SFR* is measured.

The Noise Image Method

A second and more accurate noise calculation uses a 2D image of the noise to calculate the Noise Power Spectrum, $NPS(f)$), enabling a more accurate calculation of information capacity as well as additional information metrics that characterize a camera's performance. $NPS(f)$ must be normalized by the edge variance noise, $N(x)$.

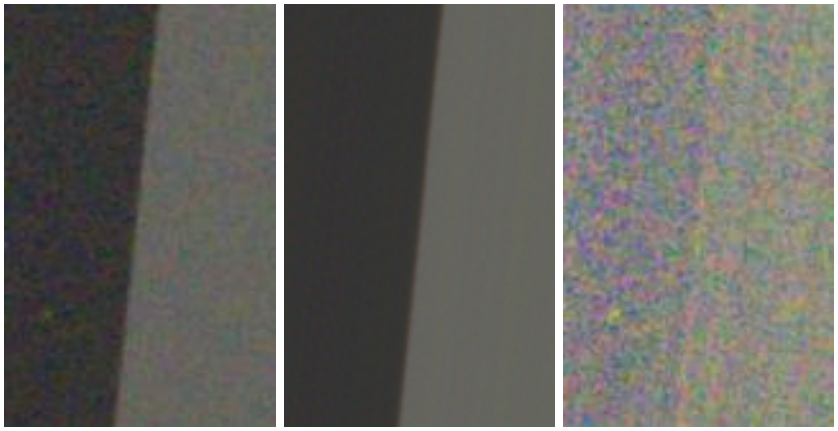
Because this method requires minimally or uniformly processed images, it is generally invalid for JPEG images from cameras, which are almost universally bilateral (nonuniformly) filtered.

The noise image

The 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 (1), we apply an inverse of the bin-

ning algorithm to set the contents of each scan line to its corresponding interleave (Inversely binned, below). Since the inverse-binned image (2) is a nearly noiseless replica of the original image (1), a noise image (3) can be created by subtracting the inverse-binned image from the original image(1) (which must be corrected for illumination nonuniformity in the direction of the edge).

The three images are shown in Figure 15. The noise image, which has a mean of zero, is displayed with an offset, lightened, and boosted in contrast for visibility.



1. Original image **2. Inverse-binned (de-interleaved)** **3. Noise image**

Figure 15. Noise image method, for a noisy (high ISO speed) image

$$\text{Noise image} = \text{Original image} - \text{Inverse-binned image} \qquad (12)$$

These images allow several additional image quality parameters to be calculated, including Noise Power (Wiener) Spectrum ($NPS(f)$) and Noise Equivalent Quanta ($NEQ(f)$), which is well-known in medical imaging systems, and described in an excellent review paper by Ian Cunningham and Rodney Shaw [16], as well as in the obscure but valuable ICRU Report 54 [17]. (ICRU is the International Commission on Radiation Units & Measurements.) These measurements are little-known outside of medical imaging, in part because they have been difficult to measure.

To repeat: the Noise Image method is not valid for bilateral-filtered images and has limited value for images that have been noise-reduced.

Noise Power Spectrum, *NPS*

$NPS(f)$, also called the *Wiener spectrum* (Figure 16), is used in the calculation of the key information metrics. The Noise Amplitude (Voltage) Spectrum, $N_V(f) = \sqrt{NPS(f)}$ is also of interest.

The 1D Noise Power or Voltage spectrum is derived from a 2D Fourier transform (FFT) of the noise image.

- Noting that $f = 0$ at the center of the 2D FFT image (after applying the MATLAB fftshift function), divide the image into several annular regions, then find the average noise power, $NPS(i)$,

for each region *i*. This procedure has been used for the Imatest Spilled Coins/Dead Leaves calculations since 2013, and has been tested thoroughly.

- Because this procedure does not maintain the invariance in energy between the spatial and frequency domains implied by Parseval's theorem, *NPS*(*f*) is normalized so that the one-dimensional integrals in frequency and spatial domain are equal.

$$\int NPS(f)\,df = \int \sigma_s^2(x)\,dx = \int N(x)\,dx \qquad (13)$$

$$NPS(i) = \frac{nps(i)\ \int N(x)\,dx}{\sum nps(i)} \qquad (13a)$$

Even though this paper focuses on metrics derived from the noise image method, the spatially dependent noise from the Edge variance method, $N(x) = \sigma_s^2(x)$, is useful for two purposes.

- to indicate of the type of image processing, depending on the presence of a peak,
- to normalize *NPS*(*f*), using Equation 13a, ensuring consistent scaling, even for small ROIs, where the de-binned image is not entirely noiseless.

Demosaicing typically causes the Noise Power Spectrum, *NPS*(*f*), to drop to about half its low frequency value at the Nyquist frequency, $f_{Nyq} = 0.5$ C/P.

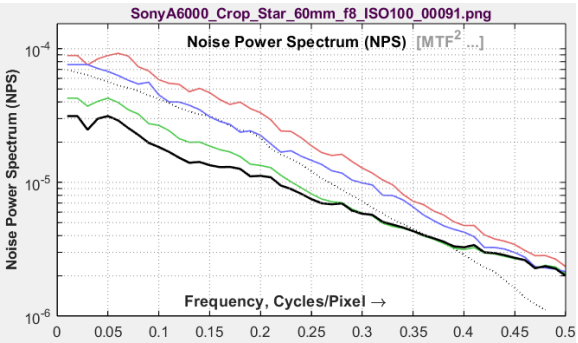


Figure 16. Noise Power Spectrum (NPS(f))

NPS(*f*) is used to calculate most of the image information metrics introduced in this paper—*NEQ*, C_n, *SNRi*, *Object Location σ*, and matched filter transfer functions.

It is a part of the ***kernel K(f)*** that appears in the equations for most of the metrics.

$$K(f) = SFR^2(f)/NPS(f) \qquad (14)$$

Equations will be written in both standard form and with *K*(*f*). Because uniform filtering affects $SFR^2(f)$ and *NPS*(*f*) identically, *K*(*f*) is not affected by uniform filtering, such as sharpening or lowpass filtering (as long as some response is maintained where $SFR^2(f)$ has significant energy).

The spatial-domain correlate of frequency domain functions divided by *NPS*(*f*) are sometimes described as "*noise-whitened*."

Noise autocorrelation, which is the inverse Fourier transform (IFT) of *NPS*(*f*), is potentially useful for evaluating the crosstalk between image sensor pixels, but the Bayer Color Filter Array (CFA) makes such measurements challenging.

Noise Equivalent Quanta, *NEQ*

NEQ(*f*) (Figure 17) is a frequency-dependent Signal-to-Noise (power) Ratio, related to the number of quanta that would result in the measured SNR when photon shot noise is dominant. It was described in 1999 by Cunningham and Shaw [16] and in 2016 by Keelan [18], and it is used in medical imaging [16, 17, 19].

$$NEQ(f) = \frac{\mu^2\, SFR^2(f)}{NPS(f)} = \mu^2\, K(f) \qquad (15)$$

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

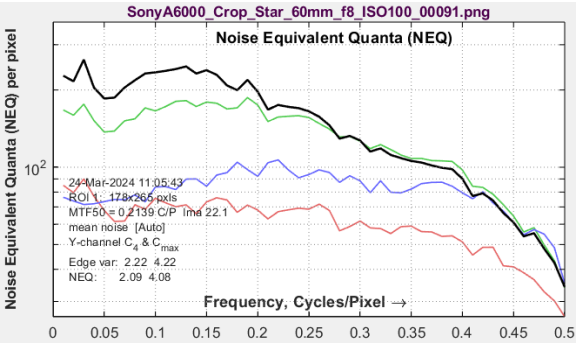


Figure 17. Noise Equivalent Quanta (NEQ(f))

In the standard definition of *NEQ*, $\mu^2 = V_{mean}^2$ (Figure 4). If *NPS* is dominated by photon shot noise, we can use the well-known property of the Poisson distribution, $SNR^2 = \bar{q}$ (where $\bar{q}$ is the mean count of the detected quanta.) to show that $NEQ(0) = \bar{q}$. For example, *NEQ*(0) = 200 corresponds to a mean of $\bar{q} = 200$ detected quanta per pixel. But because nonuniform illumination can increase NPS at the lowest spatial frequencies, causing NEQ to decrease, the maximum value of *NEQ* appears to be better for estimating $\bar{q}$.

The above equation can 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. This requires a separate (and exacting) measurement of $\bar{q}_i$ on the image sensor. The two levels of the slanted edge make *DQE* measurements particularly challenging.

NEQ(*f*) is unaffected by linear, uniform electronic filtering.

Information capacity from *NEQ*

$NEQ_{info}(f)$ (a variant of *NEQ*), calculated from (15) using $\mu = V_{P-P}/\sqrt{12}$ (to be consistent with the Edge Variance calculation for uniformly distributed levels), is signal power divided by noise power. It can therefore be substituted into the Shannon-Hartley equation to calculate information capacity, C_{NEQ}.

$$C_{NEQ} = \int_0^W \log_2\left(1 + NEQ_{info}(f)\right) df = \int_0^{0.5} \log_2\left(1 + \mu^2\, K(f)\right) df \qquad (16)$$

where $W = f_{Nyq} = 0.5$ Cycles/Pixel. [Author's note: I thought I discovered this connection, but I found it in papers on PET scanners and Digital Mammography by Christos Michail et. al. [20-21]—little known outside medical imaging.]

The key results, C_{4NEQ} and C_{maxNEQ}, are included in the Results summary (Figure 18). They are slightly different from the Edge Variance results, because $NPS(f)$ 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) =	3.87	4.57	4.02	4.66
Info capacity C_4(NEQ) =	1.61	2.26	1.72	2.36

Figure 18. Information capacity C from the two methods: Edge variance and Noise image.

Nomenclature

We generally omit the calculation type (Edge Variance or Noise Image (*NEQ*)) when referring to information capacity, *C*, because it can be deciphered from context.

For minimally or uniformly processed images, *C always* refers to C_{NEQ}, so we omit *NEQ*. We usually include the chart contrast ratio, i.e., C_n, for *n*:1 charts.

Ideal Observer SNR (*SNRi*)

SNRi (Figure 20) is a measure of the detectability of objects. It was introduced and rigorously correlated with Bayesian detection statistics in the 1996 ICRU Report 54 [17], then reintroduced to the imaging community Paul Kane [23] and Orit Skorka and Paul Kane [22]. The two-dimensional equation in [22] gives the best results.

$$SNRi^2 = \iint \left(\frac{\left|G(\nu_x, \nu_y)\right|^2 SFR^2(\nu_x, \nu_y)}{NPS(\nu_x, \nu_y)} \right) d\nu_x\, d\nu_y = \iint \left|G(\nu_x, \nu_y)\right|^2 K(\nu_x, \nu_y)\, d\nu_x\, d\nu_y \qquad (17)$$

where $G(\nu_x, \nu_y)$ is the two-dimensional Fourier transform of the rectangular object to be detected, $g(x,y)$, and $SFR(\nu)$ and $NPS(\nu)$ are defined in one dimension for spatial frequency $\nu = \sqrt{\nu_x^2 + \nu_y^2}$, which has units of Cycles/Pixel.

Objects to be analyzed are typically squares or rectangles of dimensions $w \times kw$, where $k = 1$ or 4 for a square or a 1:4 aspect ratio rectangle, respectively. Amplitude, $V_{P\text{-}P}$, is typically taken from the chart (usually with the ISO standard 4:1 contrast ratio), but can be modified for different values of Michelson contrast. The equation for the rectangular object (Figure 19) is

$$g(x,y) = V_{P-P} \cdot \text{rect}\left(\frac{x}{w}\right) \cdot \text{rect}\left(\frac{y}{kw}\right) \qquad (18)$$

where rect(x/w) = 1 for $-w/2 < x < w/2$; 0 otherwise

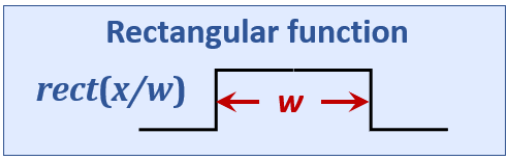


Figure 19. rect function

The 2D Fourier transform of $g\,(x, y)$ is

$$G(\nu_x,\nu_y) = kw^2\, V_{P-P}\, \frac{\sin(\pi w \nu_x)}{\pi w \nu_x}\, \frac{\sin(\pi k w \nu_y)}{\pi k w \nu_y} = V_{P-P}\, G_{rect}(\nu_x,\nu_y) \qquad (19)$$

where $G_{rect} = w\ \text{sinc}(\omega w/2) = w\ \text{sinc}(\pi w \nu)$ for $\omega = 2\pi f = 2\pi\nu$ is the Fourier transform of rect(x/w) for frequency ν. Note that G_{rect} has units of $1/\nu^2$, and since ν has units of cycles/(1D pixel), $G(\nu_x, \nu_y)$ has units of 2D pixels. (Note that pixels defined in one or two dimensions (1D or 2D) are equivalent. Using pixels as units is tricky.)

$SNRi^2$ 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 and 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 [23].

$$SNRi^2 = \Delta\nu_x\, \Delta\nu_y \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{SFR^2(i,j)\, G^2(i,j)}{NPS(i,j)} \qquad (20)$$

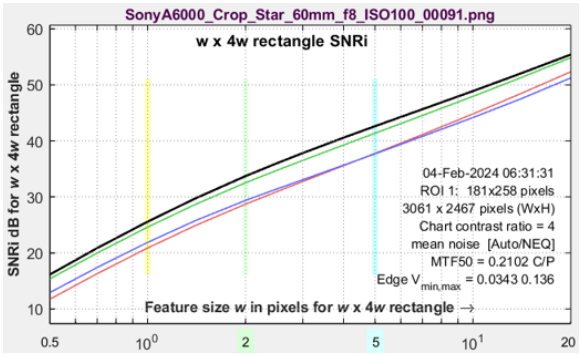


Figure 20. SNRi as a function of feature size w (in w x 4w rectangles) for a sharp, low-noise (ISO 100) image for R,G,B,Y channels

Note that like C_4, $SNRi$ is strongly affected by exposure and chart contrast. But $SNRi$ is also affected by Image Signal Processing (ISP; i.e., electronic filtering such as sharpening or lowpass filtering).

The InfoDR chart, which has slanted edges with a wide range of tones, or equivalently, exposure, allows a much more interesting of $SNRi$ (as a function of exposure) to be made.

As a result of Parseval's theorem, which states that the integrals of a Fourier transform pair, $r(x)$ and $R(\omega)$, must be equal,

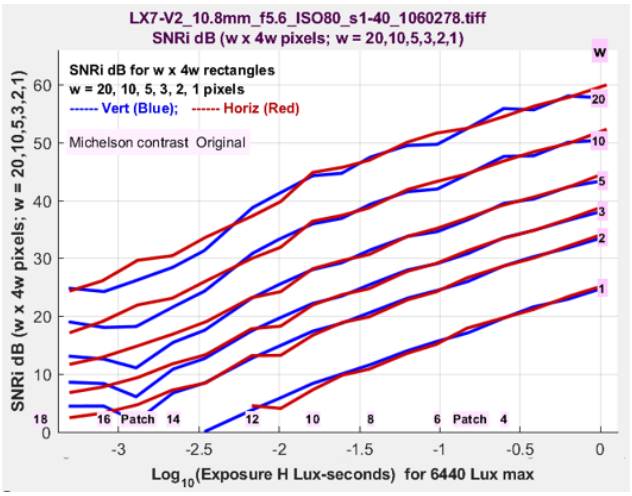


Figure 20a. SNRi for w × 4w rectangles with w = 1, 2, 3, 5, 10 & 20 pixels for compact camera.

$$\int_{-\infty}^{\infty}|r(x)|^2dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}|R(\omega)|^2d\omega = \int_{-\infty}^{\infty}|R(2\pi f)|^2df \qquad (21)$$

We observe that $SNRI^2$ is equivalent to the total (integrated) noise-whitened Signal/Noise energy of the object in the spatial domain.

Object Detection Error Probability, *ODEP*

Literature on $SNRi$ mentions that it is closely related to how well an object can be detected via Bayesian statistics, but it omitted the equation, so we derived it in the companion document, **Image quality testing based on information metrics, Appendix 2**.

$$\text{Error probability} = ODEP = \frac{1}{2}\text{erfc}\left(\frac{SNRi}{2\sqrt{2}}\right)$$

(Note that Error Probability is short for Object Detection Error Probability).

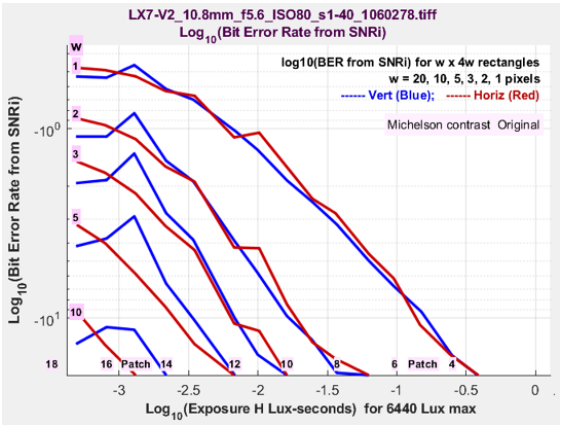


Figure 21. Log_{10} (Error Probability, ODEP) The y-axis can be linear or logarithmic.

This equation is based on an error threshold set for equal probabilities of false negatives and false positives. We added a plot of $\log_{10}$(*Error Probability*) to InfoDR (Figure 21, shown on the right, for the same camera and object geometries as Figure 20a): for $w \times 4w$ rectangular objects, where $w = 20$, 10, 5, 3, 2, and 1 pixel. Log_{10} was used because of the extreme range of error probabilities.

Task: Detection, Classification, Identification (DRI)

Figure 21 for Log_{10}(*Error Probability*) presents a dilemma: Error probabilities are vanishingly low for reasonable size objects. To resolve this, we introduced the concept of ***Task***. There are three generally recognized tasks. In order of increasing difficulty,

Task ("DRI" categorization)	
Detection	"Something is present" (Easiest; SNRi can be calculated from a large fraction of the object size)
Classification (recognition)	The most important task: It's a person, not a car, tractor, bicycle, etc." (More difficult. SNRi must be calculated from a small fraction of the object size.)
Identification	"It's Elvis!" (Most challenging)

Classification (also called recognition) is the most common task for automotive imaging, though not the only one. To predict system performance, the rectangles used to calculate *SNRi* and *ODEP* must be a small fraction of the total size of the smallest object of interest. This fraction may correspond to a feature size, for example, the hands or face of a person, which are roughly 1/10 of the person's total size.

Research is under way find the best relative size (as a fraction of total object size) to correlate *Error Probability* to measured system error rates. The relative size is dependent on the object properties and the task.

The rectangles used to calculate *ODEP* may also require different (usually lower) contrast ratios than the ISO 12233 standard C_R = 4:1 (C_{Mich} = 0.6). The contrast level can be set for the *SNRi*/*ODEP* displays.

SNRi displayed in dB per pixel
This display has not proven to be useful.

Because standard *SNRi* plots can be difficult to read (in part because *SNRi* has units of pixels), *SNRi* can also be plotted in dB per pixel (Figure 21a). It is somewhat easier to read than the standard *SNRi* image, but it is more of a *relative* measurement— useful for evaluating changes from image processing.

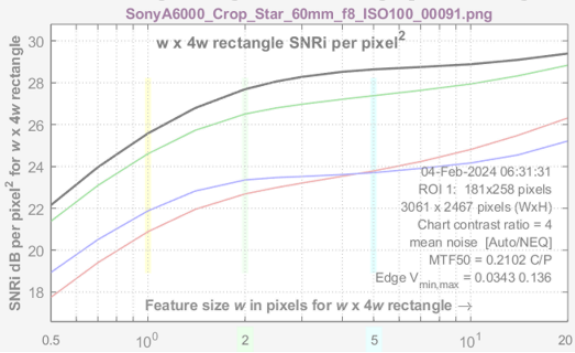


Figure 21a. SNRi per pixel for a sharp, low-noise (ISO 100) image

Object visibility and *SNRi*

One of the benefits of *SNRi* measurements is that they can predict object visibility for small, low contrast squares or rectangles. The *SNRi* prediction begs for visual confirmation.

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

We show two images, below: Figure 22 for a relatively low-noise image and Figure 23 for a noisy image (both from a camera with a Micro Four-Thirds sensor, at ISO 100 and 12800). The

sides of the squares are $w = 1, 2, 3, 4, 7, 10, 14$, and 20 pixels. The original chart contrast ratio of 4:1 (light/dark = 4), equivalent to a Michelson contrast $Contrast_{Mich}$ = (light-dark)/(light+dark) = 0.6, was used for the outer squares. The middle and inner squares have $Contrast_{Mich} = 0.3$ and 0.15, respectively.

How to use these images — The yellow numbers are the square widths in pixels. The outer (left and right) patches correspond to the *SNRi* curves for the ISO 12233-standard 4:1 contrast ratio, where, according to the Rose model [16], *SNRi* of 5 (14 dB) should correspond to the threshold of visibility.

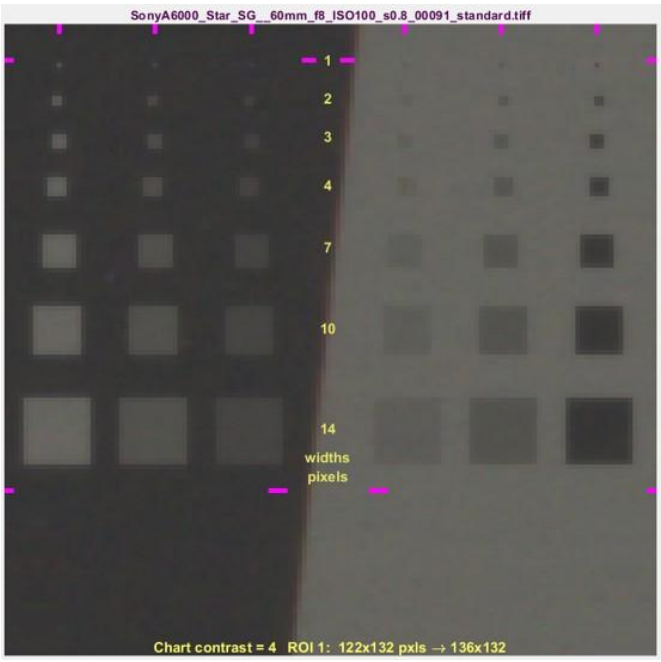


Figure 22. Low noise image, ISO 100.

Figure 23. Noisy image, ISO 12800.

The *SNRi* curve in Figure 24 is for the noisy ISO 12800 image in Figure 23, 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 Figure 23, the Y (luminance) channel *SNRi* at $w = 2$ is 15 dB; it reaches 19 dB for $w = 3$; a little above the expectation that the threshold of visibility is around 14 dB.

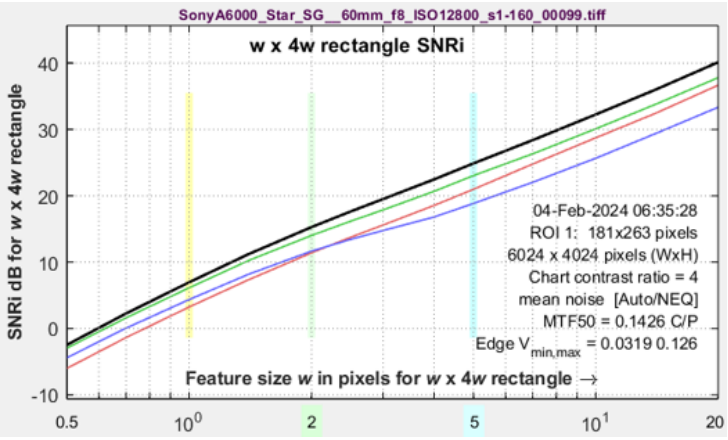


Figure 24. SNRi for noisy ISO 12800 image (above)

Only original pixels were used in these images of squares, but we used some "smoke and mirrors" tricks to make squares that have the same blur as the original image

How the squares were made

1. Expand the image if needed to make room for all the squares by adding mirrored versions of image to the sides, top, and bottom.
2. Create a horizontal mirror of the full image. This is the "mirror" part.
3. Create a mask consisting of ideal $w \times w$ squares, with 0 in the background, 1 in the squares, and sharp sides.
4. Blur the squares with the MATLAB filter2 function. This is the "smoke" part. Determining the blur kernel was challenging. We found that we did not 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, using the original image where the mask = 0, the mirrored image where the mask = 1, and blending them elsewhere.
7. Reapply the gamma encoding.

Edge SNRi and Object Location Standard Deviation (*Object Location σ*)

Object Location $\sigma = \sigma_{objLoc}$ (formerly called Edge Location σ) has been de-emphasized *for now* because it does not scale as expected. (It should increase with the object's circumference, but more slowly than *SNRi*, which is proportional to the object's *area*. We will restore it if/when we resolve the scaling issue.

Edge SNRi and its inverse, $\boldsymbol{\sigma_{objLoc}}$, are measures of how well an object can be located. The equation for *Edge SNRi* is similar to *SNRi*, described above and in [17, 22, 23].

$$Edge\ SNRi^2 = \iint \left(\frac{\left|H(\nu_x, \nu_y)\right|^2 SFR^2(\nu_x, \nu_y)}{NPS(\nu_x, \nu_y)} \right) d\nu_x\, d\nu_y$$
$$= \iint \left|H(\nu_x, \nu_y)\right|^2 K(\nu_x, \nu_y)\, d\nu_x\, d\nu_y \qquad (22)$$

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

Because *Edge SNRi* has units of spatial frequency (1/distance), we report its inverse, Object Location Standard Deviation, σ_{objLoc}, which has more intuitive units of distance. Lower is better for σ_{objLoc}.

$$Object\ Location\ \sigma = \frac{1}{Edge\ SNRi} \qquad (22a)$$

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 n:1 = 4:1 contrast ratio. *SNRi* and $1/\sigma_{objLoc}$ are both proportional to the Michelson contrast ($(n\text{-}1)/(n\text{+}1)$), and can be scaled for different contrast levels.

$$h(x,y) = V_{P-P} \cdot d\left[\text{rect}\left(\frac{x}{w}\right)\right]/dx \cdot d\left[\text{rect}\left(\frac{y}{kw}\right)\right]/dy = V_{P-P} \cdot I_I\left(\frac{x}{w}\right) \cdot I_I\left(\frac{y}{kw}\right) \qquad (23)$$

where $I_I(x/w) = d(\text{rect}(x/w)/dx$ (Figure 25) is called the "odd impulse pair," consisting of a pair of Dirac delta functions of opposite polarity separated by the object width w.

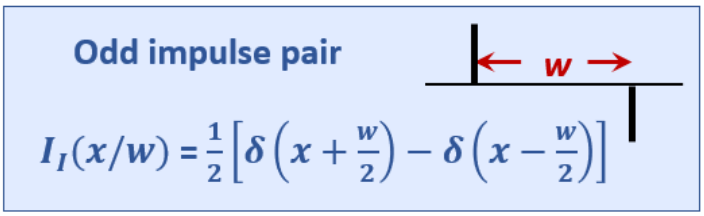


Figure 25. Odd impulse pair

$H(v_x,v_y)$ is the Fourier transform of the edges of the object to be detected, equivalent to $2\pi v\ G(v_x,v_y)$ for frequency v. Expressed in two dimensions,

$$H(v_x, v_y) = 2\ V_{P-P} \sin(\pi w v_x) \sin(\pi k w v_y) \qquad (24)$$

$H(v_x, v_y)$ is dimensionless.

Object Location $\sigma^2 = 1/Edge\ SNRi^2$ (Figure 26) is numerically calculated usin a similar equation to $SNRi^2$.

$$Edge\ Location\ \sigma^2 = 1/Edge\ SNRi^2 = 1/\left[\Delta v_x\, \Delta v_y \sum_{i=1}^{N_x}\sum_{j=1}^{N_y} \frac{SFR^2(i,j)\, H^2(i,j)}{NPS(i,j)}\right] \qquad (25)$$

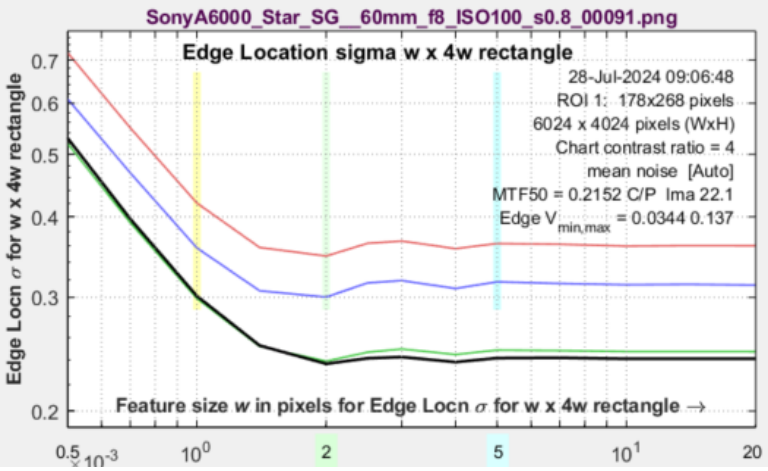


Figure 26. Object location σ for a sharp, low-noise (ISO 100) image

SNRi and *Object Location* σ are affected by signal processing (sharpening, etc.), making them useful for evaluating filtering (ISP filtering applied prior to the object classification/machine learning/AI blocks).

Line Spread Function (LSF) doublet results

Object Location σ is related to pairs of Line Spread Functions of opposite polarity called *LSF doublets*, $r(x)$ (Figure 27), which are used in several key calculations.

$$r(x) = \big(LSF(x) - LSF(x-w)\big)/\sigma \quad \text{and} \quad R(v) = \frac{H(v)\,MTF(v)}{\sqrt{NPS(v)}} \tag{26}$$

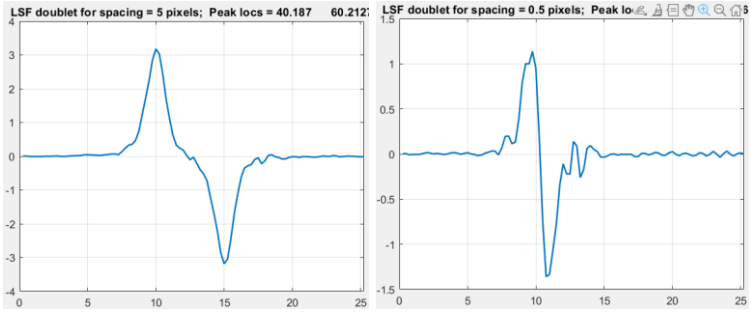


Figure 27. LSF doublets. w = 5.0 pixels (left), w = 0.5 pixels (right)
Amplitude for w = 0.5 is 1/3 as large as for w = 5.0 pixels.

As spacing *w* decreases, the peaks are shifted more from their original locations and amplitude decreases.

As a result of Parseval's theorem, *Edge SNRI²*, which is defined in frequency domain, is equivalent to the total (integrated) noise-whitened Line Spread Function doublet energy.

Effects of Image Signal Processing (ISP)

Several recent papers [24-26] state that appropriate image processing prior to Object Classification, Machine Vision or Artificial Intelligence algorithms may improve the performance (accuracy, speed, and power consumption) of AI systems. Because information capacity is independent of Image Signal Processing— at least with ISP that does not remove information, such as sharpening— it provides little guidance about filter design for optimal image processing.

SNRi has some drawbacks for predicting the quality object detection. It indicates how well an object's presence can be detected, but it says nothing about its *shape*. Shape detection is dependent on the edge detection, which is quantified by *Object Location* $\sigma = \sigma_{objLoc} = 1/Edge\ SNRi$. And there is the problem of object color. What If the object has the same color as the background? (Think of gray cars in front of gray concrete.) In such cases it is the *edge* that matters. For this reason, *Object Location σ* should be given comparable weight to *SNRi* when designing filters.

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 the detection of its edges. In practice, ISP needs to perform a multitude of tasks: detecting objects and edges greater than a minimum size and limiting interference from neighboring objects.

The results in the table below were obtained starting with an unsharpened image, applying various combinations of sharpening (Radius R, Amount A) and gaussian lowpass filtering (blurring) to the image, which was linearized prior to filtering, then restored to its original gamma encoding afterwards. *SNRi* and *Edge SNRi* are for $w \times 4w$ rectangles (w in pixels). $N_{uniform}$ was used for these images, which had no noise peak. (A peak was only present in the JPEG, which was not analyzed.) Standard (non-interpolated) binning was used.

Table 2. Results for various filters for a 24 MP Micro Four-Thirds camera, 1 inch sensor, ISO (EI) 100, *w* × 4*w* rectangle.

Filter (ISO 100)	MTF50 C/P	Obj Loc σ *pixels w* = 1	Obj Loc σ *pixels* lg *w*	SNRi dB/Pxl^2 w = 1	SNRi dB/Pxl^2 w = 5	C_{max} (NEQ)
None	0.214	0.306	0.246	25.5	28.6	4.08
σ=0.7 LPF	0.160	0.316	0.226	27.0	30.5	3.91
Sharp R2A3	0.541	0.370	0.303	23.5	26.2	3.8
Sharp R2A3, σ=0.7	0.335	0.336	0.261	25.1	27.8	3.86
Sharp R1A2 σ=0.7	0.265	0.303	0.239	25.8	28.7	4.09

Table 3. Results for various filters for a 24 MP Micro Four-Thirds camera, ISO (EI) 800, *w* × 4*w* rectangle.

Filter (ISO 800)	MTF50 C/P	Obj Loc σ pixels w = 1	Obj Loc σ pixels lg w	SNRi dB/Pxl^2 w = 1	SNRi dB/Pxl^2 w = 5	C_{max} (NEQ)
None	0.216	0.767	0.593	18.2	22.2	2.96
σ=0.7 LPF	0.156	0.654	0.476	20.4	24.5	2.87
Sharp R2A3	0.529	0.814	0.666	16.9	20.3	2.82
Sharp R2A3, σ=0.7	0.342	0.703	0.563	18.4	21.7	2.83
Sharp R1A2 σ=0.7	0.274	0.705	0.555	18.7	22.4	2.96

We can make several observations from these results.

- A pure lowpass filter (LPF $\sigma = 0.7$) improves *Object Location σ* and *SNRi*.

- Pure sharpening (R2A3) degrades both metrics. This illustrates how strong oversharpening, which is often used boost summary metrics like MTF50, can degrade performance. It can be identified by strong edge "halos" and SFR peaks (Figure 13).
- Sharpening + LPF slightly improves *Object Location σ* for small *w*, and improves other metrics somewhat less than a pure LPF. However, the sharpened + LPF images will be less susceptible to interference from neighboring objects.

The bottom line is that appropriate lowpass filtering as well as lowpass filtering with some sharpening can potentially improve performance. Of course, these are just a few of many filter combinations of potential interest. And they are only for one specific camera at two ISO speeds (Exposure Indices 100 and 800).

And as we indicated, filters should be designed for more than a single task. The pure LPF performs well in the absence of interfering objects. Some sharpening may reduce the effects of interference, and as we have shown, sharpening, when combined with lowpass filtering, causes little performance degradation.

Matched filters

A matched filter [27] (sometimes called a "noise-whitened matched filter") is a custom filter that maximizes the SNR, i.e., the detection probability, for

- a system with a specific response, and
- a specific object (or edge).

Matched filters were originally developed for impulse detection in radar (a single airplane at a large distance). They are discussed in ICRU Report 54 [17], but were mostly ignored outside medical imaging because they were not directly relevant to human vision.

For an impulse (a δ-function, i.e., a distant airplane), the matched filter transfer function is identical to the noise-whitened system response (where "noise whitening" is division by $N_V(f)$, or equivalently, $\sqrt{NPS(f)}$).

$$\mathcal{F}_{matched}(f) = SFR(f)/(\sqrt{NPS(f)}) = \sqrt{K(f)} \qquad (26)$$

A *matched filter*, $\mathcal{F}_{matched}(f)$, which optimizes *SNRi* or *Object Location σ*, has the same frequency spectrum as the system, including the edge or object. For edge or object detection,

$$\mathcal{F}_{matched}(f) = \frac{|P(f)|\,SFR(f)}{\sqrt{NPS(f)}} = |P(f)|\sqrt{S(f)} \qquad (27)$$

P(*f*) is equal to *G*(*f*) (Equation (19); the Fourier transform of the object) for *SNRi* or *H*(*f*) (Equation (24); the Fourier transform of the edge) for *Object Location σ*.

Figures 28 and 29 show the transfer functions for matched filters for *SNRi* or *Object Location σ*. A Sharpening + LPF filter designed to approximate the matched filter for *Object Location σ* (Figure 29) has a response peak near 0.20 Cycles/Pixel (Figure 30).

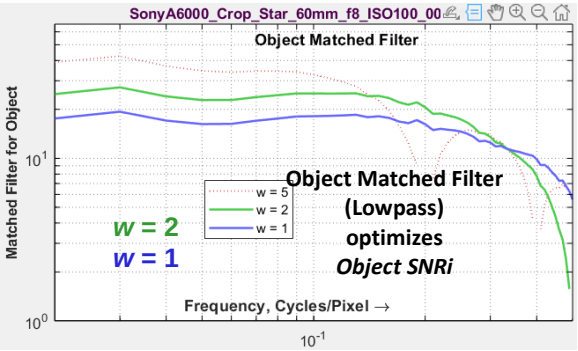


Figure 28. Matched filter for optimum object detection: Lowpass (LPF).

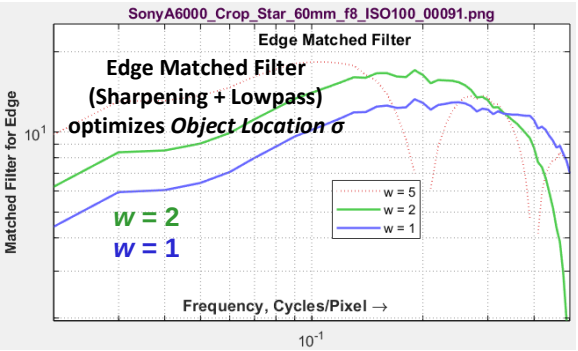


Figure 29. Matched filter for optimum edge detection. Can be approximated with sharpening + LPF.

Matched filters are optimized for a single task: detecting an object or edge of a certain size (by maximizing *SNRi* or minimizing *Object Location* σ). But real-world filters must perform a multitude of tasks: they must detect objects and edges of varying sizes, contrasts, and colors. This calls for tradeoffs, which are fortunately quite mild. Since large objects are usually detected well, filters should be designed to perform well with small objects or edges.

Figure 30 shows the transfer function of a filter with a $\sigma = 0.7$ gaussian LPF + sharpening with Radius = 1 and Amount = 2 (R1A2G07).

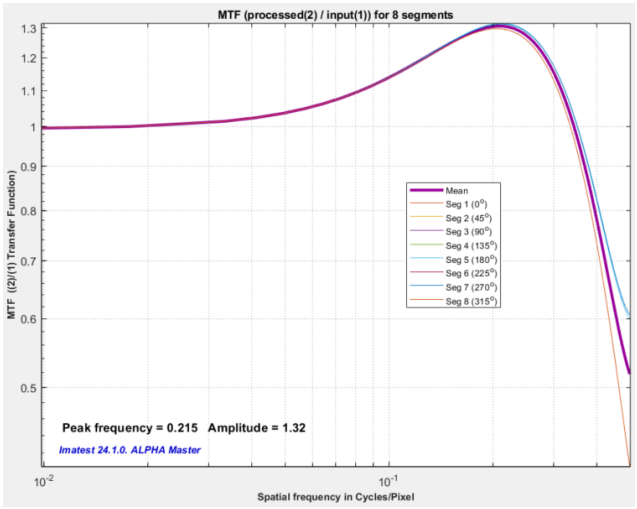


Figure 30. Filter transfer function for σ = 0.7, Radius = 1, Amount = 2; R1A2.

This transfer function approximates the edge matched filter in Figure 29. The matched filter can be realized by combining a standard lowpass filter (Bessel, Butterworth, etc.) with standard sharpening or unsharp masking (USM). Details for realizing this filter are beyond the scope of this document.

InfoDR chart, exposure, and pixel level

Most information-based performance metrics, including *NEQ(f)*, C_4, *SNRi*, and *Object Location σ*, (but not *SFR* or C_{max}) are sensitive to exposure. For this reason, we need to maintain consistent exposure when acquiring images for comparing different cameras. Linearized Digital Numbers (DNs or pixel levels) should be similar. A reasonable level needs to be established as we develop best practices for measurement.

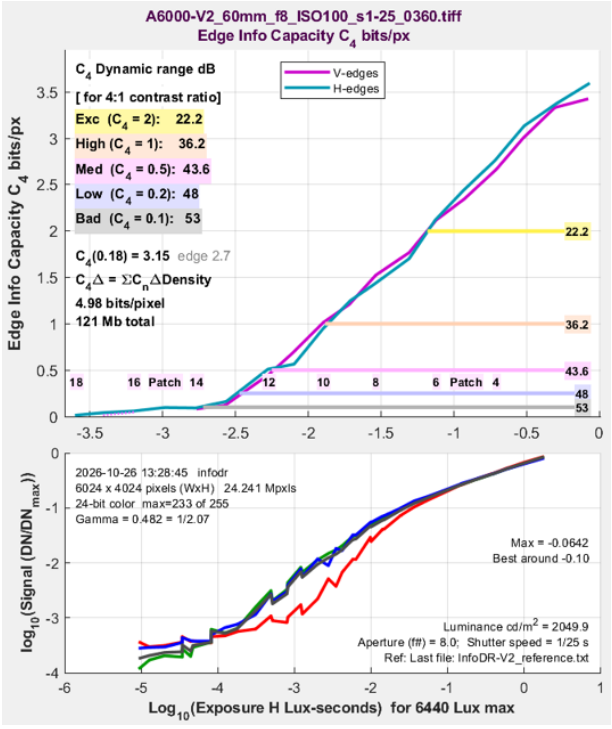


Figure 31. C₄ vs exposure and information-based Dynamic Range from InfoDR chart

This can be accomplished with the **Information-based Dynamic Range (InfoDR)** test chart, which is described in the **InfoDR** section of the white paper, **Image quality testing based on information metrics** [31] and in three documentation pages: **Part 1: InfoDR Introduction** – **Part 2: Using InfoDR** – **Part 3: InfoDR Results**. It allows C_4 information capacity and other information metrics (for a selectable contrast level), to be measured over a wide range of illumination.

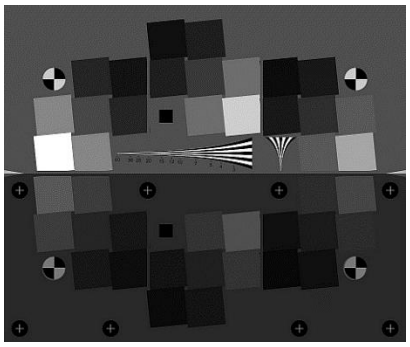


Figure 31a. InfoDR chart: 2-layer transmissive version

C_4 as a function of illumination — one of many available plots — is shown in the upper plot, above. ***Information-based Dynamic Range*** is displayed as text on the upper-left and horizontal bars on the right for C_4 levels, 2, 1, 0.5, 0.2, and 0.1 bits/pixel. Object detection tends to fail when C_4 drops below 0.1 (possibly 0.2) b/p.

Heuristic and experimental metrics

From **Wikipedia**: A **heuristic** or **heuristic technique** (*problem solving*, *mental shortcut*, *rule of thumb*) is any approach to problem solving that employs a pragmatic method that is not necessarily optimized, perfected, or rationalized, but is nevertheless "good enough" as an approximation or attribute substitution.

In the process of developing information metrics, we found several measurements, which we regard as "heuristic" or "experimental," that are potentially useful, but have not been rigorously proven. Research is needed to confirm their validity.

$C_4\Delta$ — The InfoDR plot of C_4 versus Exposure in Figure 31, shown above, has an interesting heuristic result, $C_4\Delta = \sum C_4\ \Delta$ Density, which is the area under the C_4 curve. It appears to be closely related to overall image quality. Values of $C_4\Delta$ are comparable to C_{max}, which is difficult to calcu-

late for HDR (High Dynamic Range) sensors. For the high-quality APS-C camera in Figure 31, $C_4\Delta = 4.98$ bits/pixel, while $C_{max} \cong 4.6$ bits/pixel. It may be a good substitute for C_{max} for HDR sensors. Verification is definitely needed for this metric.

$C_4 \times$ Area — is a simple way to approximate *SNRi*, which has a rather complex equation.

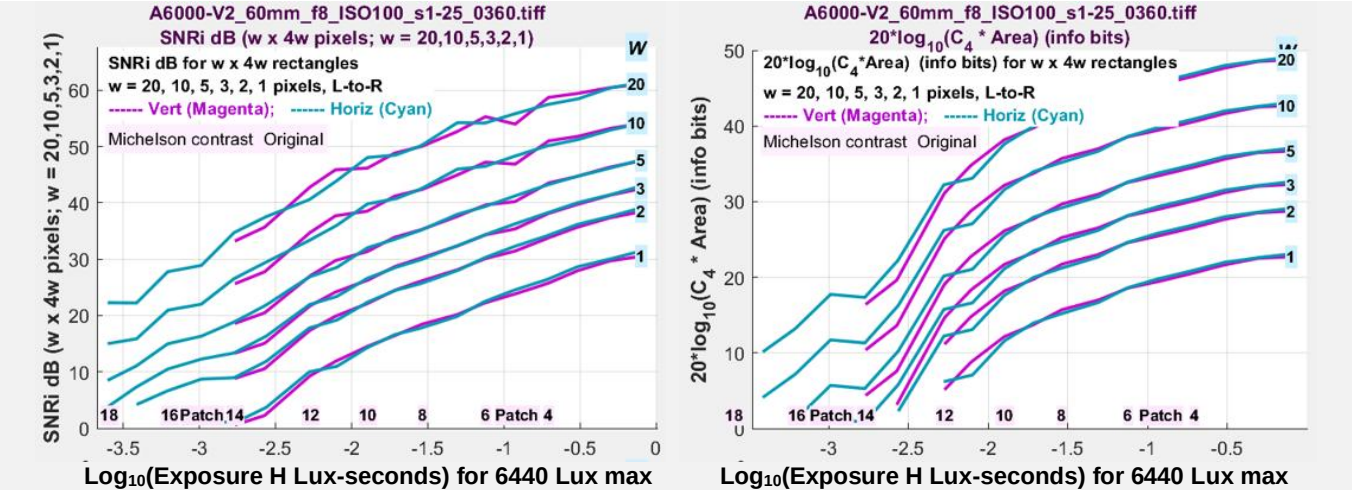


Figure 32. SNRi dB and 20×log10(C₄×Area) as a function of exposure

Summary

The basic premise behind our work on information metrics is that traditional sharpness and noise measurements are poor predictors of Machine Vision/Artificial Intelligence system performance. Also, they provide no insight into how to design optimum image processing. The new metrics and techniques described here are intended to accomplish this. They include

1. A two-step method for measuring camera noise from slanted edges:
   a. spatially dependent noise, $N(x) = \sigma_s^2(\mathrm{x})$, from the edge variance method, and
   b. noise power spectrum, *NPS*(*f*), from the noise image method, which results in a more accurate information capacity estimate and enables other information metrics to be calculated. This method only applies to minimally or uniformly processed images.
2. A set of metrics for quantifying object and edge detection performance, derived from the noise image method, including *NPS*(*f*), *NEQ*(*f*), C_n (typically C_4), *SNRi*, *Object Location σ*, and Object Detection Error Probability (*ODEP*).
3. A technique for designing matched filters for optimizing detection performance, based on a tradeoff between optimizing *SNRi* and *Object Location σ* and minimizing interference from nearby objects.

The equations for the information metrics all contain the equation kernel, $K(f) = SFR^2(f)/NPS(f)$. This unifies the results from the noise image method.

We are working on verifying that these calculations and design techniques correlate with MV/AI system performance.

The key concepts in this document are

1. Information capacity, which combines sharpness, noise, and contrast loss, is a fundamental figure merit for imaging systems that is appropriate for selecting cameras.
2. Measuring noise at the *same location* as the slanted-edges used to calculate *SFR* (*MTF*), i.e., co-locating the noise and signal measurements makes the results convenient, robust, and reduces the likelihood of error.
3. A noise peak in $\sigma_s(x)$ allows bilateral-filtered images to be distinguished from uniformly-processed images for "black box" cameras with unknown image processing, allowing the optimum noise calculation can be selected.
4. Information metrics, including, C_n, for n:1 contrast slanted edges (typically 4:1), are sensitive to exposure, making them useful for measuring low light performance, especially when the InfoDR test chart is used.
5. *SNRi* and Object Detection Error Probability (*ODEP*) are measures of how well an object can be detected. *ODEP* is most useful with a *task* (Detection, Classification, Identification) and a measurement object size relative to the total object size are defined. This is a subject of ongoing research.

Camera information capacity and related information metrics are still novel in the imaging industry. Significant effort will be required to make them better known; it will be a struggle until ISO 23654 is published. (We encourage readers to consider joining the ISO TC42 committee behind it.) A positive note: the units for *C*— information bits per pixel (or total image) for a specified ISO speed or exposure— are intuitive and easy to understand.

We would like to see information metrics, including information capacity, *SNRi*, and *Object Location σ*, become standard specifications for cameras intended for machine vision. Cameras can be characterized by their performance over a range of ISO speeds (exposure indices) and/or light (lux) levels. We are optimistic that this will lead to improved performance and reduced energy use [28].

Future work

- Continue collaborating with partners in academia and industry to correlate camera information capacity, *C*, and object detection metric, *SNRi* with the performance of Machine Vision and Artificial Intelligence systems. This may involve using image data sets like Audi A2D2 [29] to determine mean Average Precision (mAP) and Intersection over Union (IoU) [30]. Information Capacity as a Predictor of Perception Performance (*Imatest* post) describes a paper of the same name by Diarmaid Geever et. al. from the University of Galway. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=11355802
- Verify the calculation validity gain insight into the physical meaning of the *Object Location σ* metric, which is similar to 1/*SNRi*, but has not has not been subjected to the same rigorous verification [17, 23]— correlating it with Bayesian statistics.
- Become familiar with the measured numbers for the metrics (e.g., what are "good" values of *SNRi* or *Object Location σ*?)
- Transform units from native sensor native units of cycles/pixel to practical units like cycles/angle or cycles per object distance, as needed.

- Determine best practices for measuring the information capacity of High Dynamic Range (HDR) sensors.
- Determine best practices for designing the output (matched) filter: How much weight should *SNRi* and *Object Location σ* be given? How much sharpening is needed to limit external interference?
- Study the effects of demosaicing, which may include nonlinear processing that enhances edges, on information capacity.
- Continue working on the ISO 23654 standard, Photography - Digital cameras - Image Information Metrics, overseen by ISO TC42.
- Explore the correlation between *C* with the subjective visual appearance of images. This is challenging because visual perception is strongly affected by image processing (sharpening, color balance and saturation, tonal response, etc.).

References

[1] ISO/DIS 12233 Digital Cameras Resolution and spatial frequency responses, https://www.iso.org/standard/88626.html.

[2] 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.

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

[4] R. Shaw, "The Application of Fourier Techniques and Information Theory to the Assessment of Photographic Image Quality," Photographic Science and Engineering, Vol. 6, No. 5, Sept.-Oct. 1962, pp. 281-286. Available for download from https://www.imatest.com/wp-content/uploads/2022/12/Rodney-Shaw_Fourier_Information_full.pdf.

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

[6] Frédéric Cao, Frédéric Guichard, Hervé Hornung, "Information capacity: a measure of potential image quality of a digital camera," DxO Labs, 2010.

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

[8] J. C. Dainty and R. Shaw, "Image Science," Academic Press, 1974.

[9] Francis T. S. Yu, "Optics and Information Theory," Wiley, 1976.

[10] Tomasi, C., and R. Manduchi. "Bilateral Filtering for Gray and Color Images". Proceedings of the 1998 IEEE International Conference on Computer Vision. Bombay, India. Jan 1998, pp. 836–846.

[11] N. Koren, "Interpolated ISO 12233 slanted-edge SFR (MTF) calculation," www.imatest.com/2024/01/improved-slanted-edge-mtf-calculation/

[12] Peter Burns, Burns Digital Imaging, http://burnsdigitalimaging.com

[13] N. L. Koren, "Measuring camera information capacity with slanted-edges (Invited)" Electronic Imaging, 2023, pp 454--1 - 454-9, https://doi.org/10.2352/EI.2023.35.8.IQSP-454

[14] Orit Skorka et. al., "Evaluation of Signal and Noise Metrics of High Dynamic Range Image Sensors by IEEE P2020 Methodology," presented at Electronic Imaging 2024.

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

[16] 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-optics Image Science and Vision

[17] ICRU Report 54, Medical Imaging – The Assessment of Image Quality, Bethesda, MD: International Commission on Radiation Units and Measurements, 1996.

[18] Brian W. Keelan, "Imaging Applications of Noise Equivalent Quanta," 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.

[19] 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. https://doi.org/10.1118%2F1.4730287

[20] Michail C, et. al., Information Capacity of Positron Emission Tomography Scanners. *Crystals*. 2018; 8(12):459. https://doi.org/10.3390/cryst8120459.

[21] Christos M. Michail, et. al., "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.

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

[23] 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

[24] 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

[25] 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

[26] 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/mi12010073

[27] Nicholas R. Rypkema, "A Straightforward Derivation of the Matched Filter," nrr.mit.edu/sites/default/files/documents/Matched_Filter.pdf

[28] Soumya Sudhakar, "Data centers on wheels: the carbon footprint of self-driving cars," TEDxBoston, November 2022.

[29] Audi Autonomous Driving Dataset (A2D2) https://www.a2d2.audi/a2d2/en.html

[30] mAP Explained. https://www.v7labs.com/blog/mean-average-precision

[31] *Imatest* white paper: Image quality testing based on information metrics, https://docs.imatest.com/assets/uploads/IQ_testing_infometrics_whitepaper.pdf

Author Biography

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

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 [2,3] 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 pos-

sible 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.

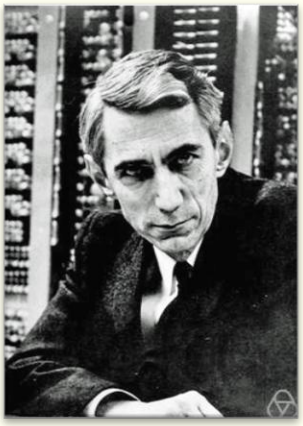
Now, suppose you have a weirdly warped coin that has a probability of 0.95 for a head (1) and 0.05 for a tail (0). Much less information is gained from the results of a flip. The equation for the information in a trial with m outcomes, where $p(x_i)$ is the probability of outcome i and $\sum_{i=1}^{m} p(x_i) = 1$, 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 nearby.) 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 [4].)

Appendix 2. Binning noise

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 {-Δ/2, Δ/2}, where Δ = 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 σ_{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 $LSF(x)$. 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

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

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

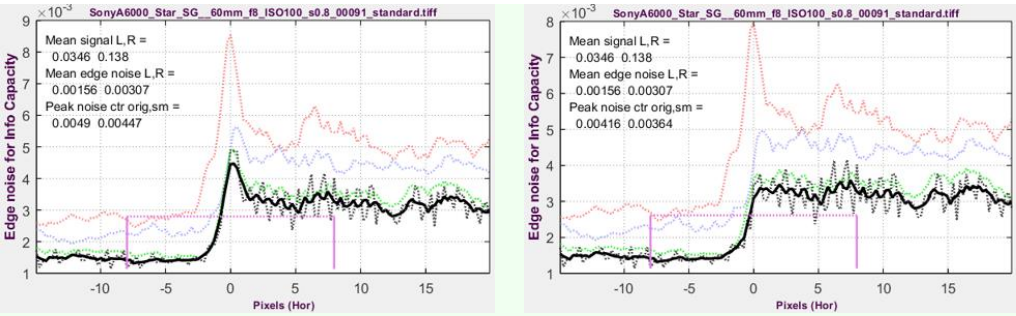


Figure A2-1. Edge noise for a Micro Four-Thirds digital camera, ISO 100, Y (Luminance) channel from raw image converted to TIFF with minimal processing.
Left: with binning noise Right: binning noise removed

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)**.

Appendix 3. Edge location variance (or standard deviation)

An additional result can be derived from the Edge Variance method: ***The edge location variance,*** $\boldsymbol{\sigma^2_{EdgeLoc}}$ **(*or standard deviation,*** $\boldsymbol{\sigma_{edgeLoc}}$**)**. It refers to a single edge, not an entire object, and must be distinguished from Object Location σ. (The nomenclature has been changed to reduce potential confusion, which was significant when $\boldsymbol{\sigma_{EdgeLoc}}$ was introduced.)

Edges are important because they are often required to distinguish objects. For example, to distinguish a gray vehicle (very popular in 2026) from a gray concrete background, edges are critical.

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

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

The actual location of an edge is affected by interference from neighboring edges (mostly the closest edge) as well as $\sigma_{edgeLoc}$. When edges are close together (small w), interference causes edge amplitudes to decrease and edge locations to shift. Displays of *LSF* amplitude and shift versus spacing are shown below for an unprocessed image and the same image moderately sharpened with Radius = 1 and Amount = 2. $\sigma_{edgeLoc}$ is not a major metric, but is of definite interest. *SNRi* and Object Location σ are more useful for quantifying system performance.

$\boldsymbol{\sigma_{edgeLoc}}$ is displayed in a rather obscure ***Imatest*** display; for **Info-related: NPS, NEQ, SNRi…** plots, select **LSF Doublet shift**. (In the future, we may make it easier to find.)

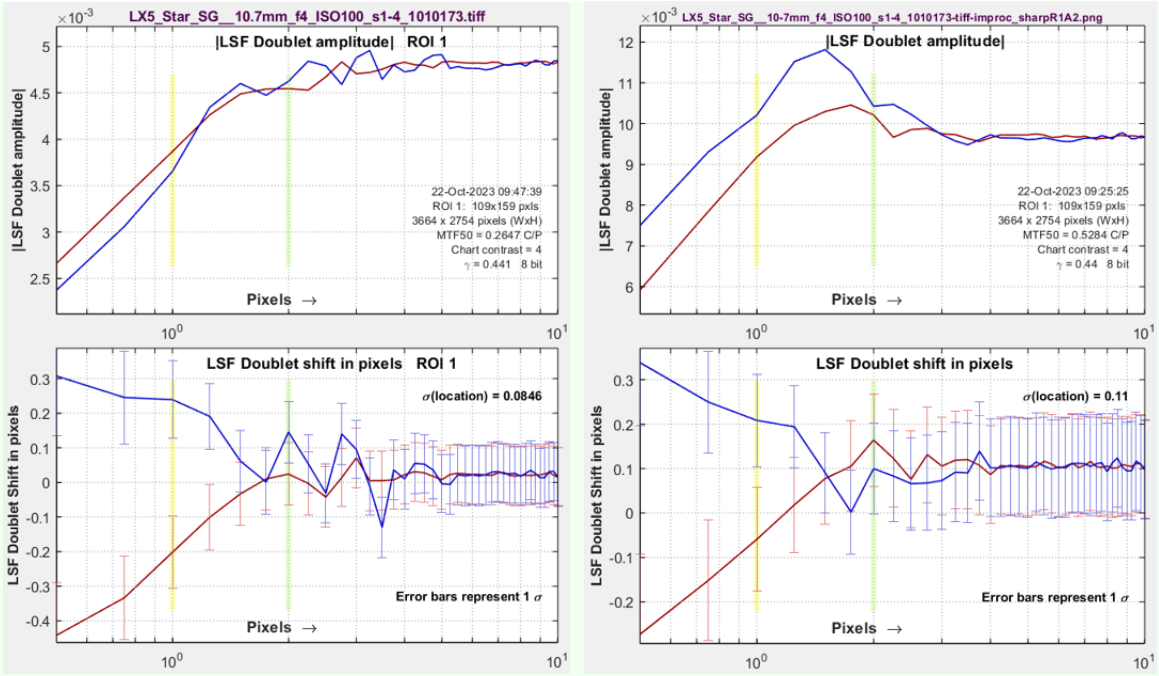
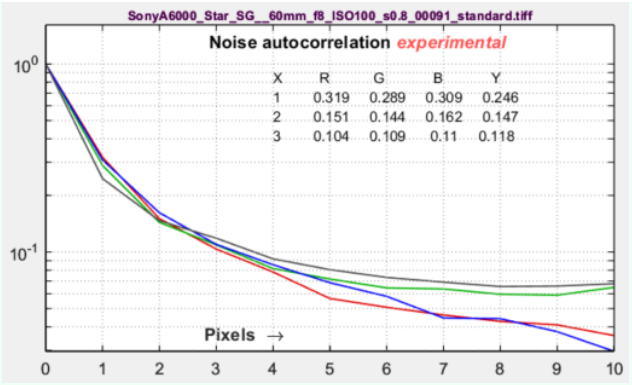


Figure A3-1. LSF doublet amplitudes and shifts
TIFF unsharpened $\sigma_{EdgeLoc}$ = 0.0846 px Sharpened with Radius = 1, Amount = 2. $\sigma_{EdgeLoc}$ = 0.0846 px

The difference between the unsharpened and sharpened (R1A2) results is not large. At a pixel spacing of 1 (labeled 10^0 on the x-axis), the difference between $\Delta LSF_{maximum} + \sigma$(Location) and $\Delta LSF_{minimum}$ - $\sigma_{edgeLoc}$ 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, $\sigma_{edgeLoc}$ 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, $\sigma_{edgeLoc}$ would have been much larger and there would be less improvement with sharpening.

Appendix 4. 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, and hence is closely related to the sensor's *SFR*, when the effects of demosaicing and fixed-pattern noise are removed and the primary noise source is photon shot noise.

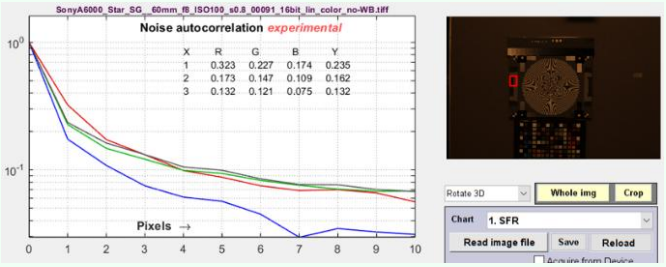


The idea behind the hypothesis is that light (potons) incident on the sensor is uncorrelated, so that if there were no crosstalk, the noise would be white.

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.

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 autocorrelation