Author's note

This is the presentation from January 21, 2024. It's a little rough. I am working on the final paper (to be published by Imaging.org), which is due February 15.

The **improvement to the ISO 12233 Edge SFR calculation** was implemented on January 14.

The features described here are available in the **Imatest Pilot Program**. They are not yet fully documented. The documentation and the final paper should be available by mid-February.

Many of the details in the presentation are in the white papers linked from **www.imatest.com/solutions/image-information-metrics**

Image Information Metrics from Slanted Edges

Norman Koren Imatest LLC

January 2024

A new set of metrics, derived from information theory, for selecting cameras and measuring and optimizing their object and edge detection performance, intended for machine vision and artificial intelligence systems

Image Information Metrics

Applications include automotive imaging (driver assistance and autonomous vehicles), robotics, security, manufacturing, and medical imaging.

Basic premise:

Conventional MTF and noise metrics are insufficient for the above tasks.

“There are reliable differences in the object detection performance between systems with the same MTF50. Far larger effects are caused by variations in the illumination conditions.” Zhenyi Liu et. al., presented by Brian Wandell, EI 2023

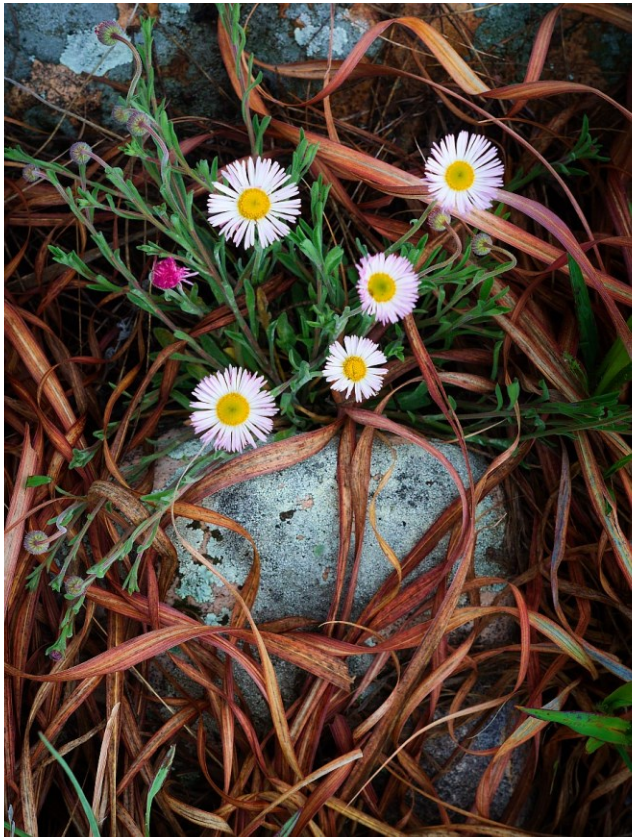
Outline of the presentation

- **Personal history: How I combined my backgrounds in photography and engineering (information theory) to create the new metrics**
- **Basic concepts of information capacity, *C***
- **The slanted-edge MTF calculation & recent improvements**
- ***Two* new methods for calculating noise (and hence *C*) in the presence of a signal — the widely used slanted edge**
- **Metrics related to *C*, including *NPS*, *NEQ*, and most importantly, *SNRi* and *Edge SNRi* (metrics for object and edge detection)**
- **Matched filters to optimize object and edge detection**

We will omit important but unrelated subjects such as Dynamic Range, Tone Mapping, stray light, and the differences between human and machine vision.

Background — photography

- **Grew up in Rochester, NY. "Kodak city" Frequently visited George Eastman House. Fascinated by both the fine prints and the cameras.**
- **Interest in photography started around age 12. Dissatisfied with sharpness of early cameras.**
- **Summer job University of Rochester Institute of Optics, 1961. MTF curves.**
- Master's degree in physics and 34 year career in magnetic recording technology
- **Photography was my primary hobby. Mastered darkroom printing; had occasional shows. Taught evening class on making high quality images from 35mm film, 1972-3.**
- **Launched normankoren.com (images and technical tutorials) in 2000, which led to founding *Imatest*.**



imatest®

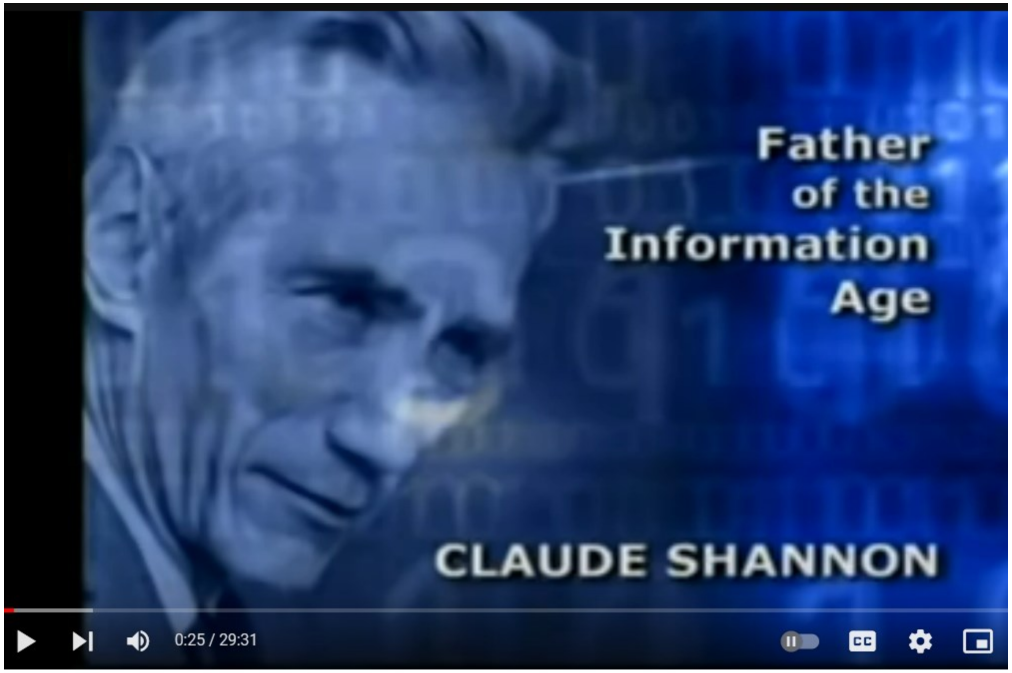
Background — engineering

- Masters degree in physics
- 34-year career in magnetic recording technology (1967-2001): modeling and disk and tape drive performance, designing read/write channels

  The math for channel analysis (especially pulse slimming) is identical to image sharpening.
- Kodak San Diego, 1985-1998: Frequently visited UCSD Center for Magnetic Recording Research (CMRR), which produced the video,

  [Claude Shannon – Father of the Information Age](www.youtube.com/watch?v=z2Whj_nL-x8).

  Became acquainted with information theory.



Claude Shannon - Father of the Information Age

www.youtube.com/watch?v=z2Whj_nL-x8

What is information?

Information, defined by Claude Shannon in his classic 1948 and 1949 papers, is a measure of the resolution of uncertainty, i.e., how much is learned from the outcome of a measurement.

For a system with n possible states, $s_1, \ldots, s_n$, with probabilities $p(s_1), \ldots, p(s_\mathrm{n})$, information can be represented as ***entropy***,

$$H(S) = \sum_{i=1}^{n} p(s_1) \log_2(1/p(s_1)) = -\sum_{i=1}^{n} p(s_1) \log_2(p(s_1))$$

Example: a "fair" coin flip. $p_1 = p_2 = 0.5$; H = entropy = .5 + .5 (information gained from the flip) = 1 "bit".

When one outcome is more probable than the other, the information gained in the trial is lower. For example,

For $p_1 = 0.95$; $p_2 = 0.05$, H = 0.95*0.074+.05*4.322 = 0.286.

The Shannon-Hartley equation

Shannon showed that an electronic channel has an *Information capacity, C*, which is the maximum rate that it can transmit information without error.

Shannon-Hartley *equation.*

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

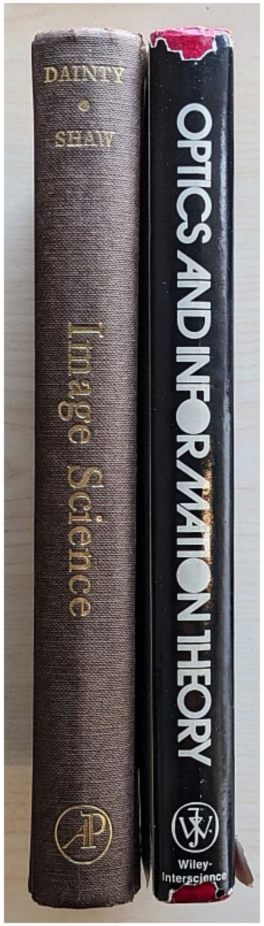
Key inputs: ***bandwidth W, average signal power S, and average noise power N*** Units of bits/pixel or bits/image

A camera is such a channel.

Lens (S(f)) → Sensor (S(f),N(f)) → Electronics (N(f))

Several books (Dainty & Shaw (1974), Francis T.S. Yu (1976)) discussed information theory. But they failed to gain traction because they didn't offer convenient ways to measure *C*.

They were *ahead of their time*.



Founded Imatest in 2003 — early work

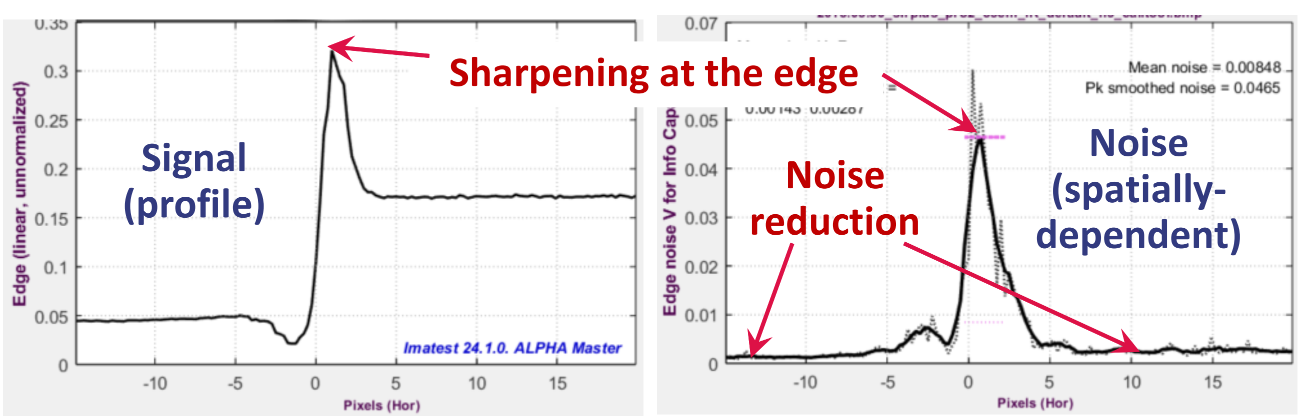
Started with standard measurements: MTF, noise, distortion, tonal response, dynamic range, color, etc.

We sought a way to compare “black box” cameras – with unknown image processing (often very different amounts of sharpening).

Information capacity was promising, but there was a problem.

Image processing can be very different near edges (often sharpened) where MTF is measured, and in flat areas (often smoothed) where noise is measured.

This increases measured values of *C*, even though information is actually *removed* from the image.



An extreme case of bilateral filtering, but *real*.

The quest

We realized that the key to obtaining reliable and convenient information capacity measurements was to measure signals and noise at the *same location*.

We didn't know how to do this.

This started us a long quest for the "holy grail" of image quality metrics:

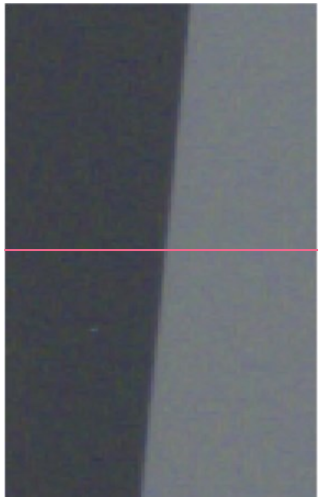
Measurements of MTF and noise made at the same location, enabling convenient, reliable calculations of information capacity

Starting in late 2022 we discovered *two* methods to accomplish this with the widely used slanted edge.

imatest®

Review of the slanted-edge algorithm

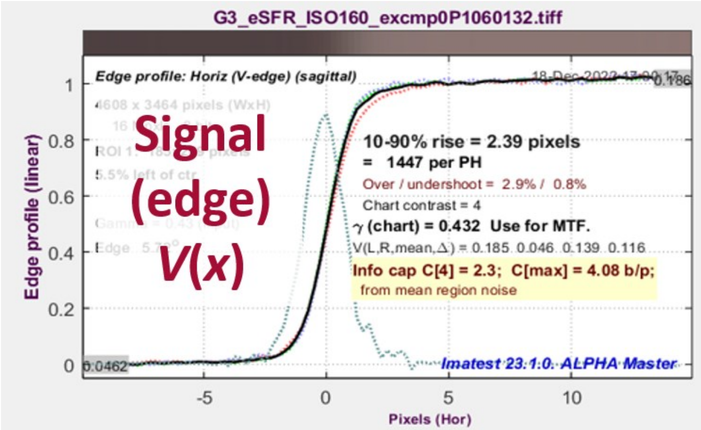
- **Linearize the image,**
- **Find the centers of each scan line**
- **Fit a polynomial curve to the centers,**
- ***New in 2024*: *Interpolate the scan lines to improve MTF accuracy, especially at high spatial frequencies.***
- **Add each shifted scan line to one of four bins to obtain a 4x oversampled average edge, which can be used to calculate *MTF* from the Fourier transform of $d\mu(x)/dx$.**



$$\mu_s(x) = V(x) = \frac{1}{L}\sum_{l=0}^{L-1} y_l(x)$$

Averaging improves SNR by

$$\sqrt{\text{samples in each bin}} \cong \sqrt{L/4}.$$



imatest®

Problem with standard ISO 12233 *MTF* algorithm
artifacts (that look like noise) at high spatial frequencies

- Inconsistencies were observed in some calculations at high spatial frequencies.
- Many authors have commented on this *MTF* measurement issue.

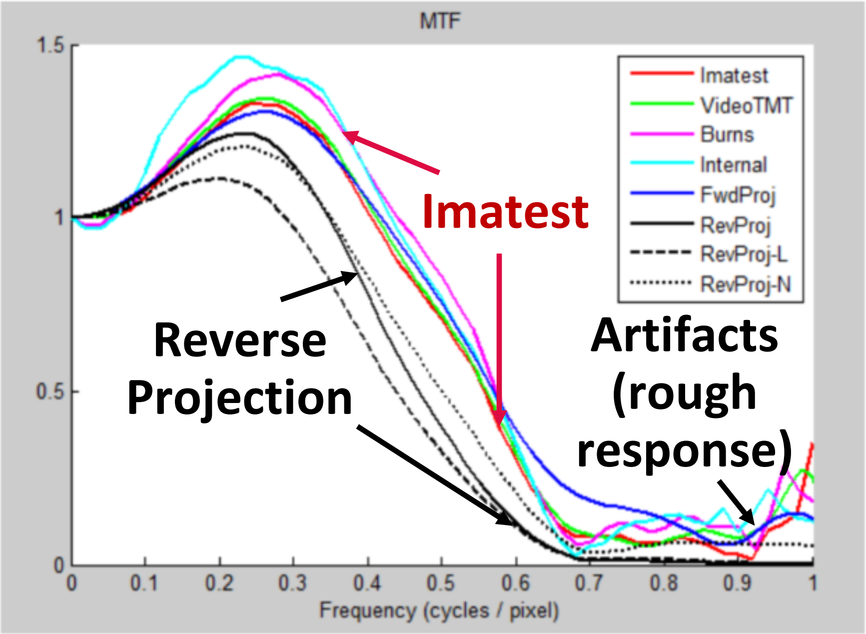


Figure 8. *MTF curves of the techniques of the 50 × 50 crop of the 5° edge.*

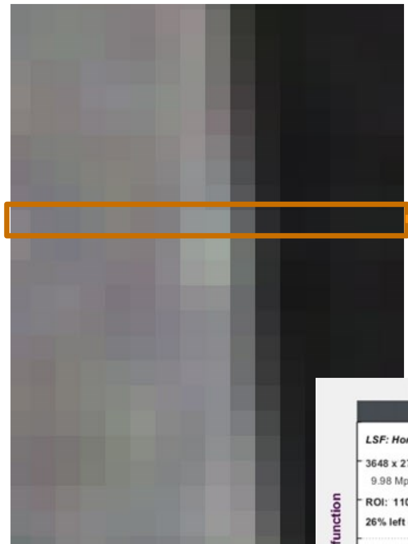
Stan Birchfield, "Reverse-Projection Method for Measuring Camera MTF," EI2017, has identified a problem and proposed a fix. Complex and covered by Microsoft patent.

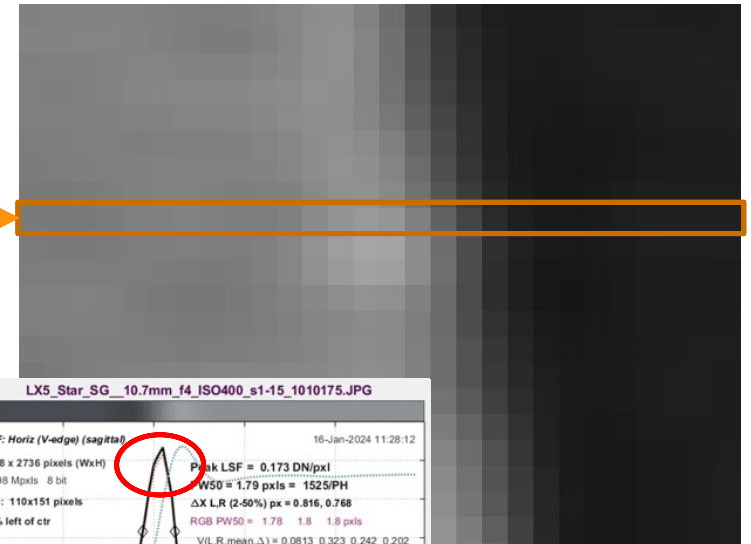
There are also relevant papers by **Kenichiro Masaoka**, **David Haefner**, and others.

Improved *MTF* calculation

needed because old MTF calculation had high frequency artifacts

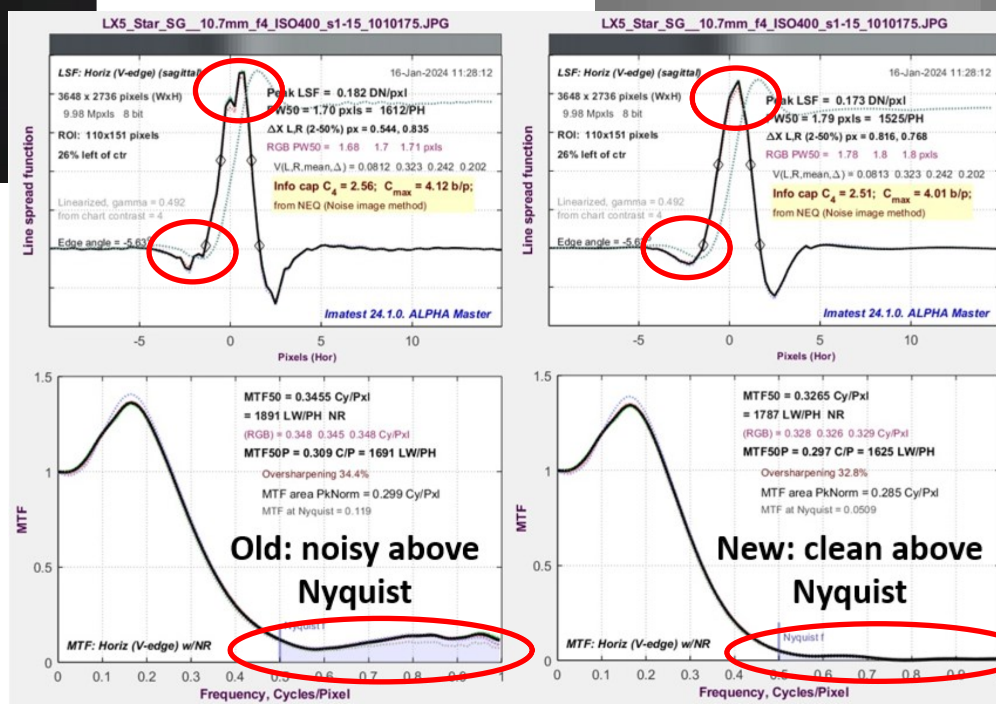
Before binning, interpolate each horizontal scan line to increase the number of pixels from *N* to 2*N*-1.





Use MATLAB interp1 or interp2, 'cubic'.

Old method: no interpolation



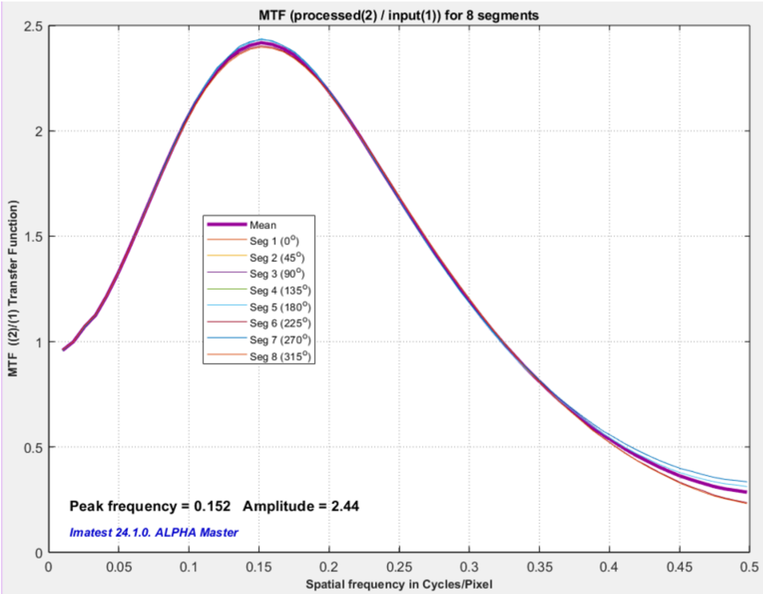
New method: interpolated

Smoother and better behaved at $f > f_{Nyq}$

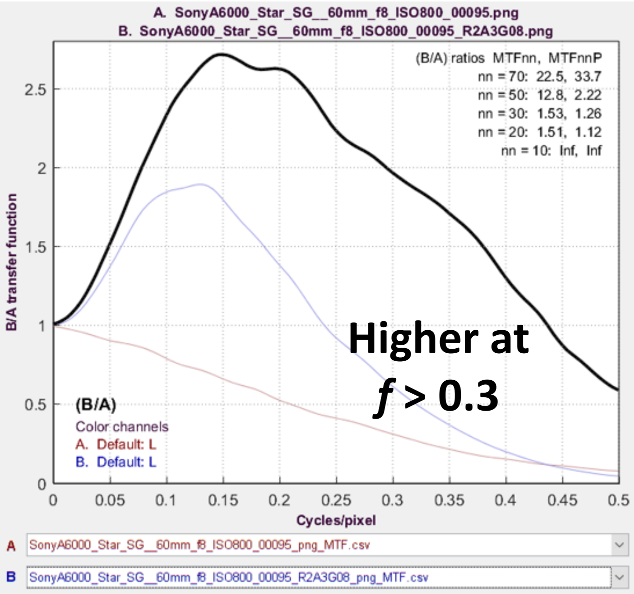
Checking the *MTF* calculations: USM + LPF

Compare filter design with *MTF*(filtered)/*MTF*(unfiltered) for Unsharp Mask sharpening R2A3 + Gaussian Lowpass Filter (LPF) with $\sigma = 0.8$.

Filter design transfer function



MTF transfer function filtered/unfiltered

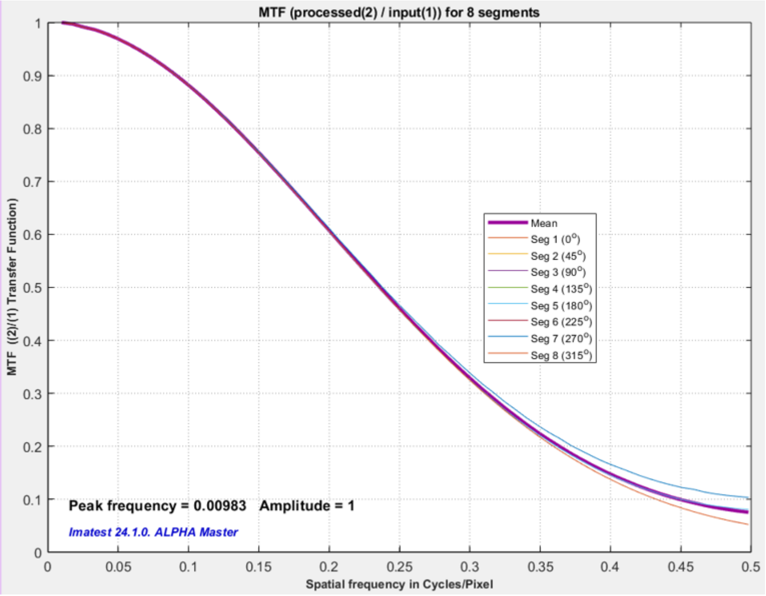


Discrepancy at *f* > 0.3 C/P. *MTF* measurements of sharpened images are good enough for simple metrics like *MTF50*, but not for *SNRi and Edge SNRi* calculations derived from $MTF(f)^2/NPS(f)$.

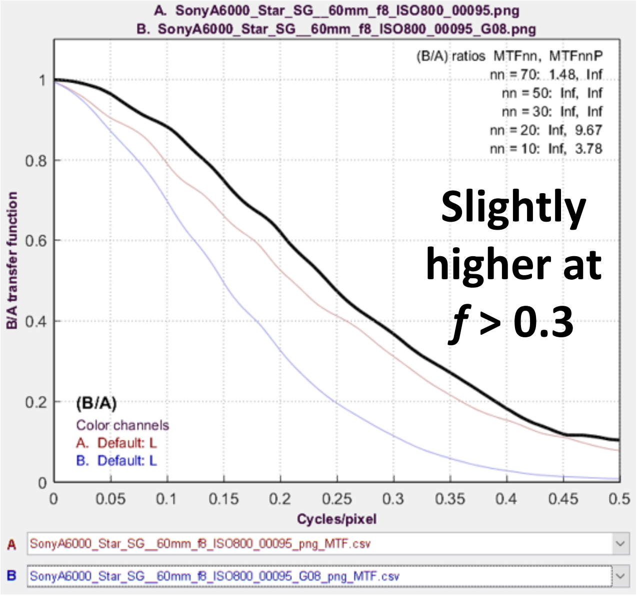
Checking the calculations: LPF

Compare filter design with *MTF*(filtered)/*MTF*(unfiltered) for Gaussian Lowpass Filter (LPF) with σ = 0.8.

Filter design transfer function



MTF transfer function filtered/unfiltered

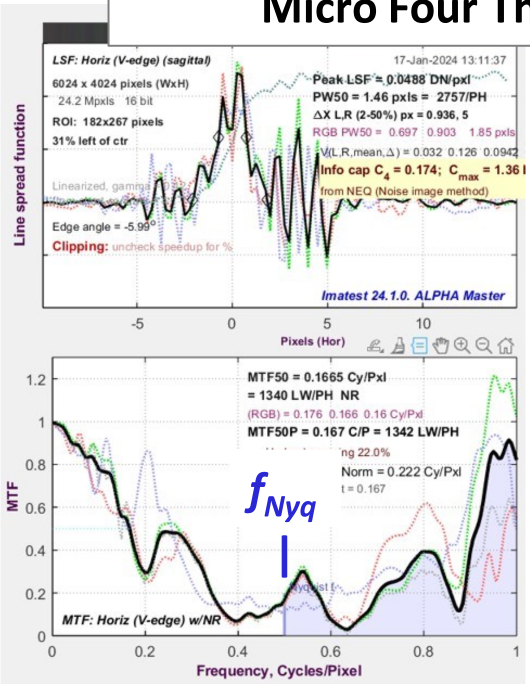


MTF **measurements of unsharpened images are good enough to design *Matched filters*, which optimize *object and edge detection*. We are working on improvements for sharpened images.**

More on the improved *MTF* calculation

BIG improvement with extremely noisy images

Micro Four Thirds camera ISO 12800



Old method

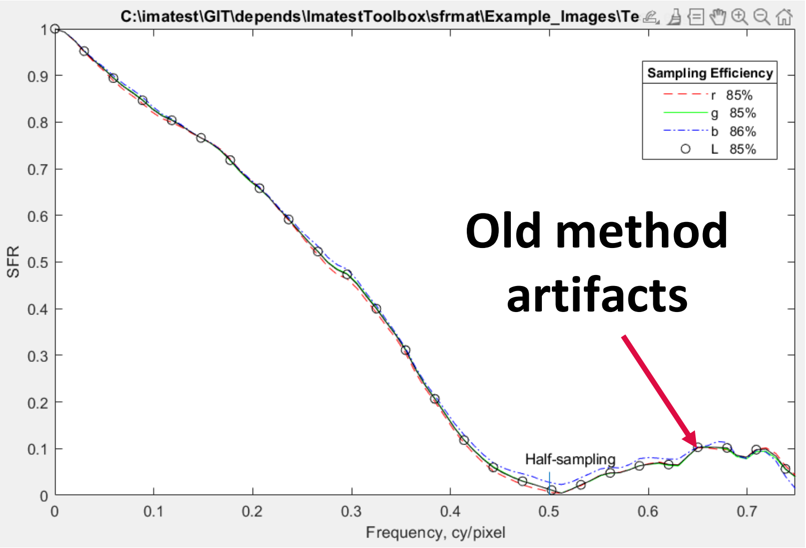
New method: interpolated

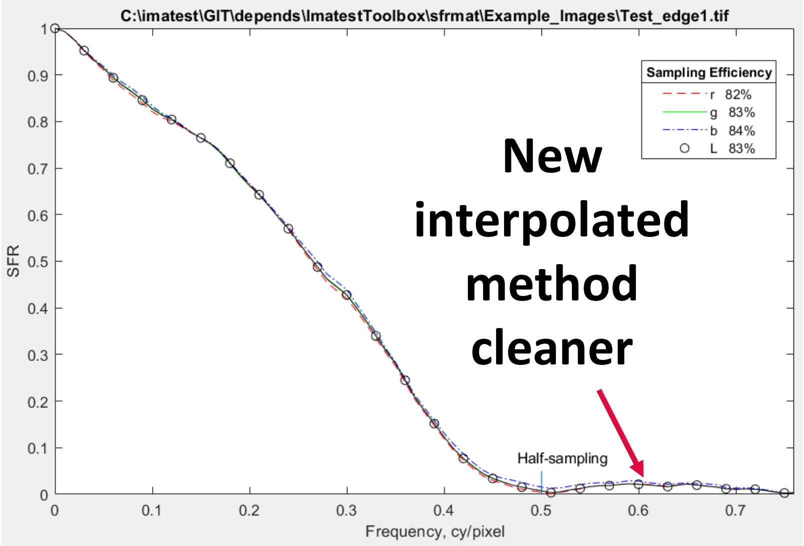
Differences are visible at $f > f_{Nyq}/2$ and strongest at $f > f_{Nyq}$.
For uniformly processed images, nearly identical to Siemens Star

We are communicating the new technique to the ISO TC42 committee for inclusion in a future ISO 12233 release.

Apply interpolation fix to SFRMAT5

- **SFRMAT5**, written by Peter Burns (burnsdigitalimaging.com) is the standard implementation of the ISO 12233 algorithm.
- It is used as the platform for implementing sample code, required by ISO for inclusion into standards.
- We will work with Peter to figure out how to distribute it. (It will eventually include sample code for the information metrics.)





Methods for measuring noise and hence information capacity C in the presence of a signal

- **Method 1**: the Edge Variance method for measuring spatially-dependent noise $N(x)$ by summing the squares of each scan line.
- **Method 2**: the Noise Image method for measuring the noise power or amplitude spectrum, $NPS(f)$ or $N_V(f)$.

Method 1: the Edge Variance method

for measuring noise near the slanted edge.

In addition to summing each scan line,

<u>sum the *squares* of each scan line</u>, $\rho_s(x) = \frac{1}{L}\sum_{l=0}^{L-1} y_l^2(x)$.

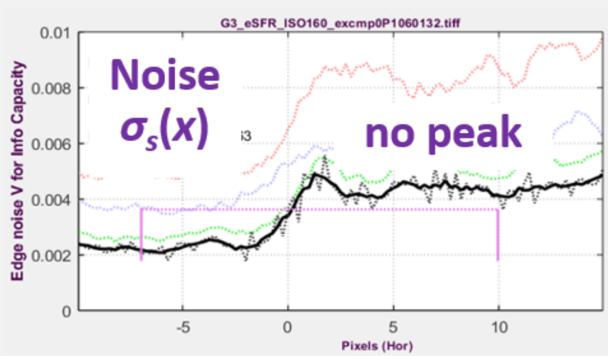
$$\sigma_s^2(\mathrm{x}) = N(x) = \frac{1}{L}\sum_{l=0}^{L-1} y_l^2(\mathrm{x}) - \left(\frac{1}{L}\sum_{l=0}^{L-1} y_l(\mathrm{x})\right)^2 = \rho_s(x) - \mu_s^2(x)$$

Variance $\sigma_s^2(\mathrm{x}) = \rho_s(x) - \mu_s^2(x)$ is the spatially dependent <u>noise power</u> $N(x)$.

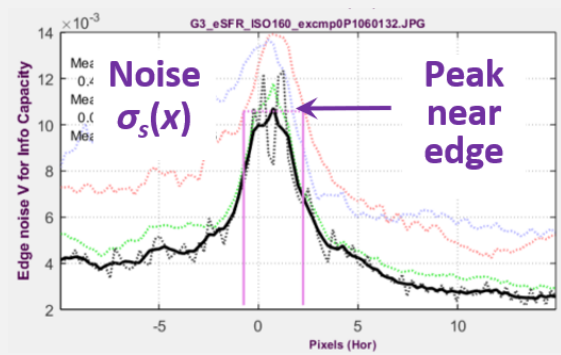
Examples of noise amplitude $\sigma_s(x) = N_V(x) = \sqrt{N(x)}$

for two different types of image processing

Uniformly-processed



Bilateral-filtered

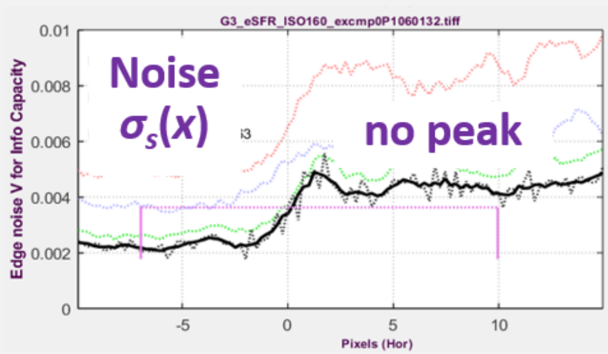


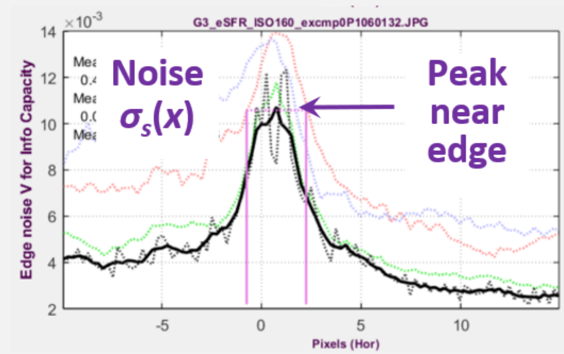
Best results when ROI length ≥ 100 pixels.

Noise power $N(x)$ for calculating information capacity, C

was not previously visible

Noise amplitude $\sigma_s(x) = \sqrt{N(x)}$





Uniformly or minimally processed images

Unsharpened or uniformly sharpened. No noise reduction

Little or no noise peak.

C is calculated from

$$N_{avg} = \text{mean}(N(x)).$$

More accurate than bilateral filtered

Bilateral-filtered images

Sharpened near the edge; noise-reduced elsewhere) JPEG images from most consumer cameras

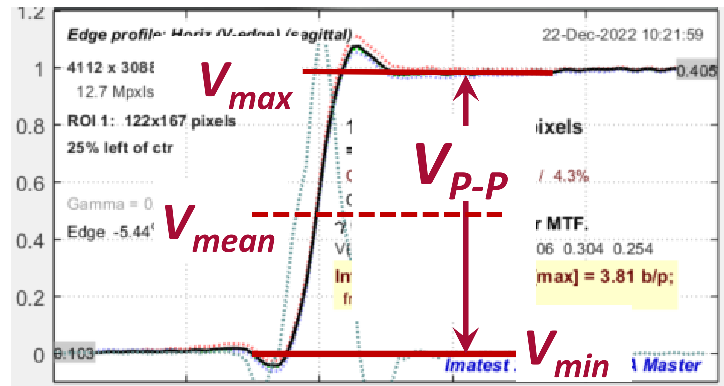
Distinct noise peak – ***identifies bilateral filtering***

C is calculated from the smoothed peak noise power, $N_{peak\text{-}smooth}$.

Less accurate than uniformly processed.

Signal power $S(f)$ for calculating C

In addition to noise, the average signal power $S_{avg}(f)$ is also needed to calculate C.



Signal power S, which is proportional to V^2 for signal amplitude V, is typically measured from charts with 4:1 contrast ratio.

Information capacity is maximum when V is ***uniformly distributed*** from V_{min} to V_{max} (a range of $V_{P\text{-}P}$). Signal frequency-dependence comes from MTF.

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

To calculate information capacity C, enter $S_{avg}(f)$, N, and bandwidth $W = f_{Nyq} = 0.5$ C/P into the Shannon-Hartley equation.

$$C = \int_0^W \log_2\left(1 + \frac{S_{avg}(f)}{N_{avg}}\right) df$$

imatest®

Information capacities C_n and C_{max}

Information capacity C measured from low-contrast chart images (to minimize saturation and nonlinear operation) is a strong function of exposure and chart contrast ratio n. For this reason the chart contrast ratio should be specified, i.e., C_n for an n:1 ratio.

C_4 is widely used for ISO-standard 4:1 charts.

Maximum information capacity, C_{max} : a stable metric for characterizing cameras

Derived from C_n , but insensitive to chart contrast and exposure.

- Extrapolate $V_{p\text{-}p}$ to $V_{max} = 1$ (smaller in some cameras),
- Adjust the signal-dependent noise power for the increased signal: straightforward for linear sensors; challenging for HDR.

Details in white papers linked from
www.imatest.com/solutions/image-information-metrics

imatest®

Consistency of information capacity

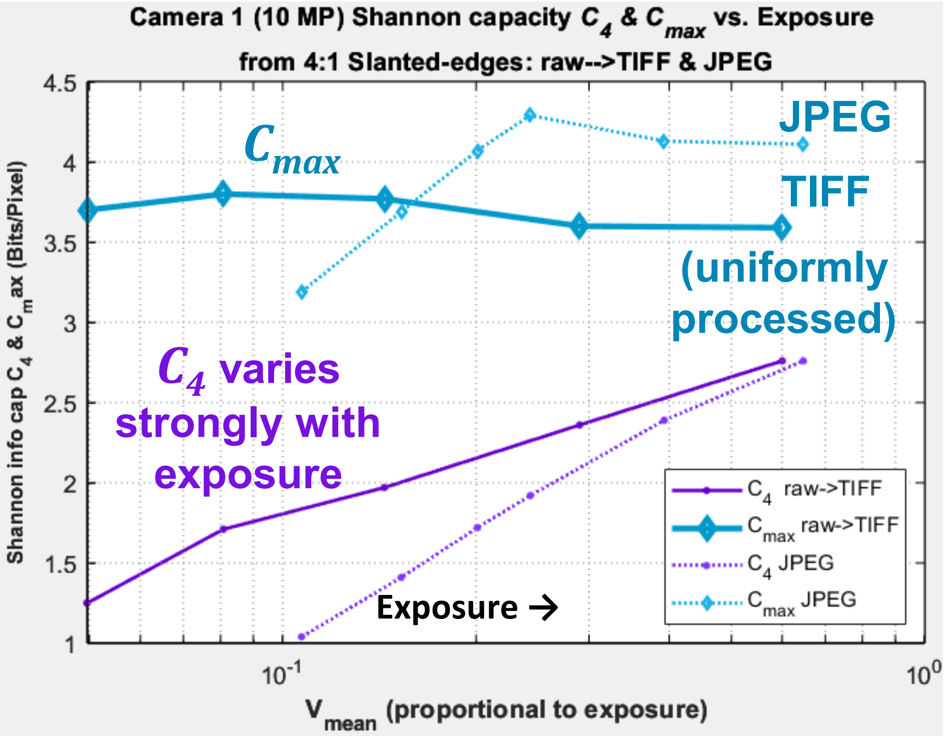
C_4 and C_{max} were measured as functions of exposure for consumer cameras.

- Minimally processed (TIFF) files are more consistent than JPEGs.
- C_4 varies as expected, increasing with exposure.
- C_{max} is nearly consistent.

C_{max} and C_4 as functions of exposure

for a 10.1 MP compact consumer camera

Solid lines: raw→TIFF;
dotted lines: JPEG



(Lowest exposure is extremely underexposed.)

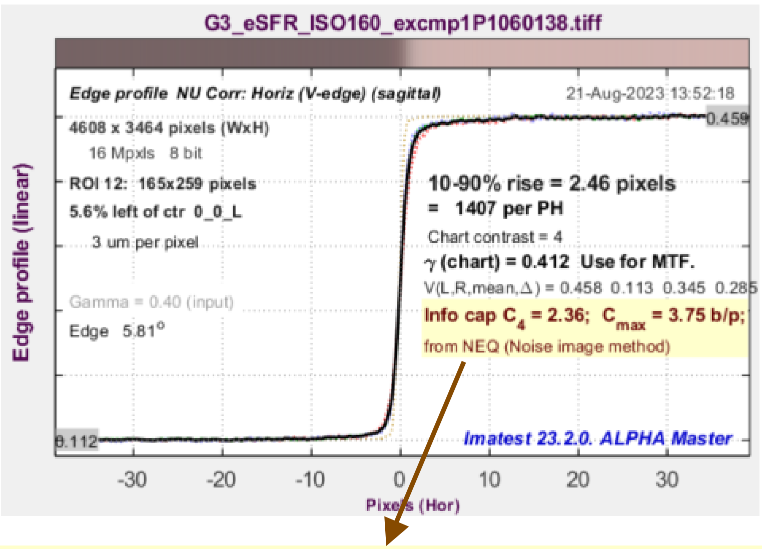
imatest®

Displays of information capacity C_4 and C_{max}

The 3D plot illustrates how C varies over the image. Mean(C_{max}) = 2.959 b/p.

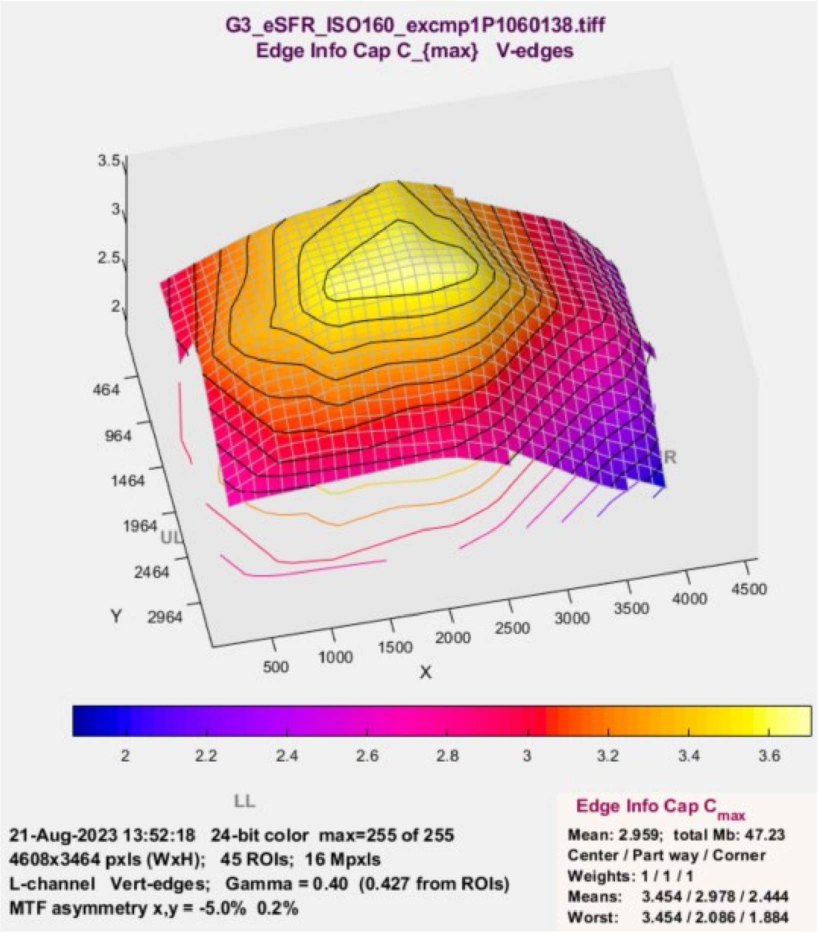
Total info capacity $C_{maxTotal}$ = mean(C_{max}) * number of pixels = 47.23 Mb

C_4 and C_{max} are displayed in the upper part of the Edge/MTF plot.



Information capacities C_4 = 2.36 b/p; C_{max} = 3.75 b/p.

from NEQ (Noise image method)

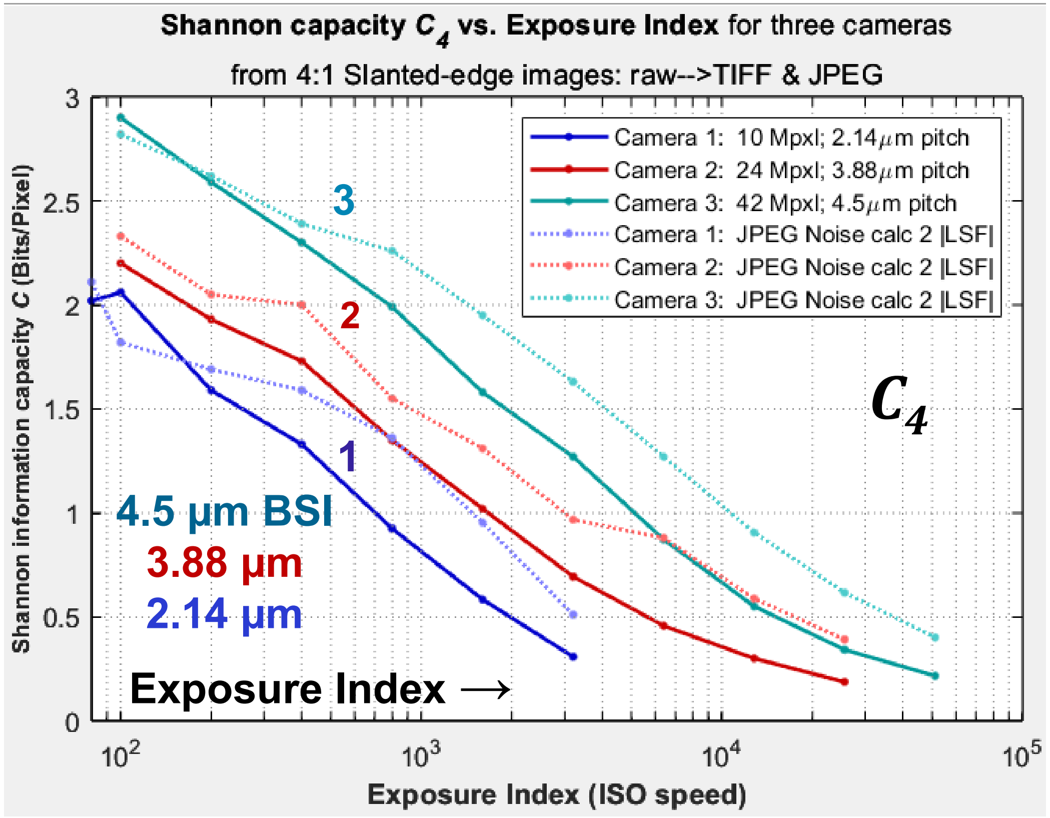


imatest®

C_4 results for three cameras

Sensors: 4.5 µm BSI, 3.88 µm, 2.14 µm

C_4 decreases with Exposure Index (ISO speed, i.e., analog gain) and increases with pixel size, as expected.

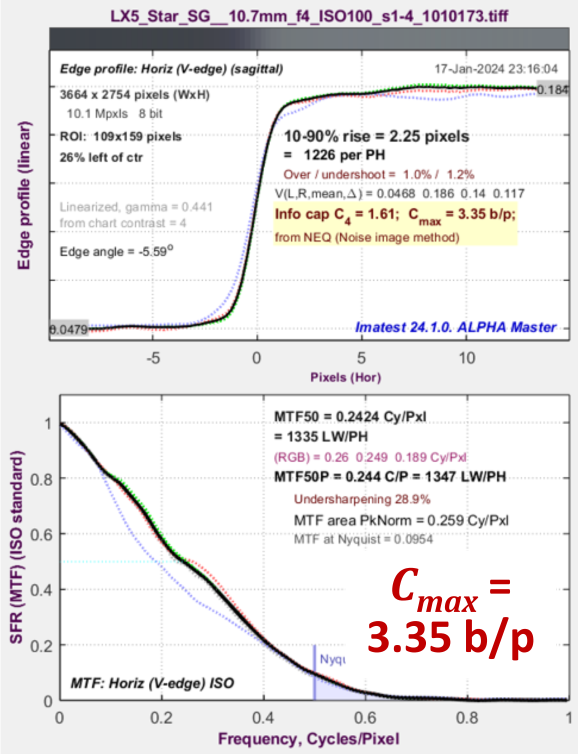


C_{max} tracks C_4, but is larger by about 2 bits/pixel.

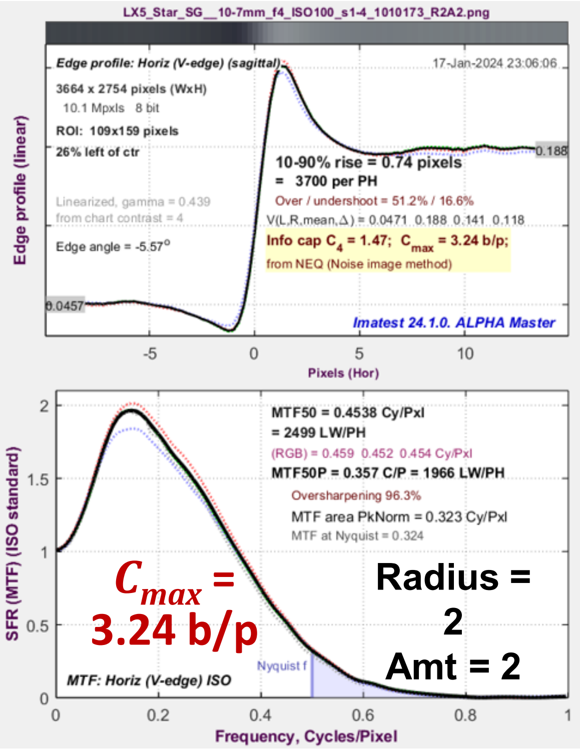
Sharpening and C_{max}

Sharpening has little effect on C_{max} because it boosts the frequency-dependent signal and noise by the same amount.

Minimally-processed TIFF



USM-sharpened TIFF



Post-processing cannot increase C.

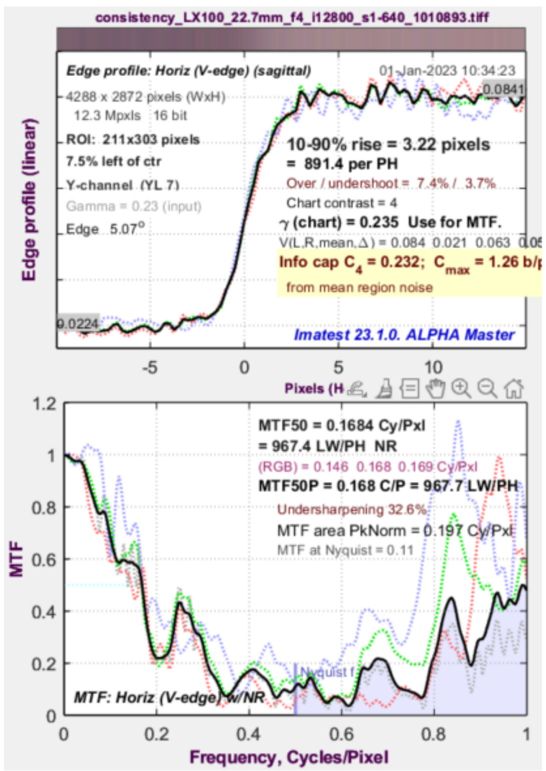
C is not useful for determining optimum image processing.

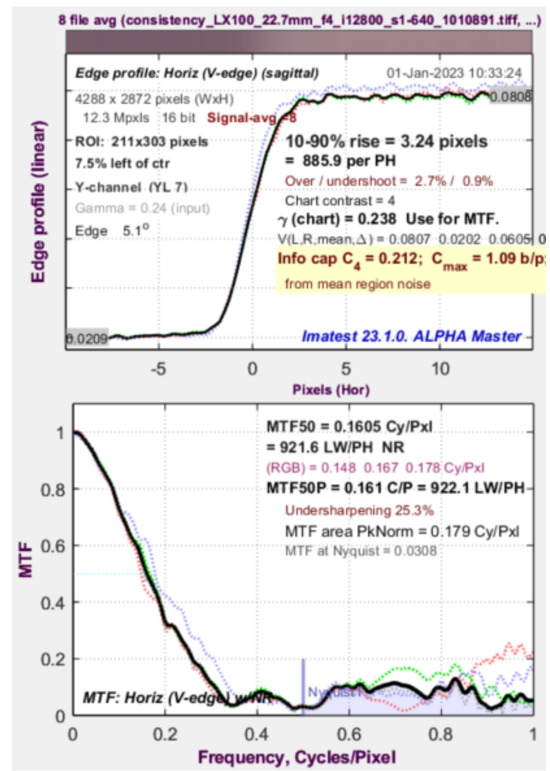
It doesn't indicate the effect of postprocessing on object or edge detection.

Signal averaging to improve quality and consistency of measurements

A classic technique that increases SNR by $\sqrt{n}$ whenever n identical images are averaged, e.g., by 3dB when n is doubled. To obtain correct information capacity measurements, noise power is multiplied by n.

Noisy image (ISO 12800; 1 inch sensor)

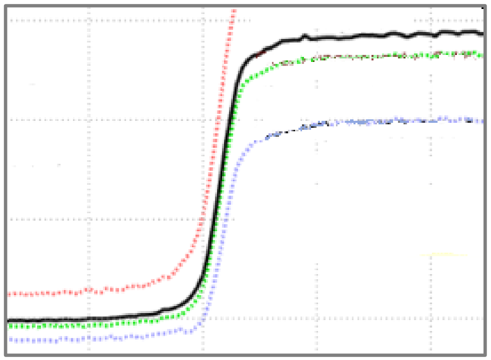




Single image Old *MTF* calculation method (no interpolation) ***n = 8 averaged***

imatest®

Method 2: The Noise image method enables several additional metrics

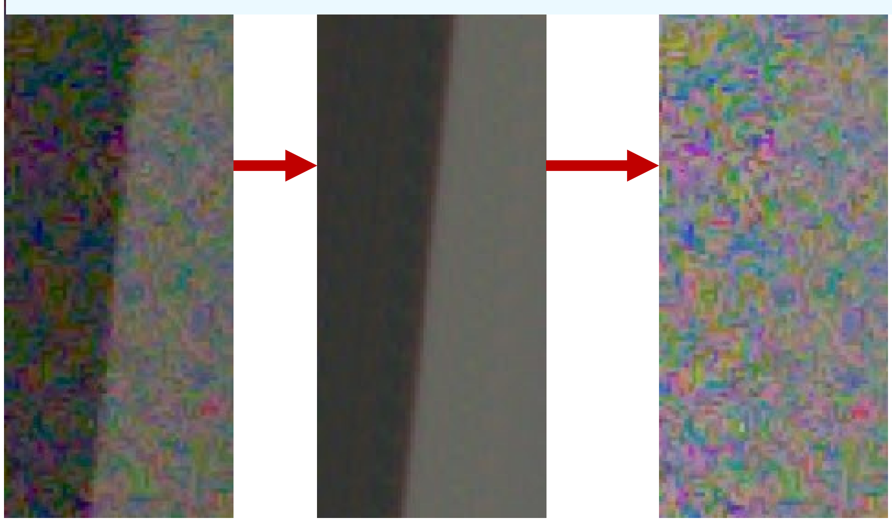


The averaged oversampled image consists of four averaged interleaves from the original bins of the ISO 12233 calculation.

De-bin the image by moving the low-noise contents of each interleave back to their locations in the original image.

(1) Original | **(2) De-binned** | **(3) Noise image = original – de-binned**

Micro 4/3 camera @ ISO 12800



The noise image (3) is the difference between the original (1) and de-binned images (2).

This method should *not* be used with bilateral-filtered images

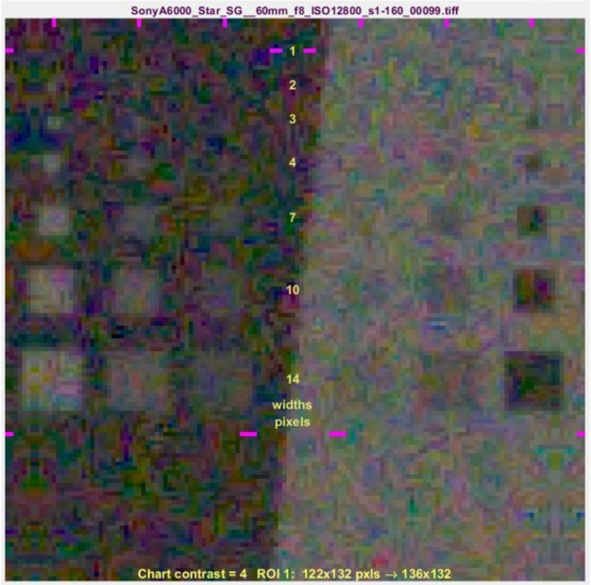
Measurements derived from the noise image

Affect camera selection and/or system performance

- ***Noise Equivalent Quanta (NEQ(f))*** — a frequency-dependent SNR, important in medical imaging.
- ***Information capacity, C_{NEQ}*, derived from NEQ** — Similar to values to ***C*** from the Edge Variance method (uses *NPS*).
- ***Ideal observer Signal-to-Noise Ratio (SNRi)*** — detectability of objects.
- ***Edge SNRi*** — detectability of edges.

Others (intermediate calculations, etc.)

- ***Noise power Spectrum (NPS(f))*** — intermediate calculation
- ***Object visibility*** — of small/low contrast objects, shown on the right. Related to ***SNRi***.
- ***Noise Autocorrelation*** ($\mathrm{IFFT}(\sqrt{NPS(f)})$) may indicate sensor crosstalk;
- ***Detective Quantum Efficiency (DQE)***



imatest®

Noise Power Spectrum $NPS(f)$

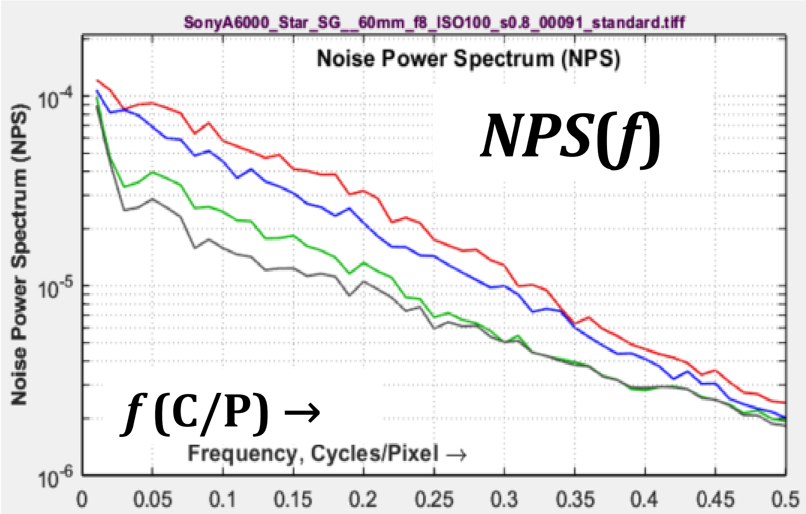
The 2D Fourier Transform (FFT) of the noise image must be transformed into 1D.

- Noting that $f = 0$ at the center of the 2D FFT image, divide it into several annular regions, and find the average noise power for each region.
- 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** $\int NPS(f)\, df = \int \sigma^2(x)\, dx = \int N(x)\, dx$

The noise amplitude (voltage) spectrum is

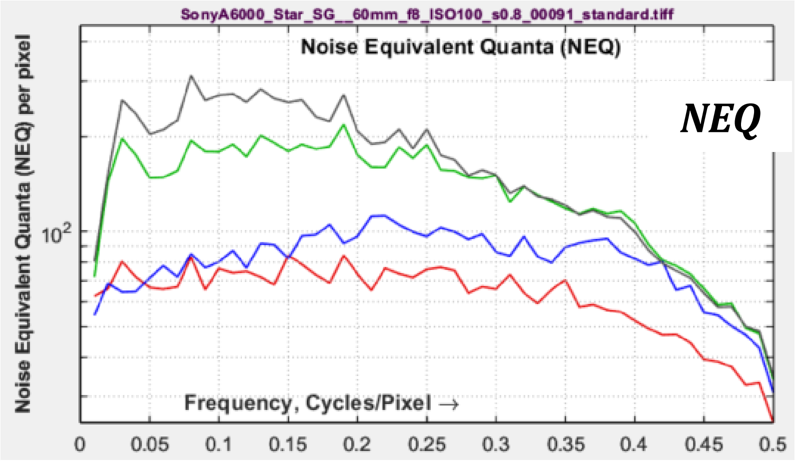
$$N_V(f) = \sqrt{NPS(f)}$$

$NPS(f)$ is used to calculate several key results.



Noise Equivalent Quanta $NEQ(f)$

***NEQ*(*f*)** Frequency-dependent Signal-to-Noise (power) Ratio, equivalent to the number of quanta that would generate the measured SNR when photon shot noise is dominant. Used in medical imaging.



$$NEQ(f) = \frac{V_{mean}^2 \, MTF^2(f)}{NPS(f)}$$

$\boldsymbol{K(f) = MTF^2(f)/NPS(f)}$ is the ***kernel*** of ***NEQ(f)*** and several information metrics to be introduced.

Because uniform filtering affects ***MTF²(f)*** and ***NPS(f)*** identically, <u>***NEQ*(*f*)** and ***K*(*f*)** are not affected by uniform filtering</u> such as sharpening or lowpass filtering.

C_{NEQ} — Alternate Information capacity derived from $NEQ(f)$

C_{NEQ} is calculated by altering the NEQ equation to represent a uniform amplitude distribution, replacing V_{mean} with $V_{P-P}/\sqrt{12}$.

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

C_{NEQ} can be thought of as a summary metric for $NEQ(f)$.

Results are similar to C from the Edge variance method; the two methods provide a good check on each other.

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

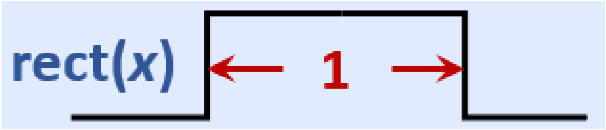
Detective Quantum Efficiency $DQE(f)$ is the ratio of $NEQ(f)$ (the number of quanta equivalent to the measured SNR) to the mean number of incident quanta. Its maximum value is 1. *Under development.*

$$DQE(f) = \frac{NEQ(f)}{\bar{q}}$$

Ideal Observer Signal-to-Noise Ratio *SNRi*

SNRi* is metric for the detectability of *objects, based on rectangular objects with sides *w* and *kw*. For

$\Delta g(x,y) = \Delta Q \cdot \mathrm{rect}(x/w) \cdot \mathrm{rect}(y/kw)$,



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

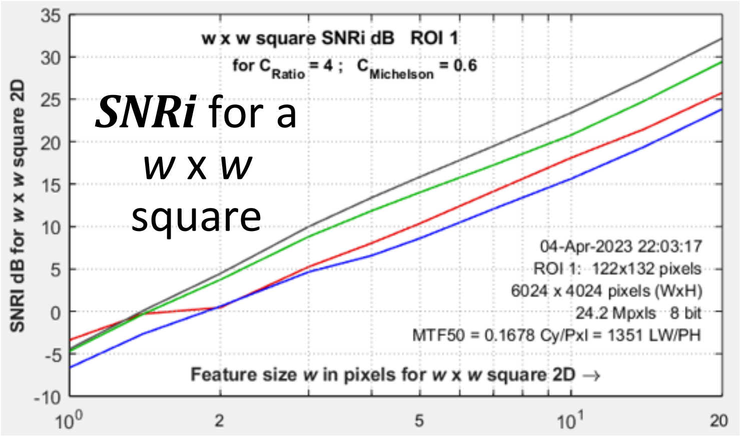
$$FFT\big(\Delta g(x,y)\big) = G(f_x, f_y) \; = kw^2 \Delta Q \frac{\sin(\pi w f_x)}{\pi w f_x} \frac{\sin(\pi k w f_y)}{\pi k w f_y}$$

$$SNRi^2 = \int_0^{f_{yNyq}} \int_0^{f_{xNyq}} \frac{|G(f_x,f_y)^2|\, MTF^2(f)}{NPS(f)} df_x \, df_y \;\; \text{where} \;\; f = \sqrt{f_x^2 + f_y^2}$$

In spatial domain, *SNRi*^{2} is the total energy of the object S/N: the basis of the *object visibility* display (next slide).

***SNRi* is proportional to the Michelson contrast of the chart ((lt-dk)/(lt+dk)).**

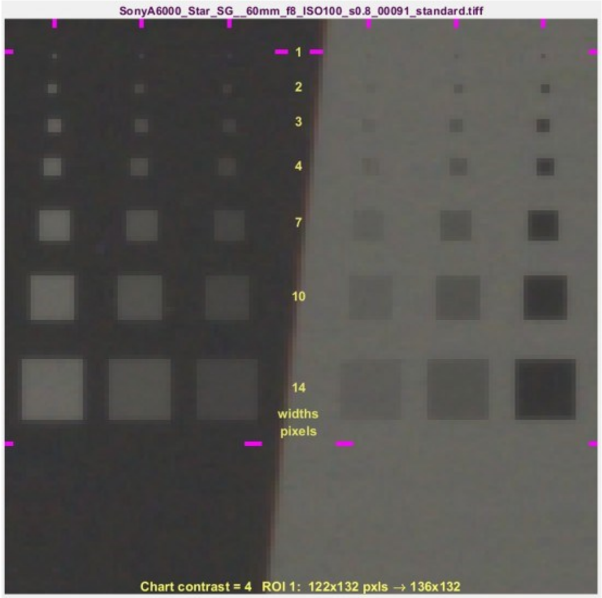
The *SNRi* plot can be difficult to interpret because it strongly increases with *w*.



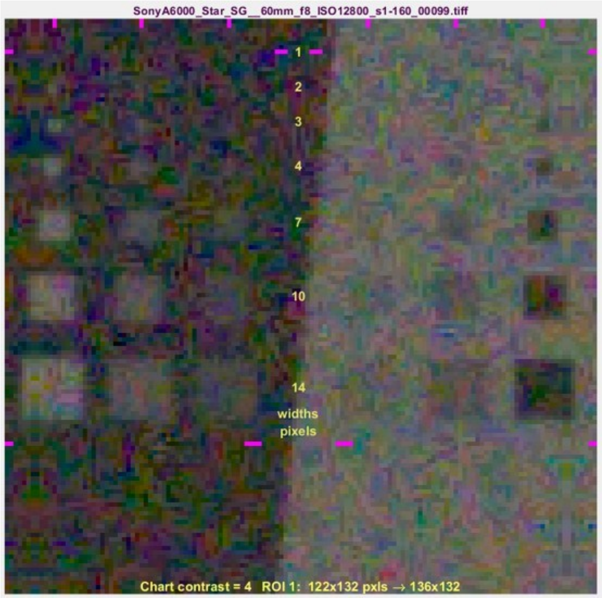
***w* in units of pixels →**

imatest®

Object visibility and *SNRi*



Low noise ISO 100

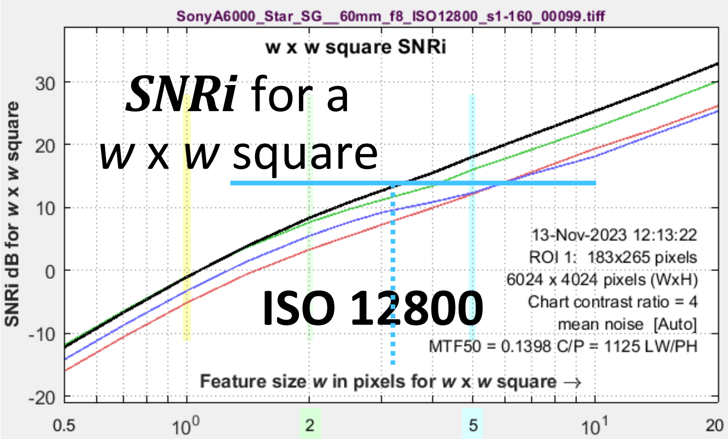


Noisy ISO 12800

Based on work by Rose*, a feature should be visible when *SNRi* ≥ 5 (14 dB).

Because objects sometimes have the same color as the background, we need to look at *edge detection*.

*See the white papers

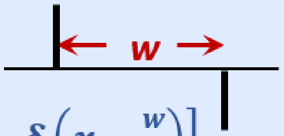


imatest®

EDGE SNRi

Edge SNRi* is new metric of the detectability of *edges*.** Equation is similar to SNRi, with the object replaced by edges, forming ***Line Spread Function doublets (pairs opposite-polarity δ-functions spaced by w).

Odd impulse pair



$$I_I(x/w) = \frac{1}{2}\left[\delta\left(x+\frac{w}{2}\right) - \delta\left(x-\frac{w}{2}\right)\right]$$

$$\Delta h(x,y) = V_{P-P} \cdot I_I(x/w) \cdot I_I(y/kw)$$

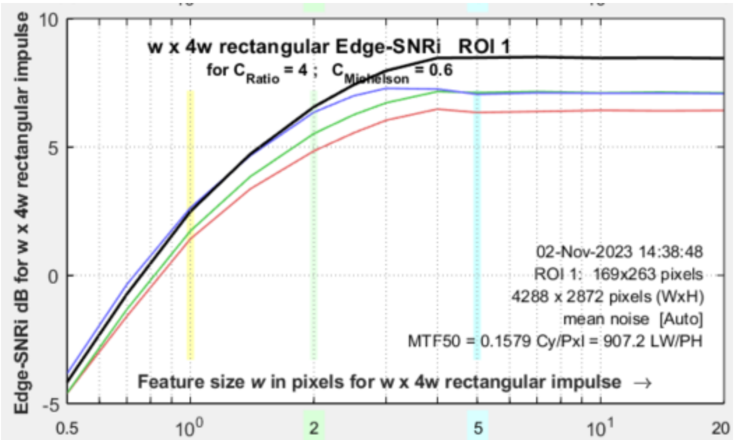
$$FFT\big(\Delta h(x,y)\big) = H(f_x, f_y) \ = 2\, V_{P-P} \sin(\pi w f_x) \sin(\pi k w f_y)$$

$$Edge\ SNRi^2 = \iint \left(\frac{\left|H(f_x, f_y)\right|^2 MTF^2(f_x, f_y)}{NPS(f_x, f_y)} \right) df_x\, df_y$$

***Edge SNRi* is our preferred metric for evaluating system performance.**

In spatial domain, *Edge SNRi²* is the energy of the LSF doublets.

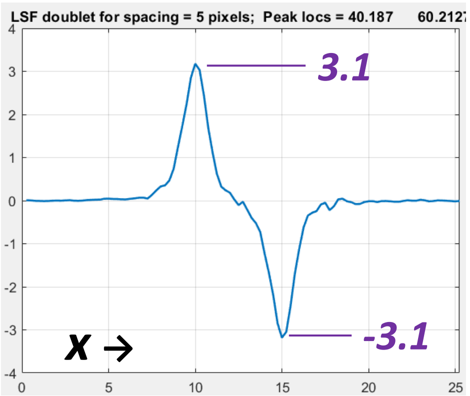
Can be improved with filtering (ISP).



imatest®

Line Spread Function (LSF) doublets

shown in spatial domain



LSF doublets

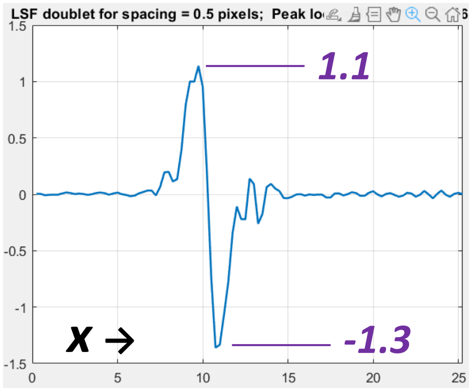
(object edges)

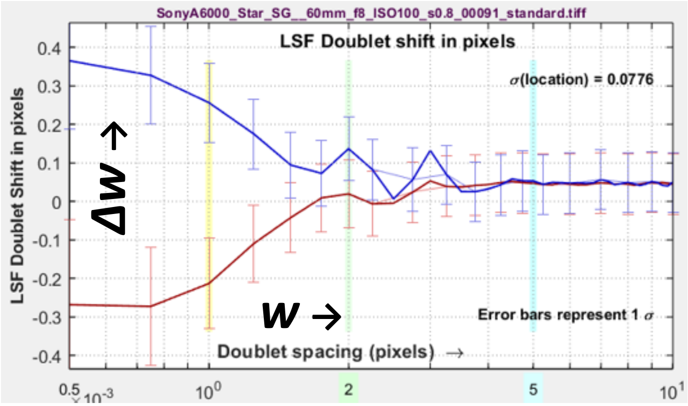
x and *w* in pixels

***w* = 5**

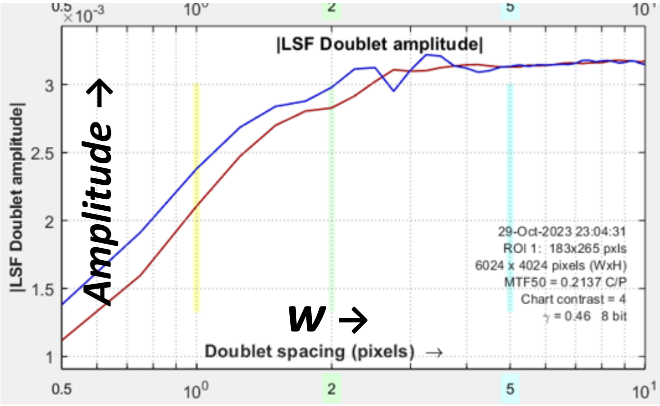
***w* = 0.5**

Amplitude is greatly reduced





Peak shift vs. spacing *w*



Amplitude vs. spacing *w*

Edge SNRi²* is the total energy of the doublets (object edges) making it a powerful metric for *edge detection.

$$Edge\ SNRi^2 \cong \iint LSF^2(x) LSF^2(y)\ dx\, dy$$

imatest®

Filtering 1

Filters are processors intended to improve a system's performance. A linear filter has a ***transfer function***, defined in frequency domain.

$$\mathscr{H}(f) = Output(f)/Input(f).$$

Typically used for sharpening (high frequency boost), noise reduction (LPF = low frequency cut), or a combination of the two.

Questions

1. Which of the image information metrics ($NEQ(f)$, C_{NEQ}, $SNRi$, $Edge\ SNRi$) are affected by filtering?
2. How to design an optimum filter the key detection metrics?
3. Will optimized filters improve MV/AI system performance? (We will need to work with researchers in academia/industry to answer this one.)

Filtering 2

To answer Q 1, define the equation ***kernel***, $\boldsymbol{K}(\boldsymbol{f}) = \boldsymbol{MTF}^2(\boldsymbol{f})/\boldsymbol{NPS}(\boldsymbol{f})$, then rewrite the key equations using ***K(f)*** (shown in ***boldface*** for emphasis).

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

$$C_{NEQ} = \int_0^W \log_2(1 + \mu^2\, \boldsymbol{K}(\boldsymbol{f}))\, df \quad \text{where } W = f_{Nyq} = 0.5\ C/P$$

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

where $G_{rect}(f)$ and $H_{doublet}(f)$ are the Fourier transforms of a rectangular object and edge (doublet).

Since filtering has the same effect on $MTF^2(f)$ and $NPS(f)$, $\boldsymbol{K}(\boldsymbol{f})$ should be unaffected by filtering, and therefore

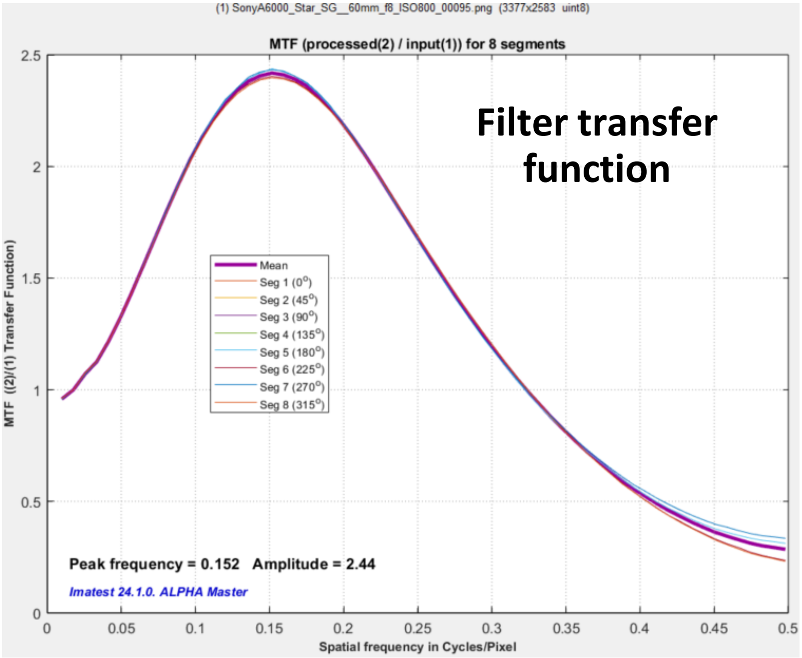
- ***NEQ(f)*** and $\boldsymbol{C_{NEQ}}$ are unaffected by filtering.
- ***SNRi*** and ***Edge SNRi*** can be strongly affected by filtering.

Filter simulation using *EDGE SNRi*

To study how filtering affects *SNRi* and *Edge SNRi*, we used the *Imatest* Image Processing module to filter raw images with combinations of

- **Gaussian* Lowpass filtering (LPF),**
- **Unsharp Masking (USM),**

We searched for filters that would enhance *SNRi* and *Edge SNRi* performance.



The filter shown on the right,
USM with R = 2 and A = 3, Gaussian LPF with σ = 0.8 was promising.

Sharpening reduces interference from neighboring objects, but increases noise. Including a well-tuned Lowpass Filter (as a part of the filter) to reduce noise is generally beneficial.

*Gaussian filters are approximations to more realistic filters, e.g., Bessel, Butterworth, etc.

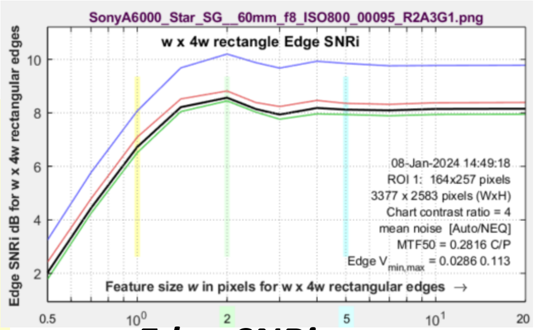
Filter simulation results

We calculated *SNRi* and *Edge SNRi* with several filters for $w \times 4w$ rectangles.

A6000 ISO 800 01/07/2023

Using *SNRi* and *Edge SNRi*, filters can be found that improve performance.

Lowpass Filtering (LPF) is always beneficial. Some sharpening is helpful.



Edge SNRi* vs. *w

Filter (Results are for a $w \times 4w$ rectangle.)	*MTF50* C/P	*Edge SNRi* $w = 1$	*Edge SNRi* large *w*	*SNRi* dB/pxl $w = 1$	*SNRi* dB/pxl $w = 5$	C_{max} (NEQ)	σ(loc.) pixels
None	**0.210**	**0.60**	**3.48**	**17.8**	**22.7**	2.79	0.15
USM R2A3	0.453	0.55	3.13	17.0	20.6	2.79	0.22
USM R2A3 + σ = 0.8	**0.321**	**6.21**	**8.04**	**19.9**	**23.2**	3.07	0.16
σ = 0.8 LPF only	0.149	2.81	5.96	**20.3**	**25.2**	2.72	0.13
USM R2A5 + σ = 0.8	**0.386**	**7.55**	**8.55**	**20.9**	22.8	3.35	0.18
USM R2A5 (extreme oversharpening)	0.527	2.23	4.55	18.1	20.6	3.04	0.25

σ indicates Gaussian Lowpass Filter (LPF)

Due to *MTF* calculation issues (which we're actively working on), these results may overstate the benefits of sharpening.

imatest®

Optimum filtering: the matched filter

- A "matched filter" is a custom filter that maximizes the SNR, i.e., detection probability, for
  - A specific object (or edge), and
  - A system with a specific response.
- Developed for impulse detection in radar. For an impulse (δ-function), matched filter frequency response is identical to the system response
- Discussed in ICRU Report 54 (an important but obscure document on medical imaging from 1996 that discusses *SNRi* and *NEQ*, and connects them to Bayesian statistics). [*Note that the Edge mf is new*.]

$$\text{Object matched filter} = |G_{rect}(f)|\ MTF(f)/N_V(f)$$

$$\text{Edge matched filter} = |H_{doublet}(f)|\ MTF(f)/N_V(f)$$

$$\text{where}\ \ MTF(f)/N_V(f) = \sqrt{K(f)}$$

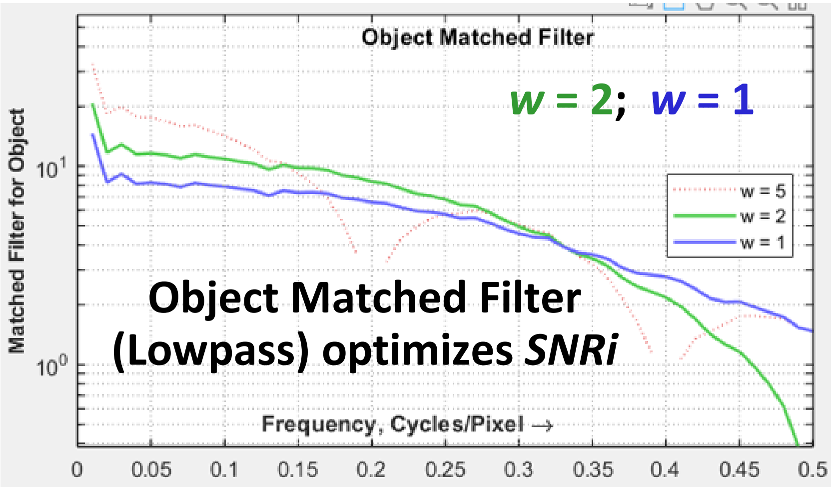
$$\text{Object or Edge}\ SNRi^2 = \iint \text{matched filter}^2\, df_x\, df_y$$

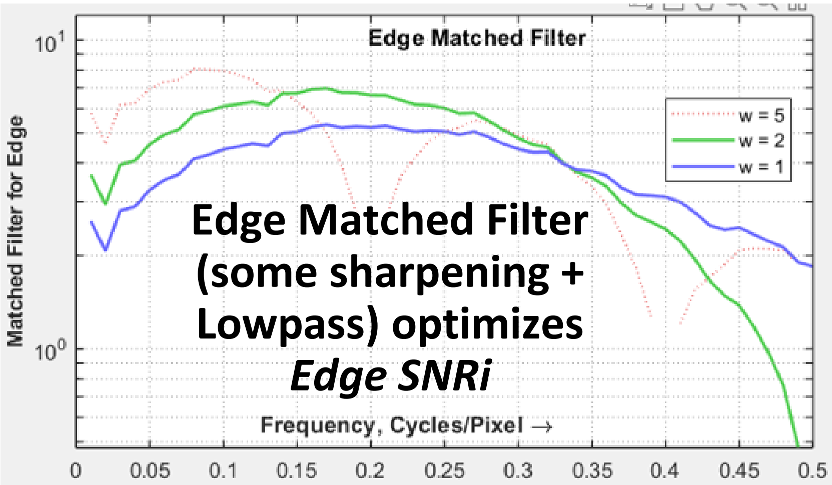
Matched filter example
calculated by *Imatest*

- Pure lowpass filter for (rectangular) optimum (object) *SNRi*.
- Sharpening (moderate) + lowpass filter for optimum *Edge SNRi*.

Best practice: Matched filters optimize a single metric: ***SNRi*** or ***Edge SNRi*** for a specific object width ***w***. In the real world, the filter must perform well for a variety of conditions, including interference from neighboring objects.

The big question: what to match to? Tradeoffs needed.





Proposed workflow

Determine the information capacity required for the application.

Select the camera with the *minimum* number of pixels that meets the requirement (along with other requirements, such as dynamic range and insensitivity to stray light).

Minimizing the pixel count should

- **Increase speed**
- **Reduce power consumption, and**
- **Reduce cost**

Find the optimum Image signal processing (ISP; typically a matched filter) that optimizes edge and/or object detection while controlling interference from neighboring objects.

imatest®

Summary – key concepts

We have discovered a mother lode of valuable metrics hidden in the slanted edge.

1. **Information capacity C is the fundamental predictor of potential MV/AI system performance — the best metric for selecting and qualifying cameras.** ***Traditional MTF and noise measurements are not sufficient.***
2. Spatially dependent noise $N(x)$ and frequency-dependent noise $NPS(f) = N_V(f)^2$ are calculated by separate methods from slanted edges, resulting in similar values of C, along with several additional metrics.
3. Because C_4, measured from 4:1 (low) contrast slanted edges, is highly sensitive to chart contrast and exposure, we developed a stable metric, C_{max}, for maximum camera information capacity.
4. Object and/or edge detection (i.e., object recognition) can be optimized with appropriate ISP (matched filter), using $SNRi$ and $Edge\ SNRi$.
5. Existing slanted-edge images can be used to obtain the new metrics. Old images do not need to be retaken.

imatest®

To do

- Partner with researchers in industry and academia to determine the correlation between image information metrics and the performance of MV/AI systems.
- Improve the slanted-edge algorithm for better results with sharpened images above 0.3 C/P.
- Document "best practices" for measuring image information metrics and designing an optimum matched filter (what to match?).
- Find a good method for characterizing information capacity in HDR sensors, where noise is not a simple function of signal.
- Better understand the numeric results for *SNRi* and *Edge SNRi*.
- Study how *C* relates to human perception. (ISP has a strong impact.)
- Work on the new ISO 23654 standard for image information metrics. We invite participation.

Above all, educate the imaging community on the benefits of information-related metrics.

Conclusion

Standard *MTF* and noise measurements are inadequate predictors of Machine Vision/AI system performance.

We have developed a powerful toolkit of measurements for predicting and optimizing MV/AI system performance.

Key metrics

Information Capacity, *C*	*SNRi*	*Edge SNRi*
Camera selection & qualification	Object detection	Edge detection
Independent of ISP	Optimize for object recognition with matched filter.	

Thank you.

More detail on the calculations can be found on
www.imatest.com/solutions/image-information-metrics

Visit our booths at Photonics West and Electronic Imaging 2024

imatest®