

# Camera Information Capacity: A Key Performance Indicator for Machine Vision and Artificial Intelligence Systems

Version 1.2, March 17, 2020

Norman Koren, Imatest LLC

[www.imatest.com](http://www.imatest.com)

We present a new image quality measurement — *information capacity* — that combines sharpness, noise, and the effects several types of artifact to form a fundamental figure of merit for image quality. It is of particular interest to Machine Vision and Artificial Intelligence, but generally applicable to a wide range of imaging.

**Abstract**— As Machine Vision (MV) and Artificial Intelligence (AI) systems are incorporated to an ever-increasing range of imaging applications— from medical image interpretation to controlling autonomous vehicles— it becomes vital to ensure that camera measurements provide the best possible prediction of system performance. At the present time it is standard practice to measure the two major factors that affect performance— sharpness and noise (or Signal-to-Noise Ratio)— as well as several additional factors, including tonal and color response, optical distortion, and sensor linearity, then to estimate system performance based on a combination of these factors. This estimate is usually based on experience, and is often more of an art than a science.

We propose a new measurement: **Camera information capacity**, based on Claude Shannon’s groundbreaking work on information theory, published in 1948 and 1949 [1],[2], which is the basis of modern electronic communications, but is still unfamiliar to imaging scientists. Shannon showed that every communication channel (which can be characterized by bandwidth and noise) has an *information capacity* that determines the maximum rate it can transmit data without error. A camera is such a communication channel, albeit with a difference: it transmits data to two-dimensional pixels instead of one-dimensional time. Since the algorithms behind Machine Vision operate on *information rather than pixels*, camera information content is critical to system performance.

Until now camera information capacity has been vaguely defined and difficult to measure. We define it to have units of bits per pixel or bits per image at a specified ISO speed, making it easy to compare very different cameras, and we describe a convenient method for measuring it using the **Imatest** program. This measurement can be used to solve some important problems, such as finding a camera that meets an information capacity specification with the minimum number of pixels— important because fewer pixels means faster processing.

## Information theory background

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

### What is information?

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

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

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

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

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

Claude Shannon was one of the genuine geniuses of the twentieth century— renowned among electronics engineers, but little known to the general public. The [medium.com](#) article, [10,000 Hours With Claude Shannon: How A Genius Thinks, Works, and Lives](#), is a great read. There are also nice articles in [The New Yorker](#) and [Scientific American](#). The 29-minute video [“Claude Shannon – Father of the Information Age”](#) is of particular interest to the author of this white paper because it was produced by the UCSD Center for Memory and Recording Research, which he visited frequently in his previous career.



Figure 1. Claude Shannon

### Channel capacity

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

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

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

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

This equation is not directly usable because bandwidth  $W$  is not well-defined, noise is not white, and it applies to one-dimensional systems (functions of time), whereas imaging systems have *two* dimensions. [Appendix I](#) shows how to convert the version of the equation for frequency-dependent (colored) noise into two dimensions (and back), resulting in Camera image capacity of

$$C = 2\pi \int_0^B \log_2 \left( \frac{S(f) + N(f)}{N(f)} \right) f df = 2\pi \int_0^B \log_2 \left( 1 + \frac{S(f)}{N(f)} \right) f df$$

Where  $C$  is the channel capacity in bits/pixel;  $B = \text{Nyquist frequency } f_{Nyq} = 0.5 \text{ cycles/pixel}$ ;  $S(f)$  is the signal power spectrum,  $N(f)$  is the noise power spectrum (which we interpret as broadband noise in the presence of  $S(f)$ ).

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 it's used downstream 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.

## The Siemens star measurement

### Selecting the test chart

For a test chart to accurately measure  $C$ , it must fulfill several requirements.

- It should have a wide tonal range, close to the maximum the camera can reproduce without saturating (clipping) well-exposed images.
- It should be easy to obtain, and, if possible, it should be an ISO standard chart.
- It should allow the measurement of  $S(f)$  and  $N(f)$  at the *same* location. This is important for two reasons.
  - because many cameras, particularly consumer cameras, use nonuniform and/or non-linear image processing to improve image appearance. The most common variety is the *bilateral filter*, which sharpens the image (boosts high spatial frequencies) near sharp features like edges, but reduces noise (i.e., smooths or lowpass filters the image) away from sharp features. This can increase the measured value of  $C$  while actually *removing* information from the image.

- because several types of artifact are absent and have little or no impact on  $C$  when  $S(f)$  and  $N(f)$  are measured at separate locations.

After considering several chart types we decided on the sinusoidal Siemens star, which is a part of the [ISO-12233-2014/2017 standard](#). The beauty of the Siemens star is that it can be divided into several radial and angular segments, where the frequency in each segment of average radius  $r$  is  $f = n_{cycles}/2\pi r$ . This allows  $S(f)$  and  $N(f)$  to be conveniently calculated by the method of [Appendix II](#).

The contrast of the star chart should be as close as possible to 50:1 (the minimum specified in the standard; close to the maximum achievable with matte media). Higher contrast can make the star image difficult to linearize in some black-box testing cases. Lower contrast is acceptable, but should be reported with the results. The chart should have 144 cycles for high resolution systems, but 72 or fewer cycles is sufficient for low resolution systems. The center marker (quadrant pattern), used to center the image for analysis, should be 1/20 of the star diameter.

### Photograph the chart

Full instructions on photographing the chart and analyzing it in Imatest are found on [Star Chart](#). The background to the measurement is on [Shannon Information Capacity](#).

Acquire a well-exposed image of the Siemens star in even, glare-free light. Exposures should be reasonably consistent when multiple cameras are tested. The mean pixel level of the linearized image inside the star should be in the range of 0.16 to 0.36 for normalized image data in the range [0,1], where 0 is the response to no light and 1 is the level where pixels saturate.

Consistent exposure (by which we mean consistent signal levels inside the Siemens star image) is critically important when comparing cameras. The levels are shown below **Imatest** results plots and in the JSON and CSV output. An example from Figure 3 is **Mean levels: linear = 0.181, pixel = 0.431**, showing the linearized and original gamma-encoded pixel levels inside the Siemens star, normalized to 1.

The reason consistency is so important is that signal  $S$  is proportional to the exposure, but noise  $N$  consists of a fixed component (from electronic noise and other sources) and a variable component (from shot, i.e., photon noise), which varies as the square root of  $S$ . Hence SNR and information capacity  $C$  are sensitive to exposure, as shown in Figure 8.

The center of the star should be located as close as possible to the center of the image to minimize measurement errors caused by optical distortion (if present).

The size of the star in the image should be set so the maximum spatial frequency, corresponding to the minimum radius  $r_{min}$ , is larger than the Nyquist frequency  $f_{Nyq}$ , and, if possible, no larger than  $1.3f_{Nyq}$ , so sufficient lower frequencies are available for the channel capacity calculation. This means that a 144-cycle star with a 1/20 inner marker should have a diameter of 1400-1750 pixels and a 72-cycle star should have a diameter of 700-875 pixels. For high-quality inkjet printers, the physical diameter of the star should be at least 9 (preferably 12) inches (23 to 30 cm).



Figure 2. Typical image of Siemens star, framed so the maximum frequency is slightly above  $f_{Nyq}$ .

Other features may surround the chart, but the average background should be close to neutral gray (18% reflectance) to ensure a good exposure (it is OK to apply exposure compensation if needed). Figure 2 shows a typical star image in a 24-megapixel (4000×6000 pixel) camera.

## Information capacity results

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

Table 1. Cameras used in the tests

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

We captured both JPEG and raw images, which were converted to 24-bit sRGB TIFF images (designated as raw/TIFF) with no sharpening or noise reduction and gamma  $\cong 2.2$ . Results with 48-bit Adobe sRGB conversion were nearly identical.

Figure 3 illustrates results for the 24-megapixel camera 2 raw/TIFF image at ISO 100. Signal voltage ( $\sqrt{S(f)}$ ), noise voltage ( $\sqrt{N(f)}$ ), and  $(S(f) + N(f))/N(f)$  in dB are all calculated from the sinusoidal star pattern. Note that  $MTF(f) = \sqrt{S(f)/S(0)}$ . Noise  $N(f)$  is more accurate than for flat patches

(found in standard test charts such as ISO 14524, ISO 15739), especially in the presence of widely-used nonlinear (bilateral filtering), which may aggressively reduce noise more in smooth areas.

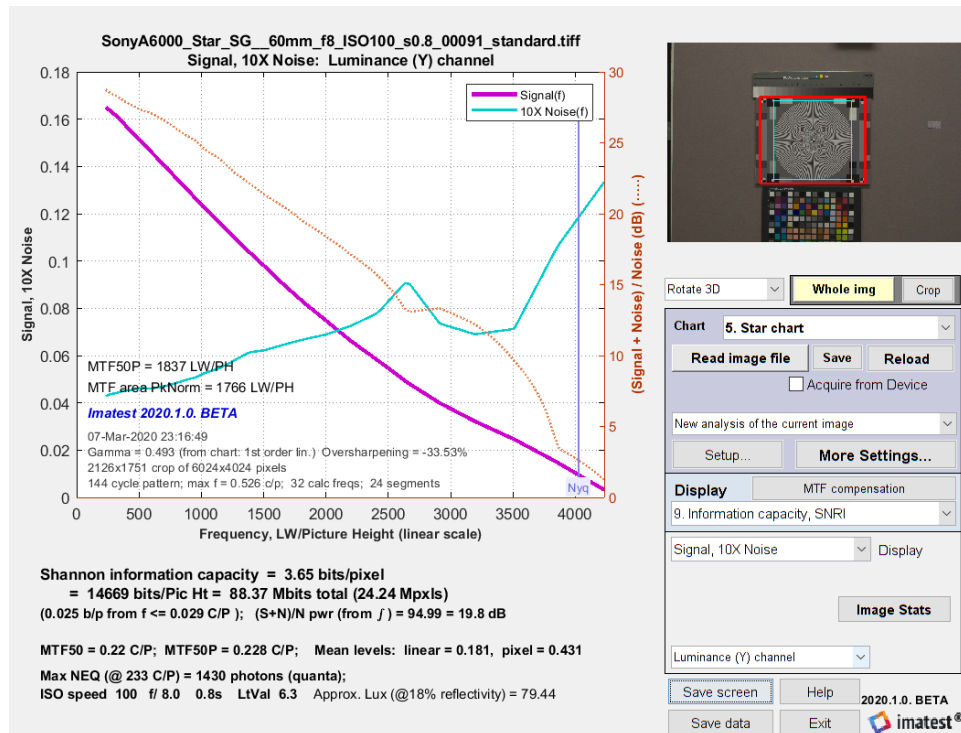


Figure 3. Signal voltage, 10x noise voltage, and (S+N)/N plot for 24 megapixel Camera 2, TIFF from raw, ISO 100, 32 frequencies, 32 radial segments. Summary results are shown below the plot.

The key result of Figure 3 is a single number: Shannon information capacity = 3.65 bits per pixel at ISO 100 (an excellent value; typical of high-quality cameras at low ISO speeds).

This number can be used to compare different cameras operating under different conditions (ISO speed, light level, aperture, etc.) and to quantify a camera's performance for Artificial Intelligence systems.

### JPEG vs. minimally-processed RAW images

Figure 4 shows results from the raw/TIFF images (solid lines) and JPEG images (dotted lines) as a function of ISO speed. The frequency dependence of the raw/TIFF results are consistent for the three cameras. The improvement between cameras 2 and 3 is much greater than expected from the ratio between pixel areas because camera 3 has a Backside-Illuminated (BSI) sensor (a greatly improved technology).

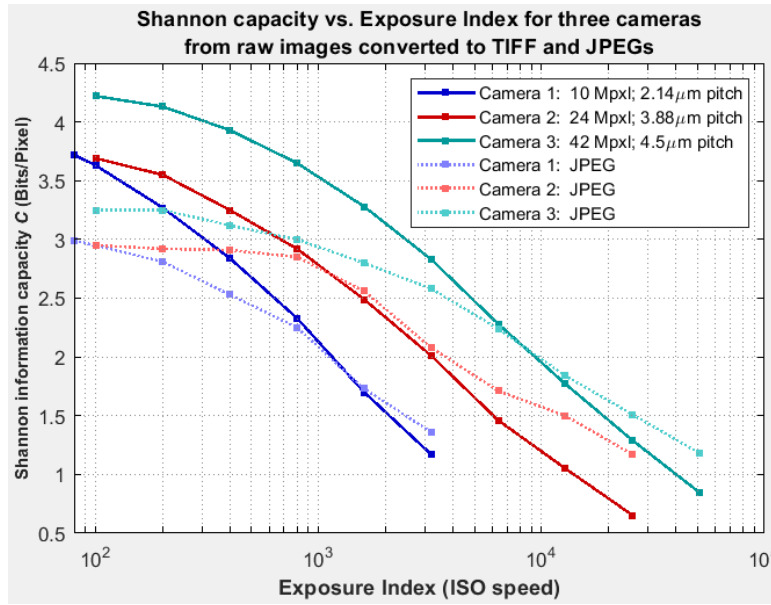


Figure 4. Information capacity for the three cameras as a function of Exposure Index: solid lines for TIFFs derived from raw images; dotted lines for JPEGs.

The striking result in Figure 4 is that sharpened JPEG images generally have *lower* information capacities than unsharpened raw/TIFF images (except at the highest ISO speeds, where JPEG image processing is different: reduced sharpening and increased noise reduction distorts  $C$ ).

To understand these results we need to look at the individual results. 10-megapixel Camera 1 at ISO 80 is a good example because it has the strongest sharpening of the three cameras tested.

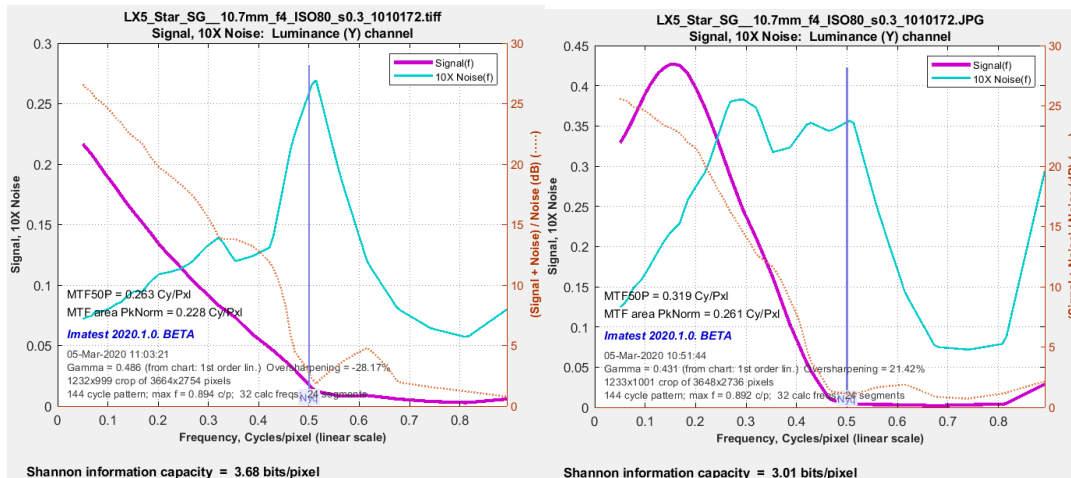


Figure 5. Shannon capacity plot, showing Signal voltage (magenta), 10x noise voltage (cyan), and  $(S+N)/N$  (brown) for 10-megapixel Camera 1 @ ISO 80. Raw/TIFF (left), JPEG (right). Note that  $MTF(f) = S(f)/S(0)$  (curve)

Figure 5 illustrates how the sharpening boost (the cause of the JPEG signal peak around 0.15 C/P) also boosts the noise, degrading  $(S(f) + N(f))/N(f)$ , leading to reduced  $C$ . The strong spatial-domain sharpening of the JPEG is visible on the slanted edge located just to the left of the star.

To further illustrate the effects of sharpening, which increases noise as well as high frequency amplitude, we ran several simulations on the 24-megapixel raw/TIFF camera 2 image (Figure 3). The baseline image has no added sharpening or noise reduction. For the others, USM is Unsharp Mask;  $R_n$  = sharpening radius;  $A_n$  = sharpening amount. Gaussian  $g$  is Gaussian (blur) filter with parameter  $g$ .

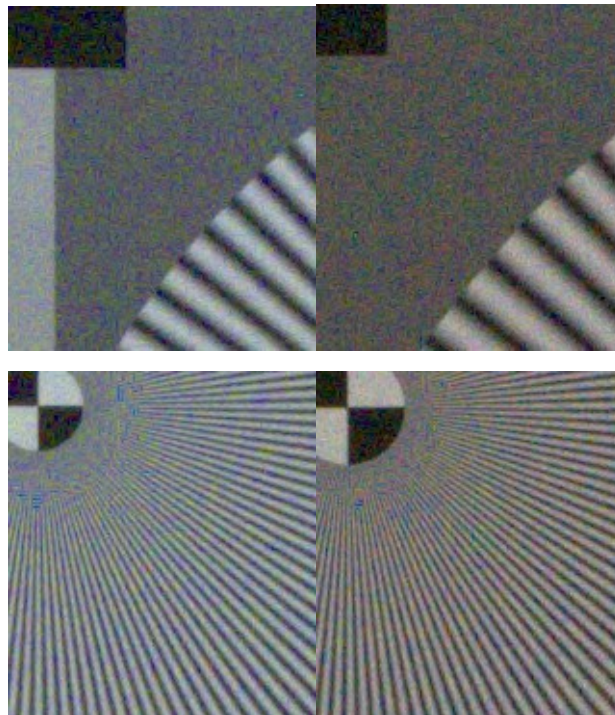
**Table 2. Capacity losses for sharpening & noise reduction**

Image	$C$	MTF50 (c/p)	MTF50P(c/p)
Baseline	3.69	0.22	0.229
USM R2 A1	3.65	0.345	0.323
USM R1 A2	3.63	0.407	0.397
Gaussian 0.7	2.99	0.162	0.168
Gaussian 1.0	2.25	0.138	0.143
USM R2A1, Gaussian 0.7	3.06	0.241	0.239

Since sharpening boosts noise along with signal, it does not increase information capacity  $C$ . In general, the measured value of  $C$  cannot be increased by any uniform image processing.

### Visual comparison of images with similar information capacity

Using the results in Figure 4, we selected images from two different cameras with similar information capacities (1.7 bits/pixel): camera 1 (2.14  $\mu\text{m}$  pixel pitch) at ISO 1600 and camera 3 (4.5  $\mu\text{m}$  BSI pixel pitch) at ISO 12800.



Camera 1, ISO 1600

Camera 3, ISO 12800

**Figure 6.** Comparison of raw/TIFF images from two different cameras with similar information capacity ( $\cong 1.7$ ), significantly degraded from ISO 100.



Noise appears very similar in the upper images, and sharpness and artifacts are similar in the lower images.

## Artifacts

A unique property of the Siemens star information capacity measurement method is that it measures the image quality loss from several types of artifacts that can arise from image capture or image processing.

### Clipping (saturation) artifacts

Clipping (or saturation) takes place when image pixels reach their limit, then get abruptly chopped off. It can occur in strongly overexposed images (Figure 7, on the right) or in images with excessive sharpening. Information is absent in clipped regions, but the sharp corner contains high frequency energy that exaggerates slanted-edge MTF measurements. Clipping is illustrated in Figure II-3 for an overexposed star.

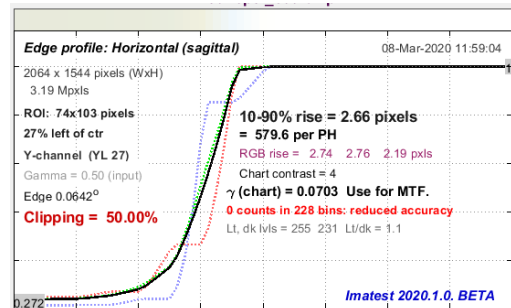


Figure 7. Clipping on a slanted-edge

We studied the effects of clipping using several images from 10-megapixel camera 1, which had the greatest software sharpening of any of the cameras tested. We used six exposure settings, from nominal-3 f-stops (EV) to nominal+2, saving each as JPEG and raw files. Shannon information capacity as a function of exposure is shown in Figure 8.

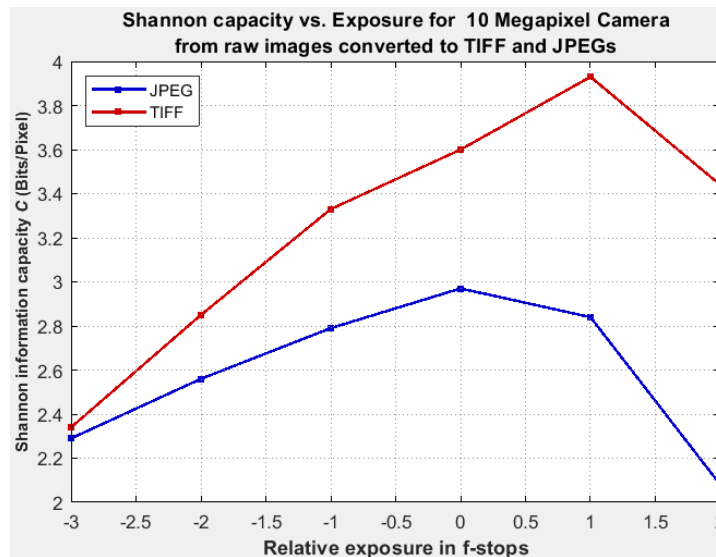
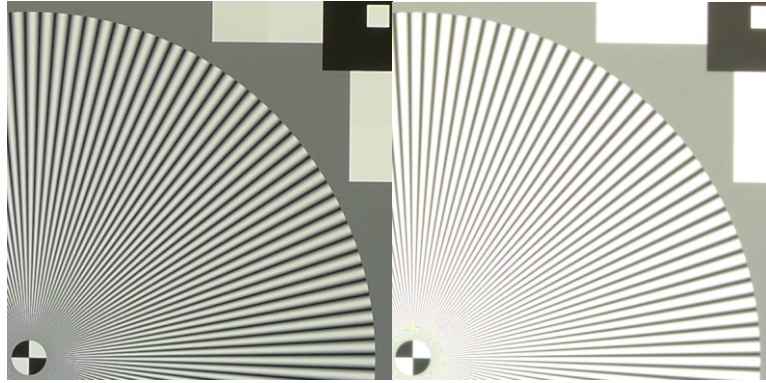


Figure 8. Shannon information capacity vs. Exposure error in f-stops.

Calculated information capacity increases with exposure up the point where clipping starts. The reason is that Signal-to-Noise ratio increases with exposure: signal increases linearly, but electronic noise is constant and shot (photon) noise increases with the square root of exposure. This is why consistent exposure is important when comparing different cameras.

The JPEG and raw/TIFF curves have a different shape because the JPEG images have reduced sharpening at the low exposures, while the minimal image processing for the raw/TIFF images is unchanged.



**Figure 9.** Portions of two JPEG images used in Figure 8.  
 Left: Rel. exposure = 0 (nominal); Right: Relative exposure = +2 (overexposed & clipped).

Clipped (overexposed) images have reduced information capacity measurements because clipped regions contain no information. Because they have zero noise, they are removed from the calculation using the method described in [Appendix II](#). The reduction in  $C$  is somewhat less than expected, most likely because the equation for  $C$  assumes a linear system, and is not accurate for highly nonlinear systems.

### **Demosaicing artifacts (mostly aliasing)**

Aliasing— low frequency artifacts such as moiré fringing that can appear when the signal reaching the image sensor has significant energy above the Nyquist frequency,  $f_{Nyq}$  (0.5 Cycles/Pixel)— can degrade image quality. Images from Bayer sensors can have significant color aliasing because the Nyquist frequency of the red and blue channels is half that of the sensor as a whole.

The amount of aliasing is strongly dependent on the demosaicing algorithm used by the raw converter. Simple algorithms such as bilinear demosaicing have severe aliasing. Most modern cameras use sophisticated algorithms that take advantage of detail from all color channels to construct the image in each individual channel.

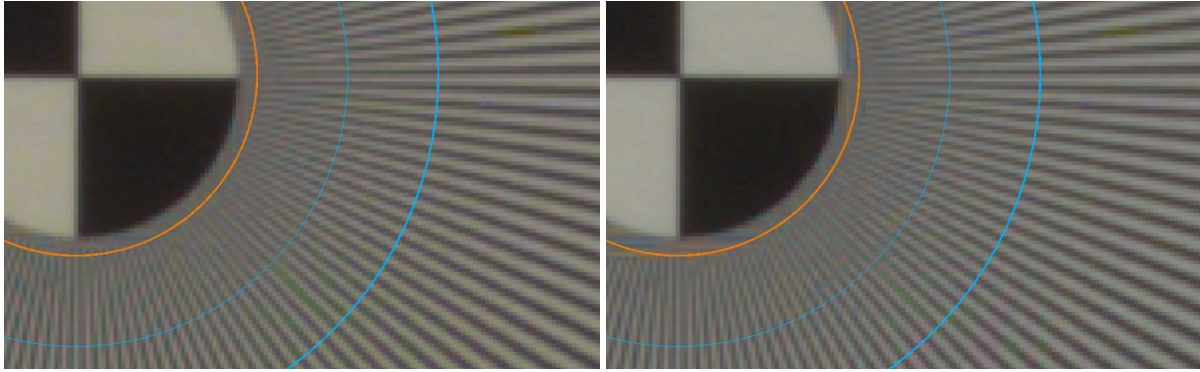
We examined the effects of demosaicing using the four algorithms available in [dcraw](#): bilinear (a simple form of linear interpolation with notoriously low quality), VNG, PPG, and AHD (in order of increasing quality) and the recommended AMaZE algorithm in [RawTherapee](#), which offers many additional algorithms. Our measurements correlated with the descriptions of algorithm quality.

**Table 3. Comparison of demosaicing algorithms**

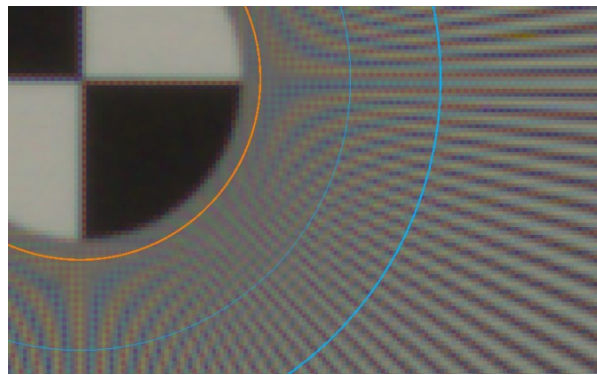
<b>Demosaicing algorithm</b>	<b><math>C</math> (bits/pixel)</b>	<b><math>MTF50P</math> (C/P)</b>
dcraw bilinear	1.8	0.191
dcraw VNG	2.51	0.219
dcraw PPG	3.14	0.229
dcraw AHD	3.69	0.229
RawTherapee AMaZE	3.95	0.236

Information capacity  $C$  is a far more sensitive measure of demosaicing quality than  $MTF50P$ .

The enlarged images below compare three of the algorithms. The **orange** circle is  $f_{Nyq}$ . The **cyan** circles are at  $0.75f_{Nyq}$  and  $0.5f_{Nyq}$ . The two highest quality algorithms (Figure 10) are difficult to distinguish, but artifacts are clearly visible with bilinear demosaicing (Figure 11), which also has reduced sharpness.



**Figure 10.** RawTherapee AMaZE demosaicing (left),  
dcraw Adaptive Homogeneity-Directed (AHD) demosaicing (right).

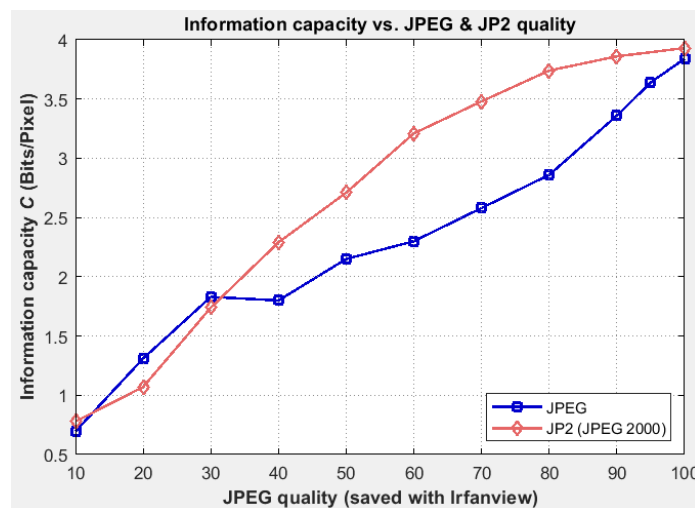


**Figure 11.** Bilinear demosaicing (notoriously low quality)

To summarize, information capacity  $C$  is an excellent way of evaluating demosaicing algorithms, which primarily differ in the amount of aliasing.

### Data compression

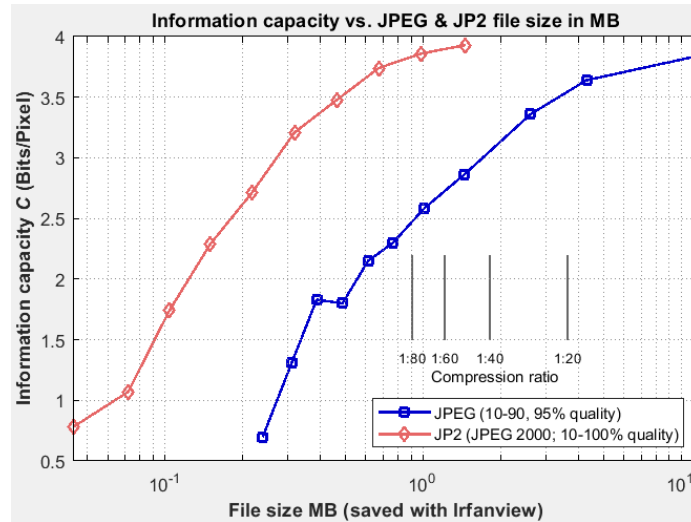
We studied the effects of data compression by loading a high quality raw/TIFF image acquired at ISO 100 from 24-megapixel camera 2 into [Irfanview](#), then saving it as JPEG and JPEG 2000 (JP2) files at quality levels from 10 to 100. Information capacity is a strong function of quality level for both file types.



**Figure 12.** Information capacity  $C$  vs. quality level.

This agrees well with [3] and [4], where several metrics for evaluating JPEG data compression were tested. The conclusion was, “the JPEG ‘quality factor’, used to specify the compression level, was found to have the best performance when predicting the results of subjective tests.”

Figure 13 shows information capacity as a function of file size for an original (uncompressed) file size of  $6024 \times 4024 \times 3$  Bytes = 72.7 MB. JPEG 2000 seems to be the clear winner here. The general trends are similar to results in section 4.7 of [5], which is based on subjective evaluations of a set of images.

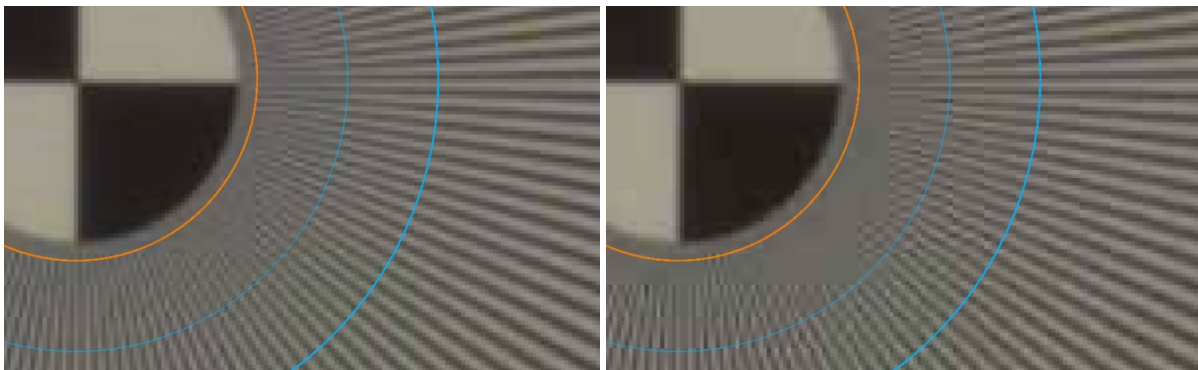


**Figure 13.** Information capacity  $C$  vs. file size in MB. Original file size = 72.7 MB ( $6024 \times 4024 \times 3$ ). Compression ratio is for comparison of results with [5], section 4.7.

We also found that  $MTF_{50P}$ , the spatial frequency in Cycles/Pixel where  $MTF$  drops to half its peak value, correlates poorly with JPEG and JPEG 2000 quality.

These results illustrate the potential of using information capacity  $C$  for evaluating image compression algorithms.

Figure 14 illustrates JPEG image degradation at quality levels of 70 and 50%.



**Figure 14.** 70% JPEG quality level (left), showing some contrast loss near  $f_{Nyq}$  (the orange circle). 50% JPEG quality loss (right), showing significant contrast loss above  $0.75 f_{Nyq}$  (the thin cyan circle).

## Artifact summary

The slanted-edge method is effective in measuring the effect of three types of artifact (clipping, aliasing (from demosaicing), and data compression) on image quality. This was a surprise benefit of measuring noise in the presence of the signal. When noise is measured in flat patches (which is quite widespread), few artifacts are present to degrade image quality.

## Additional details

A few additional details need to be listed. They are described in full detail on [Shannon information capacity](#).

### *Information capacity over the image field*

The results we have presented thus far are information capacity  $C_{star}$  for the region of the star near the center of the image, expressed in units of bits per pixel at a specified ISO speed. The total information capacity  $C_{total}$  for the image field must take variations in  $C$  into account.  $C_{star}$  can be found by one of two methods.

1. Use a grid of Siemens star charts. Such a grid is illustrated in the ISO 2014/2017 standard.  
 $C_{total} = \text{mean}(C_{star}) \times \text{megapixels}$
2. Use a chart with multiple slanted-edges, preferably one of the four **Imatest** charts with automatic Region of Interest (ROI) detection. This method (details on the [Shannon information capacity](#) page) uses a quantity called “star-equivalent information capacity,”  $C_{starEq}$  — an approximation to the information capacity that would have been calculated from Siemens stars (not recommended as a primary calculation). The equation for the total camera information capacity based on  $C_{starEq}$  measured at multiple locations in the image is

$$C_{total} = \frac{C_{star}(\text{center}) \times \text{mean}(C_{starEq})}{C_{starEq}(\text{center})} \times \text{megapixels}$$

### *Color channels*

Most of the results in this paper are for the luminance (Y) channel, where  $Y = 0.2125 \cdot R + 0.7154 \cdot G + 0.0721 \cdot B$ . The separate results for the R, G, and B channels are also of interest. For the Camera 2 raw/TIFF image, which has  $C = 3.69$  bits/pixel for the Y-channel,  $C_R = 2.90$ ,  $C_G = 3.53$ , and  $C_B = 3.07$  for the R, G, and B channels, respectively.  $C$  is higher for the Y-channel because the uncorrelated noise from the R, G, and B is combined.

Total  $C = C_R + C_G + C_B$  is nearly triple when the R, G, and B color channels are analyzed separately. But we can now think of a pixel as having 24 (RGB) bits instead of 8.

In the future we plan to analyze chroma channels for information capacity  $C$ . The most likely candidates are  $C_B$  and  $C_R$  (from  $Y C_B C_R$ ), which are derived from B-Y and R-Y (where Y is mostly green). We will use special charts (most likely stars with R-G and B-G patterns).

### *Future work*

- Standardize the camera information capacity measurement. **Imatest** works with several standards committees, including ISO TC42, IEEE P1958 (Camera Phone Image Quality), and IEEE P2020 (Automotive Image Quality). This will be a challenging process because camera information capacity  $C$  is new and unfamiliar to most committee members.

- Do more to correlate information capacity  $C$  with the subjective visual appearance of a variety of images, without and with additional image processing, beyond the indirect correlations in the section on Data compression. We may extend the model of  $C$  to include viewing conditions and the human visual system to obtain a “visual information capacity”, analogous to visual noise or acutance.
- Correlate  $C$  with performance of Machine Vision and Artificial Intelligence systems.

## Summary

Camera information capacity  $C$  has great potential value as a figure of merit for evaluating camera image quality because it combines the effects of sharpness, noise, and several types of artifact. But until recently it was not easy to measure.

The method presented here makes it convenient to measure information capacity  $C$ , signal  $S(f)$  (proportional to MTF), and noise  $N(f)$  from images of Siemens star test charts.

We once again stress that camera information capacity is a *novel* measurement. Only one aspect is traditional—measuring signal power  $S(f) = MTF(f)^2$  from the Siemens star. The rest—calculating the noise power  $N(f)$  from the Siemens star and applying  $S(f)$  and  $N(f)$  to the two-dimensional form of the Shannon capacity equation (simplified to one dimension) are entirely new.

Despite its unfamiliarity, the units of camera information capacity—information bits per pixel (or total image) for a specified ISO speed—are intuitive and easy to understand. We believe that it is a better indicator of a camera’s potential image appearance (after tuning) than any existing measurement.

In the future we would like to see information capacity — either bits per pixel or megabits total at specified ISO speeds (exposure indices) or light (lux) levels — become an integral part of a standard camera specifications.

---

## Appendix I: Transforming the Shannon-Hartley equation from one to two dimensions and back

The Shannon-Hartley equation for frequency-dependent (colored) noise, first presented in 1948 [2], is

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

Where  $C$  is the channel capacity in bits/pixel;  $B$  = Nyquist frequency  $f_{Nyq}$ ;  $S(f)$  is the signal power spectrum,  $N(f)$  is the noise power spectrum.

This equation applies to one-dimensional measurements (usually signal as a function of time), whereas images have *two* dimensions, i.e., pixels. To overcome this limitation, we need to express  $C$  as a double integral.

$$C = \iint_0^B \log_2 \left( \frac{S(f_x, f_y) + N(f_x, f_y)}{N(f_x, f_y)} \right) df_x df_y$$

where  $f_x$  and  $f_y$  are frequencies in the  $x$  and  $y$ -directions, respectively. In order to evaluate this integral, we transform  $x$  and  $y$  into polar coordinates  $r$  and  $\theta$ .

$$C = \int_0^{2\pi} \int_0^B \log_2 \left( \frac{S(f_r, f_\theta) + N(f_r, f_\theta)}{N(f_r, f_\theta)} \right) f_r df_r df_\theta$$

Since  $S(f_r, f_\theta)$  and  $N(f_r, f_\theta)$  are only weakly dependent on  $\theta$ , the double integral can be rewritten in one-dimension.

$$C = 2\pi \int_0^B \log_2 \left( \frac{S(f) + N(f)}{N(f)} \right) f df$$

This resembles the original Shannon-Hartley equation, except for the factor of  $2\pi$  in outside the integral and  $f$  inside the integral.

## Appendix II: Measuring signal and noise power from the Siemens Star.

Because nonuniform image processing is so commonly applied in consumer digital imaging systems, it is highly desirable to measure signal and noise *at the same location* in the image, i.e., to measure noise in the presence of signal. We have developed a method to accomplish this with the well-known Siemens Star chart, which is a part of the [ISO-12233-2014/2017 standard](#).

To measure signal and noise at the same location, we use an image of a sinusoidal Siemens-star test chart consisting of  $n_{cycles}$  total cycles, which we analyze by dividing the star into  $k$  radial segments (32 or 64), each of which is subdivided into  $m$  angular segments (8, 16, or 24) of length  $P_{seg}$ . The number sine wave cycles in each angular segment is  $n = n_{cycles}/m$ .

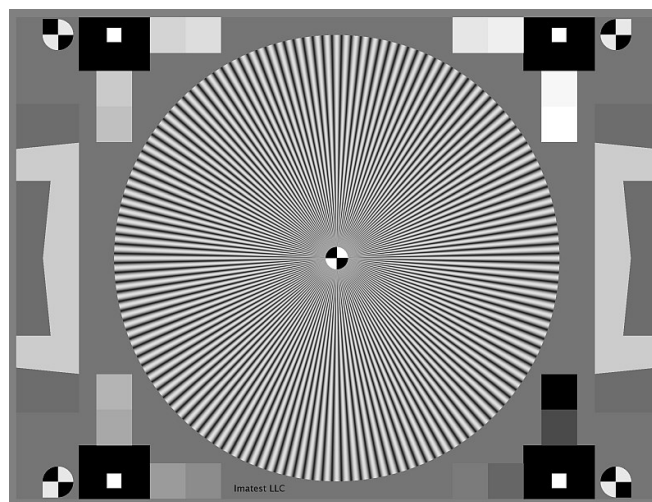


Figure II-1. 144-cycle Siemens star pattern.

For radius  $r$  in pixels, the spatial frequency in Cycles/Pixel is  $f = n_{cycles}/2\pi r$ . This means that the Nyquist frequency,  $f_{nyq} = 0.5 C/P$  is located at  $r = 46$  pixels for a 144-cycle star and 23 pixels for a 72-cycle star [6].

We calculate the ideal (pure sine wave + second harmonic) signal  $s_{ideal}(\varphi)$  for each segment from the actual (noisy) input signal  $s_{input}(\varphi)$ , using [Fourier series coefficients](#)  $a_j$  and  $b_j$ .

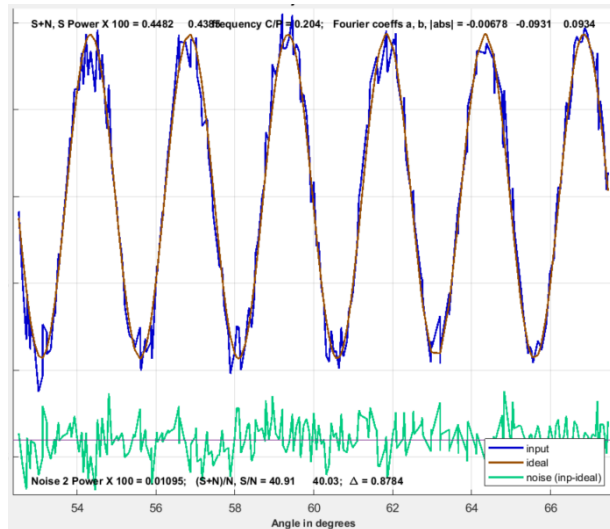
$$s_{ideal}(\varphi) = \sum_{j=1}^2 a_j \cos\left(\frac{2\pi j n \varphi}{P_{seg}}\right) + b_k \sin\left(\frac{2\pi j n \varphi}{P_{seg}}\right)$$

where

$$a_j = \frac{2}{P} \int_P s_{input}(x) \cos\left(\frac{2\pi j n x}{P_{seg}}\right) dx, \quad b_j = \frac{2}{P} \int_P s_{input}(x) \sin\left(\frac{2\pi j n x}{P_{seg}}\right) dx$$

The second harmonic term ( $j = 2$ ) term improves results for JPEG images from cameras, which frequently have “shoulders” (regions of reduced contrast) in their tonal response curves, making them difficult to linearize with a simple gamma formula.

The minimum star frequency is  $f_{min} = n_{cycles}/2\pi r_{star}$  where  $r_{star}$  is the radius of the star in pixels. The contribution of frequencies below  $f_{min}$  is calculated by extrapolation— by assuming that  $S(f) = S(f_{min})$  and  $N(f) = N(f_{min})$  for  $f < f_{min}$ . This assumption is reasonable for  $N(f)$  but somewhat pessimistic for  $S(f)$ . However, the effect on  $C$  is minimal because of the presence of  $f$  in the integrand of the equation for  $C$ .



**Figure II-2.** Illustration of  $s_{input}(\varphi)$  (blue),  $s_{ideal}(\varphi)$  (brown), and noise  $N(\varphi) = s_{ideal}(\varphi) - s_{input}(\varphi)$  (green).

The spatial domain noise in each segment is

$$N(\varphi) = s_{ideal}(\varphi) - s_{input}(\varphi)$$

The mean signal power used to calculate  $C$  is  $S(f) = \text{mean}(a_n^2 + b_n^2)/2$  for all angular segments at radius  $r$  (where  $f = n_{cycles}/2\pi r$ ). Frequency domain noise power  $N(f)$  is the mean of the variance ( $\sigma^2$ ) of the middle 80% of  $N(\varphi)$  for each angular segment (there can be irregularities near the ends).



**Saturation (clipping)**— Clipped (saturated) regions of an image contain no information as well as no noise, but have a very high Signal-to-Noise ratio. Since this can lead to erroneous results, saturated regions must be removed before calculating the noise power (the variance of  $N(\varphi)$ ). For minimum and maximum allowable pixel levels  $p_{min}$  and  $p_{max}$  (typically 0 and 1), the locations specified below should be used for calculating the noise power (the variance) of the segment

$$N'(\varphi) = N(\varphi) \text{ where } s_{input}(\varphi) > p_{min} + \varepsilon \text{ and } s_{input}(\varphi) < p_{max} - \varepsilon ; \varepsilon \text{ is a small margin.}$$

Figure II-3 is a 3D pseudocolor plot of a small pie-shaped segment of the chart image, transformed in to a rectangular area, illustrating the appearance of saturation (clipping) for an overexposed image.

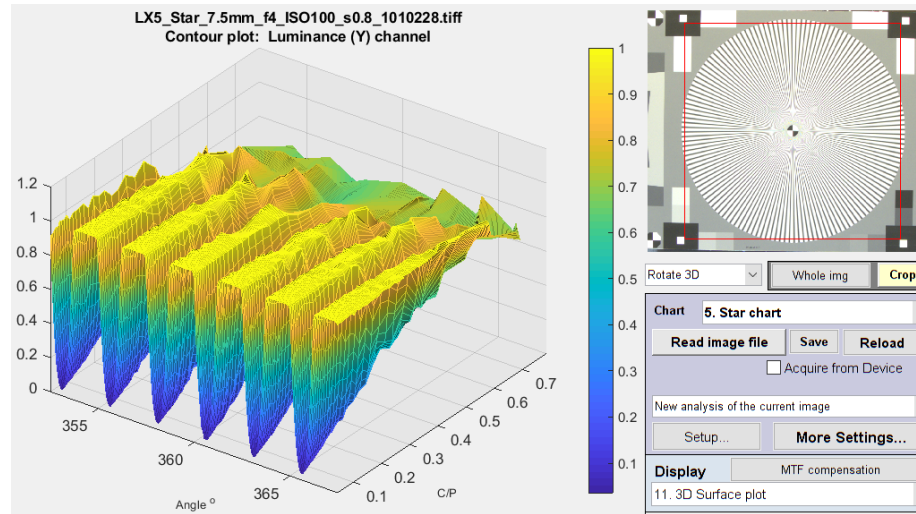


Figure II-3. 3D surface plot of overexposed image, illustrating strong clipping.

The Shannon information capacity  $C$  of this image (Figure 8) is 3.44 bits/pixel. For a comparable image, slightly overexposed, but with no saturation, it is 3.93 bits/pixel.  $C$  cannot be measured accurately in the presence of saturation because the equations assume a nearly linear system. Nevertheless, it is lower than the unsaturated image (as it should be), unlike with slanted-edges. We have seen [instances where clipping may have been deliberately introduced for pictorial effect](#), but it should be avoided in applications where information content is important.

## References

- [1] C. E. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J., vol. 27, pp. 379–423, July 1948; vol. 27, pp. 623–656, Oct. 1948. [people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf](http://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf)
- [2] C. Shannon, "Communication in the Presence of Noise", Proceedings of the I.R.E., January 1949, pp. 10-21. [fab.cba.mit.edu/classes/S62.12/docs/Shannon\\_noise.pdf](http://fab.cba.mit.edu/classes/S62.12/docs/Shannon_noise.pdf)
- [3] R. Jenkin, R. Jacobson, M. Richardson, "Use of the first order Wiener kernel transform in the evaluation of SQRIn and PIC quality metrics for JPEG compression," Proc. SPIE 5294, Image Quality and System Performance, 2003, Electronic Imaging 2004.
- [4] A. Ford, "Relationships between Image Quality and Still Image Compression", PhD Thesis, University of Westminster, UK, 1997. (not available on the internet)
- [5] E. Allen, "Image Quality Evaluation in Lossy Compressed Images", PhD Thesis, University of Westminster, UK, 2017. [westminsterresearch.westminster.ac.uk/item/q0vq5/image-quality-evaluation-in-lossy-compressed-images](http://westminsterresearch.westminster.ac.uk/item/q0vq5/image-quality-evaluation-in-lossy-compressed-images)
- [6] U. Artmann, "Quantify Aliasing a new approach to make resolution measurement more robust," Electronic Imaging 2019, Society for Imaging Science and Technology.

