Camera imaging quality testing with MTF:

A complete guide from beginner to professional

By Dan Xie, Neil Kang, Alec Liu, Thomas Tao, Henry Hu

BACKGROUND INTRODUCTION

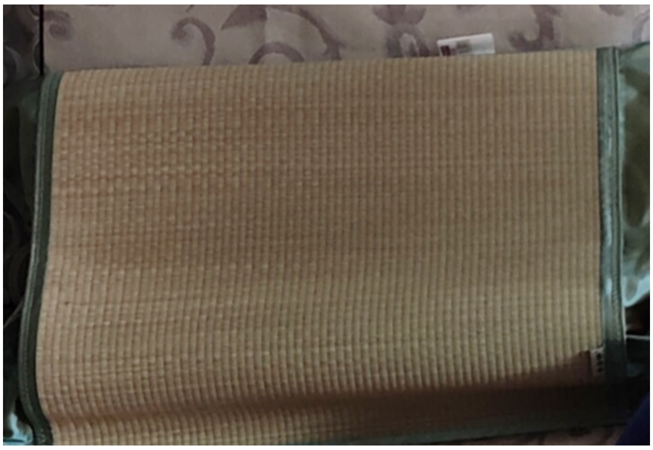
Consumers often equate camera resolution with the number of pixels, with a higher number denoting better resolution. In reality, the number of pixels is only a partial factor in camera imaging resolution, with other factors including lens specification, sensor size, auto focus accuracy and the ISP algorithm also impacting the resolution of the final image or video.

Example: The two images below were taken using the author's two smartphone cameras using the pixels setting (around 12M pixels). As you can see, the image of the right has a much lower resolution.

If the number of pixels is not a reliable indicator of camera resolution, what professional metrics can be used?

ISO12233: Photography - Electronic still picture imaging - Resolution and spatial frequency responses contains several options, of which modular transform function (MTF) is the most frequently used.

This paper explains the professional MTF in an easy-to-understand way, with illustrations. By the end, you'll even be able to use Python code to calculate MTF using the edge-based spatial frequency response (e-SFR) method.



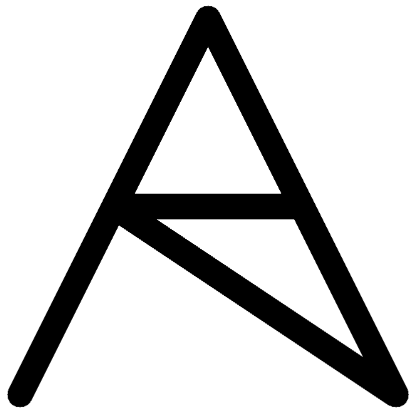
UNDERSTANDING THE BASICS OF CAMERA RESOLUTION

(1) CAMERA PIXELS NUMBER

One pixel is one sampling point on the sensor. The greater the number of pixels, the more accurately the camera can rebuild the images.

Example: Letter A on the left is drawn with 100x100 pixels, on the right 1,000x1,000 pixels.

On the other hand, more pixels do not always mean better image quality. Packing more pixels onto the same-sized sensor results in smaller individual pixels, which can lead to reduced image quality, especially increased noise.

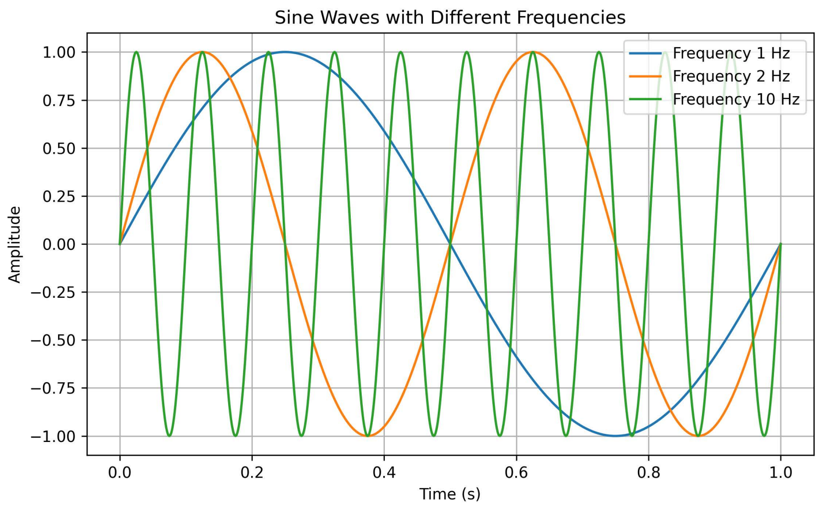


(2) SPATIAL FREQUENCY

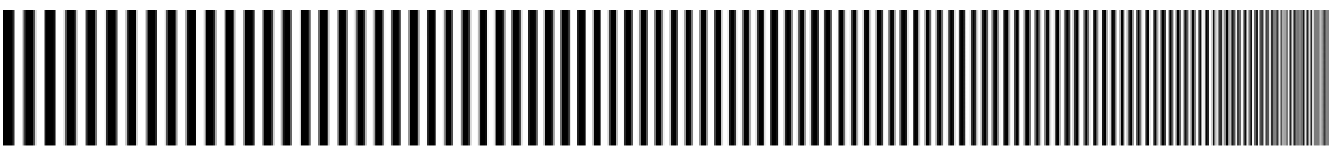
Frequency is typically described in the time domain. For example, electrical power usually has a frequency of 50 Hz or 60 Hz, meaning the AC power completes 50 or 60 cycles per second. A sine wave serves as the basic signal; in the time domain, its frequency indicates how many cycles it repeats in one second.

In the spatial domain, frequency is expressed in units such as cycles per dimension unit.

In camera imaging industry, this spatial frequency is usually defined at cycles/pixel (C/P), lines/image height and line pairs/image height.



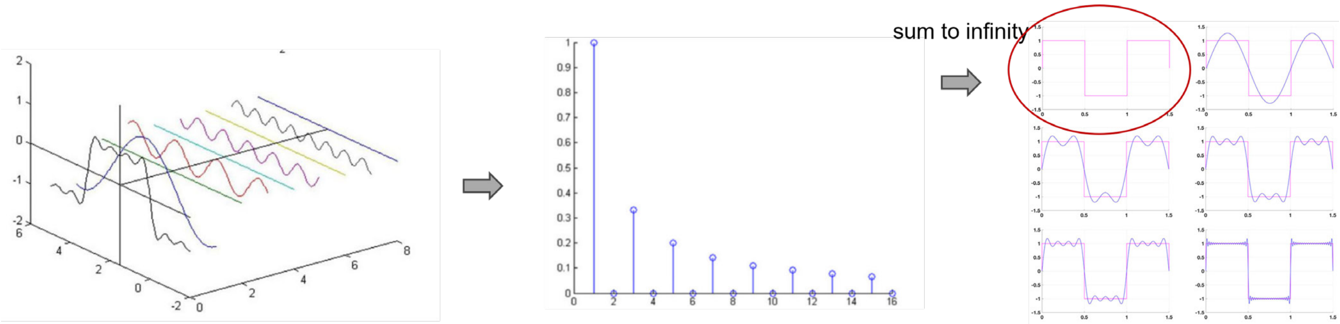
Example: Plot below shows black and white line pairs increasing in spatial frequency.



(3) FOURIER TRANSFORM

Any time or space domain signal can be the sum of different signals within the frequency domain.

The illustration below demonstrates that a rectangular signal in the time domain is the sum of an infinite series of sine waves of varying frequencies and amplitudes.



(4) NYQUIST LIMIT

The Nyquist limit arises from digital signal sampling theory, which states that at least two sampling points per cycle are required to accurately reproduce the original signal.

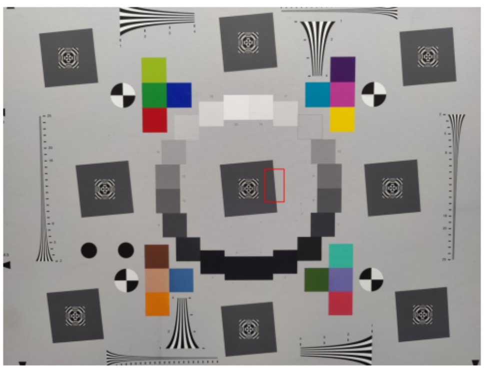
As illustrated in the plot below, a minimum of two sampling points per cycle is necessary to digitally reproduce a sine wave (Method B). For a digital camera, the Nyquist limit frequency is 1 cycle per 2 pixels, or 0.5 cycles per pixel.

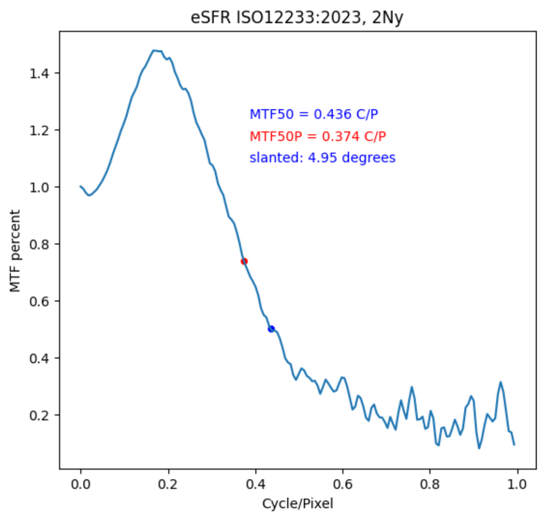
WHAT IS CAMERA MTF?

Modular transform function (MTF) is a way to express camera imaging spatial frequency in two metrics: modulation and frequency.

Frequency has already been explained, but the concept of modulation is more complex.

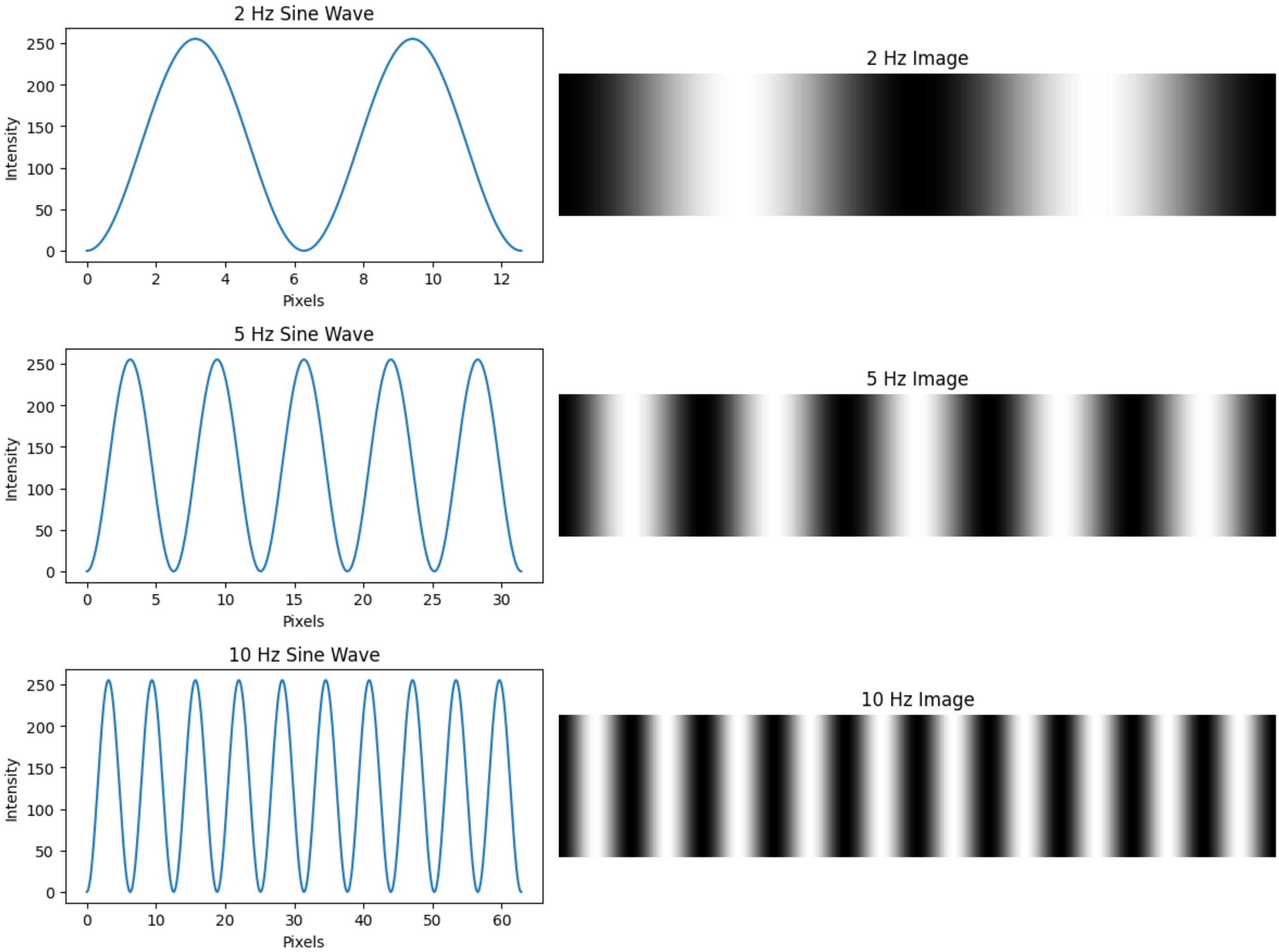
The plot below shows a typical MTF result for a camera - X axis is the camera spatial frequency and Y axis is the camera modulation.





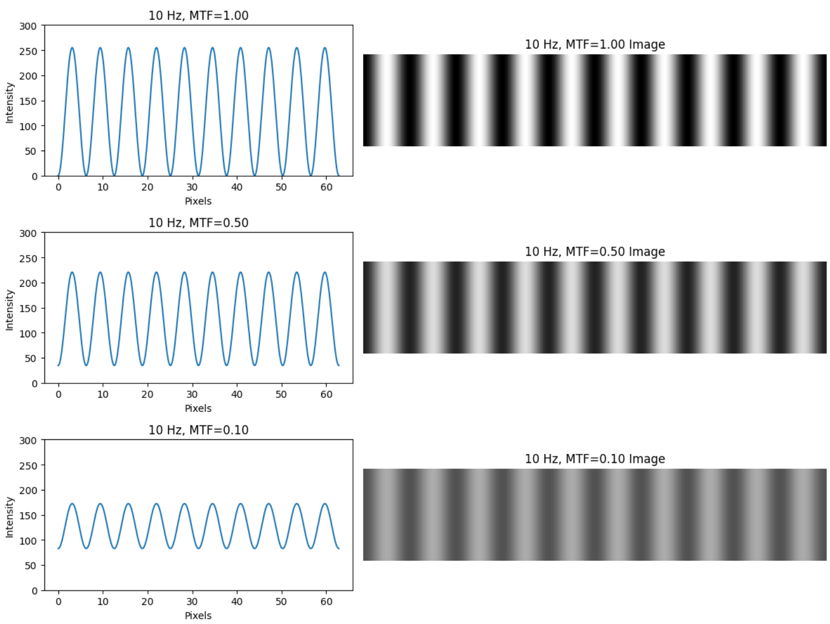
To understand modulation, let us start with a few examples. In a digital image, which is typically 8-bit, the brightest value is 255, and the darkest is 0. Now, if we use 255 as the peak (crest) of a sine wave and 0 as its lowest point (trough), we can visualize how the sine wave varies at different spatial frequencies.

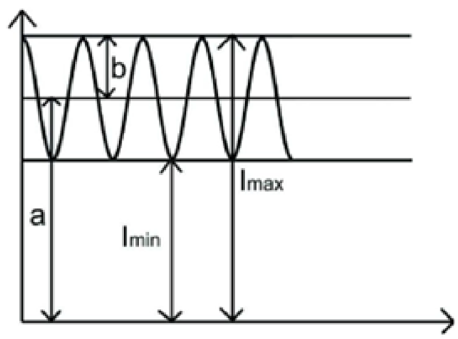
The resulting sine waves, with brightness values ranging from 0 to 255, are shown on the right.



Let’s examine how the waves will appear when we keep their frequency at 10Hz and vary the modulation values to [1, 0.5, 0.1]. They will now appear as below.

Note: the patterns below use a gamma correction of 1/2.2 and an initial intensity digital value range of (0, 255)





The formula provided above applies to the absolute modulation value. For MTF plots, the industry typically uses relative modulation values. In this context, MTF50 refers to the modulation value where the relative modulation is 50%.

$$modulation\_relative\_value = \frac{certain\ absolute\ modulation\ value}{initial\ absolute\ modulation\ value}$$

From the two examples above, you should observe the following:

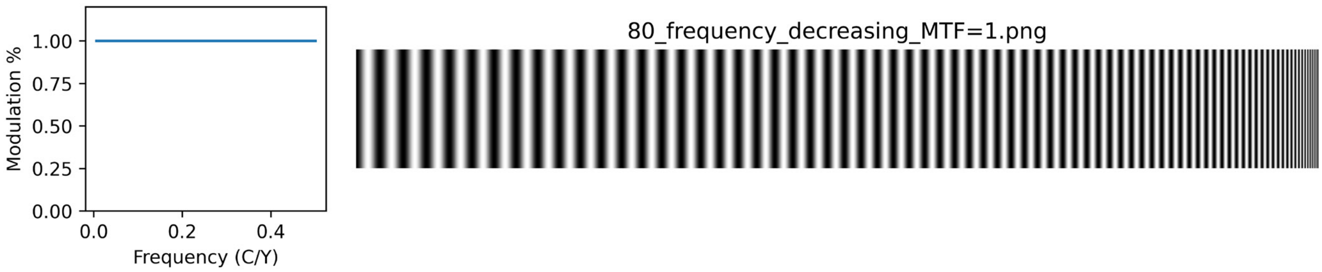
1. Higher spatial frequency: As the spatial frequency increases, the black and white stripes of the sine wave become closer together. When the frequency becomes too high, the human eye can no longer distinguish between the black and white lines
2. Lower modulation: Reduced modulation results in lower contrast between the light and dark parts of the sine wave. When modulation is too low, the black and white lines also become indistinguishable to the human eye

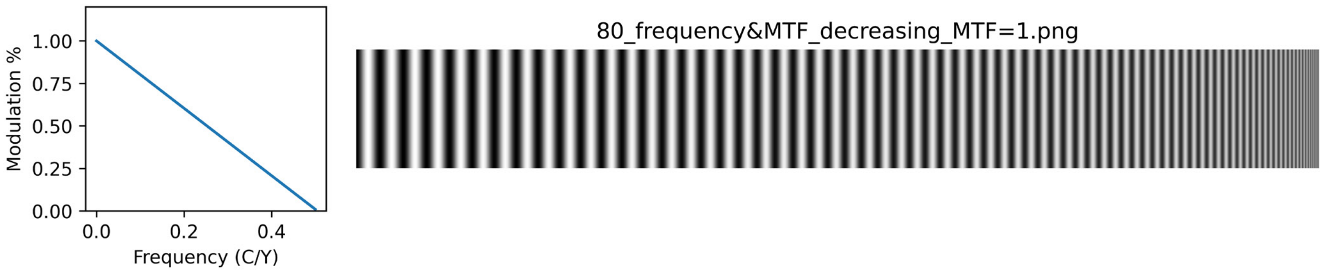
Let us now revisit the previous MTF plot. The curve in this plot illustrates the camera's resolution performance in terms of MTF. To better understand this, let us take MTF50 as an example, where the frequency value is 0.418 C/P. This can be interpreted in one of two ways:

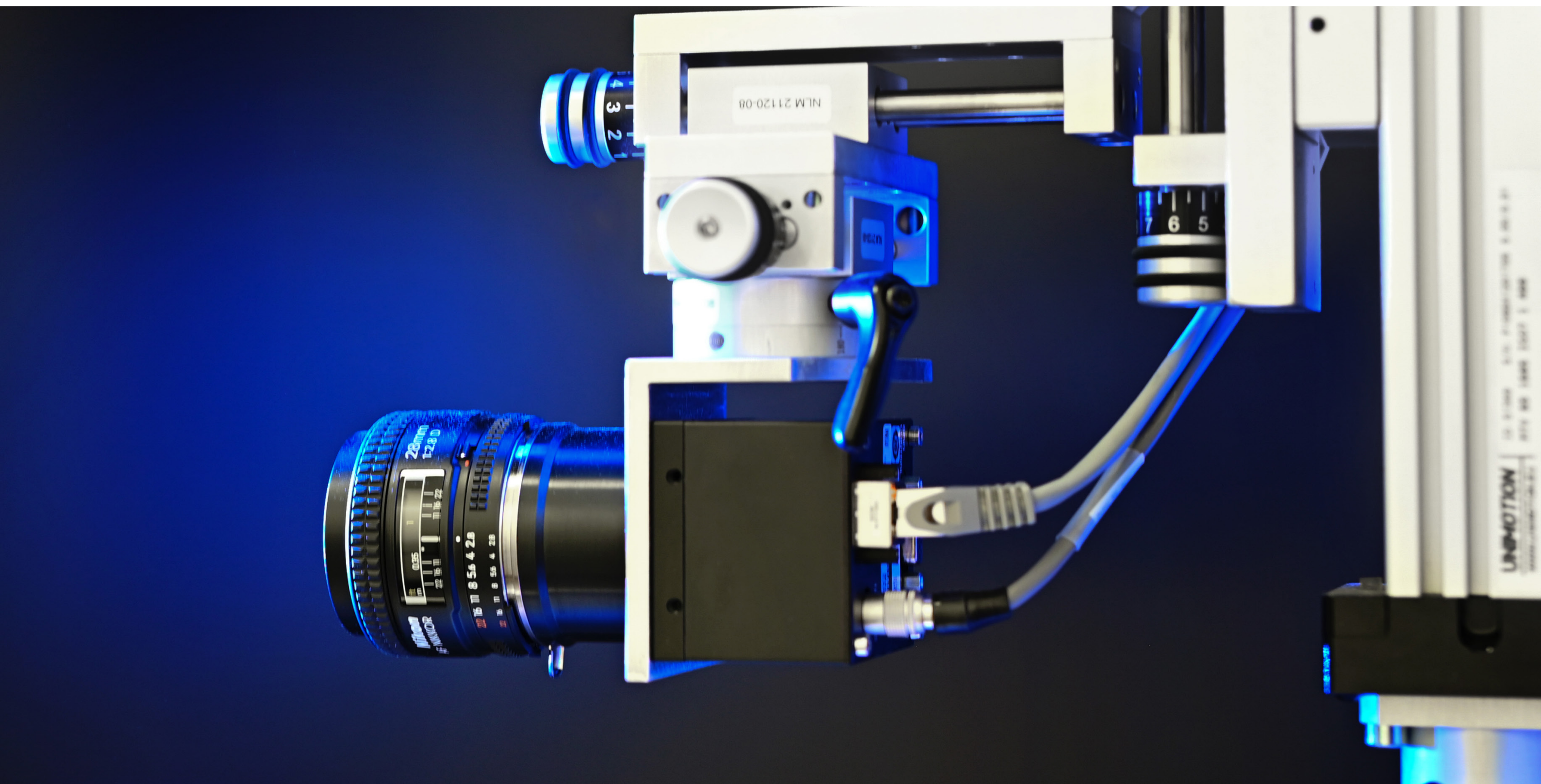
1. The camera can achieve a resolution of 0.418 C/P when its modulation is 0.5 (50% of the initial modulation value)
2. The camera reaches a 50% relative modulation level when the spatial frequency is 0.418 C/P

Finally, let us use two patterns to illustrate how the sine wave pattern changes with variations in frequency and modulation. In the plot, the left side shows the MTF curve, while the right side displays the corresponding sine wave pattern.

From the perspective of a human eye, as the modulation decreases (towards the bottom of the plot), the line pairs in high-frequency region become less distinguishable.







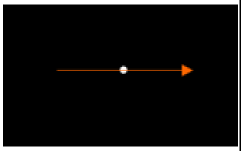
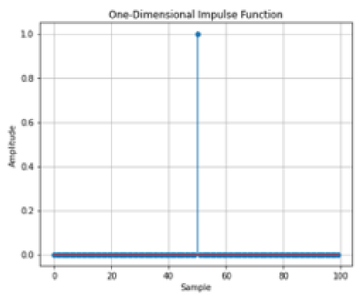
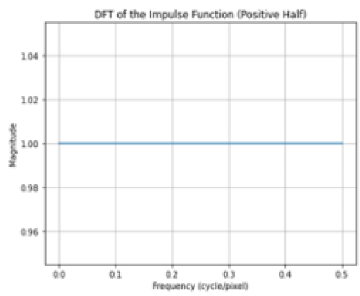
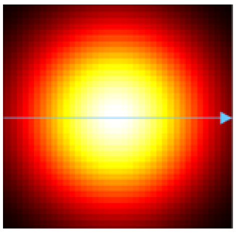
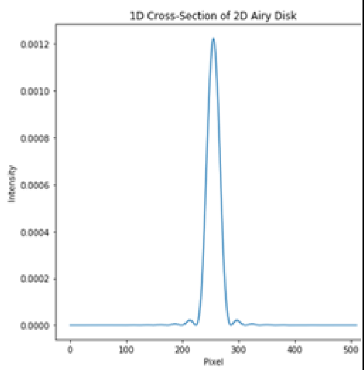
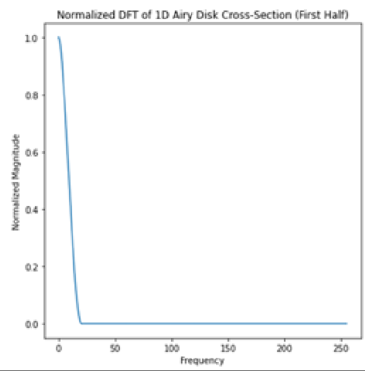
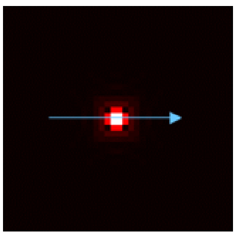
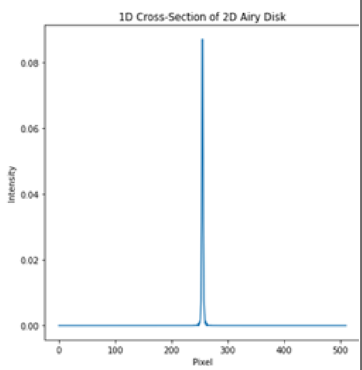
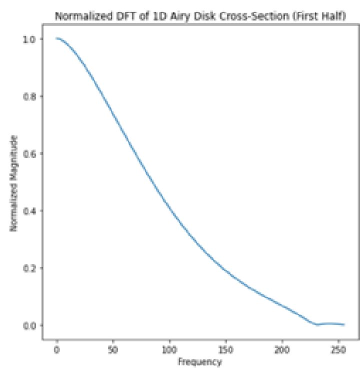
MTF MEASUREMENT THEORY

There are two methods for measuring MTF performance in a camera, as outlined in ISO 12233: the e-SFR method and sine wave-based spatial frequency response (s-SFR) method. The s-SFR method is straightforward. It uses a test chart with sine wave patterns ranging from low to high frequency. The e-SFR method uses the Fourier transform of the lens point spread function, also known as Airy disk, While the calculation process is quite complex, the test chart is quite simple.

Since the e-SFR method is more widely used in the industry, we will focus on this method.

THE ORIGIN OF THE E-SFR METHOD:

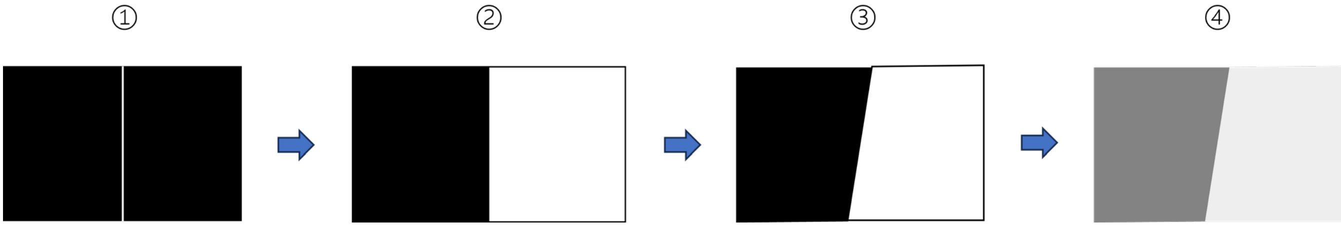
1. An infinitely small light point corresponds to an ideal impulse signal in the spatial domain. The Fourier transform of this impulse signal results in an ideal straight line in the frequency domain, as shown in the plots in the first row below
2. When a camera with a small lens captures an image of this infinitely small light point, the Airy disk phenomenon causes the light point to appear as a series of waves, like the image in the second row. Its Fourier transform is no longer a straight line but rather a decreasing curve
3. Similarly, if the camera with a larger lens is used, the result is shown in the third row. In this case, you may find the Fourier transform curve resembles MTF plot of the camera commonly used in testing

Case	Luminance intensity in the spatial domain	Luminance intensity in the frequency domain
Infinitely small light point 		
Small lens capturing an infinitely small light point 		
Large lens capturing an infinitely small light point 		

USING A SLANTED-EDGE PATTERN TO MEASURE E-SFR

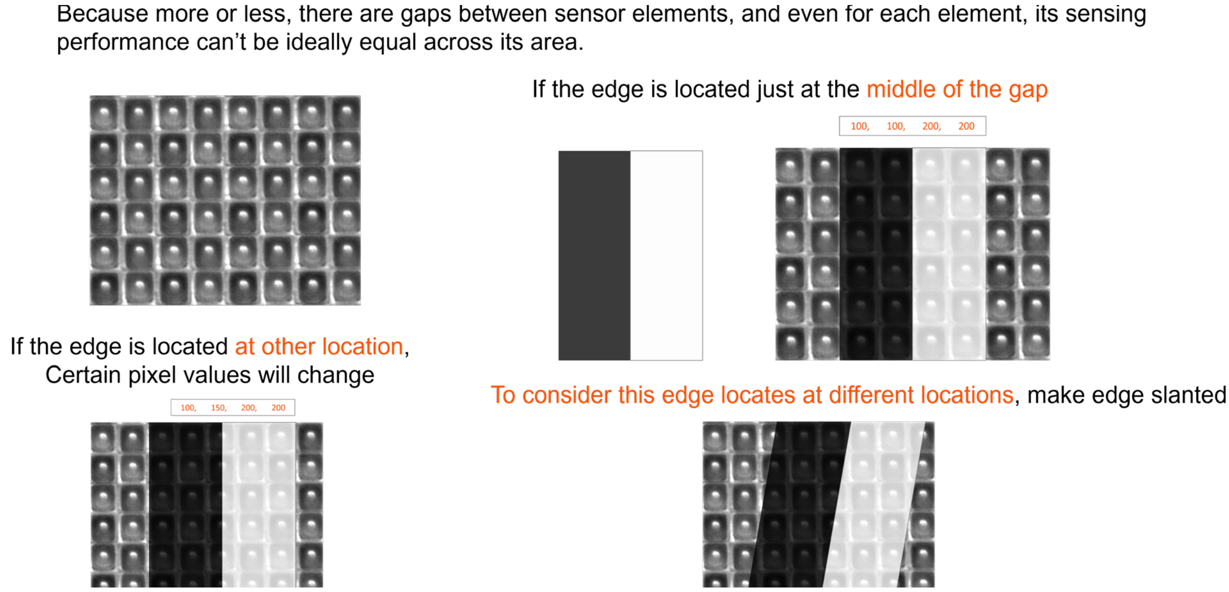
An infinitely small point is difficult to print and challenging for a camera to sample accurately, as camera sensor pixels have a finite size. A more practical way to achieve a similar outcome is to use an infinitely thin line ①, from which multiple MTF lines can be calculate and then averaged.

However, an infinitely thin line ① also has the same challenge that both printing techniques and camera sampling cannot truly replicate infinite size. Therefore, an edge pattern ② is used, as its first directive transform ideally resembles an infinitely thin line.



So, why have we intentionally slanted this edge? ③ The reason for this is that the physical sampling points (pixels) of a camera sensor are not infinitely small, and there are gaps between them.

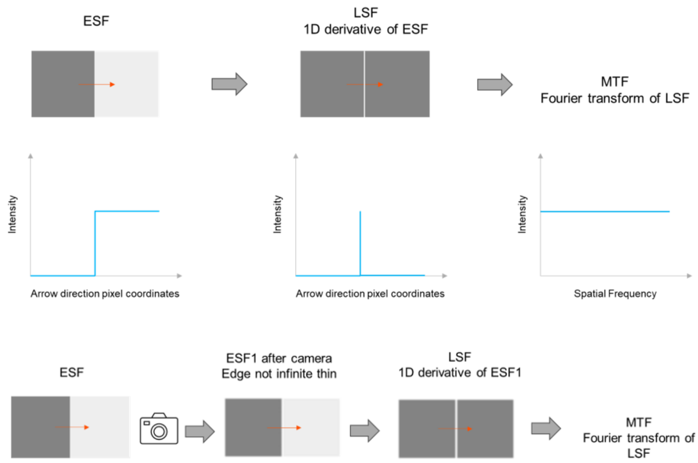
As shown below, the digital pixel values will vary depending on the precise location of the edge. By using a slanted edge (③) and averaging the values of each horizontal line, the variations caused by sampling location differences can be mitigated.

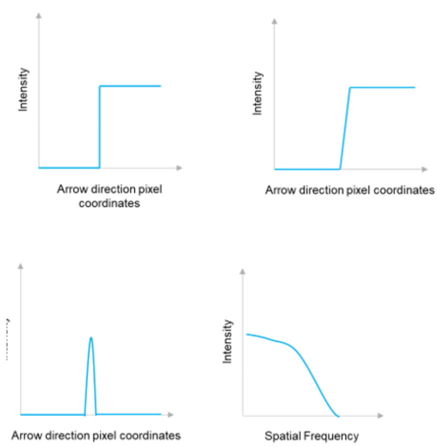


Finally, to prevent exposure saturation in both dark and light areas, a moderate contrast edge pattern ④ with a recommended contrast ratio of 4:1 is used.

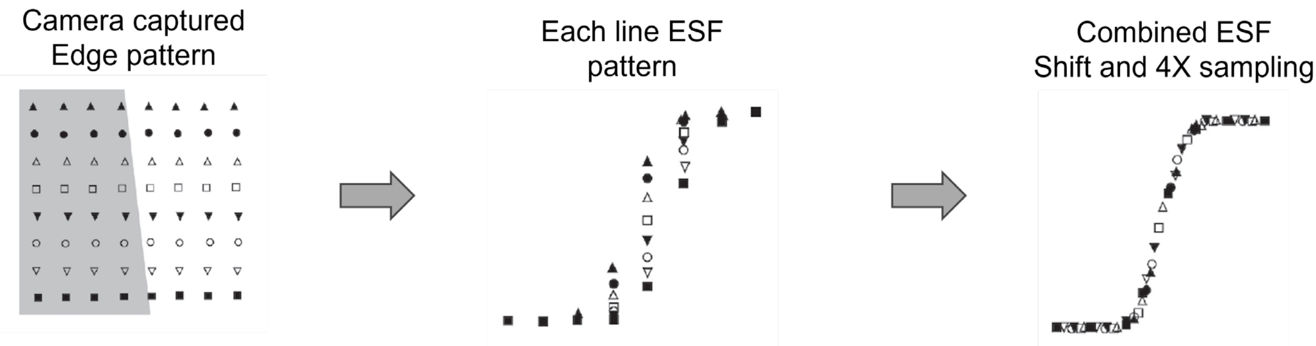
Summarizing the transformation process from an edge pattern to a camera MTF plot:

The first image shows the ideal MTF plot of the edge pattern without the influence of a camera. The second image illustrates a typical camera MTF plot when capturing this edge pattern.

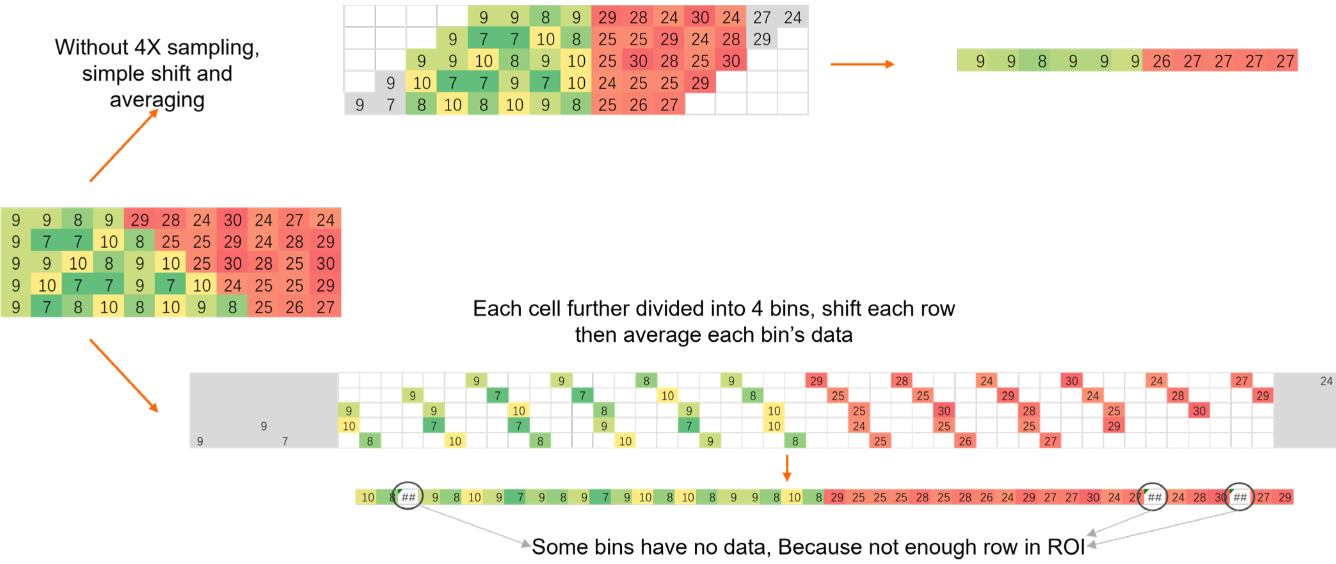




The process below illustrates the general steps of this 4x super-sampling approach.



This process involves two major steps: (1) shifting the data in each row and (2) obtaining data that is 4x the original length in a single row. To better illustrate this 4x super-sampling process, we provide a specific, simplified example below.



As shown in previous demonstrations, the averaged 4x sample data may contain blank elements. To minimize the likelihood of this issue:

- Select a greater height when choosing the region of interest (ROI) (ISO 12233 recommended a minimum of 100 rows)
- If blank elements still occur, use an additional algorithm to fill them in. Although ISO 12233 does not specify a method, common approaches include using the value of a neighboring element or the average of the values on both sides

MTF CALCULATION STEPS IN PYTHON

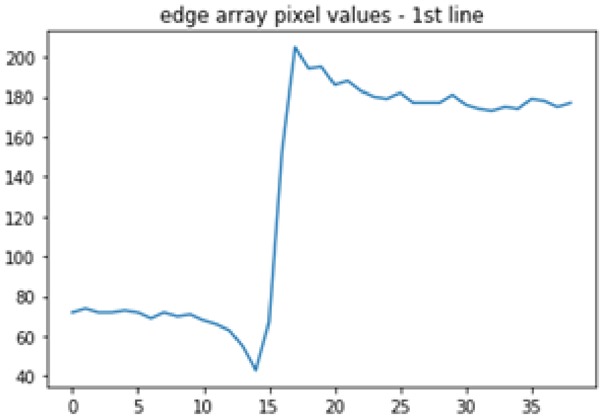
In this last section, we will apply our knowledge by outlining the key calculation steps in Python. If you are a skilled Python user, these guidelines should help you write a complete Python code to perform e-SFR MTF analysis.

SELECT ROI

a. Read the image and convert it from RGB to Y, using the appropriate formula. Typically, the ITU-R BT.709-6 formula is used: Y = 0.2126R + 0.7152G + 0.0722B

b. Use OpenCV function selectROIs( ) for manual selection of rectangular ROIs

c. Rotate the ROI to a horizontal orientation, positioning the dark area on the left and the light area on the right, to create the initial pixel value ROI data array, which we will refer to as ROI_pixel_value



INVERSE OECF TRANSFORM (LINEARIZATION)

The ROI data from step (1) is a 2D NumPy array of pixel values. To convert these values into luminance, we perform a transformation known as linearization in ISO 12233. The resulting data is referred to as 'linearized data'.

There are generally two approaches for this:

a. Calculate the opto-electronic conversion function (OECF) per ISO14524: Typically, this involves using the gray patches included in the e-SFR chart. The OECF is then used to generate a new data array: ROI_luminance.

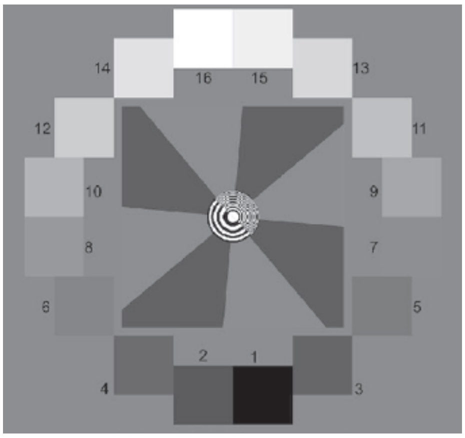
Below is an example of 16 gray patches from the ISO 12233:2023 standard e-SFR chart. To perform the calculation, the density values of these gray patches must be known. The relationship between density and luminance is given by the formula below. Note that E can be set to 1, as the luminance values in this calculation will ultimately be relative percentages, which cancels out E.

$$L_i = \frac{10^{-D_i}\ E}{\pi}$$

where

D_i is the grey-scale patch visual density;

E is the illuminance incident on the chart (measured using a cosine corrected photometer), in lux.



b. Simpler approach using luminance contrast: An easier method is to use the known luminance contrast of the test edge pattern, which is typically 4:1.

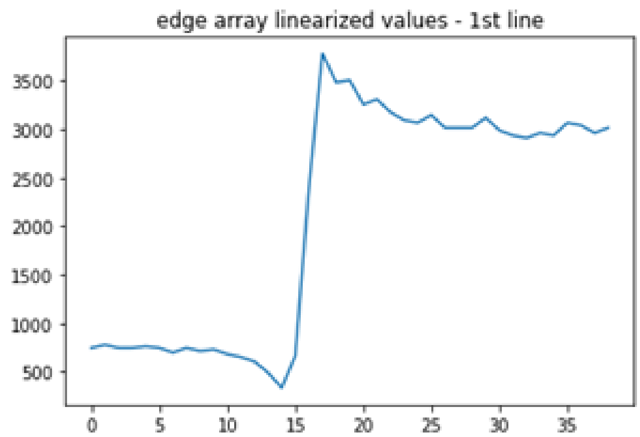
   Using this contrast, you can calculate the camera's gamma value with the formula provided below. Once the gamma value is determined, apply it to the formula to transform the ROI_pixel_value data array into the ROI_luminance data array.

$$Pixel\_Value = A * Lumiance^{gamma}$$

Where:

A can be set to 1, as the final luminance values will be relative percentages, causing A to cancel out in the calculation

Note: The OECF method is more accurate than the gamma method because the camera's transform function from luminance to pixel value is usually not a single power function.



FIRST ROUND CENTRAL LOCATION CALCULATION

ROI_luminance is a 2-dimensional NumPy array, and so the next task is to find the central location of each row in this data array.

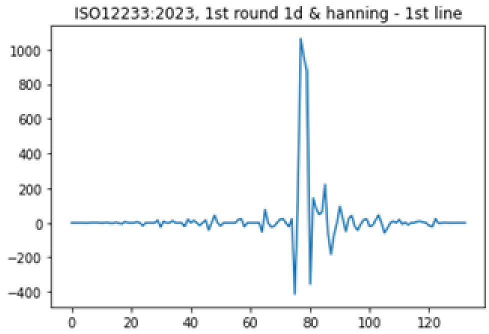
Central location is defined by the following formula in ISO 12233, however the formula in the current version is not allied with its official MATLAB code. Here is the correct formula:

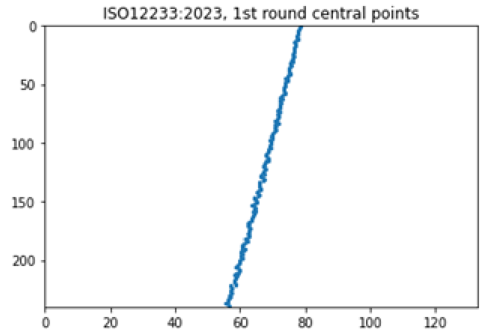
$$C(\mathrm{r}) = \frac{\sum_{p=1}^{P} p * \varphi(p,r)}{\sum_{p=1}^{P} \varphi(p,r)} - 0.5$$

- p is each row index number
- Φ is each pixel linearized value
- P is array horizontal length

Detailed steps for this first round central location calculation are outlined in the table below.

	TASK	LIBRARY/FUNCTION	NOTES
Step 1	First directive of ROI_ luminance	NumPy	[-1/2, 1/2] filter
Step 2	Apply Hann window	NumPy, hann( )	Hann window center at array center
Step 3	Calculate each line central location	NumPy	Use central location formula above
Step 4	Determine central location polynomial formula	NumPy, polyfit( )	Fifth order



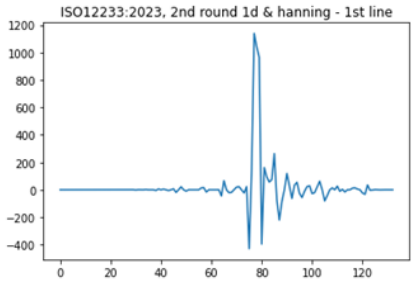


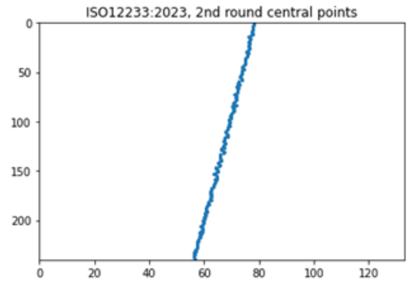
SECOND ROUND CENTRAL LOCATION CALCULATION

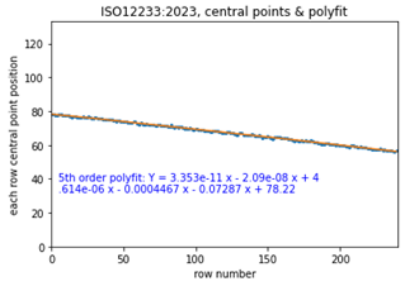
The first round central location calculation is quite rough because the Hann filter window it applied to each row of the data array using the same central location.

Ideally, however, the Hann filter should be applied at each row's individual central location. The purpose of the second central location calculation is to achieve this.

	TASK	LIBRARY/FUNCTION	NOTES
Step 1	First directive of ROI_ luminance	NumPy	[-1/2, 1/2] filter
Step 2	Apply Hann window	NumPy, hann( )	Hann window center at first round central locations
Step 3	Calculate each line central location	NumPy	Use central location formula above
Step 4	Determine central location polynomial formula	NumPy, polyfit( )	Fifth order







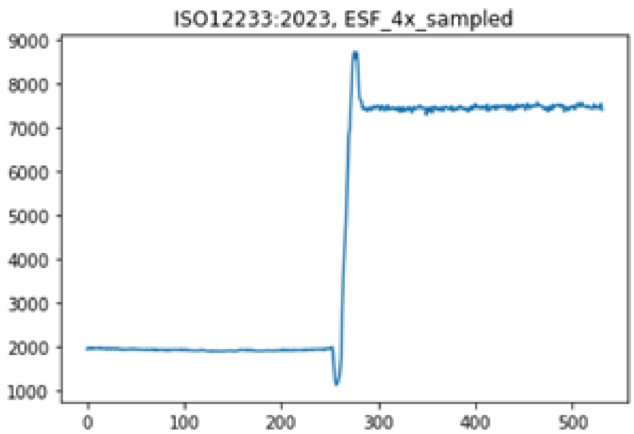
4X SUPER-SAMPLING, SHIFT AND ESF

So far, we have two important data sets: a 2-dimensional array of ROI_luminance and a list of central locations for each row of ROI_luminance.

The next task is to convert this 2D ROI_luminance array into a single array with four times the original length.

Detailed steps are listed below.

	TASK	LIBRARY/FUNCTION	NOTES
Step 1	Each row expands to 4x data	NumPy	For each original data point, place it into the first bin of the four expanded bins.
Step 2	Shift each row based on second round central location formula	NumPy, roll( )	Blank bins after shift fill with nan
Step 3	Average each column of 4x shifted array	NumPy	For bins with a nan value, decide what value to adopt



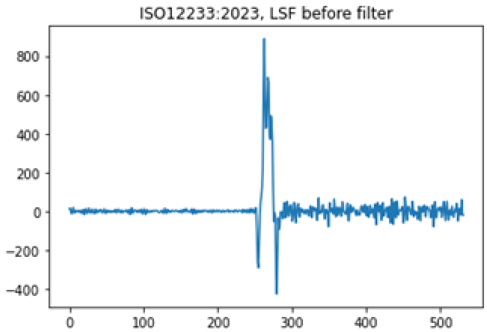
CONVERT 4X ESF TO 4X LSF

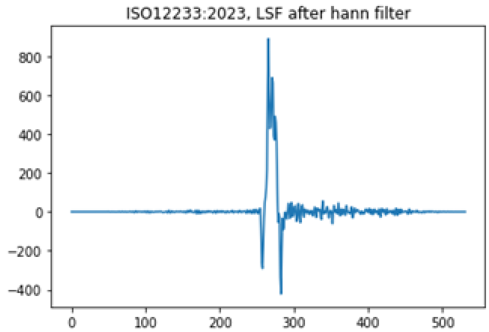
MTF is the Fourier transform of the line spread function (LSF) data array. Currently, we have a 4x length ESF data array. The derivative of the ESF is the LSF, and ISO 12233 provides the formula below to calculate this first derivative (where ESF refers to the 4x ESF', and LSF' to the 4x LSF):

$$LSF^{'}(j) = \frac{ESF'(j+1) - ESF'(j-1)}{2}, \text{ for } j = 2, \ldots, N-1$$

The table below details the steps to do this transformation.

	TASK	LIBRARY/FUNCTION	NOTES
Step 1	Calculate LSF by filter [-1/2, 0, 1/2]	NumPy	LSF[0]=LSF[1], LSF[-1]=LSF[-2]
Step 2	Roll LSF peak to center	NumPy, roll( )	-
Step 3	Apply Hann window	NumPy, hann( )	-





FOURIER TRANSFORM TO GET INITIAL MTF

Up to step (6), the 4x-length, single-row data array we obtain in the spatial domain is crucial and is referred to as LSF'W in ISO 12233.

While the process to derive the LSF is quite complex, the remaining steps are straightforward.

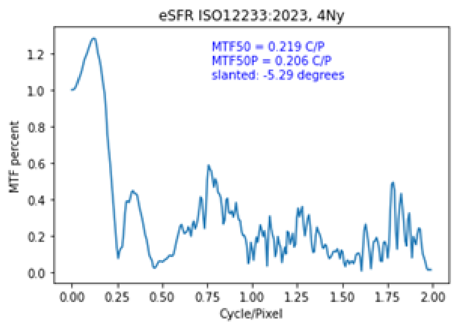
They involve performing a discrete Fourier transform (DFT), calculating the relative amplitude percentage (with the initial modulation set to 1), and applying a correction factor $D\ (k)$. The formulas below are used for this transformation."

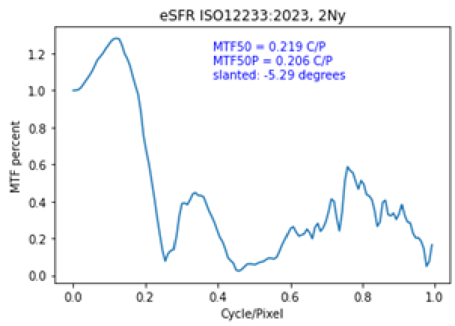
$$e\text{-}SFR(k) = D(k)\left|\frac{\sum_{n=1}^{N} LSF_W^{'}(n)\, e^{-\frac{i2\pi kn}{N}}}{\sum_{n=1}^{N} LSF_W^{'}(n)}\right|$$

$$D(k) = \min\left[\frac{1}{\operatorname{sinc}\left(\frac{2\pi k}{N}\right)}, 10\right]$$

$$\operatorname{sinc}(\varnothing) \equiv \frac{\sin \varnothing}{\varnothing}$$

$$\operatorname{sinc}(0) = 1$$



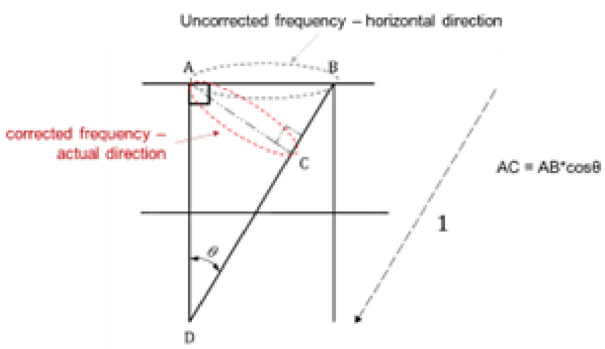


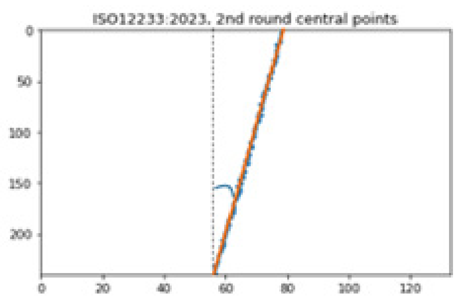
SLANTED-ANGLE CORRECTION FOR FREQUENCY AND MTF CALCULATION

The 4x super-sampling is performed in the AB direction (illustrated below), but ideally, it should be done at the AC direction.

Since 4x super-sampling impacts the final MTF frequency values, a correction is needed to adjust the frequency values from the AB to the AC ratio:

$$\overline{AC} = \overline{AB}\cos\theta$$

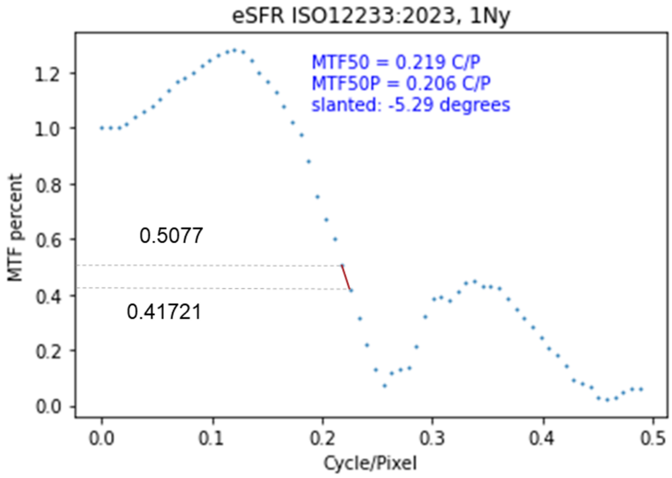




To calculate the slanted-angle (θ), use the centrals points from the second round and apply a first order polynomial fitting approach to determine the polynomial parameters. This will allow you to compute the θ value.

To precisely obtains a specific MTF value, such as MTF50, an additional step is required. First, identify the two frequency points where the modulation values are just above and below 50%, as shown in the plot below. For example, one point has a modulation value of 0.5077, and the other is 0.41721.

Next, use a first-order polynomial fit for the line connecting these two points. Using this polynomial function, you can accurately determine the MTF50 (where modulation = 0.5), which in this case is 0.219 cycles per pixel (C/P)



BIBLIOGRAPHY:

- ISO12233, Photography — Electronic still picture imaging — Resolution and spatial frequency responses
- ISO14524, Photography — Electronic still-picture cameras — Methods for measuring opto-electronic conversion functions (OECFs)
- ISO 9334, Optics and photonics — Optical transfer function — Definitions and mathematical relationships
- CIPA DC-003-Translation-2020 Resolution Measurement Methods for Digital Cameras

We are SGS – the world's leading testing, inspection and certification company. We are recognized as the global benchmark for sustainability, quality and integrity.

SGS's optical testing capabilities cover both imaging and display optics. The imaging optics laboratory assesses the quality of camera products across various industries, including mobile phone cameras, surveillance cameras, automotive cameras and medical cameras. Meanwhile, the display optics laboratory evaluates the display quality and touch functionality of products such as screens, AR/VR devices and HUDs.

By combining the comprehensive capabilities of imaging and display optics, our optical laboratories can perform thorough and professional optical quality testing on products that integrate both imaging and display technologies, such as camera monitor systems (UN-R46), under-display systems and AR devices.

Our technical personnel are actively involved in the development and updating of domestic and international testing standards. They are members of standards committees for ISO TC42, IEC TC100, IEC TC110 and IEC TC124.

When you need to be sure

SGS Headquarters
1 Place des Alpes
P.O. Box 2152
1211 Geneva 1
Switzerland

ee.global@sgs.com

sgs.com/ee