Measuring MTF with wedges: pitfalls and best practices

Norman Koren, Robert C Sumner, and Henry Koren; Imatest LLC; Boulder, CO USA

Abstract

As digital imaging becomes more widespread in a variety of industries, new standards for measuring resolution and sharpness are being developed. Some differ significantly from ISO 12233:2014 Modulation Transfer Function (MTF) measurements. We focus on the ISO 16505 standard for automotive Camera Monitor Systems, which uses high contrast hyperbolic wedges instead of slanted-edges to measure system resolution, defined as MTF10 (the spatial frequency where MTF = 10% of its low frequency value). Wedges were chosen based on the claim that slanted-edges are sensitive to signal processing. While this is indeed the case, we have found that wedges are also highly sensitive and present a number of measurement challenges: Sub-pixel location variations cause unavoidable inconsistencies; wedge saturation makes results more stable at the expense of accuracy; MTF10 can be boosted by sharpening, noise, and other artifacts, and may never be reached. Poor quality images can exhibit high MTF10. We show that the onset of aliasing is a more stable performance indicator, and we discuss methods of getting the most accurate results from wedges as well as misunderstandings about low contrast slanted-edges, which correlate better with system performance and are more representative of objects of interest in automotive and security imaging.

Introduction

Applications of digital imaging are rapidly expanding in a number of industries, including automotive, security, and medical, to mention a few. As these applications grow, new standards are being established [1] to ensure that image quality meets the specific needs of each industry.

Perhaps the most important image quality factor is sharpness or resolution, which determines how much detail an image can convey. Both are defined in terms of Modulation Transfer Function (MTF) (which is identical in practice to Spatial Frequency Response, SFR). MTF is the contrast of a sine wave pattern at spatial frequency f relative to low frequencies.

The native spatial frequency units for calculating MTF is Cycles per Pixel (C/P). MTF is typically reported in C/P or in units that can be directly derived from C/P, such as Line Widths per Picture Height (LW/PH), where f (LW/PH) = 2 f (C/P) × Picture Height. Picture Height can be arbitrarily chosen to meet ISO 16505 requirements.

Sharpness and resolution have similar, though not quite identical, meanings. Sharpness is related to overall viewer perception, and is closely associated with MTF50, the spatial frequency where MTF drops to half (50%) of its low frequency value. Resolution (often short for vanishing resolution) has many definitions. In our context it refers to the highest spatial frequency where detail is visible, and is closely related to the smallest distance between two distinguishable objects—known as the Rayleigh limit, which also corresponds to MTF10. While this is a classical theoretical value for describing the resolution of an optical system, in practice MTF10 can be difficult to measure in noisy real-world situations.

This paper focuses on the ISO/FDIS 16505 standard [2], which specifies automotive mirror replacement via Camera Monitor Systems (CMS). Omissions and ambiguities in the resolution measurement method defined in this standard have been a source of frustration for engineers. Details of the standard are covered in the “Handbook of Camera Monitor Systems- the Automotive Mirror-Replacement Technology based on ISO 16505” [3], which we will refer to as the “CMS Handbook”.

We begin by describing hyperbolic wedge and edge-based MTF measurements in depth, then we compare results for the two, pointing out reasons for the discrepancies and best practices for obtaining reliable measurements.

Resolution test charts

We focus on two test patterns for measuring MTF: the slanted-edge and the hyperbolic wedge, both of which are included in the ISO 12233:2014 standard [4] and referenced in ISO 16505.

The hyperbolic wedge consists of a converging set of black bars on a white background that linearly increase in spatial frequency. Figure 1 (a crop of the ISO 12233:2000 chart) contains wedges with 5 and 9 bars. Both the ISO 12233:2000 and 2014 revisions specify wedge contrast to be between 40:1 and 80:1, with no real justification for such high contrast, which frequently causes saturation and clipping problems and is not representative of real objects that need to be distinguished. We prefer the contrast to be around 10:1.

Figure 1. Hyperbolic wedges and slanted-edge from the ISO 12233:2000 target.

The wide bar shown in Figure 1, which can be used as a low frequency reference for wedge MTF calculations, contains slanted-edges on either side. The tilt angle is approximately 5 degrees. In the ISO 12233:2000 standard the contrast ratio of the edge was also specified as between 40:1 and 80:1. This high contrast caused images to saturate or clip in many practical situations, compromising MTF measurement accuracy. As a result, the ISO 12233:2014 revision specifies a relatively low 4:1 slanted-edge contrast.

Hyperbolic wedge MTF measurements

ISO 16505 chose to use hyperbolic wedges for MTF for the following reason, stated in Annex E.1.

...
“Edge enhancement is a well-known technology among others techniques but such a processing will strongly affect the reproducibility of the SFR measurement. Along the discrete sampling of image; the SFR measurement improperly used can lead to incorrect results of limit resolution measurement. … Therefore, a traditional resolution measurement method using black and white hyperbolic resolution chart is advised to be used to evaluate the resolution (MTF) performance of the CMS.”

We will show that the hyperbolic wedge is at least as susceptible to edge enhancement and has a number of other issues that affect the accuracy of hyperbolic wedge MTF measurements: the MTF definition, sub-pixel positioning, sharpening (edge enhancement), clipping (saturation), and nonlinear tonal response (gamma).

We begin by describing the ISO 16505 calculations, which do not produce a plot of MTF versus spatial frequency. Such plots are invaluable for characterizing systems and identifying potential problem areas. We then describe calculations in the Imatest program [5].

**ISO 16505 MTF and spatial frequency calculations**

An MTF calculation based on hyperbolic wedges is not defined in ISO 12233:2014. Instead, it defines a visually-determined resolution limit as the spatial frequency where the number of visible bars drops below the low frequency count. This frequency is also called the “onset of aliasing”, $f_{\text{alias}}$, (though this nomenclature applies primarily to high resolution systems where the bar count is limited more by aliasing than by contrast loss.)

Although ISO 16505 does not explicitly define a wedge-based MTF (it is vaguely described in section 3.2 and Annex D.3), it is defined in equations (12) and (13) in the CMS Handbook, 3.8.3. For signal amplitude $I$ measured on a scan line across the wedge at a location with spatial frequency $f$ (Figure 2),

\[
M(f) = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}} \tag{1}
\]

\[
\text{MTF}(f) = \frac{M(f)}{M(0)} \tag{2}
\]

A problem with this definition is that it is only correct for sine waves [6], while each cross-section of a black-to-white hyperbolic wedge is representative of a square wave, which has a fundamental Fourier component that is larger than its amplitude by a factor of $4/\pi$ [7].

\[
X_{\text{square}}(i) = \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\sin(2\pi(2k-1)/i)}{2k-1} \tag{3}
\]

The reproduced bars change from square waves at low spatial frequencies to sine waves at spatial frequencies where $M(f)/M(0)$ drops below about 0.7, i.e. at high spatial frequencies where harmonic content (3rd, 5th, and higher orders) is sufficiently suppressed so the bars appear to be sinusoidal.

MTF at high spatial frequencies, where signal amplitude $I$ is sinusoidal, would be correct if MTF were normalized to $\pi/4$, i.e., if equation (2) were rewritten as $\text{MTF}(f) = \pi M(f)/4M(0)$. But MTF at low spatial frequencies would be incorrect (too low) if this equation were used.

Another problem with implementing Equation (1) is that $I_{\text{max}}$ and $I_{\text{min}}$ are affected by noise, interference from adjacent bars, and by the sampling phase of the bars with respect to the pixel boundaries. Figure 2 is an example of the cross-section of a noisy wedge with 9 bars at approximately $0.7 \times$ Nyquist frequency. There are some similar images in Figures E.5 of ISO 16505, where several distinctly different candidates for $I_{\text{max}}$ and $I_{\text{min}}$ are visible. The ISO standard is silent about how to determine which local extrema to use for the above calculation. Is it most appropriate to take the largest maximum and lowest minimum? Or is the mean of the maxima and minima more appropriate? The Imatest calculation, described in Equations (6)-(8), circumvents this issue and also resolves the square wave error.

An additional problem with the ISO 16505 MTF definition is that there is little mention of linearizing the image. Most images have gamma encoding applied, where gamma is typically around 0.5 for common color spaces like sRGB, which are designed for display at gamma = 2.2). Gamma encoding can be expressed as

\[
P = L^\gamma \quad \text{where } \gamma \text{ is the encoding gamma.} \tag{4}
\]

(Actual tonal response may be somewhat more complex.) Equations (1) and (2) are not correct (even for sine waves) unless the image is linearized, i.e., an approximate inverse of gamma encoding is applied.

Sharpening, which is widely applied in automotive (and most other types of) imaging adds an additional complication for gamma-encoded images. Sharpening is a linear operation that typically involves subtracting shifted and amplitude-reduced replicas of the image from the original image. It is usually applied after gamma-encoding. $I_{\text{max}} + I_{\text{min}}$, which is used to normalize $M(f)$ in Equation (1) is not strongly affected by sharpening because $I_{\text{max}}$ and $I_{\text{min}}$ tend to move in opposite directions. But when the image is linearized, the interaction between bars shifts $I_{\text{max}} + I_{\text{min}}$, resulting in a modest error. This error is not present in slanted-edge MTF calculations, which are effectively normalized by the large light and dark areas far from the edges, and hence are relatively unaffected by sharpening artifacts.

There are no wedge-based MTF plots in ISO 16505. Instead, Annex E.3 of the standard recommends deriving MTF from a cross-section of the wedge at a spatial frequency of interest.
procedure can be a tedious and is subject to error. It is not suited for high volume or automated testing.

**Imatest MTF and spatial frequency calculations**

MTF can be calculated more accurately from a cross section of a wedge by first finding the spatial frequency of the bars, then deriving a new set of coefficients based on Fourier coefficients. Spatial frequency is initially calculated in units of Cycles per Pixel (C/P), which can be converted into Line Widths per Picture Height (for any specified height) with the simple equation,

\[
f (\text{LW/PH}) = 2 f (\text{C/P}) \times \text{Picture Height}
\]

We illustrate the calculation for a vertical wedge.

1. Find the boundaries of the wedge (taking care to ignore interfering patterns like tic marks or numbers) at each value of y.
2. For signal amplitude I measured at each scan line across the wedge, find mean(I) inside the boundaries to use as a threshold for finding locations where I − mean(I) crosses from positive to negative or negative to positive. These crossings are used to determine where the number of detected bars drops below the low frequency value, corresponding to \( f = f_{\text{alias}} \).
3. For each scan line (y-location) where the number of detected bars equals the low frequency value \( f = f_{\text{alias}} \), the mean of the positive-to-negative and negative-to-positive crossing intervals is the period of the bars in pixels, equal to the inverse of the spatial frequency \( f \) in cycles/pixel (C/P).
4. For hyperbolic wedges, spatial frequency varies linearly with distance. Use a linear regression fit at each scan line for \( f < f_{\text{alias}} \) to find a first order equation for the spatial frequency \( f(y) \) for use in all calculations and plots. [Note: we obtained better results using \( f < 0.9 f_{\text{alias}} \) for the regression fit.]

Because Equation (1) is strictly correct only for sine waves and because \( l_{\text{max}} - l_{\text{min}} \) is subject to errors from noise, sampling phase, and interference from neighboring bars, we replace \( l_{\text{max}} - l_{\text{min}} \) in Equation (1) with the Fourier coefficients of I inside the wedge, where boundaries \( x_1 \) and \( x_2 \) represent a discrete number of detected periods.

\[
C_{\cos} = \int_{x_1}^{x_2} I(x) \cos(2\pi f(y)) \, dx / (x_2 - x_1)
\]

\[
C_{\sin} = \int_{x_1}^{x_2} I(x) \sin(2\pi f(y)) \, dx / (x_2 - x_1)
\]

\[
M(f) = \sqrt{c_{\cos}^2 + c_{\sin}^2} / l_{\text{max}} \Delta l_{\text{min}}
\]

Using \( \sqrt{c_{\cos}^2 + c_{\sin}^2} \) instead of \( l_{\text{max}} - l_{\text{min}} \) is far more accurate because it represents the sine coefficients, derived from an integral (a sort of synchronous detection) that reduces the effects of noise at signals at frequencies other than \( f(y) \). Equation (8) is substituted into Equation (2) for all the Imatest wedge-based MTF calculations that follow.

**Wedge-based MTF**

The following example illustrates many of the measurement difficulties arising from real-world images. It consists of a modified ISO 12233:2000 chart (with an added low-contrast slanted-edge) that has been converted from a linear raw format using gamma = 0.5 (Equation (4); pixel level = brightness^gamma; similar to widely-used color spaces such as sRGB).

No sharpening has been applied. The image is unevenly illuminated and the white background is saturated near the center (the bottom of the crop). Even illumination would have been desirable, but uneven illumination is very common, especially with wide-angle lenses that exhibit strong light falloff.

Figure 3. Crop of image (originally 3120 pixels high) used to obtain the results in Figures 4-8.

The unsharpened image is shown.

Figure 4. Wedge-based MTF from Figure 3 (no sharpening applied) using Imatest software
Figure 4 shows the MTF from the wedges on the left and right of Figure 3, with a portion (at the bottom) of the upper thick central slanted bar used for a low frequency reference.

The low and high frequency wedges are shown rotated on the top. Edge boundaries (red), onset of aliasing $f_{alias}$ (red), and Nyquist frequency $f_{nyq}$ (blue) are displayed. There is some roughness in the unsmoothed (thin dotted black) MTF curve in the main plot, caused by the uneven illumination at frequencies where the two wedges overlap. The smoothed (thick black) MTF curve is more reliable. The red curve displays the normalized bar count. The (thin dotted) unsmoothed curve is strongly affected by noise, and is not reliable. The (thick red) smoothed curve is much more reliable. The MTF result is affected by highlight saturation (though the sharpened image is much more affected).

Note that the onset of aliasing $f_{alias}$ (also called the resolution limit), where the bar count drops from its maximum value, is at a lower frequency than the 10% MTF frequency ($MTF_{10}$).

Note also that the $MTF$ response is quite flat at spatial frequencies where $MTF$ is between 0.07 and 0.11. Measurements made along a flat curves like this are especially susceptible to small changes in noise and signal processing, which can produce large changes in the measurement. In this example, a small change in the system response could cause $MTF_{10}$ to be anywhere between 2400 and 4800 LW/PH. This is frequently observed for the low-valued tails of $MTF$ curves, which is why we conclude that $MTF_{10}$ is not a robust measurement.

An important limitation to wedge-based MTF measurements is that results can be sensitive to sub-pixel shifts. We used FiveFocal Imager [8] to simulate several sub-pixel shifts for a fairly sharp 1280x720 pixel image where unsharpened $MTF_{50}$ (which is far enough from the Nyquist frequency to be relatively unaffected by the shifts) was between 0.277 to 0.291 C/P. $MTF_{10}$ at Nyquist frequency (0.5 C/P; the most sensitive measurement) varied from 0.189 to 0.395. $MTF_{10}$ varied from 0.670 to 0.743 C/P.

In Figure 5 the image shown in Figure 3 has been sharpened using the MATLAB `imsharpen` function with radius = 2 and amount $= 1.8$. This type of sharpening is extremely common. Sharpening is generally beneficial for images intended for human vision (machine vision is another story), but excessive sharpening can cause artifacts—halos near edges—that can cause errors in image interpretation.

The $MTF$ curve (black line) has a slight sharpening peak and is significantly more extended—$MTF_{50}$ has increased from 0.183 to 0.320 C/P and $MTF_{10}$ has increased from 0.413 to 0.558 C/P. The onset of aliasing $f_{alias}$ (resolution limit) has increased from 0.339 to 0.408 C/P. With a little added noise (not shown), $MTF$ remains above the 10% level, which means $MTF_{10}$ is not calculated.

The situation where $MTF$ never drops to the 10% level and $MTF_{10}$ is not calculated is addressed in ISO 16505 Annex D.3.

In some CMS system or measuring environment, the signal amplitude of the CMS output displayed image captured using an evaluation reference camera might not decrease down to 10 %. ... All such effects prevent an exact determination of the spatial frequency point where modulation intensity might decrease down to 10 % of the original chart image. In such case, the limiting resolution of the CMS shall be determined by way of “Visual Resolution” evaluation as defined in ISO 12233, combined by the complementary parallel bar chart for distinguishability limit verification.”

This recommendation has a number of shortcomings. The ISO 12233 visual resolution is, in fact, the onset of aliasing (where the bar count starts dropping). Features above this frequency may not be reliably distinguished, even if $MTF$ is above 10%. The $MTF$ curve flattens out at high spatial frequencies, often where $MTF$ is in the vicinity of 10%, and we have seen situations where a ramp in the response causes $MTF_{10}$ to be unreasonably high—sometimes well beyond the Nyquist frequency $f_{nyq}$ (0.5 C/P). We recommend the following equation for the MTF limit (so named to avoid confusion with resolution limit, etc.).

$$MTF \text{ limit} = \min(MTF_{10}, f_{alias}, f_{nyq})$$

Finally, we need to point out that the onset of aliasing $f_{alias}$ (the resolution limit) is the one measurement that can be performed conveniently with hyperbolic wedges but not with slanted-edges.

**Slanted-edge MTF**

The slanted-edge is one of the two test patterns recommended by ISO 12233:2014 for MTF calculations. The other, the Siemens Star, is not discussed here because it is much less efficient in the use of test chart space. Low contrast slanted-edges are sensitive to signal processing, but are sufficiently resistant to saturation so the effects of strong signal processing, i.e., spatial and temporal peaks caused by sharpening, are clearly visible. Strong sharpening is common across a variety of industries, and is often applied excessively, resulting in visible “halos” near edges that can potentially cause image features to be misinterpreted. These halos can be masked by saturation in high contrast hyperbolic wedge measurements, making wedges appear to be unaffected by such sharpening.

The test image we have been analyzing contains both low and high contrast slanted-edges. The high contrast edges are of interest here because their contrast is similar to the hyperbolic wedges, but
the low contrast edges are of greater interest because they clearly reveal sharpening artifacts.

For accurate slanted-edge MTF (or SFR) calculations, the image must be linearized, i.e., the encoding gamma ($\gamma$) must be removed. For low contrast slanted-edges, the linearization procedure in Equations (10) and (11), which is based on the OECF approximation in Equation (4), can produce excellent MTF results.

If the edge contrast $C_{\text{chart}} = \frac{L_{\text{light}}}{L_{\text{dark}}}$ (for chart reflectance or transmittance $L$) is known and the image pixel contrast $C_{\text{pixel}} = \frac{P_{\text{light}}}{P_{\text{dark}}}$ (for pixel levels $P$) is measured, the image can be linearized with excellent accuracy using the following equations.

\[ C_{\text{pixel}} = C_{\text{chart}}^\gamma; \quad \gamma = \log(C_{\text{pixel}})/\log(C_{\text{chart}}) \]  
\[ P_{\text{linear}} = P^{1/\gamma} \quad \text{linearizes the pixel levels } P \]  

For example, for chart contrast = 4 and mean light and dark pixel levels (for the edge region away from the edge itself) of 180 and 90, $\gamma = \log (180/90)/\log (4) = 0.5$. Section 3.13 and Annex E.1 of ISO 16505 contains some misunderstandings about linearization, which is quite straightforward and doesn’t need to be precise for high quality MTF measurements.

Figure (6) shows the average edge response and MTF curve for a high-contrast slanted-edge from the same unsharpened image used for Figure (4).

![Figure 6. High contrast slanted-edge for Figure 3 (unsharpened)](image)

The average edge (on top) shows some flattening due to saturation. The MTF curve and the MTF50 value of 1215 LW PH (0.195 C/P) are close to the wedge-based MTF curve in Figure (4), where MTF50 = 0.183 C/P. The low-contrast MTF curve had a similar shape with MTF50 = 0.212 C/P.

Figure (7), for the low-contrast slanted-edge from the same sharpened image used for Figure (5), clearly illustrates the effects of strong sharpening. The halo will be visible at large magnifications, though it may not be objectionable or even visible under typical viewing conditions. This is far from the most extreme sharpening we’ve seen. Because artifacts from excessive sharpening can degrade the image (and cause misidentification of features) while improving MTF10 and MTF50 metrics, slanted-edge MTF plots should be used in characterizing imaging systems to determine if the image has been excessively sharpened. A single MTF number from any test chart is not sufficient.

![Figure 7. Low-contrast slanted edge from Figure 3 (sharpened)](image)

Figure 8 illustrates results for a high-contrast slanted edge from the same image as Figure 7. Highlight saturation almost completely...
Summary and Recommendations
This paper has largely focused on MTF measurements derived from hyperbolic wedges and slanted-edges. Some summary comments are general and some are specific to ISO 16505.

General comments
1. MTF10 measured from wedges is not a robust indicator of imaging system performance (resolution limit). It is subject to large variations, and bad images can have high MTF10. The minimum of \([MTF10, f_{alias}, f_{eqy}]\) (Equation (9)) is a much better indicator.
2. Wedge-based MTF measurements will be improved if wedge contrast is reduced from ≥40:1 (typically the highest contrast the media will support) to around 10:1, where saturation and clipping are much less severe. Lower contrast wedges are more representative of real objects that need to be distinguished in practical situations.
3. For a complete characterization of an imaging system, sharpening overshoot artifacts must be measured. This is best done with a low contrast slanted-edge. 4:1 contrast, specified in ISO 12233:2014, is a good value.

ISO 16505 comments
1. Spatial frequency calculations in ISO 16505 can be complex and difficult to follow. Units are LW/PH (Line Widths per Picture Height), but the Picture Height is not generally the height (smaller dimension) of image under test: it is derived from a square image representing the size of the monitor within the image under test. Spatial frequency units (used for MTF10) also involve a mirror magnification and image aspect ratio, which can result in confusion because they are so different from standard units.
2. Where necessary, pass-fail thresholds and other specifications should be derived using the many parameters referenced in ISO 16505 \((W_{crop}, H_{crop}, W_{monitor}, H_{monitor}, \alpha_{monitor}, \alpha_{mirror}, V_{eye}, \text{etc.})\). Angular spatial frequency units (cycles/milliradian, cycles/degree, etc.) can be used for intermediate calculations, then converted to LW/PH in the actual display.
3. Figure 27 and Annex D.1 and E.3 of the standard contain a proposed test chart. In addition to reducing the wedge contrast, we would add a low contrast (4:1) slanted edge and remove the two slanted parallel bar charts, which require precise scaling, which is difficult to achieve in practice and poorly defined in the standard. They add nothing to properly done wedge and edge measurements.
4. Section 7.8.5 of the standard contains several issues. Figure 34 shows a high contrast slanted-edge chart, which would be highly susceptible to saturation, and hence is not useful for observing sharpening artifacts. A 4:1 edge contrast, in conformance with ISO 12233:2014, is preferred. The statement about 6 meter lens-to-test chart distance is confusing. When the distance is \(≥200 \times \text{the lens focal length} \) the image is effectively at infinity, and greater distances offer no advantage—but they require gigantic test charts. 5 degree edge rotation is not necessary, though 0, 45, and 90 degrees should be avoided; most MTF software applies a cosine correction for larger edge angles.
5. Annex E of the standard describes a procedure for correlating hyperbolic wedge (MTF) and slanted-edge (SFR) measurements. Given the many differences in how these methods respond to noise, nonlinearities, and other perturbations, we see little benefit in this procedure. We recommend making both wedge and slanted-edge measurements.

References

Author Biography
Norman Koren received his BA in physics from Brown University (1965) and his Masters in physics from Wayne State University (1969). He worked in the computer storage industry simulating digital magnetic recording systems and channels from 1967-2001. He founded Imatest LLC in 2003 to develop software to measure the quality of digital cameras and imaging systems
Robert C Sumner holds a BS in Cinema and Photography from Ithaca College and a MS in Electrical Engineering from Boston University. He joined Imatest as an Imaging Science Engineer in 2016.
Henry Koren helped with Imatest from the very beginning. He received a BS in Computer Science and Engineering from UC San Diego in 2009. He is the Director of Engineering, focusing on manufacturing testing, at Imatest.

Robert C Sumner holds a BS in Cinema and Photography from Ithaca College and a MS in Electrical Engineering from Boston University. He joined Imatest as an Imaging Science Engineer in 2016.
Henry Koren helped with Imatest from the very beginning. He received a BS in Computer Science and Engineering from UC San Diego in 2009. He is the Director of Engineering, focusing on manufacturing testing, at Imatest.