SANDIA REPORT

SAND2015-20779
Unlimited Release
Printed January 2015

Security camera resolution measurements: Horizontal TV lines versus modulation transfer function measurements

Gabriel C. Birch, John C. Griffin

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation,
a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's
National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@adonis.osti.gov
Online ordering:	http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.fedworld.gov
Online ordering:	http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2015-20779
Unlimited Release
Printed January 2015

Security camera resolution measurements: Horizontal TV lines versus modulation transfer function measurements

Gabriel C. Birch
Systems Advanced Concepts Engineering
Sandia National Laboratories
1515 Eubank Blvd. SE
Albuquerque, NM 87123
gcbirch@sandia.gov

John C. Griffin
Systems Advanced Concepts Engineering
Sandia National Laboratories
1515 Eubank Blvd. SE
Albuquerque, NM 87123
johgrif@sandia.gov

Abstract

The horizontal television lines (HTVL) metric has been the primary quantity used by division 6000 related to camera resolution for high consequence security systems. This document shows HTVL measurements are fundamentally insufficient as a metric to determine camera resolution, and propose a quantitative, standards based methodology by measuring the camera system modulation transfer function (MTF), the most common and accepted metric of resolution in the optical science community. Because HTVL calculations are easily misinterpreted or poorly defined, we present several scenarios in which HTVL is frequently reported, and discuss their problems. The MTF metric is discussed, and scenarios are presented with calculations showing the application of such a metric.

Contents

Figures

1 Introduction

The ability of a camera to resolve targets of interest must be quantified and tested to assure that a system meets specifications required by a physical security system. Typically, Sandia National Labs (SNL) Division 6000 has used horizontal television lines (HTVL) of resolution as the standard metric of camera resolution for high consequence security systems. However, there are significant issues with using HTVL as an objective, comparable metric of camera resolution. Specifically, HTVL requires the use of a human to subjectively determine where the transition occurs from resolvable lines to unresolvable lines, called the just resolvable contrast. In addition, HTVL measurements using pixels back projected onto targets do not account for degradations caused by aberrations in the optical system, and HTVL to pixel conversions are often done incorrectly due to the confusing nature of the HTVL metric.

To rectify these problems, we propose a metric called the modulation transfer function (MTF). The MTF is a measurement of an imaging system's resolving contrast as a function of object size (typically measured in line pairs per mm on the detector). The MTF metric is a quantitative, repeatable, and system agnostic metric that does not involve a human. The MTF, or its mathematical relatives such as the point spread function or optical transfer function, is a very common and widely accepted metric to quantify imaging system resolution in the optical science community.

Fundamental concepts used by both the HTVL metric and the MTF method are discussed. Next, several typical scenarios measuring or using HTVL are presented. Each of these scenarios include numerical examples of their use. Additionally, the positive and negative qualities of each method are discussed. Finally, the proposed MTF method is discussed.

2 HTVL definition

Horizontal television lines of resolution (HTVL) is a metric that has been used in an attempt to quantify the resolution of an imaging system. HTVL is an old metric, originating with analog televisions and display devices. In the CCTV community, HTVL was adopted to describe the resolving power of an analog camera.

The use of HTVL becomes more confusing when applied to pixelated sensors. One reason for this is the use of analog television terminology and applying these terms to digital sensors. To avoid these issues, we will begin by defining resolution as used by the HTVL metric.

Based on the fundamental source, "CCTV Networking and Digital Technology," by Damjanovski [3], the television resolution is defined differently for the vertical and horizontal directions. This is shown in figure 1.

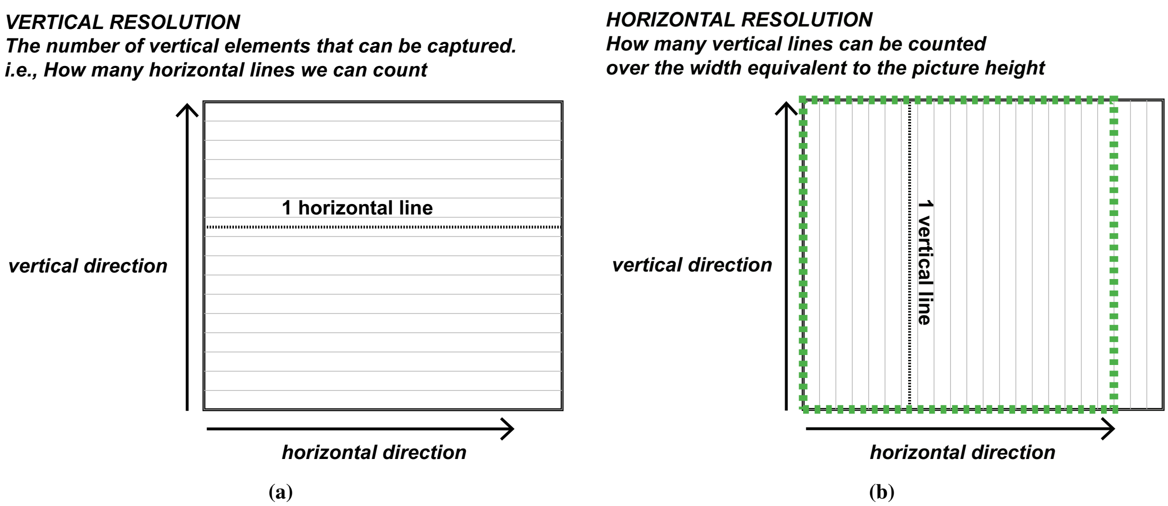

Figure 1. Definitions of (a) vertical television resolution and (b) horizontal television resolution. The green dashed box shows the width equivalent to the sensor height.

Vertical television resolution is defined by the number of vertical elements that can be captured by a camera or television. That is, the vertical resolution is equal to the number of horizontal lines distinguishable by a human viewing an image taken by the system. This yields our first fundamental concept, shown below.

Fundamental Concept 1: Vertical TV Resolution

The vertical TV resolution of a camera is equal to the number of horizontal lines distinguishable by a human viewing an image taken by the system. [3]

Misinterpretations can arise when using the terms "vertical lines" and "horizontal lines." For the vertical TV resolution, the number of *horizontal* lines are counted in the vertical dimension. In television systems, the number of horizontal lines was traditionally fixed at a standard resolution (e.g., NTSC 480i resolution). Measurements of vertical resolution for analog television quickly showed that the number of horizontal lines available to display information and the actual measured resolution were not the same [3]. Restated, televisions were not capable of displaying the Nyquist spatial frequency based on the number of display elements (i.e., even though a television specification sheet stated it was capable of displaying 200 cycles across the screen, measurements showed it could only display 140 cycles). This discrepency between theoretical television resolution and measured television resolution is known as the

Kell factor, and while it does not have a direct translation to CCTV cameras, it is an important factor necessary to frame the discussion in this text.

HTVL is defined by the number of horizontal elements that can be captured by a camera or television, with elements counted only over the width equivalent to the height. That is, the horizontal television resolution is equal to the number of vertical lines distinguishable by a human viewing an image taken by the system, and only counting resolvable lines over a dimension in the horizontal direction equal to that of the vertical direction. Again, misinterpretation can arise due to the counting of *vertical* lines when measuring the horizontal television resolution. This yields our second fundamental concept, shown below.

Fundamental Concept 2: Horizontal TV Resolution

The horizontal television resolution is equal to the number of vertical lines distinguishable by a human viewing an image taken by the system, and only counting resolvable lines over a dimension in the horizontal direction equal to that of the vertical direction. [3]

Therefore, for square elements where the horizontal and vertical spacings are equal, horizontal television resolution is also equal to the number of horizontal lines distinguished by a human across the entire vertical direction. This definition of HTVL is also described in another fundamental text, "Digital Video and HD: Algorithms and Interfaces," by Poynton [11]. Translation of what a vertical or horizontal line means when working with pixels in a sensor is discussed later in section 4.2.

3 Optical Concepts

The technical discussion to follow applies the well established terminology defined and used in the optical science communities. The imaging of an object involves three fundamental components: the object of interest, a lens, and an imaging sensor. The physical space and relationship between these is divided into image space and object space, as depicted in figure 2. Image space is defined as the space between the camera sensor and lens, while object space is defined as the space between the lens and the object being imaged. Images are formed at the sensor in image space. Sensors can also be back projected into object space to determine the pixels per target height or other metrics of interest. This calculation is performed using similar triangles, and is discussed in several example cases later in this text.

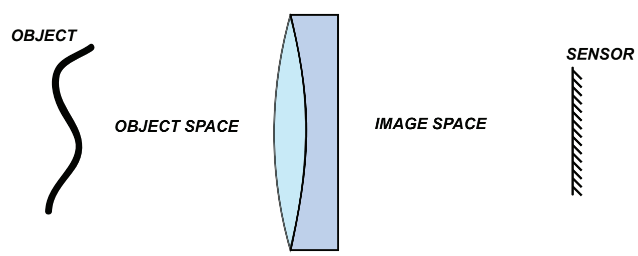

Figure 2. Diagram of object and image space, and their relation to a lens.

The ratio of a lens focal length and lens diameter is commonly called the F/# (F-number). This is shown in equation 1,

$$\mathrm{F/\#} = \frac{\mathrm{f}}{\mathrm{D}} \tag{1}$$

where f is the focal length of the lens, and D is the diameter of the pupil (commonly the diameter of the lens is used, as this is typically the entrance pupil).[1]

All lenses have a physical limit to the smallest spot they can produce on the sensor plane. This is a fundamental property of a lens, related to the wave like nature of photons, and the spot size cannot be reduced unless very exotic systems are implemented. This fundamental smallest image is called the diffraction limited spot size. The equation to determine the diameter of this diffraction limited spot, given a circular optic, is shown in equation 2,

$$\text{Diffraction limited spot size} = 2.44\lambda \text{F/\#} \tag{2}$$

where λ is the wavelength of the photon passing through the optical system and F/# is defined as shown in equation 1.

The diffraction limited spot size is important to understand when working with pixels in the 1-10 micrometer size. To see this directly, let us assume we have a lens with an F/# of 5, a wavelength of 0.5µm (blue-green light), and a pixel size of 5µm, a reasonable size for a security camera. Calculating the diffraction limited spot size of this lens system yields

$$\text{Diffraction limited spot size} = 2.44 \times 0.5\mu\text{m} \times 5 \tag{3}$$
$$\text{Diffraction limited spot size} = 6.1\mu\text{m} \tag{4}$$

Recall that the single pixel size in this example is 5µm. The diffraction limited spot size of this F/5 system is greater than the size of a single pixel in this sensor, and therefore the number of pixels in the camera does not provide any meaningful information regarding the resolving power of this system. This pixel size is not unreasonably small. Consider the rear facing camera on the iPhone5. This camera contains a Sony sensor (IMX145 derivative) with a pixel size of 1.4µm, nearly three times smaller that that of the previous example. If pixel size was used solely to calculate the resolution of this system, the predicted resolution would be much better than the measured resolution, since the lens of such a system is likely limiting the resolution, not the sensor.

This diffraction limited spot size is the smallest physical spot that can be formed by the lens. If the lens is imperfect it is said to contain optical aberration. Aberration increases the diameter of the minimum resolvable spot, and can further degrade the final image quality (and therefore degrades system resolution).

This yields our third and last fundamental concept, shown below.

Fundamental Concept 3: Diffraction Limited Spot Size

All lenses have a diffraction limited spot size, which is the smallest theoretical spot the lens can produce. This is defined as

$$\text{Diffraction limited spot size} = 2.44\lambda \text{F/\#}$$

where λ is the wavelength of the photon passing through the optical system and F/# is the ratio of the focal length to the diameter of the pupil. If the smallest theoretical spot is larger than a single pixel, the imaging system resolution is limited by the lens, not the image sensor. <u>Therefore, more pixels does not necessarily mean greater resolution.</u>

4 HTVL calculations

As described in section 2, HTVL is a metric assigned to a camera based on the number of light and dark vertical lines that can be resolved per picture height [3, 11]. This is shown in figure 3, where the example would have 12 HTVL of resolution (since there are 12 alternating black and white vertical lines that are resolvable over a span equal to that of the picture height).

[1] Strictly speaking, the F/# is a ratio that specifies the cone of light in *image space* for an object at infinity. Using this definition, the F/# is a ratio of the effective focal length over the diameter of the entrance pupil. It is possible that the largest physical element in the optical system is not the entrance pupil. When working in *object space* it is possible to scale the F/# by (1-m), where m is the magnification of the optical system.

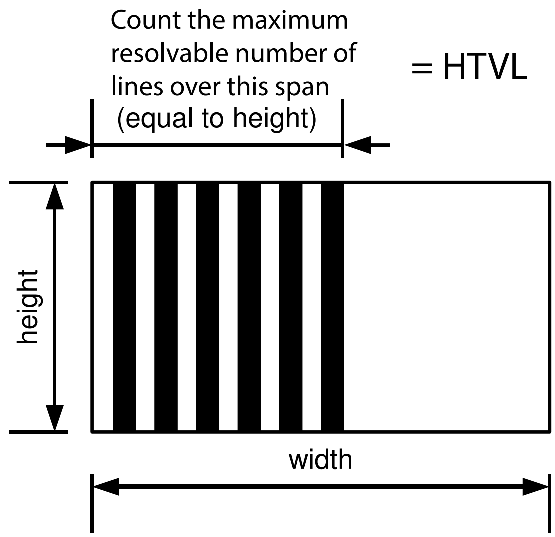

Figure 3. Diagram of the definition of HTVL.

HTVL can be measured or obtained several ways. First, it can be determined by measuring a standard test target. Second, HTVL can be inferred from the number of pixels on a sensor and back projected to a target. Finally, an HTVL per foot at a target of given distance can be calculated. This final application of HTVL is typically used to determine if a system meets resolution requirements. Note that HTVL and HTVL per foot are different quantities, with the major differences highlighted in the following three examples.

4.1 Case 1- Measuring system HTVL from a test target

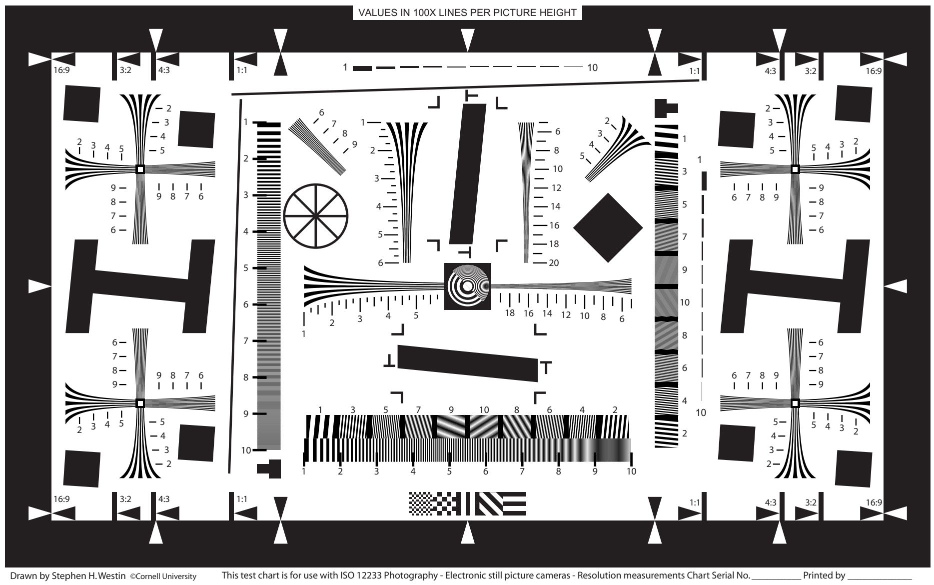

Figure 4. Camera resolution target from the ISO 12233 standard [13].

A frequently used technique to qualify a camera is to image a test target, such as the ISO 12233[2] target [5] shown in figure 4. This test positions the camera to frame the test target such that the target height fills the image height within the field of view of the camera. A human operator then views patterns on the target and determines when these bars have just resolvable contrast (i.e., just turn gray). The top of each chart has scaling information that can be used to determine the size of bars on the target. For example, in figure 4 the chart states "Values in 100x lines per picture height." Figure 5 shows how HTVL is calculated from this standard ISO resolution chart given the marker number (1 in the example) and the scaling information.

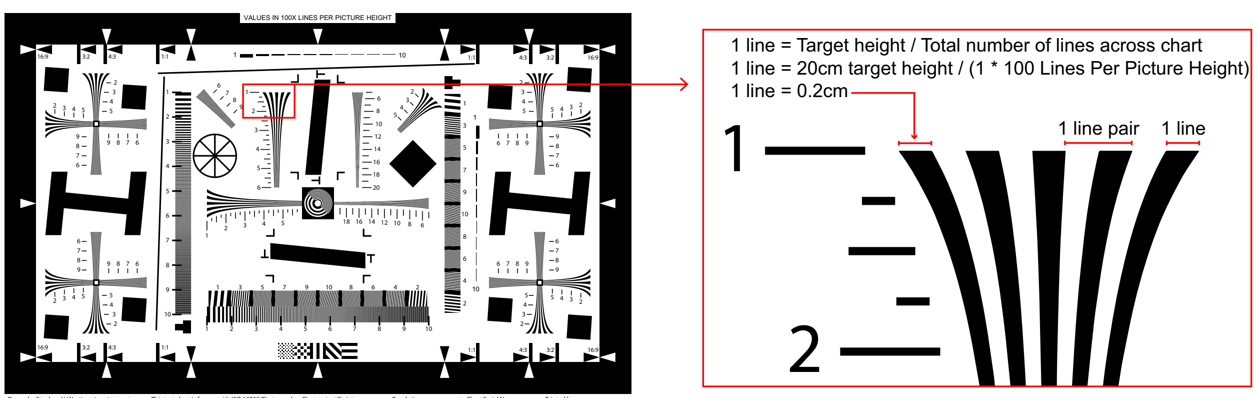

Figure 5. Method used to determine HTVL across the target height given the ISO resolution target.

[2]Note that the target shown in figure 4 is from an older version of the ISO 12233 standard. As of 2014, the ISO 12233 standard contains procedures for one qualitative test and two quantitative tests. The quantitative testing utilizes targets, including a sinusoidal Siemens star, to measure the MTF of the optical system as discussed in section 5.

Given a target height of 20cm and the information about values given in 100 lines per picture height, it is possible to determine both the size of the individual bars being measured and the HTVL resolvable by the system under test. It is important to note that this HTVL number is in $\frac{\text{lines}}{\text{Target Height}}$ and not in line *pairs*. Therefore, if a target loses all contrast at the 2 marker, the second major line shown on the right image in figure 5, the limiting resolution is measured as 200 lines across the target height, or restated as 100 line pairs across the target height. Similarly, if a target loses all contrast at the 8 marker, the limiting resolution is 800 lines across the target height or 400 line pairs across the target height.

This method accounts for the entire camera system: lens, sensor software, monitor displaying the image, and human eye visually assessing the location of just resolvable lines. However, there are significant issues with this test. Primarily, the determination of the location where lines become just resolvable is performed by a human. Humans are a statistical system, and the very best utilization of this test would present images of the test target to multiple humans and account for their statistical variation in determining the just resolvable line location.

Additionally, image processing and image enhancement inherently present on most modern digital cameras can yield spurious answers in terms of system resolution. Cameras typically perform significant enhancement to images such as de-noising, edge enhancement, and other contrast altering algorithms. Some cameras even specifically look for black and white edges and artificially enhance their contrast. It is important to realize that these changes do not alter the true resolution of the imaging system (i.e., no new information is gathered after the image is taken). Rather, *these modifications to the image are educated guesses* using algorithmic reconstruction of targets given fixed assumptions about scene content.

Practically, this level of image enhancement can affect a human's ability to determine the correct location of just resolvable lines, with the human often incorrectly specifying this location of zero contrast at higher lines per picture height than the camera is truly capable of imaging. Figure 6 shows three images captured from a Samsung 6004 camera with a stock lens viewing the ISO12233 resolution target (the chart shown in figure 4). This camera has the ability to turn sharpening on and off, and specify the level of sharpness if enabled. These images show the vertical wedge test target at varying levels of sharpness.

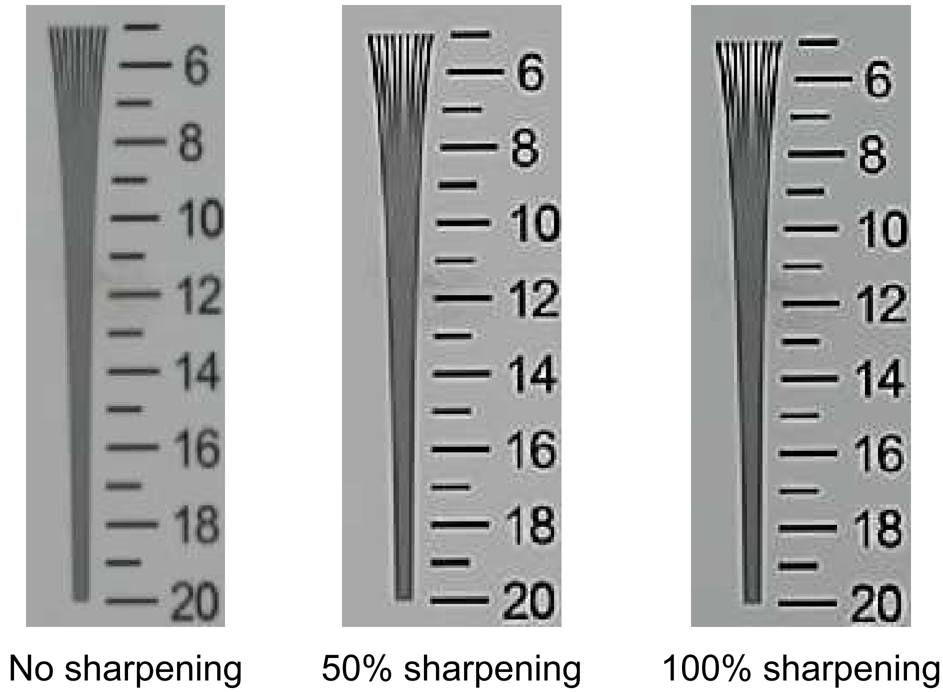

Figure 6. Three images of a target within the ISO12233 test chart imaged with the Samsung 6004 camera with varying degrees of sharpening applied by the camera. Note the difference in perceived 'just resolvable lines' between images.

The sharpened images could easily influence a human observer into thinking the just resolvable line aligns with the 800 lines per target height marker, rather than somewhere closer to the 700 lines per target height marker. It is important to realize this effect is a software toggle and does not truly represent a change in resolution of the imaging system.

4.2 Case 2- HTVL calculated from a data sheet and back projected pixels

Given a vendor data sheet for a camera, one can theoretically calculate the HTVL for a sensor[3]. The inputs needed for this calculation are the number of pixels in the horizontal and vertical dimensions of the sensor focal plane, and knowledge of the aspect ratio of the sensor focal plane. Recall that aspect ratio is defined as

$$\text{Aspect Ratio} = \frac{\text{Sensor Width}}{\text{Sensor Height}} \tag{5}$$

Figure 7 shows three sensors with different numbers of pixels and aspect ratios.

[3]Note that this is not strictly adhering to the definition of HTVL as noted in fundamental concept 2, but is a very common use of a vendor data sheet.

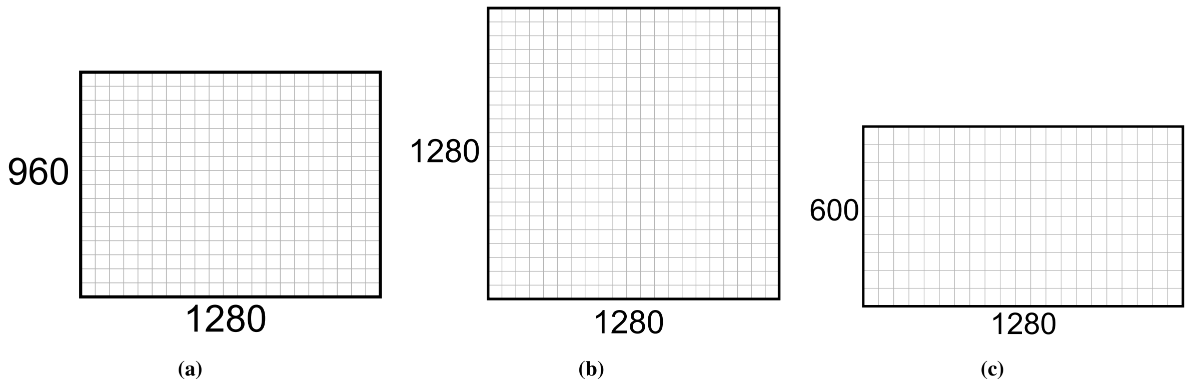

Figure 7. Three sensors with unique numbers of pixels in the horizontal and vertical dimensions and differing aspect ratios.

When calculating the HTVL of a given sensor the number of pixels in the horizontal dimension are taken as the total number of vertical lines (recall figure 1). Then, to calculate the HTVL, this number of vertical lines is multiplied by the the inverse aspect ratio. This is shown in equation 6.

$$\text{HTVL} = \text{Vertical lines (i.e., horizontal pixels)} \times \frac{\text{Sensor Height}}{\text{Sensor Width}} = \text{Vertical lines} \times \frac{1}{\text{Aspect Ratio}} \tag{6}$$

Therefore, using equation 6 we can find the HTVL for the sensors given in figure 7. For figure 7 (a) we have

$$\text{HTVL for figure 7 (a)} = 1280 \text{ vertical lines} \times \frac{3}{4} \tag{7}$$

$$\text{HTVL} = 960 \text{ HTVL} \tag{8}$$

For figure 7 (b) we have

$$\text{HTVL for figure 7 (b)} = 1280 \text{ vertical lines} \times \frac{1}{1} \tag{9}$$

$$\text{HTVL} = 1280 \text{ HTVL} \tag{10}$$

For figure 7 (c) we have

$$\text{HTVL for figure 7 (c)} = 1280 \text{ vertical lines} \times \frac{600}{1280} \tag{11}$$

$$\text{HTVL} = 600 \text{ HTVL} \tag{12}$$

Therefore, given the number of pixels in the vertical and horizontal direction, the maximum HTVL a sensor is capable of delivering is equal to the number of pixels in the vertical dimension. Because of this conversion of 1 pixel to 1 HTVL of resolution, back projection of the imaging system sensor pixels onto a target at some distance is frequently done to assure an adequate number of pixels are covering a specific target.

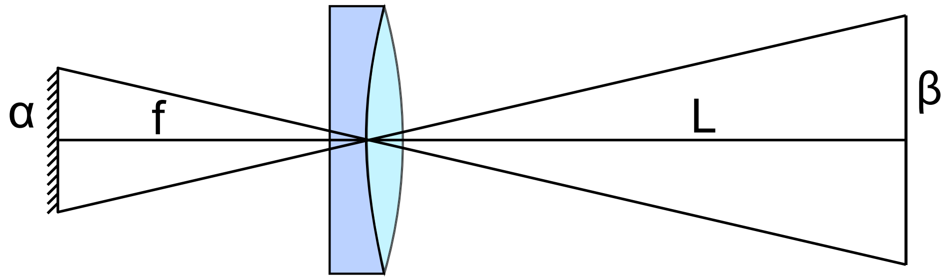

Figure 8. Diagram of similar triangles used to back project pixels to a target of interest at a given distance.

Figure 8 shows the similar triangle diagram used to back project pixels onto a target. Given the size of a pixel, α, the focal length of the lens, f, and the distance to the target, L, it is possible to calculate the back projected size of a pixel, β.

$$\frac{\alpha}{\mathrm{f}} = \frac{\beta}{\mathrm{L}} \tag{13}$$

$$\beta = \frac{\alpha \mathrm{L}}{\mathrm{f}} \tag{14}$$

For example, given a 5µm pixel, a 50mm focal length lens, and a target distance of 100ft, it is possible to determine pixel size on target given equation 14.

$$\frac{5\,\frac{\mu\mathrm{m}}{\mathrm{pixel}}\,\frac{10^{-3}\mathrm{mm}}{\mu\mathrm{m}}}{50\mathrm{mm}} = \frac{\beta}{100\mathrm{ft}\,\frac{12\mathrm{in}}{1\mathrm{ft}}\,\frac{25.4\mathrm{mm}}{1\mathrm{in}}} \tag{15}$$

$$\beta = \frac{0.005\,\frac{\mathrm{mm}}{\mathrm{pixel}} \times 30480\mathrm{mm}}{50\mathrm{mm}} \tag{16}$$

$$\beta = \frac{3.0480\mathrm{mm}}{\text{pixel on target at 100ft}} \tag{17}$$

$$\beta = \frac{3.0480\mathrm{mm}}{\text{pixel on target at 100ft}}\,\frac{1\mathrm{in}}{25.4\mathrm{mm}}\,\frac{1\mathrm{ft}}{12\mathrm{in}} \tag{18}$$

$$\beta = \frac{0.01\mathrm{ft}}{\mathrm{pixel}} \tag{19}$$

Taking the inverse of equation 19 yields the pixel per foot value commonly used to verify a camera meets specifications.

$$\frac{1}{\beta} = \frac{100\ \text{pixels}}{\mathrm{ft}} \tag{20}$$

There are several issues with using this approach to extract a resolution metric for the camera system. Most critically, this method assumes the camera system is limited by the number of pixels in the sensor. This means there is no account given to degradations caused by aberrations present in the lens. As shown in section 3 and captured in fundamental concept 3, it is entirely possible to have a physically perfect lens that still creates spots larger than the size of single pixels in a camera.

Additionally, recall the concept of the Kell factor. The Kell factor states that the real, measured resolution of a television system is some multiple less than one of the theoretical maximum display resolution. When an optical system is not sensor limited, resolution behaves just as the Kell factor would imply; the measured resolution in an image is less than the theoretical predicted resolution based on the number of pixels in the sensor. However, **there is no widely agreed upon use of the Kell factor in CCTV camera systems.** Some definitions of HTVL, specifically the Sandia National Labs division 6000 definition, do not account for this additional factor. Other sources recommend including the Kell factor in calculations to predict the 'real' resolution of the system [8, 10]. Regardless, no significant agreement exists.

Because the Sandia National Labs division 6000 definition does not include the Kell factor or a conceptual equivalent, all imaging systems are assumed to be limited by the sensor, not the optical system. Again, referring back to fundamental concept 3, this has been shown to be a poor assumption. Therefore, this method of calculating camera resolution offers little information on the actual resolving power of the imaging system.

4.3 Case 3- HTVL per foot

A common application of the HTVL metric is using the measured or calculated HTVL value and back projecting this value into object space. Typically, this is reported as an HTVL per foot calculation, and used to verify that a system design meets customer requirements. However, there is significant confusion when using a back projected number of pixels on target and converting this value to HTVL per foot on target. In "The Design and Evaluation of Physical Protection Systems," by Mary Lynn Garcia, it is stated

> "Extensive testing at Sandia National Laboratories has shown that a minimum of 6 TV lines of horizontal resolution (8 pixels) is required to accurately classify a 1 ft target." [6]

Though several issues arise when attempting to derive this statement, the most immediate problem with asserting 6 HTVL equals 8 pixels is that this implies that the effective resolution of an image is equatable to pixel count. To the contrary, it is well established and discussed previously in this work (specifically in fundamental concept 3) that the final resolution of an imaging system is a function not only of number of pixels but also other components such as the lens, wavelength, and algorithms applied to the captured image. Therefore, the equivalence of 6 HTVL and 8 pixels for a 1 foot target does not accurately describe a resolution metric.

However, even with the incorrect assumption that the pixel count per target length accurately describes the resolution, the equivalence of 6 HTVL and 8 pixels is not clear. To highlight this, we begin by asserting the quote from Garcia is either stating

$$\frac{4}{3}\text{ Aspect Ratio} \times \frac{6\text{ HTVL}}{1\text{ foot target }\textbf{width}} = \frac{8\text{ horizontal pixels}}{1\text{ foot target }\textbf{width}} \tag{21}$$

or

$$\frac{4}{3}\text{ Aspect Ratio} \times \frac{6\text{ HTVL}}{1\text{ foot target }\textbf{height}} = \frac{8\text{ vertical pixels}}{1\text{ foot target }\textbf{height}} \tag{22}$$

Equations 21 and 22 assumes several conditions such as square pixels, a single line is equivalent to a pixel, and an aspect ratio of 4:3 (a ratio typical of cameras produced at the time of publication of the Garcia book). Since the directionality of the 1 foot target (i.e., either vertical or horizontal) is not specifically stated by Garcia, we must assume these are the two possible scenarios being discussed. These two scenarios are shown in figure 9.

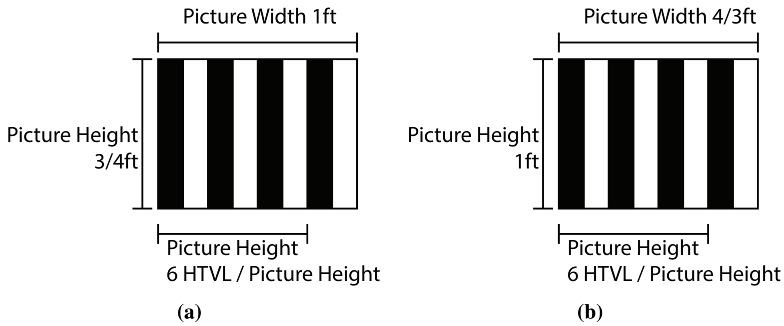

Figure 9. Two scenarios possible from the statement made in "The Design and Evaluation of Physical Protection Systems," by Mary Lynn Garcia regarding resolution and HTVL per foot relationships. Figure (a) corresponds to equation 21, while figure (b) corresponds to equation 22.

Given these conditions, an obvious problem arises; we are unable to specify a conversion between HTVL and pixels per foot on target if the direction of the target distance (i.e., vertical or horizontal) is not given. Restated, depending on the orientation of the target it would appear that there are either 6 pixels per 3/4 foot, or 6 pixels per 1 foot. Though this is only a 2 pixel per foot difference, it is a 33% difference in required resolution as specified by Garcia for accurate classification.

This problem of directionality highlights a significant error present when using HTVL as a resolution metric; projection of HTVL per foot implies a measurement made in object space, and therefore **is independent of the aspect ratio of the sensor used to perform the measurement.** Without specifying the directionality of the target dimension, two interpretations of pixels on target exist and differ by 33%. This factor is significant, and has immediate consequences on the number or type of cameras recommended for physical security systems.

5 Modulation transfer function calculations

We propose a more complete metric for understanding the resolution of an imaging system, the modulation transfer function (MTF). The MTF is a measurement of a system's image contrast as a function of spatial frequency (i.e., object size), and is a fundamental property of an imaging system. It is a vector as compared to a scalar value such as HTVL.

Technically, the MTF is the modulus of the optical transfer function (OTF), a complex quantity. The MTF can be measured a variety of ways, either using slanted bars, point sources, by directly sampling images of specialized test targets, or through other test patterns such as known noises targets[4]. In physical terms, the MTF can be viewed as a measure of how an optical system responds to a bar chart grating consisting of black and white line pairs of increasing spatial frequency (e.g., black and white line pairs that are progressively placed more closely together). The MTF is a useful metric, as it is an unbiased, system agnostic metric calculated without the input from a human observer.

Figure 10 shows a typical MTF measurement for an optical system.

[4]The MTF, OTF, pupil size, and image of a point source (called the point spread function), are all related. Deeper understanding of how these quantities can be determined from one another are discussed in several fundamental sources [7, 9].

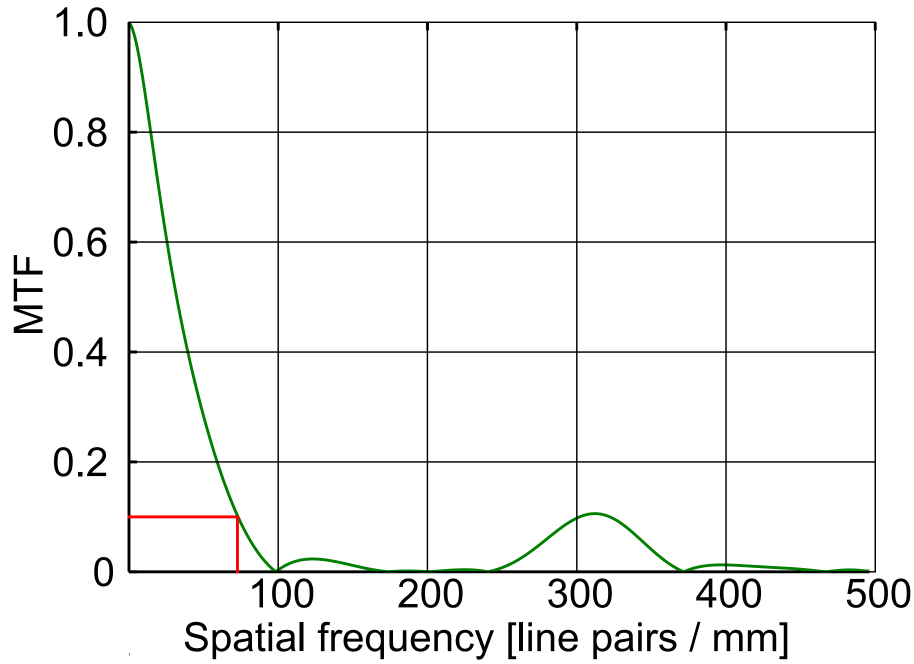

Figure 10. Plot of a measured modulation transfer function. The red lines show the location of the 10% contrast (i.e., 0.1 MTF value) and the corresponding spatial frequency value. Note that the units of line pairs per millimeter are in *image space* values (i.e., line pairs per millimeter at the detector).

The MTF vector provides significant information regarding the performance of an optical system. In figure 10 we can extract several parameters of interest. First is the cutoff frequency of the system, or the point where contrast goes to 0%. In this figure we see several zero contrast points (one at 100 LP/mm, a zero region from 175 LP/mm to 250 LP/mm, the 375 LP/mm zero, and approximately all frequencies past 475 LP/mm). In general, the cutoff frequency of this system would be said to be located at the first zero location, the 100 LP/mm location.

Additional information is contained in the sharp discontinuities seen at the 100 LP/mm and 375 LP/mm locations. Any discontinuity at 0% contrast indicates that this is a *contrast inversion frequency*. Essentially, white targets will become black and black targets will become white at frequencies greater than this discontinuity[5]. This is rarely seen in commercial systems due to the inclusion of anti-aliasing low-pass filters to reduce this effect.

MTFs are useful because they are system agnostic. That is, the MTF curve from figure 10 can be directly compared to a second system MTF without any additional knowledge. No scale factors, compensation for differing optics, or other information is required to compare MTF vectors.

Measured MTFs can also be compared to the theoretical best MTF, called the diffraction limited MTF. This diffraction limited MTF is related to the diffraction limited spot size discussed in section 3. Any system with performance close to that of the diffraction limited system is said to be near diffraction limited, and is usually regarded as a well designed system. Figure 11 shows a lens MTF compared to the diffraction limited equivalent MTF for the same lens.

[5]This is a simplification of the true phenomenon. It is possible to have multiple contrast inversions at frequencies higher than the initial contrast inversion, and therefore multiple changes from black to white or white to black.

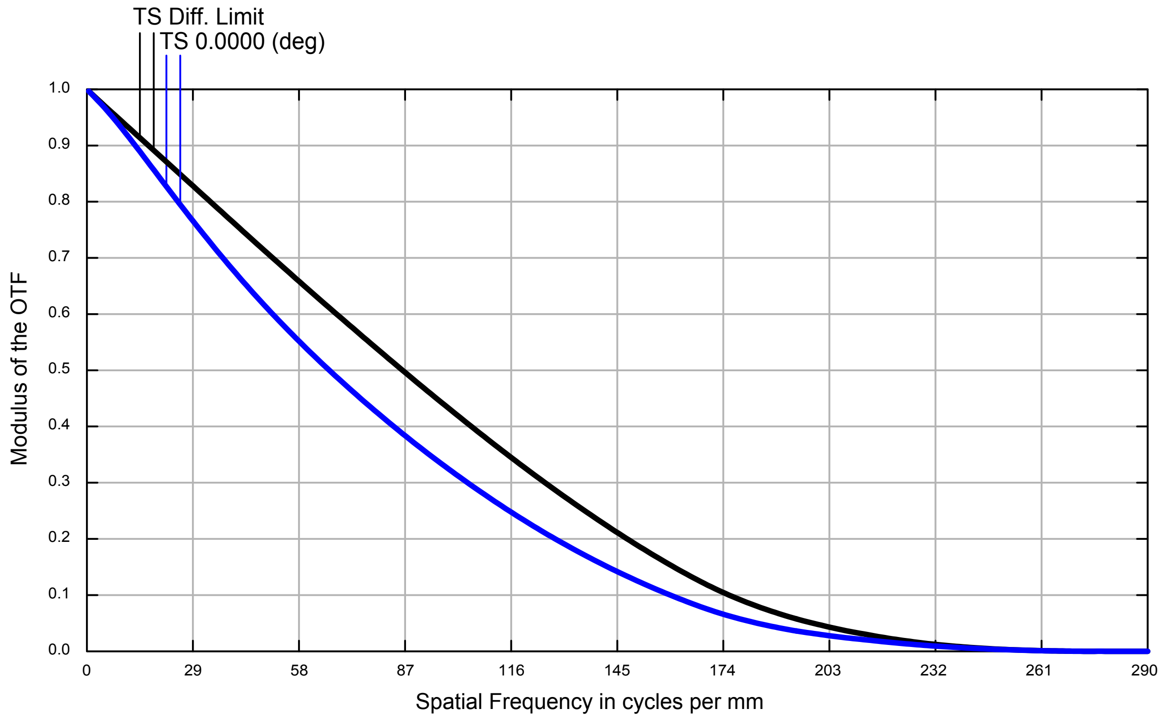

Figure 11. A predicted MTF from an optical design software. The blue line shows the MTF of a given optical system, while the black line shows the diffraction limited MTF, the theoretically perfect MTF this lens is capable of achieving given the input illumination, lens focal length, and entrance pupil diameter.

In this example the blue line is the MTF for the given lens and the black line is the diffraction limited MTF of the same lens. Note how the MTF is worse in the blue line; the mid range spatial frequencies are reduced compared to the diffraction limited MTF. Also note that the cutoff frequency is similar, found in both the diffraction limited and the real system at approximately 261 cycles per millimeter. Knowing the diffraction limited MTF of a system is useful, as this tells us the theoretically best MTF possible given wavelength, focal length, and diameter of the lens. To emphasize, it is not possible for traditional imaging systems to have an MTF greater than the diffraction limited MTF.

A useful value that can be extracted from the MTF is the spatial frequency at which contrast becomes 10%. This value of 10% contrast is often used as an approximate value where the human visual system has difficulty determining contrast. This 10% contrast metric is derived from the Rayleigh criterion and fundamental work performed by Lord Rayleigh examining the ability to resolve two targets closely spaced together using incoherent illumination or self-illumination [2]. The two overlapping targets used in the Rayleigh criterion example are shown in figure 12.

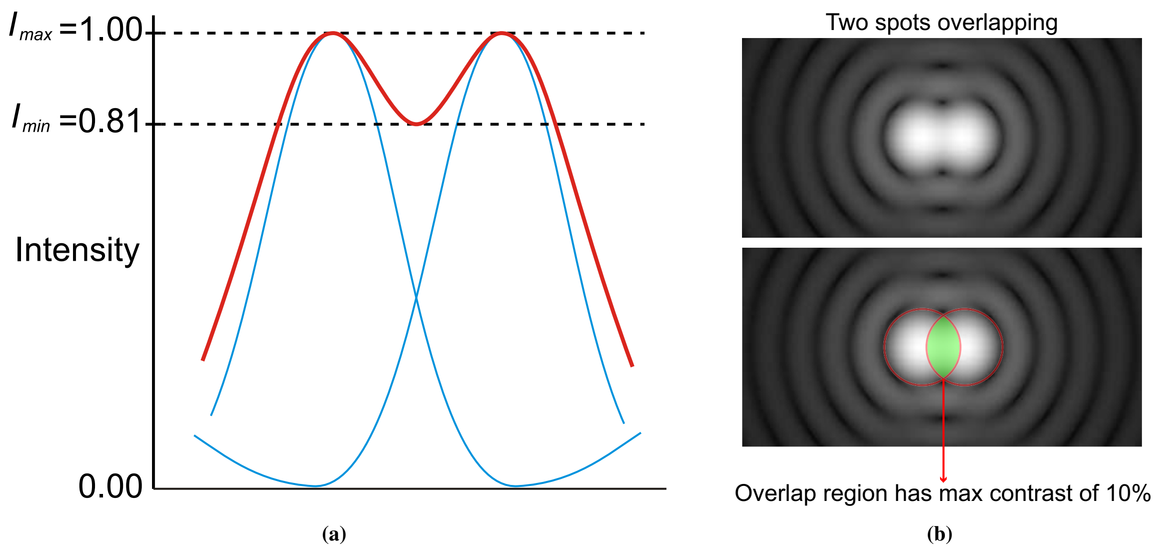

Figure 12. Overlap of two points and the summed intensity, where (a) shows intensity profiles of two points overlapping (blue lines) and their summed intensity (red line), and (b) shows the same phenomenon in an image. Given this spacing of two points, the resulting contrast of 10% was found to be the just resolvable contrast value by Lord Rayleigh.

Specifically, contrast (i.e., MTF value or modulation percent) of these just separable spots is

$$\text{Contrast} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{1.0 - 0.81}{1.0 + 0.81} \tag{23}$$

$$\text{Contrast} = 0.10 = 10\% \tag{24}$$

In the example shown in figure 10, the MTF10 value for this system is $75 \frac{\text{Line Pairs}}{\text{mm}}$. This is a value reported in image space, on the detector. Similar to the back projection of HTVL into object space discussed in section 4.3, it is possible to use the MTF10 value as a method of verifying a design meets requirements. Given the MTF10 value it is possible to use similar triangles to determine the line pairs per foot on a target at a given distance. Figure 13 shows the necessary parameters for this conversion.

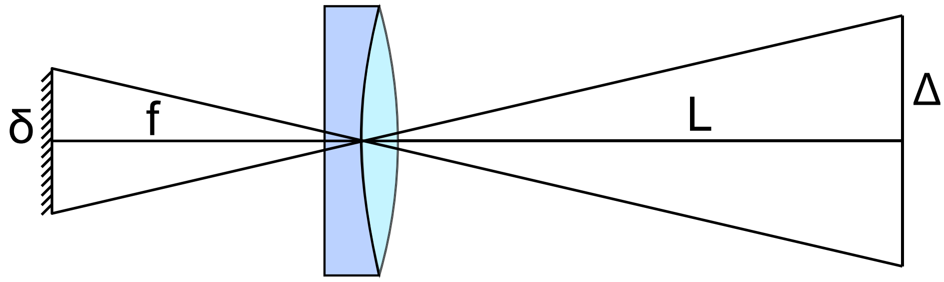

Figure 13. Diagram of an imaging system and the values needed to convert between MTF10 $\frac{\text{Line Pairs}}{\text{mm}}$ in image space and $\frac{\text{Line Pairs}}{\text{ft}}$ in object space.

We define δ as

$$\delta\left[\frac{\text{mm}}{\text{Line Pairs}}\right] = \frac{1}{\text{MTF10}} \tag{25}$$

Given the focal length of the imaging system, f, and the distance from the imaging system to the target of interest, L, we can use similar triangles to calculate Δ, the line pairs per distance at a given distance, L, shown in equation 26.

$$\frac{\delta}{\text{f}} = \frac{\Delta}{\text{L}} \tag{26}$$

Solving for Δ yields

$$\Delta = \frac{\delta \text{L}}{\text{f}} \tag{27}$$

Therefore, given an MTF value, focal length of the imaging system, and distance to target, it is possible to convert to an object space line pair per foot. It is important to note that this estimate is for a perfect system transmitting in a lossless medium, and should therefore be used to estimate performance in the best case. Further measurements of the MTF at non-optimal weather conditions should be performed to understand the performance space a camera system is capable of delivering under typical variations of the environment.

Given the MTF10 value of of $75\frac{\text{Line Pairs}}{\text{mm}}$, this converts to a δ value of $0.0133\frac{\text{mm}}{\text{Line Pairs}}$. For a 50mm focal length system imaging a target at 100ft from the camera, we find

$$\frac{0.0133\frac{\text{mm}}{\text{Line pair}}}{50\text{mm}} = \frac{\Delta}{100\text{ft}\frac{12\text{in}}{1\text{ft}}\frac{25.4\text{mm}}{1\text{in}}} \tag{28}$$

$$\Delta = \frac{0.0133\frac{\text{mm}}{\text{Line pair}} \times 30480\text{mm}}{50\text{mm}} \tag{29}$$

$$\Delta = \frac{8.1278\text{mm}}{\text{Line pair on target at 100ft}} \tag{30}$$

$$\Delta = \frac{8.1278\text{mm}}{\text{Line pair on target at 100ft}}\frac{1\text{in}}{25.4\text{mm}}\frac{1\text{ft}}{12\text{in}} \tag{31}$$

$$\Delta = \frac{0.0267\text{ft}}{\text{Line pair}} \tag{32}$$

Taking equation 32 and inverting it, we can determine the line pair per foot on a target at a given distance

$$\frac{1}{\Delta} = \frac{37.5 \text{ Line pairs}}{\text{ft}} \tag{33}$$

A system must have a minimum of 1 line pair (i.e., at least the ability to resolve one full bright and dark cycle) on a target to meet the Nyquist frequency, so this system would be well above the minimum required line pairs on target as required by the Nyquist sampling theorem.

Because MTF10 is system agnostic, human independent, and captures the resolution of the entire imaging system we propose the MTF10 value be used in place of the current HTVL metric. Using MTF10 per foot value in place of the HTVL per foot value is a significantly more robust and useful way to determine if enough resolving elements are back projected onto the target.

6 Conclusion and Implications

In this report we have shown several scenarios calculating and using HTVL. First, measurements of HTVL using wedge charts require human interpretation of just resolvable contrast locations, which can lead to subjective results of resolution. Second, pixels on target does not equate to resolution since it fails to account for the degrading effects of the lens if data is only used from the specification sheet. This is a non-trivial failure; real resolution can be substantially worse than the resolution predicted using pixels on target. Finally, HTVL per foot is an ambiguous measurement because the directionality of the distance (i.e., either vertical or horizontal) is meaningful. Ultimately, HTVL is no longer the appropriate method to use when quantifying camera resolution.

Use of a non-biased, system agnostic method, the modulation transfer function (MTF), is proposed. The MTF metric does not use a human to determine possible subjective features such as just resolvable contrast. Instead, MTF is measured via deterministic software. Using values of MTF10, it is possible to back project into object space and determine the number of line pairs per foot on a target at a given distance. This number is a true resolution metric, and can be used to compare systems or determine if a system meets requirements.

Utilizing MTF measurements for camera resolution is not only accepted by the optical science community, but also by the International Organization for Standardization (ISO). ISO-12233:2014 is the primary document that discusses camera resolution, and the 2014 edition of this standard outlines a sinusoidal Siemens star measurement to directly measure the MTF of a camera system. Moving towards a modern metric like the MTF, and specifically following the ISO-12233 standard, would enable SNL to perform quantitative, standards based testing accepted throughout the world by both international committee and optical scientists.

The ultimate goal of imaging system testing should be a complete characterization of the optics, sensor, electronics, network, monitor, and human to determine a detection metric. This concept of complete characterization is not new; research by the Night Vision Research Lab recommends this process when characterizing total system probability of detection [4, 12]. Additionally, this process is well established by the medical imaging community and the Food and Drug Administration; the entire imaging chain of the medical imaging process is of interest, from imaging device to performance of a radiologist, and can be used to determine which system is better suited for a given task [1]. Measuring the MTF of a camera is the first step in the process of complete characterization.

To summarize, by utilizing a quantitative, standards based metric such as the MTF, SNL will be able to leverage the significant efforts made in other fields, both in the U.S. government Department of Defense domain, as well as the commercial and academic domains, and move SNL on a course of cutting edge testing and qualification.

References

[1] H.H. Barrett and K.J. Myers. *Foundations of Image Science*. Wiley Series in Pure and Applied Optics. Wiley, 2013.

[2] M. Born, E. Wolf, A.B. Bhatia, P.C. Clemmow, D. Gabor, A.R. Stokes, A.M. Taylor, P.A. Wayman, and W.L. Wilcock. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 1999.

[3] V. Damjanovski. *CCTV: Networking and Digital Technology*. Elsevier/Butterworth Heinemann, 2005.

[4] Ronald G. Driggers, James S. Taylor, Jr., and Keith Krapels. Probability of identification cycle criterion (n50/n90) for underwater mine target acquisition. *Optical Engineering*, 46(3):033201–033201–6, 2007.

[5] International Organization for Standardization. Technical Committee Photography. *ISO-12233:2014 - Photography: Electronic Still-picture Cameras - Resolution Measurements*. ISO, 2014.

[6] M.L. Garcia. *Design and Evaluation of Physical Protection Systems*. The Design and Evaluation of Physical Protection Systems Series. Elsevier Science, 2007.

[7] J.W. Goodman. *Introduction to Fourier Optics*. Electromagnetics Series. McGraw-Hill, 1996.

[8] E.M. Harwood. *Digital CCTV: A Security Professional's Guide*. Elsevier Science, 2007.

[9] E. Hecht. *Optics*. Pearson Education, 2012.

[10] H. Kruegle. *CCTV Surveillance: Video Practices and Technology*. CCTV Surveillance Series. Elsevier Science, 2011.

[11] C.A. Poynton. *Digital Video and HD: Algorithms and Interfaces*. Morgan Kaufmann Series in Computer Graphics and Geometric Modeling. Morgan Kaufmann, 2012.

[12] Richard H Vollmerhausen and Eddie Jacobs. The targeting task performance (ttp) metric a new model for predicting target acquisition performance. Technical report, DTIC Document, 2004.

[13] Stephen H. Westin. ISO-12233 vector test chart, April 2010, `http://www.graphics.cornell.edu/~westin/misc/res-chart.html`.

Sandia National Laboratories