Analysis and Evaluation of Sampled Imaging Systems

Tutorial Texts Series

- *Nanotechnology: A Crash Course*, Raúl J. Martín-Palma, Akhlesh Lakhtakia, Vol. TT86
- *Direct-Detection LADAR Systems*, Richard D. Richmond, Stephen C. Cain, Vol. TT85
- *Optical Design: Applying the Fundamentals*, Max J. Riedl, Vol. TT84
- *Infrared Optics and Zoom Lenses*, Second Edition, Allen Mann, Vol. TT83
- *Optical Engineering Fundamentals, Second Edition*, Bruce H. Walker, Vol. TT82
- *Fundamentals of Polarimetric Remote Sensing*, John Schott, Vol. TT81
- *The Design of Plastic Optical Systems*, Michael P. Schaub, Vol. TT80
- *Radiation Thermometry: Fundamentals and Applications in the Petrochemical Industry*, Peter Saunders, Vol. TT78
- *Matrix Methods for Optical Layout*, Gerhard Kloos, Vol. TT77
- *Fundamentals of Infrared Detector Materials*, Michael A. Kinch, Vol. TT76
- *Practical Applications of Infrared Thermal Sensing and Imaging Equipment, Third Edition*, Herbert Kaplan, Vol. TT75
- *Bioluminescence for Food and Environmental Microbiological Safety*, Lubov Y. Brovko, Vol. TT74
- *Introduction to Image Stabilization*, Scott W. Teare, Sergio R. Restaino, Vol. TT73
- *Logic-based Nonlinear Image Processing*, Stephen Marshall, Vol. TT72
- *The Physics and Engineering of Solid State Lasers*, Yehoshua Kalisky, Vol. TT71
- *Thermal Infrared Characterization of Ground Targets and Backgrounds, Second Edition*, Pieter A. Jacobs, Vol. TT70
- *Introduction to Confocal Fluorescence Microscopy*, Michiel Müller, Vol. TT69
- *Artificial Neural Networks: An Introduction*, Kevin L. Priddy and Paul E. Keller, Vol. TT68
- *Basics of Code Division Multiple Access (CDMA)*, Raghuveer Rao and Sohail Dianat, Vol. TT67
- *Optical Imaging in Projection Microlithography*, Alfred Kwok-Kit Wong, Vol. TT66
- *Metrics for High-Quality Specular Surfaces*, Lionel R. Baker, Vol. TT65
- *Field Mathematics for Electromagnetics, Photonics, and Materials Science*, Bernard Maxum, Vol. TT64
- *High-Fidelity Medical Imaging Displays*, Aldo Badano, Michael J. Flynn, and Jerzy Kanicki, Vol. TT63
- *Diffractive Optics–Design, Fabrication, and Test*, Donald C. O'Shea, Thomas J. Suleski, Alan D. Kathman, and Dennis W. Prather, Vol. TT62
- *Fourier-Transform Spectroscopy Instrumentation Engineering*, Vidi Saptari, Vol. TT61
- *The Power- and Energy-Handling Capability of Optical Materials, Components, and Systems*, Roger M. Wood, Vol. TT60
- *Hands-on Morphological Image Processing*, Edward R. Dougherty, Roberto A. Lotufo, Vol. TT59
- *Integrated Optomechanical Analysis*, Keith B. Doyle, Victor L. Genberg, Gregory J. Michels, Vol. TT58
- *Thin-Film Design: Modulated Thickness and Other Stopband Design Methods*, Bruce Perilloux, Vol. TT57
- *Optische Grundlagen für Infrarotsysteme*, Max J. Riedl, Vol. TT56
- *An Engineering Introduction to Biotechnology*, J. Patrick Fitch, Vol. TT55
- *Image Performance in CRT Displays*, Kenneth Compton, Vol. TT54
- *Introduction to Laser Diode-Pumped Solid State Lasers*, Richard Scheps, Vol. TT53
- *Modulation Transfer Function in Optical and Electro-Optical Systems*, Glenn D. Boreman, Vol. TT52
- *Uncooled Thermal Imaging Arrays, Systems, and Applications*, Paul W. Kruse, Vol. TT51
- *Fundamentals of Antennas*, Christos G. Christodoulou and Parveen Wahid, Vol. TT50
- *Basics of Spectroscopy*, David W. Ball, Vol. TT49
- *Optical Design Fundamentals for Infrared Systems, Second Edition*, Max J. Riedl, Vol. TT48
- *Resolution Enhancement Techniques in Optical Lithography*, Alfred Kwok-Kit Wong, Vol. TT47
- *Copper Interconnect Technology*, Christoph Steinbrüchel and Barry L. Chin, Vol. TT46
- *Optical Design for Visual Systems*, Bruce H. Walker, Vol. TT45
- *Fundamentals of Contamination Control*, Alan C. Tribble, Vol. TT44
- *Evolutionary Computation: Principles and Practice for Signal Processing*, David Fogel, Vol. TT43
- *Infrared Optics and Zoom Lenses*, Allen Mann, Vol. TT42

Analysis and Evaluation of Sampled Imaging Systems

Richard H. Vollmerhausen
Donald A. Reago, Jr.
Ronald G. Driggers

Tutorial Texts in Optical Engineering
Volume TT87

SPIE PRESS
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Vollmerhausen, Richard H.
Analysis and evaluation of sampled imaging systems / Richard H. Vollmerhausen, Donald Reago, and Ronald G. Driggers.
p. cm. -- (Tutorial texts in optical engineering ; v. TT87)
Includes bibliographical references and index.
ISBN 978-0-8194-8077-4 (alk. paper)
1. Imaging systems--Image quality. 2. Image processing--Statistical methods. 3. Fourier analysis. 4. Sampling (Statistics) I. Reago, Donald. II. Driggers, Ronald G. III. Title.
TK8315.V6495 2010
621.36'7--dc22

2009053629

Published by

SPIE
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: +1 360.676.3290
Fax: +1 360.647.1445
Email: books@spie.org
Web: http://spie.org

The content of this book and CD-ROM reflects the work and thought of the author(s). Every effort has been made to publish reliable and accurate information herein, but the publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

The authors make no warranties that the software programs contained on the enclosed CD-ROM are free of error, or are consistent with any standard of merchantability, or will meet the requirements for any particular application. CD-ROM Copyright © 2010 RHV Consulting, LLC, 760 Jacktown Rd., Lexington, VA 24450.

Printed in the United States of America.

Introduction to the Series

Since its inception in 1989, the Tutorial Texts (TT) series has grown to more than 85 titles covering many diverse fields of science and engineering. The initial idea for the series was to make material presented in SPIE short courses available to those who could not attend and to provide a reference text for those who could. Thus, many of the texts in this series are generated by augmenting course notes with descriptive text that further illuminates the subject. In this way, the TT becomes an excellent stand-alone reference that finds a much wider audience than only short course attendees.

Tutorial Texts have grown in popularity and in the scope of material covered since 1989. They no longer necessarily stem from short courses; rather, they are often generated by experts in the field. They are popular because they provide a ready reference to those wishing to learn about emerging technologies or the latest information within their field. The topics within the series have grown from the initial areas of geometrical optics, optical detectors, and image processing to include the emerging fields of nanotechnology, biomedical optics, fiber optics, and laser technologies. Authors contributing to the TT series are instructed to provide introductory material so that those new to the field may use the book as a starting point to get a basic grasp of the material. It is hoped that some readers may develop sufficient interest to take a short course by the author or pursue further research in more advanced books to delve deeper into the subject.

The books in this series are distinguished from other technical monographs and textbooks in the way in which the material is presented. In keeping with the tutorial nature of the series, there is an emphasis on the use of graphical and illustrative material to better elucidate basic and advanced concepts. There is also heavy use of tabular reference data and numerous examples to further explain the concepts presented. The publishing time for the books is kept to a minimum so that the books will be as timely and up-to-date as possible. Furthermore, these introductory books are competitively priced compared to more traditional books on the same subject.

When a proposal for a text is received, each proposal is evaluated to determine the relevance of the proposed topic. This initial reviewing process has been very helpful to authors in identifying, early in the writing process, the need for additional material or other changes in approach that would serve to strengthen the text. Once a manuscript is completed, it is peer reviewed to ensure that chapters communicate accurately the essential ingredients of the science and technologies under discussion.

It is my goal to maintain the style and quality of books in the series and to further expand the topic areas to include new emerging fields as they become of interest to our reading audience.

James A. Harrington
Rutgers University

Contents

Preface

This tutorial is written for the design engineer or system analyst interested in quantifying the performance of electro-optical imagers. Advancing technology in detector arrays, flat panel displays, and digital image processing provide new opportunities to expand imaging applications and enhance system performance. Technical managers and design engineers are faced with evaluating the cost, weight, and performance of an ever-expanding selection of technology options. This book provides the theory and procedures for performance assessment.

This text supersedes *Analysis of Sampled Imaging Systems*, which was published by SPIE Press in 2000. Part I updates the earlier work. Part II discusses performance evaluation of electro-optical imagers. Part III provides computer programs and up-to-date information on detector arrays, optics, and display options. This book provides the theory, procedures, and information needed to evaluate and compare the performance of available imaging technologies.

Our prior work *Analysis of Sampled Imaging Systems* focused on the mathematical formulism needed to analyze sampled imagery. That book described the sampled imager response (SIR) function. SIR quantified sampled imager aliasing as well as the system transfer response. Fourier transform theory was used to describe and quantify sampling artifacts such as display raster, blocky images, and the loss or alteration of image detail due to aliasing.

However, the metrics provided by the earlier book were “rules of thumb.” Sampled imager design rules were based on experience and experimentation. No theory existed to quantify the effect of aliasing on visual task performance. The earlier work provided guidance on how to optimize sampled imagers by minimizing aliasing. *Analysis of Sampled Imaging Systems* did not provide a procedure to quantify the impact of aliasing on performance.

In the intervening years since the first book, we have discovered that the effect of aliasing on targeting performance is predictable by treating aliasing as noise. This book presents a resolution metric that predicts the effect of imager blur, noise, and sampling on the probability of correctly identifying targets. This new publication includes quantitative procedures for evaluating target acquisition performance.

Part I of this book includes all of the pertinent material from *Analysis of Sampled Imaging Systems*. The first five chapters remain substantially the same as the previous work. These chapters introduce sampling concepts and describe the differences between shift-variant and shift-invariant systems. Chapter 2 on Fourier optics is extensively rewritten. The errors associated with assuming separability in Cartesian coordinates are discussed, and examples are provided. The blurs associated with vibration and electronic stabilization are described. In Chapter 3, additional examples are added to better describe the SIR function. The focus of Chapters 1 through 5 remains the same, however. These chapters provide the mathematical tools needed to analyze sampled imagers.

Part II of this new book includes Chapters 6 through 10. This new material describes electro-optical imager evaluation. Chapter 6 describes target identification experiments. These experiments quantify visual task performance. Chapter 7 describes a resolution metric that predicts the probability of identifying targets. Chapter 7 also discusses the relationship between imager resolution and field performance. Chapter 8 explains aliasing as noise theory. For some years, we have known how to predict the effect of noise on target acquisition. Aliasing as noise theory predicts the effect of aliasing on target acquisition. Chapters 9 and 10 provide details on analyzing thermal imagers and imagers of reflected light, respectively.

Part III provides computer programs that implement the theory. These programs calculate the resolution of thermal and reflected light imagers. The programs are used to evaluate expected target acquisition performance and to compare imagers and assess the benefit or penalty of design changes.

Information is also provided to help make realistic assessments of imager performance. The book discusses optical performance and provides the characteristics of typical, good, and ideal lens systems. The book also contains information on a variety of display formats and interfaces, as well as detailed information on available focal plane arrays (FPAs). The information is presented in written form and is also coupled to the computer programs.

Particular emphasis is placed on theory and practice for the wide variety of available infrared FPAs. Technologies represented include InSb, HgCdTe, QWIP, and uncooled thermal arrays. Information is provided on the quantum efficiency, blur, crosstalk, and noise characteristics of each technology. The detector and array dimensions of available FPAs are provided. The availability of current information on optics, display, and FPA subassemblies allows the model user to make quick and realistic performance assessments of electro-optical imager designs.

Richard H. Vollmerhausen
Donald A. Reago, Jr.
Ronald G. Driggers
February 2010

Acronyms and Abbreviations

2afc	two alternative forced choice
A	ampere
Å	angstrom
AAN	aliasing as noise
AR	antireflection
A/W	units (amperes/watt) for spectral current responsivity
BDI	buffered direction injection
BLIP	background-limited photoconductor
CCD	charge-coupled device
CMOS	complementary metal-oxide semiconductor
CRT	cathode ray tube
CSF	contrast sensitivity function
CTF	contrast threshold function
CTIA	capacitive transimpedance amplifier
DC	direct current
DI	direct injection
DLHJ	double-layer heterojunction
erf	error function
E field	electric field
E-stab	electronic stabilization
ezoom	electron zoom
FF	fill factor
FOV	field of view
FPA	focal plane array
FPN	fixed pattern noise
FWHM	full width at half maximum
HDTV	high-definition television
HgCdTe	mercury cadmium telluride
HUD	head-up display
IFOV	instantaneous field of view
InSb	indium antimonide
IR	infrared
LACE	local-area contrast enhancement
LCD	liquid crystal display
LOS	line of sight
LSI	linear and shift invariant

LWIR	long-wave infrared
MHz	megahertz
MKS	meter-kilogram-second
MRT	minimum resolvable temperature
MTF	modulation transfer function
MWIR	midwave infrared
NEP	noise equivalent power
NETD	noise equivalent temperature difference
NIR	near infrared
NUC	nonuniformity correction
OTF	optical transfer function
Pf	picofarad
PID	probability of correct identification
psf	point spread function
PSQ	square of the Pearson's coefficient
PV	photovoltaic
QE	quantum efficiency
QWIP	quantum well infrared photoconductor
RC	resistor-capacitor
rms	root mean square
ROIC	readout integrated circuit
RSS	square root of the sum of squares
SGR	sky-to-ground ratio
SIR	sampled imager response
SIT	system intensity transfer
SMAG	system magnification
S/N	signal-to-noise ratio
SOM	specific object model
SRH	Shockley-Read-Hall
SSD	signal spectral density
str	steradian
SWIR	short-wave infrared
TCR	temperature coefficient of resistance
tgt	target
TOD	triangle orientation discrimination
TTP	targeting task performance
UAV	unmanned aerial vehicle
VFOV	vertical field of view
W	watt
WFOV	wide field of view

Chapter 1
The Sampling Process

This chapter provides an introduction to several topics. First, the physical processes involved in sampling and displaying an image are discussed. In a staring sensor, the image is blurred by the optics, and then the detector array both blurs and samples the image. The detector samples are used to construct the displayed picture. The physical processes that occur in a sampled imager are described.

When analyzing sampled systems, it is convenient to conceptually separate the preblur, sampling, and postblur (display blur) attributes of the system. These three steps in the generic sampling process are described, and the importance of each step discussed.

Next, the system properties known as *linearity* and *shift invariance* are described. Systems that are both linear and shift invariant can be characterized by a transfer function. Circuits and optical systems, for example, are characterized by their modulation transfer functions (MTFs). This chapter describes the transfer function concept using an electrical low-pass filter as an example. For linear shift-invariant systems, the transfer function provides a quantitative way of characterizing system behavior.

Sampled systems are linear, so Fourier transform theory can be used to analyze sampled systems. However, sampled systems are not shift invariant. It is the lack of the shift-invariance property that differentiates sampled systems from other systems and makes sampled systems more difficult to analyze. The lack of shift invariance in a sampled system is illustrated. Also, the reasons that a sampled system cannot be assigned a transfer function are explained.

At the end of this chapter, three different mathematical ways of representing the sampling processes are presented. These three derivations correspond to three different physical views of the sampling process. The physical basis of the mathematical techniques used in Chapter 3 is described, and the role of the display in a sampled system is put into perspective. For simplicity, many of the examples in this chapter are one-dimensional, but the discussion and conclusions apply to two-dimensional imagery.

1.1 Description of a Sampled Imager

Figure 1.1 shows a camera imaging a scene and an observer viewing the sampled image on a display. The components of the camera are shown in Fig. 1.2. A lens images the scene onto a staring detector focal plane array (FPA). Diffraction and aberrations blur the image. In the figure, the image on the FPA is upside-down as well as blurred because the optical system inverts the image.

The detector array consists of a number of rows and columns of individual detectors, as shown in the inset in Fig. 1.2. Photodetection occurs over the active area of these individual detectors. The incoming light generates photo-electrons. Since each detector integrates photo-signal over a finite active area, the detector itself also blurs the image. Individual points in the optical image are summed together if they fall on the same detector, and this summing of adjacent points causes a blur.

The detector photocurrent is integrated (in a capacitor, charge well, or by some other mechanism) for a period of time. Periodically (generally, every sixtieth of a second in the U.S. or every fiftieth of a second in Europe), the resulting signal charge is read out, and the integrator is reset. The amount of charge from each detector depends directly on the intensity of light falling on that detector. The charge output from each detector represents a sample of the lens- and detector-blurred scene intensity. Note that, in the example shown in Fig. 1.2, the active detector area does not cover the entire FPA. This array does not have a 100% fill factor.

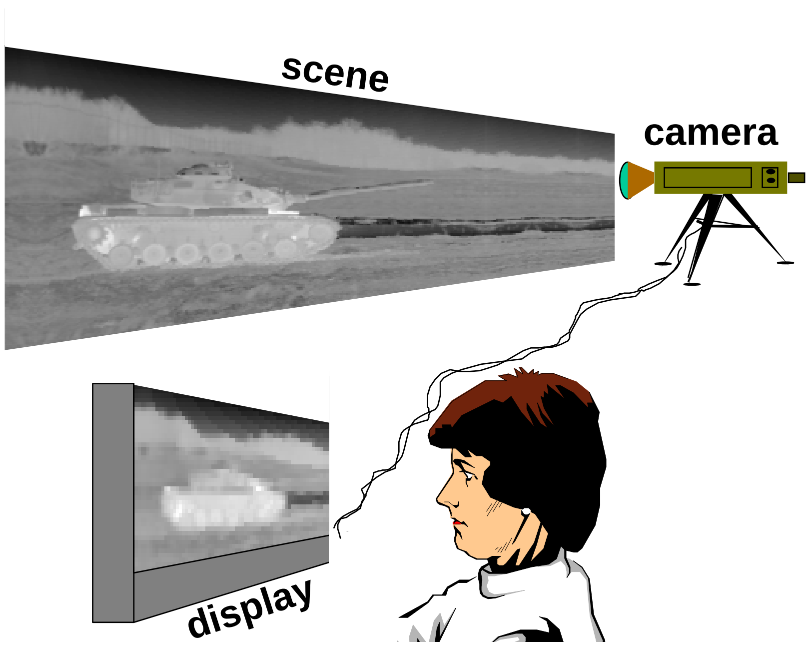


Figure 1.1 Observer viewing the display of a sampled image.

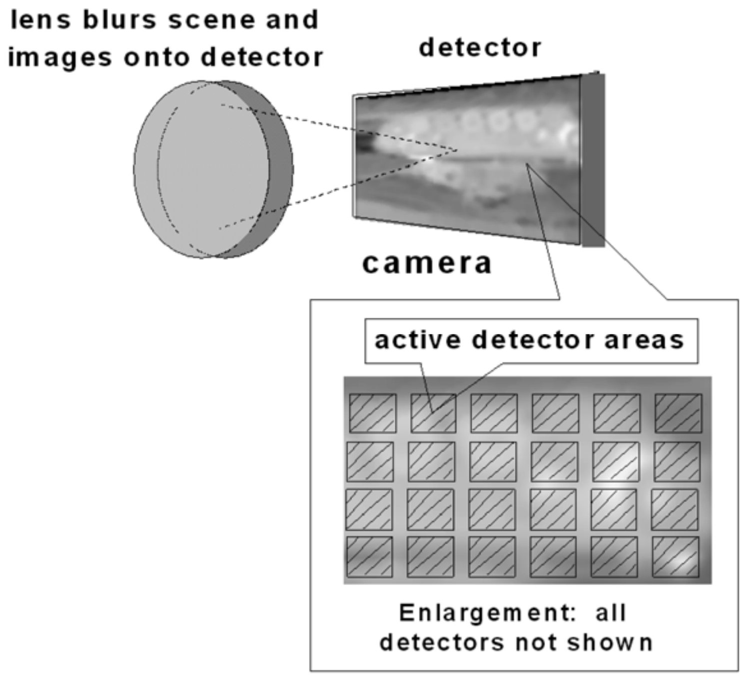


Figure 1.2 Components of a camera. The lens images a scene onto the detector array. Optical diffraction and aberrations in the lens blur the image. The detector further blurs the image by integrating signal over the active detector area. Each detector provides one sample of the blurred scene.

The two images in Fig. 1.3 summarize the action of the camera. Conceptually, the optical and detector blurs are lumped together and called the presample blur. The image with optical and detector presample blur applied is shown in the left-hand image in Fig. 1.3. The detectors then convert the light intensity at specific locations in the blurred image to electrical signals. The electrical signals represent image samples. In the right-hand image, the white dots indicate the locations where the blurred image is sampled by the detector array.

A display device is used to *reconstruct* the image from the detector (electrical) samples. The display device consists of an array of individual display pixels. A "display pixel" is an individual light-emitting area on the display surface. In the simplest case, the number and location of pixels in the display correspond to the number and location of detectors in the FPA. The brightness of each display pixel is proportional to the photo-signal from the corresponding detector.

Figure 1.4 shows the display. An individual display pixel is shown in the upper left-hand corner of the image. These pixels happen to be square. In this example, there is one display pixel for each sensor sample shown in Fig. 1.3. The intensity of each pixel shown in Fig. 1.4 is proportional to the photo-intensity at the corresponding sample location in Fig. 1.3.

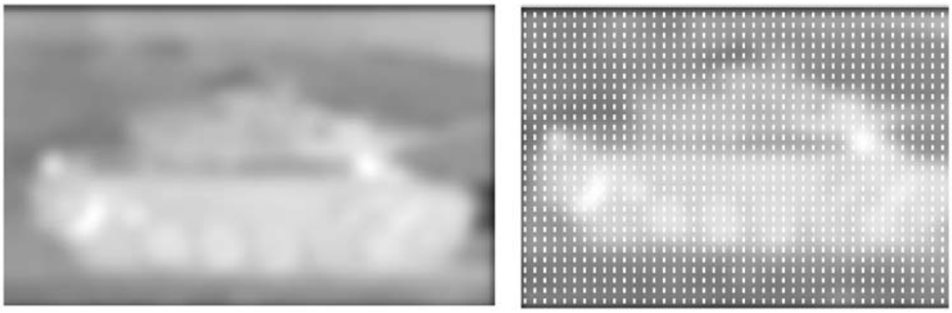


Figure 1.3 Left-hand image is blurred by the optics and the detector. Right-hand image shows the location of detector samples as white dots.

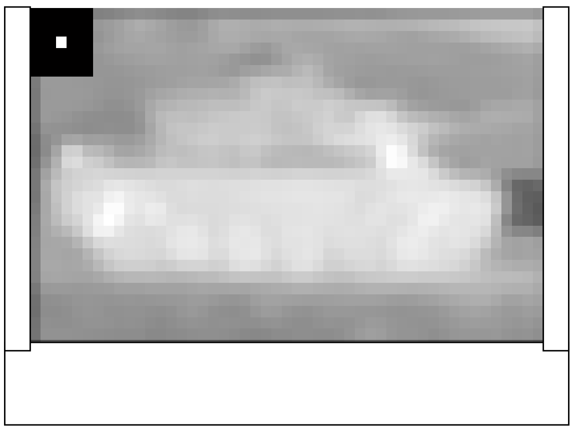


Figure 1.4 Display of a sampled image. An individual pixel is shown in the upper left-hand corner. Each pixel on the display is illuminated in proportion to the amplitude of the sample from the corresponding location in Fig. 1.3.

1.2 Description of the Sampling Process

Sampled imaging systems are characterized by a three-step process. First, the original image of Lena shown in Fig. 1.5(a) is degraded by a presample blur. The blur is caused by the combined effects of optical diffraction and aberrations, the finite size and shape of the detector, camera vibration, motion smear, and other effects. The presample blurred image is shown in Fig. 1.5(b). Next, the blurred image is sampled. That is, some electronic mechanism is used to find the amplitude of the blurred image at discrete points. In this example, the blurred image in Fig. 1.5(b) is sampled at the points shown as white dots in Fig. 1.5(c). The third step is reconstruction of the image. Each sensor sample controls the intensity of a display pixel (or a group of display pixels), as shown in Fig. 1.5(d). The shape and size (intensity distribution) of the display pixels determine the postsample blur. The shape and size of an individual display pixel are shown in the upper left-hand corner of Fig. 1.5(d). The postsample blur (reconstruction blur) also includes any postsampling electronic filtering and eye blur.

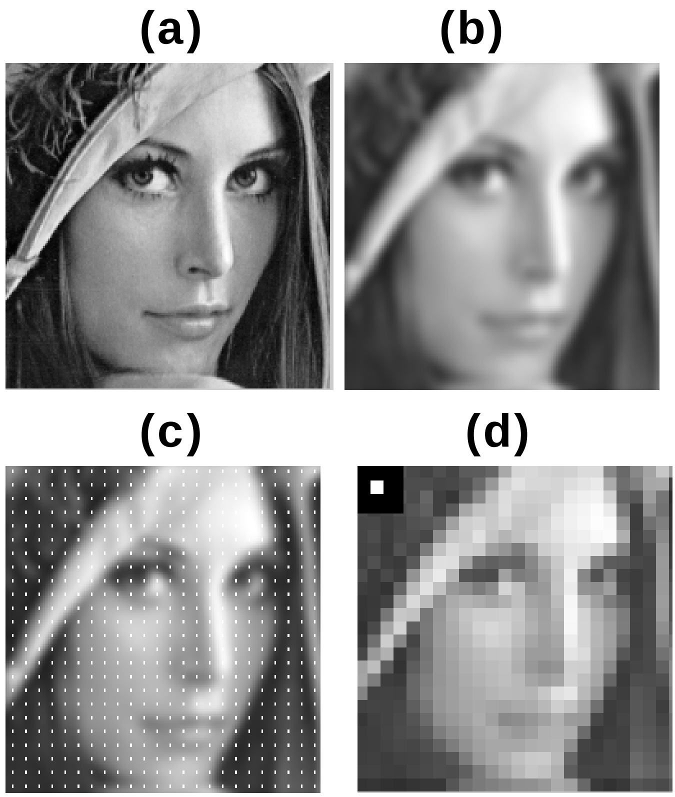


Figure 1.5 (a) Original picture. (b) Presample blur applied to original picture. (c) Picture showing locations of samples. (d) Picture showing reconstructed image with a single display pixel shown in the upper left-hand corner.

All three stages of the sampling process are necessary: preblur of the image, sampling, and postblur or reconstruction. Figure 1.6(a) shows the result of sampling the original image without prefiltering [that is, Fig. 1.5(a) is sampled rather than Fig. 1.5(b)]. In this case, aliasing hurts the final, displayed image. Figure 1.5(d) looks more like Fig. 1.5(a) than Fig. 1.6(a) does.

Figure 1.6(b) shows the image samples displayed as points rather than as the large pixels used in Fig. 1.5(d). In Fig. 1.6(b), the image is not blurred by the display pixel, and the image cannot be integrated by the eye. To get a good image, display reconstruction using a postsample blur is necessary.

As an illustration, move Fig. 1.6(b) close to the eye so that only points are seen. Now, slowly move the figure away from the eye. Lena begins to appear as the figure moves away because eye blur acts as a reconstruction filter.

Rules can be established for determining the optimum relationship between preblur, sample spacing, and postblur. A well-sampled imaging system is one in which the spacing (in milliradians or millimeters) between image samples is small compared to the width of the presample blur. In this case, sampling artifacts are not apparent in the image.

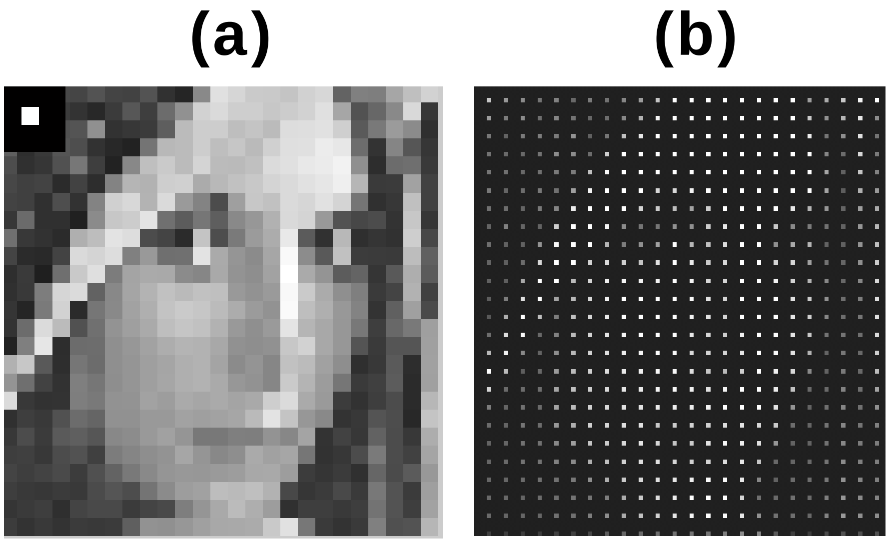


Figure 1.6 The picture in (a) shows the original image in Fig. 1.5(a) sampled without pre-blur and then reconstructed in a similar manner to Fig. 1.5(d). Figure 1.5(d) looks more like Fig. 1.5(a) than Fig. 1.6(a) does because some presample blurring is necessary. The picture in (b) is constructed with pixels that are spatially much smaller than the pixel pitch. The pixels in (b) do not blur the image, and the picture is very hard for the eye to integrate. Postblur of the image samples is also necessary.

An undersampled imaging system is one in which the sample spacing is a large fraction of the presample blur. Depending on scene content, the artifacts caused by undersampling can limit the performance of the imaging system. Poor sampling can corrupt the image by generating localized disturbances or artifacts. The corruption results in shifting object points, lines, and edges. Poor sampling can also modify the apparent width of an object or make a small object or detail disappear. That is, a fence post imaged by an undersampled sensor can be seen as thicker, thinner, or slightly misplaced.

An optimum postblur depends on the sample spacing and other factors. If the display pixel is large compared to the sample spacing, then the image will be blurred. If the display pixel is small compared to the sample spacing, then the shape of the pixel itself might be visible. Display pixels that are individually visible because of their size or intensity distribution will add spurious content to the image. The spurious response due to poor image reconstruction can seriously degrade sensor system utility.

Visible raster or a "pixelated" display makes it difficult for the observer to visually integrate the underlying sensor imagery. This is certainly true for the picture in Fig. 1.6(b), for example. However, the picture shown in Fig. 1.5(d) is also degraded by the sharply demarcated display pixels. The picture of Lena in Fig. 1.7 is generated using the same samples as were used to generate the picture in Fig. 1.5(d). The picture in Fig. 1.7 is improved because the display pixel shape provides a better match to the original image between sample points.

Figure 1.7 Picture of Lena reconstructed using the same samples as were used in Fig. 1.5(d). This picture is improved because the display pixel shape provides a better match to the original image between sample points.

A well-designed system is the result of a trade-off between sample rate or spacing, presample blur, and display reconstruction. Too much preblur can limit performance by degrading resolution while wasting detector count. Too little preblur can cause significant aliased content that limits imager performance. A display blur that is too small can introduce spurious content such as visible raster and pixelation effects that can ruin the displayed image. A display blur that is too large will limit performance by degrading resolution. Achieving good performance with a sampled imager requires trade-offs between the pre- and postsample blurs and the sample spacing.

1.3 Linearity and Shift Invariance

Systems that are both linear and shift invariant (LSI systems) can be analyzed in a very special way. A transfer function (or system response function) can be defined for an LSI system. The transfer function completely describes system behavior. For example, most circuits have a transfer response function that describes their electrical behavior. A well-corrected optical telescope has an optical transfer function that characterizes the image. LSI analysis is so common, and the ability to find a single function that completely describes a system is so typical, that it is ingrained in the modern engineering psyche. This section describes LSI theory and explains why it does not apply to sampled systems.

Systems are *linear* if superposition holds. That is, if

input A yields $\Rightarrow$ output A,

and

input B yields $\Rightarrow$ output B,

then

input A + input B yield $\Rightarrow$ output A + output B.

Some linear systems are *shift invariant*. *Shift invariance* means that if

input (t) yields ⇒ output (t),

then

input ($t - \tau$) yields ⇒ output ($t - \tau$).

In this example, t is time and τ is a fixed time offset. Shift invariance means that if an input is delayed τ seconds then the output is delayed τ seconds, but the shape of the output depends only on the shape of the input and does not change with the time delay. Sampled systems are not shift invariant. Note that the concept described here can be also applied to space and spatial offsets. In that case, t and τ can be replaced by x and x_o.

Another feature of LSI systems is that the output of an LSI system contains only the same frequencies as were input. This is not true for sampled systems.

A simple resistor-capacitor (RC) low-pass filter circuit is shown at the top of Fig. 1.8. An input sinusoid and the resulting output are shown at the bottom of the figure. This circuit is both linear and shift invariant. The output sine wave is the same frequency as the input sine wave. If two sine waves are input, the output will be the sum of the individual outputs. If the input sine wave is delayed τ seconds, then the output sine wave is delayed by τ seconds.

In Fig. 1.9, the output of the low-pass filter is sampled. The output is reconstructed using a sample and hold circuit. That is, the sample value is held constant for the entire sample period. At the bottom of Fig. 1.9, the sampled output is shown along with the input and the presampled output. [Note that the

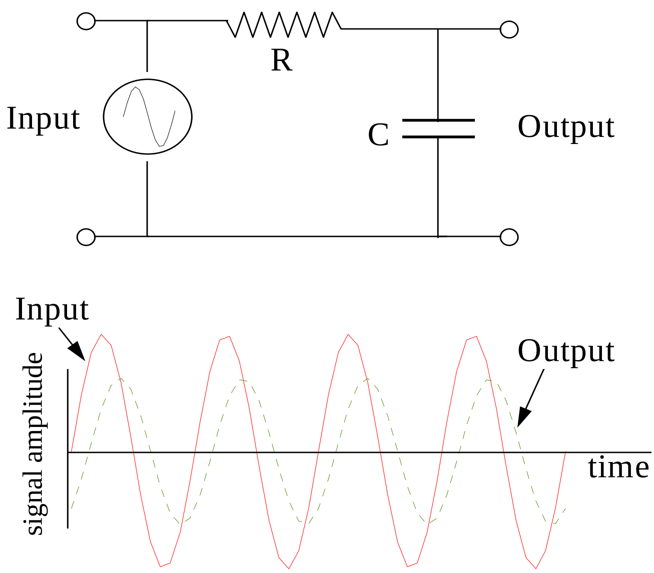


Figure 1.8 A low-pass filter is an example of a linear shift-invariant circuit. Output is at the same frequency as the input. If input shifts in time, then the output shifts in time by the same amount without changing shape or amplitude.

sampled output is shown shifted back in time (to the left) by one half of a sample period. In reality, the sampled output would be delayed (moved to the right in the figure) by one half of a sample period. The sampled output is shown shifted back in time in order to show how its shape matches the presample output.]

The sampled output shown in Fig. 1.9 is not a sine wave at the same frequency as the input. The irregularly shaped output contains many different frequencies. Also, the sampled system is not shift invariant. As illustrated in Fig. 1.10, as the input moves in time, the position of the samples on the waveform changes. The shape of the sampled output changes depending on how the samples line up with the input. The shape of the sampled output depends on when the input occurs. The sampled system is not shift invariant.

A transfer function can be defined for a system that is both linear and shift invariant. A sinusoidal steady state analysis completely defines the response of an LSI system. The response of an LSI system is defined by the relationship between the amplitudes A and B and the angles α and β for the system driving (forcing) function *dr* and response (output) function *re*.

$$dr(t) = A\cos(\omega t + \alpha) \tag{1.1}$$

$$re(t) = B\cos(\omega t + \beta), \tag{1.2}$$

where ω is 2π times the frequency in Hertz (Hz).

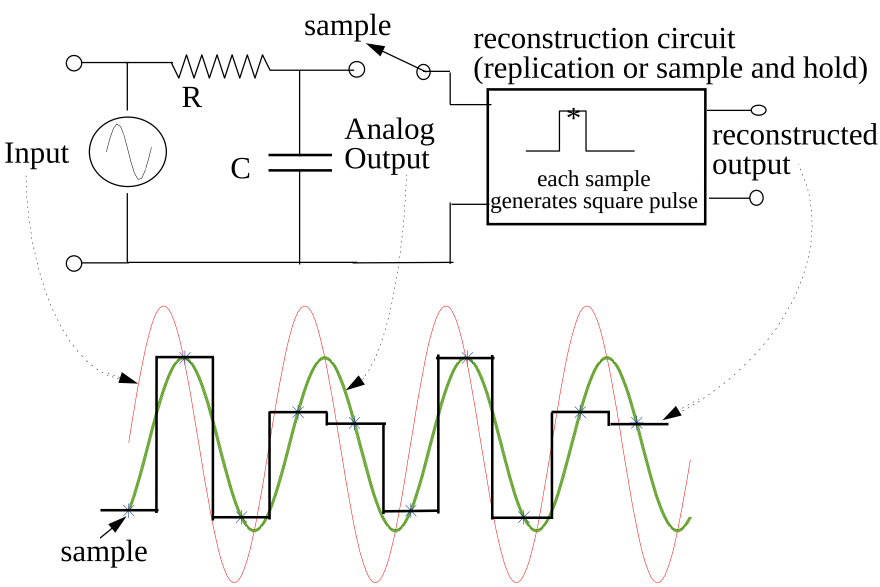


Figure 1.9 Low-pass filter circuit with sample and hold on the filter output. The output waveform has many frequencies that are not in the input spectrum.

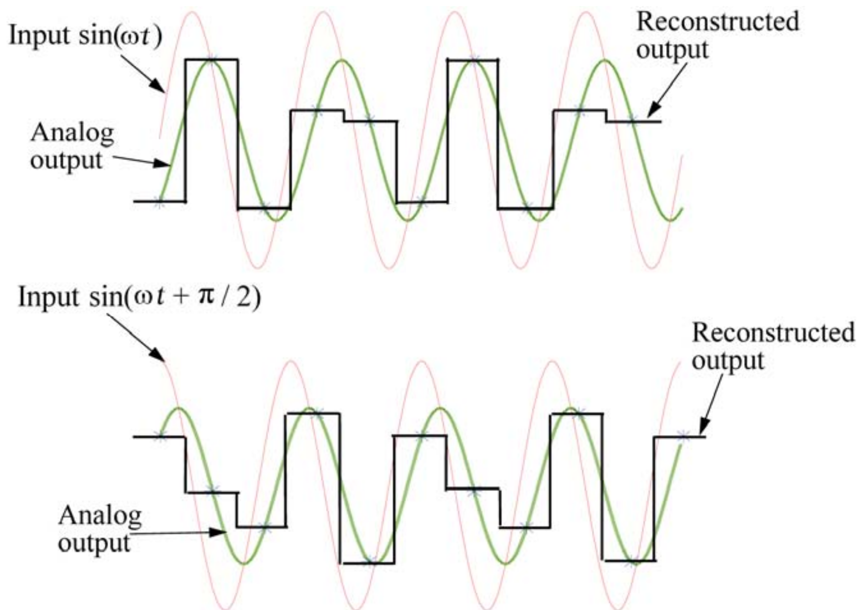


Figure 1.10 The waveform on top is sampled and reconstructed. At the bottom, the same input waveform is shifted and then sampled and reconstructed. The output waveforms on the top and bottom have different shapes. Sampling is not shift invariant. The shape of the reconstructed output changes depending on the phase (timing) of the input.

If an LSI system is driven by a sinusoid, then the output will be a sinusoid at the same frequency, but the output will generally have a different amplitude and phase. Furthermore, if the relationship between the amplitude and phase of $dr(t)$ and $re(t)$ can be found for every angular frequency ω, then that relationship completely describes both the periodic and aperiodic responses of the system.

Using complex notation, with j representing the square root of –1, the following functions represent the Fourier transforms of the driving and response functions:

$$Dr(\omega) = A(\omega)e^{j\alpha(\omega)} \tag{1.3}$$

$$Re(\omega) = B(\omega)e^{j\beta(\omega)}. \tag{1.4}$$

The transforms of the driving and response functions are related to each other by the system transfer function $H(\omega)$:

$$Dr(\omega) = H(\omega)\, Re(\omega), \tag{1.5}$$

where $H(\omega)$ is a complex function of the real variable ω. The system transfer function $H(\omega)$ completely characterizes the LSI system. Given an input $in(t)$, the output $o(t)$ is predictable from the equation

$$O(\omega) = H(\omega)\, In(\omega), \tag{1.6}$$

where $In(\omega)$ and $O(\omega)$ are the Fourier transforms of the input and output, respectively.

The sinusoidal steady state analysis described by Eqs. (1.1) through (1.6) is widely used in electrical circuit analysis and Fourier optics. For example, this analysis applies to the low-pass filter shown in Fig. 1.8. Given an input $In(\omega)$, the output $O(\omega)$ is described by

$$O(\omega) = \frac{1}{1 + j\omega RC} In(\omega) . \tag{1.7}$$

Another term used to describe LSI systems is *constant parameter*. A system that is described by a linear differential equation with constant coefficients will always be linear and shift invariant. For example, the low-pass filter is described by the following differential equation:

$$\frac{Q}{C} + R\frac{dQ}{dt} = in(t) \tag{1.8}$$

or, alternately,

$$\frac{1}{C}\int i dt + iR = in(t) , \tag{1.9}$$

where Q is charge, and i is current. In this example, both the resistance R and the capacitance C are constant so the circuit has constant parameters. The circuit is LSI.

Unfortunately, this kind of steady state analysis does not apply to sampled systems. As shown by the Fig. 1.9 example, a sine-wave input to a sampled system does not produce a sine-wave output at the same frequency. The frequency content of the output depends on the reconstruction method and the number of samples taken. Furthermore, as shown in Fig. 1.10, sampled systems are not shift invariant. The output depends on the relative position (or sample phase) of the samples compared to the input signal.

A transfer function such as $H(\omega)$ in Eq. (1.5) or (1.6) cannot be defined for a sampled system. However, as shown in the following chapters, a substitute for the system transfer response can be found for use with sampled systems.

1.4 Signal Reconstruction

The degree to which a sampled output matches the original presample signal depends on the signal reconstruction technique as well as on the sample rate. Figure 1.9 shows a sine wave reconstructed using a sample and hold. The sampled output is created by centering a rectangular pulse at each sample location. Each rectangular pulse is the width of a sample period, and the height of the pulse is the sample value. This type of signal reconstruction is called *replicated* because the sample value is used again and again.

Compare the sampled output in Fig. 1.9 to the sampled output in Fig. 1.11. In Fig. 1.11, linear interpolation is used. Each sample is connected by a straight line. The sample rate has not changed, but the output generated using linear interpolation appears to be a better representation of the input.

In order to make the subsequent example clearer, consider another way of implementing linear interpolation. This can be accomplished by centering a triangular pulse at each sample location. The pulse height is equal to the sample value, and the triangular half-width is equal to a sample period. Since adjacent triangles overlap, all triangle values at each time location are summed.

Now compare Fig. 1.12 to Figs. 1.9 and 1.11; the sampled output in Fig. 1.12 is a much closer approximation to the presampled signal. Again, the sample rate has not changed; the improvement derives from the reconstruction function used. In this case, the reconstruction function is a sinc wave [that is, a $\sin(x)/(x)$] multiplied by a Gaussian function. The reconstruction function shown in the box in Fig. 1.12 is centered at each sample location. The amplitude at each sample location is equal to the sample value at that location. The sampled output is generated by summing the overlapping reconstruction functions at all times.

The sampled output in Fig. 1.12 is a better representation of the presampled signal than the sampled output in Fig. 1.9. The difference between the sampled outputs derives from the reconstruction technique, not the sample rate. The reconstruction function used in Fig. 1.12 provides a better approximation of the original signal between the samples.

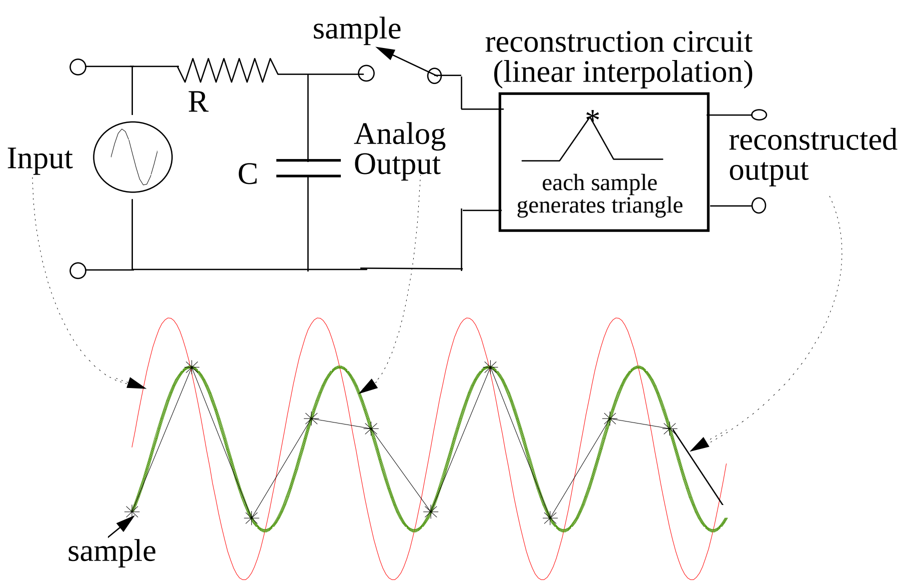


Figure 1.11 Low-pass filter circuit sampled at the same locations as in Fig. 1.9. In this figure, the output is reconstructed by linear interpolation. That is, the samples are connected with straight lines. The sampled output in this figure more closely matches the presampled waveform than does the sampled output in Fig. 1.9. Linear interpolation reconstructs the sampled waveform better than sample replication.

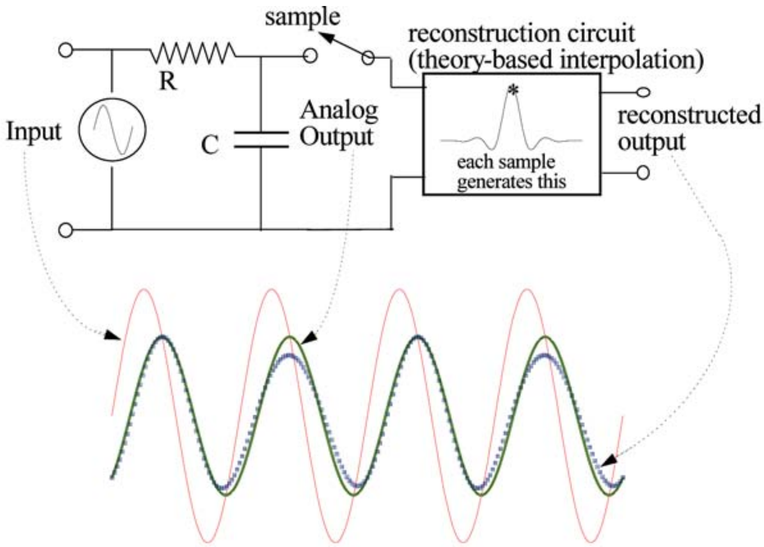


Figure 1.12 Low-pass filter circuit sampled at the same locations as in Fig. 1.9. In this figure, the output is reconstructed by theory-based interpolation. The sampled output in this figure more closely matches the presampled waveform than does the sampled output in Fig. 1.9 or 1.11. The interpolation used in this figure is better than either linear interpolation or replication.

Once the samples are taken, other information about the presampled signal is lost. However, if we know that the presampled signal was essentially bandlimited, then using replication to interpolate between samples cannot be correct. The resulting signal is certainly not bandlimited. Knowing something about the presampled signal, a better interpolation function is generated, and the fidelity of the sampled output is improved.

The quantitative benefit of improved interpolation techniques is described in future chapters. For the present, the reader is asked to consider two concepts. First, although the sample rate is an important factor in determining the fidelity of a sampled process, it is only one factor. The method used to reconstruct the signal (that is, the method used to interpolate between sample values) is also very important. Second, in most practical circumstances, nothing is known about the presample signal except the samples themselves. However, much is known about the presample filtering. In Fig. 1.12, for example, the characteristics of the low-pass filter are known. For images, the characteristics of the optical and detector components of the camera are known. Knowledge of the presample filtering can be used to predict the value of the presample signal between samples, thereby improving the fidelity of the output.

1.5 Three Ways of Viewing the Sampling Process

This section describes three different physical views of the sampling process. The three views each provide a different insight into the effects of sampling on imager performance. The mathematical derivations associated with each physical view quantify how the presample (sensor) blur, the sample rate, and the postsample (display) blur all affect the displayed image.

1.5.1 The displayed image as the sum of its parts

A sampled image is made up of individual display pixels, each illuminated in proportion to the FPA detector sample located at a corresponding spatial location. The image is mathematically described as the sum of the individual pixel intensities.

In Fig. 1.13, the function $f(x)$ is sampled at uniformly spaced intervals. If N samples are taken with spacing X, an approximation $g(x)$ to $f(x)$ can be constructed as follows:

$$g(x) = \sum_{n=0}^{N-1} f(nX)r(x - nX), \tag{1.10}$$

where $r(x)$, the reconstruction function, represents the intensity distribution (or shape and size) of an individual display pixel. Figure 1.14 shows the function $f(x)$, the reconstructed output $g(x)$, and the rect function used as $r(x)$ to generate the $g(x)$ shown. The selection of reconstruction function $r(x)$ and the sample interval X are fundamental to the fidelity with which $g(x)$ approximates $f(x)$. Different functions for $r(x)$, or a smaller spacing X, would lead to different functions $g(x)$ that better approximate $f(x)$. Equation (1.10) shows that $g(x)$ is just the sum of the pixel shapes, each placed at its proper location in space and each weighted by the corresponding sample value.

The Fourier transform of $g(x)$ is $G(\xi)$, where ξ is the frequency in cycles per milliradian or cycles per millimeter. In the following expression, $F(\xi)$ is the Fourier transform of $f(x)$, and $R(\xi)$ is the transform of $r(x)$. $R(\xi)$ is the MTF for the intensity pattern associated with an individual display pixel. The Fourier transform of $r(x - nX)$—the Fourier transform for a display pixel located at location (nX)—is $R(\xi)\exp^{-j2\pi\xi nX}$. Therefore,

$$G(\xi) = \sum_{n=0}^{N-1} f(nX)R(\xi)e^{-j2\pi\xi nX}. \tag{1.11}$$

Each display pixel is placed at the proper location, and its intensity is weighted by the corresponding sample value. The Fourier transform $G(\xi)$ is just the sum of the transforms of the individual display pixels.

Figure 1.15 shows the Fourier transforms of the space domain functions $f(x)$ and $r(x)$. In this example, $f(x)$ is Gaussian, and $r(x)$ is a rect function. Figure 1.16 shows the Fourier transform $G(\xi)$ of $g(x)$. Notice that, due to the sampling and reconstruction with a rect function, $G(\xi)$ contains frequencies not in $F(\xi)$.

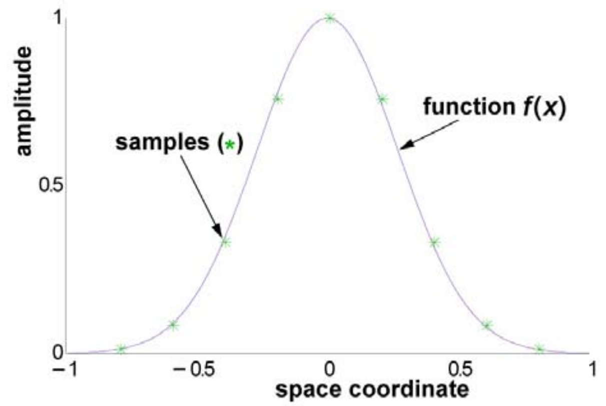


Figure 1.13 The spatial function $f(x)$ is sampled at the points shown by an asterisk.

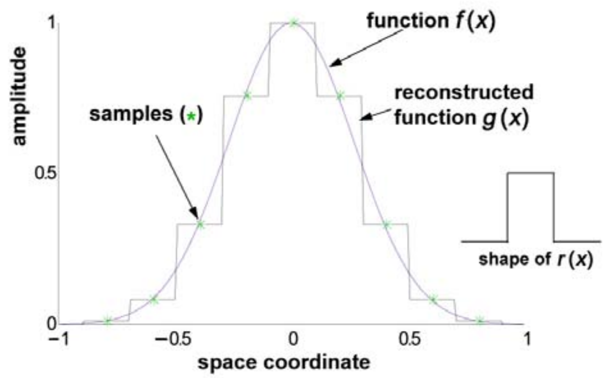


Figure 1.14 The function $f(x)$ is sampled at the asterisk. The sample spacing is five per distance unit. The reconstructed function $g(x)$ is formed by placing a reconstruction function $r(x)$ at each sample location with a height equal to the sample value.

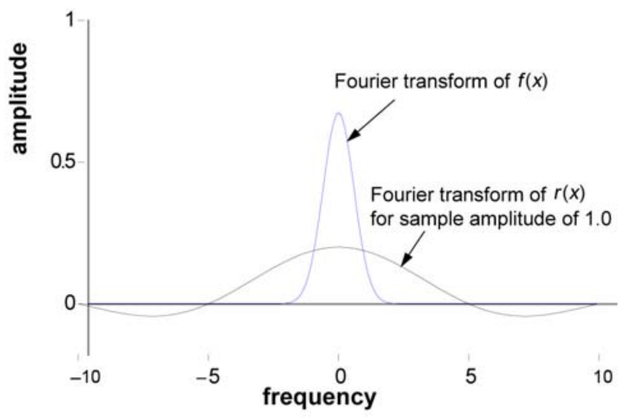


Figure 1.15 Fourier transforms $F(\xi)$ of $f(x)$ and $R(\xi)$ of $r(x)$. The function $f(x)$ is Gaussian, and $r(x)$ is a rect function. The Fourier transform shown for $r(x)$ is for a unit amplitude rect function.

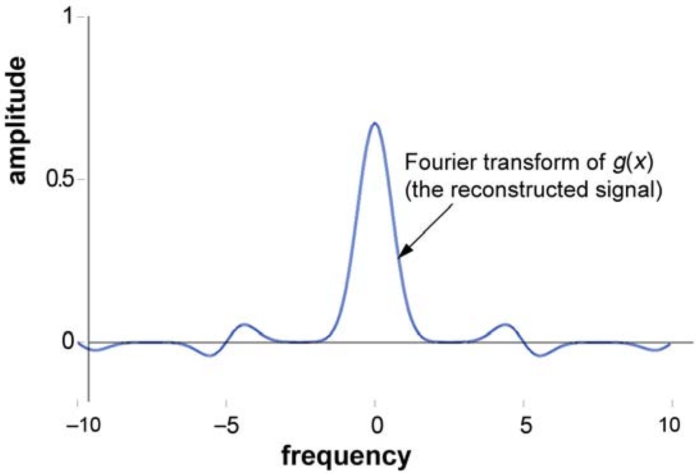


Figure 1.16 The Fourier transform $G(\xi)$ of $g(x)$. Notice that $G(\xi)$ contains frequencies not in $F(\xi)$.

Equation (1.11) illustrates why a system transfer function cannot be assigned to a sampled system. In Eq. (1.11), $R(\xi)$ is not acting to weight or filter the input spectrum $F(\xi)$; that is, $G(\xi)$ does not equal $F(\xi)R(\xi)$. $G(\xi)$ is a sum of $R(\xi)$ terms, not a product of $F(\xi)$ and $R(\xi)$. Furthermore, frequencies not in $F(\xi)$ can exist in the output if they exist in $R(\xi)$.

The above derivations illustrate that the sampled output is just a sum of reconstruction functions. The shape of the reconstruction function is fundamental in establishing the fidelity of the reconstructed signal. The samples control the frequencies in the output only by weighting the $\exp^{-j2\pi\xi nX}$ phase terms. Frequencies in $R(\xi)$ can be canceled or enhanced by the action of the weighted phase terms.

Equation (1.11) is the easiest and most straightforward way of obtaining the Fourier transform of a sampled image. Other analysis techniques, however, provide more insight into the relationship between the display, the sensor samples, and the presample blur.

1.5.2 The display as a filter of the image samples

Equation (1.11) can be factored as shown in Eq. (1.12):

$$G(\xi) = R(\xi)\left[\sum_{n=0}^{N-1} f(nX)e^{-j2\pi\xi nX}\right]. \tag{1.12}$$

Let

$$S(\xi) = \left[\sum_{n=0}^{N-1} f(nX)e^{-j2\pi\xi nX}\right]. \tag{1.13}$$

Now, $R(\xi)$ is a function of frequency that multiplies $S(\xi)$ (the function of frequency and X in the brackets). Figure 1.17 shows $S(\xi)$, $R(\xi)$, and the product $G(\xi)$. $G(\xi)$ is, of course, the same function as was shown in Fig. 1.16.

$S(\xi)$ is a periodic function of frequency that can be thought of as representing the Fourier transform of the samples of $f(x)$. The display MTF $R(\xi)$ acts upon $S(\xi)$ to produce the displayed image.

In this view of the sampling process, an array of delta functions spaced X distance apart samples the intensity distribution $f(x)$. After sampling, the delta functions are viewed as a "bed of nails." That is, the delta functions are very bright, very small points of light at the sample locations on the display. The integrated intensity of each point of light is equal to the image intensity at that location. The display pixel shape blurs these points of light to create the final displayed image.

The delta functions are pulses with vanishingly small width and height such that the area under the delta function is one. Multiplying $f(x)$ by a delta function at location nX produces a delta function with area $f(nX)$. This is true because the height of the delta function is scaled by the value of the intensity distribution at each location. Since the area under each delta function pulse is unity before the multiplication, the area of the product is the sample value.

The Fourier transform of a delta function of amplitude $f(nX)$ at location nX is $f(nX)\exp^{-j2\pi\xi nX}$. Therefore, Eq. (1.13) can be viewed as the Fourier transform of the samples of $f(x)$.

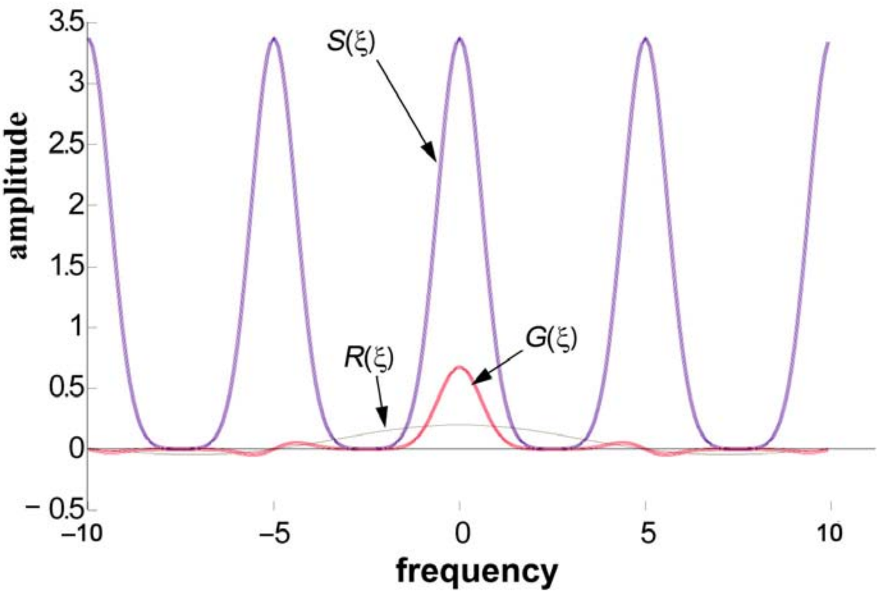


Figure 1.17 Graph showing the same $R(\xi)$ as in Fig. 1.15 and the same $G(\xi)$ as in Fig. 1.16. The amplitude (ordinate) scale has changed between these figures. $S(\xi)$, the Fourier transform of the "samples," is also shown in this figure. $G(\xi)$ results from multiplying $S(\xi)$ by $R(\xi)$.

Equation (1.12) is derived using delta functions in the following manner:

$$g(x) = \sum_{n=0}^{N-1} f(x)\delta(x - nX) * r(x), \tag{1.14}$$

where (*) denotes convolution. By the nature of the delta function,

$$f(x)\delta(x - nX) = f(nX)\delta(x - nX), \tag{1.15}$$

so that

$$g(x) = \sum_{n=0}^{N-1} f(nX)\delta(x - nX) * r(x). \tag{1.16}$$

A convolution in the space domain is a multiplication in the frequency domain. Taking the Fourier transform of both sides of Eq. (1.16) and knowing that the Fourier transform of $\delta(x{-}nX)$ is $\exp^{-j2\pi\xi nX}$, Eq. (1.12) gives the Fourier transform for $G(\xi)$.

Thus far, the sampling process has been dealt with in a physical and, hopefully, intuitive manner. However, Eqs. (1.12) and (1.16) have a limited utility in system analysis because they require the sample values. Also, the relationship between the pre- and postsample blurs and the sample rate and phase is not explicit in Eqs. (1.12) and (1.16).

1.5.3 The display as a filter of the sampled image

A different and more useful expression for the Fourier transform is found by starting with Eq. (1.14). The new expression for $G(\xi)$ does not involve the sample values.

In all of the above sampling examples, the $x = 0$ origin is the first sample point. The function $f(x)$ is defined as the intensity distribution starting at the origin. The function $f(x)$ is simply whichever spatial distribution is being sampled. The above expressions for $g(x)$ and $G(\xi)$ do not lack generality because, in those expressions, $f(x)$ is only known by its sample values. If the imaged scene shifts then $f(x)$ is redefined, and new sample values are obtained at $f(nX)$, $n = 0, 1, 2,\ldots N - 1$.

In the next view of the sampling process described below, we want to explore the behavior of $g(x)$ and $G(\xi)$ as the sample phase varies. The function $f(x)$ is a specific function in image space. The samples of $f(x)$ are taken anywhere in space. As sample phase or position varies, the function $f(x)$ does not change; only *where* it is sampled changes.

Equation (1.14) is rewritten to explicitly permit the function $f(x)$ to be located anywhere in sample space. This is done by letting $f(x)$ be offset in space by a distance x'. The origin at $x = 0$ is still a sample point:

$$g(x) = \sum_{n=0}^{N-1} f(x - x')\delta(x - nX) * r(x). \tag{1.17}$$

The Fourier transform is taken before the delta functions multiply $f(x)$. Since a multiplication in the space domain is a convolution in the frequency domain,

$$G(\xi) = \left[\sum_{n=0}^{N-1} F(\xi) e^{-j2\pi\xi x'} * e^{-j2\pi\xi nX} \right] R(\xi), \tag{1.18}$$

where (*) denotes convolution. In Eq. (1.17) for the space domain, the function $f(x)$ multiplies the delta functions, and the products are convolved with the reconstruction function $r(x)$. In Eq. (1.18) for the frequency domain, the spatial products become frequency convolutions, and the spatial convolutions become frequency products.

Equation (1.18) provides a general expression for $G(\xi)$ based on $F(\xi)$, $R(\xi)$, the sample spacing X, and the offset x'. However, this equation involves a messy convolution. In order to simplify Eq. (1.18), all space is sampled. The answer from sampling all space will equal the answer from sampling a finite interval only if $f(x)$ is zero outside the sampled interval. In the expression for $G(\xi)$ below, $F(\xi)$ is the Fourier transform of $f(x)$ limited to the sampled interval. That is, the Fourier transform $F(\xi)$ is the transform of the *sampled part* of the image.

This distinction is important only if $f(x)$ is thought to represent a entire scene, only part of which is sampled. In the derivation below, $f(x)$ must represent only the sampled part of the scene, and $F(\xi)$ is the Fourier transfer of that limited (windowed) $f(x)$.

If all space is sampled, then

$$g(x) = \sum_{n=-\infty}^{\infty} f(x - x')\delta(x - nX) * r(x). \tag{1.19}$$

The Fourier transform of an infinite set of Dirac delta functions spaced X apart in the spatial domain is an infinite set of Dirac delta functions spaced $1/X$ apart in the frequency domain:

$$\Im\left[\sum_{n=-\infty}^{\infty} \delta(x - nX) \right] = \sum_{n=-\infty}^{\infty} \delta(\xi - n/X), \tag{1.20}$$

where $\Im[\ ] \equiv$ Fourier transform operation.

Therefore,

$$G(\xi)=\left[\sum_{n=-\infty}^{\infty} F(\xi)e^{-j2\pi\xi x'} * \delta(\xi - n/X)\right]R(\xi)$$

$$G(\xi)=R(\xi)\sum_{n=-\infty}^{\infty} F(\xi-n\nu)e^{-j2\pi(\xi-n\nu)x'}$$

$$G(\xi)=R(\xi)e^{-j2\pi\xi x'}\sum_{n=-\infty}^{\infty} F(\xi-n\nu)e^{j2\pi n\nu x'}, \tag{1.21}$$

where

$$\nu = \frac{1}{X}.$$

The phase term $e^{-j2\pi\xi x'}$ represents a translation in space of the entire imaging process by a distance x'. This phase term can be ignored with no loss of information about the sampled system. The $e^{-j2\pi n\nu x'}$ term can be simplified. Let x'/X = integer M plus remainder d/X, where $0 \le d \le X$. Also, remember that ν is $1/X$ (the sample spacing).

The output spectrum is

$$G(\xi)=R(\xi)\sum_{n=-\infty}^{\infty} F(\xi-n\nu)e^{j2\pi n\nu x'}$$

$$G(\xi)=R(\xi)\sum_{n=-\infty}^{\infty} F(\xi-n\nu)e^{jn2\pi x'/X}$$

$$G(\xi)=R(\xi)\sum_{n=-\infty}^{\infty} F(\xi-n\nu)e^{jn2\pi M}e^{jn2\pi d/X} \tag{1.22}$$

$$G(\xi)=R(\xi)\sum_{n=-\infty}^{\infty} F(\xi-n\nu)e^{jn2\pi d/X}$$

$$G(\xi)=R(\xi)\sum_{n=-\infty}^{\infty} F(\xi-n\nu)e^{jn\phi},$$

where $\phi \equiv$ sample phase in radians.

This view of the sampling process is illustrated in Fig. 1.18. The presample image spectrum is replicated at every integer multiple of the sample frequency. This is shown in Fig. 1.18(a). Figure 1.18(a) also shows the display MTF. The Fourier transform of the image is found by multiplying the replicated spectrum by the display MTF. The resulting image spectrum is shown in Fig. 1.18(b).

The phase of each adjacent replica changes by the sample phase. If the display MTF does not filter out the replicas of $F(\xi)$ adjacent to the baseband, then the displayed image will vary with sample phase.

Equation (1.22) relates the displayed image spectrum to the spectrum of the presampled function, the sample spacing, the sample phase, and the display

MTF. Equation (1.22) quantitatively describes the relationship between the various stages of the sampling process. This equation allows us to explore design tradeoffs in a sampled imager.

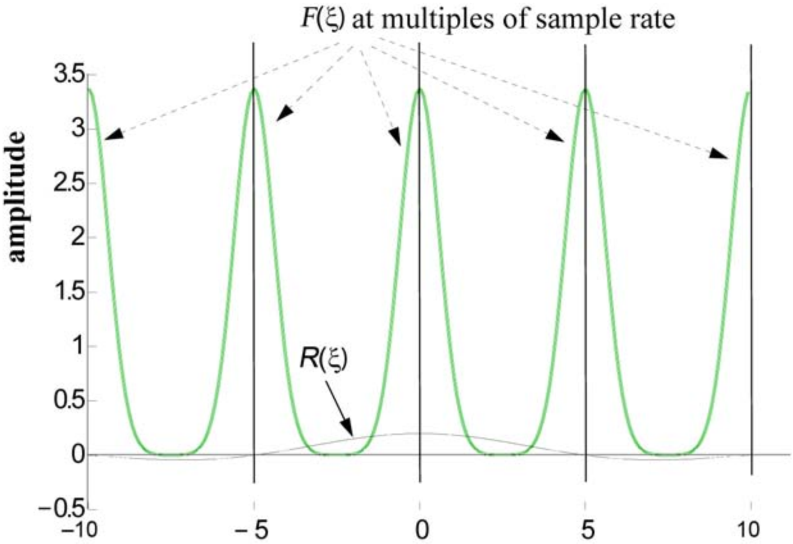


Figure 1.18(a) Sampling replicates the spectrum $F(\xi)$ at every integer multiple of the sample frequency. In this case, the sample rate is five per unit distance. The MTF of the display $R(\xi)$ then multiplies the replicated spectra. The result is the image spectrum $G(\xi)$ shown in the next figure.

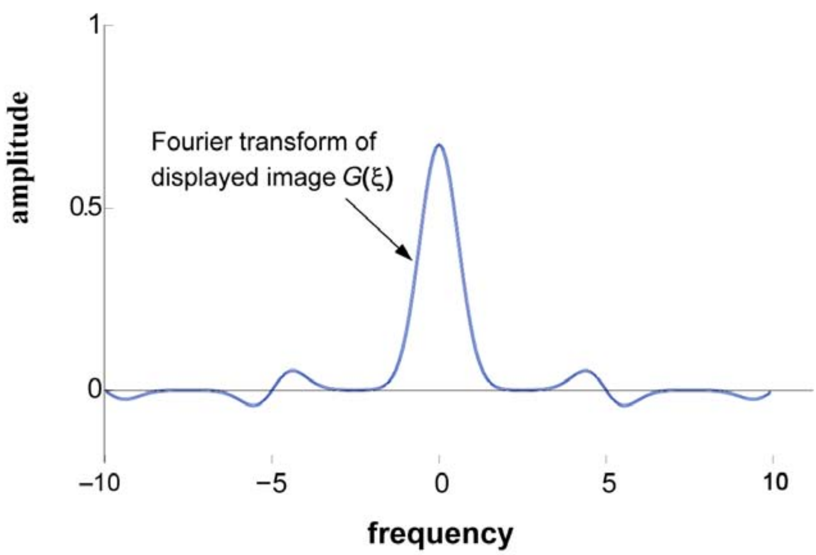


Figure 1.18(b) The Fourier transform of the displayed image $G(\xi)$ obtained by multiplying replicas of $F(\xi)$ at each multiple of the sample frequency by the display MTF $R(\xi)$. The $G(\xi)$ shown here and in Fig. 1.16 are the same.

1.6 The Sampling Theorem

The sampling theorem represents an ideal limit. It substitutes ideal pre- and postfilters for the sensor and display MTF, respectively. As a result, real system behavior is not characterized. The sampling theorem is not used directly to analyze the performance of electro-optical imagers.

Nonetheless, the sampling theorem provides insight into the ideal design. Also, a discussion of the sampling theorem illustrates the important contribution of the reconstruction function $r(x)$. This section describes the sampling theorem and gives an example of a near-ideal reconstruction of a sampled waveform. Also, some of the common misconceptions about the dictates of the sampling theorem are discussed.

1.6.1 Theory

If the Fourier transform $F(\omega)$ of $f(x)$ has no components at or above frequency $f_{samp}/2$, the function is entirely reconstructed by the series

$$f(x) = \sum_{n=-\infty}^{\infty} f(n/f_{samp}) \frac{\sin(\pi x f_{samp} - n\pi)}{(\pi x f_{samp} - n\pi)}. \tag{1.23}$$

The function $f(x)$ is sampled with sample frequency f_{samp}. As discussed in Section 1.5, $F(\omega)$ is replicated at multiples of the sample frequency, as shown in Fig. 1.20. If the replicas of $F(\omega)$ do not overlap the baseband, meaning that $F(\omega)$ is bandlimited to half the sample frequency, then $F(\omega)$ can be exactly reconstructed by using an ideal filter. This is also shown in Fig. 1.20.

An ideal filter is a rect function, with MTF of one (1.0) out to half the sample frequency and zero beyond. The Fourier transform of a rect function is a $\sin(x)/x$ or sinc wave. Convolving the sampled data with a sinc wave in the spatial domain provides an ideal reconstruction filter in the frequency domain. Equation (1.23) represents the convolution in space. Each sample at location n/f_{samp} is convolved with the function $\sin(\pi\, x\, f_{samp})/(\pi\, x\, f_{samp})$. Since the samples are delta functions at the sample locations, Eq. (1.23) represents the convolution.

The sampling theorem reconstruction is illustrated in Fig. 1.21. Two samples are indicated by asterisks in Fig. 1.21(a). For each sample, a sinc wave is generated with peak amplitude equal to the sample value. The period of each sinc wave is such that it goes through zero amplitude at every other sample location. The sum of all of the sinc waves generated for all sample values is the reconstructed signal. The solid dark curve in Fig. 1.21(b) shows the sum of the sinc waves for the two samples in Fig. 1.21(a).

A bandlimited function is uniquely determined from its values at a sequence of equidistant points that are $1/f_{samp}$ apart. The series in Eq. (1.23) is used to reconstruct the function. Each term in the series is a sample function, also referred to as a sinc function or sinc wave. For each term in the series, the sinc

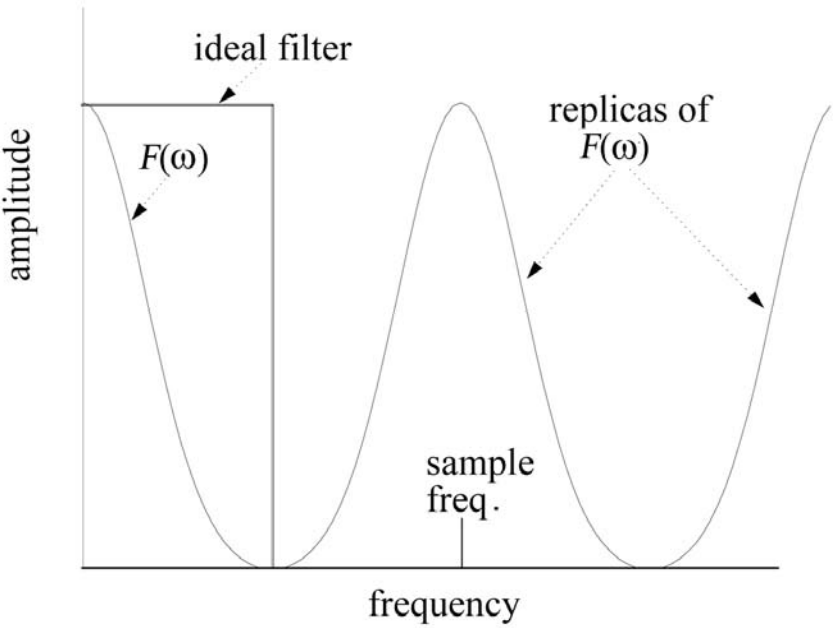


Figure 1.20 If the replicas of $F(\omega)$ do not overlap, then an ideal filter that cuts off at half the sample frequency will exactly recover the original function.

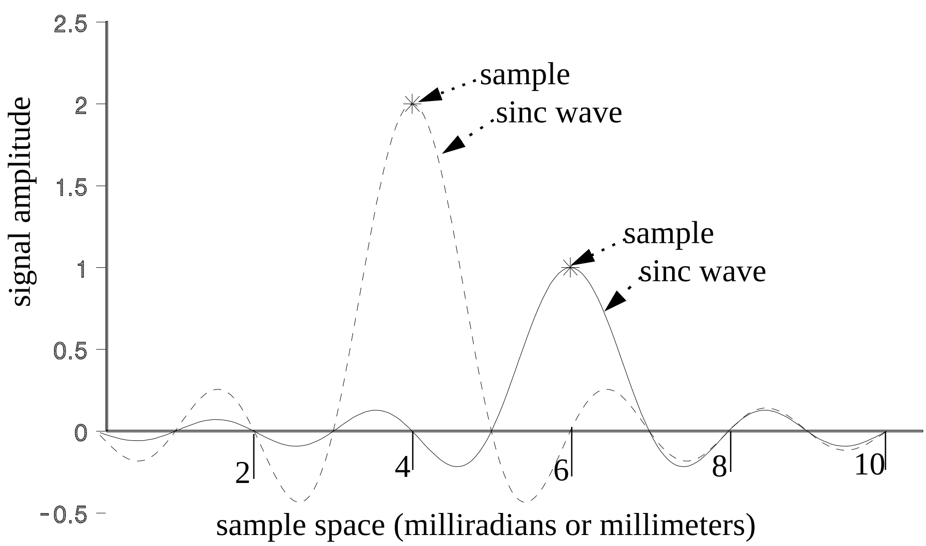


Figure 1.21(a) Example showing sinc wave reconstruction required by the sampling theorem. In this example, the sample spacing is 1.0. Two samples are shown by asterisks at locations (4,2) and (6,1). Note that the samples at positions 1, 2, 3, 5, 7, 8, etc., are not illustrated. For each sample, a sinc wave is generated with the peak equal to the sample value and period such that the sinc wave goes through zero amplitude at all other sample points.

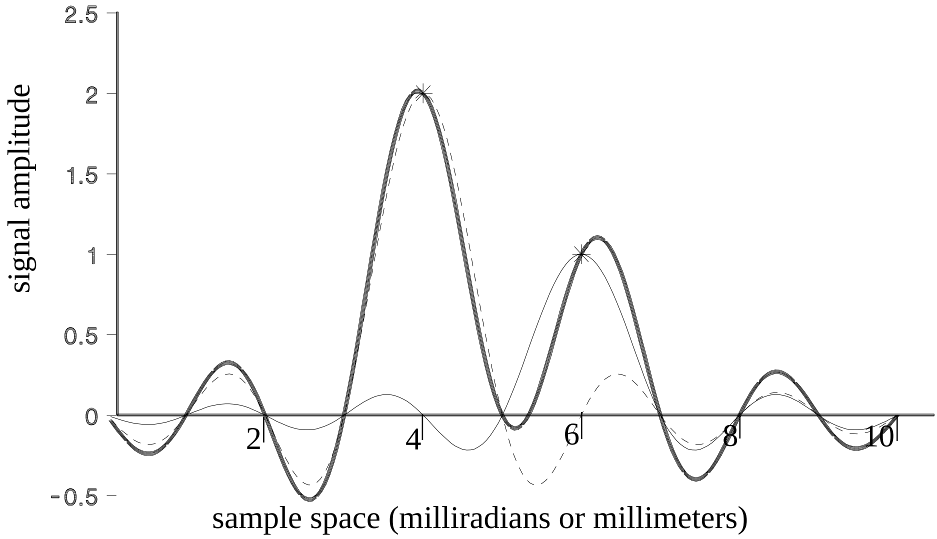


Figure 1.21(b) To generate the reconstructed waveform, all sinc waves are added together. The sum of the two sinc waves shown in Fig. 1.21(a) is shown here as the solid dark line. This example used only two samples. In the general case, sinc waves are generated for all samples, and the sum of all sinc waves is the reconstructed waveform.

wave amplitude is equal to the sample value, and the period is such that the sinc wave crosses zero at all other sample points. The function $f(x)$ is sampled over all space, and each sinc wave extends in both directions over all space.

1.6.2 Example

If a sine wave is infinitely extended, then the Fourier transform is a delta function at the sine-wave frequency. Just over two samples per cycle is adequate to reconstruct the waveform. The sampling theorem does not suggest, however, that a single cycle of a sine wave can be reconstructed by taking two samples. The Fourier transform of a single cycle of sine wave is shown in Fig. 1.22; the sine wave has a period of one unit (milliradians or millimeters). Even using the first zero as the "bandlimit," the sample rate would be 4.0 per cycle.

If the sample rate is just over two samples per cycle, an extended portion of the sine-wave function must be sampled and reconstructed in order to accurately replicate any one cycle. The following example illustrates this concept.

Figure 1.23 shows ten cycles of a sine wave. Fig. 1.24 is the Fourier transform of those ten cycles. Taking the first zero in Fig. 1.24 as an approximation to the "bandlimit," the function is sampled 2.2 times per cycle. The asterisks in Fig. 1.23 show the sample points.

Figure 1.25 shows the result of sampling and reconstructing the sixth period near the middle. Figure 1.26 shows the result of generating sinc waves for all 22 samples and adding them together. The sixth cycle is reconstructed mainly by contributions from neighboring cycles of the sine wave.

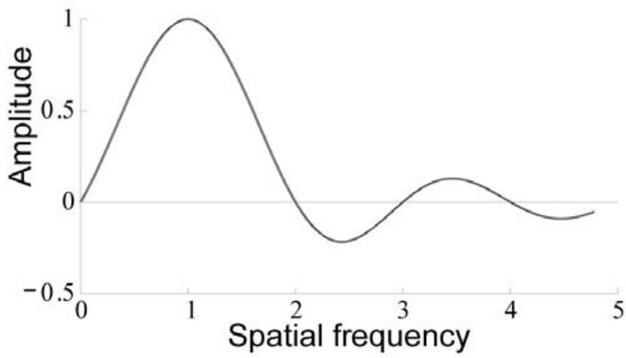


Figure 1.22 The Fourier transform of a single cycle of a sine wave with period of one. The transform is a single frequency only if the sine wave is infinitely extended.

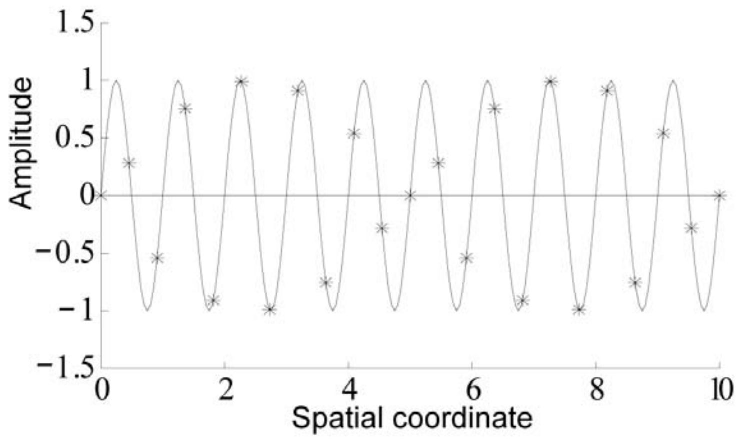


Figure 1.23 Ten cycles of a sine wave. The function is sampled 2.2 times per cycle at the asterisks.

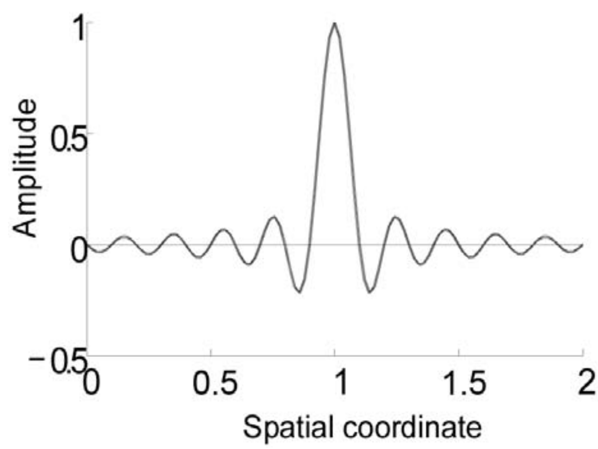


Figure 1.24 Fourier transform of the ten sine-wave periods shown in Fig. 1.23.

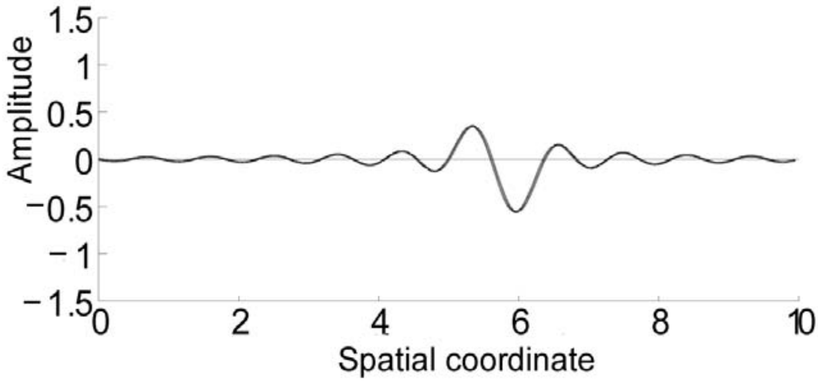


Figure 1.25 The sixth cycle is not properly reconstructed using only the samples from that period.

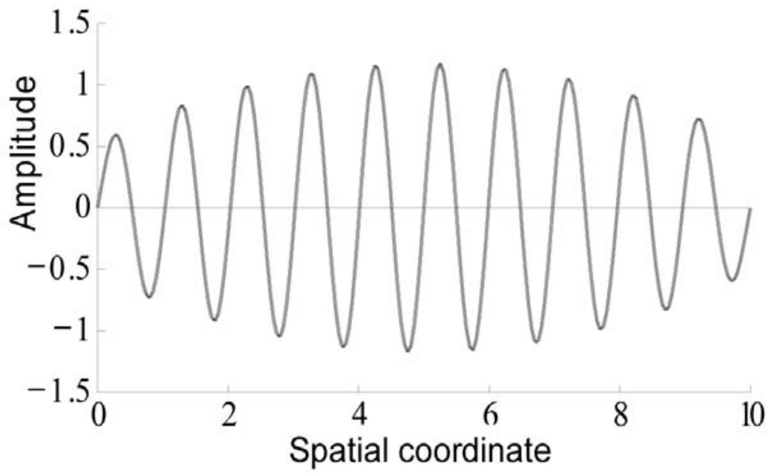


Figure 1.26 Reconstruction of Fig. 1.23 using all 22 samples. Note that the sixth cycle is now properly reconstructed. However, the end cycles are still low in amplitude.

The sixth cycle is poorly sampled due to the sample phase. It reproduces well in the final result because it is surrounded by other sampled periods of the sine wave. The edge cycles of the original ten sine-wave periods are also poorly sampled. The edge cycles do not reconstruct to the full amplitude because they have neighbors on only one side.

The sampling theorem states that a sine wave can be accurately reconstructed by taking just over two samples per cycle for an extended number of cycles. Each cycle of the sine wave is reconstructed by combining the sinc waves from all of the samples. Even if a large number of periods are sampled at just over two samples per cycle, the edges of the function will not be adequately reconstructed because half of the contributions from neighboring cycles are missing.

1.6.3 Discussion

The sampling theorem states that given a sufficient sample rate, a local region of a bandlimited function is reconstructed from both local and remote samples of the function. The sampling theorem does not suggest that two or three samples will reconstruct a single cycle of a sine wave. Nor does the sampling theorem support the assertion that just over two samples per cycle of a three- or four-bar pattern are sufficient to reconstruct the pattern.

An assumption often made when modeling sampled imagers is that frequency content at up to half the sample rate is accurately imaged by the system. This assumption is incorrect if the imager has significant MTF response beyond half the sample rate. Aliasing occurs if there is frequency content beyond half the sample rate. The sampling theorem does not suggest that a sampled signal is reconstructed accurately to half the sample rate. That assumption ignores aliasing.

The sampling theorem assumes that the presample image is ideally bandlimited and that an ideal filter is used for image reconstruction at the display. However, for practical reasons, ideal filters are not implemented. The sampling theorem does not provide a basis for evaluating real systems; it represents an ideal goal.

Bibliography

Bracewell, R. N., *The Fourier Transform and Its Applications*, McGraw-Hill, New York (1986).

Gaskill, J., Linear Systems, *Fourier Transforms and Optics*, Wiley, New York (1978).

Goodman, J. W., *Introduction to Fourier Optics*, McGraw-Hill, San Francisco (1968).

Holst, G. C., *Sampling, Aliasing, and Data Fidelity,* SPIE Press, Bellingham, WA (1998).

Wozencraft, J. M. and I. M. Jacobs, *Principles of Communication Engineering*, Wiley, New York (1965).

Chapter 2
Fourier Integral Representation of an Optical Image

This chapter describes optical transfer functions. The concepts of linearity and shift invariance were introduced in Chapter 1. This chapter continues that discussion by applying those concepts to optical imaging components and systems.

Images are two-dimensional and are accurately described by two-dimensional Fourier integrals. Common practice, however, is to analyze or measure horizontal and vertical frequency response and then use the results to characterize imager performance. This chapter describes the errors that result from assuming that an imager is characterized by its horizontal and vertical frequency response.

Throughout this tutorial, the mathematics is at the introductory calculus level. However, the descriptive arguments require some familiarity with the concepts of Fourier analysis and complex functions.

2.1 Linear Shift-Invariant Optical Systems

In Fig. 2.1, a simple optical system is imaging a clock onto a screen. For simplicity, unity magnification is assumed. If each point in the scene is blurred by the same amount, then the system is shift invariant. If the image intensity profile equals the sum of the individual blurs from each point in the scene, then the system is linear.

The optical blur is called the *point spread function* (psf). The psf is illustrated in the lower left corner of the image. Each point source in the scene becomes a psf in the image. The psf is also called the *impulse response* of the system. Each point in the scene is blurred by the optics and projected onto the screen. This process is repeated for each of the infinite number of points in the scene. The image is the sum of all of the individual psf's.

Two considerations are important here. First, the process of the lens imaging the scene is linear and, therefore, superposition holds. The image is accurately represented by the sum of the psf resulting from the lens that is imaging each

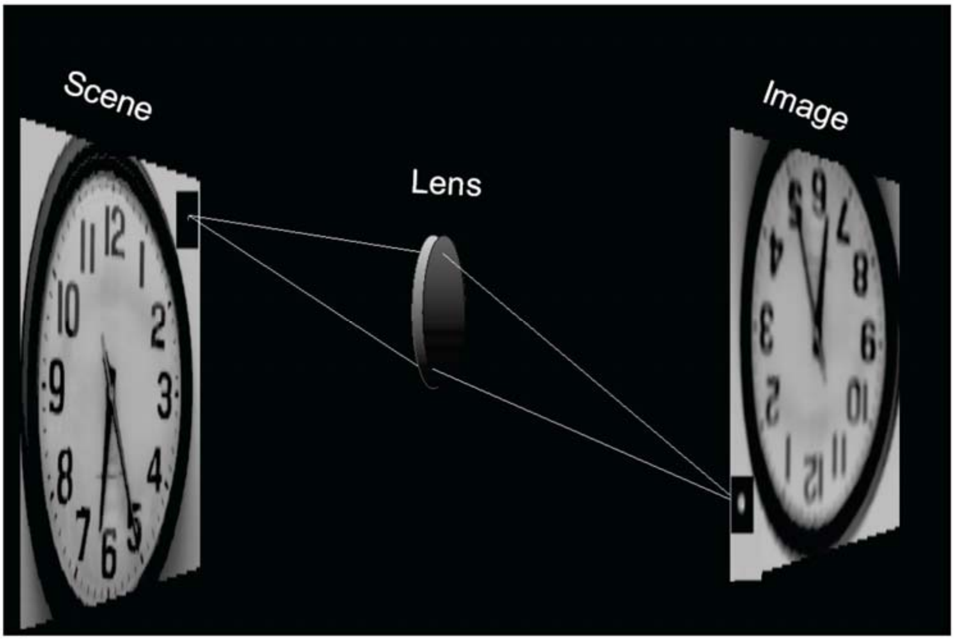


Figure 2.1 Clock being imaged by a lens onto a screen; a point source in the scene (upper right) becomes a point-spread-function blur in the image (lower left).

individual scene point. Second, it is assumed that the shape of the optical blur (that is, the shape of the psf) does not depend on position within the field of view.

In most optical systems, the psf is not constant over the entire field of view. Typically, optical aberrations vary with field angle. The optical blur is generally smaller at the center of an image than it is at the edge. However, the image plane can generally be subdivided into regions within which the optical blur is approximately constant. A region of the image with approximately constant blur is sometimes called an *isoplanatic patch*. Optical systems are linear and shift-invariant over isoplanatic regions of the field of view.

The image within an isoplanatic patch can be represented as a convolution of the psf over the scene. If $h(x,y)$ represents the spatial shape (the intensity distribution) of the psf, then $h(x-x',y-y')$ represents a psf at location (x',y') in the image plane. The units of x and y are milliradians (mrad). Let $s_{cn}(x',y')$ describe the brightness of the scene, and $i_{mg}(x, y)$ describe the brightness of the image:

$$i_{mg}(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x-x', y-y')\, s_{cn}(x', y')\, dx'\, dy' . \tag{2.1}$$

Each point in the scene radiates independently and produces a point psf in the image plane with corresponding intensity and position. The image is a linear superposition of the resulting psf's. Mathematically, that result is obtained by convolving the optical psf over the scene intensity distribution to produce the image. Since a convolution in space corresponds to a multiplication in frequency, the optical system is a spatial filter:

$$I_{mg}(\xi,\eta) = H(\xi,\eta)S_{cn}(\xi,\eta), \tag{2.2}$$

where

$I_{mg}(\xi,\eta)$ = Fourier transform of image,
$S_{cn}(\xi,\eta)$ = Fourier transform of scene, and
$H(\xi,\eta)$ = the optical transfer function (OTF).

ξ and η are spatial frequencies in the x and y direction, respectively. The units of ξ and η are cycles per milliradian (mrad^{-1}).

The OTF is the Fourier transform of the psf $h(x,y)$. However, in order to keep the image intensity proportional to scene intensity, the OTF of the optics is normalized by the total area under the psf blur spot:

$$H(\xi,\eta) = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)e^{-j\xi x}e^{-j\eta y}\,dx\,dy}{\int_{-\infty}^{\infty} h(x,y)\,dx\,dy}. \tag{2.3}$$

The MTF of the optics is the magnitude $|H(\xi,\eta)|$ of the function $H(\xi,\eta)$.

Note that the relationship in Eq. (2.2) applies between the scene and the image plane of a well-corrected optical system. The optical system is considered to be "well-corrected" because the psf (the optical blur) is reasonably constant over the image plane.

Optical systems often have multiple image planes. The first image becomes the scene that is imaged by the second set of optical elements. For example, the image in Fig. 2.1 might be re-imaged by another lens, as shown in Fig. 2.2. In this case, each point in the original image is blurred by the psf of the next set of optics. If the OTF of the second lens is $H_2(\xi,\eta)$, then

$$I_{mg}(\xi,\eta) = H_2(\xi,\eta)H(\xi,\eta)S_{cn}(\xi,\eta). \tag{2.4}$$

The total system MTF is the product of the individual MTFs. One caution is necessary here: Diffraction is caused by the limiting aperture in the imager. In a system with multiple image planes, diffraction MTF is applied only once.

The transfer function between scene and display is the product of optics MTF, detector MTF, display MTF, and the MTF of other factors that blur the

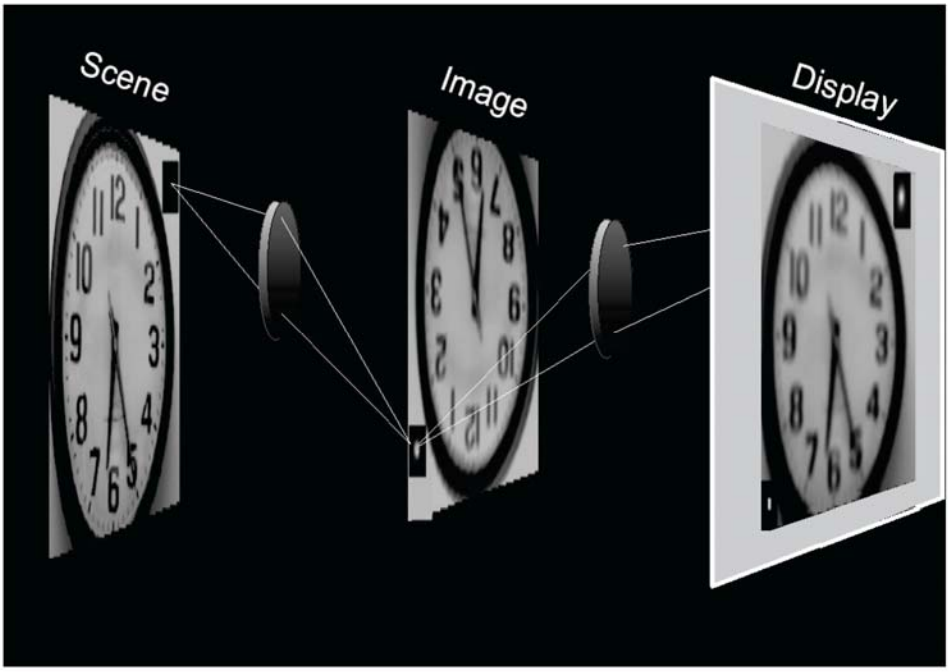


Figure 2.2 The picture is further blurred by imaging with a second lens. The OTF from the scene to the display is the product of the individual lens OTFs.

image. Any blurring of the image can be treated as an MTF as long as the blur is constant over the entire image. For example, the active area of a detector acts as an optical PSF. The larger the active detector area, the more blurred is the image. Light falling anywhere on the detector area is summed together. The detector area convolves with the scene to blur the image in the same way that the optical psf blurs the image.

The MTF of the detector is the Fourier transform of the detector photosensitive area. The display MTF is the Fourier transform of a display pixel-intensity pattern. In the absence of sampling artifacts, the Fourier transform of the displayed image is the Fourier transform of the scene multiplied by the product of optics, detector, display, and other component MTFs.

2.2 Equivalence of Spatial and Frequency Domain Filters

Equations (2.1) and (2.2) describe the filtering process in the space domain and the frequency domain, respectively. In space, the output of an LSI system is the input convolved with the system impulse response (in this case, the optical psf). Consider the example shown in Fig. 2.3. The system is a simple lens imaging the transparency of a four-bar target. Given that the lens blur is the same across the field of view, the system is LSI. The output image is the transparency intensity convolved with the lens psf.

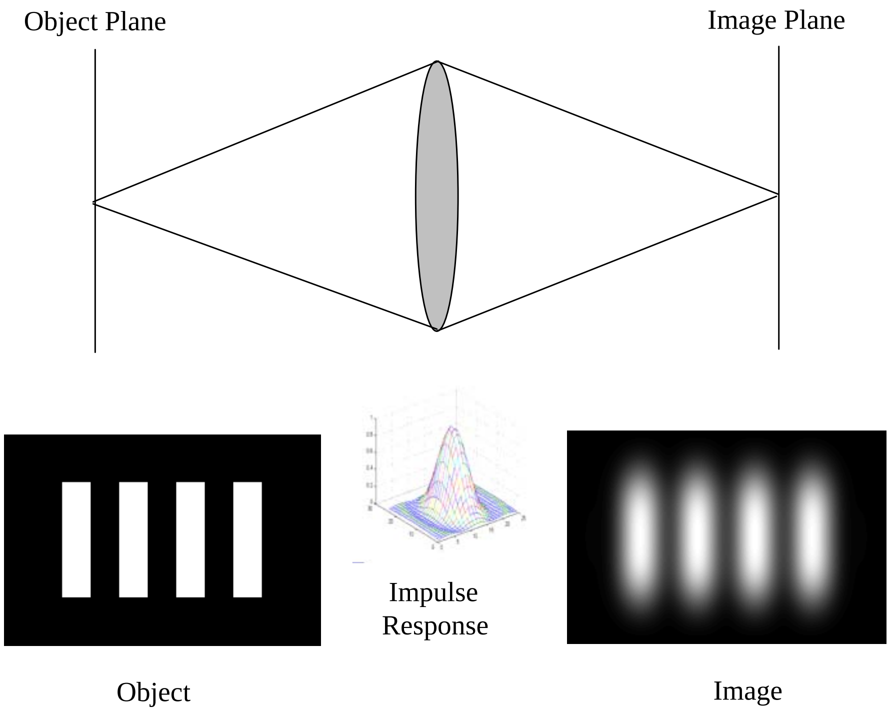


Figure 2.3 Spatial filtering in an optical system.

Figure 2.4 illustrates frequency domain filtering. The two-dimensional Fourier transform of the scene intensity is taken. The input spectrum clearly shows the fundamental harmonic of the four-bar target in the horizontal direction. The higher-order harmonics are difficult to see because they have less amplitude than the fundamental. The Fourier transform of the image is obtained by multiplying the Fourier transform of the scene by the Fourier transform of the psf. The output image is found by taking the inverse transform of the product. The resulting image is identical to that given by the spatial convolution of the psf in the space domain.

In Fig. 2.4, the zero-frequency component of the input-, transfer-, and output-frequency spectrums has been removed so that the higher-frequency components are visible. Otherwise, all that would be seen is a bright point in the middle of the picture.

LSI imaging system analysis can be performed using two methods: spatial domain analysis and frequency domain analysis. The results given by these analyses are identical, although frequency domain analysis has an advantage. Equations (2.1) and (2.3) both involve double integrals. However, an imager has many component MTFs. Using Eq. (2.1) involves calculating double integrals for line-of-sight blur, diffraction blur, optical aberration blur, detector blur, digital filtering blur, display blur, and eyeball blur. Using Eq. (2.3) involves

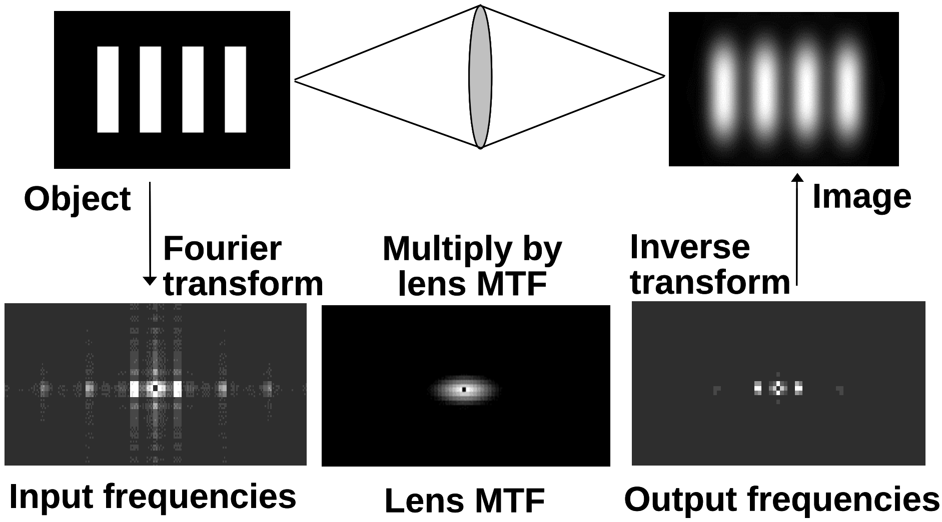


Figure 2.4 Frequency domain filtering in an optical system.

double integrals to find the Fourier transform of the scene and a second double integral to find the spatial image. Intervening calculations involve multiplying the various component MTFs. Fourier domain analysis is used because it provides accurate results with reduced computation.

2.3 Reducing LSI Imager Analysis to One Dimension

It is common to analyze imagers separately in the horizontal and vertical directions. The two-dimensional imager MTF is assumed to be the product of horizontal and vertical MTFs. This assumption reduces two-dimensional Fourier integrals to two one-dimensional Fourier integrals. The one-dimensional treatment, therefore, saves computation.

The separability assumption is almost never satisfied, even in the simplest cases; assuming separability virtually always leads to some error in the result. Nonetheless, the majority of scientists and engineers use the product of horizontal and vertical frequency response as the imager MTF. This section discusses some of the errors that result from this common simplification.

Separability in Cartesian coordinates requires that a function of (x,y) can be expressed as the product of a function of x times a function of y.

$$f(x,y) = f_x(x)\, f_y(y)\ . \tag{2.5}$$

If Eq. (2.5) is true, then the Fourier transform is also separable, and Eq. (2.6) holds.

$$F(\xi,\eta) = F_x(\xi)\, F_y(\eta)\ . \tag{2.6}$$

The optical psf is a combination of the diffraction and geometric aberrations. The on-axis optical blur can be characterized by a function that is separable in polar coordinates. The detector psf is a rectangular shape that is separable in Cartesian coordinates if the detector edges are properly aligned to the coordinate axes. The collective psf of the detector and the optics is not separable in either polar or Cartesian coordinates.

Consider the diffraction psf given by Eq. (2.7). J_1 is a Bessel function of the first kind of order one. D is the optics aperture diameter in meters. Diffraction wavelength is λ meters. Cartesian coordinates (x,y) are centered on the diffraction spot. The factor of 1,000 converts milliradians to radians.

$$psf(x,y) = \left(2J_1\left(\frac{\pi D\sqrt{x^2+y^2}}{1000\lambda}\right) \Bigg/ \frac{\pi D\sqrt{x^2+y^2}}{1000\lambda} \right)^2 . \tag{2.7}$$

The MTF H_{diff} of the psf in Eq. (2.7) is given by Eq. (2.8). (See the book by O'Neill in the bibliography.) A plot of H_{diff} for D equal to 0.1 m and λ equal to 4 μm (microns) is shown in Fig. 2.5. Only positive ξ and η are shown because the MTF is symmetrical about both frequency axes.

$$\chi = 1000\sqrt{\xi^2+\eta^2} \quad \text{cycles per radian.}$$

$$H_{diff}(\xi,\eta) = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{\chi\lambda}{D}\right) - \frac{\chi\lambda}{D}\sqrt{1-\left(\frac{\chi\lambda}{D}\right)^2}\right] \quad \text{for } \frac{\chi\lambda}{D} < 1. \tag{2.8}$$

$$H_{diff}(\xi,\eta) = 0 \quad \text{for } \frac{\chi\lambda}{D} \geq 1.$$

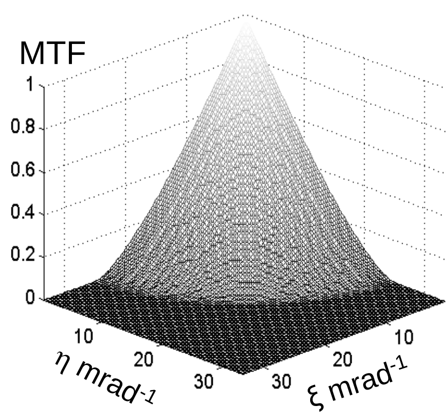


Figure 2.5 MTF of diffraction blur.

Neither Eq. (2.7) nor (2.8) is separable in Cartesian coordinates. Nonetheless, Eq. (2.9) is often used to represent diffraction MTF:

$$
\begin{aligned}
H_{x-diff}(\xi) &= \frac{2}{\pi}\left[\cos^{-1}\left(\frac{1000\xi\lambda}{D}\right)-\frac{1000\xi\lambda}{D}\sqrt{1-\left(\frac{1000\xi\lambda}{D}\right)^2}\,\right] \text{ for } \frac{1000\xi\lambda}{D}<1. \\
H_{x-diff}(\xi) &= 0 \quad \text{for } \frac{1000\xi\lambda}{D}\geq 1. \\
H_{y-diff}(\eta) &= \frac{2}{\pi}\left[\cos^{-1}\left(\frac{1000\eta\lambda}{D}\right)-\frac{1000\eta\lambda}{D}\sqrt{1-\left(\frac{1000\eta\lambda}{D}\right)^2}\,\right] \text{ for } \frac{1000\eta\lambda}{D}<1. \\
H_{y-diff}(\eta) &= 0 \quad \text{for } \frac{1000\eta\lambda}{D}\geq 1. \\
H_{xy-diff} &= H_{y-diff}(\eta)H_{x-diff}(\xi).
\end{aligned}
\tag{2.9}
$$

Figure 2.6 plots the absolute difference between the Eq. (2.8) MTF and the MTF from Eq. (2.9). In the figure, differences of less than 0.01 are zeroed out in order to give perspective to the error plot. The two-dimensional MTF is up to 0.08 larger than the separable MTF for frequencies less than the band limit. Also, the separable MTF allows frequency content that is beyond the diffraction bandlimit.

Two measures are useful in quantifying MTF. The first is Shade's equivalent bandwidth N_e, which is the total sine-wave power passed by the MTF. The second metric is integrated signal I_{sig} to the first zero. Bandwidth beyond the first zero is generally not useful for imaging.

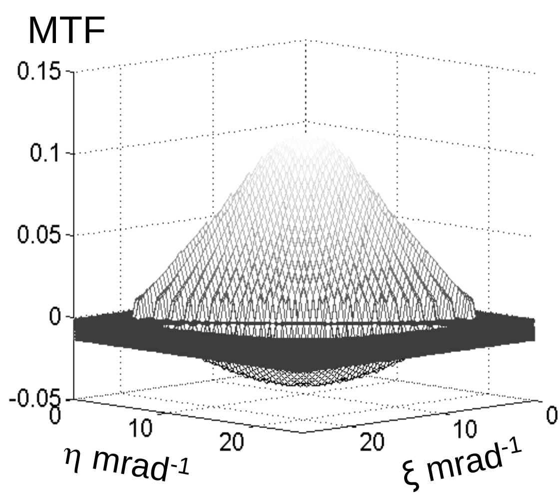


Figure 2.6 Diffraction MTF minus the separable approximation. Diffraction cutoff is at 10 mrad^{-1}.

$$N_e = \iint MTF^2(\xi,\eta)\, d\xi\, d\eta \,. \tag{2.10}$$

$$I_{sig} = \iint_{\substack{\text{first}\\ \text{zero}}} MTF(\xi,\eta)\, d\xi\, d\eta \,. \tag{2.11}$$

The ratio of N_e for separable versus two-dimensional diffraction MTF is 0.79. That is, the separable MTF predicts 21% less energy transfer than physically occurs. The I_{sig} ratio is 0.85. This indicates 15% less signal than is actually passed by the imager. Furthermore, 5% of the separable MTF signal is beyond the bandlimit. So the separable approximation passes signal beyond the diffraction limit but underestimates in-band transfer by 20%. These comparisons are for a 0.1-m aperture and diffraction wavelength of 4 μm. However, since both the nonseparable and separable MTFs scale exactly, these same ratios hold for all aperture and wavelength combinations.

There are many cases where a 0.08 absolute error in MTF is unimportant. The same is true for a 21% error in the signal energy passed by the imager. The separable prediction that 5% of the signal is beyond the diffraction bandlimit is likely unimportant in most cases. However, any of the errors might lead to false conclusions, depending on the goal of the analysis.

The next example discusses errors associated with modeling diamond-shaped detectors. Figure 2.7 illustrates a square detector and the same detector rotated by 45 deg. The rotated detector has a diamond shape. The two-dimensional MTF of the diamond detector is the MTF of the square detector rotated by 45 deg in the frequency domain. However, the equivalence of these two detector shapes in terms of blur properties is obscured by our experience when testing imagers.

Generally, vertical and horizontal edges or slits are used to test imager MTF. If the detector edge is parallel to the test slit, then the first zero in the MTF occurs at a frequency equal to the inverse of the detector width. However, if either the test slit or the detector is rotated 45 deg, then the first zero in the MTF occurs at $\sqrt{2}$ higher frequency. Since it is common to use horizontal and vertical slits during imager test, experience suggests that rotating square or rectangular detectors extends the frequency cutoff of the imager.

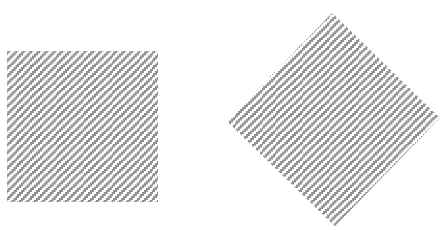


Figure 2.7 Square detector is at the left, and diamond detector is at the right. Rotating the square detector by 45 deg makes it a diamond detector.

Equation (2.12) gives the MTF of a square detector in which the detector edges of length d are parallel to the axes of the Cartesian coordinate system. The two-dimensional MTF in Eq. (2.12) is separable.

$$H_{det} = [\sin(\pi\xi d)/(\pi\xi d)][\sin(\pi\eta d)/(\pi\eta d)]\,. \tag{2.12}$$

Equation (2.13) gives the MTF of the same detector rotated by 45 deg. In this case, the two-dimensional MTF is not separable in ξ and η.

$$\begin{aligned} H_{det} &= [\sin(\pi\xi' d)/(\pi\xi' d)][\sin(\pi\eta' d)/(\pi\eta' d)], \\ \text{where}\quad \xi' &= \xi\cos(45\text{ deg}) - \eta\sin(45\text{ deg}), \text{ and} \\ \eta' &= \xi\sin(45\text{ deg}) + \eta\cos(45\text{ deg}). \end{aligned} \tag{2.13}$$

Equation (2.14) gives the directional frequency response along the horizontal and vertical axes for the diamond detector. This is the same frequency response as that along the diagonal of the square detector. Equation (2.14) is sometimes used as the basis for a separable MTF for the diamond detector.

$$\begin{aligned} H_{diamond} &= [\sin(\pi\xi d')/(\pi\xi d')]^2[\sin(\pi\eta d')/(\pi\eta d')]^2, \\ \text{where}\quad d' &= \frac{d}{\sqrt{2}}. \end{aligned} \tag{2.14}$$

Figures 2.8 and 2.9 show the two-dimensional MTF represented by Eqs. (2.13) and (2.14), respectively. Figure 2.10 shows the difference plot obtained by subtracting Eq. (2.14) from (2.13). Figure 2.10 shows the error that results when using Eq. (2.14) to represent the MTF of a rotated square detector.

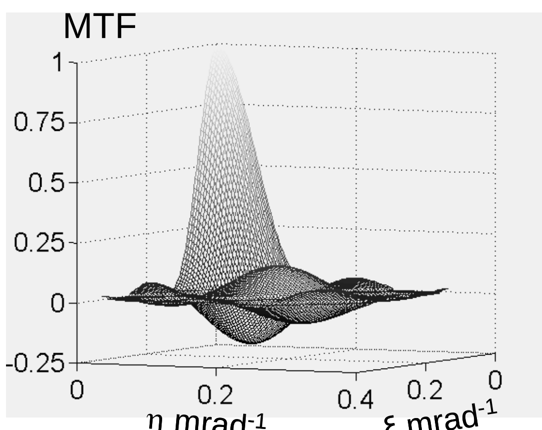


Figure 2.8 Two-dimensional MTF of a rotated square detector. The MTF is given by Eq. (2.13).

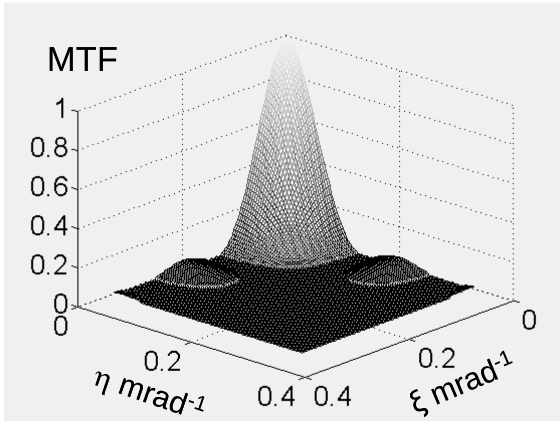


Figure 2.9 Plot of $H_{diamond}$ given by Eq. (2.14)

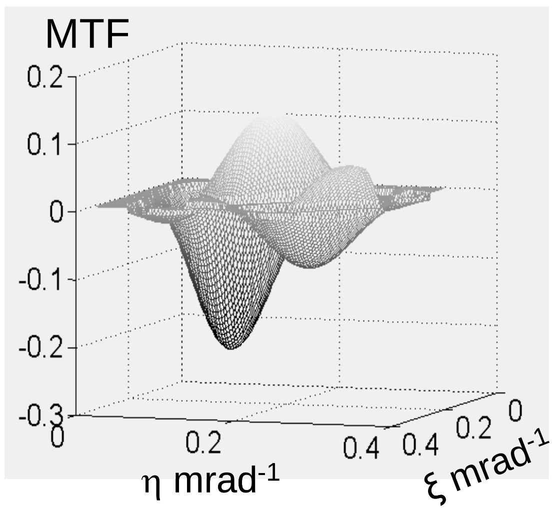


Figure 2.10 Two-dimensional MTF error associated with using Eq. (2.14) to represent a rotated square detector.

As expected, the diamond MTF has unrealistically good frequency response at high spatial frequencies, but also has poorer than actual frequency response at midband. $H_{diamond}$ passes 18% more sine-wave energy, and I_{sig} is 9% larger than the actual two-dimensional MTF of the rotated detector.

Another common simplification is to use real MTF and avoid dealing with complex numbers. Complex numbers have both real and imaginary parts, so complex MTF data tables are twice as large as real data tables. However, only a symmetrical psf has real MTF.

If $f(x)$ is symmetrical about $x = 0$, then the sinusoidal terms in the Fourier transform cancel, and only a real $F(\xi)$ results. The lens at the top of Fig. 2.11 forms a blur that is symmetrical both horizontally and vertically. The resulting

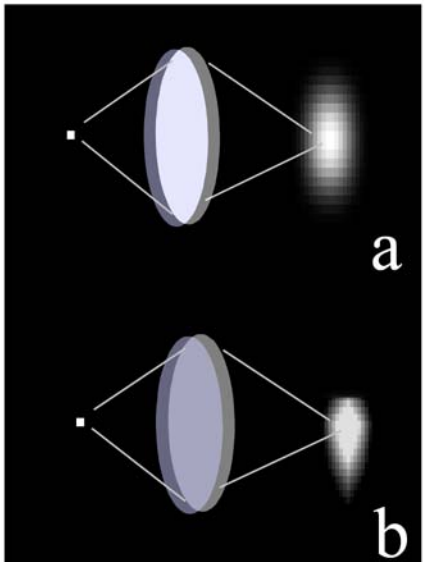


Figure 2.11 (a) The lens forms a blur that is symmetrical about both a horizontal and a vertical axis. The resulting MTF is real. (b) The blur is symmetrical about a vertical axis, but not about a horizontal axis. The resulting horizontal MTF is real, but the vertical MTF is complex.

MTF is real. The lens at the bottom of Fig. 2.11 forms a blur that is symmetrical about a vertical axis, and the horizontal MTF is real. However, the blur at the bottom is not symmetrical about a horizontal axis; therefore, the vertical MTF is complex.

Two simplifications are commonly used when modeling electro-optical imagers. The imager MTF is real and equal to the product of horizontal and vertical frequency response. Both simplifications introduce error into the analysis.

2.4 Perspectives on One-Dimensional Analysis

The purpose of this section is to describe the pros and cons of one-dimensional analysis. There are many negatives. For a tutorial, however, one-dimensional analysis provides the simplest way of describing complex concepts. In Chapter 1, one-dimensional examples were used to describe sampling and signal reconstruction. Most of the concepts covered in the remainder of this book are described in one dimension, and most of the examples use functions of a single variable. Concepts are easier to understand in one dimension, and one-dimensional math is less cluttered. Once the concept is understood, the leap to two dimensions is more intuitive.

There is a dichotomy in difficulty between describing a two-dimensional function and performing two-dimensional analysis. Consider the detector MTF discussion in the last section. Two-dimensional Fourier transforms are easy to implement numerically. On the other hand, three figures (Figs. 2.8, 2.9, and 2.10) are needed to plot the two MTF functions and to show the difference plot. Figure

2.12 shows optics, detector, display, and total system MTF all on the same plot. The shape of each function, the relative importance of each MTF at each spatial frequency, detector cutoff frequency—all of this information and more is conveyed by the one-dimensional plots. Figure 2.12 is not simply easier to generate; it is visually more meaningful.

The one-dimensional treatment has advantages when explaining concepts or comparing different functions. However, one-dimensional plots can be misleading. Figure 2.13 shows two images of a tank. To the left is the original image. To the right, the image is blurred with a Gaussian kernel. Figure 2.14 shows the two-dimensional Fourier transforms of the images in Fig. 2.13. The transform of the original image is to the left, and the transform of the blurred image is to the right. Figure 2.15 shows the intersection of the two-dimensional plots with the horizontal and vertical planes. That is, Fig. 2.15 is "one-dimensional" plots of the Fourier transforms. The horizontal and vertical plots do not represent the scene in any real sense. No actual scene is separable.

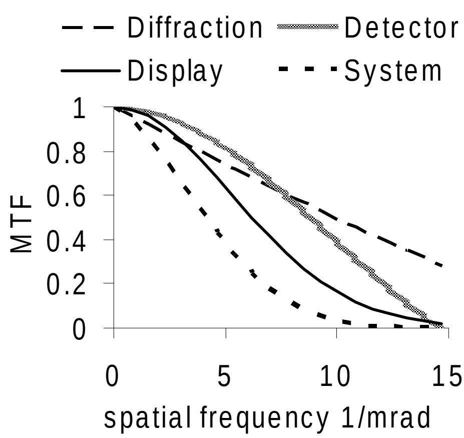


Figure 2.12 This figure illustrates that one-dimensional MTF plots are easier to interpret and better for comparing functions. Figures 2.8 and 2.9 display more information about each function, but a third plot (Fig. 2.10) is needed to show the difference.

Figure 2.13 Original picture of tank to the left and blurred image to the right. The Fourier transforms of these images are shown in Figs. 2.14 and 2.15.

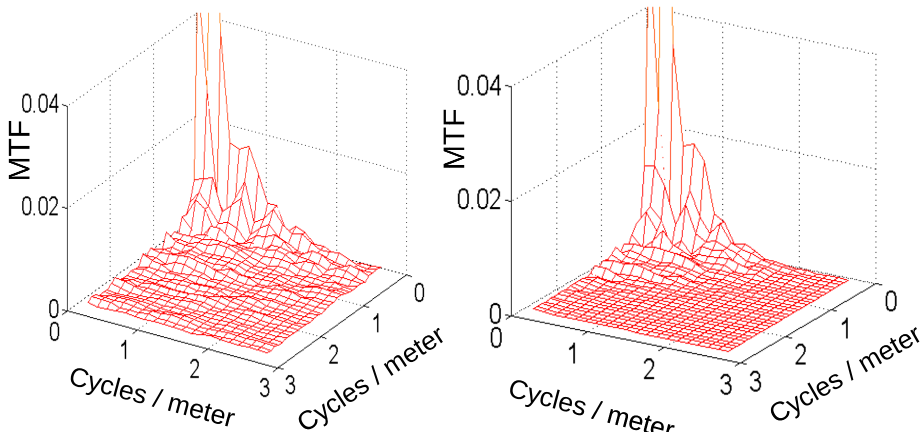


Figure 2.14 Two-dimensional Fourier transforms of the tank images shown in Fig. 2.13. To the left is the Fourier transform of the original; to the right is the transform of the blurred image. Note the expanded ordinate scale.

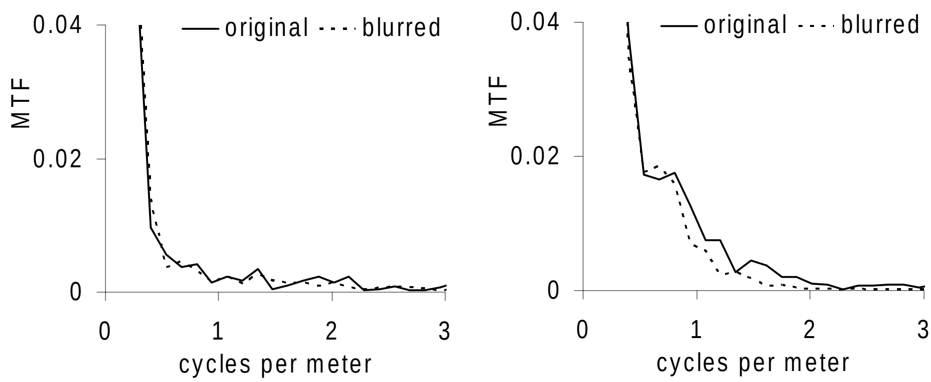


Figure 2.15 Intersections of Fig. 2.14 Fourier transforms with horizontal surface to the left and vertical surface to the right. Note the expanded ordinate scale.

Furthermore, the horizontal and vertical plots present the frequency data in a misleading way. It appears that low frequencies completely dominate the Fourier transforms. This is true even though the ordinate scales are expanded to show the high-frequency amplitudes. In the two-dimensional plot, the low frequencies are a small area. Low-frequency amplitudes are high, but the encompassed volume is limited because the area covered in the frequency domain is small. High-frequency amplitudes are small, but high frequencies cover a large area in the frequency domain. If the frequency domain is viewed as a room, then low frequencies are in the corner of the room, while medium and high frequencies cover the whole floor. The one-dimensional plots do not provide an accurate insight into imagery frequency content.

The one-dimensional treatment is useful for illustrating concepts and showing comparisons. Whenever possible, however, one-dimensional analysis should be avoided. For imager analyses to be meaningful, the mathematical functions must accurately represent the physical quantities they describe. Each compromise in accurately describing the system degrades the value of the analytical result. Furthermore, although many of these compromises were necessary in the past, they are not justified today. The analyst is no longer bound by hand computation, and everyone has access to fast computers. Many of the compromises still made today are simply historical holdovers.

One area in which the one-dimensional treatment is ubiquitous is modeling human target acquisition performance. When collecting psychophysical data on vision, experimenters use horizontal or vertical sine-wave gratings or bars in the same manner that imagers are tested. The available visual data tends to fit with the separable model.

This section is summarized as follows: The one-dimensional treatment is useful for illustrating concepts and for comparing two or more functions. One-dimensional examples are used frequently in this tutorial. The one-dimensional treatment is widely used for target acquisition models in which data on observer vision is limited. However, be aware that one-dimensional plots can be misleading. They show one slice through an irregular volume and do not represent the entire MTF. Furthermore, one-dimensional plots exaggerate the amount of low-frequency content in the image. Lastly, although one-dimensional descriptions are useful in many cases, one-dimensional analysis leads to errors. No real imager has separable horizontal and vertical MTFs.

2.5 Imager Modulation Transfer Functions

2.5.1 Imager components

Diffraction MTF H_{diff} for a circular aperture is given by Eq. (2.15). D is the aperture diameter, and λ is wavelength; both dimensions are in meters. Spatial frequencies ξ and η are in cycles per milliradian (mrad^{-1}).

$$\chi = 1000\sqrt{\xi^2 + \eta^2} \quad \text{cycles per radian.}$$

$$H_{diff}(\xi,\eta) = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{\chi\lambda}{D}\right) - \frac{\chi\lambda}{D}\sqrt{1-\left(\frac{\chi\lambda}{D}\right)^2}\right] \quad \text{for } \frac{\chi\lambda}{D} < 1. \tag{2.15}$$

$$H_{diff}(\xi,\eta) = 0 \quad \text{for } \frac{\chi\lambda}{D} \geq 1.$$

Figure 2.16 shows diffraction MTF H_{block} versus radial frequency for circular apertures with circular blockages. The abscissa is a fraction of the diffraction

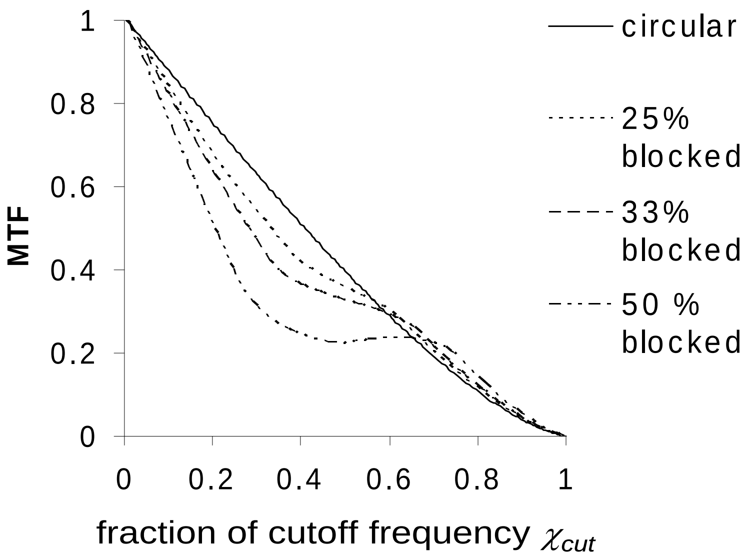


Figure 2.16 MTF of circular apertures that are partially blocked by circular obstructions. The obstructions are 0.25, 0.33, or 0.5 the diameter of the aperture. The abscissa is a fraction of the diffraction cutoff. The plots represent MTF versus radial frequency.

cutoff frequency, and the ordinate is MTF. The blockages are centered in the aperture and have diameters of 0.25, 0.33, or 0.50 of the aperture diameter. Note that very high frequencies are somewhat enhanced compared to an open aperture. However, the collecting area of the blocked apertures is smaller than the open aperture. For circular partially blocked apertures, two-dimensional MTF $H_{blocked}(\xi,\eta)$ is found by calculating a fraction of the diffraction cutoff χ_{cut} using Eq. (2.16). Then $H_{blocked}(\chi_{cut})$ is obtained from Fig. 2.16.

$$\chi_{cut} = 1000\lambda\sqrt{\xi^2+\eta^2}\Big/D. \tag{2.16}$$

The two-dimensional diffraction MTF H_{rect} for a rectangular aperture with horizontal dimension W meters and vertical dimension L meters is given by Eq. (2.17). This two-dimensional MTF is separable in Cartesian coordinates.

$$\begin{aligned}
\xi_{cut} &= 1000\lambda/W,\\
\eta_{cut} &= 1000\lambda/L,\\
H_W(\xi) &= (1-\xi/\xi_{cut}) \text{ for } \xi < \xi_{cut} \text{ and zero otherwise,}\\
H_L(\eta) &= (1-\eta/\eta_{cut}) \text{ for } \eta < \eta_{cut} \text{ and zero otherwise,}\\
H_{rect}(\xi,\eta) &= H_W(\xi)\,H_L(\eta).
\end{aligned} \tag{2.17}$$

Equation (2.18) gives the two-dimensional MTF H_{det} for rectangular detectors. Detector horizontal dimension d_h and vertical dimension d_v are in milliradians (mrad). Angular dimension in milliradians is found by dividing detector dimension in meters by focal length in meters and then multiplying by 1,000. θ defines the angle between the horizontal detector edge and the horizontal coordinate axis. Counterclockwise is positive θ. H_{det} is separable only if θ is a multiple of $\pi/2$.

$$\begin{aligned} H_{det} &= [\sin(\pi\xi' d_h)/(\pi\xi' d_h)][\sin(\pi\eta' d_v)/(\pi\eta' d_v)], \\ \text{where}\quad \xi' &= \xi\cos(\theta) - \eta\sin(\theta) \text{ and} \\ \eta' &= \xi\sin(\theta) + \eta\cos(\theta). \end{aligned} \tag{2.18}$$

Aberrations dominate optical MTF in the visible and near-infrared spectral regions and can be important in the mid- and longwave infrared. Glare can also be a problem with some types of optics. These MTFs are strongly field-angle dependent and tend to be symmetrical about only one axis. Analysis is done over a limited range of field angles where the blurs are fairly consistent. Generally, some combination of Gaussian and exponential MTF is used to represent aberrations.

MTF for a Gaussian blur is given by Eq. (2.19). The MTF is for separable spatial blur amplitude of $1/e$ at α_H mrad horizontally and α_V mrad vertically.

$$H_{Gaussian}(\xi,\eta) = e^{-\pi^2\xi^2\alpha_H^2}\, e^{-\pi^2\eta^2\alpha_V^2}. \tag{2.19}$$

MTF H_{exp} for the exponential blur in Eq. (2.20) is given by Eq. (2.21).

$$h_{\exp}(x,y) = e^{-\alpha\sqrt{x^2+y^2}}. \tag{2.20}$$

$$H_{\exp}(\xi,\eta) = \frac{\alpha^2}{(\alpha^2 + 4\pi^2\sqrt{\xi^2+\eta^2})}. \tag{2.21}$$

2.5.2 Line-of-sight jitter

This section discusses the MTF associated with line-of-sight (LOS) jitter. Formulas for vibration MTF are provided. Also, the effect of eye tracking on vibration blur is discussed and quantified.

The MTF associated with simple periodic motion is found in the Burle *Electro-Optics Handbook*. It is important to note, however, that the MTF of the blur caused by periodic motions at two or more different frequencies is not the sum of the individual MTFs. When multiple vibration frequencies are present, superposition holds for the position of the LOS, but not for the sum of intensities. This is illustrated in Fig. 2.17. The point of light at the top left is vibrating up and down one thousand times per second. The resulting blur is shown to the right. The point of light in the middle left is vibrating up and down one hundred times per second, and the resulting blur is shown to its right. At the bottom, the motions are summed. Note that the motions sum, and the intensities do not. The final blur depends on the summed positions.

Equation (2.22) gives the OTF for a periodic vibration with period t_P. The MTF is the magnitude of OTF. Equation (2.22) is similar to Eq. (8-12) in the Burle *Electro-Optics Handbook*, except that the Burle equation has only the sinusoid term. Equation (2.22) is for linear vibration along the horizontal axis. The angular offset x in milliradians is described as a function $f(t)$, where t is time in seconds. For linear motion along a different axis, the x motion is $f(t)\cos(\alpha)$, and the y motion is $f(t)\sin(\alpha)$. α is the angle between the direction of vibration and the horizontal axis. Positive α is counterclockwise.

$$OTF(\xi) = \frac{1}{t_P}\int_0^{t_P}\{\cos[2\pi\xi f(t)] + i\sin[2\pi\xi f(t)]\}dta\,. \tag{2.22}$$

Derivation of Eq. (2.22) is straightforward. At any instant in time, blur location x is equal to $f(t)$. Equation (2.23) gives the Fourier transform of the intensity of the point source at that instant in time. This is the Fourier transform of a delta function. Equation (2.22) integrates Eq. (2.23) over a period to obtain the OTF of the blur.

$$OTF_{point} = e^{-2\pi i\xi f(t)}\,. \tag{2.23}$$

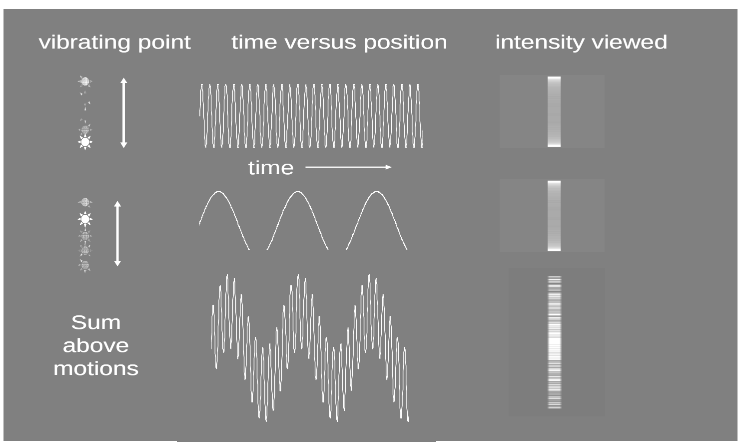


Figure 2.17 When the LOS is vibrating at two different frequencies, the MTF of the individual blurs do not sum. The positions sum, but the blur intensity distribution depends on the total motion.

If the vibration is not along a straight line, then Eq. (2.24) is used to find $OTF(\xi,\eta)$:

$$OTF(\xi,\eta) = \frac{1}{t_P}\int_0^{t_P} \{\exp[-i2\pi\xi f_x(t)]\exp[-i2\pi\eta f_y(t)]\}dt\,. \qquad (2.24)$$

Note that Eqs. (2.22) and (2.23) require knowledge of $f(t)$. The vibration power spectrum is not sufficient because phase information is also needed.

If only the magnitude of the Fourier transform of $f(t)$ is known, Eqs. (2.22) and (2.23) cannot be used. In that case, we employ the Gaussian approximation that is already in many performance models. We assume that the Fourier transform $F(\xi)$ is the sum of discrete frequencies and that only the magnitude of each frequency component is known.

$$F(\xi) = \sum_{n=1}^{N} a_n \delta\left(\xi - \frac{n}{t_P}\right), \qquad (2.25)$$

where a_n is the magnitude of the n^{th} of N components. δ is the Dirac delta function. We calculate the root-mean-square (rms) blur radius β using Eq. (2.26):

$$\beta = \sqrt{\frac{\sum_n a_n^2}{2N}}\,. \qquad (2.26)$$

Equation (2.27) approximates the vibration blur MTF. Figure 2.18 shows an example; this example uses the sum of ten sine waves with random amplitudes and phases. The Gaussian is a good approximation at low spatial frequencies for complex spectra. Figure 2.19 shows the Gaussian approximation to a single sine-

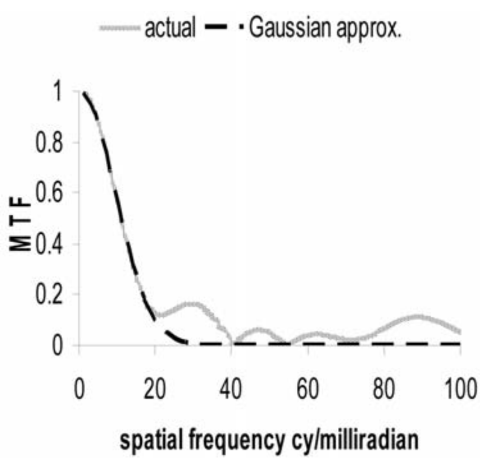


Figure 2.18 Example showing Gaussian approximation to the sum of ten sinusoids with random amplitude and phase.

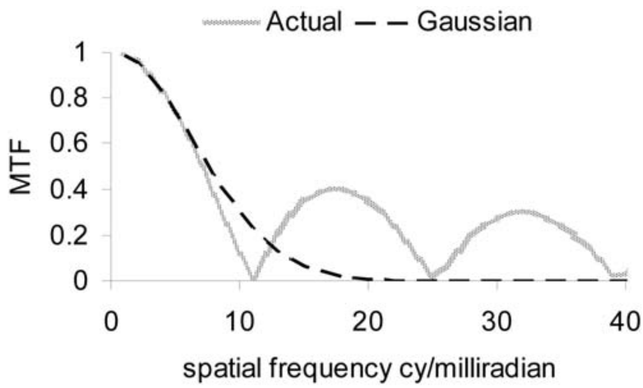


Figure 2.19 Gaussian approximation (dashed line) to the Bessel MTF of a single sine wave. Approximation is fairly accurate to the first zero of the Bessel MTF.

wave vibration. The approximation is good to the first zero of the Bessel function. The examples here support the decision to use the Gaussian approximation.

$$MTF_{Gaussian} = \exp(-\beta^2\, \xi^2)\,. \tag{2.27}$$

2.5.2.1 Reduction of line-of-sight jitter by eye tracking

For predictable motion, observers have full visual acuity for crossing velocities of up to 30 deg per sec or more. Translation of the image does not cause blur at the eye if the observer has time to initiate ocular tracking. Only motion not tracked by the eye causes a blur.

The oculomotor system is a low-pass filter with a corner frequency of 1.5 Hz. For predictable stimuli, phase lags are minimal for frequencies below about 2 Hz. Figures 2.20 and 2.21 show the gain and phase lag for the eye-tracking simple sinusoids. In both figures, the abscissa is frequency in Hertz. In Fig. 2.20, the ordinate is gain in decibels; in Fig. 2.21, the ordinate is phase lag in degrees. The data are taken from experimenters A, B, and D, as cited in Boff and Lincoln's *Engineering Data Compendium*. Figure 2.22 shows phase lag while tracking complex motion; gain does not change substantially from that shown in Fig. 2.20. The small dashed line shows results for the sum of several low-frequency sine waves, and the long dashed line shows results for bandlimited Gaussian random motion.

The oculomotor system acts like a 0.33-sec integrator. For predictable motion, the integration looks both forward and backward in time. When tracking complex motion, the integration looks only backward in time. The gray lines in Figs. 2.20 and 2.21 show gain and phase lag, respectively, of a 0.33-sec integrator. The filter output at time t is the integral of the input signal from t – 0.165 to t + 0.165 sec. In Figs. 2.20 and 2.21, notice that the model behavior matches measured data rather well.

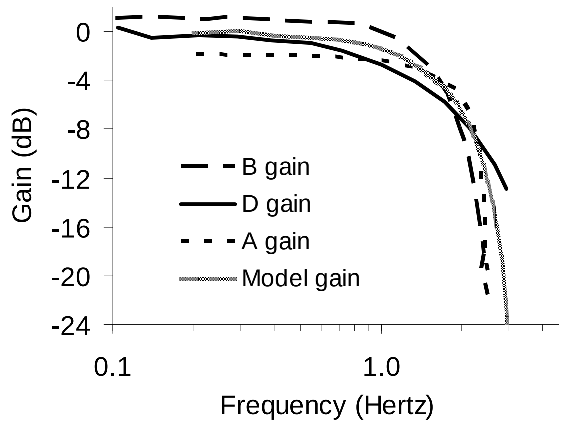


Figure 2.20 Plot showing gain of the oculomotor system while tracking sine waves. The data are from various sources.

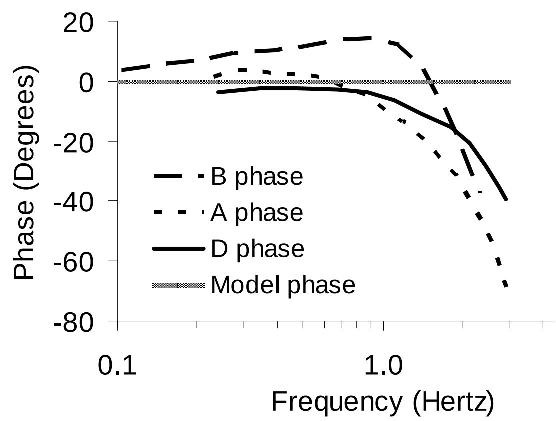


Figure 2.21 Plot showing phase lag when tracking single sine waves. Since the motion is predictable, phase lag is either zero or slightly leads the track point.

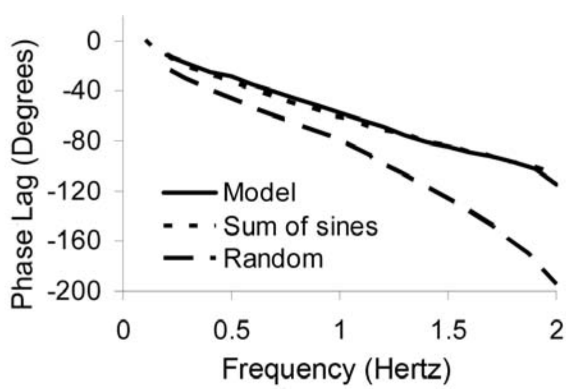


Figure 2.22 Plot showing phase lag when tracking the sum of multiple low-frequency sine waves (small dashes) and bandlimited random motion (large dashes). The eye lags the target because the motion is not predictable.

When tracking complex motion, the filter output is the integral of input motion from time t – 0.33 sec to t sec. The gain at each frequency is not affected by the change from noncausal to causal integration. Figure 2.22 shows that the phase lag associated with causal integration matches measured data when stimulus motion is not predictable.

Equations (2.28) and (2.29) give the MTF of oculomotor-filtered LOS jitter. Equation (2.28) relates perceived jitter $g(t)$ to physical LOS jitter $f(t)$ when the eye tracks predictable low-frequency motion. We use Eq. (2.29) when low-frequency motion is complex. Equation (2.30) gives vibration OTF for perceived jitter $g(t)$.

$$g(t) = f(t) - \tfrac{1}{0.33}\int_{t-0.165}^{t+0.165} f(t)\,dt\,, \tag{2.28}$$

$$g(t) = f(t) - \tfrac{1}{0.33}\int_{t-0.33}^{t} f(t)\,dt\,, \tag{2.29}$$

$$OTF(\xi) = \frac{1}{t_P}\int_0^{t_P} \{\cos[2\pi\xi g(t)] + i\sin[2\pi\xi g(t)]\}dt\,. \tag{2.30}$$

2.5.2.2 Effect of temporal sampling on line-of-sight jitter

Temporal sampling by the imager is usually not important. However, stroboscopic effects can occur that lead to large errors when using Eqs. (2.28) through (2.30). This section describes the situation in which temporal sampling is important.

Imager frames typically occur 50 or 60 times per second. The eye incorporates spatial information from multiple frames into a high-resolution image. The long time constant associated with higher-order vision results in smooth apparent motion and a temporal response to jitter that is much longer than a frame time. Due to limitations in electron-well capacity of FPAs, however, only a fraction of a frame time is typically used to integrate signal. If no synchronism exists between the LOS jitter and the frame capture sequence, then vibration MTF is not affected. The detector randomly samples jitter position, and temporal sampling is not important.

Synchronism between LOS jitter and detector timing does affect vibration MTF. Only the jitter positions that occur during detector integration contribute to vibration MTF. If the full range of LOS position is not sampled, then vibration MTF changes significantly. Figure 2.23 shows a detector integration sequence $p(t)$ with frame time T_d and integration time W. Equation (2.31) predicts stroboscopic effects when the detector integration sequence is synchronized to LOS jitter.

$$OTF(\xi) = \frac{T_d}{W t_P}\int_0^{t_P} \{\cos[2\pi\xi f(t)\,p(t)] + i\sin[2\pi\xi f(t)\,p(t)]\}dt. \tag{2.31}$$

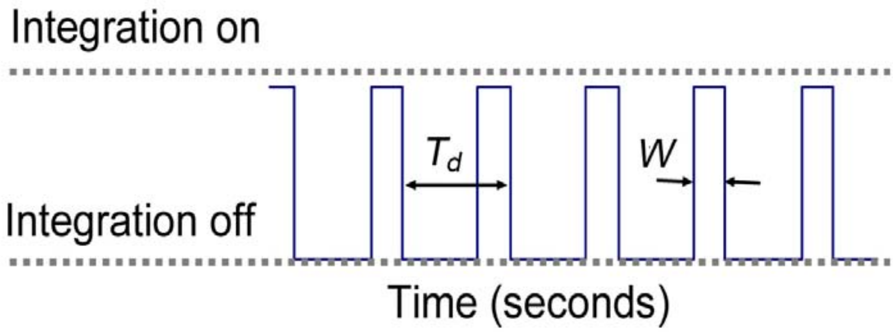


Figure 2.23 Detector integration time sequence $p(t)$.

Notice that Eq. (2.31) uses $f(t)$, the unfiltered LOS motion, rather than $g(t)$, the LOS motion as filtered by the oculomotor system. Including the oculomotor filter might corrupt the calculation of stroboscopic effects; this spatial filter does not exist in the sensing hardware. If an oscillation with a spatial frequency below 2 Hz is present, then Eq. (2.31) yields pessimistic results because eye tracking is not considered.

2.5.3 Electronic stabilization

Equations (2.30) and (2.31) apply to traditional stabilization systems in which the sensor output is displayed without electronic correction. However, if a residual stabilization error signal is available in real time, then electronic stabilization (E-stab) is possible. The stabilization error signal might be available from inertial sensors on the gimbal, or the error signal might be derived from image processing. E-stab uses the error signal to calculate how much the image should move on the display in order to cancel LOS jitter. Display interpolation provides subpixel registration of the image from frame to frame.

Equations (2.30) and (2.31) do not apply to E-stab because the entire jitter period is not sampled. One blur associated with E-stab is due to target jitter during detector integration. Assume M detector samples during a vibration period t_p. Equation (2.32) gives the OTF_{E_det} associated with sample N.

$$OTF_{E\_det}(\xi, N) = \frac{1}{W}\int_{NT_d}^{NT_d+W}\{\cos[2\pi\xi(f(t)-f(NT_d))]+i\sin[2\pi\xi(f(t)-f(NT_d))]\}dt. \tag{2.32}$$

Obviously, OTF_{E_det} depends on the relationship between $f(t)$ and a particular detector sample. It is impossible to know *a priori* whether to use the best MTF, the worst MTF, or the average of the two. The average is selected. The OTF associated with the detector integration is OTF_{det}:

$$OTF_{det}(\xi) = \frac{1}{M+1} \sum_{N=0}^{M} OTF_{E\_det}(\xi, N)\,. \tag{2.33}$$

A probability density function $p(x)$ describes residual jitter at the display. $p(x)$ is the probability that the target offset is x. Equations (2.34) and (2.35) give the $OTF_{residual}$ associated with $p(x)$:

$$\int_{-\infty}^{\infty} p(x)\,dx = 1. \tag{2.34}$$

$$OTF_{residual} = \int_{-\infty}^{\infty} p(x) \exp(-i2\pi\xi x)\,dx. \tag{2.35}$$

It is possible that frame-to-frame correlation exists in the stabilization error function. In that case, an error probability density function is not adequate to calculate the vibration MTF associated with residual jitter at the display. Instead, we generate an error function $e(t)$ based on frame-to-frame behavior. The error function is continuous because interpolation is used between discrete errors at each frame time. The eye perceives smooth apparent motion when viewing the quickly strobed images on the display. The oculomotor system filters the physical error function $e(t)$ to create the perceived error function $e'(t)$.

We use Eq. (2.36) when simple low-frequency oscillations are part of the display jitter:

$$e'(t) = e(t) - \frac{1}{0.33} \int_{t-0.165}^{t+0.165} e(t)\,dt\,. \tag{2.36}$$

We use Eq. (2.37) when the low-frequency component is random or consists of many low-frequency sinusoids:

$$e'(t) = e(t) - \frac{1}{0.33} \int_{t-0.33}^{t} e(t)\,dt\,. \tag{2.37}$$

The filtered error function $e'(t)$ is then used to calculate display jitter $OTF_{residual}$; see Eq. (2.38). To develop a statistically meaningful answer, generate $e(t)$ several times and use an average $OTF_{residual}$. The total OTF associated with E-stab ($OTF_{E\text{-}stab}$) is then the product of detector blur and residual blur as given by Eq. (2.39).

$$OTF_{residual}(\xi) = \frac{1}{t_P} \int_{0}^{t_P} \{\cos[2\pi\xi e'(t)] + i \sin[2\pi\xi e'(t)]\} dt. \tag{2.38}$$

$$OTF_{E\text{-}stab} = OTF_{det}\, OTF_{residual}\,. \tag{2.39}$$

2.5.4 Motion blur

The blurs discussed in Sections 2.5.2 and 2.5.3 result from an inability of the eye to track the object. During scene-to-sensor motion, blur also arises in the hardware. Motion degrades both pre-MTF and post-MTF. Preblur arises from scene motion across the detector. Postblur arises from image motion across the display.

If the scene moves across the detector while signal is integrated, a blur occurs. Equation 2.40 gives the MTF H_{trans} associated with the scene moving across the detector during signal integration. Detector integration time is t_d, and φ is angular rate in milliradians per second. α is the angle of motion relative to the horizontal axis.

$$H_{trans}(\xi,\eta) = \frac{\sin[\pi\varphi t_d \xi \cos(\alpha)]}{[\pi\varphi t_d \xi \cos(\alpha)]} \frac{\sin[\pi\varphi t_d \eta \sin(\alpha)]}{[\pi\varphi t_d \eta \sin(\alpha)]}. \tag{2.40}$$

The eye has full visual acuity up to angular rates of 30 deg per sec and some resolution capability to rates of about 60 deg per sec. Normal video rate in the U.S. is 30 frames and 60 fields per sec. This means that motion of 30 deg per sec creates an angular offset between video fields of a half deg. A half-degree offset is easily visible to the human eye. The perception of smooth motion when viewing video is created by the behavior of human vision.

When the eye tracks a moving object on a video monitor, smooth ocular pursuit occurs. The eye tracks at a uniform angular rate, even though the video image is presented briefly each 1/60 of a second. With a cathode ray tube (CRT), the image decays quickly. However, with many flat panel displays, the image is optically present through all or much of the field time. This means that the image is present as the eye moves to the next expected position of the object.

Blur occurs at the display in exactly the same way as it occurs at the detector. In this case, the eye is the detector. The image is not moving smoothly over the display, but the eye is. That causes a noticeable blur if the image presentation time is extended beyond a few milliseconds.

With a CRT, only the exponential decay of the phosphor is present. Generally, a CRT display has very good (very high) motion MTF because the display phosphors are very quick. The slowest phosphors are generally associated with color displays, and decay time is between 2 and 3 milliseconds (msec). This causes little blur.

Active liquid crystal displays (LCDs) have circuitry to hold each pixel on throughout the field time. Motion MTF H_{LCD} is modeled using Eq. (2.40) with τ_{LCD} equal to the pixel hold time. In this case, φ_d is angular rate to the eye as the target moves across the display. φ_d equals φ multiplied by imager magnification.

$$H_{trans}(\xi, \eta) = \frac{\sin[\pi \varphi_d \ \tau_{LCD} \xi \cos(\alpha)]}{[\pi \varphi_d \ \tau_{LCD} \xi \cos(\alpha)]} \frac{\sin[\pi \varphi_d \ \tau_{LCD} \eta \sin(\alpha)]}{[\pi \varphi_d \ \tau_{LCD} \eta \sin(\alpha)]}. \quad (2.41)$$

2.5.5 Field replication

Certain imagers collect imagery 30 times per sec but display at 60 Hz to avoid flicker. This is done by replicating the display fields. That is, each image is displayed twice, spaced at 1/60-sec intervals. This display technique creates motion artifacts. The artifacts might appear as a double image or might just appear as a blur.

Certain types of uncooled imagers use this field replication. Also, second-generation scanned sampled imagers use this technique when in the 30-Hz mode. All second-generation imagers have both 30- and 60-Hz modes. The 60-Hz mode is used during search or any time the scene moves rapidly through the field of view. The 60-Hz mode avoids field replication and therefore avoids the distracting motion artifacts.

Field replication is generally seen as a double image. However, many people perceive the artifact as a blur. Blur tends to be less bothersome than a double image, but we have no method of predicting the impact of a double image on performance. Field replication is therefore optimistically modeled as a blur. The motion MTF $H_{replicate}$ associated with field replication is given by Eq. (2.42) using τ_{field} equal to a field time.

$$H_{replicate}(\xi,\eta) = \frac{\sin[\pi \varphi_d \ \tau_{field} \xi \cos(\alpha)]}{[\pi \varphi_d \ \tau_{field} \xi \cos(\alpha)]} \frac{\sin[\pi \varphi_d \ \tau_{field} \eta \sin(\alpha)]}{[\pi \varphi_d \ \tau_{field} \eta \sin(\alpha)]}. \quad (2.42)$$

2.5.6 Analog electronic filters

Electronic circuits are different from other imager components. They are *causal*; an output cannot precede the input that causes it. The impulse response for a circuit must be one sided. When a lens blurs a point source of light, the blur spreads in all directions. When an analog circuit blurs a point source, the blur spreads in one direction.

In Figure 2.24(a), a linear array of detectors scans the image of a bar. The detector signals are amplified using analog electronics. The signal from a single detector is shown as a dashed line in Fig. 2.25. The solid line shows a theoretical output where the amplifier is perfect. The real signal is delayed in time. Edge intensities rise and fall exponentially rather than instantaneously. The image of the bar is shown in Fig. 2.24(b). The image is blurred horizontally, and all blur is to the right. The blur is not symmetrical. An unsymmetrical blur cannot be represented by a real MTF.

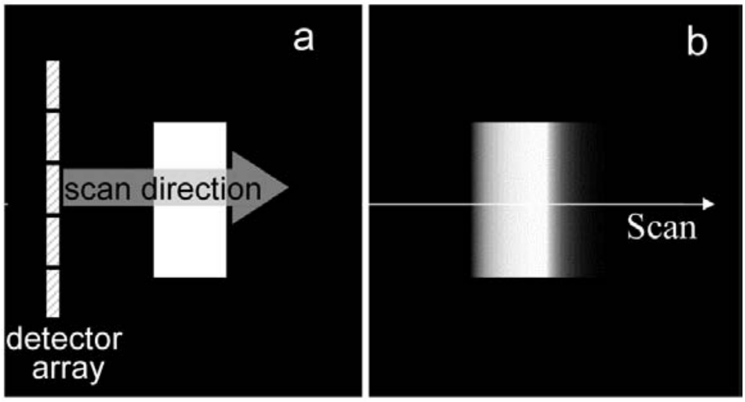


Figure 2.24 (a) A detector array is scanned over the image of a bar pattern. (b) The image is blurred horizontally due to electronic filtering of the detector signals. All electronic blur is to the right. The blur is not symmetrical in space.

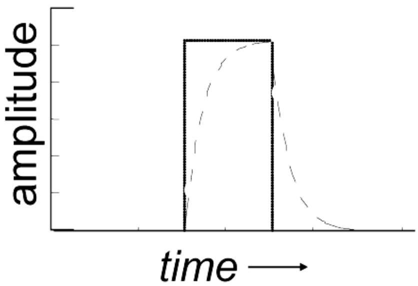


Figure 2.25 Dashed line shows detector signal versus time. Solid line shows detector signal without any delay in the electronics. All blur is forward in time; therefore, all blur is to the right in Fig. 2.24.

2.5.7 Display MTF

Display MTF is the Fourier transform of the pixel shape. Typically, pixels are either approximately rectangular or Gaussian. The MTFs are given by Eqs. (2.18) and (2.19), respectively. To use these formulas, however, the angle subtended by the display pixel in object space must be known.

Most models represent all spatial frequencies in object space. That is, angles are defined by target range and distance from the imager. The observer views the display from only a fraction-of-a-meter distance. There are several ways to convert linear pixel dimensions into object-space angles.

System magnification (SMAG) is illustrated in Fig. 2.26. The target subtends an angle at the imager equal to target size divided by range. In this case, angle is in radians. The image of the target subtends a larger angle at the observer's eye. The angle subtended by the target image at the eye divided by the angle subtended by the target at the imager is the SMAG. Instead of selecting an arbitrary target, the vertical field of view can be used to define SMAG. The value of SMAG is the same when calculated either way.

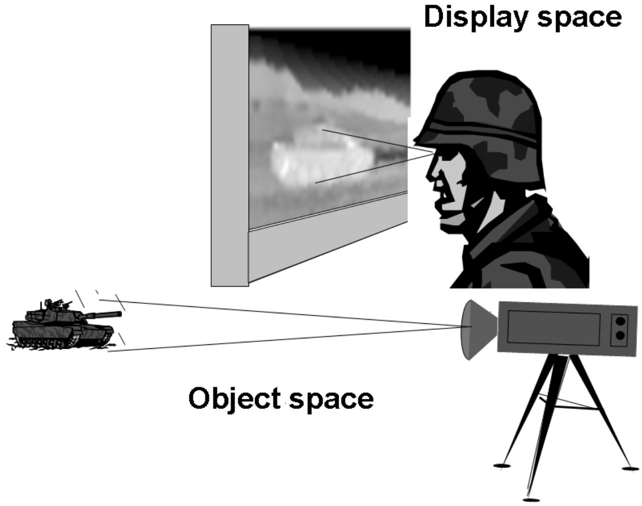


Figure 2.26 System magnification (SMAG) is the angle subtended by the target image at the eye divided by the angle subtended by the target at the imager.

$$SMAG = \frac{\text{vertical size of picture} \big/ \text{viewing distance}}{\text{vertical field of view in radians}} . \tag{2.43}$$

SMAG is used to relate pixel angular dimension x_{obs} at the observer's eye to angular dimension x in object space. Angular frequency in display space ξ_{obs} is also related to angular frequency ξ in object space using SMAG. In this case, angle is generally expressed in mrad, and frequency in mrad^{-1}.

$$x = \frac{x_{obs}}{SMAG} . \tag{2.44}$$

$$\xi = \xi_{obs}\, SMAG. \tag{2.45}$$

Probably the easiest way to relate linear pixel size to milliradians subtense in object space is to use the formula in Eq. (2.46). In this formula, pix_h is horizontal pixel size, and h_d is display height. By display height we mean the dimension of the active picture. Both dimensions are in centimeters. VFOV is the imager vertical field of view in degrees. Multiplying by 17.45 converts VFOV into milliradians. The angular subtense p_h of the pixel in object space is the fraction pix_h/h_d of the field of view in milliradians.

$$p_h = VFOV\,17.25\,pix_h \,/\, h_d\,. \tag{2.46}$$

If the pixel is rectangular and oriented along the horizontal axis, then the horizontal MTF H_{pix} is obtained by substituting p_h for d_h in Eq. (2.18). The vertical MTF and the MTF for other shapes are found in a similar fashion.

$$H_{pix} = \frac{\sin(\pi\,\xi\,p_h)}{(\pi\,\xi\,p_h)}. \tag{2.47}$$

Bibliography

Boff, K. R. and J. E. Lincoln, *Engineering Data Compendium: Human Perception and Performance*, Human Engineering Division, Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, OH (1988).

Boff, K. R. and J. E. Lincoln, *Engineering Data Compendium: Human Perception and Performance*, Vol. 1, Sections 1.617 and 940, Human Engineering Division, Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, OH (1988).

Boreman, G. B., *Basic Electro-Optics for Electrical Engineers*, SPIE Press, Bellingham, WA (1998).

Driggers, R., P. Cox, and T. Edwards, *Introduction to Infrared and Electro-Optical Systems*, Artech House, Norwood, MA (1999).

Electro-Optics Handbook, TP-135, 117–120, Burle Industries, Inc., Lancaster, PA (1974).

Goodman, Joseph W., *Introduction to Fourier Optics*, Roberts & Co. Publishers, Greenwood Village, CO (2005).

Gaskill, J., *Linear Systems, Fourier Transforms, and Optics*, Wiley, New York, (1978).

O'Neill, E., *Introduction to Statistical Optics*, Addison-Wesley, Reading, MA (1963).

Overington, I., *Vision and Acquisition*, Crane, Russak, New York (1976).

Pinson, L., *Electro-Optics*, Wiley, New York (1985).

Vollmerhausen, R. and K. Mayes, "The effect of helicopter vibration on the viewing of panel mounted displays," *SID Digest* **XXVI**, 651–654 (1995).

Vollmerhausen, R. and T. Bui, "The effect of sensor field replication on displayed imagery," *SID Digest* **XXV**, 667–670 (1995).

Chapter 3
Sampled Imager Response Function

This chapter describes the sampled imager response (SIR) function. The SIR has two parts: it quantifies the transfer response of a sampled imager, and it quantifies aliasing.

The procedure for finding the SIR function parallels the procedure for finding the transfer response for a nonsampled imager. The SIR function is the Fourier transform of the imager point spread function (psf). As discussed in Chapter 1, the psf of a sampled imager is not shift invariant; it varies depending on sample phase. For this reason, the SIR function is more complicated than the transfer function. The SIR function contains information about both the transfer response of the system and about aliasing.

The SIR function depends on the sensor presample MTF, the sample spacing, and the postsample or display MTF. These imager characteristics are known to the design engineer or system's analyst. The SIR function does not depend on sample values; it characterizes the process by which the samples are taken and displayed.

Since the sampling artifacts produced by an imager depend on the scene being imaged, one might question a mathematical process that quantifies sampling artifacts without including an explicit description of the scene. In that regard, we rely on assumptions identical to those used for nonsampled imagers.

MTF is used to characterize a nonsampled imager. MTF is the Fourier transform of the psf. It describes the blur produced in the image by a point in the scene. The importance of good MTF is not established until the frequency content of the scene is known. The impact or importance of sensor blur depends on scene frequency content.

Nonetheless, MTF has proven to be a good indicator of the overall utility of an imager. The ability of an imager to resolve scene detail is important because of the possibilities it provides. Good MTF is not needed for every task or on every occasion. But the characterization of the imager's response to a point source provides a good indication of the quality of images that are expected in a variety of circumstances.

A similar logic applies to sampled imagers. We cannot know how each detail in the scene will be corrupted by the sampling process until the scene itself is specified. However, the *tendency* of the imager to produce visible display raster or corrupt scene details can be characterized.

In this chapter, a response function for sampled imagers is derived by examining the image formed on the display by a point source of light in the scene. The SIR function provides a quantitative method to characterize both the quality of the sampled imager's transfer response and its tendency to generate sampling artifacts or aliasing.

The SIR is developed in two parts. In this chapter, the scene presents a constant stimulus in angle space. In Fig. 3.1, the clock grows in size in proportion to the distance to the imager. The image on the FPA is not range dependent. The clock is the same size on the display regardless of range. Since the imager samples in angle space, using a constant angular stimulus lets us focus on sampling behavior and ignore radiometry and other complications.

Figure 3.2 shows the behavior of a real object. An object of fixed linear size grows smaller in angle space as range increases. The displayed image gets smaller at long range. The effect of constant linear size is discussed in Chapter 8.

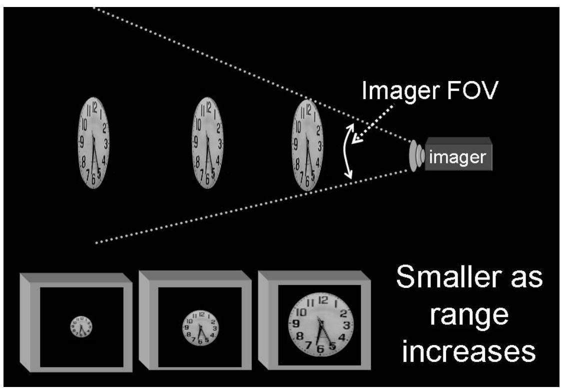


Figure 3.1 Illustration of the scene presenting a constant stimulus in angle space; the clock face grows with range. The size of an object on the display does not depend on range.

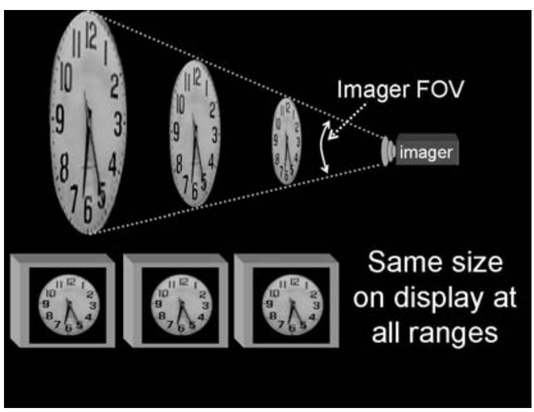


Figure 3.2 Illustration of the actual behavior of real objects. Linear dimension is constant, and the displayed image gets smaller as range increases.

3.1 Fourier Transform of a Sampled Image

Figure 3.3(a) shows a picture of a clock. Figure 3.3(b) shows an intensity trace taken along the white line drawn in Fig. 3.3(a). Figure 3.4(a) shows the clock blurred by a lens, and Fig. 3.4(b) shows the resulting intensity trace. It is the blurred intensity pattern in Fig. 3.4(b) that is sampled. In a real sensor, the image would be blurred by the optics and detectors and perhaps by other factors such as line of sight jitter or motion blur. In this conceptual example, the optical blur is a stand-in for all of the possible presample blurs.

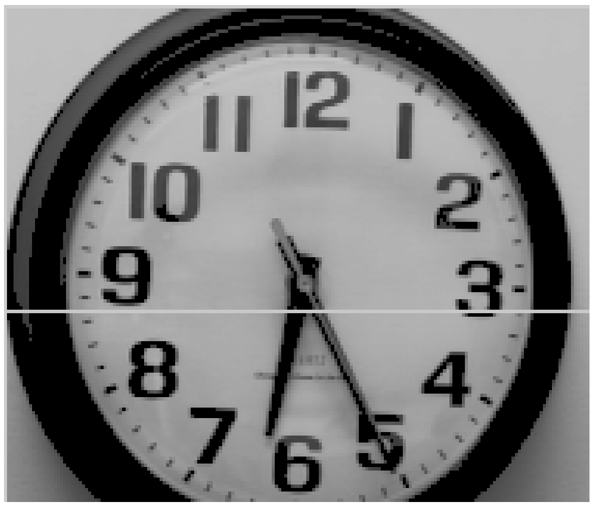


Figure 3.3(a) Clock picture before it is blurred by a lens. The plot in Fig. 3.3(b) shows intensity along the white line in Fig. 3.3(a).

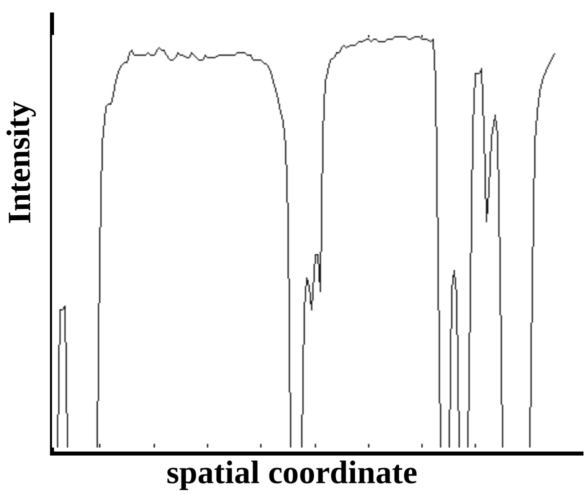


Figure 3.3(b) Plot of the intensity along the white line shown in Fig. 3.3(a). Notice the high frequency content indicated by sharply rising and falling lines and significant intensity variation over a small spatial interval.

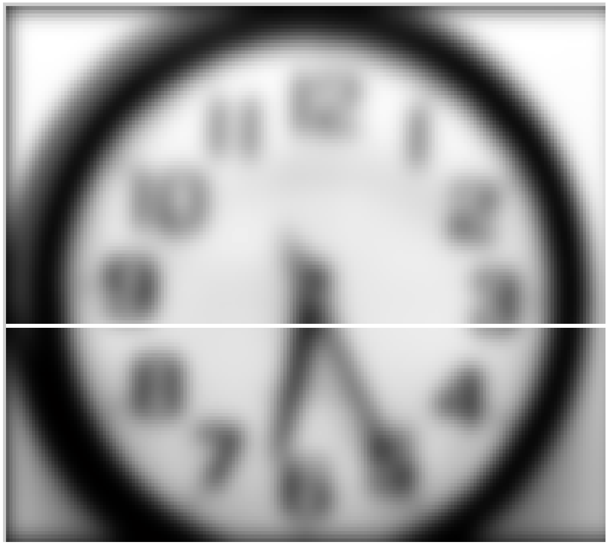


Figure 3.4(a) Clock imaged and blurred by a lens. Intensity along the white line is shown in Fig. 3.4(b).

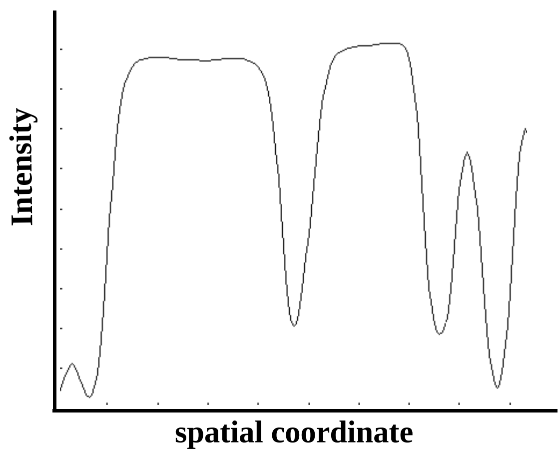


Figure 3.4(b) Intensity plot of the blurred image taken along the white line shown in Fig. 3.4(a). Notice that intensity varies less in each spatial interval; the intensity rise and fall are less sharp.

Figure 3.5 shows the presampled image $f(x)$ sampled in the x direction with spacing X mrad. The image is sampled throughout space. Only a finite number of samples are nonzero because the image itself is finite. In the derivation below, $f(x)$ represents only the sampled part of the scene, and $F(\xi)$ is the Fourier transform of that windowed $f(x)$.

The $x = 0$ origin is a sample point. The origin of the scene is offset by a distance d; this allows the sample phase to be explicitly included in the derivation. We want the mathematics to tell us what happens when the presample image is moved with respect to the sample points.

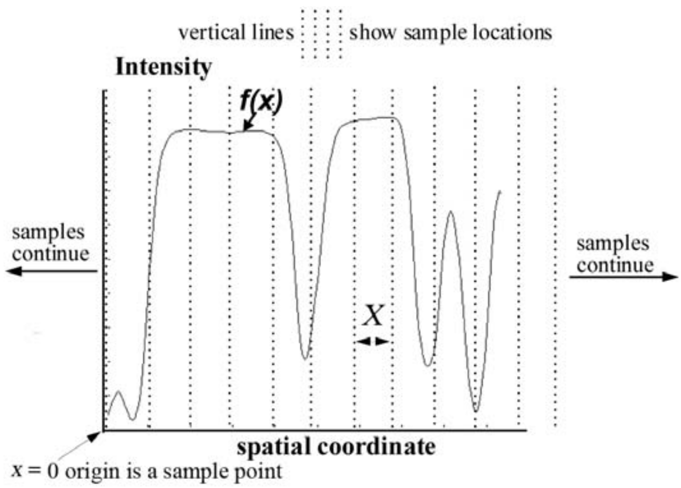


Figure 3.5 Blurred image $f(x)$ is sampled at points indicated by dashed vertical lines. Sample spacing is X mrad. Sampling continues throughout all space. The $x = 0$ origin is a sample point.

The presampled image $f(x)$ is formed by convolving the psf of the lens $h(x)$ with the scene $s_{cn}(x)$.

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} h(x-x')s_{cn}(x'-d)dx' \\ f(x) &= h(x)*s_{cn}(x-d), \end{aligned} \tag{3.1}$$

where * indicates convolution.

The image is sampled by multiplying $f(x)$ by a comb of delta functions. The displayed image is then reconstructed by convolving the display pixel shape with the delta function samples. This process is described in Section 1.5. Let $p_{ix}(x)$ be the intensity distribution associated with a single display pixel and $P_{ix}(\xi)$ be its Fourier transform. The displayed image is represented by

$$i_{dsp}(x) = \left\{ [h(x)*s_{cn}(x-d)] \sum_{n=-\infty}^{\infty} \delta(x-nX) \right\} * p_{ix}(x). \tag{3.2}$$

Because a convolution in the space domain is a multiplication in the frequency domain, Eq. (3.3) is equivalent to Eq. (3.2).

$$I_{dsp}(\xi) = \left[H(\xi)S_{cn}(\xi)e^{-j2\pi\xi d} * \sum_{n=-\infty}^{\infty} \delta(\xi - n\nu) \right]]P_{ix}(\xi), \tag{3.3}$$

where ν is the sample frequency ($1/X$) with units of cycles per milliradian (mrad^{-1}).

$$I_{dsp}(\xi) = P_{ix}(\xi) \sum_{n=-\infty}^{\infty} H(\xi - n\nu) S_{cn}(\xi - n\nu) e^{-j2\pi(\xi - n\nu)d}$$
$$I_{dsp}(\xi) = P_{ix}(\xi) \sum_{n=-\infty}^{\infty} F(\xi - n\nu)\, e^{-j2\pi(\xi - n\nu)d}. \tag{3.4}$$

$F(\xi)$ equals $H(\xi)\ S_{cn}(\xi)$ because $f(x)$ is defined by Eq. (3.1) to be the convolution of $h(x)$ with $s_{cn}(x)$. The Fourier transform of the presampled image is the MTF of the sensor multiplied by the Fourier transform of the scene. The displayed frequency spectrum is illustrated in Fig. 3.6. The sampling process replicates the presample image spectrum $F(\xi)$ at each integer multiple of the sample frequency. The displayed spectrum is the product of the replicated spectra multiplied by the display pixel MTF.

The Fourier transform of the sampling artifacts is the product of the display MTF and all of the replicated spectra except the one located at the frequency origin (the one for n = 0). Figure 3.6 shows only the amplitude of the spectra; phase is not shown. Each replicated spectrum, however, varies in phase from the adjacent spectrum by the sample phase increment $2\pi d/X$. $\Delta\phi$ is the change in phase between replicas.

$$\begin{aligned} \Delta\phi &= 2\pi\left[\xi - (n-1)\nu\right]d - 2\pi(\xi - n\nu)d. \\ \Delta\phi &= 2\pi\nu d. \\ \Delta\phi &= 2\pi d\ /\ X. \end{aligned} \tag{3.5}$$

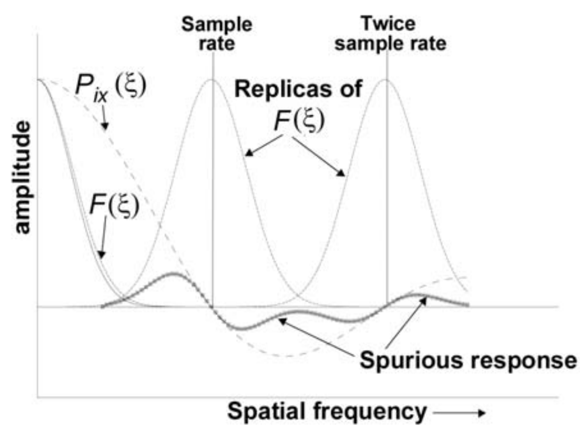


Figure 3.6 The presample image spectrum $F(\xi)$ is replicated at all integer multiples of the sample frequency. The displayed image spectrum is the product of the display MTF $P_{ix}(\xi)$ and all of the replicated spectra including the original $F(\xi)$. The Fourier transform of the sampling artifacts is that part of the image spectrum that is found by multiplying the display MTF by all of the replicated spectra except the one located at the origin. This product is called the spurious response of the sampled imager.

The magnitude of each replicated spectrum does not depend on sample phase. However, as sample phase d changes, the phase relationship between the replicated spectrums varies, and that changes the image. This is why the image is sample-phase dependent. Artifacts appear, disappear, or change size or position because of variation in sample phase.

In Eq. (3.4), the sampled signal contains the same transfer response as a nonsampled imager. That is, the n equal zero term is $Pix(\xi)H(\xi)S_{cn}(\xi)$, which is the product of the sensor and display MTFs multiplied by the Fourier transform of the scene. This term represents the desired frequency content in the image.

In this tutorial, the entire portion of the image spectrum that results from sampling, other than the baseband spectrum itself, is called *spurious response.* The spurious response is the sum of all of the terms in Eq. (3.4), except for the $n = 0$ term. The spurious response is the Fourier transform of the sampling artifacts in the image.

The term spurious response is a synonym for *aliasing*. Aliasing correctly refers to the entire portion of the displayed image spectrum that results from the replicated or "aliased" spectra. We have defined another term for the aliased response because using the word aliasing to represent the entire spurious spectrum that results from sampling could be confusing.

It has become common practice among engineers to use the term "aliasing" to refer only to overlap in the frequency domain between the sample-generated replica spectra and the baseband spectrum. Ideally, aliasing that does not overlap the baseband spectrum [does not overlap $F(\xi)$] can be removed by postsample filtering. However, even near-ideal filters are difficult to implement, and removing such artifacts as visible raster from an image often entails degrading the baseband response. In most sampled imagers, aliased frequency content can and does exist at frequencies above the half-sample frequency.

In order to avoid confusion, the entire portion of the image spectrum that results from sampling, other than the baseband spectrum, is referred to as *spurious response.* In frequency space, the spurious response is the Fourier transform of the sampling artifacts.

3.2 The Sampled Imager Response Function

The SIR function $R_{sp}(\xi)$ for a sampled imager is found by examining the impulse response of the system. The SIR is the Fourier transform of the image formed on the display by a point source in the scene. The procedure is identical to that used with nonsampled systems, but the results differ because sampled systems are not shift invariant.

The function sampled is $h(x)$, the psf of the presampled image. Therefore, based on Eq. (3.4),

$$R_{sp}(\xi) = P_{ix}(\xi) \sum_{n=-\infty}^{\infty} H(\xi - n\nu) e^{-j2\pi(\xi - n\nu)d} . \tag{3.6}$$

The replicas of $H(\xi)$ that are centered at two or more times the sample frequency are generally filtered out by the display and eyeball MTF because of their high frequency. In most practical systems, only the replicas adjacent to the baseband with $n = \pm 1$ contribute visible sampling artifacts. This typical situation is represented in Fig. 3.7. Also, the $e^{-j2\pi\xi d}$ phase factor is common between terms. This phase factor is dropped because it provides no insight into system response. Equation (3.6) is simplified by writing out the sum and dropping the terms with $n > 1$.

$$
\begin{aligned}
R_{sp}(\xi) &\approx P_{ix}(\xi)\sum_{n=-1}^{1} H(\xi - n\nu)e^{j2\pi n\nu d}, \\
R_{sp}(\xi) &\approx P_{ix}(\xi)H(\xi) + P_{ix}(\xi)H(\xi-\nu)e^{j2\pi\nu d} + P_{ix}(\xi)H(\xi+\nu)e^{-j2\pi\nu d}, \\
R_{sp}(\xi) &\approx P_{ix}(\xi)H(\xi) + P_{ix}(\xi)H(\xi-\nu)e^{j\phi} + P_{ix}(\xi)H(\xi+\nu)e^{-j\phi},
\end{aligned}
\tag{3.7}
$$

where the sample phase ϕ equals $2\pi d/X$. In some cases, the terms with $n = \pm 2$ or higher contribute to the spurious response. In those cases, Eq. (3.7) needs to be modified to include those terms.

The SIR function $R_{sp}(\xi)$ has two parts: the transfer term and the spurious response terms. The first term in Eq. (3.7) is the transfer response of the imager. This transfer response does not depend on sample spacing. It is the only term that remains for small sample spacing. A sampled imager has the same transfer function as a nonsampled (that is, a very well-sampled) imager.

However, a sampled imager always has the additional response terms, which we refer to as *spurious response*. These spurious response terms in Eq. (3.7) are filtered by the display MTF $P_{ix}(\xi)$ in the same way that the transfer response is filtered. However, the position of the spurious response terms on the frequency axis depends on the sample spacing. If the sample spacing is large (the sample

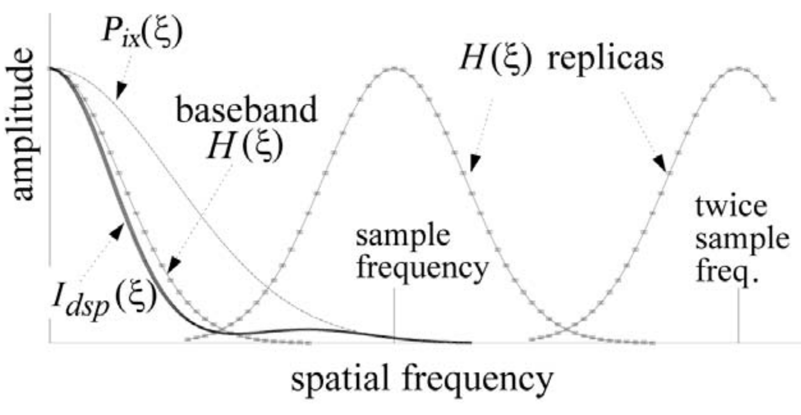


Figure 3.7 The plot shows replicated spectras of the presample MTF $H(\xi)$. Display MTF $P_{ix}(\xi)$ multiplies presample MTF $H(\xi)$ to create the transfer response $I_{dsp}(\xi)$. $P_{ix}(\xi)$ multiplies the replicas of $H(\xi)$ to create the aliasing or spurious response.

frequency is small), then the spurious response terms lie close to the baseband in the frequency domain. In that case, the spurious response is difficult to filter out and might even overlap the baseband. If the sample spacing is small (the sample frequency is high), then the spurious response terms lie far from the baseband in the frequency domain, and the spurious response is filtered out by the display and eyeball MTFs.

Figures 3.6 and 3.7 show only the amplitudes of the transfer and spurious responses; however, the sample phase in Eq. (3.7) must not be ignored. The phase relationship between the transfer response and the spurious response depends on the sample phase. This means that the displayed image changes depending on the exact relationship between the sensor and the scene. A small angular movement of the sensor changes the sample phase, which results in a different display frequency spectrum. A sampled imager is not shift invariant. As the sensor is panned across the scene, the displayed image changes depending on sample phase. The change in display frequency spectrum (the change in the intensity pattern on the display) is due to the variation in sample phase in Eq. (3.7).

Eye MTF is an important factor in limiting the visibility of both the transfer response and the spurious response. When analyzing a sensor and display system, $P_{ix}(\xi)$ should generally include the eye MTF. However, when presenting examples in a book, many factors are difficult to control. In many of the examples that follow, the display pixels are large in order to minimize the impact of the unknown viewing conditions.

3.3 Examples of Sampled Imager Response Functions

Following are some examples that illustrate the kind of information that can be gleaned from the SIR function. In these examples, the amplitudes of the transfer and spurious response terms are plotted, and the relationship between the response function plots and image characteristics is discussed.

Typically, only the positive-frequency plane is shown in plots. Since the image is real, the transfer and spurious terms must be symmetrical about the frequency origin. In some cases, however, the $n = -1$ aliasing term spreads to the positive-frequency axis and vice versa. In these cases, it is sometimes clearer to plot negative as well as positive frequencies.

In the following examples, $P_{ix}(\xi)$ is real; that is, the display pixel shape is symmetrical. A real MTF can take on negative values. Multiplying a complex function by a negative real value changes the phase of the function by 180 deg. In the following figures, plotting a portion of a transfer or spurious term as negative indicates a 180-deg phase shift.

3.3.1 Example 1: The pictures of Lena in Chapter 1

The original picture of Lena shown in Fig. 1.5(a) is blurred to produce Fig. 1.5(b). The blur is generated by convolving with both a Gaussian and a rect function. The Gaussian falls to 10% amplitude at a half width of four pixels. The rect function is eight samples wide. The image is downsampled 8:1 both horizontally and vertically. The reconstruction in Fig. 1.5(d) is done with square display pixels.

Figure 3.8 shows the transfer and spurious response terms for the picture of Lena in Fig. 1.5(d). The display MTF used to make Fig. 1.5(d) is also shown in Fig. 3.8. The large square display pixels lead to considerable out-of-band spurious response in the image. The high-frequency content permitted by these display pixels makes it difficult for the eye to integrate the underlying image.

The picture in Fig. 1.7 is constructed using the display MTF shown in Fig. 3.9. The transfer and spurious response terms associated with Fig. 1.7 are also shown in that figure. The display MTF now removes the high-frequency content. The picture in Fig. 1.7 is better than the picture in Fig. 1.5(d). Figure 1.7 does not look as good as Fig. 1.5(b) because the spurious response overlaps the transfer response, corrupting the image. Also, the display MTF adds additional blur to the image, lowering the transfer response.

The improvement between Figs. 1.5(d) and 1.7 results from removing image frequency content beyond the half-sample frequency. The "display MTF" in Fig. 3.9 actually results from a combination of digital image processing plus display pixel characteristics. This is explained in Chapter 4, which covers display interpolation.

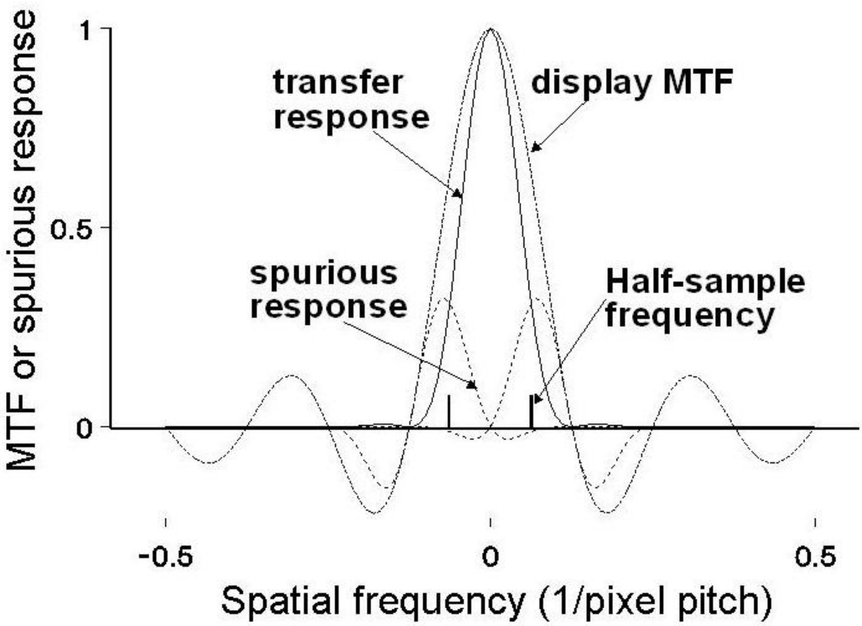


Figure 3.8 Transfer and spurious response and display MTF for the Lena picture shown in Fig. 1.5(d). The large square display pixel leads to considerable spurious response beyond the half-sample rate. The frequency content beyond half the sample rate represents the blocky distinct display pixel edges seen in Fig. 1.5(d).

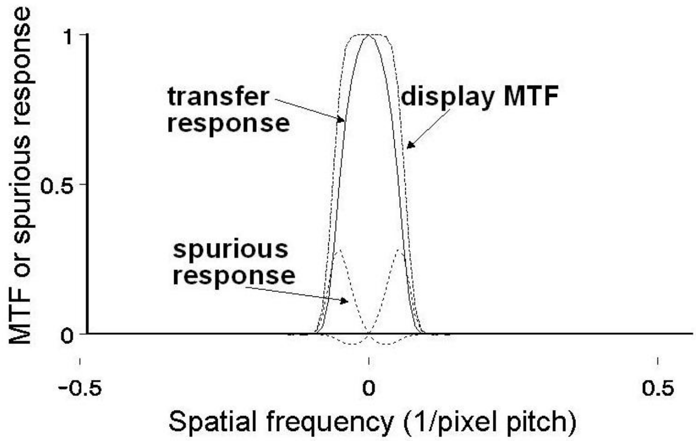


Figure 3.9 Transfer and spurious response and display MTF for the Lena picture shown in Fig. 1.7. The display MTF filters out most frequency content beyond the half-sample rate.

3.3.2 Example 2: Effect of changing sample rate

This example starts with the picture of the clock shown in Fig. 3.1. The horizontal and vertical MTFs of the charge-coupled device (CCD) camera are shown in Fig. 3.10. The sampling rate of the CCD camera is 1.67 cy/(mrad)$^{-1}$, both horizontally and vertically.

This example illustrates the impact of changing sample rate while holding other sensor and display parameters constant. In this example, the presample and postsample MTFs are held constant, and only the sample rate is varied.

Presample MTF does not change simply because fewer samples are used to reconstruct the image. Holding the presample MTF constant is easy. However, in order to hold the postsample MTF constant while varying sample spacing, the display pixel shape cannot change. Since the display pixels must be large enough to fill in the image at the lowest sample rate, the pixels must overlap at the higher sample rates. This is illustrated in Figs. 3.11(a) and (b).

In Fig. 3.11(a), each square pixel is aligned side by side. At the lowest sample rate, the pixel size (width and height) equals the pixel pitch. In Fig. 3.11(b), the large display pixels are still used to reconstruct the image; this keeps the postsample MTF constant. However, in the cases where the samples are closer together, the display pixels overlap. This overlap of adjacent pixels when the sample spacing is less than the pixel dimension is illustrated in Fig. 3.11(b). Overlap between adjacent display pixels blurs the image; the blur increases as the overlap increases.

The display pixels are large enough that they are contiguous at the lowest sample rate. At the higher sample rates, the display pixels overlap. This keeps the display pixel size and shape constant between examples that use different sample spacings. A single display pixel is shown in the upper right corner of each of Figs. 3.12, 3.15, and 3.18.

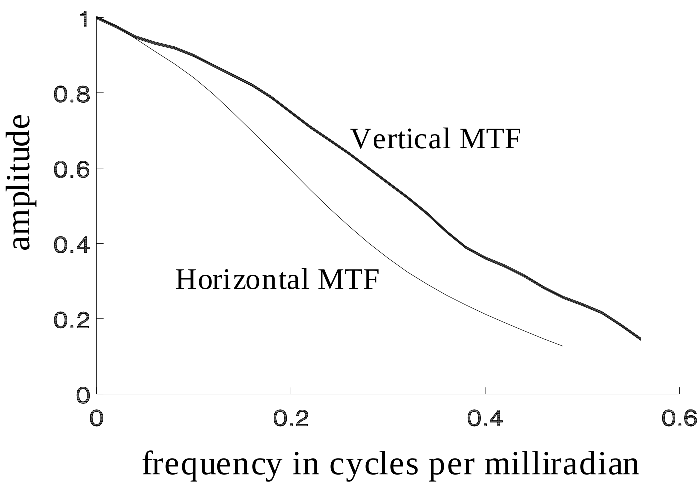


Figure 3.10 Measured horizontal and vertical MTFs for the camera used in the clock example.

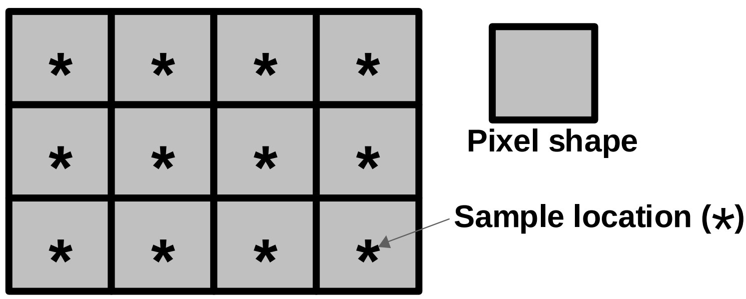


Figure 3.11(a) The clock images were reconstructed with square pixels. The pixel size is such that, at the lowest sample rate, the pixels just touch at each edge.

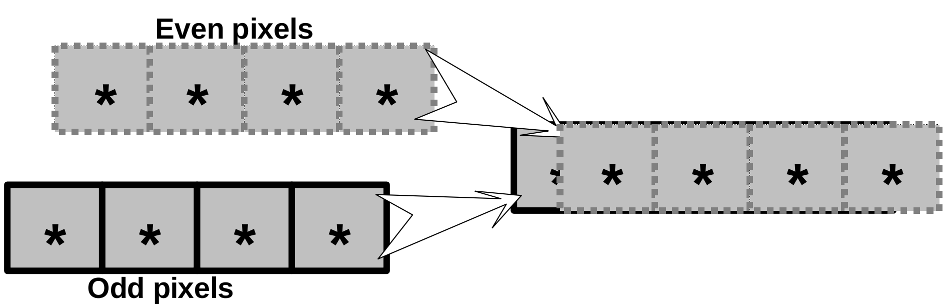


Figure 3.11(b) The higher-sample-rate images are reconstructed with display pixels that overlap. The overlap is illustrated above in one dimension, but in the examples the pixels overlap in both the horizontal and vertical directions. Each sample is reconstructed with a large square display pixel, and when the samples are close together, the large display pixels overlap, blurring the image. Keeping the pixel size constant keeps the MTF constant between the examples, even though the sample rate is varied.

The display pixel has been made large in order to minimize the impact of viewing condition. The display pixel shown at the upper right of each picture should appear large and square to the viewer. Eye MTF has not been factored into the display MTF in this example.

Figure 3.12 shows the clock image downsampled by two in each direction. That is, every other sample from the original image shown in Fig. 3.1 is discarded both horizontally and vertically. One quarter as many samples are used to generate Fig. 3.12 as are used to generate Fig. 3.1. The sample rate is now 0.83 cy/mrad in both directions. In Fig. 3.12, unlike Fig. 3.1, large square display pixels are used. The display pixels overlap because they are large compared to the sample spacing. The sample spacing is one-quarter of a display pixel dimension in each direction. The image in Fig. 3.12 is blurry.

Figure 3.13 shows the camera MTF and the replicated spectrum centered at the new sample rate of 0.83 cy/mrad. Replicated spectra at other multiples of the sample frequency are present but not shown. The display MTF is also shown in Fig. 3.13. The display pixel is essentially square with a size in object space of 4.8 mrad; therefore, the display MTF is a sinc wave that cuts off (first MTF zero) at about 0.2 cy/mrad.

The horizontal and vertical transfer and spurious responses associated with Fig. 3.12 are shown in Fig. 3.14. The picture in Fig. 3.12 is blurry because the MTF of the large display pixel creates a significant drop in the transfer response. The transfer response is dominated by the 0.2-cy/mrad cutoff of the display. However, sampling artifacts are not apparent. As seen in Fig. 3.14, spurious response is almost zero everywhere on the frequency axis. The MTF replica that causes the spurious response is centered at the sample frequency of 0.83 cy/mrad and is therefore well filtered by the MTF of the large display pixel.

Figure 3.12 Clock image sampled at 0.83 cy/mrad and displayed using large square pixels as shown in the upper right-hand corner. The large display pixels overlap, blurring the image.

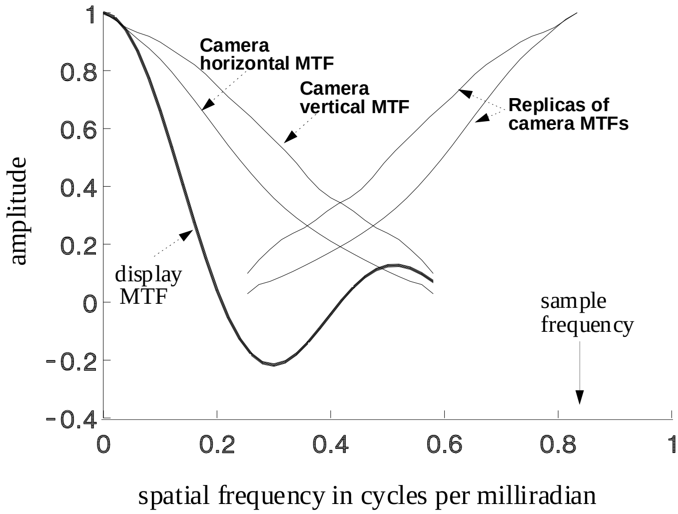


Figure 3.13 Graph showing replicas of camera MTFs located at multiples of 0.83 cy/mrad (the sample rate in Fig. 3.12). The display MTF is also shown.

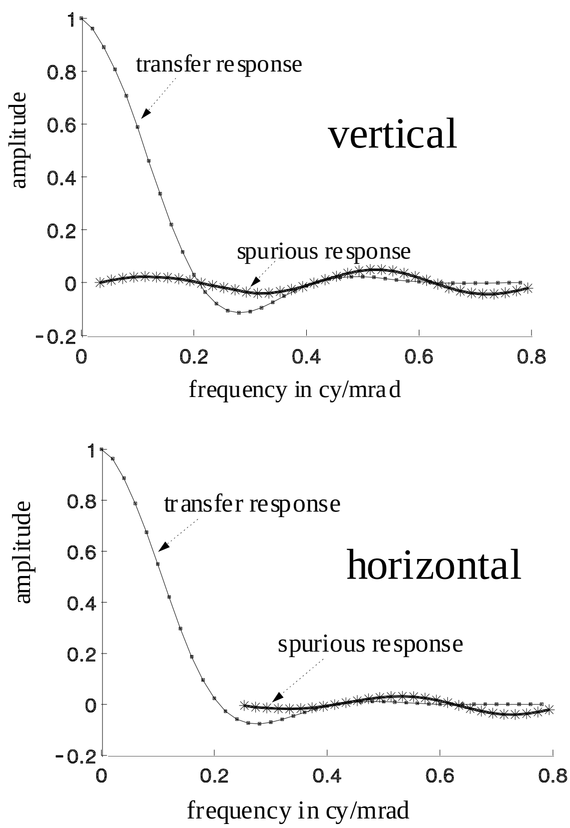


Figure 3.14 Transfer and spurious response for the camera under the sampling and display conditions associated with Fig. 3.12.

In Fig. 3.15, the clock picture in Fig. 3.1 is now downsampled by four in each direction. The sample spacing is one-half of a pixel dimension in each direction. The display pixels still overlap, but not as much as in the previous example shown in Fig. 3.12. The shape and size of the display pixel has not changed, but the sample spacing has been increased by a factor of two in each direction. The image is clearly interpretable, but sampling artifacts have appeared. Some edges are jagged, pixel intensities on the numbers are incorrect, and some of the minute markings are missing or out of place.

Figure 3.16 shows the camera horizontal MTF, the presample MTF replicas due to sampling at the new sample rate, and the display MTF. In order to reduce clutter in the figure, the vertical MTF and its replicas are not shown. The sample rate is now 0.42 cy/mrad, and replicas of the presample MTF are spaced closer together than in Fig. 3.13 because they are spaced at multiples of the smaller sample frequency.

Figure 3.17 shows the horizontal and vertical transfer and spurious responses for the sample spacing used in Fig. 3.15. The transfer response is the same as in the previous example, but the spurious response has changed; it is worse. Since spurious response is now at lower frequencies due to the lower sample rate, it is not filtered out as effectively by the display MTF.

Also, notice that the spurious response shown in Fig. 3.17 is in two parts. This is because the $P_{ix}(\xi)\ H(\omega + \nu)\ e^{-i\phi}$ term from the negative side of the frequency axis contributes. The term spreads over to the positive-frequency axis when the sample spacing gets small. As a shorthand notation in the figure, the spurious term from the MTF replica centered at the sample frequency on the positive half of the frequency axis is labeled as $e^{+i\phi}$. The spurious term from the MTF replica centered on the negative frequency axis is labeled $e^{-i\phi}$.

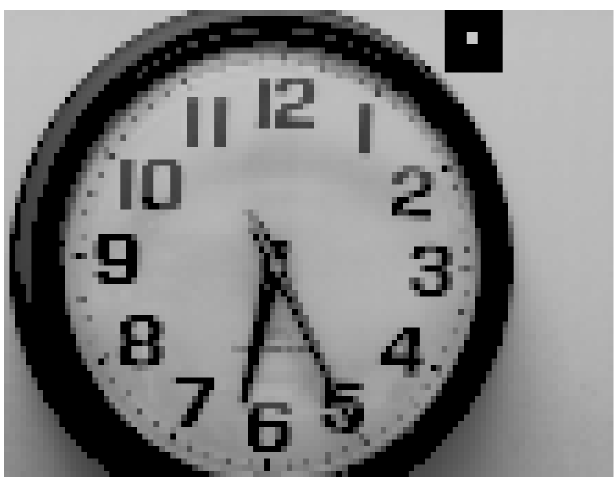


Figure 3.15 Clock image in Fig. 3.1 downsampled by four in each direction; sample frequency is now 0.42 cy/mrad. The display pixel has not changed from Fig. 3.12. Sampling artifacts have appeared. The intensity and thickness of lines varies across the picture, and some of the minute markings are missing completely.

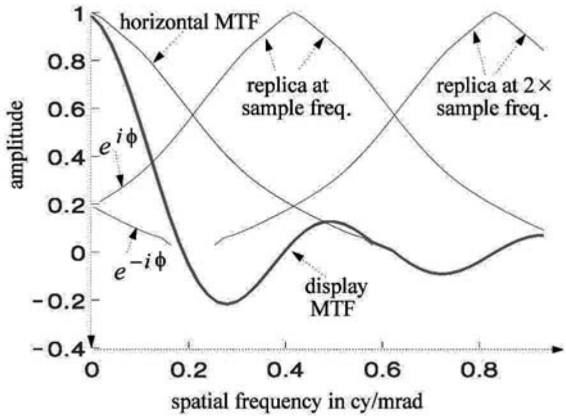


Figure 3.16 Horizontal camera MTF and replicas resulting from sampling at 0.42 cy/mrad. Since the first replica is now centered at a lower frequency, it is not filtered as effectively by the display MTF.

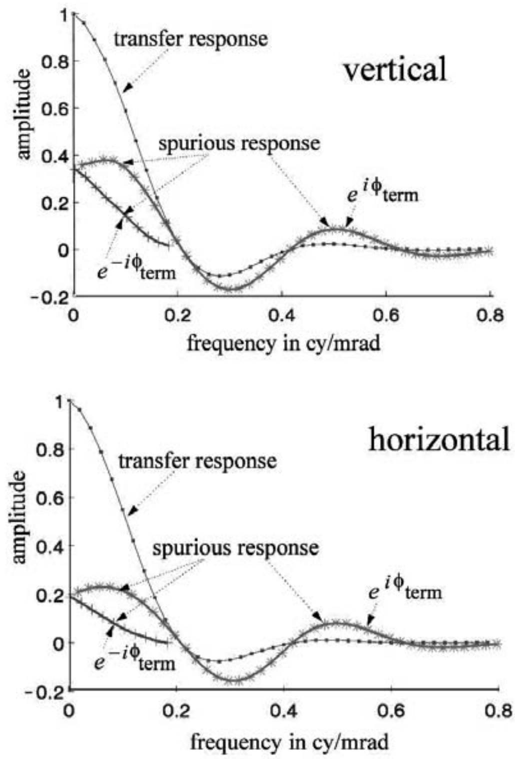


Figure 3.17 Transfer and spurious response for the camera under the sampling and display conditions associated with Fig. 3.15. Notice that the spurious response is in two parts, because the $P_{ix}(\xi)H(\xi+\nu)e^{-i\phi}$ term from the negative side of the frequency axis contributes. The baseband content is somewhat corrupted due to the overlap of spurious and transfer response. The blocky nature of the image is due to the higher-frequency spurious content.

In Fig. 3.15, the minor changes in detail (variations in intensity, line widths, etc.) are due to the overlap between the spurious response and the transfer response. (This overlap between spurious and transfer responses is shown in Fig. 3.17.) Furthermore, due to the sinc-wave nature of the display MTF, the spurious response beyond the half-sample rate is not removed. The spurious response beyond the half-sample frequency contributes the blocky look and jagged edges in the picture.

Figure 3.18 shows the clock image downsampled eight times in both directions; the sample spacing is now equal to the size of the display pixel in each dimension. Again, the transfer response has not changed; this is because the camera MTF and display pixel MTF are the same as in the two previous clock examples. Because of the lower sample rate, the blockiness and jagged edges are more apparent than in the previous example. Also, the corruption of the baseband is much worse; parts of the image are completely missing or substantially changed. These picture characteristics can be explained by looking at Figs. 3.19 and 3.20.

Figure 3.19 shows that the replicated spectra is now centered at multiples of 0.21 cy/mrad, the new sample frequency. The horizontal and vertical transfer and spurious responses are shown in Fig. 3.20. The low-frequency spurious response is about as large as the transfer response, causing significant corruption of the basic image information. Also, the high-frequency spurious response is quite large, meaning that the sharp pixel edges are quite visible.

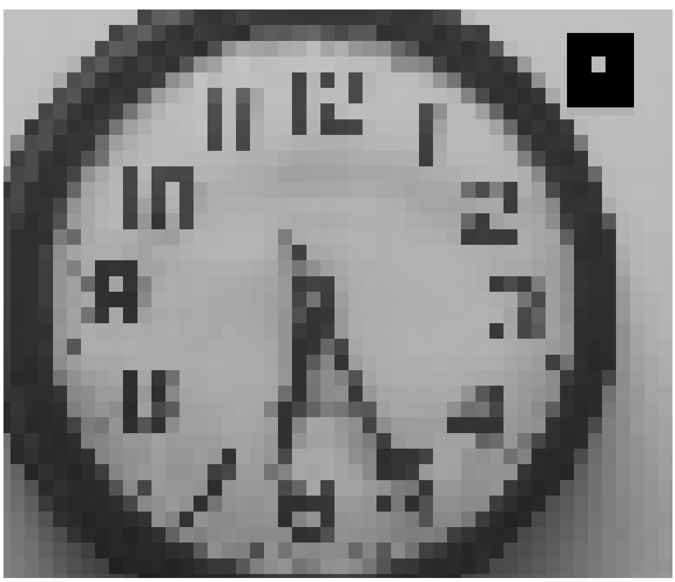


Figure 3.18 Clock image in Fig. 3.1 downsampled by eight in each direction. Sample frequency is now 0.21 cy/mrad. The sample pitch now equals the pixel size, whereas in Figs. 3.12 and 3.15, adjacent pixels overlapped. Because of the lower sample rate, parts of the image are missing or substantially changed due to aliasing. Also, the blocky nature of the display pixels is now visible.

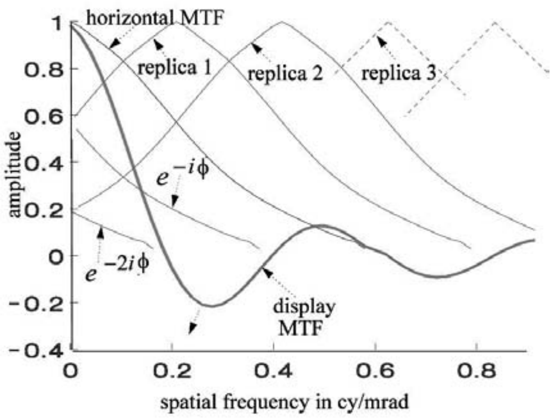


Figure 3.19 Graph showing horizontal camera MTF and replicas resulting from sampling at 0.21 cy/mrad. The display MTF cannot filter out the spurious response because the MTF replicas that lead to spurious response are now at low frequencies. The MTF replica at twice the sample rate on the negative half of the frequency axis is labeled $e^{-2i\phi}$, and the replica at the sample frequency on the negative half of the frequency axis is labeled $e^{-i\phi}$.

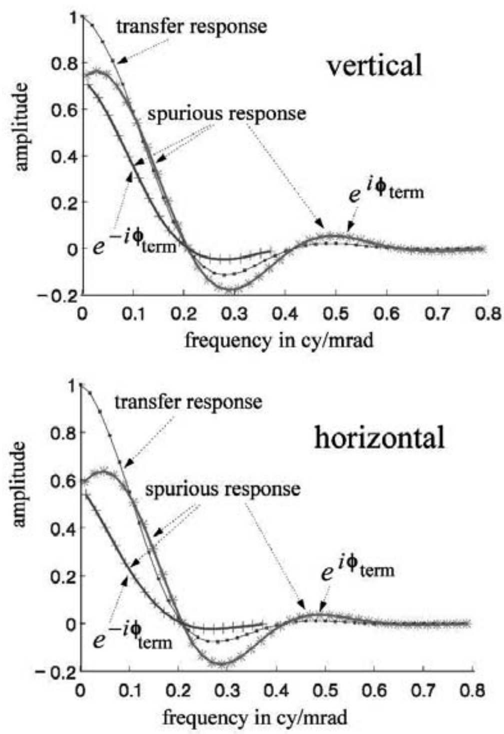


Figure 3.20 Transfer and spurious response for the camera under the sampling and display conditions associated with Fig. 3.18. Note that the spurious response seriously overlaps the transfer response and therefore corrupts the image content. The blockiness (the sharply defined pixels edges) in the image is indicated by the higher-frequency spurious content.

Comparing Figs. 3.12, 3.15, and 3.18, the images do not look more blurry as the sample rate decreases. On the contrary, although the pictures become increasingly more corrupted, the images appear sharper as sample rate decreases due to the increased high-frequency content of the spurious response.

3.3.3 Example 3: Midwave thermal imager

This example illustrates the application of the SIR function to a midwave thermal imager. The presample blur includes optical diffraction and detector shape. The imager spectral band is 3 to 5 μm, and the diffraction wavelength is 4 μm. The diameter of the objective lens is 2.2 cm. Focal length is 5 cm. The detector size is 31 μm square, and detector pitch is 38 μm. The detector angular subtense is 0.62 mrad. The sample rate is 1.32 samples per mrad. With a 256 × 256 detector staring FPA, the field of view of the system is 11 × 11 deg.

The imager presample MTF is shown in Fig. 3.21(a). The three curves represent MTF due to diffraction, MTF due to integrating light over the active detector area, and the product. The product is the imager presample MTF. Only the horizontal spectrum is shown since the vertical spectrum is identical. Notice that the diffraction cutoff is well beyond the half-sample frequency.

Figure 3.21(b) shows the replicas created by sampling for $n = \pm 1, \pm 2, \pm 3, \pm 4$. The imager is poorly sampled, and all eight replicas contribute to visible aliasing. Figure 3.21(c) shows the display MTF. The display pixel shape is the same as was used for the detector.

The spurious response is shown in Fig. 3.21(e). Spurious response is found by multiplying the replicas in 3.21(b) by display MTF. Transfer response is shown in Fig. 3.21(d). The imager transfer response is found by multiplying the replica at $n = 0$ by display MTF.

Figure 3.21(f) shows the root of the sum of squares (RSS) of all of the spurious response terms at each spatial frequency. The terms overlap in frequency, and some metric of total aliasing is desired. If the image samples have random amplitudes and phases, then the RSS is a good approximation for the spurious amplitude at each frequency.

The diffraction cutoff is well beyond the half-sample frequency, and a large amount of in-band aliasing results. Because of the display MTF, the thermal image also has a large amount of signal at frequencies beyond the half-sample frequency. The in-band aliasing cannot be filtered out. However, the out-of-band aliasing and resulting artifacts can be removed.

The high-frequency aliasing is filtered out using interpolation and a higher-resolution display. Figure 3.22 shows the same signals as Fig. 3.21 except that display interpolation is used. Figure 3.22(c) shows both the new display MTF and the MTF from Fig. 3.21(c) for comparison. The spurious response in Fig. 3.22(e) is now limited to the half-sample frequency.

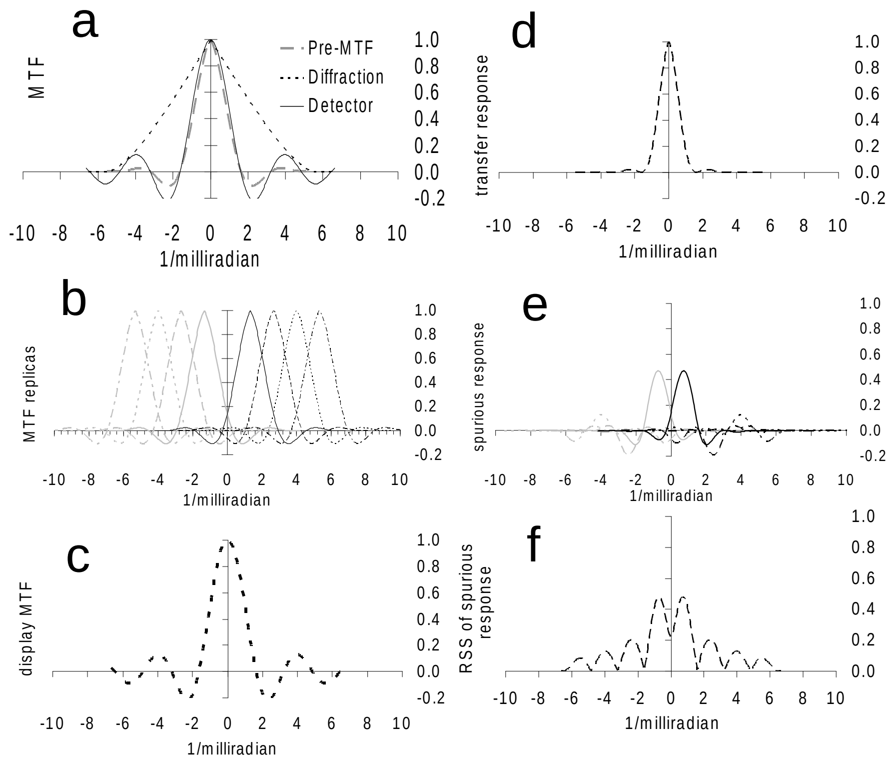


Figure 3.21 (a) Presample MTF, (b) MTF replicas, (c) display MTF, (d) transfer response, and (e) spurious response for the example thermal imager. The RSS of spurious terms at each spatial frequency is shown in (f). These signals are associated with (b) and (e) in Fig. 3.23.

Figure 3.23 shows a simulation of a target at 200 m in (a) through (c). The target is at 400 m in (d) through (f). To make the pictures easier to see and compare, the 400-m images are enlarged to the same size as the 200-m images. In (a) and (d), the pristine image is blurred but not downsampled. So (a) and (d) show the effect of blur only. In (b) and (e), the image is blurred, downsampled, and reconstructed using the display MTF shown in Fig. 3.21(c). In (c) and (f), the image is blurred, downsampled, and reconstructed using the display MTF plus interpolation. The interpolated MTF is shown in Fig. 3.22(c).

In Fig. 3.23(b) and (e), note the degradation caused by out-of-band aliasing. The distinct pixels capture the eye and hide the underlying image. The images in (c) and (f) have the same number of samples but appear much cleaner. Note also, however, that the in-band aliasing degrades the images. This is seen by comparing (c) to (a) and (f) to (d). Figures 3.23(c) and (f) are not as good as (a) and (d), respectively. In (c), compare the road wheels, the back edge of the turret, and the details at the top rear of the vehicle to the same details in (a). In-band aliasing is corrupting image (c). Note also that sampling has a greater effect at the longer range. This is because the imager samples in angle space. The number of samples on target decreases with range.

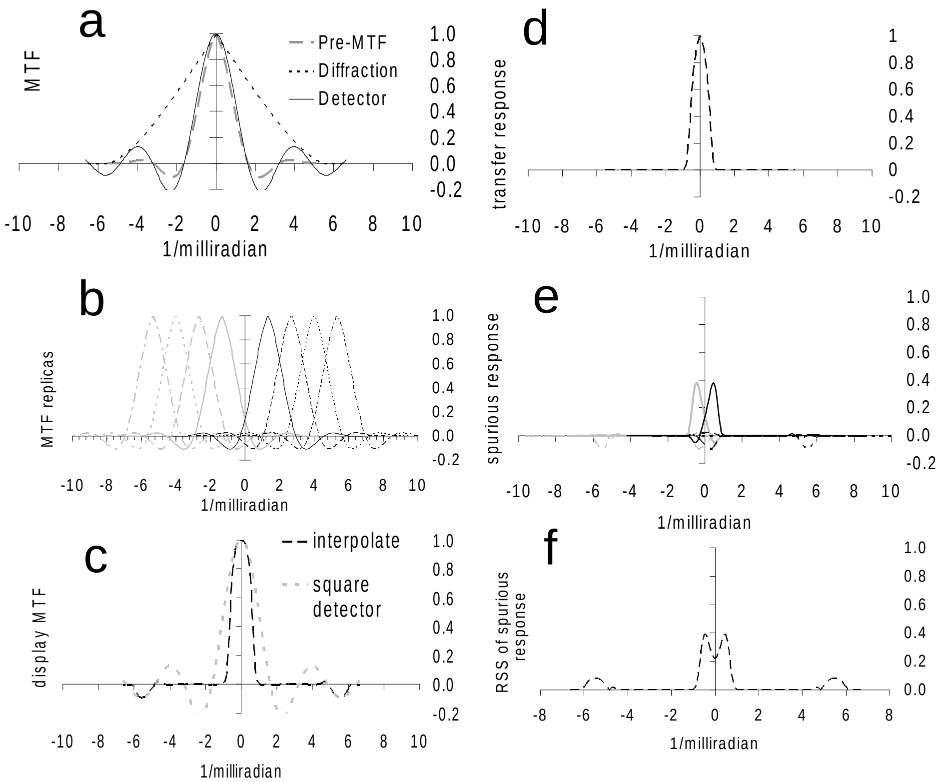


Figure 3.22 This figure shows the same signals as Fig. 3.21, except that display interpolation is used. The new display MTF is shown in (c). The reduced spurious response is shown in (e) and (f). These signals are associated with (c) and (f) in Fig. 3.23.

The SIR functions and RSS plots quantify the amount and type of aliasing present in the displayed image. The spurious signal beyond the half-sample frequency in Figs. 3.21(e) and (f) is clearly visible in Figs. 3.23(b) and (e). The SIR functions also show that changing the display does not affect in-band aliasing.

3.3.4 Example 4: Two-dimensional SIR example

In this example, presample MTF is again based on diffraction and detector blur. In this case, however, the sample rate is sufficient to avoid any in-band aliasing. This means that detector spacing is small compared to the diffraction blur size. A well-sampled imager requires either a large *f/#* or small detectors.

Diffraction wavelength is 4 μm. Aperture size is 20 cm. The diffraction cutoff frequency is 50 mrad^{-1}. Detector pitch is therefore 0.01 mrad to provide a 100-mrad^{-1} sample rate. Taking 15 μm as a reasonable detector size in the midwave infrared, the focal length is 1.5 m. This means that the optical *f/#* is

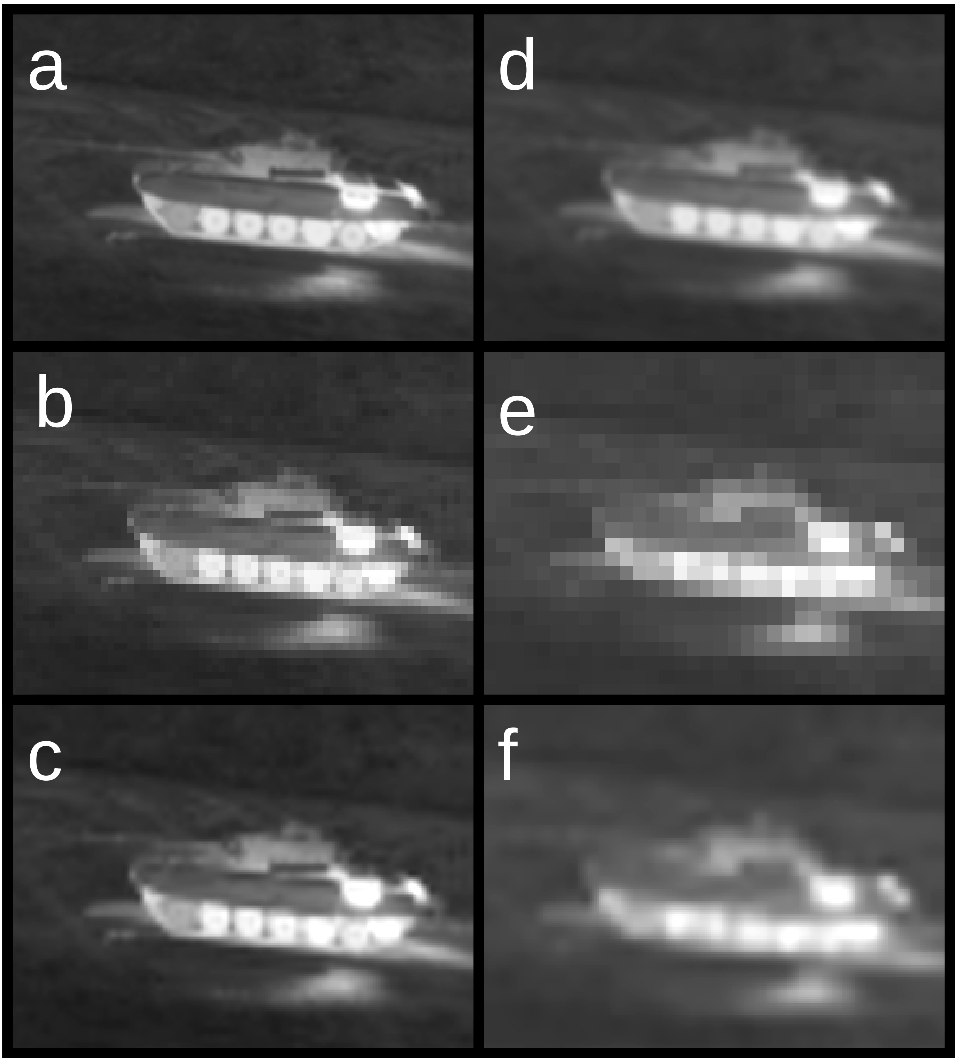


Figure 3.23 Simulation of a target at 200 m (left column) and at 400 m (right column). The pictures in (a) and (d) are blurred but not downsampled. The right-hand pictures are enlarged to make comparisons easier. In (b) and (e), the spurious response beyond the half-sample frequency makes the target hard to see. The sharp pixel edges capture the eye and hide the underlying image. In (c) and (f), the images are improved by removing aliasing beyond the half-sample frequency. Note that in-band aliasing degrades the images more at longer range.

7.5. With a 640 × 480 FPA, the ultranarrow field of view is 0.28 × 0.37 deg. In this example, a flat panel display with square pixels is used. The display MTF is the same as the detector MTF.

Figure 3.24(a) shows the presample MTF. Sample replicas are shown in (b), and display MTF is shown in (c). Figure 3.24(d) shows the combination of transfer and spurious response. All signal below 50 mrad^{-1} is transfer response, and all signal at higher frequencies is spurious.

Figure 3.25 simulates a target at 15 km. Figure 3.25(a) shows the target blurred but not sampled. Figure 3.25(b) shows the target blurred, sampled, and then displayed with the display MTF represented in Fig. 3.24(c). Note the spurious signal beyond the half-sample frequency in Fig. 3.24(d). This results in the visible pixel edges in Fig. 3.25(b).

In Fig. 3.25(c), the picture is generated using display interpolation to remove the high-frequency aliasing beyond the half-sample frequency. The same number of samples is used to generate Figs. 3.25(b) and (c). Figure 3.25(c) is almost identical to Fig. 3.25(a). The blur is very slightly worse than in Fig. 3.25(a) because of the interpolation. However, no sampling artifacts are visible.

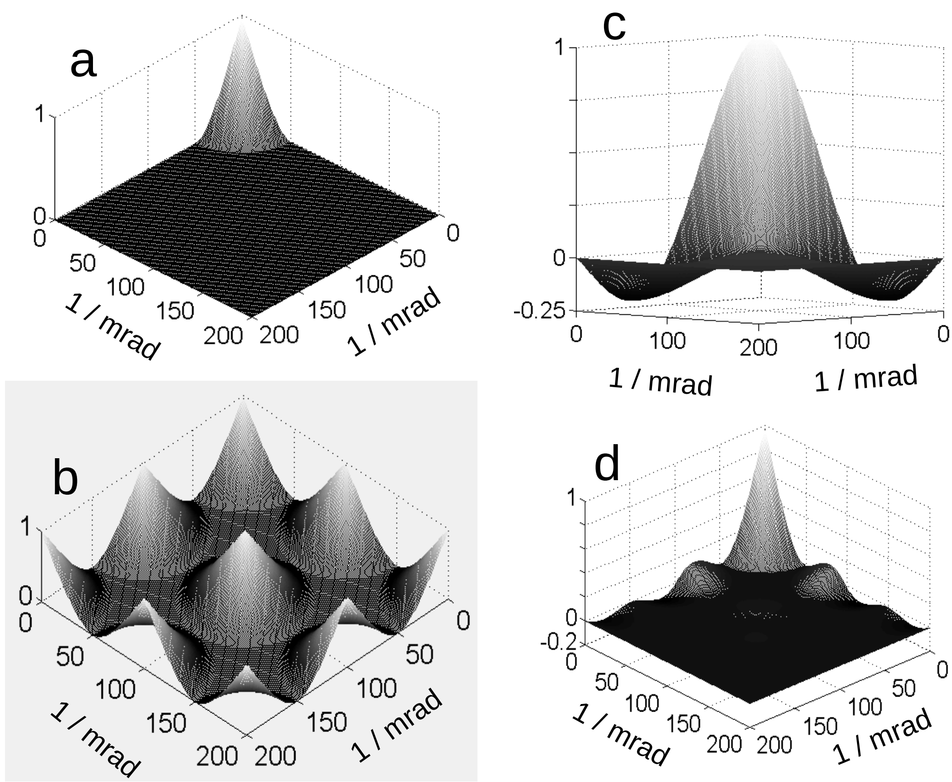


Figure 3.24 (a) Presample MTF. (b) Sample replicas. Note that replicas exist off the horizontal and vertical axes. Whether these are visible depends on the display MTF. In this case, display MTF in (c) multiplies the replicas in (b) to generate the SIR in (d). The main spurious response beyond the half-sample frequency is on the horizontal and vertical axes. This is because the detector edges are aligned to the coordinate axes. A small amount of aliasing is present at (100,100) because a small "bump" in display MTF exists near that frequency.

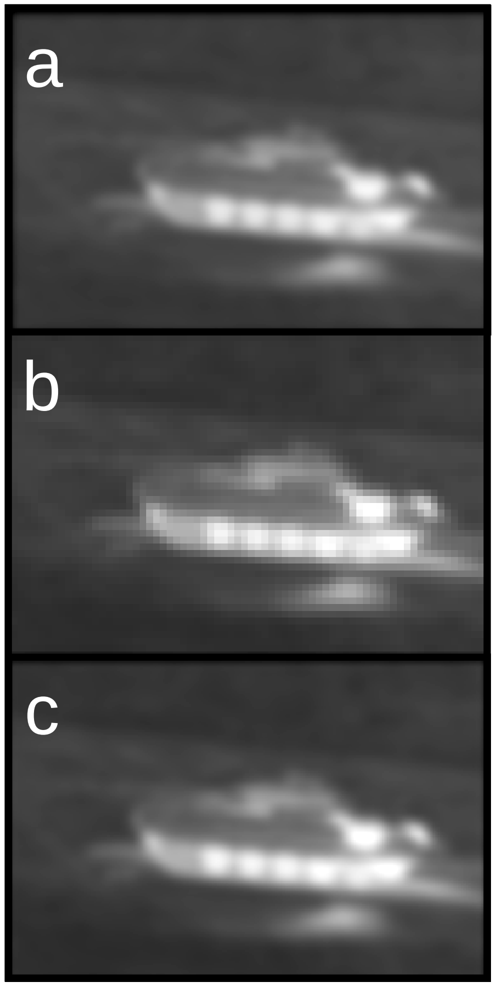


Figure 3.25 Target viewed with ultranarrow field-of-view thermal imager. Target is at 15 km. (a) The target is blurred but not downsampled. (b) The target is blurred, downsampled, and reconstructed using square pixels. (c) The image is reconstructed using display interpolation. Comparing (c) to (a) shows the effect of in-band aliasing; none is present. Comparing (b) to (c) shows the effect of out-of-band aliasing. The square display pixels in (b) cause aliasing beyond the half-sample frequency. This creates distracting artifacts in the viewed image.

Note that a presample MTF replica exists at (100,100), but that particular aliasing is mostly suppressed by the display MTF. The display pixels are oriented with edges along the coordinate axes. Figure 2.8 illustrates what happens to MTF if a square shape is rotated 45 deg. If display pixels are oriented diagonally, the aliasing at (100,100) is visible. Figure 3.26 shows the SIR when the display pixels are diagonal. The visible pixel edges are now diagonal also.

The one-dimensional treatment is accurate as long as detectors and display pixels are oriented along the horizontal and vertical axes. Fortunately, diagonal orientation of display pixels is unusual.

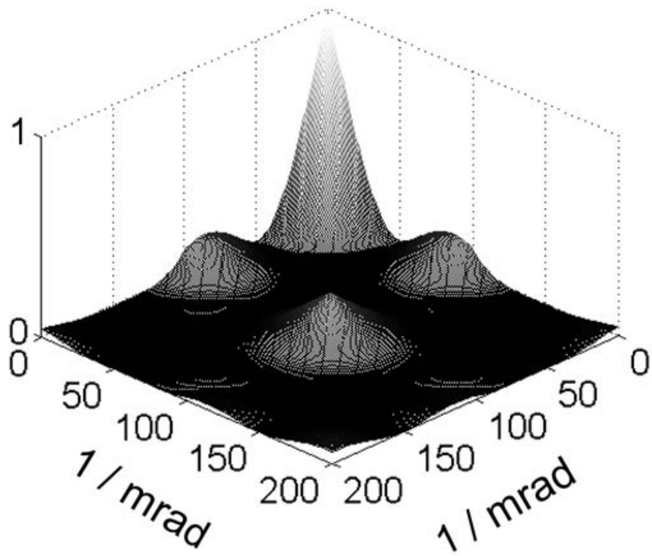


Figure 3.26 SIR when display pixels are oriented diagonally. Aliasing off the horizontal and vertical axes is now much more visible.

Bibliography

Vollmerhausen, R., "Display of Sampled Imagery," in *Electro-Optical Imaging: System Performance and Modeling*, L. Biberman, Ed., SPIE Press, Bellingham, WA (2000).

Vollmerhausen, R., "Impact of display modulation transfer function on the quality of sampled imagery," *Proc. SPIE* **2743**, 12–22 (1996).

Wittenstein, W., J. C. Fontanella, A. R. Newberry, and J. Baars, "The definition and the OTF and the measurement of aliasing for sampled imaging systems," *Optica Acta* **29**(1), 41–50 (1982).

Wozencraft, J. M. and I. M. Jacobs, *Principles of Communication Engineering*, Wiley, New York (1965).

Chapter 4
Sampled Imager Optimization

In theory, a sampled imager is optimized by removing all aliasing. One applies presample and postsample filters at half the sample frequency; one avoids in-band spurious response by removing any frequency content beyond the half-sample frequency before sampling; and one avoids out-of-band spurious response by removing any frequency content beyond the half-sample frequency on the display.

It is possible to create a postsample filter that is close to ideal. How close to ideal and how to do it are the main topics of this chapter. If the correct interpolation function is used, a display with multiple pixels for each sensor sample provides a good postfilter.

Unfortunately, adding preblur beyond that provided by the optics and detector is as likely to hurt performance as to help. There are multiple reasons for this. One is that an "ideal" prefilter in the same sense as interpolation is not possible. The postfilter is applied by digital manipulation of the samples. Once the samples are taken, it is too late to apply a prefilter.

Adding preblur by using slower optics, defocusing, adding fiber optic plates, or other optical schemes degrades in-band MTF. Removing frequency content beyond the half-sample frequency is good. Blurring the image to do it is probably counterproductive.

A third problem with adding preblur is that in-band aliasing might not be all that important, anyway. Blur is undesirable. Noise is undesirable. Aliasing might be undesirable, depending on the target, the task, the imager, and the range.

Chapters 6, 7, and 8 discuss performance evaluation. The trade-off between blur and aliasing can be made once a scenario is selected and target contrast, size, and range are known. There are certainly times when degrading optical blur improves range performance. Unfortunately, that kind of trade-off is scenario dependent.

Chapter 5 discusses interlace and dither. Both remove in-band aliasing by providing better sampling of the scene. In the remainder of this chapter, the benefit of display interpolation is discussed.

If the display provides one pixel per sensor sample, then the postfilter characteristics, $P_{ix}(\xi)$ in Eq. (3.6), are set by the design of the display. The display pixel shape and spacing, in combination with the eye MTF, determine the relative amount of transfer and spurious response in the final image. The pixel shape and spacing might be such as to provide a pleasing image, but generally the result is suboptimal.

It is possible to remove essentially all of the out-of-band spurious response with little degradation to the transfer response of the sensor. Image interpolation can remove much of the “bad” without affecting the “good.”

Interpolation requires multiple display pixels for each sensor sample. Interpolation connotes having a display with multiple pixels for each detector in the FPA. Electronic zoom (ezoom) connotes using multiple display pixels for each detector sample but losing field of view. Ezoom uses multiple display pixels for each sensor sample, so some of the sensor samples cannot be displayed.

Figure 4.1(b) illustrates interpolation. The display in Fig. 4.1(b) has twice as many pixels horizontally and vertically as has the display in 4.1(a). Although there are twice as many display pixels, each pixel is half the size. Figure 4.1(c) illustrates ezoom. With ezoom, the number of display pixels does not increase. Instead, multiples of the original display pixels are used for each sensor sample. This means that some of the original image is not displayed. It also means that the portion of the field of view that is displayed is larger.

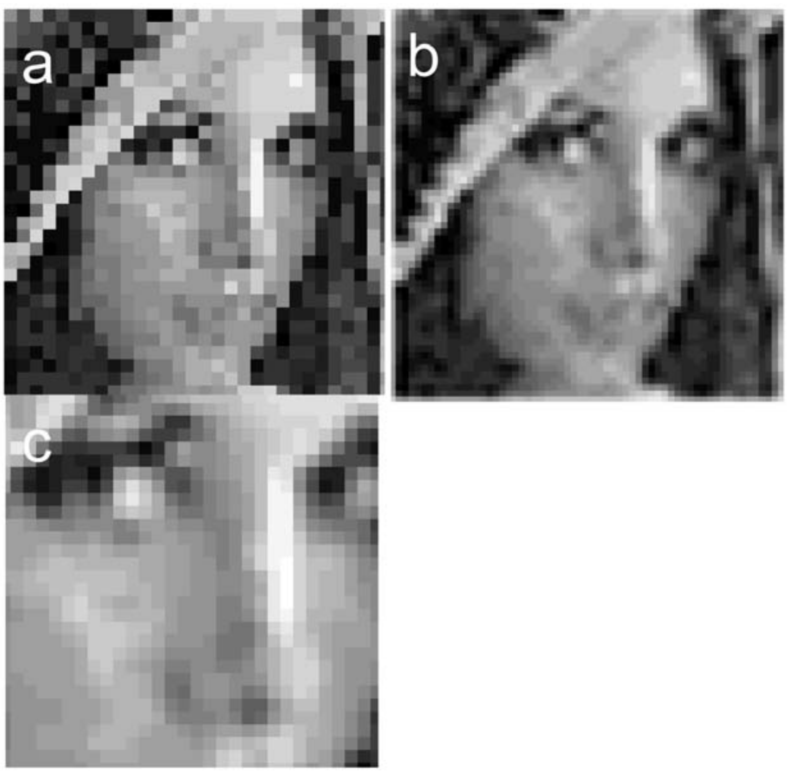


Figure 4.1 The picture in (b) uses interpolation. The picture in (c) uses ezoom. The pixels in (b) are half the size of the pixels in (a). Twice as many pixels horizontally and vertically are required to show the whole image. In (c), the pixels are the same size as in (a). Only half of the image can be displayed, and that half is twice as large.

Interpolation requires image processing, and the display must have more than one pixel per sensor sample. There is a system cost for interpolation. Quite often, however, interpolation provides the least expensive way to obtain the needed performance. The alternative might be increasing detector count, which requires both a bigger FPA and more display pixels.

Interpolation enhances the fidelity of the displayed image. From a spatial-domain viewpoint, the reconstructed image should match the presampled image between samples. Multiple smaller display pixels are used to generate the correct intensity pattern. From a frequency-domain viewpoint, the interpolation filters out aliasing above the half-sample frequency while passing baseband frequencies.

4.1 Interpolation Implementation

Interpolation functions that have a width of only a few samples and have the desired frequency characteristics can be generated by applying a Gaussian window to the sinc function. Some interpolation functions with good filtering characteristics are shown in Figs. 4.2(a) through (c). Their frequency transforms are shown in Fig. 4.3.

Figure 4.4 illustrates interpolation implementation. The interpolation function is centered at a point for which the interpolation value is desired. Since the interpolation function is six samples wide, the three samples to each side of the interpolated point will contribute to the interpolated value. The coefficients that multiply each sample value are found by centering the interpolation function on the interpolated point and finding the amplitude of the function at the sample locations. For example, the interpolated value associated with the location halfway between samples 0 and 1 are as follows:

$$\begin{aligned}\text{Value at } 0.5 = {} & 0.558\,(\text{value of sample } 0 + \text{value of sample } 1)\\ & - 0.088\,(\text{value of sample } {-1} + \text{value of sample } 2)\\ & + 0.011(\text{value of sample } {-2} + \text{value of sample } 3).\end{aligned}$$

The interpolation points for the entire image are found by convolving a discrete kernel over the image data. Some convolution coefficients for finding the midpoints between the image samples are shown in Table 4.1. These coefficients are based on the interpolation functions shown in Figs. 4.2(a) through (c). Convolution kernels using these coefficients generate new data points and are used to double the size of an image.

As an example, the six-sample-wide interpolation will be used to double the size of an image [*Img*].

$$[Img] = \begin{bmatrix} Img11 & Img12 & Img13 & \ldots & \ldots & Img1N \\ Img21 & Img22 & Img23 & \ldots & \ldots & Img2N \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ ImgM1 & ImgM2 & ImgM3 & \ldots & \ldots & ImgMN \end{bmatrix}$$

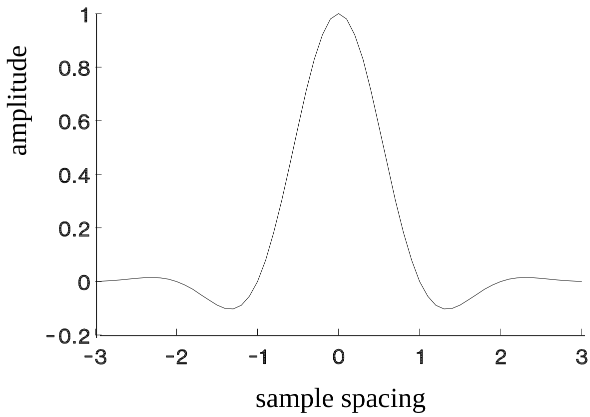


Figure 4.2(a) Six-sample-wide reconstruction function used to interpolate sampled data.

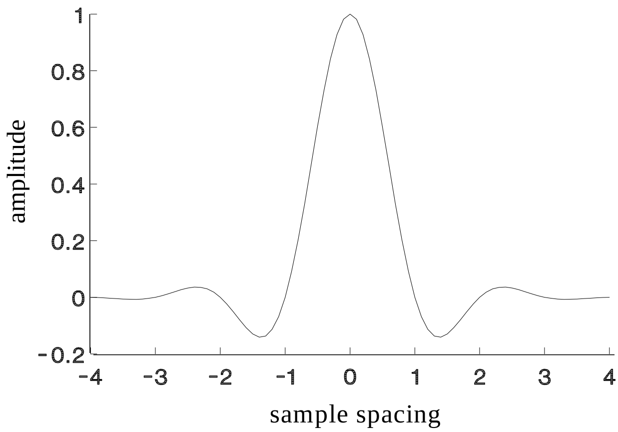


Figure 4.2(b) Eight-sample-wide reconstruction function used to interpolate sampled data.

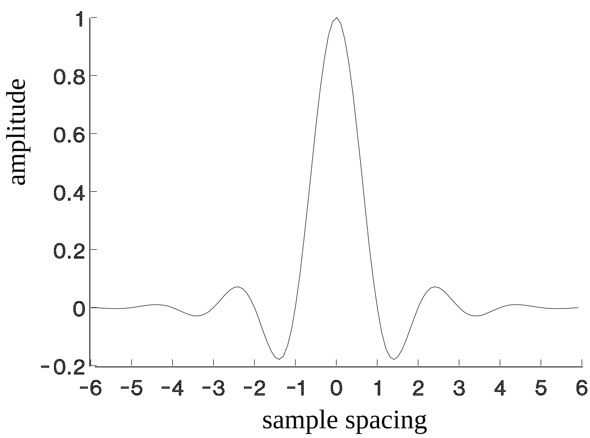


Figure 4.2(c) Twelve-sample-wide reconstruction function used to interpolate sampled data.

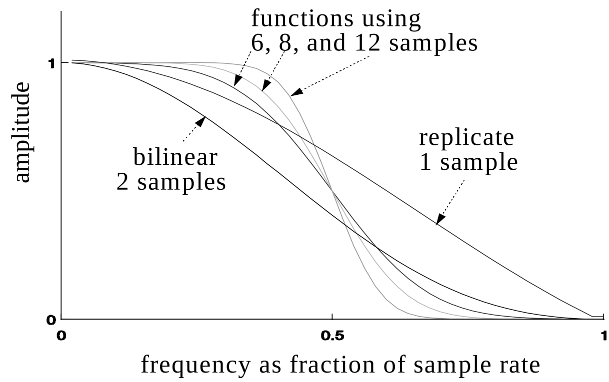


Figure 4.3 Filter characteristics for the pixel replication, linear interpolation, and reconstruction functions shown in Figs. 4.2(a) through (c).

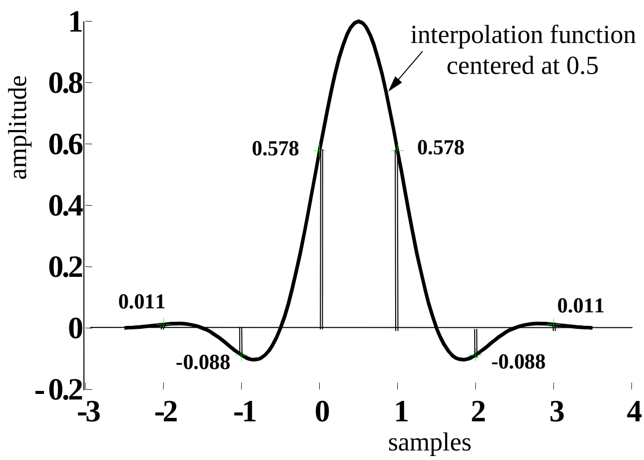


Figure 4.4 Coefficients used to interpolate the midpoints between samples. The interpolation function is centered at the point for which an interpolation value is needed. The amplitude of the interpolation function at the adjacent sample locations gives the coefficients to use for interpolation. For example, to obtain the interpolated value for the 0.5 location, multiply 0.578 by the sum of the values for sample 0 and sample 1, subtract 0.088 multiplied by the values of samples –1 and 2, then add the product of 0.011 by the values of samples –2 and 3.

Table 4.1 Kernel coefficients applied to data on both sides of the interpolation point.

Type/size	Kernel coefficients					
linear	0.5					
6 samples	0.58	–.089	0.011			
8 samples	0.604	–0.13	0.032	–0.006		
12 samples	0.6213	–0.1704	0.0693	–0.0276	0.01	–0.003

As an example, the six-sample-wide interpolation will be used to double the size of an image [*Img*].

$$[Img] = \begin{bmatrix} Img11 & Img12 & Img13 & \ldots & \ldots & Img1N \\ Img21 & Img22 & Img23 & \ldots & \ldots & Img2N \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ ImgM1 & ImgM2 & ImgM3 & \ldots & \ldots & ImgMN \end{bmatrix}$$

An array [*Img′*] is generated with twice the image elements both horizontally and vertically (the final image contains four times as many display elements).

$$[Img'] = \begin{bmatrix} Img11 & 0.0 & Img12 & 0.0 & Img13 & 0.0 & \ldots & \ldots & 0.0 & Img1N \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & \ldots & \ldots & 0.0 & 0.0 \\ Img21 & 0.0 & Img22 & 0.0 & Img23 & 0.0 & \ldots & \ldots & 0.0 & Img2N \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & \ldots & \ldots & 0.0 & 0.0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ ImgM1 & 0.0 & ImgM2 & 0.0 & ImgM3 & 0.0 & \ldots & \ldots & 0.0 & ImgMN \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & \ldots & \ldots & 0.0 & 0.0 \end{bmatrix}$$

For the six-sample-wide interpolation, the horizontal interpolation kernel is [*KH*] and the vertical kernel is [*KV*].

$$[KH] = [0.011 \;\; 0.0 \;\; -0.089 \;\; 0.0 \;\; 0.58 \;\; 1.0 \;\; 0.58 \;\; 0.0 \;\; -0.089 \;\; 0.0 \;\; 0.011\,]$$
$$[KV] = [0.011 \;\; 0.0 \;\; -0.089 \;\; 0.0 \;\; 0.58 \;\; 1.0 \;\; 0.58 \;\; 0.0 \;\; -0.089 \;\; 0.0 \;\; 0.011\,]^T,$$

where T indicates the transpose of the array. The horizontal convolution is performed first in order to generate complete horizontal rows. After performing the horizontal convolution with [*KH*], the vertical convolution with [*KV*] is performed. After the vertical convolution is performed, the matrix *Interp* contains the final double-sized image

$$[Interp] = \{[Img'] * [KH\,]\} * [KV],$$

where * means convolution.

The interpolation process can be thought of as performing a near-optimum reconstruction and then resampling at a higher frequency so that more display pixels are used in the final image. Interpolation provides a way of estimating the shape of the presampled image between sensor samples. If no interpolation is used, the shape of the image between samples depends on the inherent shape of the display pixels. When interpolation is used, the inherent display pixel shape has less effect on the final image.

The response function for a sampled imager is now extended to include the interpolation function MTF. Based on Eq. (3.7),

$$\begin{aligned}&\text{Transfer response} = Z(\xi)P_{ix}(\xi)H(\xi)\,.\\&\text{Spurious response} = Z(\xi)P_{ix}(\xi)H(\xi-\nu)e^{-j\phi} + Z(\xi)P_{ix}(\xi)H(\xi+\nu)e^{j\phi}.\end{aligned} \tag{4.1}$$

$Z(\xi)$ includes the effect of resampling. For example, the MTF for linear interpolation is $\mathrm{sinc}^2(\pi\xi X/2)$, where ξ is frequency in cycles per milliradian and X is the sample spacing in milliradians. That MTF is associated with convolving a triangle over the sample points and is appropriate if many interpolated points are used between each of the original sample points. However, if a single new midpoint value is interpolated between each original sample, then the effect of resampling on $Z(\xi)$ must be considered.

The MTF for discrete linear interpolation can be found in the following manner. As discussed at the end of Chapter 1, the spatial distribution corresponding to the Fourier transform of the samples is a set of Dirac delta functions with areas equal to the sample values. Linear interpolation is represented by a set of three delta functions that are convolved over the sample points to yield the interpolated image. One delta function is unity magnitude, the other two are one-half value, offset to either side of the unity delta function by half a sample width.

$$\begin{aligned}&\text{New samples} =\\&\left[\sum_n f(x)\delta(x-nX)\right] * \left[\delta(x)+0.5\delta(x-X/2)+0.5\delta(x+X/2)\right]\end{aligned} \tag{4.2}$$

where $*$ means convolution.

A convolution in the spatial domain corresponds to a multiplication in the frequency domain. The Fourier transform associated with linear interpolation is the transform of the three delta functions. The Fourier transform of the three delta functions is multiplied by one half to normalize response to unity at zero frequency.

$$\begin{aligned}Z(\omega)_{linear} &= 0.5\,[1+0.5e^{-i\omega X/2} + 0.5e^{\,i\omega X/2}\,]\\&= 0.5+0.5\cos(\omega X/2).\end{aligned} \tag{4.3}$$

The MTF for the other interpolation functions are found in a similar fashion.

$$\begin{aligned}Z(\omega)_{\,replicate} &= \cos(\omega X/4)\\Z(\omega)_{\,k-samp} &= 0.5 + \sum_{j=1}^{k/2} a(j)\cos[(2j-1)\omega X/2],\end{aligned} \tag{4.4}$$

where $a(j)$ are the coefficients shown in Table 4.1.

Figures 4.5(a) and (b) show the filtering characteristics of the discrete interpolation functions. The filtering functions in Fig. 4.5(a) are cyclic (periodic) with a period that is double the original sample frequency. The replication MTF is periodic with a period four times the sample frequency.

A good interpolation function removes the first replicated sideband in the original data, while not significantly degrading the transfer response of the system. Replication is not effective at suppressing out-of-band spurious response and also has a degrading effect on the transfer response. Linear interpolation is much better than replication in terms of suppressing out-of-band spurious response. Linear interpolation does, however, degrade the transfer response. The 6-, 8-, or 12-sample-wide interpolation functions are extremely effective at rejecting out-of-band spurious response while minimally degrading transfer response.

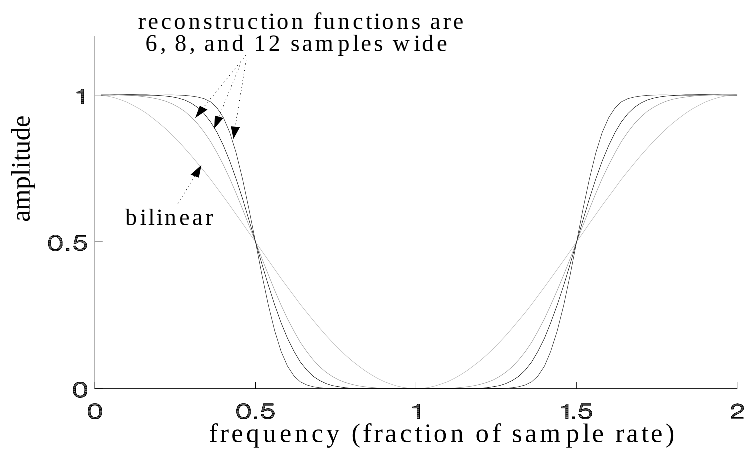


Figure 4.5(a) MTF of discrete interpolation functions. A good interpolation function removes the spurious response centered at the sampling frequency with minimum effect on the transfer response.

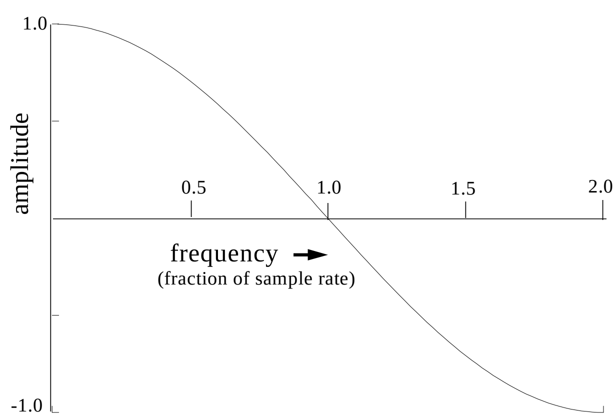


Figure 4.5(b) MTF associated with pixel replication. This filter function does not provide good rejection of the spurious response term that is centered at the sample frequency.

The picture shown in Fig. 4.6 is a CCD camera image of a pen-and-ink drawing of a helicopter. Figures 4.7, 4.8, and 4.9 show the helicopter enlarged eight times by using replication, bilinear interpolation, and the six-sample interpolation function, respectively.

The spurious response in the replicated image masks the underlying picture. The blocky images make it difficult for the visual system to integrate the picture information. Bilinear interpolation provides a much better image than does pixel replication. When compared to the images created with bilinear interpolation, the images created with the broader six-sample interpolation function are sharper and have fewer spurious artifacts.

Figure 4.6 CCD camera image of a pen-and-ink drawing of a helicopter.

Figure 4.7 Helicopter drawing magnified eight times by pixel replication.

Figure 4.8 Helicopter drawing magnified eight times by bilinear interpolation. Spurious artifacts are considerably reduced when compared with the pixel replication shown in Fig. 4.7.

Figure 4.9 Helicopter drawing magnified eight times using the six-sample interpolation function shown in Fig. 4.4. Picture is sharper and has fewer spurious artifacts than pictures generated with pixel replication or bilinear interpolation.

Sampled imagery is optimized by a postfilter that passes spatial frequencies beyond the half-sample frequency and rejects higher frequencies. Although the combination of display hardware and eyeball might provide a decent postfilter, interpolation and ezoom provide a way to control postfilter characteristics and optimize the resulting image.

Replication is easiest to implement in terms of the required electronic processing. However, the resulting display characteristics are very poor. Replication does not filter out frequencies beyond the half-sample rate. Bilinear interpolation provides a better result than replication. However, neither the transfer nor the rejection characteristics of bilinear interpolation are optimal. Interpolations using at least six coefficients provide near-ideal bandpass characteristics while rejecting aliasing beyond the half-sample frequency.

Bibliography

Arguello, R., H. Kessler, and H. Sellner, "Effect of sampling, optical transfer function shape, and anisotropy on subjective image quality," *Proc. SPIE* **310** 24–33 (1981).

Barbe, D. and S. Campana, "Imaging Arrays Using the Charge-Coupled Concept," in *Advances in Image Pickup and Display*, B. Kazan, Ed., Vol. 3, 171–296 (1977).

D'Agostino, J., M. Friedman, R. LaFollette, and M. Crenshaw, "An experimental study of the effects of sampling on FLIR performance," *Proc. IRIS Specialty Group on Passive Sensors* **291**, Infrared Information Analysis Center, ERIM, Ann Arbor (1990).

Flynn, D. and B. Sieglinger, "Formalism for analyzing spatial, spectral, and temporal effects in a hardware-in-the-loop system involving a pixelized projector and a passive imaging sensor," *Proc. SPIE* **2741**, 46–68 (1996).

Goodman, J., *Fourier Optics*, McGraw-Hill, San Francisco, p. 9 (1968).

Holst, G., "Effects of phasing on MRT target visibility," *Proc. SPIE* **1488**, 90–98, (1991).

Howe, J., L. Scott, S. Pletz, J. Horger, and J. Mark, "Thermal model improvement through perception testing," *Proc. IRIS Specialty Group on Passive Sensors* **II**(1) 1–29, Infrared Information Analysis Center, ERIM, Ann Arbor (1989).

Huck, F., S. Park, D. Speray, and N. Halyo, "Information density and efficiency of two-dimensional sampled imagery, *Proc. SPIE* **310**, 36–42 (1981).

Legault, R., "The Aliasing Problems in Two-Dimensional Sampled Imagery," in *Perception of Displayed Information*, L. Biberman, Ed., 279–312, Plenum, New York (1973).

Meitzler, T., T. Cook, G. Gerhart, and R. Freeling, "Spatial aliasing effects in ground vehicle IR imagery," *Proc. SPIE* **1689**, 226–241 (1992).

Owen, P. and J. Dawson, "Resolving the differences in oversampled and undersampled imaging sensors: Updated target acquisition modeling strategies for staring and scanning FLIR systems," *Proc. SPIE* **1689**, 251–261 (1992).

Park. S. and R. Hazra, "Aliasing as noise: A quantitative and qualitative assessment," *Proc. SPIE* **1969**, 54–65 (1993).

Park, S. and R. Schowengerdt, "Image sampling, reconstruction, and the effect of sample-scene phasing," *Applied Optics* **21**(17), 3142–3151 (1982).

Schade, O., "Image Reproduction by a Line Raster Process," in *Perception of Displayed Information*, L. Biberman, Ed., 233–277, Plenum, New York (1973).

Schade, O., "Image gradation, graininess, and sharpness in television and motion-picture systems," *J. Soc. Motion Pict. Telev. Engs.* **61**(2), 97–164 (1953).

Sequin, C., "Interlacing in charge coupled imaging devices,"*IEEE Transactions on Electron Devices* **20**(6), 535–541 (1973).

Vollmerhausen, R., “Display of sampled imagery,” *IRIA-IRIS Proc.: Meeting of the IRIS Specialty Group on Passive Sensors* **1**, 175–192 (1990).

Vollmerhausen, R., “Impact of display modulation transfer function on the quality of sampled imagery,” *Proc. SPIE* **2743**, 12–22 (1996).

Vollmerhausen, R., R. Driggers, and B. O’Kane “The influence of sampling on recognition and identification performance,” *Opt. Eng.* **38**(5), 763–772 (1999).

Vollmerhausen, R., and R. Driggers, “Modeling the target acquisition performance of staring array sensors,” 1998 *IRIS Passive Sensors Proc.* 440000-82-X(1), **1**, 211–224, Infrared Information Analysis Center, ERIM, Ann Arbor (1998).

Vollmerhausen, R. H., “Display of Sampled Imagery,” Ch. 25 in *Electro-Optical Imaging: System Performance and Modeling,* L. Biberman, Ed., SPIE Press, Bellingham, WA (2000).

Vollmerhausen, R. H., R. G. Driggers, and D. L. Wilson, “Predicting range performance of sampled imagers by treating aliased signal as target-dependent noise,” *J. Opt. Soc. Am. A* **25**, 2055–2065 (2008).

Vollmerhausen, R. H., “Predicting the effect of gain, level, and sampling on minimum resolvable temperature measurements,” *Opt. Eng.* **48**(7), 076401 (2009).

Wittenstein, W., F. C. Fontanella, A. R. Newberry, and J. Baars, “The definition and the OTF and the measurement of aliasing for sampled imaging systems,” *Optica Acta* **29**(1), 41–50 (1982).

Chapter 5
Interlace and Dither

Interlace and dither reduce sensor sample spacing without increasing detector count. A high-resolution image frame is comprised of two or more lower-resolution fields or subimages taken sequentially in time. Between each field or subimage, a nodding mirror or other mechanical means is used to move the locations where the scene is sampled. Interlace and dither achieve high resolution while minimizing FPA complexity.

The mechanical operation of a dither or microscan mirror is illustrated in Fig. 5.1. A mirror is inserted between the objective lens and the FPA. The angular motion of the mirror about the scan axis is very small. Image translation on the FPA is subpixel (that is, less than a sample spacing). The figure indicates a horizontal translation of the image, but the mirror can be made to scan about any axis. The subpixel image translation provided by the scan mirror allows changing the sample phase of the FPA relative to the scene.

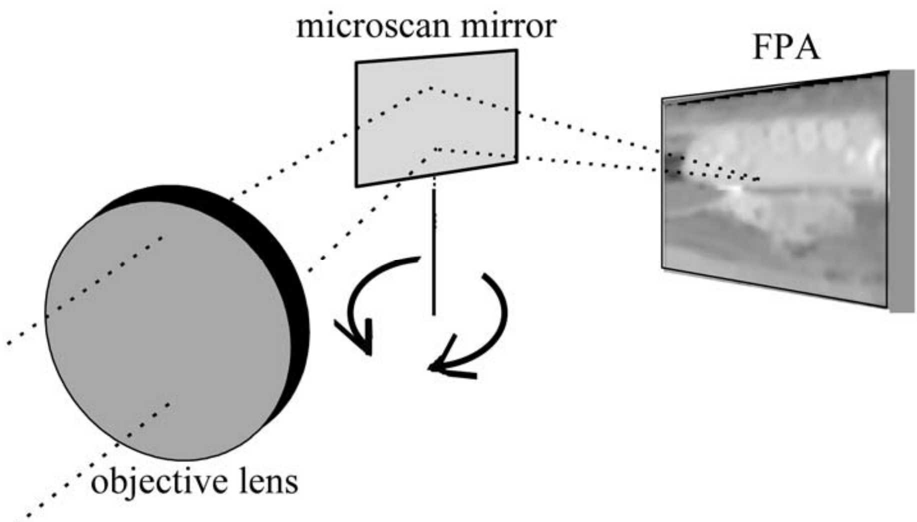


Figure 5.1 Operation of a microscan mirror. The mirror permits changing sample phase by subpixel translation of the image relative to the FPA.

Interlace generally implies that the subimages or field images are taken and displayed in time synchronism. That is, the pixels from subimages taken at different times are not combined and then displayed, but rather the time sequencing of the sensor field images is maintained at the display. The reduced-resolution field images are combined into a high-resolution image by the human visual system. Dither, on the other hand, generally implies that the field images are combined to form a higher-resolution image prior to the display.

Interlace is used to improve sensor sampling without increasing pixel rate or electronic throughput of the system. It takes advantage of the eye's ability to integrate multiple fields of imagery, presented in time sequence, into a higher-resolution scene.

Video display rates must be 50 or 60 Hz in order to avoid perceptible flicker, but each 50- or 60-Hz image does not need to display every pixel. Flicker is avoided in most circumstances by displaying every other pixel in each field. Flicker can occur when the image contains lines that are one pixel wide, such as lines in a graph. In that case, local regions of the image do not have approximately the same intensity between fields. Most natural scenes, however, do not contain such constructs, and the sensor presample blur mitigates the problem when it does occur.

Standard video uses two fields per frame and vertical interlace. That is, the video display is a vertical raster of lines, and every other line is displayed in every other field. In the U.S., the field rate is 60 Hz and the frame rate is 30 Hz. In Europe, the standard rate is 50-Hz field and 25-Hz frame. In this chapter, the U.S. standard is assumed.

Figure 5.2 illustrates video interlace. Each 1/60 of a second an interlaced sensor gathers every other horizontal row of scene samples. The video display shows every other horizontal line. Every other video line is collected by the sensor in each field. Every other video line is displayed during each field. Interlace is used because the full-resolution image needs to be produced and displayed only 30 times per second.

With an interlaced sensor and display, the human visual system integrates the full-resolution image whether the image is stationary or moving relative to the sensor. The 30-Hz update of pixel information is more than adequate to support the perception of smooth apparent motion. Exceptions to smooth apparent motion can occur. If the scene is comprised of very simple high-contrast structures, then image breakup can sometimes be seen during scene-to-sensor motion. However, for natural complex scenes, such breakup is very rare. For a human observer, interlace provides full-resolution imagery at half the pixel throughput.

The terms dither or microscanning are generally applied only to staring imagers. Although vertical interlace can certainly be accommodated by a staring imager, it is also common for the field subimages to be combined into a high-resolution image prior to the display. Figure 5.3 illustrates diagonal dither. At the

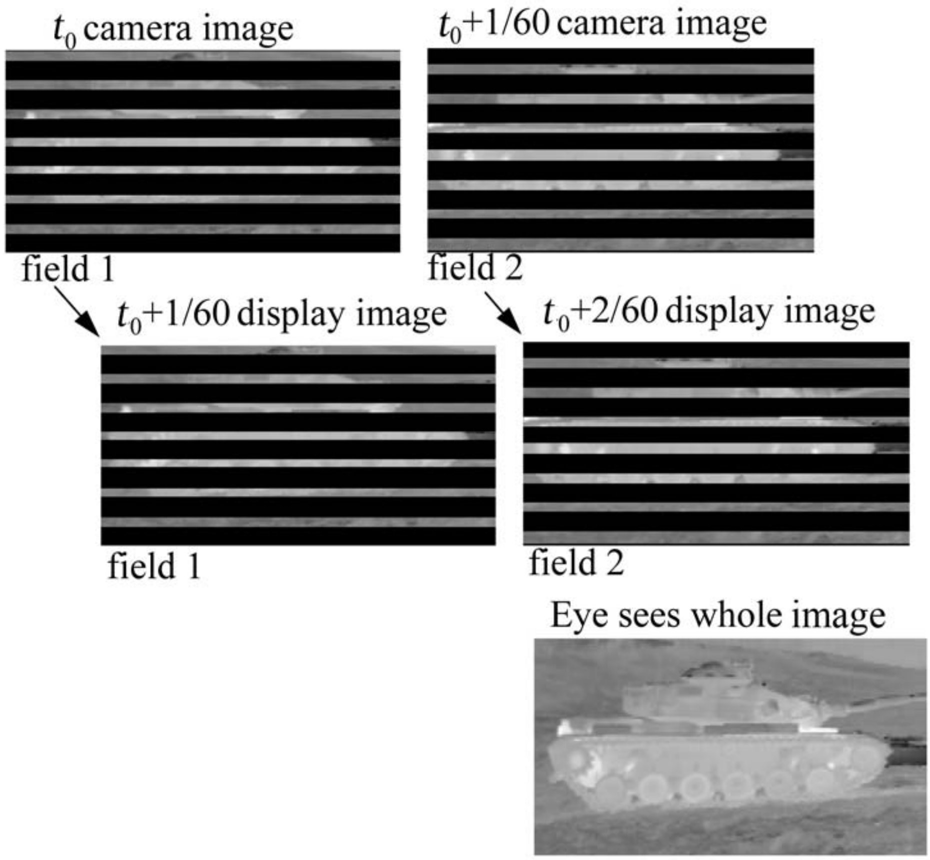


Figure 5.2 Illustration of interlace. At top, the sensor or camera collects the first field of imagery consisting of alternating horizontal lines. The first field is displayed 1/60 of a second later. The camera then collects the lines not collected in the first field, and these are subsequently displayed.

top of the figure, imagery is collected by the low-resolution FPA. A mirror is used to move the scene diagonally across the FPA, so that the sample locations are changed. A second subimage is then collected. The two subimages are combined in the electronics to form a higher-resolution image that is then displayed. Combining the high-resolution image prior to the display gives more flexibility in the dither sampling pattern. Extra samples can be added horizontally or diagonally as well as vertically, and the standard video display is still used. Combining pixels from fields taken at different times, however, can lead to artifacts in the displayed imagery when there is scene-to-sensor motion.

The sampling advantages of interlace and dither are described in the next section. The sampling advantages are described for a static scene; there is no scene-to-sensor motion. Section 5.2 describes the impact of dither on sensor resolution and sensitivity. The final section of this chapter describes the effect of motion.

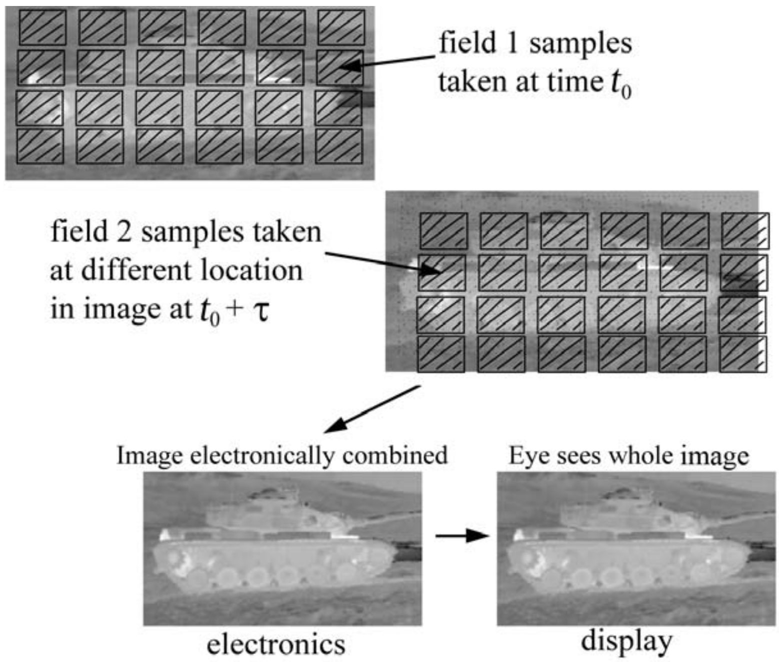


Figure 5.3 Illustration of dither. At the top, imagery is collected by an FPA. The scene is moved slightly across the array by a nodding mirror. At time τ later, imagery is collected at different sample locations. The combined samples from the two fields or subimages are electronically combined to form a higher-resolution image that is subsequently displayed.

5.1 Sampling Improvement with Static Scene

An FPA sampling grid is shown in Fig. 5.4. The sampling grid is a rectangular array of points. For these examples, the sample spacing is taken as 1 mrad. Common interlace and dither geometries are shown in Figs. 5.5 through 5.7.

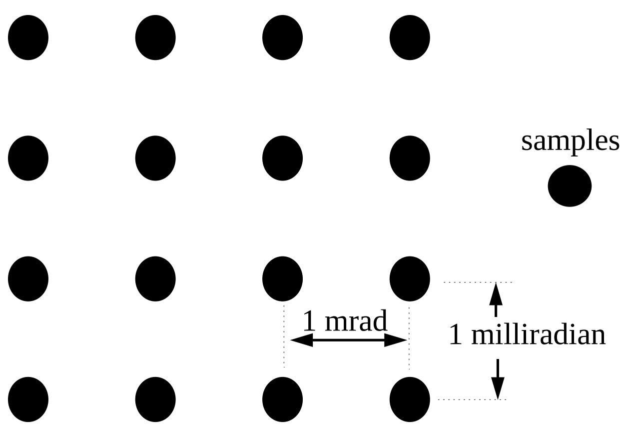


Figure 5.4 Sampling grid for each field or subimage. The example sample spacing is 1 mrad.

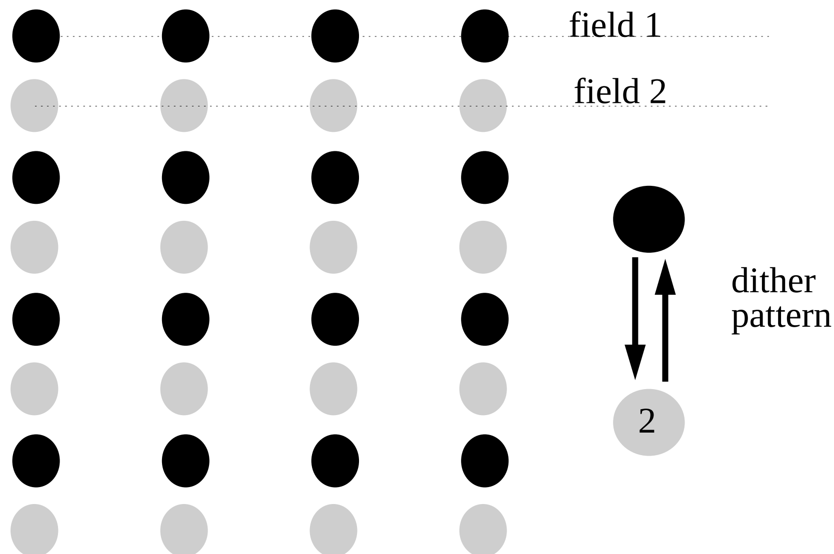


Figure 5.5 Vertical interlace or dither pattern adds extra vertical samples. The samples at black-dot locations are taken during field 1. The samples at locations of gray dots are taken during field 2. The horizontal sample spacing is still 1 mrad, but the vertical sample spacing is now 0.5 mrad.

Figure 5.5 shows vertical dither or interlace. The field 1 or subimage collects samples at the locations shown by the black dots. At time τ later, the field-2 samples are taken at the locations of the gray dots. Vertical dither adds vertical samples, while the horizontal sample rate remains the same as in Fig. 5.4.

Figure 5.6 shows two of several ways of dithering both horizontally and vertically. Subimage 1 is collected by the FPA with the microscan mirror in a stationary mode. The microscan mirror moves the sample locations to position 2. While the mirror is moving, the FPA samples from subimage 1 are read out. The scan mirror stabilizes at position 2, and subimage 2 is collected. The time between sampling subimage 1 and subimage 2 is τ seconds. The mirror then moves the sample locations to position 3, and subimage 3 is collected. The process is repeated for subimage 4. If the scene is static, the effective sample spacing both horizontally and vertically is 0.5 mrad when all of the subimages are combined. If the scene is static, the horizontal and vertical dither patterns are mathematically separable. That is, the horizontal and vertical increase in sampling can be treated separately.

Figure 5.7 illustrates diagonal dither, which improves image sampling both horizontally and vertically. However, as is shown below, neither the horizontal nor vertical sample rate is doubled. The diagonal dither cannot be treated as separable.

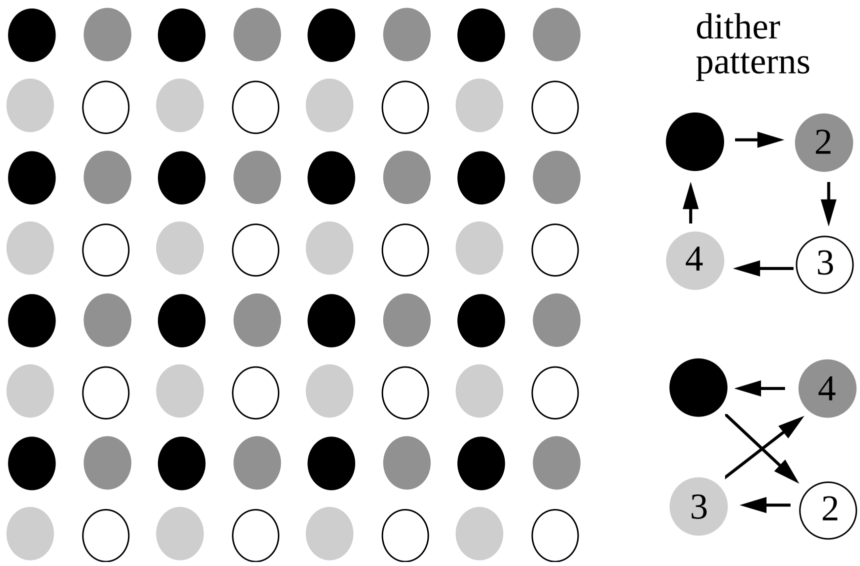


Figure 5.6 The scene is dithered both horizontally and vertically. The black, gray, light gray, and white dots represent samples taken during different fields. Sample spacing when all of the subimages are combined is 0.5 mrad both horizontally and vertically. The dither can be accomplished in multiple ways; only two of them are shown. If the scene is static, then all of the dither patterns provide the same result.

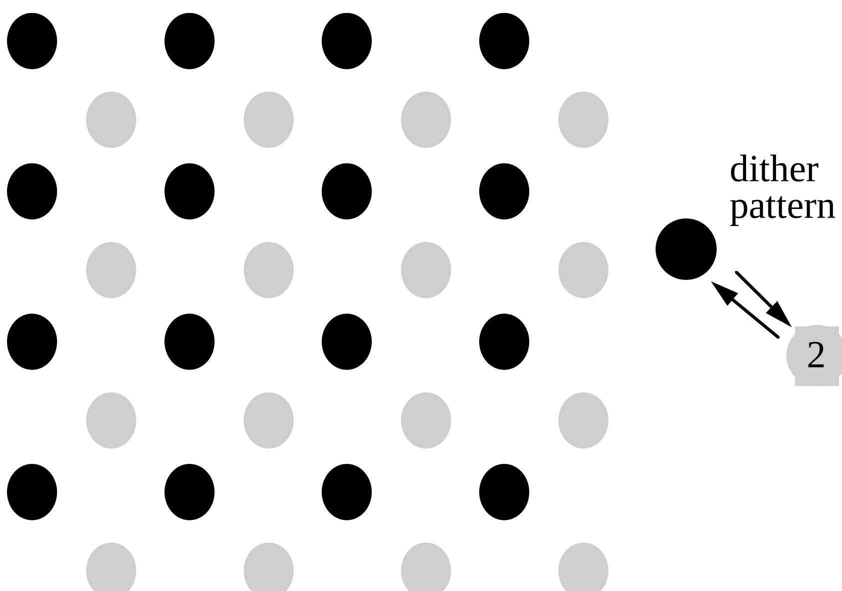


Figure 5.7 Diagonal dither. Black dots show sample locations during field 1. Gray dots show sample locations during field 2. Scene sampling has improved but neither horizontal nor vertical sampling is doubled.

Horizontal or vertical dither doubles the sample rate in the respective axis. Horizontal dither is used as an example. Recall from Chapter 3 that an image is sampled by multiplying by a comb of impulse functions. The displayed image is then reconstructed by convolving the display pixel shape over the delta function samples. Remember that $p_{ix}(x)$ is the intensity distribution associated with a single display pixel, and $P_{ix}(\xi)$ is the Fourier transform. Remember also that $h(x)$ is the presample blur, $s_{cn}(x)$ is the scene intensity distribution, and X is the sample spacing. For the present example, X is the sample spacing in one field or subimage. The display is generated by summing two sample sets taken τ time apart. In a manner similar to Eq. (3.2), the displayed image is represented by

$$\begin{aligned} i_{dsp}(x) = {} & \left\{ [h(x) * s_{cn}(x-d)] \sum_{n=-\infty}^{\infty} \delta(x-nX) \right\} * p_{ix}(x) \text{ samples taken at } t_0 \\ + & \left\{ [h(x) * s_{cn}(x-d)] \sum_{n=-\infty}^{\infty} \delta(x-nX-X/2) \right\} * p_{ix}(x) \text{ samples taken at } t_0+\tau. \end{aligned} \tag{5.1}$$

As long as the scene is static, $s_{cn}(x)$ is not a function of time. The samples can be described by a single sum:

$$\sum_{n=-\infty}^{\infty} \delta(x-nX) + \sum_{n=-\infty}^{\infty} \delta(x-nX-X/2) = \sum_{n=-\infty}^{\infty} \delta(x-nX/2). \tag{5.2}$$

The result shown in Eq. (5.3) is identical to Eq. (3.2) but with the X sample spacing halved. The sample rate is doubled by horizontal or vertical interlace or dither.

$$i_{dsp}(x) = \left\{ [h(x) * s_{cn}(x-d)] \sum_{n=-\infty}^{\infty} \delta(x-nX/2) \right\} * p_{ix}(x). \tag{5.3}$$

The improved sample rate can be illustrated in the Fourier domain. Equations (3.4) through (3.7) give the sampled response for an image generated with a single field. Equation (5.4) shows the sampled response that results from adding two fields or subimages. The expression in Eq. (5.4) is based on a static scene and on a change in sample phase ϕ of 180 deg or π radians between fields.

$$\begin{aligned} R_{sp}(\xi) = {} & P_{ix}(\xi) \sum_{n=-\infty}^{\infty} H(\xi-n\nu) S_{cn}(\xi-n\nu) e^{-jn\phi} \quad \text{(field 1)} \\ + \quad & P_{ix}(\xi) \sum_{n=-\infty}^{\infty} H(\xi-n\nu) S_{cn}(\xi-n\nu) e^{-jn(\phi+\pi)} \quad \text{(field 2).} \end{aligned} \tag{5.4}$$

For n odd, the two terms in Eq. (5.4) are of equal amplitude but 180 deg out of phase. For n even, the two terms are equal. Therefore,

$$R_{sp}(\xi) = 2P_{ix}(\xi) \sum_{n=-\infty}^{\infty} H(\xi - n2\nu) S_{cn}(\xi - n2\nu) e^{-j2n\phi}, \qquad (5.5)$$

and the sample frequency effectively doubles.

Diagonal dither is illustrated in Fig. 5.7. In order to describe diagonal dither, the sampling expression in Eq. (3.2) is expressly shown as two-dimensional in Eq. (5.6). To simplify the expression, $f(x,y)$ is substituted for the convolution of $h(x,y)$ and $s_{cn}(x,y)$; that is, $f(x,y)$ is the presample image. Separability of the presample image function is still assumed, so that $f(x,y) = f_x(x) f_y(y)$.

$$i_{dsp}(x,y,t_0) = \left\{ \left[f_x(x) \sum_{n=-\infty}^{\infty} \delta(x - nX) \right] * p_{ixx}(x) \right\} \left\{ \left[f_y(y) \sum_{m=-\infty}^{\infty} \delta(y - mY) \right] * p_{ixy}(y) \right\}. \qquad (5.6)$$

Equation (5.6) represents the set of samples taken during the first field of the diagonal dither. The second set of samples is represented by Eq. (5.7), where the image is moved rather than defining different sample locations. In Eq. (5.1), the samples from the second set are taken at different locations in space relative to the first set of samples. In Eq. (5.7), the image is moved relative to the sample set; the location of the sample set is not changed. The representations are equivalent.

$$i_{dsp}(x,y,t_0+\tau) = \left\{ \left[f_x(x - X/2) \sum_{n=-\infty}^{\infty} \delta(x - nX) \right] * p_{ixx}(x) \right\} \left\{ \left[f_x(y - Y/2) \sum_{m=-\infty}^{\infty} \delta(y - mY) \right] = * \, p_{ixy}(y) \right\}. \qquad (5.7)$$

Chapter 3 provides the details for finding the Fourier transforms of Eqs. (5.6) and (5.7). The sample phase is assumed to be zero for this example; this reduces equation complexity without affecting the conclusions.

$$I_{dsp}(\xi,\eta,t_0) = P_{ixx}(\xi) \sum_{n=-\infty}^{\infty} F_x(\xi - n\nu_x) P_{ixy}(\eta) \sum_{m=-\infty}^{\infty} F_y(\eta - m\nu_y). \qquad (5.8)$$

$$I_{dsp}(\xi,\eta,t_0+\tau)=P_{ixx}(\xi)\sum_{n=-\infty}^{\infty}F_x(\xi-n\nu_x)e^{-j2\pi(\xi-n\nu_x)X/2}$$
$$P_{ixy}(\eta)\sum_{m=-\infty}^{\infty}F_y(\eta-m\nu_y)e^{-j2\pi(\eta-m\nu_y)Y/2}. \qquad (5.9)$$

$$I_{dsp}(\xi,\eta)=P_{ixx}(\xi)\sum_{n=-\infty}^{\infty}F_x(\xi-n\nu_x)P_{ixy}(\eta)\sum_{m=-\infty}^{\infty}F_y(\eta-m\nu_y)$$
$$[1+(-1)^n(-1)^m e^{-j\pi\xi X}e^{-j\pi\eta Y}]. \qquad (5.10)$$

In Equation 5.10, the $e^{-j2\pi n\nu xX/2}$ factor reduces to $(-1)^n$ because $\nu_x = 1/X$.

The Fourier transform in Eq. (5.10) is not separable; that is, it cannot be factored into a function of ξ and a function of η. It follows from Fourier transform theory that the space function is not separable, either. That is, even if the presample function $f(x,y)$ is separable, the result of diagonal dither is almost always not separable.

The Fourier transform in Eq. (5.10) is separable only if either $f_x(x)$ or $f_y(y)$ is constant. For example, if the scene is a vertical bar pattern, the diagonal dither provides the same sampling performance as double sampling in the horizontal direction. If the scene is a horizontal bar pattern, then the vertical sampling rate is effectively doubled. However, for other than these bland or very regular scenes, the sampling is not effectively doubled on either axis.

If a diagonally dithered sensor is tested in the lab using horizontal and vertical bars, slits, or edges, the conclusion might be reached that diagonal sampling doubles the effective sampling rate in both directions. That is not the case. A sensor can be characterized by tests using horizontal and vertical bars only if the sensor system MTF is separable. The image resulting from the diagonal dither process is not separable. A diagonal dither does not provide a doubled sample rate in both the horizontal and vertical directions.

5.2 Resolution and Sensitivity

Dither improves sample rate for a static scene. Dither might also improve S/N. But dither cannot improve the presample MTF or the sensor transfer response. Figure 5.8 shows the transfer response (dark line) and spurious response for an undersampled sensor. The figure abscissa is frequency normalized to the sample rate, and the ordinate gives amplitude. The sensor is undersampled because the spurious response overlaps the transfer response of the system. This sensor has in-band aliasing, which cannot be corrected at the display. In-band aliasing can only be removed by better sampling. As shown at the bottom of the figure, adding dither doubles the sample rate and eliminates the in-band aliasing.

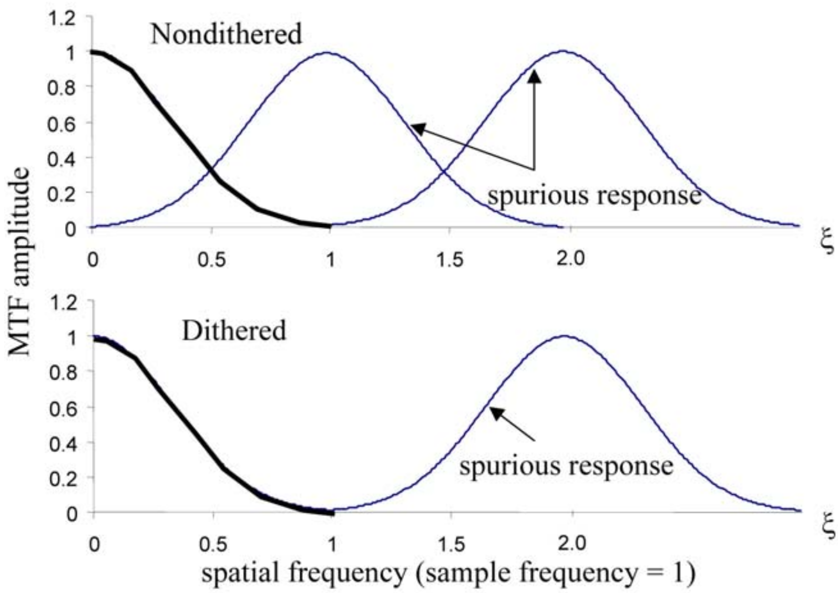


Figure 5.8 The top graph shows sampled response when dither is not used. The spurious response overlaps the transfer response (shown as a dark line). The in-band spurious response (spurious response below 0.5 on the frequency axis) cannot be removed by display processing. The bottom graph shows sampled spectrum when dither is used. Spurious response no longer overlaps the baseband; in-band aliasing is avoided.

The presample MTF and resulting transfer response can be degraded if the dither mirror is not stationary during the time the detector integrates scene signal. Scene-to-sensor motion during the detector dwell time causes motion blur. MTF can also be reduced if there is a discrepancy between the sample position on the FPA and the corresponding pixel position on the display. However, given that the microscanner positions the focal plane correctly and stabilizes the image on the detector for the signal integration period, the dithering action increases only the sampling rate.

For a well-sampled imaging system, no benefit is seen with dithering. However, for an undersampled imaging system, dithering provides a reduction or elimination of in-band aliasing.

Dither may or may not change the sensitivity of an imaging system. The detector FPA cannot be collecting signal during dither mirror motion, as the motion blur would be excessive. However, some detectors cannot integrate signal for the entire 1/60 of a second available at the typical display rate.

For example, assume that an InSb imager with an *f*/3 optical system has a typical integration time of 1 msec. The integration time is limited to the time it takes for the integration capacitor to fill up with photon-generated electrons. At the typical 60-Hz display rate, the imager has 17 msec to collect samples. The example InSb sensor can capture multiple subimages during each 1/60 of a second without loss of sensitivity.

One possible timing diagram for an InSb sensor with dither is shown in Fig. 5.9. Standard video display format is assumed; there are 60 display fields per second with 2:1 vertical interlace resulting in a 30-Hz frame rate. This example provides a 2:1 improvement in both horizontal and vertical sampling.

Two possibilities for the spatial sampling pattern are shown in Fig. 5.10. The detector integrates signal for about 1 msec at position 1 and time 0. During this period, the dither mirror is stationary. After the detector circuit stops integrating signal, the dither mirror moves horizontally to a second position and stabilizes. The second position changes the horizontal sample phase by 180 deg; that is, the image is translated by one-half pixel horizontally. While the mirror is changing positions, the InSb array must output to digital memory all of the samples collected at position 1. The mirror stabilizes at position 2, and the detector collects samples at that position at time 8 msec. These samples are subsequently read out of the FPA. The two subimages are combined in memory, and an image with improved horizontal resolution is displayed during the first display field.

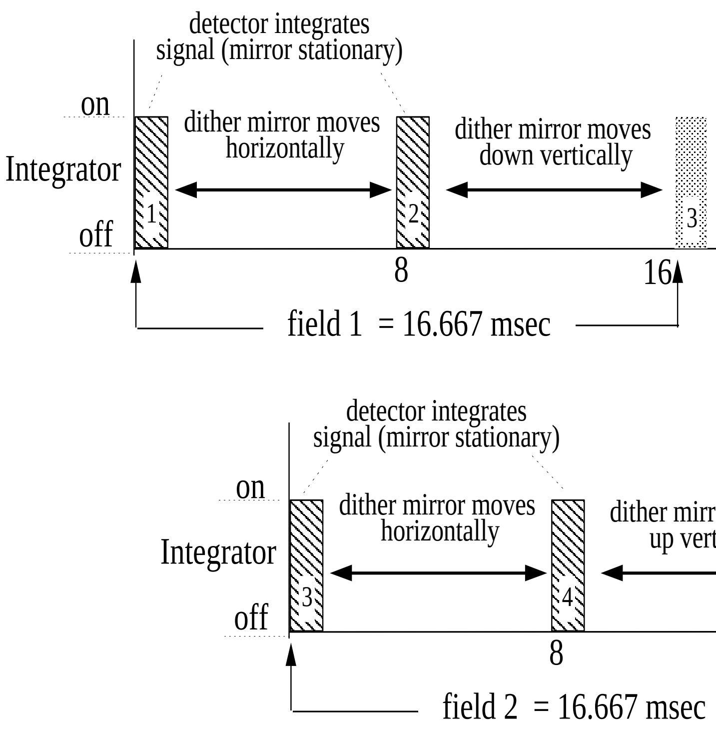


Figure 5.9 Dither timing diagram for InSb sensor showing detector integrating signal for 1 msec at each dither location. The dither mirror moves sample locations during dead time between detector samples. The spatial scan pattern is shown in Fig. 5.10. The spatial location of each sample is indicated by the number. In this example, samples from two dither locations are used to generate one field of displayed video. A frame of video consists of the two fields shown. A video frame contains samples from four dither locations.

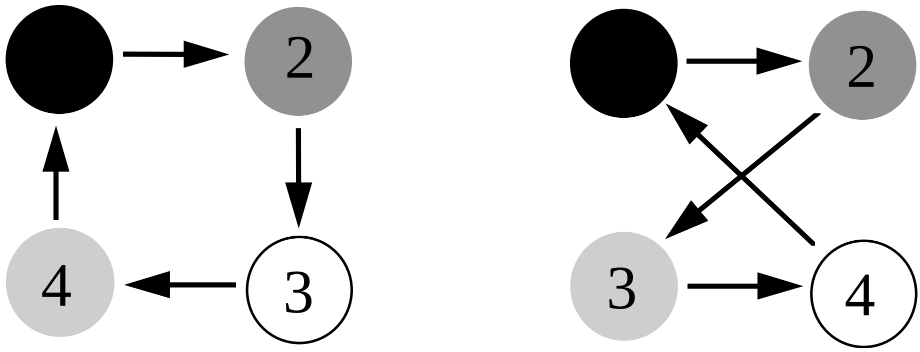


Figure 5.10 Dither patterns used to improve horizontal resolution and provide vertical interlace. The black, dark gray, light gray, and white dots represent sample locations. The numbers in the dots represent location sequence.

The image is then dithered by one-half pixel vertically, and two subimages are collected for the next display field. Again, the second display field has improved horizontal resolution, because two subimages are combined electronically before the display. Improved vertical resolution is provided by the vertical interlace.

In order for the dithering to be effective, the microscan mirror must change positions accurately and stabilize quickly. Furthermore, the InSb array must be capable of outputting all of the FPA samples at a 120-Hz rate. Given these capabilities, however, the extra samples provided by dither are achieved without loss in sensitivity.

In contrast, a PtSi system with the same optics integrates for 16 msec. Even with a zero microscanner movement time, the detector integration time for a two-point field dither would be cut to 8 msec. This factor of two reduction in detector integration time results in a square root of two reduction in S/N. Allowing realistic time for mirror motion, the sensitivity loss for PtSi would be substantial.

With today's technology, dither is most needed when using a low-resolution InSb or HgCdTe FPA with up to 320 × 240 elements. These arrays have high quantum efficiencies and limited photoelectron integration times. These qualities make dither attractive, because the improved sampling provided by dither can be achieved without significant loss of sensitivity.

5.3 Effect of Scene-to-Sensor Motion

Dither can lead to motion artifacts such as doubled or jerky images. To avoid motion artifacts, imagery should be collected and displayed with the same time sequence. That is, imagery collected in one sensor field should not be displayed more than once. Imagery collected in multiple sensor fields at different times should not be combined into one display field. Dither is not yet common in fielded systems. But the problems associated with using different update rates for the sensor and display are well documented.

There can be significant practical advantage to capturing imagery at a rate that is different from the display rate. Movies at the local cinema work exactly this way. A movie is recorded by shooting 24 pictures a second; each picture is projected twice to avoid perceptible flicker. This technique saves a lot of film.

Projecting each picture twice results in motion artifacts. Camera operators are taught to pan the camera slowly, and slow shutter speeds are used to blur the image. Keeping the audience from noticing the motion artifacts is part of the art of film making. The next time you are at the theater, look closely; you will see a double or jerky image as the camera pans a scene.

The "standard" next-generation thermal imager originally operated at a 30-Hz frame rate and was not interlaced. Display flicker was avoided by displaying each image twice; the first field of a standard video display was new sensor imagery, and the second field was a replication of the first field.

Noninterlace operation at 30 Hz was chosen for compatibility with digital image processing. Noninterlace operation provides a fully sampled "snapshot" for image processing, regardless of scene motion. Sensor operation at 30 Hz was chosen, rather than at 60 Hz, because it provides better sensitivity for the processor, the mechanical scanner is easier to build and is more efficient, and system throughput would be doubled. The doubling of throughput requires doubling the bandwidth of system components.

As with the movies, however, this display technique leads to motion artifacts. Unlike the movies, clear vision during rapid image motion can be critical in many applications. Two military examples in which an operator must perform a visual task during scene-to-sensor motion include using a targeting sensor to rapidly search an area and using a pilotage sensor to fly a helicopter. During the execution of these tasks, the operator sees doubled, jerky, or blurred images, and these artifacts can be very pronounced. Normal eye and head movement make the motion artifacts particularly annoying in night helicopter pilotage systems that use head-tracked sensors and helmet-mounted displays.

The motion artifacts described in this section are real and can be quite annoying to system users. Based on experiments and strong feedback from users, all next-generation imagers either operate at 60 Hz or are equipped with a "30- or 60-Hz" switch so that the user can select a mode without these artifacts. First-generation thermal imagers built with 4:1 interlace and displayed at 2:1 were discarded after the bothersome artifacts disappointed the customers.

In order to discuss the expected magnitude of the motion artifact problem, consider the two dither implementations illustrated in Figs. 5.9 and 5.10. Remember that a video field is 16.67 msec in duration, and there are two video fields per video frame. Image samples taken 8 msec apart are combined to form video field 1. Another set of image samples taken 8 msec apart are combined to form field 2.

This dither implementation captures two images and combines them into a single picture, making a double exposure. If scene motion occurs, the result is a blurred or double image. Two views of the moving object are presented by each display field.

Whether the blur or double image is bothersome or even noticeable depends on many factors. Those factors include motion rate, the field of view of the sensor, the transfer function of the sensor, the field of view of the display at the eye, display blur, and the nature of the task being performed. To calculate the visual offset angle V_{off} in degrees at the eye, let P_{rate} be the pixels per second motion on the display, let P_{ang} represent the angle in degrees of one display pixel at the eye, and let τ be the time between subimages (8 msec in the above example):

$$V_{off} = P_{rate}\, P_{ang}\, \tau. \tag{5.11}$$

Based on experience with a variety of military systems and tasks, limiting τ to 8 msec or less results in imagery that is suitable for most purposes. However, the artifacts resulting from τ values between 16 and 33 msec make imagery unsuitable for some military tasks.

It should be emphasized that none of the discussion in Section 5.3 applies to interlace. Interlaced sensors acquire and display image samples in a time-synchronous manner. The artifacts described in this section arise from a disconnect between the rate at which imagery is acquired by the sensor and the rate at which it is displayed to the observer.

Bibliography

Bui, T. and R. H. Vollmerhausen, "Sensor System Psychophysics" Chapter 26 in *Electro-Optical Imaging: System Performance and Modeling*, L. Biberman, Ed., SPIE Press, Bellingham, WA (2000).

Friedenberg, A., "Microscan in infrared staring systems," *Opt. Eng*. **38,** 1745–1749 (1997).

Mikoshiba, A., G. Izumi, T. Nakamura, and K. Igarashi, "Appearance of false pixels and degradation of picture quality in matrix displays having extended light-emission periods," *SID Digest* **XXIII**, 659–662 (1992).

Vollmerhausen, R. and T. Bui, "The effect of sensor field replication on displayed imagery," *SID Digest* **XXVI**, 667–670 (1995).

Chapter 6
Quantifying Visual Task Performance

This chapter describes an identification task and associated test methodology. Image quality is quantified by the ability to identify objects in a test set. Chapter 7 describes an imager resolution metric that predicts the ability of observers to perform the identification task. The resolution metric predicts the impact of imager blur, noise, and sampling on target identification.

This chapter also discusses the use of diverse target sets to establish operational utility. In a target acquisition sense, the focus is on the average probability of identifying many objects in the scene. The ability to identify a specific target is not addressed. The focus on diverse target sets greatly simplifies the analysis presented in Chapter 7. Also, if the goal of imager analysis is to select a good imager for a particular type of task, target sets provide a better way to quantify performance than specific target analysis.

Using standard target sets to check visual acuity is quite common. In Fig. 6.1, the patient is reading an eye chart in the optometrist's office. The doctor tells her to read the smallest line she can. Many others have tried to read this or similar charts. The doctor knows the letter size that represents average vision. Reading the chart provides a quantitative method of comparing vision among observers.

Using high-contrast letters is not the best way to check vision, however. A scene consists of many luminance levels. The eye achieves an integrated view of objects by connecting lines and surfaces. These lines and surfaces do not share a particular brightness throughout their extent. For example, the background immediately behind a target might not be uniform, and yet the eye sees the target silhouette. Perspective is gained from converging lines even if they vary in luminance with increasing range. Slight changes in hue or texture can provide an excellent cue as to the distance and orientation of an object. Acute vision requires not only the ability to discriminate small details that happen to have good contrast, but also the ability to discriminate small differences in shades of gray.

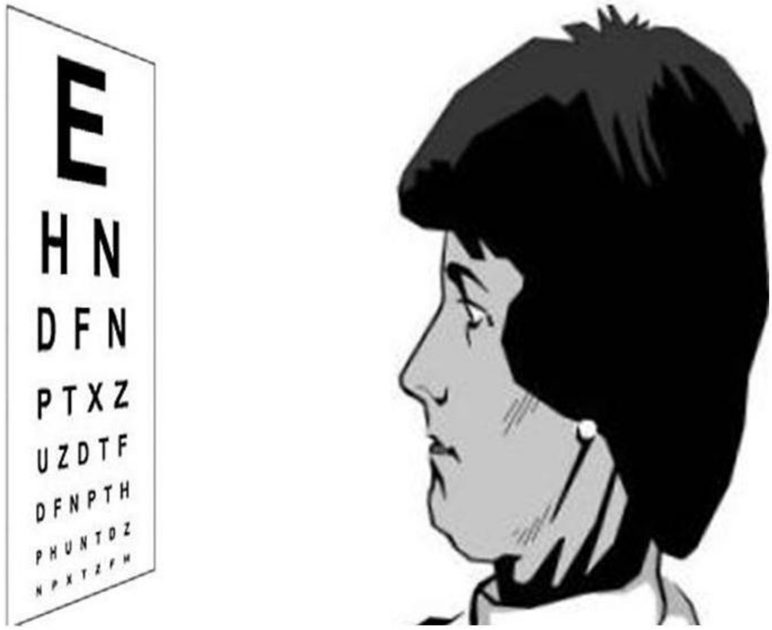


Figure 6.1 Patient reads an eye chart in the optometrist's office. This is a common way of checking vision against a known standard.

As an example, consider the picture of Goldhill in Fig. 6.2. The picture has an average modulation contrast of 0.22. The three-bar charts to the right have contrasts of 0.04, 0.08, 0.16, and 0.32. At the bottom of Fig. 6.2, noise is added and the pictures blurred. High-contrast details are still visible, but low-contrast details disappear. This is clearly illustrated by the bar charts. But also look at the windows in the houses; notice how the walking person blends into the bush behind him. A quantification of visual performance requires that resolution be measured for all shades of gray in the image.

The eye chart in Fig. 6.3 checks grayscale as well as angular resolving power. Like the original eye chart, the characters are various sizes. However, various contrasts are also represented. In this case, the chart is moved to various ranges. At each range, the observer tries to read all of the letters and numbers on the chart. The fraction of correct calls is large at short range and decreases as range increases. Plotting the fraction of correct calls versus range effectively characterizes visual resolution.

Using sets of real objects provides at least a sense of greater realism. The visual mechanisms underlying object identification are not understood. It is possible that character recognition does not sufficiently quantify visual performance. Figure 6.4 shows an observer using an imager to view the same set of vehicles at multiple ranges. If the vehicles in the set are carefully chosen, some identifications are simple and some challenging. The diversity of spatial features makes target sets preferable over characters.

Figure 6.2 Photograph of Goldhill beside bar charts with equivalent contrasts. The bottom pictures are blurred and noisy. Notice that blur and noise degrade high contrast less than they degrade low contrast. Using the degradation of high-contrast bars to estimate image quality does not work.

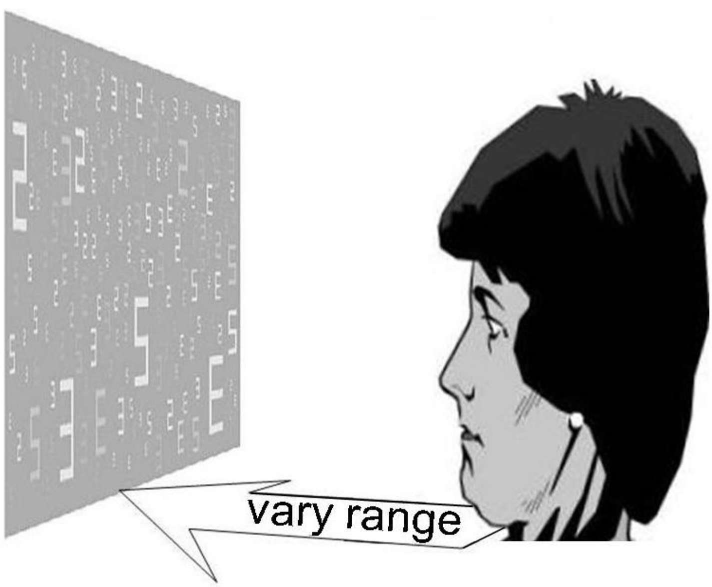


Figure 6.3 Eye chart in which the characters vary in contrast as well as size. The probability of correctly reading the characters quantifies visual resolution.

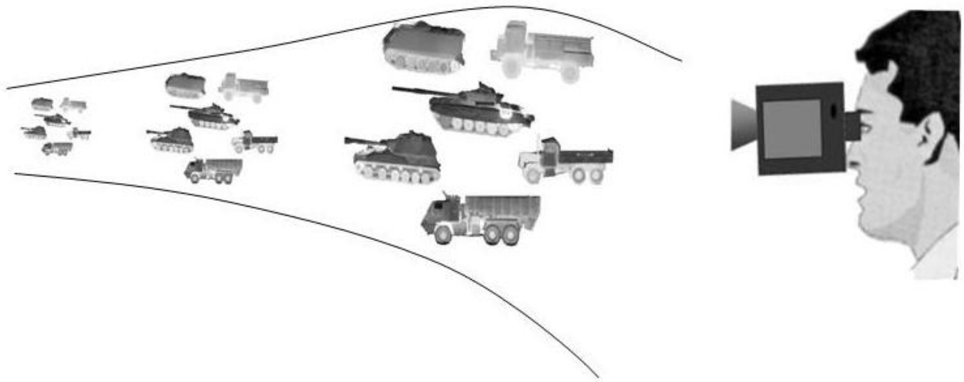


Figure 6.4 Observer using an imager to view the target set. The same set is positioned at three different ranges. The observer tries to identify each vehicle in the target set.

Section 6.2 discusses factors that affect target identification. Section 6.3 describes target signature measurement. Section 6.4 describes experimental procedures and discusses the attributes of a good test set. Section 6.5 provides procedures for field testing imagers. The meaning and importance of taking data at multiple ranges is also discussed. Sections 6.6 and 6.7 discuss selection of test sets. Before describing those details, however, the utility of a resolution metric is discussed. Section 6.1 describes the relationship between resolution models and expected operational performance.

Detection, recognition, and identification are general usage words. Each of us conceptualizes and adapts the meaning of these words depending on circumstance. This chapter discusses only target identification, which is associated with a specific experimental procedure.

6.1 Specifying and Evaluating Field Performance

Generally, hardware is purchased for use at an unknown time and place. The types of targets and localities might be known. Imagers are used on ships at sea, as security cameras in stores, and as gun sights on tanks. Certainly each operational use has different requirements.

However, the size, structure, orientation, and distinctive markings of the target are seldom known in advance. If the imager is used outdoors, then the weather is a dominant variable. Even if the exact target and target orientation are known in advance, the frequency spectrum seen by the imager is range dependent. All of these factors suggest that optimizing an imager for viewing one target at one range is suboptimal.

Hardware is purchased for future use. The operator, the place, the mission, and the target are unknown. Nonetheless, the inevitable cost-versus-weight-versus-performance trades must be made now. Image resolution is helpful in making these trades.

For example, an aircraft prime selects a 4-in diagonal display for the thermal imagery. The electro-optical engineers want a 6-in display to improve targeting performance. The program manager decides that a 6-in display might fit but needs to understand the benefit. The program manager is told that the larger display extends object identification range by ten percent. This may or may not impress the program manager. But the information allows the program manager to weigh range performance along with other factors.

The benefit of the 6-in display is couched in terms of a relative increase in range. The target, scenario, and environment are unknown. Improved resolution lets the observer perform a visual task at ten percent greater range. No target specifics are needed to understand the benefit.

Relative range performance is often an effective means of communicating operational requirements. For example, users of unmanned aerial vehicle (UAV) imagery are satisfied with the image quality. The problem is that flight altitude is too low, and the air vehicle is in jeopardy. The requirement might be to provide the “same imagery” but from a higher altitude. Knowledge of current imager resolution is sufficient to understand the technical requirements for the new hardware.

The users of a weapon sight might complain that the weapon “outshoots” the optical sight by a factor of two. A night pilotage system provides good imagery, but only on clear nights. A targeting sight provides good identification range against tanks, but only if they are exercised and hot. In each case, user experience provides a resolution baseline that defines good imagery. The design of the current hardware is known. The resolution of that hardware is known. A hardware upgrade should provide that same resolution at farther ranges or under different environments.

Evaluate the resolution of the current hardware. Agree on the new requirement for greater range, higher altitude, or operation against colder targets. Specify the same imager resolution but under the new operational scenario.

There are cases in which absolute range performance is desired. Will the rifle sight support positive identification of the target at 700 m? Will the targeting pod allow the pilot to see civilians in the target area? At what range does the tank sight allow a positive friend-or-foe decision? These are critical questions faced daily by people who deserve an honest answer. We do not know.

Resolution models do not predict absolute range performance. No existing model predicts the probability that a specific target is identified at a certain range. However, as the above examples illustrate, the ability to quantify relative performance is a useful tool.

The rest of this chapter is devoted to describing what is predicted by the resolution model. The factors that affect target identification are discussed. In particular, the role of the confuser set in establishing range performance is highlighted. The design of field tests is also discussed.

6.2 Factors That Influence Target Identification

Large high-contrast objects are easier to see than small low-contrast objects. We see and identify mountains farther than we do buildings, buildings farther than cars, and cars farther than people. Of course, the atmosphere can modify the relationship between size and range. On a rainy day, a car one mile in the distance is visible, but the mountains behind it are obscured. But it is clear that clouds and mist have reduced the contrast of the mountains. Range depends on contrast as well as size.

Although these intuitive concepts work well on the grand scale of car versus mountain, they do not work at all for car versus car or tank versus tank. Target acquisition tasks seldom focus on car versus mountain discriminations. More likely, the requirement is to discriminate which car or which tank, or to discriminate between car and tank. When comparing car to car, tank to tank, or car versus tank, the size and contrast are seldom important features.

In Operation Desert Storm, Egypt had both T62 Russian tanks and M60 American tanks. Iraq had T72 Russian tanks that looked very much like the T62. Figure 6.5 shows these three tanks. A friend-versus-foe decision is harder between the T62 and T72 than between the T62 and M60. In experiments, expert observers often mistake the T62 for the T72 and *vice versa*. The range at which a reliable identification of the T72 occurs depends on whether the alternative is a T62 or an M60. The question “At what range is the T72 reliably identified using this weapon sight?” is not answerable without defining the confuser target.

Target discriminations are always comparisons. Is it target A or target B or target C? The question is not whether the observer knows a T72 when he sees it. He does. The question is how well he can see the target. Does he see it well enough to differentiate between a T72 and an M60? Does he see it well enough to differentiate between a T72 and a T62? Identification range depends as much on the confuser set as on the target.

On a visit to the optometrist, the patient reads a line on the eye chart. The first letter is an L. That one is easy. The next letter is a W. Also easy. But the next letter might be a D or perhaps an O. If there were no O’s in the alphabet, the letter would be a D. If there were no O’s in the alphabet, the patient would read the line correctly. That would not make his eyesight better. It would make the visual task easier.

Figure 6.5 Pictures of the T62 and T72 Russian tanks and the M60 American tank.

Range performance is established using target sets. The same set is placed at each range. All of the targets in the set are identified. A procedure that uses different targets at each range changes task difficulty at each range. At one range the target set is L and W. At the next range the set is L and D. At the third range the set is D and O. Even though the observer understands that the letter will be an L, a W, a D, or an O, task difficulty still varies at each range. Task difficulty is kept constant by using the entire target set at all ranges.

6.3 Measuring Target Signatures

Using an imaging radiometer, let us find the radiance in watts per square meter (W m^{-2}) for each member of the target set. The images are collected at close range. As illustrated in Fig. 6.6, we segment the target from the background, find a background box with area equal to the target projected area, and calculate the RSS of target variance and the difference of the target and background mean. See Eq. (6.1).

$$C_{tgt} = \sqrt{\left(\text{target mean } - \text{ background mean}\right)^2 + \sigma_{tgt}^2}\ . \tag{6.1}$$

Assuming that the target set consists of N_{asp} aspects of N_{tgt} targets, Eq. (6.2) yields C_{tgt} for the set. C_{tgt} is just the average for the set. Target area is represented by presented area. Find the solid angle Ω_{tgt} subtended by each target. The target area is then Ω_{tgt} multiplied by range to target in meters. The average area of the target set L_{tgt} is used to find aliasing noise.

$$C_{tgt} = \frac{\sum_{m=1}^{N_{asp}} \sum_{n=1}^{N_{tgt}} C_{tgt\_m\_n}}{N_{asp}\, N_{tgt}}. \tag{6.2}$$

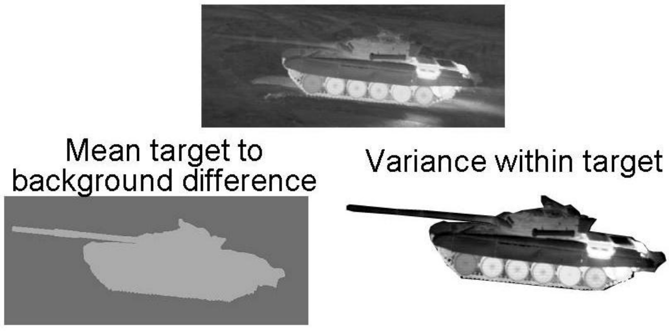


Figure 6.6 A radiometer collects radiance in W m^{-2} and calculates RSS of the variance within the target and the difference of target and background means.

6.4 Experimental Procedure

This section describes metric validation experiments. A robust resolution metric predicts the impact of blur, noise, contrast, system magnification, and sampling on target identification. Metric validation involves a series of experiments. In some experiments, the imagery results from viewing the targets at different ranges. This is what happens in field tests using actual imagers. In other cases, pristine imagery gathered at close range is degraded by adding blur, adding noise, lowering the contrast, downsampling, interpolating the imagery, or changing displays. In some experiments, one type of degradation is applied. In other experiments, multiple degradations are applied. One experiment applies different levels of blur. Another experiment applies different levels of noise. A third experiment applies both blur and noise. The experimental matrix represents a wide range of imager characteristics and target contrasts. The metric must predict the results of all of the experiments.

Figure 6.7 illustrates the first experimental constraint. The observer is isolated from the imagery collection. He views an image on the display and must identify the object based on that image. He has no situation awareness other than the training described in the following paragraphs.

The observer is trained on the target set. Only objects from the known set are presented. Target images from each range are presented randomly and one target at a time. Object identification experiments are forced choice. The observer must select one choice from a list of known objects.

Fifteen to twenty observers perform each experiment. The large number of observers is needed to average out variations in eyesight. The considerable variation in target identification capability must also be averaged by including a number of observers.

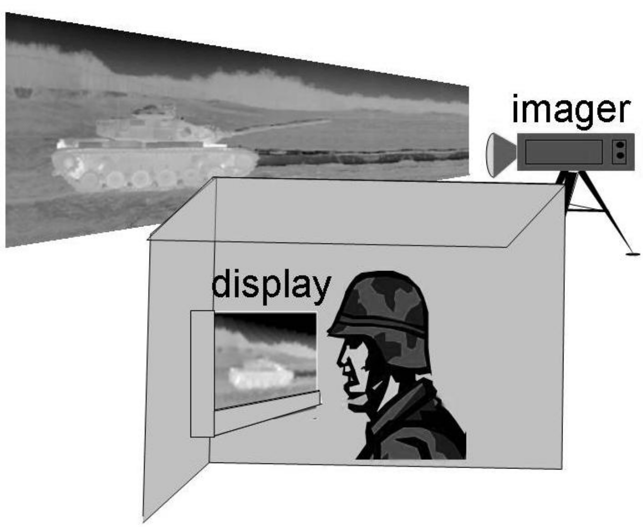


Figure 6.7 The observer has no situation awareness other than knowledge of the target set.

The target set used in many experiments is shown in Figs. 6.8(a) and (b). The set consists of 12 aspects of each of 12 tactical vehicles. The 12 vehicles are shown in Fig. 6.8(a). The 12 aspects are illustrated in Fig. 6.8(b) using the T55 Russian tank. The average area of this 144-vehicle-aspect set is 9.7 m^2.

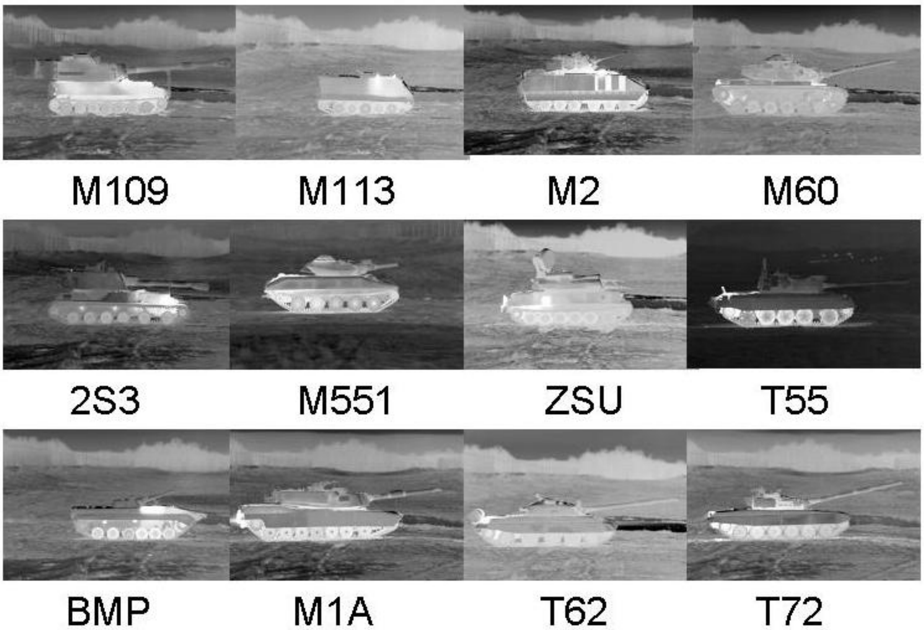


Figure 6.8(a) Twelve-vehicle target set. The vehicles include tanks, self-propelled Howitzers, and armored personnel carriers.

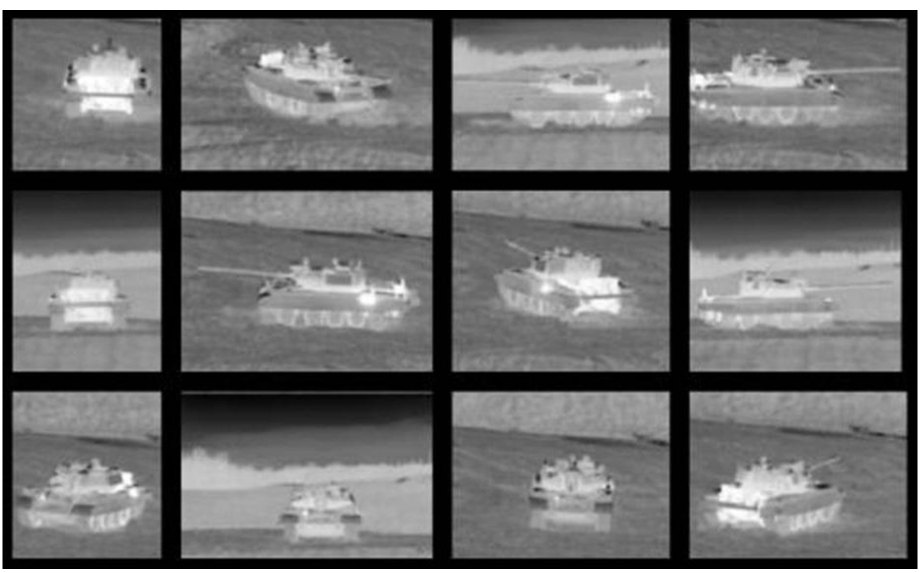


Figure 6.8(b) This figure illustrates the 12 target aspects used for each target in Fig. 6.8(a).

In the tactical vehicle sets, some of the vehicles look alike, some look somewhat alike, and some look distinct from the rest of the set. The same target set is used at all ranges. High probability is obtained at close range because the observers see the small details that differentiate similar targets. As range increases, targets that look alike are confused. At long range, only the unique targets are identified. The features that differentiate targets are small, medium, and large.

Images of the tactical vehicles are degraded by applying blur, adding noise, reducing contrast, etc. The degraded images are presented to the observers. Probability of correct identification (PID) is calculated using Eq. (6.3). Figure 6.9 shows example plots of PID versus range. The long-dash line is for the 12-target set. PID is the average probability at each range. That is, for N_{obs} observers and for an N_{tgt} set each with N_{asp} aspects,

$$PID = \frac{\text{correct calls}}{N_{tgt}\, N_{asp}\, N_{obs}}. \tag{6.3}$$

Figure 6.9 also shows PID for a selected nine-target set. In the 12-target set, the M109 and the 2S3 look alike. They are both self-propelled Howitzers. The T55, T62, and T72 Russian tanks also look alike. They represent the same generation of Russian design philosophy. We create the nine-vehicle set by removing the M109, T62, and T55. In other words, we remove the vehicles that look most like other members of the set. The PID range increases for the nine-vehicle set. The nine-target set is less confusing.

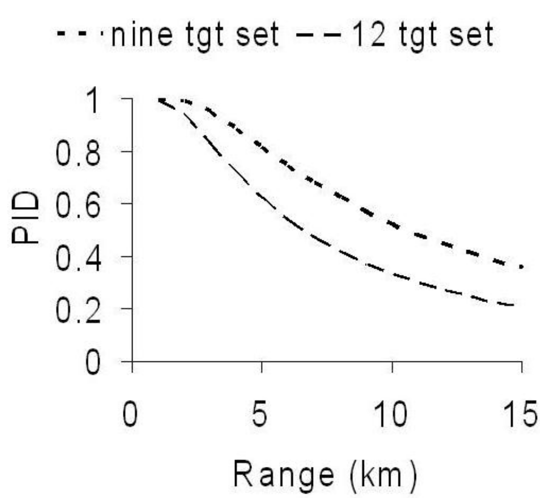


Figure 6.9 PID on ordinate versus range in kilometers on the abscissa. The nine-target (tgt) set is easier to identify because the targets in the set look less alike.

The procedures described in this section are used to validate a resolution metric. The metric is then used to evaluate the expected performance of an imager. The next section describes field test procedures to verify imager performance.

6.5 Field Test Procedure

The 12-vehicle set shown in Fig. 6.8 is not practical for field testing an imager. The full set is only needed for metric validation. The set of three aspects of eight tactical vehicles shown in Fig. 6.10 is more practical. The three aspects for each vehicle are chosen from one of the options shown in Fig. 6.11. The first option is front, right-rear-oblique, and left side. The second option is rear, left-front-oblique, and right side. Each option provides some view of the entire vehicle.

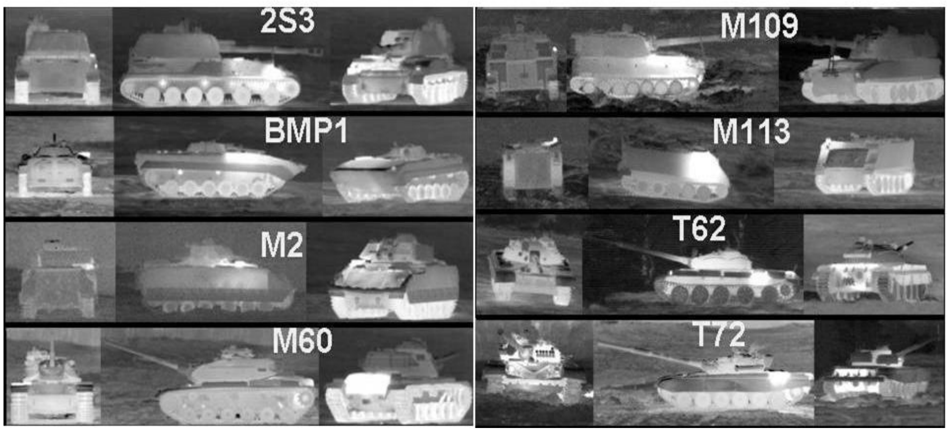


Figure 6.10 Tactical vehicle target set consisting of three aspects each of eight vehicles. Note that some vehicles, such as the M109 and 2S3, look alike. Other vehicles, such as the M113, have a distinct look.

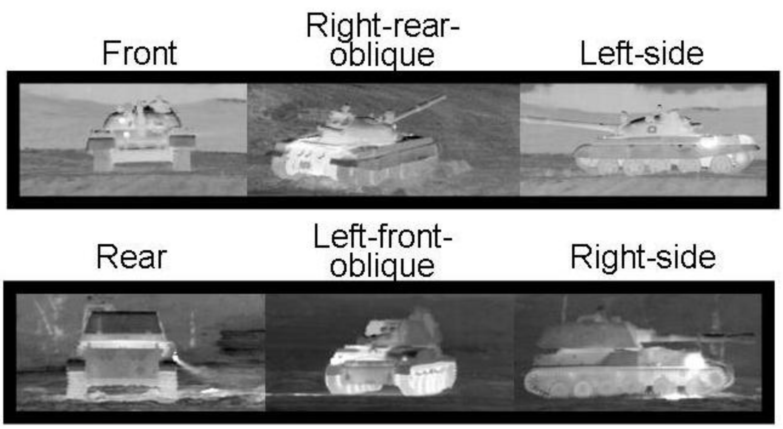


Figure 6.11 Top and bottom show two options for the three aspects of each target in the eight-vehicle set.

All of the vehicles in the set share common features: all are tracked, and all are bigger than a car and smaller than a tractor trailer. However, some of the vehicles in this set look very much alike; the T62 and T72 are highly confusable, and the M109 and 2S3 self-propelled Howitzers look quite similar. On the other hand, the M113 is a small cold box. It looks distinct from the other vehicles in the set. The set shares overall scale. But some members of the set are differentiated by small features. Other members of the set are identifiable by a unique silhouette. The spatial features that differentiate members of the target set are small, medium, and large.

The average area of this target set is 9.7 m^2. The vehicles should be sufficiently exercised that small variations in the atmosphere and other environmental factors are unimportant. It is likely that average target contrast is at least three to four Kelvin effective blackbody temperature. However, the target contrast should be measured during the field test.

All vehicle aspects are placed at each range. The ranges are selected to give a distribution of probabilities. At each range, N_{obs} trained and tested observers try to identify all of the vehicles. Probability of correct identification is the fraction of correct calls out of the total number of attempts. PID at each range for N_{asp} aspects of N_{tgt} targets is given by Eq. (6.3). In this case, the product N_{asp} N_{tgt} equals 24.

Figure 6.12 shows a typical PID-versus-range curve. At close range, even small spatial features are visible, and the entire target set is reliably identified. At slightly longer range, vehicles such as the T62 and T72 get confused. The M109 and 2S3 are also difficult to differentiate, especially at some aspects. However, the M113 is reliably identified even at long range. It has a unique shape and is not confused with other vehicles. It is the smallest and coldest vehicle in the set. However, that is irrelevant; it is the only box on the field.

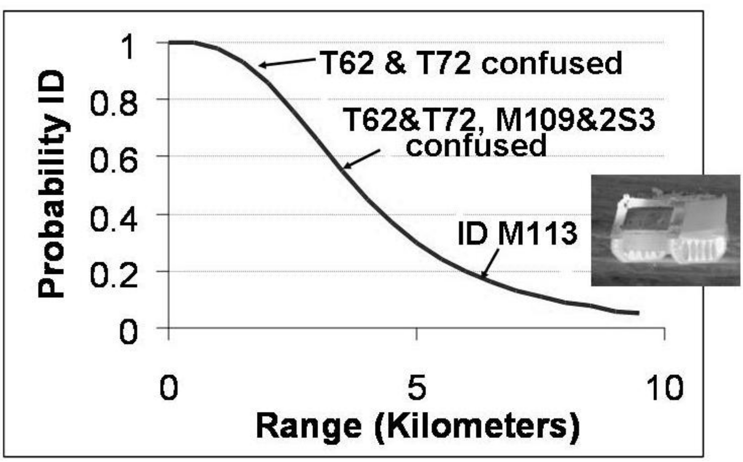


Figure 6.12 PID-versus-range curve for the eight-vehicle set. At near range, probability is less than 1.0 because the T62 and T72 are confused. At long range, the M113 is still reliably identified. PID is an average and does not apply to any one member of the set.

The PID curve does not represent the probability of identifying each object. PID is the fraction of objects identified. The PID curve does not apply to any one vehicle or vehicle aspect. A PID of 0.9 means that some vehicles at some aspects are confused. A high PID does not mean reliable target identification. A PID of 0.2 means that one-fifth of the objects are still identifiable. A low PID does not indicate unreliable target identification.

Consider how Fig. 6.12 applies to the employment of Egyptian armor in Desert Storm. The 0.9-PID range does not infer a high probability of making a correct friend-versus-foe decision for an Egyptian tank crew. In fact, PID falls below 1.0 precisely because the T62 and T72 are so confusable.

Each part of the PID range curve is equally meaningful. The fact that the imager does not allow reliable discrimination between the T62 and T72 at 2 km is meaningful. However, the fact that the M113 is still identifiable at 5 km is also meaningful. Each piece of experimental data contributes to knowledge of field performance.

When specifying performance, we describe the eight-vehicle test set. Equation (6.3) defines PID. Probability data for five or six ranges are obtained. Ranges are selected so that maximum probability is above 0.8 and minimum probability is below 0.3. Field data are corrected to compare to model predictions per the instructions in the next chapter. Corrected PID-versus-range data are compared to model predictions. The pass requirement for the field evaluation is that data and predictions compare with at least a 0.9 value for the square of the Pearson's coefficient (PSQ).

6.6 Test Sets Other Than Tactical Vehicles

Using tactical vehicles as a test set might not be convenient. Many factors make the use of tactical vehicles expensive and difficult to schedule. These factors include access to Army bases, the cost of renting the vehicles and the test site, plus the need for trained observers. Also, although the nature of the test objects is theoretically unimportant, using tactical vehicles might seem inappropriate to some users or customers.

Creating and verifying alternative test sets is a difficult process. To understand part of the difficulty, consider what would happen if a jeep and a truck were added to the eight-vehicle set. The PID-versus-range curve would be distorted. This would happen because the added vehicles would be too easy to distinguish from each other and from the rest of the target set. Adding unique targets to the set increases diversity in one sense—but the wrong kind of diversity.

It is true that the M113 stands out, more so from the side than from the front. This provides the small fraction of spatial cues that are visible at long range. But the fraction of cues visible at each range must be balanced, gradually decreasing as range increases. The T62 and T72 get confused at close range. The M109 and 2S3 get confused at slightly longer range. However, in both cases, the PID equals 0.5. The T62 is confused with the T72, but not confused with the M109. The

M109 is confused with the 2S3, but not with other vehicles. In both cases, the PID is well above chance at midrange. Eventually, the Russian tanks and self-propelled Howitzers get confused with the M60 and other vehicles and each other. These multiple interactions result in the PID-versus-range curve shown in Fig. 6.12.

The chart to the left in Fig. 6.13 illustrates the concept of a resolution test set. To the left, the letters have various sizes and contrasts. At each range, the character PID depends on the imager bandwidth and noise characteristics. To the right, the letters are all the same size and contrast. The amplitude of the Fourier transform of all of the letters is the same. For the letters on the right, the spatial cue differentiating the letters resides within a limited frequency spectrum. Predicting PID for the letters to the right depends on how the Fourier transform of the letters interacts with the imager blur and noise. In a good test set, there must not be a disproportionate number of big letters, small letters, dim letters, bright letters, etc.

The PID-versus-range curve in Fig. 6.12 is always obtained if the imagery is limited by white noise. When performance is limited by blur, contrast, or by a combination of factors, the target set must be properly constructed.

Both of the tactical vehicle test sets were chosen and validated without any modifications. Feeling confident, the 12 faces in Fig. 6.14 were selected as an alternative test set. The resolution model failed to predict PID. Further research showed that PID for the faces in Fig. 6.14 could be predicted by the resolution metric. However, it was necessary to include the Fourier spectrum of the faces in the model.

Papers on the specific object model (SOM) are listed in the bibliography. The SOM includes the Fourier spectrum of the target in the resolution calculation. This allows calculation of PID for specific targets. Eventually, it might be possible to use arbitrarily selected target test sets. However, those procedures are not yet sufficiently defined and validated.

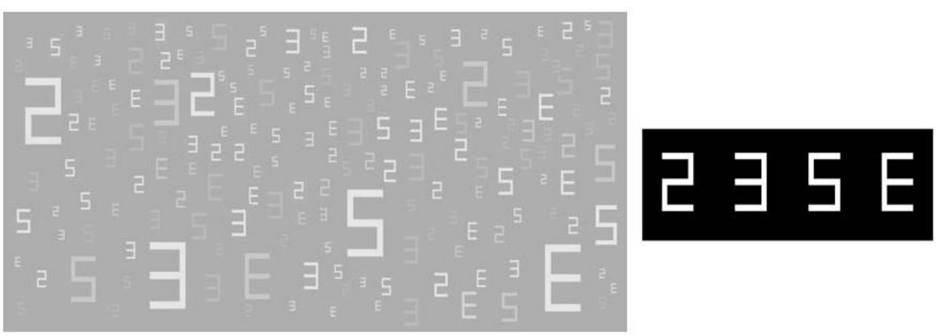


Figure 6.13 The chart on the left illustrates the test set concept. The letters are big, small, bright, and dim. Size and contrast distributions are uniform. On the right, the test objects have a defined Fourier spectrum. The spatial cues differentiating the letters on the right are not diverse.

Figure 6.14 Set of 12 faces that look sufficiently diverse to be a good test set. However, with this test set, range performance can only be modeled by using the frequency spectrum of the faces as target contrast. In other words, the target set is not sufficiently diverse that imager resolution alone predicts performance.

For the present, a suitable test set must be validated per the procedure in Section 7.5. The procedure verifies that the selected target set has the diversity of spatial cues necessary to stress imager resolution.

6.7 Field Testing Using Bar Targets

Bar targets are a convenient alternative to object sets. Bar targets are certainly "specific objects." However, a specific object model is not required. Resolution is defined by whether the observer sees the bar-space-bar-space-bar-space-bar. A validated minimum resolvable temperature (MRT) model is described at the end of Chapter 9.

The shortcomings of using bar charts include the subjective nature of the test and variation in MRT due to control settings. MRT also varies with sample phase. Figure 6.15 shows two plots of MRT versus spatial frequency. One plot is for optimum-sample phase. The other plot is for worst-case sampling. Optimizing sample phase during field testing is probably not possible. Furthermore, platform jitter might actually improve sampling. Another problem is that MRT tends to be flat to beyond the half-sample frequency. Since bar-chart contrast is difficult to control in the field, only the high-contrast frequency cutoff can be measured. Nonetheless, bar-chart testing provides a quick way to test resolution in the field. Also, the need to develop and validate a suitable target test set is obviated.

The MRT model is used to find thermal contrast in Kelvin versus bar frequency in cycles per milliradian. A bar frequency ξ_{cut} is selected for testing. Bar pattern design is illustrated in Fig. 6.16. Bar width is L_{bar} meters, and length is generally seven times the width. L_{bar} is related to ξ_{cut} and range R_{ng} in kilometers by Eq. (6.4).

$$L_{bar} = \frac{R_{ng}}{2\xi_{cut}}. \tag{6.4}$$

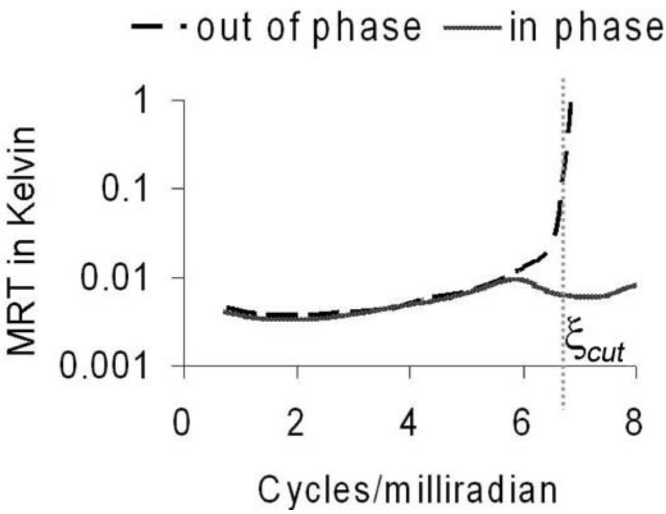


Figure 6.15 Plot of MRT with sample phase optimized (in-phase) and worst case (out of phase) sampling.

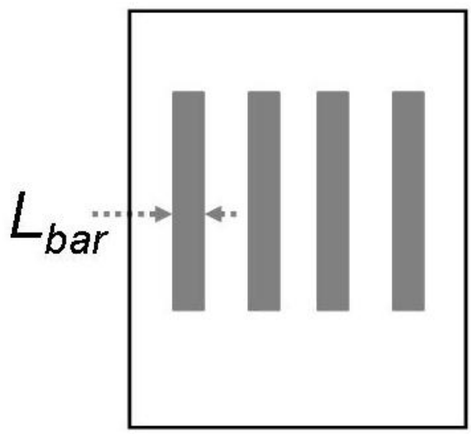


Figure 6.16 Illustration of four-bar pattern for field testing.

Chapter 7
Evaluating Imager Resolution

This chapter describes a procedure for evaluating imager resolution. A resolution metric is presented that predicts the effect of imager blur, noise, and sampling on the probability of correctly identifying targets. The targeting task performance (TTP) metric predicts the results of the target identification experiments described in Chapter 6. The resolution metric quantifies the ability to identify objects and, in general, to discriminate scene details.

Target acquisition task performance depends on how well information is coupled from the scene to the observer. System performance cannot be predicted without considering the characteristics of human vision. Imager noise is unimportant if it is too fine grained to bother the observer. Imager resolution is wasted if it far exceeds human acuity. Figure 7.1 shows an imager connected to a display and viewed by an observer. Observer vision is treated as a "black box." Psychophysical data are used to characterize visual thresholds and noise. The imager is evaluated by determining how well input spatial information is coupled to the human visual system.

Experimentally, resolution predicts the probability of identifying objects in a known target set. Mathematically, resolution is calculated by stimulating the imager with a broad frequency spectrum. The TTP metric is an integral of all of the spatial frequency content that exceeds visual thresholds. TTP quantifies how well input spatial information is coupled to the observer. Poor coupling reduces the probability of identifying objects. Good coupling improves the probability of identifying objects. TTP is validated by the fact that it predicts experimental object identification probabilities.

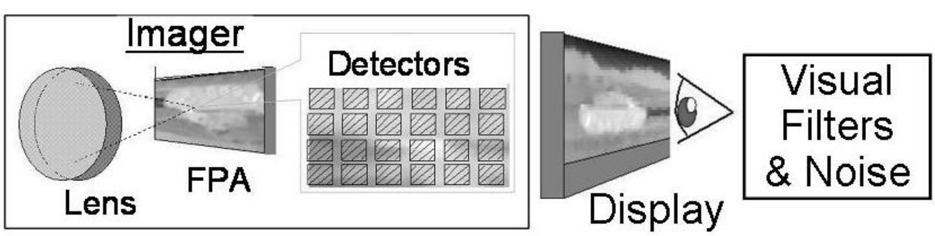


Figure 7.1 The resolution metric quantifies how well scene information is coupled to observer vision.

Section 7.1 discusses the logic of using a broad frequency stimulus to mathematically represent diverse target sets. Section 7.2 describes how to model imager gain and level and other features of the user interface. Gain and level controls are needed because the display has dynamic range limitations. Section 7.2 describes how the combined impact of all system component blurs and noise is established by referring each degradation to a single point in the signal chain.

In Section 7.3, the observer is added to the imaging chain. The observer is treated as a "black box." The Fourier domain blur and noise characteristics of the box are defined by psychophysical measurements. Section 7.3 describes how the contrast threshold function of an imaging system (CTF_{sys}) is determined. The CTF_{sys} is used in the TTP resolution metric described in Section 7.4. The procedure used to verify target set characteristics is described in Section 7.5.

7.1 Imager Evaluation Procedure

Available target acquisition models use a flat or "white" frequency spectrum to represent the diverse target set. The spectrum is shown in Fig. 7.2. The Weiner restoration filter also assumes a flat spectrum. This might seem illogical, since the Fourier transform of a real scene or image is essentially never flat. Both natural scenes and manmade objects tend to have a substantial low-frequency component in the Fourier transform.

But the typical transform of targets does not drive either target acquisition or image restoration. In both cases, the objective is not to see that there is a scene; the objective is to see what is in the scene. The goal is to identify as many objects in the scene as possible. A good imager resolves small details and presents a true grayscale. The imager is not a correlation receiver designed to optimize S/N for a typical scene. Rather, we want the imager to transmit whatever is in the scene to the display with as little degradation as possible.

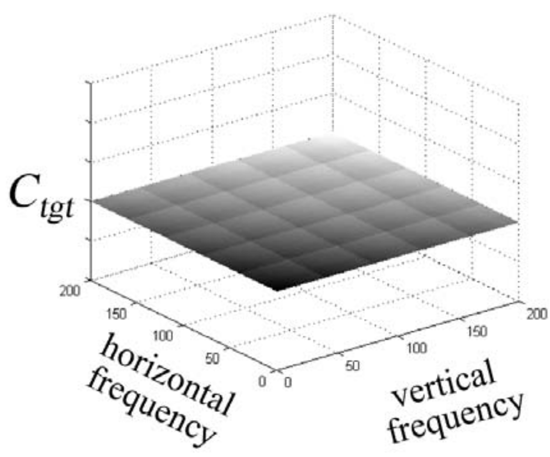


Figure 7.2 Flat target spectrum that has the same signal spectral density at all spatial frequencies. The signal spectral density (SSD) is C_{tgt} W mrad^2.

The flat spectrum represents individual points in the scene. The goal is to optimize the imager blur circle in order to transduce scene information with the best possible fidelity. The imager is optimized to present the small differences of shape and intensity that convey the uniqueness of individual scene elements.

C_{tgt} is set to equal the average RSS contrast as measured in Section 6.3. The signature used for evaluation has constant signal amplitude at all spatial frequencies. The spectrum is "white." Completely uncorrelated noise is known as white noise. For white noise, the power spectral density is a constant, and the autocorrelation function is an impulse function centered at zero. C_{tgt} is the signal spectral density (SSD) with units of watts per square cycles per milliradian (W/mrad^{-2} or W mrad^2).

Furthermore, the SSD is constant with range. Figure 7.3 illustrates the logic of this assumption. Imagers are used to view small objects such as knives, guns, or flashlights at close range. Medium-sized objects such as tanks and trucks are viewed at a few hundred meters or a kilometer or two. Large objects such as buildings and towers are viewed at longer distances. At each range, the objects have a variety of shapes, sizes, and contrasts.

For each C_{tgt}, TTP represents bandwidth or information capacity. TTP quantifies the capability of the imager to present scene detail in a form suitable for visual interpretation. The TTP metric quantifies imager resolution versus target contrast.

Creating a target set to test the TTP resolution metric is not straightforward. The spatial cues that differentiate targets are spread throughout the spatial frequency domain. The target spectrum is not white. The cue spectrum is white. The target set must have a distribution of small, medium, and large discrimination features. All features are seen at close range, and PID is high. At midrange, the smaller features are not visible, and PID drops. At still farther ranges, only the large cues are visible. The PID-versus-range curve depends on the correct distribution of spatial cues.

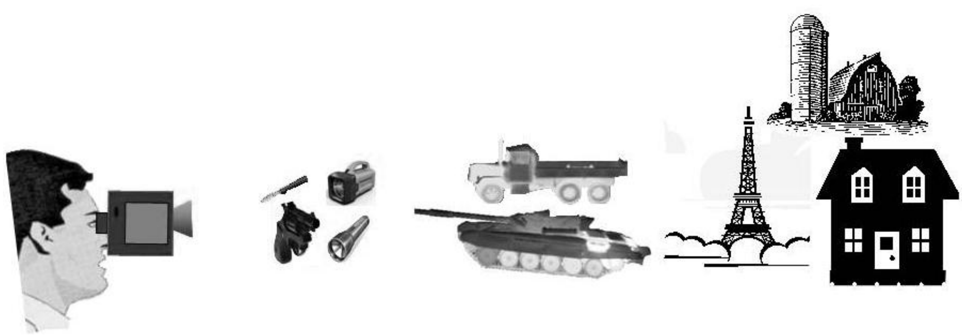


Figure 7.3 Small objects are viewed at close range, medium-sized objects at midrange, and large objects at long range. Each size category has a distribution of contrasts. The range-independent target signature represents different target sets at different ranges. C_{tgt} is varied to establish imager capability for different scene contrasts.

The spatial cues that differentiate one target from another might be obvious. However, representing discrimination cues in the Fourier domain is an unsolved problem. One tank hatch is somewhat square, the other tank hatch more rounded. The Fourier transforms of the hatches do not reveal the “roundedness” of a hatch. Nor is it clear how much the hatch shape contributes to differentiating between the tanks.

Section 7.5 describes the procedure for verifying that a test set has the correct distribution of spatial cues. Any set of objects works if the imagery is degraded by white noise. However, when imagery is degraded by blur or a combination of blur and noise, the target set must have cue diversity. The tactical vehicle sets described in Chapter 6 have the needed characteristics. Other test sets must be validated before they can be used in laboratory or field tests.

7.2 Modeling Gain, Level, and the User Interface

In Figure 7.1, the imager output is presented on a display and viewed by a human observer. Imager gain and level controls are used to put significant scene details within the dynamic range of the display. Varying imager gain alters the relationship between imager blur and noise and the bandwidth, internal noise, and quantization characteristics of the display. This section describes how to model the display/user interface.

Figure 7.4 shows the first step in quantifying the imager-to-display interface. The mapping between scene radiance in W mrad^{-2} and display luminance in foot Lambert (fL) must be known. The system gain S_{gain} in fL/W mrad^{-2} depends on both imager and display control settings. S_{gain} also depends on many imager and display design details as well as the state of any automatic gain mechanisms.

As a matter of practicality, it is much easier to estimate scene signature and display luminance than it is to search out all of the imager and display design details. Even with complete schematics of the hardware, the controls are seldom calibrated, and the state of automatic gain mechanisms is unknown.

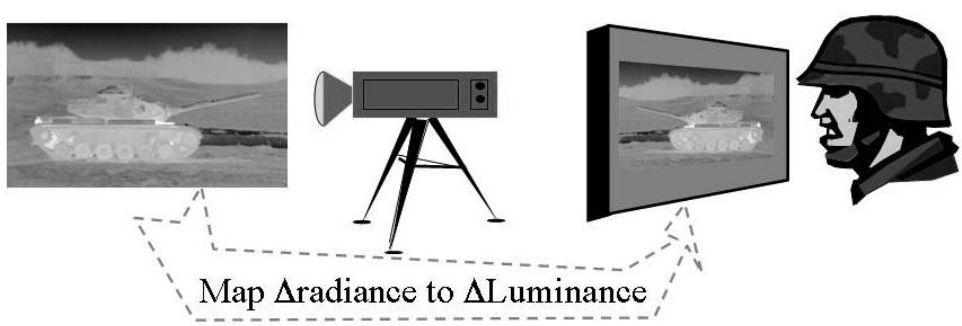


Figure 7.4 In order to quantify performance, the relationship between a delta luminance on the display and a delta radiance in the scene must be established.

In experimental situations, both the target radiance and display luminance are measured. In other situations, scene radiance and display luminance are estimated based on previous experience. For example, radiance from a poor thermal target is generally about 0.2 nanowatts per square meter (nW m^{-2}) in the 3.6- to 5-μm midwave infrared (MWIR). In the 8- to 10-μm long-wave infrared (LWIR), a poor thermal scene has a radiance near 1 nW m^{-2}. Typical and hot targets are ten and one-hundred times more radiant, respectively. These radiances correspond to 0.1, 1, and 10 K effective blackbody temperature for poor, typical, and hot targets, respectively.

Displays are set to photopic luminance of 50 fL or more during the day. At night, an operator trying to use a display while maintaining dark adaptation sets the display to 0.1–0.3 fL. In a dim room, the display might be 5–10 fL. These estimates of target radiance and display luminance are used to establish S_{gain} for typical operational scenarios.

Figure 7.5 illustrates how imager gain and level are used to put the target radiance in the dynamic range of the imager. As shown in Fig. 7.5(a), thermal signatures are low contrast. Everything radiates in the infrared spectral bands—background as well as target. Target radiance appears as a small difference on top of a large thermal flux. Imager level is used to remove most of the thermal pedestal. Removing the thermal pedestal enhances display contrast, as shown in Fig. 7.5(b). Imager level puts the target signature within the display dynamic range. Imager gain expands target radiance to cover a substantial portion of display grayscale.

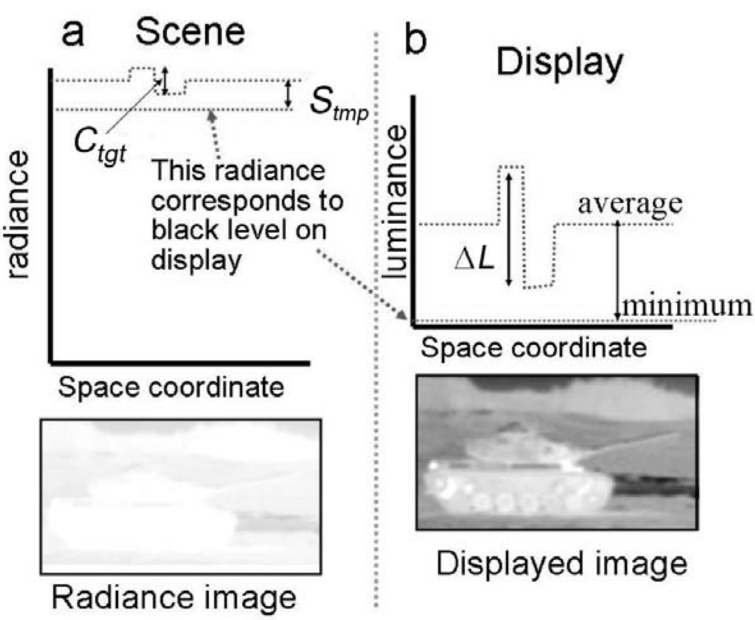


Figure 7.5 (a) Target radiance is low contrast. (b) Imager gain and level are used to put target radiance within the dynamic range of the display.

Of course, display brightness and contrast controls are also used to adapt the display to the scene. In the current discussion, the origin of the gain and level adjustments is not important.

S_{tmp} is the radiance that elevates display luminance from black to average. If $L_{display}$ is average display luminance, then S_{gain} is given by Eq. (7.1). ΔL is the change in display luminance that results from the target-to-background radiance difference C_{tgt}. S_{gain} is also calculated using C_{tgt} and ΔL:

$$S_{gain} = \frac{S_{tmp}}{L_{display}} = \frac{C_{tgt}}{\Delta L}. \tag{7.1}$$

Image gain state alters the relationship between Γ_{disp} and Γ_{imager}. This is illustrated in Fig. 7.6. In Fig. 7.6(a), the target is hot and imager gain is low. Detector noise is not greatly amplified by the imager electronics. The signal presented to the display is clean, and any type of display noise or imperfection dominates the displayed image. Figure 7.6(a) shows the clean signal, small imager noise, and input-independent display noise.

In Fig. 7.6(b), the target is cold and imager gain is high. Detector noise is now amplified by the imager electronics. The input to the display is noisy. Imager noise dominates the constant display noise.

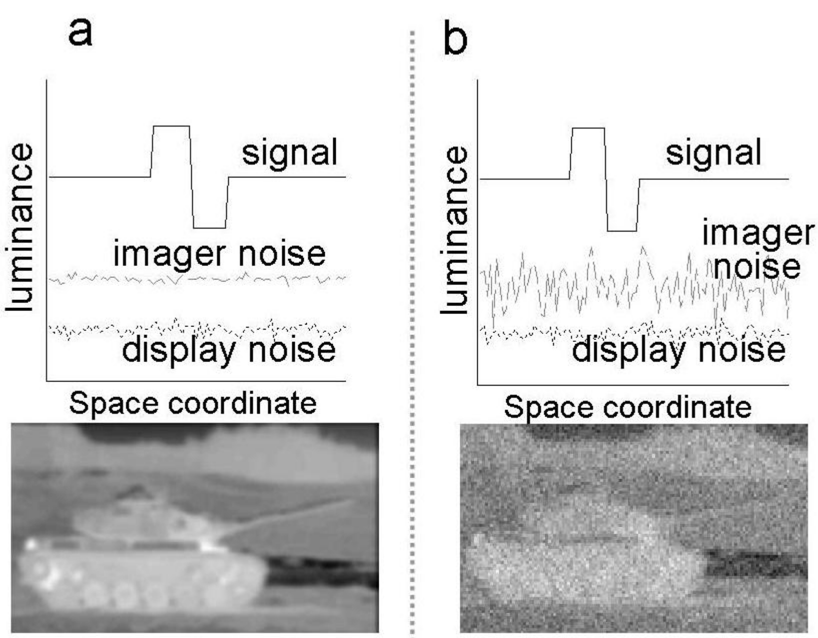


Figure 7.6 (a) Imager gain is low. There is little noise on the displayed signal, so any display noise dominates. (b) Imager gain is high, and a noisy signal is presented to the display. Here imager noise dominates.

The relationship between imager noise and display noise depends on gain and level settings. In Fig. 7.7, the scene radiance is blurred by line-of-sight jitter, optical diffraction, and summation over the detector. The detector array then adds photodetection noise and samples the scene. Alias noise is also added here; imager electronics applies a variable gain, and digital filtering is applied. Interpolation is an example of digital filtering. The display adds noise and filters all of the input signals. The display output in fL is viewed by the eye.

Different gains and blurs are applied to the various noise sources. A single point in the signal chain is picked, and the noise terms are scaled to RSS at that point. The display output is picked because the gain in fL/W mrad^2 is known. Also, psychophysical data on the visual system depends on display luminance and spatial frequency.

Each signal and noise source is converted to modulation on the display. Modulation contrast is defined as maximum signal minus minimum signal divided by maximum signal plus minimum signal. Signal modulation is demonstrated in Fig. 7.8. The arrow shows peak-to-peak for this sine wave. Modulation of this sine wave is 0.5 because the peak-to-peak signal difference is 5 fL and, therefore, is signal average.

The definition of modulation contrast is modified slightly in order to correctly RSS the noise terms. Contrast is defined as peak-to-peak signal difference divided by twice the average display luminance. Dividing all quantities by the same value allows the noise terms to be added in quadrature.

$$\text{modulation} \approx \frac{|\text{peak-to-peak}|}{2\,L_{display}}. \tag{7.2}$$

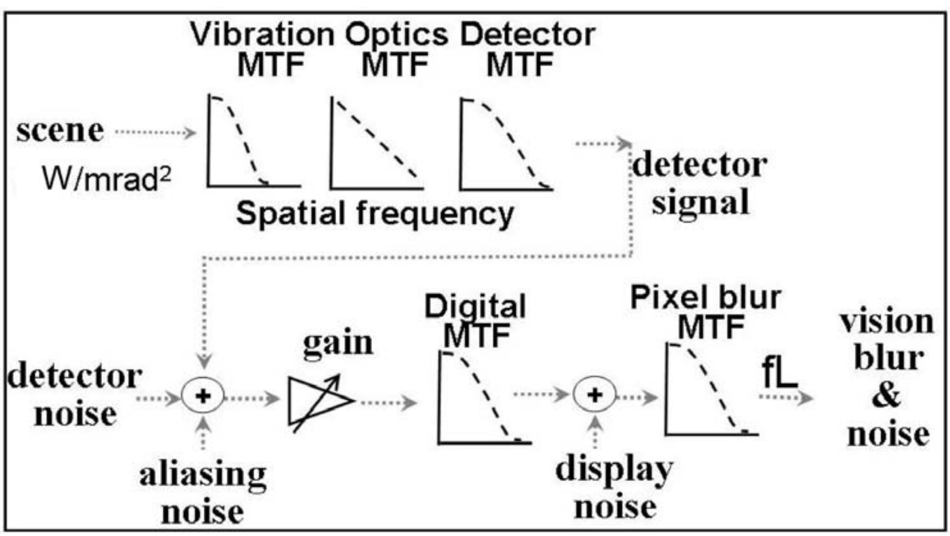


Figure 7.7 Schematic showing the relationship between various signal and noise sources.

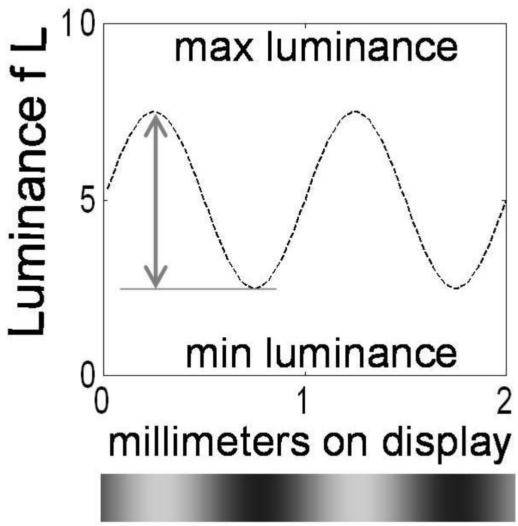


Figure 7.8 The top of the figure shows an intensity trace of the displayed sine wave shown at the bottom. Modulation of the signal is peak-to-peak divided by twice the average display luminance.

7.3 Observer Vision

The observer is treated as a “black box” and is represented by measured Fourier domain characteristics such as transfer response and frequency-dependent detection thresholds. Details on the psychophysical data used in the model are provided in the appendix. Only the essential concepts are summarized here.

One problem associated with all of the available psychophysical data is that separability in Cartesian coordinates is assumed. That is, data are provided for the observer viewing horizontal patterns and vertical patterns. It is therefore necessary to accommodate the available data by developing a separable imager model.

Nothing about the eye is separable in Cartesian coordinates. However, the use of separability is supported by the following argument. The known data support a horizontal and a vertical eye model. We suggest that the eye is sufficiently characterized by taking the geometric mean of the horizontal and vertical performance. In this section, separability in Cartesian is assumed.

A simple, engineering model of the eye and visual cortex is shown in Fig. 7.9. This figure shows the MTF associated with the eyeball and visual cortex. The figure also shows points where noise is injected into the visual signal.

Eyeball MTF and information on the visual cortex bandpass filters are available from psychophysical experiments. Vision science tells us that there is a high-resolution achromatic feature set in the visual cortex. Other experiments demonstrate that it is cortical noise that limits this feature set, not retinal noise. Therefore, noise on the display must be filtered by the eyeball MTF and bandpass filters before being added in quadrature with eye noise.

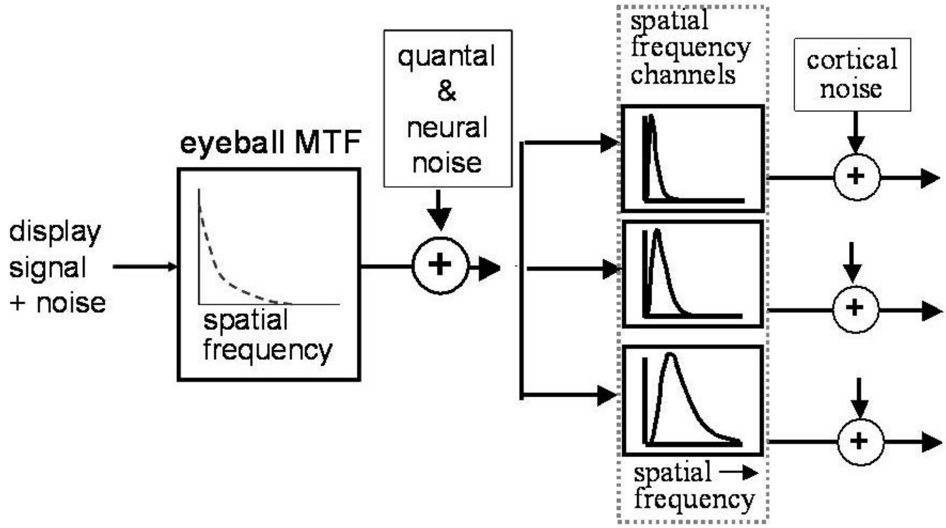


Figure 7.9 Engineering model of the eye showing the spatial filters and noise sources acting on the display signal.

The best available psychophysical data are the CTF. The inverse of CTF is the contrast sensitivity function (CSF). CTF and its inverse CSF quantify the spatial frequency response of human vision. A sine-wave pattern is presented to an observer, and a response is solicited as to whether the sine wave is visible. In Fig. 7.10, the observer is viewing a sine-wave pattern. While holding constant the average luminance to the eye, the contrast of the bar pattern is lowered until it is no longer visible to the observer. That is, the dark bars are lightened and the light bars darkened, holding the average constant, until the bar-space-bar pattern disappears. A decrease in contrast from left to right is shown at the top right in the figure. The goal of the experiment is to measure the amplitude of the sine wave that is just visible to the observer. In practice, techniques are used to arrive at actual eye threshold; the observer is not simply asked whether the bars are visible. A typical experimental technique is described in the appendix.

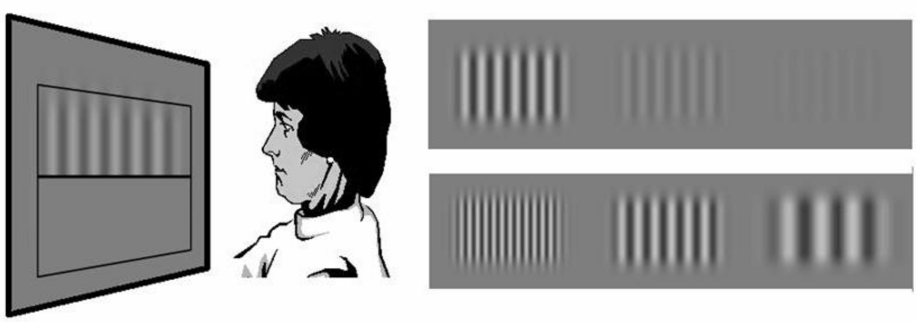


Figure 7.10 Experimental setup for measuring CTF/CSF. Top right shows variation in contrast. Bottom right shows variation in spatial frequency.

Contrast threshold is the sine-wave amplitude at which the observer is correct half of the time, independent of chance. There is a finite probability of seeing the sine wave at reduced contrast, and there is some chance of not seeing the sine wave at contrasts above threshold. The function $\delta(C/CTF)$ describes the probability of seeing a signal with contrast C when eye threshold is CTF. Like CTF itself, the function δ is available from published data.

The contrast threshold function is the best available measure of human vision. The observer model starts with measured naked-eye thresholds and then estimates the threshold elevation that results from adding imager blur and noise. That is, the observer looks at the world through the imager. The model predicts the system CTF (CTF_{sys}) that would be measured with the imager between the eye and the display. This is illustrated in Fig. 7.11.

Only equations for horizontal $CTFH_{sys}(\xi)$ are described. The equations for vertical $CTFV_{sys}(\eta)$ follow an identical logic. Presenting both would be redundant. Also, most dependencies are dropped, and only the variables of integration are shown.

Adjusting naked-eye CTF to account for blur is straightforward. At the top of Fig. 7.12, the sine wave is at observer threshold. In the middle, a lens in front of the eye blurs the image. The sine wave is now below threshold. At the bottom, the display contrast is increased until the observer is again at threshold. The increase of display contrast must just overcome the loss of contrast caused by the lens. Therefore, $CTFH_{sys}$ is given by Eq. (7.3):

$$CTFH_{sys}(\xi) = \frac{CTF(\xi)}{MTF_{lens}(\xi)}. \tag{7.3}$$

Contrast loss occurs due to glare as well as to blur. In Fig. 7.13, sunlight glare on the display might severely degrade the imagery. Let M_{dsp} represent contrast loss. L is average display luminance in fL. L_{min} is minimum luminance.

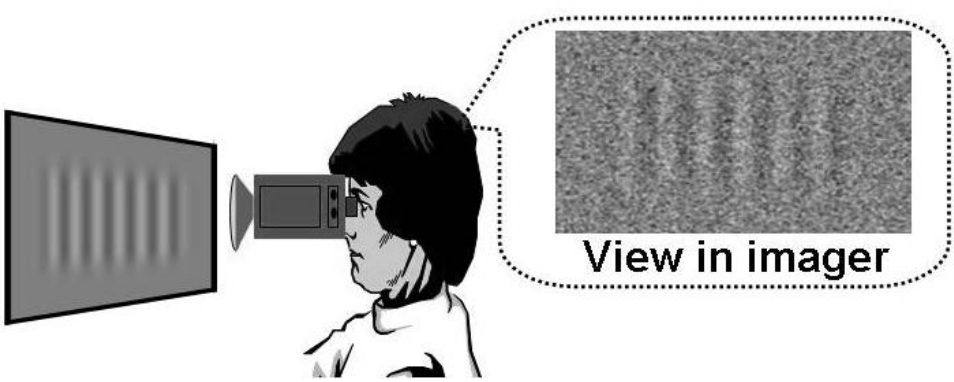


Figure 7.11 An observer views the display through the imager in order to establish the CTF_{sys}.

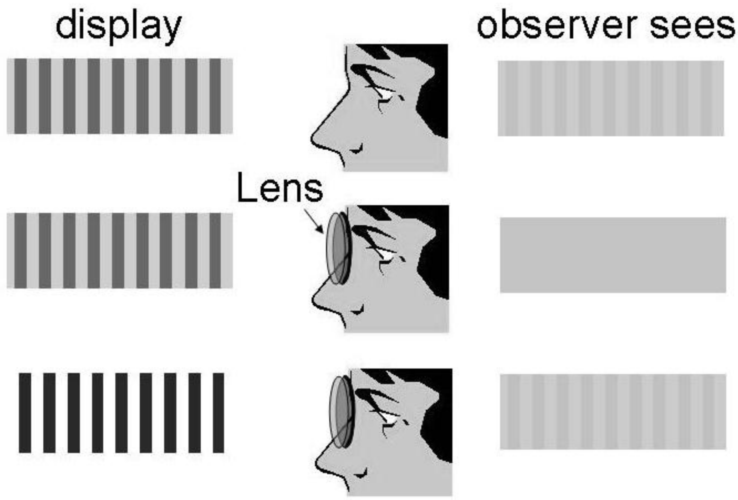


Figure 7.12 The sine wave at the top is at observer threshold. In the middle, the lens reduces visible contrast, and the sine wave is no longer visible. At the bottom, the sine wave contrast is increased just enough to overcome the MTF loss in the lens.

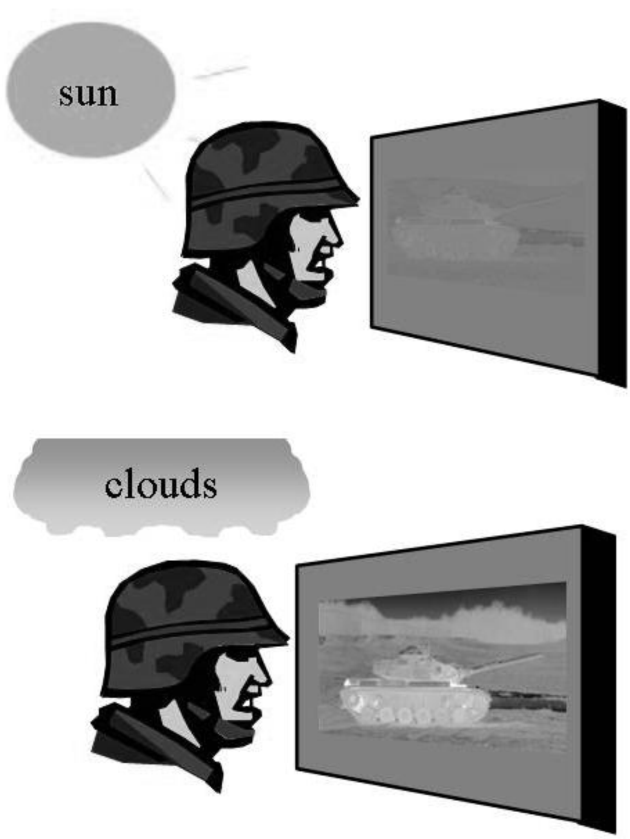


Figure 7.13 Sunlight glare on the display degrades the observed image.

To measure these luminances, we point a photometer at various locations on the display to find the minimum and average luminances. The modulation loss due to both glare and any brightness offset of the display black level is given by Eq. (7.4). M_{dsp} degrades $CTFH_{sys}$ as shown in Eq. (7.5).

$$M_{dsp} = \frac{L - L_{min}}{L}. \tag{7.4}$$

$$CTFH_{sys}(\xi) = \frac{CTF(\xi)}{M_{dsp}\, MTF_{lens}(\xi)}. \tag{7.5}$$

Noise is filtered by eyeball MTF and the visual cortex filters and then added in quadrature with eye noise. The number of bandpass filters is finite. However, the number is large enough that we assume that a filter exists for each spatial frequency presented to the observer. Some scalloping loss is being ignored, but the error is not large. The effect of the bandpass filters is to limit the noise that affects each spatial frequency in the image. A numerical fit for the bandpass filters is provided in the appendix.

In Eq. (7.6), n_{eye} is cortical noise, $\sigma_H(\xi)$ is imager noise filtered by the eyeball and visual cortex bandpass filters, and σ_H is the noise affecting threshold at grating frequency ξ. Using Weber's law, eye noise is proportional to display luminance. In Eq. (7.7), α is an empirically established proportionality constant.

$$CTFH_{sys}^2(\xi) = \frac{CTF^2(\xi)}{M_{dsp}^2\, MTF^2(\xi)} \left[\frac{n_{eye}^2(\xi) + \alpha_{temp}^2 \sigma_H^2(\xi)}{n_{eye}^2(\xi)} \right]. \tag{7.6}$$

$$CTFH_{sys}^2(\xi) = \frac{CTF^2(\xi)}{M_{dsp}^2\, MTF^2(\xi)} \left(1 + \frac{\alpha^2 \sigma_H^2(\xi)}{L^2} \right). \tag{7.7}$$

Psychophysical data such as eyeball MTF and naked eye CTF vary between observers. Adjusting model predictions based on known observer characteristics is, of course, sensible. Assuming equal skill at interpreting imagery, a person with poor eyesight does not acquire targets as well as a person with good eyesight.

However, fitting model predictions to data based on assumed variations in the observer obscures all model errors. The psychophysical data used in this book are based on young healthy people with known good vision. Parameters such as α are established empirically. Once established, however, the same value is used consistently for all types of imagers and to predict target acquisition probabilities and bar pattern thresholds. The observer parameters should be changed when better vision data becomes available. However, observer vision characteristics are not used on an experiment-by-experiment basis as a handle to adapt model predictions to data.

The observer model pertains to cone vision of foveated targets. Cones mediate vision down to about 0.01 fL. An observer at night might set his display as low as 0.1 fL. This low display luminance is a compromise between maintaining dark adaptation and effective viewing of display information. Therefore, the display luminance levels of interest here vary from about 0.1 fL to several hundred fL. However, there is little variation in visual thresholds above about 100 fL.

CTF_{sys} is now known and can be used to evaluate image quality:

$$CTF_{sys}(\xi,\eta) = \sqrt{CTFH_{sys}(\xi)CTFV_{sys}(\eta)} \,. \tag{7.8}$$

However, it should be remembered that the form of Eq. (7.8) results from psychophysical data that is insufficient to create a two-dimensional model. Eye performance is represented by the geometrical mean of horizontal and vertical performance. The application of Eq. (7.8) is limited to evaluating imager resolution. The separable model is not suitable for modeling target or scene structure.

7.4 Predicting Probability of Identification

The ability of an observer to resolve a target at range R_{ng} kilometers is quantified by the TTP metric Φ, which is the resolution in cycles per meter of the surface at range R_{ng}. CTF is a modulation threshold. C_{tgt} must be expressed as modulation. T_{con} is C_{tgt} expressed as modulation on the display:

$$T_{con} = \frac{C_{tgt}}{2\,S_{tmp}} \,. \tag{7.9}$$

$$\Phi = \left[\iint \delta\left(\frac{T_{con}}{CTF_{sys}(\xi,\eta)} \right) \frac{T_{con}}{CTF_{sys}(\xi,\eta)} \frac{d\xi\, d\eta}{R_{ng}^2} \right]^{1/2} . \tag{7.10}$$

Assuming separable mathematics, Φ is given by Eq. (7.11). As discussed earlier, Eq. (7.11) must be used when imager noise or aliasing is significant. This is because available psychophysical data allows us to calculate only $CTFH_{sys}$ and $CTFV_{sys}$.

$$\Phi = \left[\int \sqrt{\delta\left(\frac{T_{con}}{CTFH_{sys}(\xi)}\right)\frac{T_{con}}{CTFH_{sys}(\xi)}}\,\frac{d\xi}{R_{ng}} \int \sqrt{\delta\left(\frac{T_{con}}{CTFV_{sys}(\eta)}\right)\frac{T_{con}}{CTFV_{sys}(\eta)}}\,\frac{d\eta}{R_{ng}} \right]^{1/2}. \tag{7.11}$$

The degree of difficulty in identifying a target depends on how much the target looks like other objects in the target set. Φ84 is the value of Φ that results in 0.84 probability of task performance. Φ84 quantifies task difficulty. It is determined empirically for each target set. For the 12-vehicle set shown in Fig. 6.8, Φ84 is 14.8. For the eight-vehicle set shown in Fig. 6.10, Φ84 is 12.1. PID is a function of the ratio Φ to Φ84.

The error function (*erf*) relates resolution ratio Φ/Φ84 to PID. According to the central limit theorem, the normal distribution applies if many small independent effects additively contribute to the observation. This is true even if the mechanisms underlying the phenomena are unknown.

$$PID(\Phi / \Phi 84) = erf\,(\Phi / \Phi 84) = \frac{2}{\sqrt{\pi}} \int_0^{\Phi/\Phi 84} e^{-t^2}\, dt\,. \tag{7.12}$$

Predicting probability using the error function is consistent with basic model assumptions. The Φ defined by Eqs. (7.10) and (7.11) quantifies image quality by integrating suprathreshold image content. The thresholds associated with feature detection are established by a multichannel frequency-selective process in the visual cortex. The nature of visual feature formation and thresholding is not important in this discussion. It is only important that many independent spatial features contribute to a cognitive decision such as target identification.

To predict PID versus range, the following procedure is used. The naked eye CTF is degraded by imager blur and noise to establish $CTFH_{sys}$ and $CTFV_{sys}$. Φ is found at each range by a numerical integration corresponding to Eq. (7.11). PID is predicted using Φ and Φ84 in Eq. (7.12).

Ideally, a target acquisition model would not need Φ84. The model itself should comprehend the features that discriminate one object from another. At the current time, however, understanding and quantifying the features that discriminate one target from another is an unsolved problem.

To measure Φ84, we view the target set with any imager in the spectral band of interest. We place the target set at various ranges and record observer responses. We use Eq. (7.11) to calculate Φ and Eq. (7.12) to find the Φ84 that provides a good fit between PID data and predictions.

Φ84 is different between the emissive and reflective spectral bands. Target cues such as heat from the engine or tailpipe are not visible in the reflective bands. Within a given spectral band, however, Φ84 is a characteristic of the target set. Once known, the target set is used as an "eye chart" to verify the effective resolution of the imager.

The TTP metric was originally derived empirically. Published experiments had demonstrated that spatial frequency content contributed to scene understanding after exceeding contrast threshold. Furthermore, other experiments demonstrated that visual cognition was "bottom up" in the sense that spatial tuning did not adapt to stimulus. It was assumed that integrating excess contrast would be an effective image-quality predictor. Φ_{try} in Eq. (7.13) was calculated for $0.1 \leq r \geq 1$ and $-1 \leq q \geq 1$. The selection $r = 0.5$ and $q = 0$ was based on least average error between predicted and measured PID.

$$\Phi_{try} = \left\{ \begin{array}{l} \int \left[\delta\left(\dfrac{T_{con}}{CTFH_{sys}(\xi)}\right) \dfrac{T_{con}}{CTFH_{sys}(\xi)} \right]^r \dfrac{\xi^q \, d\xi}{R_{ng}} \\ \qquad \int \left[\delta\left(\dfrac{T_{con}}{CTFV_{sys}(\xi)}\right) \dfrac{T_{con}}{CTFV_{sys}(\eta)} \right]^r \dfrac{\eta^q \, d\eta}{R_{ng}}^{1/2} \end{array} \right\}. \tag{7.13}$$

TTP appears to represent image information content. Consider an information receiver that is a spatial frequency analyzer with resolution $d\xi \, d\eta$. Information content is the number of independent signal amplitudes discriminated by the frequency analyzer. At each spatial frequency, C_{tgt}/CTF_{sys} amplitudes are discriminated with probability δ. Equation (7.10) represents information content provided to the spatial feature set in the visual cortex.

7.4.1 Comparing experimental data to model predictions

Two aspects of experimental data must be corrected to compare to model predictions. Observers might select the correct target simply by chance. With eight vehicles, P_{chance} equals 0.125. With 12 vehicles, P_{chance} equals 0.083. The resolution model predicts PID independently of chance. Equation (7.14) corrects model predictions to compare to field or data:

$$PID_{field} = PID_{model}\left(1 - P_{chance}\right) + P_{chance}\,. \tag{7.14}$$

The second correction involves expert versus typical observers. We know from decades of field testing that some observers are simply good. Generally, good observers know that they are good, and so do their peers. But observer skill varies greatly. In a group of twenty trained, tested, and apparently well-motivated observers, some will be good, others mediocre, some not so good. When the imagery presents sufficient detail, good observers tend not to make mistakes. The average observer, however, makes mistakes that do not correlate to any physical measure of image resolution. These mistakes are simply "human error." This is a frustrating aspect of dealing with human performance. But the "ten percent mistake rate" is consistently observed if the identification task is difficult.

The mistake rate occurs for tactical vehicle identification. However, the mistake rate is not seen in experiments involving character or number discrimination. The mistake rate appears to result from difficult identification tasks.

Equation (7.15) is used to correct model predictions for typical observers. Equation (7.14) gives field probability for expert observers. It should be emphasized that "typical" represents the average over any group of twenty observers.

$$PID_{field} = PID_{model}\left(0.9 - P_{chance}\right) + P_{chance}\,. \tag{7.15}$$

7.5 Test Sets Other Than Tactical Vehicles

Generally, the purpose of a field test is to check system operation on the vehicle platform and in realistic dynamic environments. Problems with component mounting, electrical interface, stabilization, pointing and tracking, electromagnetic interference, and myriad other factors become apparent only when the imager is integrated into the system and operated in the field. The field test provides a demonstration of imager capability under realistic conditions.

Verifying system performance by measuring PID versus range requires an appropriate target set. Design and validation of the needed target set characteristics is the subject of this section. However, creating a target set and measuring range performance is a difficult proposition. In addition to validating the target set per the instructions below, the field exercise must provide a statistically significant PID database. Generally, validating operation by measuring PID versus range is difficult and expensive.

Imager and system operation can also be validated by a combination of laboratory measurements and field demonstrations. Imager resolution depends on hardware characteristics such as MTF, detector noise levels, and imager gain. Quantitative hardware characteristics are easy to measure in the laboratory but difficult to ascertain in the field. The combination of laboratory measurements and field demonstrations to discover integration and interface problems is an alternative to PID-versus-range measurements.

If the decision is made to measure PID versus range, then an appropriate test set is required. As discussed in Chapter 6 and earlier in this chapter, test sets must have sufficient cue diversity to stress the resolution capability of the imager. A suitable test set must be validated per the following procedure:

1. Assemble a test set. Objects in the set should share scale in the sense that a group of tactical vehicles, a selection of cars, or a collection of handheld objects share scale. The set should contain objects that look alike and objects that are distinct. The goal is to provide diversity in the shape and size of discrimination cues.
2. Select any imager in the spectral band of interest.
3. Collect imagery at various ranges. Use the entire target set at each range. Collect size and contrast data per Section 6.3.
4. Obtain average PID versus range using observers trained to identify the target set.
5. At each range, the target set size and average RSS contrast are known. Imager blur and noise characteristics are known. Calculate Φ at each range using Eq. (7.11).
6. Using Eq. (7.12), find $\Phi84$ that best fits predicted PID to data. The square of the Pearson's coefficient (PSQ) should exceed 0.9. If no good fit is found, the target set is not suitable for resolution testing.
7. However, a good fit often occurs even if the test set is not suitable. Even if a good fit between model and data is found, further testing is necessary.
8. Degrade the imagery by applying additional blur and/or noise. Retest to find the new PID-versus-range curve.
9. Using the $\Phi84$ found in step 6, the PSQ should exceed 0.9 for both data sets. If the $\Phi84$ predicts all of the PID data, then the test set is suitable for resolution testing.

Bibliography

Driggers, R., R. Vollmerhausen, and K. Krapels, "Target identification performance as a function of temporal and fixed pattern noise," *Opt. Eng.* **40**, 443–447 (2001).

Vollmerhausen, R., "Incorporating display limitations into night vision performance models," *IRIS Passive Sensors* **2**, 11–31 (1995).

Vollmerhausen, R. H., "Modeling the Performance of Imaging Sensors," Chapter 12 in *Electro-Optical Imaging: System Performance and Modeling*, L. Biberman Ed., SPIE Press, Bellingham, WA (2000).

Vollmerhausen, R. H., E. Jacobs, and R. Driggers, "New metric for predicting target acquisition performance," *Opt. Eng.* **43**, 2806–2818 (2004).

Vollmerhausen, R. H., E. Jacobs, J. Hixson, and M. Friedman, "The Targeting Task Performance (TTP) Metric; A New Model for Predicting Target Acquisition Performance," Technical Report AMSEL-NV-TR-230, U.S. Army CERDEC, Fort Belvoir, VA (2005).

Vollmerhausen, R. and A. L. Robinson, "Modeling target acquisition tasks associated with security and surveillance," *Appl. Opt.* **46**, 4209–4221 (2007).

Vollmerhausen, R. H., S. Moyer, K. Krapels, R. G. Driggers, J. G. Hixson, and A. L. Robinson, "Predicting the probability of facial identification using a specific object model," *Appl. Opt*. **47**, 751–759 (2008).

Vollmerhausen, R. H., R. G. Driggers, and D. L. Wilson, "Predicting range performance of sampled imagers by treating aliased signal as target-dependent noise," *J. Opt. Soc. Am. A* **25**, 2055–2065 (2008).

Vollmerhausen, R. H., "Representing the observer in electro-optical target acquisition models," *Opt. Exp.* **17**(20), 17253 (2009).

Chapter 8
Quantifying the Effect of Aliasing on Visual Task Performance

Aliasing degrades imagery and affects visual task performance. This chapter describes a model that predicts the effect of sampling on target identification. Aliasing is treated as noise. The combined effect of aliasing and detector noise degrades the system contrast threshold function (CTF_{sys}). The degraded (elevated) CTF_{sys} lowers the TTP resolution. The effect of aliasing on PID is predicted by the TTP metric.

Details on the thermal and reflective models are described in Chapters 9 and 10, respectively. Those chapters provide details on calculating CTF_{sys} for different types of imagers. This chapter explains modeling concepts.

Aliasing acts like noise because of the imaging task. We are interested in quantifying expected or average performance over many object identification attempts. At all ranges, the imager is presented with a diverse target set. The objects are placed randomly in the field of view. At each range, the task is repeated many times. The size and placement of spatial features varies from target to target. There are many targets in many different scenes. Aliasing acts like noise because of the combination of multiple targets, target diversity, random target placement, and the random sample phase of various target details.

The aliased signal is different from detector noise in two ways. First, aliasing disappears as the target contrast disappears. The amplitude of aliasing depends on target contrast. Second, the image corruption due to aliasing gets worse with increased range. This is because sampling is constant in angle space, and targets become poorly sampled as range increases.

Total noise is the quadrature sum of detector noise and aliasing. In order to sum the two noises, they must be properly scaled. Section 8.1 explains how signal and detector noise are spatially normalized in the thermal and reflective models. Section 8.2 describes the different temporal treatments of noise in detectivity versus photon-counting models. As far as the eye is concerned, noise is noise. However, the disparate treatment of noise in the two types of models results in different calibration constants. Section 8.3 describes the behavior of target signatures in imager coordinates. Section 8.4 compares and contrasts aliasing noise and detector noise. Sections 8.3 and 8.4 explain why aliasing noise is range dependent.

8.1 Model Treatment of Spatial Noise

The imager model assumes a constant angular stimulus. The scene is constant in angle space. This is illustrated in Fig. 8.1(a). The radiance of the bar patterns is expressed in cycles per milliradians, not cycles per meter. The model calculates imager resolution in angle space. At each range R_{ng}, linear resolution Φ in cycles per meter (m^{-1}) is calculated based on angular resolution Φ_{angle} in cycles per milliradian ($mrad^{-1}$). See Eq. (8.1). R_{ng} is range in kilometers.

$$\Phi = \frac{\Phi_{angle}}{R_{ng}}. \tag{8.1}$$

Assuming a constant angular stimulus when modeling imagers makes sense. Both detector noise and blur are constant in angle space. The linear size of imager blur grows in proportion to range, as does the scene radiance pattern equivalent to imager noise. The size increase of noise and blur is illustrated in Fig. 8.2. Imager properties are constant in angle space; imager resolution is determined in angle space. Once angular resolution is known, linear resolution at each range is calculated.

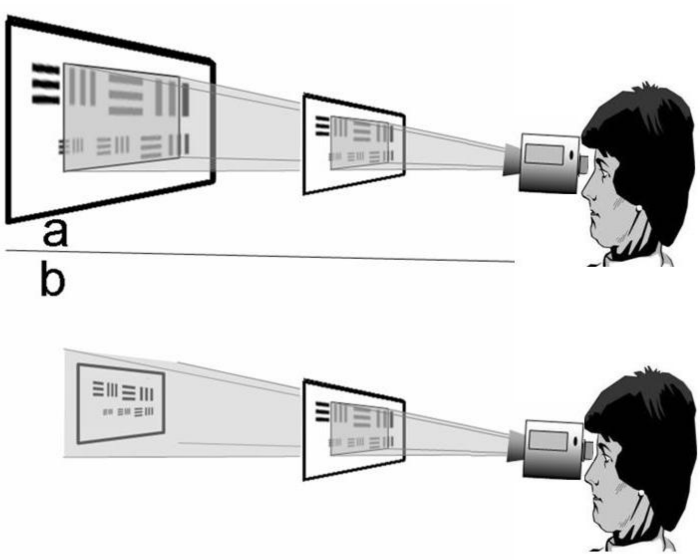


Figure 8.1 (a) The viewed object grows as range increases. This is assumed by the imager model. (b) Real objects have constant linear size.

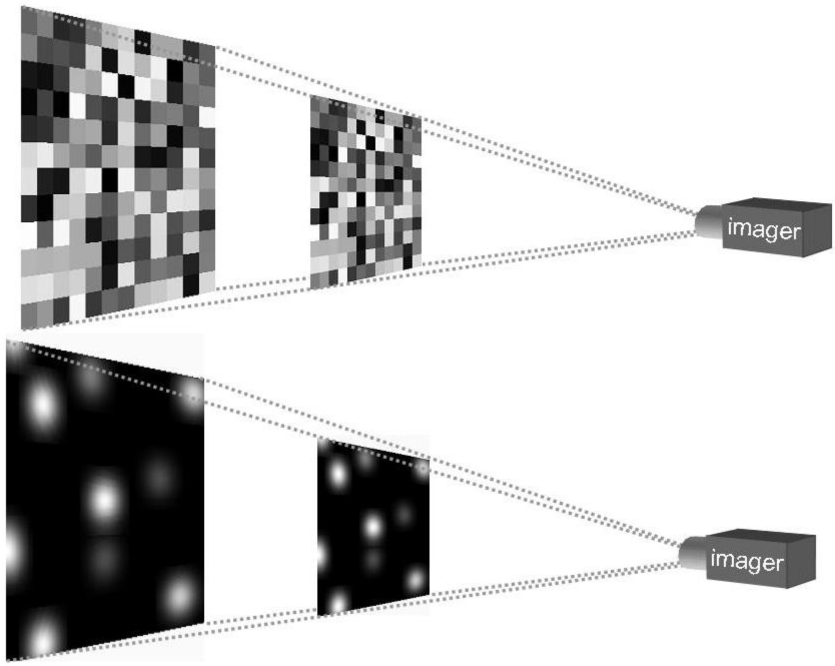


Figure 8.2 Both the blur at the bottom and the detector noise at the top grow in proportion to range. Imager properties are constant in angle space.

Both signal and detector noise are normalized to one milliradian horizontally and vertically. In Fig. 8.3, one milliradian intercepts a larger surface area as range increases. It should also be noted that model assumptions limit application to ranges with at least several detector instantaneous fields of view (IFOV) on the viewed object. The radiometry assumptions used in this book fail when target surface subtends less than an IFOV. The logic of a resolution model, however, is even more stringent. Resolution requires multiple pixels on target.

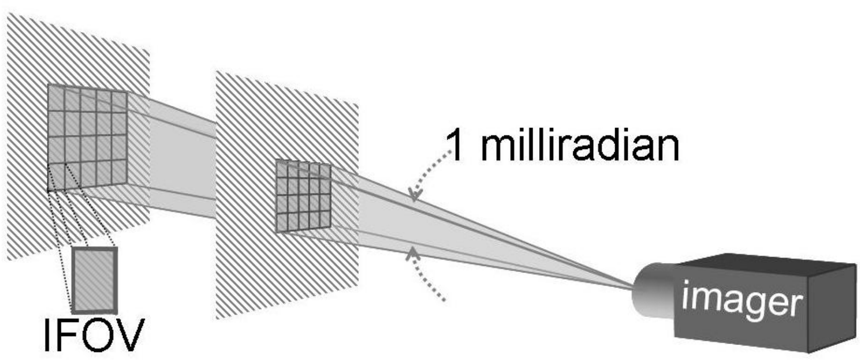


Figure 8.3 Signal and noise are normalized over one milliradian horizontally and vertically. Also, the model requires at least a few pixels (i.e., a few IFOV) subtended by the object surface.

8.2 Treatment of Temporal Noise In Detectivity Versus Photon-Counting Models

Both detectivity and photon-counting models are prevalent. The detectivity model is often used for thermal imagers. The photon-counting model is generally used for imagers of reflected light. Both models spatially normalize signal and noise over one milliradian horizontally and one milliradian vertically. However, the models treat temporal noise differently. This section discusses the different ways these two models treat temporal signals.

Each model is self-consistent in the treatment of signal and noise. A problem arises, however, when the imager noise must be compared to external noises. An example is the relationship between detector noise and eye noise; the eye is affected by the amplitude of the noise on the display and not by the origin of the noise. Nonetheless, the scaling factor relating eye noise to detector noise in the thermal model is different from that in the reflected light model. This is because signal and noise are calculated differently in the two models.

Detectivity models are based on representing signal and noise by flux (i.e., watts). Noise equivalent power (NEP) is the watts on the detector that generates an rms output equal to the detector noise. By convention, rms is the standard deviation of the noise measured for one second. By definition, flux (power) implies a rate. On the other hand, photon-counting models *integrate* (i.e., they count) photo-electrons for a frame time or an eye-integration time. A parallel can be drawn between integrating photo-electrons and joules. A parallel cannot be drawn between integrating photo-electrons and watts without properly accounting for time. An integrated flux model treats noise differently from the way a flux model treats noise.

Consider an example. The detector in Fig. 8.4 integrates 5-μm light for t_{eye} seconds. The value of t_{eye} is 0.039 sec. QE is quantum efficiency. Let's assume that one million photo-electrons are collected. There are one thousand noise electrons because noise is the square root of photo-electrons. Signal to noise SN_{photon} at time t_{eye} is given by Eq. (8.2):

$$SN_{photon} = \frac{1,000,000}{\sqrt{1,000,000}} = 1000 : 1. \tag{8.2}$$

Now consider the same example but calculate signal to noise using power (flux). Energy per photon for 5-μm light is 4×10^{-20} J. The signal S_{watt} (watts) needed to generate $1,000,000/t_{eye}$ electrons per second is obtained by dividing the number of electrons per second by QE to get photon flux, and then multiplying the photon flux by the energy per photon.

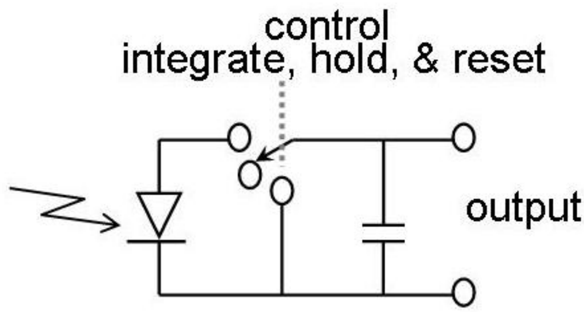


Figure 8.4 Detector signal is integrated for t_{eye} seconds. Signal is the average number of electrons on the capacitor. Noise is the rms variation between samples.

$$S_{watt} = \frac{(4\times 10^{-20})(1{,}000{,}000)}{QE\, t_{eye}}. \tag{8.3}$$

To calculate noise, we take an approach that parallels the detectivity model. This approach parallels the detectivity model because detectivity equals 1/*NEP*. We calculate NEP representing the rms noise due to signal S_{watt}. By convention and by the definition of NEP, the rms is taken over one second, as in Eq. (8.4):

$$NEP = \frac{4\times 10^{-20}}{QE}\sqrt{\left(\frac{1{,}000{,}000}{t_{eye}}\right)}\,\text{electrons}. \tag{8.4}$$

The 1,000,000/t_{eye} is signal electrons per second. The square root gives the number of noise electrons per second. The number of electrons divided by QE is photon flux (photons per second). Multiplying the joules per photon by the number of photons per second gives noise power in watts. Since NEP represents noise rms over a second, an adjustment is made to calculate signal to noise at t_{eye} seconds. Noise equivalent power is multiplied by $\sqrt{1/t_{eye}}$. That is, NEP is noise power in watts per root-Hertz. NEP is multiplied by the square root of the (single-sided) frequency bandwidth to get the rms noise associated with an integration of t_{eye} seconds:

$$SN_{det} = \frac{S_{watt}}{NEP\sqrt{\dfrac{1}{t_{eye}}}} = 1{,}000:1. \tag{8.5}$$

In the reflectivity model, E_{av} is the integrated photo-electron signal. The signal at time t_{eye} is $t_{eye}\,E_{av}$. The one-second rms noise is $\sqrt{E_{av}}$. The noise at time t_{eye} is $\sqrt{t_{eye}E_{av}}$. The photon-counting model calculates noise n_{photon} as shown in Eq. (8.6):

$$n_{photon} = \sqrt{t_{eye}\,E_{av}}\,. \tag{8.6}$$

The detectivity model calculates noise n_{flux} using Eq. (8.7):

$$n_{flux} = NEP\sqrt{\frac{1}{t_{eye}}}. \tag{8.7}$$

Notice that n_{flux} grows as t_{eye} decreases. NEP is rms noise averaged for one second. NEP/$\sqrt{t_{eye}}$ is the one-second rms of noise through a filter with bandwidth $1/t_{eye}$. We want the noise rms through the $1/t_{eye}$ bandwidth filter after t_{eye} seconds. The noise after t_{eye} seconds through a $1/t_{eye}$ bandwidth filter is simply NEP.

Both types of model set luminance proportional to flux; display luminance is proportional to E_{av} in the photon-counting model and proportional to S_{watt} in the flux model. Referring back to the discussion of eye noise in Section 7.3, we are looking for a proportionality constant α_{prop} to RSS eye noise and detector noise yielding total noise n_{total}. Equations (8.8) and (8.9) show the RSS of noise terms for the photon-counting and flux models, respectively. L is display luminance.

$$n_{total}^2 = \frac{n_{eye}^2 + \alpha_{prop}^2\, t_{eye}\, E_{av}}{L^2}. \tag{8.8}$$

$$n_{total}^2 = \frac{n_{eye}^2 + \alpha_{prop}^2\, NEP^2}{L^2}. \tag{8.9}$$

α_{prop} has the value 862. However, the $\sqrt{t_{eye}}$ factor [the t_{eye} factor in the numerator of Eq. (8.8)] is incorporated into the proportionality constant in the photon-counting model. α in the reflective models is 169.6, and t_{eye} is not explicitly included in the equations. These conventions are used in this book in order to be consistent with the literature. However, using different proportionality constants does make it appear as though the eye treats noise differently in the different models. That is not the case.

8.3 Relating Target and Imager Coordinate Systems

Figure 8.5 illustrates the relationship between an imager field of view and a target at three different ranges. As range increases, the target subtends a smaller portion of the imager field of view. The squares show imager samples fixed in angle space. The linear spacing between samples grows as range increases. However, target characteristics are fixed in linear space. As range increases, the sample spacing on target increases, and the target becomes poorly sampled.

Figure 8.6 shows a clock imaged with a CCD camera. Every other pixel is discarded, so the clock pictures are poorly sampled. At close range, the numerals are easy to read, and the poor sampling is not apparent. As range increases, however, the numerals become highly distorted, and the poor sampling is apparent. The camera does not change. The clock does not change. The sample spacing increases, making the effect of aliasing on the ability to read the numerals range dependent.

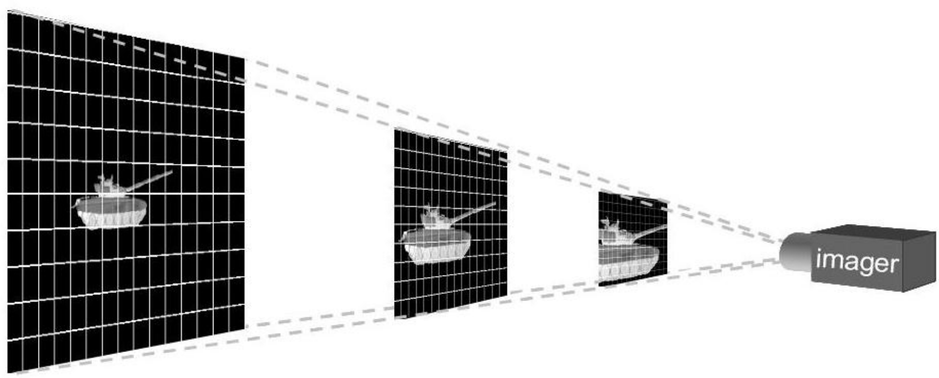


Figure 8.5 Field of view is fixed in angle space. The squares show imager samples. As range increases, target details shrink in angle space. The target becomes poorly sampled.

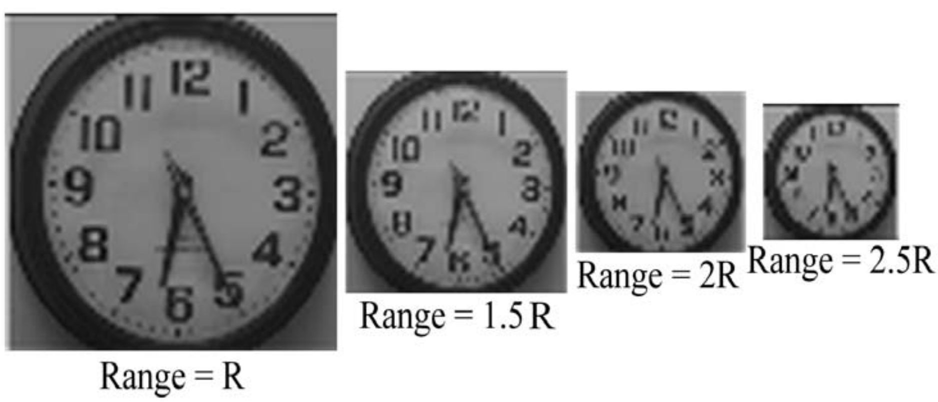


Figure 8.6 Pictures of a clock face taken with a CCD camera. Undersampling is exaggerated by discarding every other pixel in all of the pictures. The numerals become hard to read at longer range because sample spacing on the clock face becomes wider at increased range.

In an optical system, one expects blur to limit range performance. The imager has a certain angular resolution. As range increases, the target becomes angularly smaller and therefore more blurry. In the absence of significant noise, we expect blur to dominate range performance. But this is not always the case. In many practical circumstances, aliasing dominates range performance. The clock is extremely undersampled to illustrate a point. No commercial CCD camera is that poorly sampled. Figure 8.7 presents a more realistic example.

Consider a midwave thermal imager with diffraction wavelength of 4 μm. The FPA has 640 × 480 detectors, each 20 μm square. Fill factor is near one. The *f*/4 camera has a 7.3-cm aperture diameter with a 29.3-cm focal length. The field of view is 1.9 × 2.5 deg. The NETD is 0.04 K. However, the target is hot, and noise is unimportant in this example.

Figure 8.7(a) shows a target at 1.1 km. The target is highly resolved, and small details are highly visible at this close range. Figure 8.7(c) shows the same target at 4.3 km. The picture has been enlarged to the same size as the picture in Fig. 8.7(a). At 4.3 km, the target appears blurry, and details are hard to discern.

Figures 8.7(b) and (d) are pictures using the same imager. However, in these pictures, both vertical interpolation and horizontal dither are used. That is, each image is effectively 1280 × 960 pixels. A picture is taken. The camera is rotated half a pixel to the right horizontally and another picture taken. The camera is moved half a pixel down vertically and a third picture is taken. Finally, the camera is moved half a pixel to the left horizontally and a fourth picture is taken. The four pictures are aligned on a 1280 × 960-pixel grid. The vertical interlace and horizontal dither create images with exactly the same MTF as in Figs. 8.7(a) and (c). The pictures in Fig. 8.7(b) and (d) have better sampling.

The close-range picture in Fig. 8.7(b) appears to be the same as the picture in 8.7(a). The target is close and well sampled by the imager even without interlace and dither. However, comparing the pictures in 8.7(c) and (d) illustrates that sampling has a significant effect at the longer range. In Fig. 8.7(c), the image appears to be blurry. Figure 8.7(d) shows that most of the image corruption is actually caused by aliasing. Certainly the longer-range target is more blurry than the close-range target, but target identification is more limited by poor sampling at 4.3 km than by blur.

Normally, a 640 × 480 imager with *f*/4 optics is not considered "undersampled." Sampling artifacts are not visible when viewing typical scenes. However, the 4-μm diffraction cutoff of 18 cycles per milliradian is well beyond the half-sample frequency of 7.3 cycles per milliradian. At distances where expert observers identify targets, image corruption due to aliasing becomes apparent and limits performance.

The examples in Figs. 8.6 and 8.7 illustrate that the impact of aliasing is range dependent. None of the equations in Chapter 3 have range dependence. This is because all of the SIR analyses are done in angle space. The scene stimulus grows in size in proportion to range. The current discussion describes real targets of constant linear size.

Figure 8.7 In (a) and (b), the target is at 1.1 km. In (c) and (d), the target is at 4.3 km; (c) and (d) are enlarged for comparison. The MTF is the same in all images. However, in (b) and (d), the sampling is better than in (a) and (c). Notice that image quality in (d) is substantially better than in (c). This improvement results solely from better sampling.

In order to focus on the geometry of coordinate transformations, monochromatic radiation is assumed. Dealing with the spectral nature of target signatures is discussed in Chapters 9 and 10. The examples use thermal imagers, emissive targets, and meter-kilogram-second (MKS) units. In this discussion, surfaces are assumed to be Lambertian. Target and scene radiances in watts per square meter (W m^{-2}) are differences relative to the minimum flux in the field of view. In the thermal spectral bands, everything radiates. Thermal images are actually differences in radiance due to small temperature and emissivity changes across the field of view. A discussion of the nature of thermal signatures is provided in Chapter 9. For the present discussion, radiances can be thought of as absolute.

Also, for the moment, atmospheric absorption and scattering are ignored. The atmosphere certainly has a strong and often dominant effect on target signatures. However, the current discussion focuses on geometrical factors.

The imaging concept is illustrated in Fig. 8.8. Range R_{ng} in kilometers is a large distance compared to focal length f_l in centimeters. Many of the formulas are approximate, but the approximations are very good, provided that the FPA dimensions are small compared to the imaged scene.

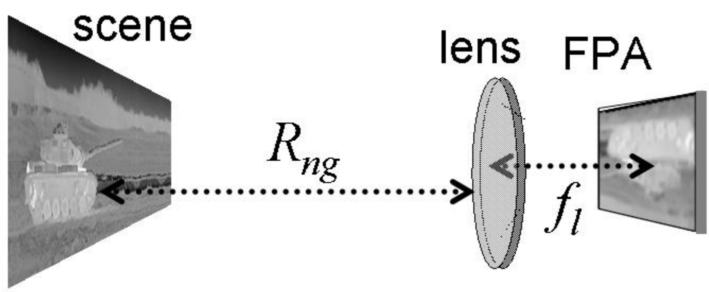


Figure 8.8 Lens images scene onto an FPA. The formulas in this chapter are accurate provided that R_{ng} is a large distance compared to f_l.

In imager coordinates, horizontal angle is expressed as x mrad, and spatial frequency is expressed as ξ mrad^{-1}. In target coordinates, length is expressed as x_m meters. When range is R_{ng} kilometers, x and x_m are related by Eq. (8.10). Figure 8.9 illustrates the relationship between angle and linear dimension. Using ξ_m as spatial frequency in cycles per meter, ξ and ξ_m are related by Eq. (8.11). One cycle per meter at one kilometer is equivalent to one cycle per milliradian. One cycle per meter at two kilometers becomes two cycles per milliradian. At two kilometers, one meter subtends half a milliradian. The angle gets smaller, and angular frequency increases.

$$x = \frac{x_m}{R_{ng}}. \tag{8.10}$$

$$\xi = \xi_m R_{ng}. \tag{8.11}$$

Figure 8.10 shows the image of a square target 1 m on edge. The target radiates 1 W m^{-2}. At 1 km, the target subtends 1 mrad. In imager coordinates, the dotted line shows the Fourier transform at 1-km range. The solid black and gray lines show what happens to the horizontal Fourier transform with the target at 2 and 4 km, respectively.

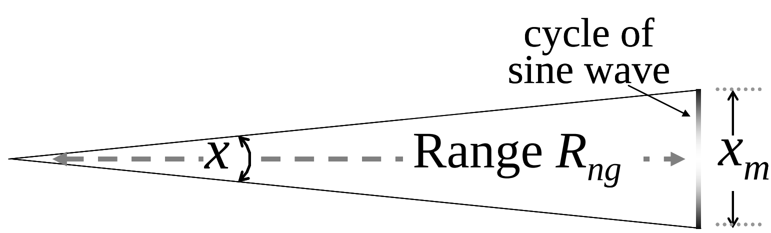


Figure 8.9 Relationship between linear dimension on target x_m and angular dimension at the imager x.

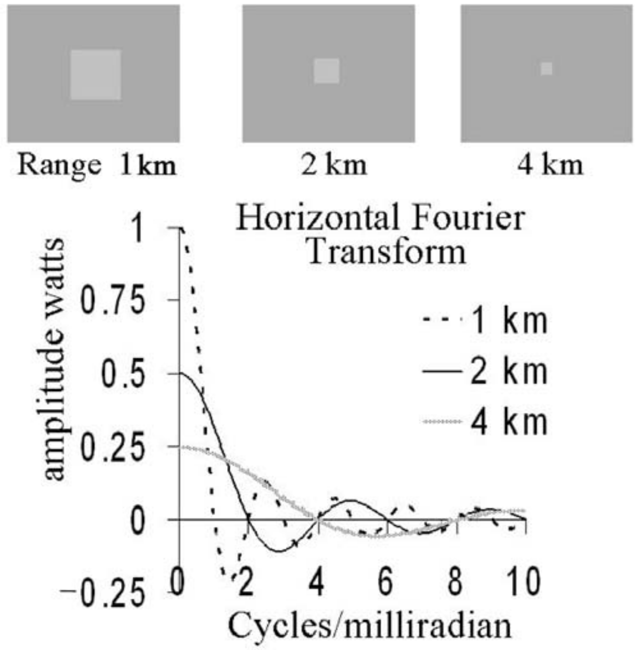


Figure 8.10 Square target at 1, 2, and 4 km. The graph shows horizontal Fourier transforms in imager coordinates. Low-frequency content decreases as range increases. Target signal spreads to higher frequencies.

Ignoring atmosphere, the peak intensity of the target does not change with range. In the space domain, the peak amplitude of the 4-km target is the same as that of the 1-km target. In Fig. 8.10, the 1-km target to the left and the 4-km target to the right are equally bright. Consider a small solid angle *dS* steradian (str). As R_{ng} increases, the energy reaching the imager aperture from each point on the target surface decreases as R_{ng}^2. However, R_{ng}^2 more of the target surface is subtended by *dS*. The two effects exactly compensate.

In the frequency domain, target spectrum spreads to higher spatial frequencies as range increases. Low-frequency signal decreases with range. Note that Fig. 8.10 shows the horizontal Fourier transform in imager coordinates. For a square, the same transform applies vertically. At 4 km, the amplitude actually drops by a factor of 16. The effect of range on target signature is found by multiplying the two-dimensional target frequency spectrum by imager MTF.

Let $S_{cn}(\xi_m,\eta_m)$ represent the Fourier transform of target radiance. The target has constant linear dimension; the target is some object. $G(\xi,\eta,R_{ng})$ is the angular radiance distribution at the imager aperture. $G(\xi,\eta,R_{ng})$ and $S_{cn}(\xi_m,\eta_m)$ are related by Eq. (8.12):

$$G(\xi,\eta,R_{ng}) = \frac{S_{cn}(\xi/R_{ng},\eta/R_{ng})}{R_{ng}^2}. \tag{8.12}$$

Ignoring sampling, the Fourier transform $F(\xi,\eta,R_{ng})$ of the image is given by Eq. (8.13). $P_{ix}(\xi,\eta)$ is the display MTF. $H(\xi,\eta)$ is the presample MTF including detector, optics, and vibration. Equation (8.13) also holds for sampled imagers if the pre- and postfilters are effective. That is, both H and P_{ix} must filter out all frequency content at and above the half-sample rate.

$$F(\xi,\eta,R_{ng}) = P_{ix}(\xi,\eta)\,H(\xi,\eta)\,\frac{S_{cn}(\xi/R_{ng},\eta/R_{ng})}{R_{ng}^2}. \tag{8.13}$$

In Eq. (8.13), $F(\xi,\eta,R_{ng})$ has units of watts per square cycles per milliradian (W mrad^2). P_{ix} and H are unitless. S_{cn} has units of watts per square cycles per meter (W m^2). The $R_{ng}{}^2$ in the denominator converts the W m^2 to W mrad^2.

If aliasing is present, then $F(\xi,\eta,R_{ng})$ is given by Eq. (8.14). Horizontal sample rate is ν, and vertical sample rate is γ. Equation (8.14) is the Fourier transform of the target radiant distribution in imager coordinates. Although the summations theoretically go to $\pm\infty$, the display and eyeball MTF limit the aliasing terms to n_{max} and m_{max}:

$$F(\xi,\eta,R_{ng}) = P_{ix}(\xi,\eta)\sum_{n=-n_{max}}^{n_{max}}\sum_{m=-m_{max}}^{m_{max}}\cdot \left\{H(\xi-n\nu,\eta-m\gamma)\,\frac{S_{cn}[(\xi-n\nu)/R_{ng},(\eta-m\gamma)/R_{ng}]}{R_{ng}^2}\right\}. \tag{8.14}$$

In Eq. (8.15), $F_T(\xi_m,\eta_m,R_{ng})$ is the Fourier transform of the displayed image transposed to the target plane:

$$F_T(\xi_m,\eta_m,R_{ng}) = P_{ix}(\xi_m R_{ng},\eta_m R_{ng})\sum_{n=-n_{max}}^{n_{max}}\sum_{m=-m_{max}}^{m_{max}}\cdot \left\{H(\xi_m R_{ng}-n\nu,\eta_m R_{ng}-m\gamma)\,\frac{S_{cn}\left[\xi_m - n\nu/R_{ng},\eta_m - m\gamma/R_{ng}\right]}{R_{ng}^2}\right\}. \tag{8.15}$$

These equations provide several insights:

1. In image space, the peak target signal remains constant as range increases. This is true until the target subtends less than a single-detector instantaneous field of view. In the Fourier domain, peak amplitude drops

as range increases. The target does not get dimmer with range; it gets angularly smaller.

2. Similarly, the aliasing terms in Eqs. (8.14) and (8.15) do not get dimmer as range increases. Aliasing intensity remains constant. The Fourier amplitude drops as R_{ng}^2 because aliasing only corrupts areas of the displayed image that contain target signal.
3. The presence or absence of aliasing is completely controlled by the imager. It is true that the sample spacing on target increases with range. However, the imager presample MTF provides more aggressive filtering of the target as range increases. The two effects are proportional. At long range, a more aggressive prefilter is needed to avoid aliasing. Since imager blur is fixed in angle space, it provides a more aggressive filter at longer range. If aliasing is present, the amplitude of the aliased signal is proportional to target contrast. However, the imager can remove all aliasing.

Equations (8.14) and (8.15) are useful for providing insights into the range-dependent properties of target signatures. Furthermore, if the goal is to predict whether a specific target is visible or identifiable at a particular range, the range-dependent signature is used in the analysis. The spatial cues that discriminate one target from another or from background are fixed in linear space. Predicting the visibility of those cues at range requires the use of Eq. (8.14) or some equivalent formulation.

8.4 Spatial Scaling of Aliasing Noise

Aliasing is added in quadrature (RSS) with detector noise. In order to RSS the aliased signal with detector noise, it must be properly scaled. The specific procedures for calculating alias signal and detector noise are provided in subsequent chapters. The problem addressed here is how to properly scale aliasing for combining with other imager noise terms.

A target of area L_{tgt} m^2 subtends a solid angle of L_{tgt}/R_{ng}^2. The solid angle subtended by the target is range dependent. Aliasing is generated over the target area. In order to properly scale aliasing noise to detector noise, the aliasing noise term is divided by L_{tgt}/R_{ng}^2. That is, aliasing is normalized to a square milliradian like the other noise and signal terms. The normalization is range dependent because the noise is caused by imager and target interaction.

Equation (8.16) gives the aliased noise spectrum $A(\xi,\eta)$. Randomness is assumed, so we can RSS any overlap of aliasing terms. The term with both m and n equal to zero is the imager transfer response. Transfer response is not included in aliasing noise. Generally, significant signal from aliasing terms with m or n indices above 2 or 3 is removed by the postsample MTF H_{post}.

$$A(\xi,\eta) = \left[\sum_{\substack{m=-\infty \\ m,n\neq 0}}^{\infty} \sum_{n=-\infty}^{\infty} H_{pre}^{2}(\xi - n\nu, \eta - m\gamma) \right]^{1/2} H_{post}(\xi,\eta), \tag{8.16}$$

where

ξ = horizontal spatial frequency in cycles per milliradian (mrad^{-1}),
ν = horizontal sampling frequency (mrad^{-1}),
η = vertical spatial frequency (mrad^{-1}),
γ = vertical sampling frequency (mrad^{-1}),
H_{pre} = presample imager MTF, and
H_{post} = postsample imager MTF.

Total imager noise Γ_{imager} is the quadrature sum of detector noise and aliasing. Detector noise SSD is Γ_{det} with units of $\sec^{1/2}$ W mrad^{-2}. Equation (8.17) gives total noise averaged for one second. The details on incorporating aliasing noise into the thermal and reflected light CTF models are provided in Chapters 9 and 10, respectively.

$$\Gamma_{total}(\xi,\eta) = \left[\Gamma_{det}^{2} + \frac{R_{ng}^{4}}{L_{tgt}^{2}} C_{tgt}^{2} A^{2}(\xi,\eta) \right]^{1/2}. \tag{8.17}$$

Aliasing is generated over the entire target area. When the target is either large or close, the total aliased signal is large. However, the total signal from the target is also large. The target gets angularly smaller with an increase in range. As the displayed target area decreases, the aliased signal also decreases. However, the Fourier domain signal from the target is both decreasing in amplitude and spreading in frequency. Aliasing has a greater effect at longer range.

C_{tgt} is independent of range. The model target grows in size, as shown in Fig. 8.1(a). Aliasing has a constant effect on a target defined in angle space. However, the linear size of real targets is constant. Furthermore, the discrepancy that arises from the difference between model and real objects cannot be addressed after finding imager resolution. The amplitude relationship between detector noise and aliasing must be established before determining angular resolution. The range-dependent nature of aliasing must be incorporated into the angle-space model. Otherwise, the proper relationship between aliasing and detector noise is not maintained.

The parameter L_{tgt} is a scaling factor that can generally be associated with target size. That is, L_{tgt}/R_{ng}^2 is the solid angle presented to the imager. However, L_{tgt} actually represents the scale of the spatial features used to discriminate the targets. For example, consider two billboards. One billboard is twice the size of the other. However, the words on both billboards are equal in size and font. The large billboard cannot be read at twice the distance of the small billboard. Target size in this case is letter size, not billboard size.

Bibliography

Vollmerhausen, R. H., R. G. Driggers, and D. L. Wilson, "Predicting range performance of sampled imagers by treating aliased signal as target-dependent noise," *J. Opt. Soc. Am. A* **25**, 2055–2065 (2008).

Vollmerhausen, R. H., R. G. Driggers, and D. L. Wilson, "Predicting range performance of sampled imagers by treating aliased signal as target-dependent noise: erratum," *J. Opt. Soc. Am. A* **26**, 2418–2418 (2009).

Vollmerhausen, R., "Display of Sampled Imagery," in *Electro-Optical Imaging: System Performance and Modeling*, L. Biberman, Ed., SPIE Press, Bellingham, WA (2000).

Chapter 9
Thermal Imager Topics

This chapter provides details on calculating signal and noise in thermal imagers. A detectivity model is used to calculate thermal imager contrast threshold function (CTF_{sys}). CTF_{sys} is used in the targeting task performance (TTP) metric to calculate imager resolution.

Thermal imagers sense heat energy with wavelengths between 3 and 12 μm. The 3- to 5-μm band is called midwave infrared (MWIR), and the 8- to 12-μm band is called longwave infrared (LWIR). Figure 9.1 shows typical atmospheric transmission for a 1-km horizontal path. There are three transmission windows from 3 to 4.2 μm, 4.4 to 5 μm, and 8 to 13 μm.

Everything near room temperature radiates in the infrared. The emissivity of natural objects is generally above 70%. Most manmade objects are also highly emissive. Thermal sensors derive their images from small variations in temperature and emissivity within the scene. Typically, the thermal scene is very low contrast.

Figure 9.2 shows the spectral radiant exitance from blackbodies at 300 K and 303 K. The difference between the two curves is also shown. The difference is small; however, a 3-K contrast represents good thermal imaging conditions.

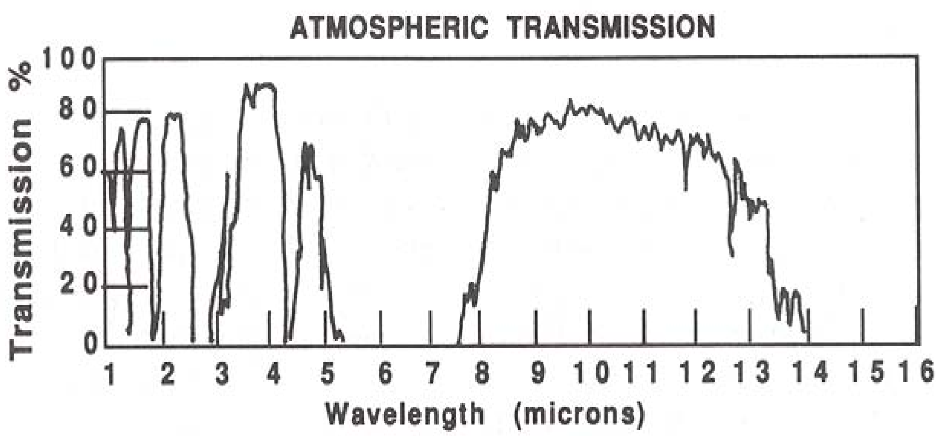


Figure 9.1 Atmospheric transmission over a 1-km path.

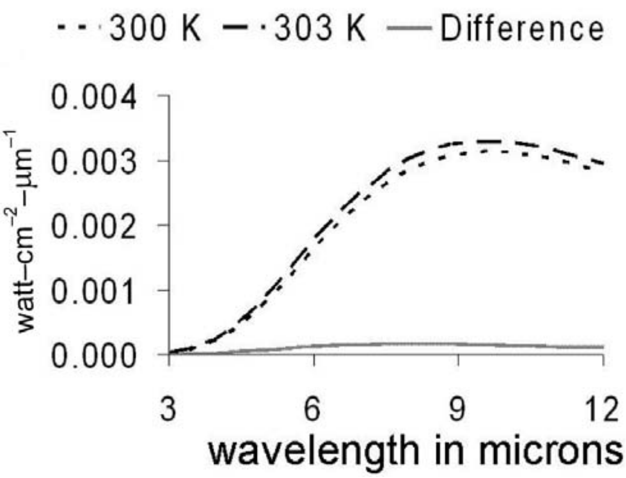


Figure 9.2 Thermal radiant exitance from 300- and 303-K blackbodies. Although the difference represents good thermal contrast, the relative difference is small.

Figure 9.3(b) shows an 8- to 10.5-μm LWIR image. Picture contrast is enhanced by displaying differences within the scene. The minimum pixel radiance is set to picture black level. The RSS target contrast is 0.004 watts per square centimeter (W cm^{-2}). Section 6.3 describes RSS contrast measurement. The in-band target contrast is the same as the difference plot in Fig. 9.2.

Figure 9.3(a) shows scene radiance displayed with an absolute intensity scale. Each pixel intensity represents scene radiance at that point. The scene is very low contrast. Figure 9.4 shows a plot of intensities along the dashed line in Fig. 9.3(a). The dashed line crosses the hottest part of the vehicle. Even this relatively hot surface has only a 5% contrast.

Although the typical thermal scene is low contrast, exceptions do exist. For example, the radiance difference between sky and ground can be quite large on a clear day. Also, the classic “burning tank” can overload a thermal imager. In general, however, thermal sensors are designed to map small differences in the scene’s radiant energy into a usable image.

Scene radiance depends on a number of factors. The spectral radiance of an object depends on its surface temperature and emissivity. Radiant emittance also depends on the nature of the light being reflected or scattered from the surface. The apparent spectral radiance of an object as seen by an imager is also affected by the spectral transmission of the atmosphere. Many factors affect the effective thermal contrast of the scene.

Thermal cameras are designed to handle the large signal pedestal associated with emissive imagery. These contrast-enhancement features must be considered when evaluating thermal imager resolution using CTF_{sys} and the TTP metric.

Section 9.1 defines effective blackbody temperature (T_{ebb}), a radiometric unit used in target acquisition models. The relationship between T_{ebb}, surface temperature T_{surf}, and pyrometer temperature T_{pyro} is explained.

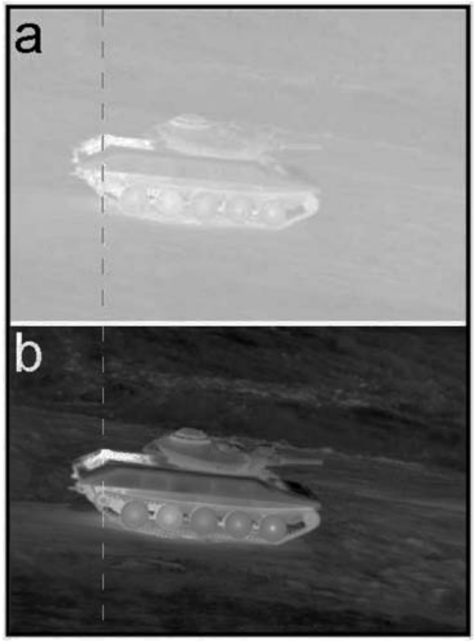


Figure 9.3 Picture of a tank. (a) Radiance is plotted on an absolute scale. (b) Only difference radiance is displayed. This enhances scene contrast.

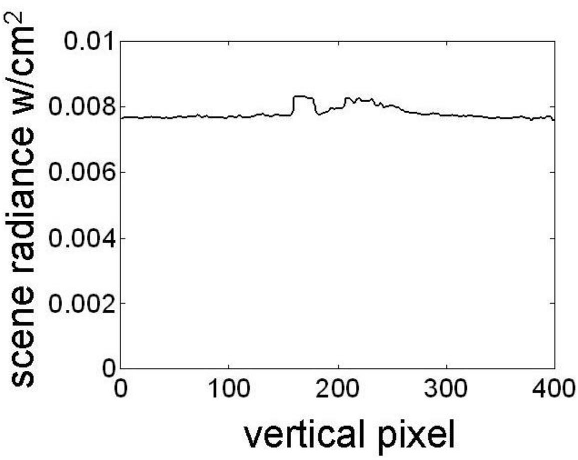


Figure 9.4 Plot of intensity along the dashed line shown in Fig. 9.3(a). The plot shows that even hot parts of the target are low contrast.

The impact of imager noise on human vision is described in Chapter 7. The impact is the same for any kind of imager. Only the amplitude of the noise counts, not its origin. However, there are two types of hardware models; these models treat signal and noise in fundamentally different ways. Section 9.2 describes signal and noise in detectivity models. Section 9.3 derives CTF_{sys} for thermal imagers.

The effect of aliasing on CTF_{sys} is described in Section 9.4. Predicting the average probability of identifying members of a target set is discussed in Section 9.5. Section 9.6 discusses how to model local-area contrast enhancement (LACE). LACE is needed because of large radiance swings in the scene. However, the resolution model does not consider target or background spatial features. The technique for incorporating the benefit of LACE into the resolution model is outlined. Section 9.7 provides a minimum resolvable temperature model.

9.1 Effective Blackbody Temperature

Hot objects radiate energy. Temperature is an intuitive measure of target thermal contrast. This section defines effective blackbody temperature (T_{ebb}). T_{ebb} is a radiometric unit that is proportional to the radiant exitance in the spectral band of the imager. The relationship between T_{ebb} and radiant exitance is not the same as the relationship between T_{pyro} and radiant exitance. Both T_{pyro} and T_{ebb} are discussed in order to clarify the differences.

Imagers respond to radiance. It is common practice, however, to associate surface radiance with the temperature of a blackbody generating the same radiance in the spectral band of the imager. Suppose that the radiance of a surface is measured. The radiometer operates in the 8–12-μm spectral band. The radiance is one watt per square centimeter per steradian (W cm^{-2} str^{-1}). Suppose also that a blackbody at temperature T_1 radiates 1 W cm^{-2} str^{-1} in the spectral band. Then the measured surface has a "temperature" T_1.

Pyrometers estimate absolute surface temperature T_{surf} by sensing radiance. Pyrometer temperature T_{pyro} is the temperature of a blackbody generating the same radiance in some spectral band. Pyrometers are used to control the temperature in blast furnaces, to sense the temperature of superheated steam in a boiler, and to establish whether radiant heating recently installed in a floor is working properly. The absolute temperature of the surface is the objective of the measurement.

Thermal imagers and pyrometers have different uses. Target acquisition generally makes use of small differences in scene radiance. The example in Fig. 9.3 illustrates a "hot" target for a thermal imager. Thermal imagers generally operate over a very limited temperature span compared to pyrometers. Furthermore, the relationship between T_{pyro} and radiance is nonlinear. T_{pyro} is not used in target acquisition modeling.

A third "temperature" is defined for use in modeling thermal imagers. T_{ebb} Kelvin (K) is a radiance unit. Also, T_{ebb} is associated with a specific target spectrum. T_{pyro} is discussed before T_{ebb} in order to highlight the differences.

Blackbody spectral radiant exitance W_b is a function of wavelength λ and temperature T. Wavelength is in microns, and temperature is in K:

$$W_b(\lambda, T) = \frac{C_1}{\lambda^5 \left[e^{C_2/\lambda T} - 1 \right]}. \tag{9.1}$$

W_b units are W cm^{-2} μm^{-1}. C_1 is 11934, and C_2 is 14388. For a spectral band λ_1 to λ_2, radiant exitance $W(T)$ is given by Eq. (9.2):

$$W(T) = \int_{\lambda_1}^{\lambda_2} \frac{C_1}{\lambda^5 \left[e^{C_2/\lambda T} - 1 \right]} d\lambda. \tag{9.2}$$

Figures 9.5 and 9.6 plot T versus $W(T)$ for the 8–12-μm band and the 3–5-μm band, respectively. Figures 9.7 and 9.8 plot the radiant exitance corresponding to a 1-K temperature difference. The –40–60° Celsius (C) scale represents terrestrial environments.

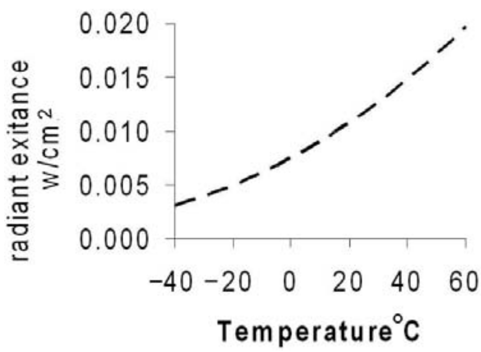


Figure 9.5 Radiant exitance of a blackbody versus temperature for the 8–12-μm spectral band.

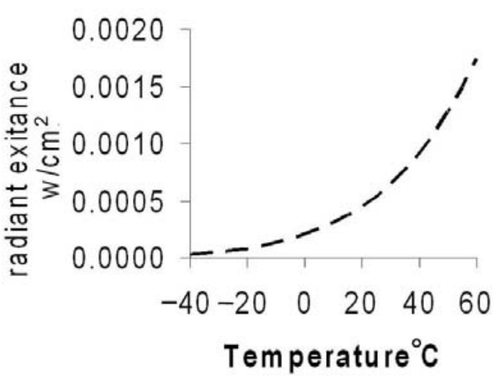


Figure 9.6 Radiant exitance of a blackbody versus temperature for the 3–5-μm spectral band.

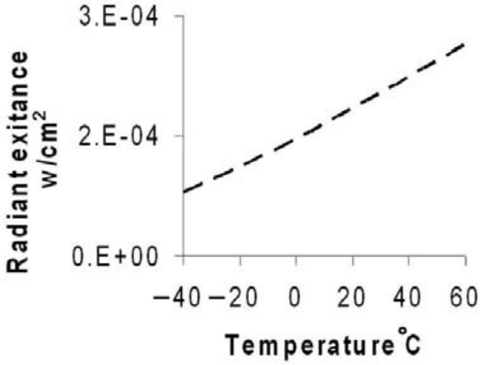


Figure 9.7 Radiant exitance in the 8–12-μm spectral band for a 1-K temperature difference centered on the surface temperature.

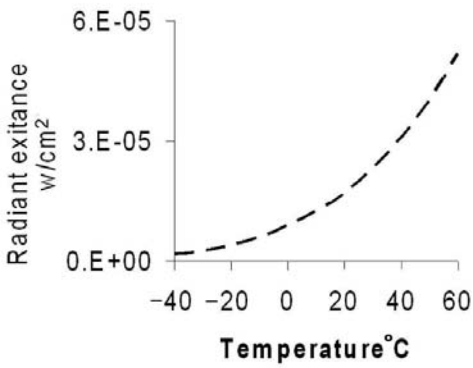


Figure 9.8 Radiant exitance in the 3–5-μm spectral band for a 1-K temperature difference centered on the surface temperature.

The temperature T in Figs. 9.5 through 9.8 represents ideal T_{pyro}. The relationship between T_{pyro} and radiant exitance is not linear. Quantifying target contrast as a delta T_{pyro} requires specifying the center temperature. If targets and backgrounds are blackbodies, and if the absolute temperatures of target and background are the only significant factors in establishing target radiant contrast, then managing the nonlinearity might be worthwhile. Neither eventuality is true. The imager responds to radiant flux, and many factors affect radiant contrast.

Figure 9.9 illustrates the T_{ebb} concept. The radiometer operates in some thermal spectral band. In the laboratory, the voltmeter is zeroed on the 300-K blackbody. The voltage V_K associated with a 1-K change in blackbody temperature is recorded. Radiometer response is linear. Twice as much radiation in the spectral band results in a $2V_K$ voltage reading. A surface radiating less than a 300-K blackbody results in a negative reading. The radiometer is taken to the field to measure the "temperature" of various scene elements. The voltmeter reading divided by V_K measures the effective blackbody temperature (T_{ebb}) of each object. T_{ebb} has units K.

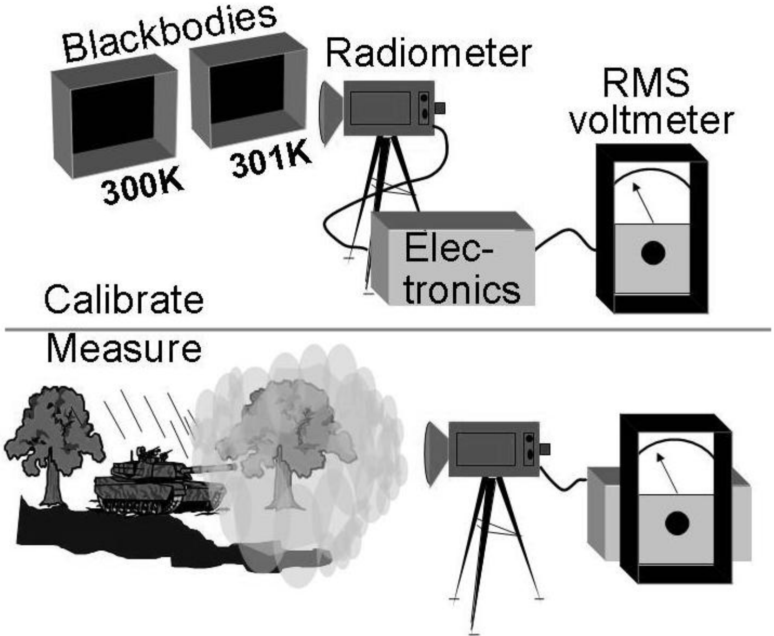


Figure 9.9 Illustration of the concept of T_{ebb}.

The procedure just described is only approximate. If W_{tgt} is in-band target radiant exitance as sensed by the imager, then Eqs. (9.3) through (9.5) define T_{ebb}. The target-to-background spectrum is represented by the rate of change of Planck's equation with respect to temperature evaluated at 300 K. $S(\lambda)$ is the spectral response of the imager normalized to peak.

$$W_{ebb} = \int_{\lambda_1}^{\lambda_2} \left.\frac{\partial W_b(\lambda, T)}{\partial T}\right|_{T=300} S(\lambda)\, d\lambda. \tag{9.3}$$

$$W_{ebb} = \int_{\lambda_1}^{\lambda_2} \frac{S(\lambda)\, C_1 C_2 e^{C_2/300\lambda}}{300^2\, \lambda^6 \left[e^{C_2/300\lambda} - 1 \right]^2} d\lambda. \tag{9.4}$$

$$T_{ebb} = \frac{W_{tgt}}{W_{ebb}}. \tag{9.5}$$

T_{ebb} has several convenient properties. Unlike T_{pyro}, T_{ebb} is proportional to power on the detector. Furthermore, T_{ebb} inherently represents a rate of change for terrestrial temperature variations. The spectrum associated with T_{ebb} is shown in Fig. 9.10. Since the actual target-to-background spectrum is seldom known, this spectrum is used in many thermal models. The spectrum shown represents a 1-K target contrast. For different target contrasts, the spectrum amplitude simply scales.

9.2 Signal and Noise in the Detectivity Model

Noise in a thermal detector can be found from first principles. Noise depends on background flux, quantum efficiency, dark current, and readout circuit amplifier noise. Calculating noise from first principles requires a large amount of information about the detector and its environment. Generally, noise calculations or measurements are done by the manufacturer.

The manufacturer measures the noise with the detector packaged in a dewar and viewing a 300-K thermal background. The measured noise pertains only to the cold shield and background used during the measurement. A significant change in detector flux environment requires either a new measurement or a revised noise calculation. However, providing one noise amplitude simplifies communication with the detector array user.

Spectral detectivity (D_λ) is used to specify the noise in a thermal detector:

$$D_\lambda = \frac{1}{NEP_\lambda}. \tag{9.6}$$

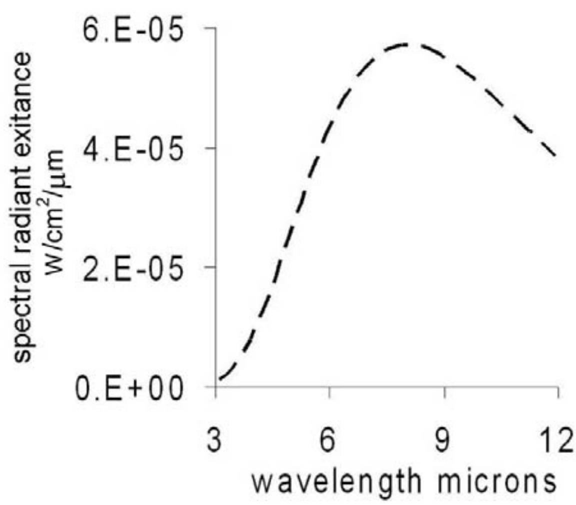


Figure 9.10 Spectral radiant exitance associated with T_{ebb}. This is a 1-K target-to-background spectrum.

NEP_λ is the spectral noise-equivalent power. It is the monochromatic signal power necessary to produce an rms signal to noise of unity. Spectral *D*-star (D_λ*) is a normalization of D_λ to unit area and bandwidth.

$$D_\lambda^* = D_\lambda \left(A_{det} \Delta f \right)^{1/2} . \tag{9.7}$$

In Eq. (9.7), Δf is the temporal bandwidth, and A_{det} is the active area of a single detector. In the thermal model, detector performance is characterized by peak spectral *D*-star $D_{\lambda peak}$ and relative detector response $S(\lambda)$. $S(\lambda)$ is the detectivity at each wavelength divided by $D_{\lambda peak}$.

The spectral radiant power E_{fpa} on the FPA is calculated as follows:

$$E_{fpa} = \frac{\pi \tau L_{scene}}{4 F\#^2}, \tag{9.8}$$

where L_{scene} is scene radiance in W cm^{-2}str$^{-1}\mu$m^{-1}, τ is transmission of the optics and atmosphere, and *F#* is the *f*-number of the optics. The parameters τ, L_{scene}, and E_{fpa} are all functions of wavelength λ. The signal to noise in one pixel (SN_{pix}) and for an eye integration time can now be calculated.

$$SN_{pix} = \frac{D^*_{\lambda\ peak}}{\left(\frac{A_{det}}{t_{eye}} \right)^{1/2}} \int_{\lambda_1}^{\lambda_2} A_{det}\ E_{fpa}(\lambda)\ S(\lambda)\ d\lambda. \tag{9.9}$$

$$SN_{pix} = \left(\frac{D^*_{\lambda\ peak} \left(t_{eye} A_{det} \right)^{1/2} \pi \tau}{4F\#^2} \right) \left(\int_{\Delta\lambda} L_{scene}(\lambda)\ S(\lambda)\ d\lambda \right). \tag{9.10}$$

Per the discussion in Section 9.1, L_{scene} is represented by some multiple Γ of the rate of change of Planck's blackbody equation with respect to temperature evaluated at 300 K. Scene radiance is integrated over the spectral band of the imager:

$$L_{scene} = \Gamma \left. \frac{\partial W_b(\lambda, T)}{\partial T} \right|_{T=300} . \tag{9.11}$$

$$\delta_{scene} = \int_{\lambda_1}^{\lambda_2} \frac{\partial W_b(\lambda, T)}{\partial T} \Big|_{T=300} S(\lambda) d\lambda.$$

$$SN_{pix} = \frac{\Gamma \, \delta_{scene} \, D^*_{\lambda peak} \left(t_{eye} A_{det}\right)^{1/2} \pi \, \tau}{4 \, F\#^2}. \tag{9.12}$$

We want to normalize signal and noise over a milliradian horizontally and vertically. In one square radian, the signal to noise would increase by an amount $[F_o^2/(H_{pit}V_{pit})]^{1/2}$ where F_o is the effective focal length of the afocal or objective lens. H_{pit} and V_{pit} are detector horizontal and vertical pitch, respectively. The signal-to-detector noise in one square milliradian is

$$SN_{det} = (1 \times 10^{-3}) \, \Gamma \, \delta_{scene} \, F_o \sqrt{\eta_{stare} \, t_{eye}} \, D^*_{\lambda \; peak} \, \pi\tau \Big/ 4 \, F\#^2. \tag{9.13}$$

η_{stare} is an efficiency factor and includes the detector area fill factor. Also, due to limitations in photo-electron storage capacity, the FPA might integrate signal for t_{int} seconds rather than a full field time. The efficiency factor is adjusted for detector integration time. T_{CCD} is the number of fields or frames per second.

$$\eta_{stare} = \frac{t_{int} \, T_{CCD} \, H_{det} \, V_{det}}{H_{pit} \, V_{pit}}. \tag{9.14}$$

Equation 9.13 gives the signal to noise for temperature difference Γ. Noise modulation at the display is needed to find CTF_{sys}. Setting signal to noise to unity, Γ_{det} is noise standard deviation in units of K mrad sec$^{1/2}$.

$$\Gamma_{det} = \frac{4 \, F\#^2}{(1 \times 10^{-3}) \; \delta_{scene} \, F_o \cdot \sqrt{\eta_{stare} \, t_{eye}} \, D^*_{\lambda \; peak} \; \pi \, \tau}, \tag{9.15}$$

where

$F\#$ = f-number of the optics,

Δf = temporal bandwidth in Hertz,
A_{det} = active area of a single detector = $H_{det}\, V_{det}$,
$D^*\lambda_{peak} = D\lambda^*$ at wavelength of peak response,
$S(\lambda)$ = response of detector at wavelength λ relative to peak response,
E_{fpa} = W cm^{-2} μm^{-1} on the detector array,
L_{scene} = W cm^{-2}str^{-1}μm^{-1} from the thermal scene,
τ = transmission of optics,
E_{det} = watts on a single detector,
$W_b(\lambda,T)$ = Planck's equation for blackbody radiation,
T = temperature,
Γ = amplitude of apparent blackbody temperature difference,
SN_{pix} = pixel signal-to-noise ratio,
T_{CCD} = field rate (probably 60 Hz), and
t_{int} = detector integration time < = $1/T_{\mathrm{CCD}}$.

9.3 Thermal Imager Contrast Threshold Function

Calculating $\mathrm{CTF_{sys}}$ requires that detector noise be expressed as display luminance noise. This, in turn, requires a mapping between radiometric temperature changes in the scene and the matching luminance changes on the display. The gain through the imager must be established in terms of foot-Lamberts per Kelvin.

Scene contrast temperature (S_{tmp}) is the delta effective blackbody temperature in the scene needed to raise display luminance from black to average. Recall that the thermal image arises from small variations in temperature and emissivity within the scene. These small variations are superimposed on a large background flux. Zero luminance on the display corresponds to the minimum scene radiant energy, not to zero radiant energy. S_{tmp} is not the absolute background radiometric temperature. It is the temperature contrast needed to raise the display luminance from zero to average.

With display noise modulation established, $CTFH_{sys}$ and $CTFV_{sys}$ are calculated using Eqs. (9.16) and (9.17), respectively. CTF_{sys} is the geometric mean of $CTFH_{sys}$ and $CTFV_{sys}$.

$$CTFH_{sys}(\xi) = \frac{CTF\left(\xi / SMAG\right)}{M_{dsp} H_{sys}(\xi)} \left(1 + \frac{\alpha_{thermal}^2\, \Gamma_{det}^2\, QH_{hor}(\xi)\, QV_{hor}}{S_{tmp}^2} \right)^{1/2}, \tag{9.16}$$

$$CTFV_{sys}(\eta) = \frac{CTF\left(\eta / SMAG\right)}{M_{dsp}V_{sys}(\eta)}\left(1 + \frac{\alpha^2_{thermal}\ \Gamma^2_{det}\ QH_{ver}\ QV_{ver}(\eta)}{S^2_{tmp}}\right)^{1/2}. \quad (9.17)$$

$$QH_{hor}(\xi) = \int \left| B(\xi' / \xi) H_{elec}(\xi') H_{dsp}(\xi') H_{eye}\left(\xi' / SMAG\right)\right|^2 d\xi', \quad (9.18)$$

$$QV_{hor} = \int \left| V_{elec}(\eta) V_{dsp}(\eta) H_{eye}\left(\eta / SMAG\right)\right|^2 d\eta, \quad (9.19)$$

$$QH_{ver} = \int \left| H_{elec}(\xi) H_{dsp}(\xi) H_{eye}\left(\xi / SMAG\right)\right|^2 d\xi, \quad (9.20)$$

$$QV_{ver}(\eta) = \int \left| B(\eta' / \eta) V_{elec}(\eta') V_{dsp}(\eta') H_{eye}\left(\eta' / SMAG\right)\right|^2 d\eta', \quad (9.21)$$

where

$\alpha_{thermal}$ = 862 root-Hz (proportionality constant for detectivity models),
ξ = horizontal spatial frequency in $(\text{mrad})^{-1}$,
η = vertical spatial frequency in $(\text{mrad})^{-1}$,
SMAG = system magnification,
$CTF(\xi/SMAG)$ = naked-eye contrast threshold function,
S_{tmp} = scene temperature that raises display luminance from black to average,
$B(\xi \text{ or } \eta)$ = visual cortex eye filters from the appendix,
$H_{eye}(\xi \text{ or } \eta)$ = eyeball MTF; see the appendix,
$H_{elec}(\xi)$ = horizontal electronics MTF,
$V_{elec}(\eta)$ = vertical electronics MTF,
$H_{dsp}(\xi)$ = horizontal display MTF,
$V_{dsp}(\eta)$ = vertical display MTF,
$H_{sys}(\xi)$ = horizontal system MTF,
$V_{sys}(\eta)$ = vertical system MTF,
QH_{hor} = horizontal noise bandwidth for $CTFH_{sys}$,
QV_{hor} = vertical noise bandwidth for $CTFH_{sys}$,
QH_{ver} = horizontal noise bandwidth for $CTFV_{sys}$, and
QV_{ver} = vertical noise bandwidth for $CTFV_{sys}$.

The length of CTF bar patterns is always in excess of one degree. See Fig. 9.11. Although blurring the length of the pattern does degrade CTF to some extent, the loss of modulation is small. Blurring of the bar length is therefore ignored.

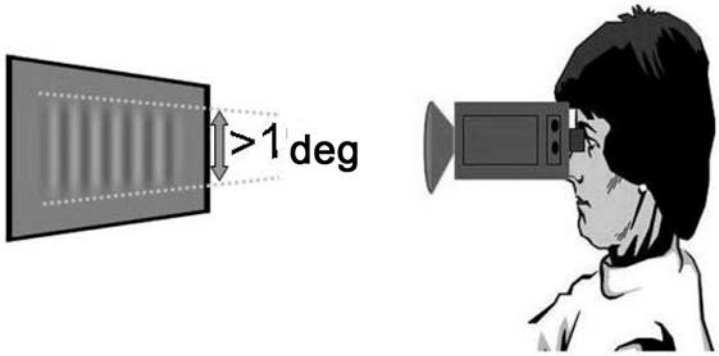


Figure 9.11 Long bars are used during CTF testing. Degradation of CTF due to blurring of the length of the bars is ignored.

9.4 Adding Aliasing Noise

To model the effect of aliasing on CTF_{sys}, we calculate the RSS of the aliased signal and the detector noise. Both detector noise and aliased signal are expressed as modulation on the average display luminance. Aliasing increases noise and therefore degrades (elevates) CTF_{sys}. Using the degraded CTF_{sys} results in poorer resolution of the target and reduced PID.

The aliased signal is different from detector noise in two ways. First, aliasing disappears as the target contrast disappears. Second, the image corruption due to aliasing gets worse with increased range. This is because sampling is constant in angle space, and targets become poorly sampled as range increases.

Equations (9.22) and (9.23) give the aliasing spectrums for horizontal and vertical signals, respectively. These two equations also involve infinite sums. However, the display and eye filter out high spatial frequencies. Only a few of the terms are numerically significant.

$$A_H(\xi) = \left[\sum_{n=-\infty}^{\infty} H_{pre}^2(\xi - n\nu) \right]^{1/2} H_{elec}(\xi)\, H_{dsp}(\xi), \tag{9.22}$$

$$A_V(\eta) = \left[\sum_{n=-\infty}^{\infty} V_{pre}^2(\eta - n\gamma) \right]^{1/2} V_{elec}(\eta)\, V_{dsp}(\eta), \tag{9.23}$$

where

V_{pre} = vertical presample MTF,
H_{pre} = horizontal presample MTF,
ν = horizontal sample frequency in $(\text{mrad})^{-1}$, and
γ = vertical sample frequency in $(\text{mrad})^{-1}$.

Contrast threshold function data are modulation contrast. For consistency, T_{con} is used for target contrast. T_{con} equals modulation contrast of the displayed target image.

$$T_{con} = \frac{C_{tgt}}{2 S_{tmp}} . \tag{9.24}$$

Like detector noise, the aliased signal is filtered by the eye and visual cortex. It is then summed in quadrature with eye noise.

$$QH_{alias}(\xi) = \int \left| A_H(\xi) B(\xi' / \xi)\, H_{eye}\left(\xi' / SMAG\right) \right|^2 d\xi' . \tag{9.25}$$

$$QV_{alias}(\eta) = \int \left| A_V(\eta) B(\eta' / \eta)\, H_{eye}\left(\eta' / SMAG\right) \right|^2 d\eta' . \tag{9.26}$$

$CTFH_{sys}$ and $CTFV_{sys}$ are given by Eqs. (9.27) and (9.28). Aliasing is fixed pattern noise. In order to RSS the aliased signal with detector noise, the signal must be properly scaled temporally as well as spatially. The aliased signal uses the α associated with photon counting. α is $\alpha_{thermal}$ multiplied by the $\sqrt{t_{eye}}$. The model uses the noise in one square milliradian. The ratio L_{tgt}/R_{ng}^2 scales the aliasing noise spatially. L_{tgt} is the target area in square meters. R_{ng} is range in kilometers.

$$CTFH_{sys}(\xi) = \frac{CTF\left(\xi / SMAG\right)}{M_{dsp} H_{sys}(\xi)} \cdot \left(1 + \frac{\alpha_{thermal}^2\, \Gamma_{det}^2\, QH_{hor}(\xi)\, QV_{hor}}{S_{tmp}^2} + \frac{\alpha^2\, T_{con}^2 R_{ng}^2\, QH_{alias}}{L_{tgt} S_{tmp}^2} \right)^{1/2} . \tag{9.27}$$

$$CTFH_{sys}(\eta) = \frac{CTF\left(\eta / SMAG\right)}{M_{dsp}V_{sys}(\eta)} \cdot \left(1 + \frac{\alpha_{thermal}^2 \, \Gamma_{det}^2 \, QH_{ver}\,QV_{ver}(\eta)}{S_{tmp}^2} + \frac{\alpha^2 \, T_{con}^2 R_{ng}^2 \, QV_{alias}}{L_{tgt} S_{tmp}^2}\right)^{1/2} . \qquad (9.28)$$

In these equations, $\alpha_{thermal}$ equals 862 root-Hz, and α equals 169.6 root-Hz.

L_{tgt} is the scale size of spatial discrimination cues. Generally, L_{tgt} equates to the average size of objects in the identification set; however, this is not always the case. Perhaps ships have the same hull and superstructure but differ in armament. L_{tgt} is the average size of the armament, not the size of the ship.

9.5 Predicting Range Performance

The TTP metric (Φ) at range R_{ng} km is given by Eq. (9.29). This is the separable form of TTP as explained in Chapter 7. $\delta[C_{tgt}/CTFH(\xi)]$ is the probability of seeing a sine wave with contrast C_{tgt} at spatial frequency ξ when eye threshold is $CTFH(\xi)$.

$$\Phi = \left[\int \sqrt{\delta\left(\frac{C_{tgt}}{CTFH_{sys}(\xi)}\right)\frac{C_{tgt}}{CTFH_{sys}(\xi)}} \, \frac{d\xi}{R_{ng}} \int \sqrt{\delta\left(\frac{C_{tgt}}{CTFV_{sys}(\eta)}\right)\frac{C_{tgt}}{CTFV_{sys}(\eta)}} \, \frac{d\eta}{R_{ng}} \right]^{1/2} . \qquad (9.29)$$

Equation (9.30) gives the average PID members of a target set. Task difficulty for the set is $\Phi84$.

$$PID(\Phi / \Phi84) = erf\,(\Phi / \Phi84) = \frac{2}{\sqrt{\pi}} \int_0^{\Phi/\Phi84} e^{-t^2}\, dt . \qquad (9.30)$$

To predict PID versus range, the following procedure is used. The naked-eye CTF is degraded by imager blur and noise to establish $CTFH_{sys}$ and $CTFV_{sys}$. Φ is found at each range by a numerical integration corresponding to Eq. (9.29). PID is predicted using Φ and $\Phi84$ in Eq. (9.30).

9.6 Modeling Contrast Enhancement and Boost

Neither a target nor a realistic scene is in the resolution model. Eq. (9.29) quantifies imager resolution. Target contrast in Eq. (9.29) is flat across spatial frequencies. Equation (9.30) provides PID against a target set that is preselected to have diverse spatial cues. TTP quantifies how well spatial details are coupled from the scene to the human observer.

Quite often, however, the goal of image processing is to correct problems caused by the peculiarities of scene radiance. Local-area contrast enhancement (LACE), for example, corrects problems that arise in the scene and provides significant performance enhancement. That performance enhancement can be predicted using the TTP metric. However, the nature of the resolution model must be considered when modeling LACE.

Figure 9.12(a) illustrates why LACE is useful. The atmosphere radiates in the thermal spectral bands. The signature from the target at long range is on a radiance pedestal that is much larger than the pedestal for the target at close range. To prevent saturation, imager gain must be reduced. Low imager gain results in poor rendering of the targets.

In Fig. 9.12(b), a high-pass filter removes the large variation in radiance. Now the imager gain is increased, and the targets are much easier to see. Imager gain and level are optimized for the targets, not for the background.

LACE is generally implemented using a high-pass filter. However, modeling LACE by incorporating a high-pass filter into $H_{elec}(\xi)$ or $V_{elec}(\eta)$ is not appropriate. No information about the variation in scene radiance is input to the model. Therefore, the high-pass filter appears to boost high frequencies beyond the flat spectrum that is optimum.

The benefit of LACE is an increase in imager gain and improved scene contrast. Target details fill more of the display dynamic range, and display contrast is enhanced. The benefit of LACE is modeled by changing S_{tmp} in Eqs. (9.16) and (9.17).

Figure 9.13(a) is a thermal image of four tanks. Radiance is added, increasing radially from the top left corner. The intensity grows to 72 K at the bottom right corner. S_{tmp} equals 36 K. In Fig. 9.13(b), a high-pass filter is applied to remove the radiance ramp. The large variation in radiance is removed, but target contrast has not improved. S_{tmp} still equals 36 K. In Fig. 9.13(c), S_{tmp} is lowered to 6 K. Now vehicle details are visible. The high-pass filter allows the gain to increase. The improvement results from the increased gain.

Although LACE usually involves the use of a high-pass filter, sometimes a more sophisticated method such as Retinex or bilateral filtering is used. Whatever the choice of implementation, the benefit is modeled by increasing gain (decreasing S_{tmp}).

Note that the model is quite capable of predicting the benefit or penalty of digital boost. For example, incorporating a digital boost filter to compensate for diffraction blur helps predicted performance, provided that noise is not excessive. Unlike the variation of scene radiance, optical blur is a model input. The model calculates the benefit of overcoming a known blur.

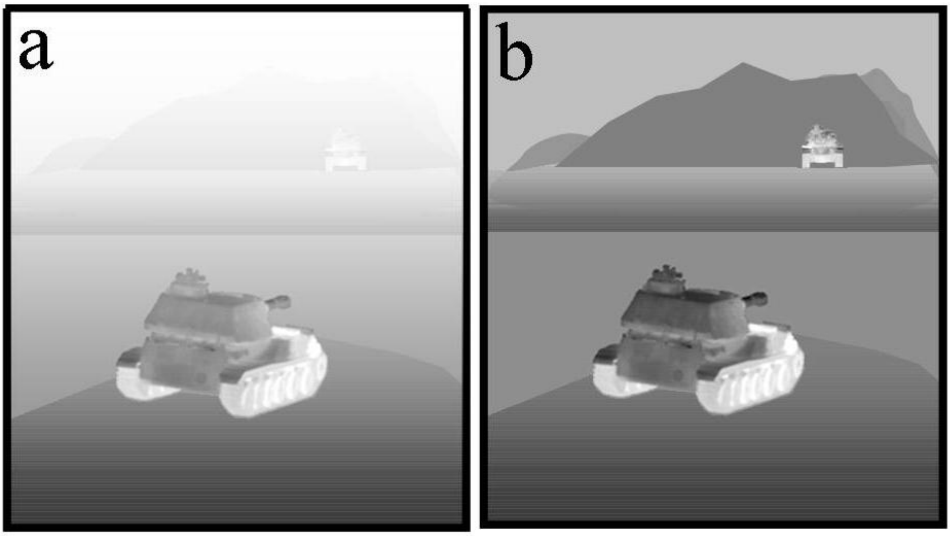


Figure 9.12 (a) Atmospheric path radiance creates a large signal swing over the imager field of view. Imager gain is reduced in order to prevent saturation. (b) A high-pass filter removes the low-frequency radiance component. Increased imager gain enhances the view of the distant target.

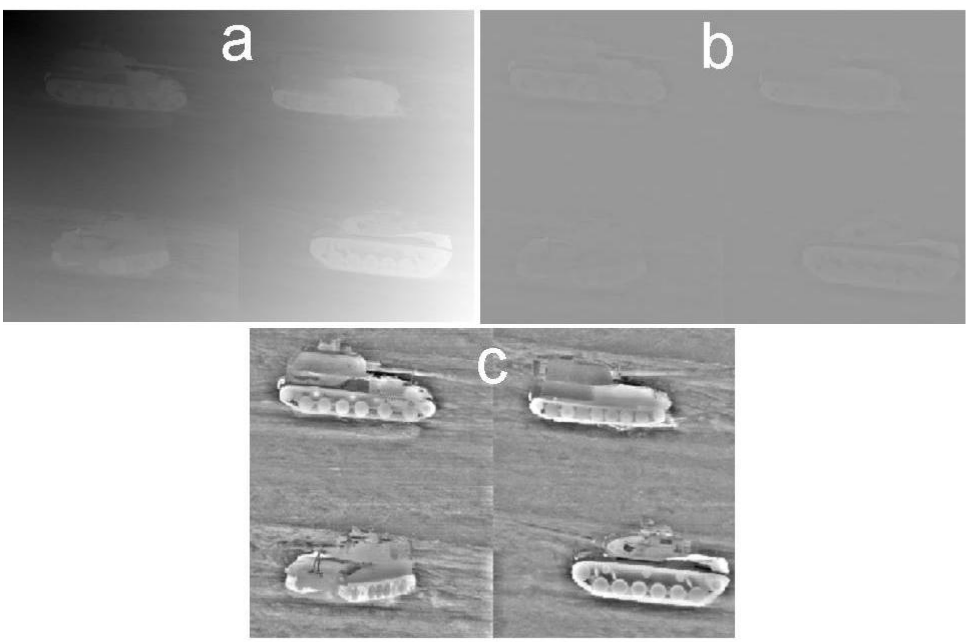


Figure 9.13 (a) A large radiance ramp is applied radially. Gain is low so that the radiance ramp does not saturate the picture. (b) A high-pass filter removes the ramp. Gain has not changed from (a), and target details are still low contrast. In (c), gain is increased so that target details become visible.

9.7 Minimum Resolvable Temperature

This section describes a model to predict the minimum resolvable temperature (MRT) performance of thermal imagers. MRT is commonly used to check system performance. Four-bar patterns of various sizes are viewed simultaneously. The minimum temperature between bar and space for bar visibility is the MRT. Good imager performance is associated with low MRT.

MRT setup is shown in Fig. 9.14. Plates with four bars cut out are placed in front of the blackbody. The temperature between the plate and blackbody is varied. The imager views the bar pattern through a collimator. The observer calls the pattern at the minimum temperature where all of the bars are visible. The procedure is conducted using a variety of bar sizes. The plot of minimum temperature versus bar size is called the MRT. A typical MRT plot is shown in Fig. 9.15.

Two problems plague laboratory MRT. First, it is difficult to achieve consistent results. The operator is permitted but not mandated to change gain, level, and sample phasing for each bar pattern viewed. MRT is supposed to represent the best achievable sensor performance. However, optimizing performance at each bar frequency is not mandated. Imagers lacking manual gain and level cannot be optimized for each bar pattern. MRT procedure varies from place to place, imager to imager, and engineer to engineer.

The second problem is that the procedure that is actually used during testing is not recorded along with the MRT data. The operator does not indicate whether the data is associated with optimizing gain and level, nor does the operator record any information about imager gain state. Therefore, MRT data is not associated with a specific procedure or known imager state.

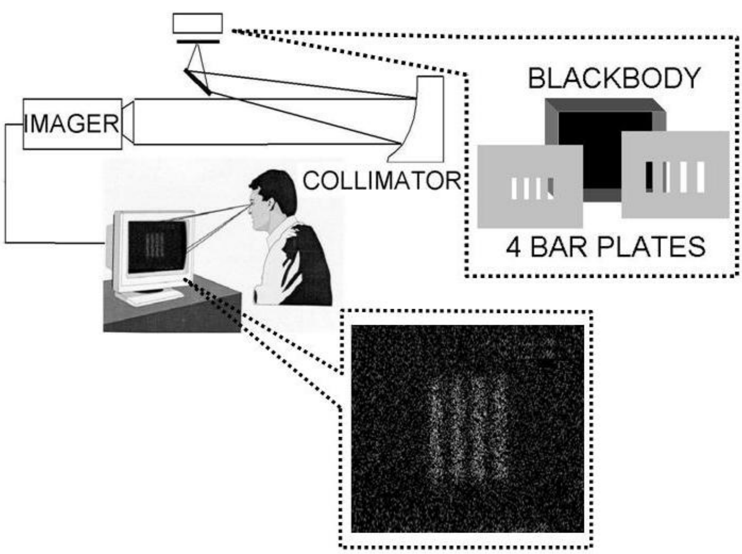


Figure 9.14 Observer views a four-bar pattern through a thermal imager. Blackbody temperature is varied to change bar contrast. Observer calls temperature where bars are barely visible.

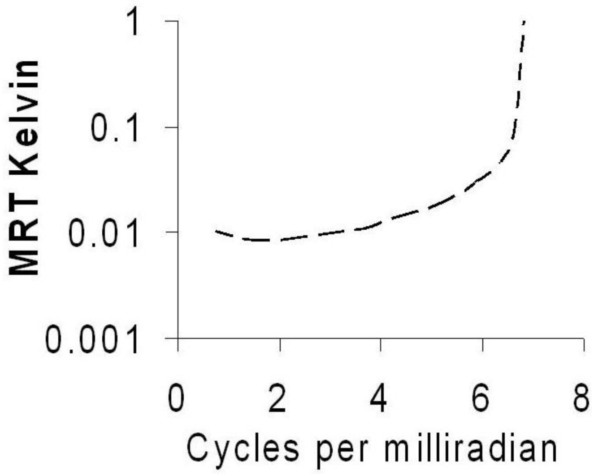


Figure 9.15 Typical MRT plot. Abscissa is bar-pattern period. Ordinate is minimum temperature where the bars are visible.

Inconsistency in MRT measurements is generally blamed on the subjective nature of bar-pattern testing. Judging when bars are visible is subjective. Most observers establish a mental standard and quickly become self-consistent. Learning consistency between operators takes longer and requires a desire to cooperate.

Recently, the triangle orientation discrimination (TOD) measurement has become popular because it is a forced-choice test. Various orientation and contrast TOD patterns are shown in Fig. 9.16. In a TOD experiment, the contrast of the triangles is varied in a manner similar to varying the contrast of bar patterns. In this case, however, the observer must call the orientation of the triangle. Using forced choice rather than a subjective call improves the chances of finding an accurate threshold.

The TOD procedure mitigates the error associated with the subjective nature of calling bar patterns. However, the TOD procedure shares all of the other problems associated with MRT. During measurements, the gain and level state of the imager are not specified or documented.

Section 9.7.1 presents MRT prediction theory. Section 9.7.2 adapts MRT theory for sampled imagers. MRT results vary depending on imager control settings. Section 9.7.3 suggests improvements to the typical MRT procedure to make data more predictable and consistent from measurement to measurement.

9.7.1 Predicting minimum resolvable temperature

To predict MRT, the sine-wave model is adapted for bar patterns. The model for predicting the effect of blur and noise on sine-wave grating detection was described in Section 9.3.

Figure 9.16 Triangle orientation discrimination (TOD) patterns. The operator must call the correct orientation. Using a forced-choice test reduces the errors associated with a subjective judgment such as calling bar visibility.

Figure 9.17 plots display intensity for an imager viewing a 0.125-cycle-per-milliradian ($mrad^{-1}$) four-bar pattern. The superimposed curve represents bar signal after blurring by the eyeball MTF and visual cortex bandpass filter. The amplitude difference between a center bar and the adjacent space closer to the edge of the bar pattern is what determines bar-pattern visibility. The amplitude difference between the locations indicated by the arrows is found and then compared to the threshold needed for sine-wave detection.

The Fourier spectrum of the cortical signal is found by multiplying the Fourier transform of the four-bar pattern by the various filters. Amplitudes are found by taking inverse Fourier transforms. MRT is given by Eq. (9.35) using A_{bar}, A_{space}, and S_L from Eqs. (9.32), (9.33), and (9.34), respectively. The difference $A_{bar} - A_{space}$ substitutes for H_{sys} in Eq. (9.16). The following definitions are used:

W = bar pattern period in mrad,
$\xi_0 = 1/(2W)$,
$H_W(\xi)$ = the bar width MTF = $\sin(\pi\xi W)/(\pi\xi W)$,
$H_L(\xi)$ = the bar length MTF = $\sin(7\pi\xi W)/(7\pi\xi W)$, and
S_L = fractional intensity due to blur of bar length.

$$H_{four\text{-}bar} = W\ H_W(\xi)[2\cos(2\pi W\xi) + 2\cos(6\pi W\xi)]. \qquad (9.31)$$

$$A_{bar}(\xi_0) = W\int_{-\infty}^{\infty} H_{sys}(\xi)\, B(\xi/\xi_0)\, H_{four-bar}(\xi)\cos(2\pi W\xi)\, d\xi\,. \qquad (9.32)$$

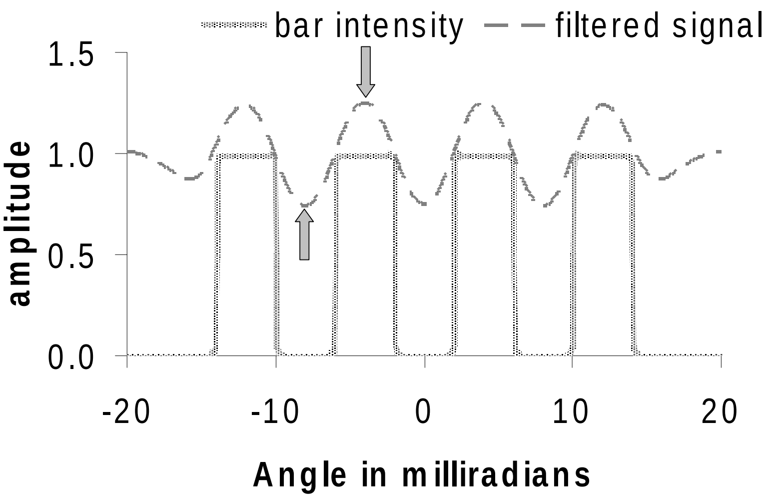


Figure 9.17 Display intensity of 0.125-mrad^{-1} four-bar pattern. The dashed line is bar intensity filtered by eyeball MTF and visual cortex band-pass filter. The filtered signal is raised to superimpose on the bar intensity for easy comparison. Threshold intensity is the difference in amplitude of the filtered signal at locations indicated by arrows.

$$A_{space}(\xi_0) = W \int_{-\infty}^{\infty} H_{sys}\left(\xi\right) B(\xi / \xi_0) H_{four-bar}\left(\xi\right) \cos(4\pi W \xi)\, d\xi\,. \tag{9.33}$$

$$S_L(\xi_0) = L \int_{-\infty}^{\infty} H_{sys}\left(\xi\right) H_{eye}\left(\xi / SMAG\right) H_L\left(\xi\right) d\xi\,. \tag{9.34}$$

$$MRT(\xi_0) \big/ 2S_{tmp} =$$

$$\frac{CTF\left(\xi_0 / SMAG\right)}{\left[A_{bar}(\xi_0) - A_{space}(\xi_0)\right] S_L(\xi_0)} \cdot \left(1 + \frac{\alpha_{thermal}^2 \Gamma_{det}^2 \; QH_{hor}(\xi)\, QV_{hor}}{S_{tmp}^2}\right)^{1/2}. \tag{9.35}$$

S_{tmp} K is the change in scene temperature that raises display luminance from black to average. Imager gain is L/S_{tmp} with units fL/K. Knowing S_{tmp} and L, variations in display luminance are related directly to variations in scene radiant temperature.

In Eqs. (9.32) and (9.33), note that $B(\xi)$ is used to filter the four-bar Fourier transform, but $H_{eye}(\xi)$ is not. We are comparing bar threshold to CTF measured with gratings that are essentially a single frequency. Eyeball MTF degrades threshold during CTF measurements, but the visual cortex filters do not. However, because of the Fourier spectrum of the bars, the visual cortex filters do affect bar threshold. Since linearity is assumed, bringing the visual cortex filters forward to the display is allowed mathematically.

The S_L factor is needed because high-frequency MRT bars are short. High-frequency four-bar patterns are difficult to see partly because of reduced spacing between the bars and partly because of the short length. CTF gratings are long and have constant length, regardless of frequency. S_L accounts for the effect of bar length on threshold.

Equation (9.35) gives MRT when gain and level are not optimized for each bar size. If gain and level are adjusted, bar modulation covers a greater portion of the display dynamic range, as shown in Fig. 9.18. The imager is adjusted such that the bar modulation fills half of the dynamic range of the display. Since S_{tmp} is defined as the temperature difference that raises the display from black level to average luminance, the following relationship holds at threshold:

$$S_{tmp} = MRT\left[A_{bar}(\xi) - A_{space}(\xi)\right]S_L \, . \tag{9.36}$$

Using Eq. (9.36) in Eq. (9.35) and solving for MRT, the equation for laboratory MRT is given by Eq. (9.37):

$$MRT(\xi) = \frac{2CTF(\xi / SMAG)\,\alpha^2_{thermal}\,\Gamma^2_{det}\quad QH_{hor}(\xi)\,QV_{hor}}{\left[A_{bar}(\xi) - A_{space}(\xi)\right]S_L\sqrt{1 - 4\,CTF^2(\xi / SMAG)}} \, . \tag{9.37}$$

MRT for constant gain and linear intensity transfer is found using Eq. (9.35). Equation (9.37) is used when gain and level are optimized for each bar pattern.

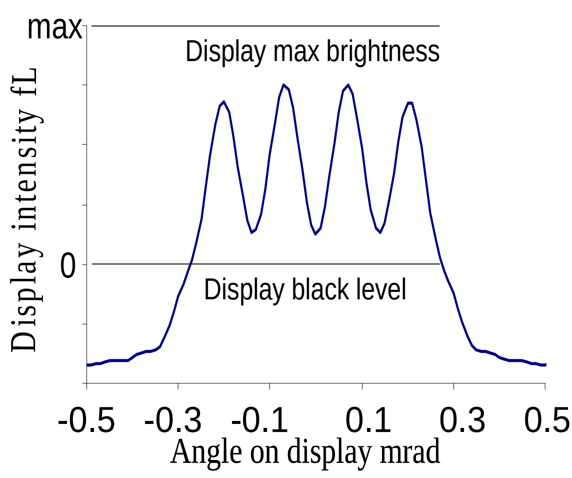


Figure 9.18 Plot of bar intensity on the display showing that bar modulation occupies half of the display dynamic range.

Note that Eq. (9.37) assumes sufficient imager gain to fill the display dynamic range with detector noise. An imager that is gain limited does not achieve the MRT predicted by Eq. (9.37).

9.7.2 Predicting sampled imager minimum resolvable temperature

The Fourier transform of the displayed image $H_{view}(\xi)$ of a sampled bar pattern is calculated by replicating the presample frequency content at multiples of the sample frequency ν and then filtering with the postsample MTF. Eyeball MTF is not in the postfilter because we are predicting contrast on the display. The eye/brain bandpass filter is included because the MRT stimulus (a four-bar pattern) is different from the stimulus used to generate CTF data (a sine wave). Equation (9.38) assumes that only the first three sampling replicas contribute to the image in practical situations. The sum over n is theoretically taken from $-\infty$ to $+\infty$.

$$H_{view}(\xi) = H_{post}(\xi)\, B(\xi) \sum_{n=-3}^{3} H_{four\text{-}bar}(\xi - n\nu)\, H_{pre}(\xi - n\nu)\, e^{in\phi} . \quad (9.38)$$

The sample phase ϕ is equal to $2\pi x\nu$, where x is the offset of sample position from the bar center. MRT is found for both zero sample phase and 180-deg sample phase. That is, MRT is found for both the case in which a sample point is at the center of the bar and the case in which the sample points have moved by half a sample spacing. The imaginary terms in Eq. (9.38) cancel, resulting in Eqs. (9.39) and (9.40) for a sample phase of 0 and 180 deg, respectively.

$$H_{view}(\xi,0) = H_{post}(\xi)\, B(\xi) \sum_{n=-3}^{3} H_{four\text{-}bar}(\xi - n\nu)\, H_{pre}(\xi - n\nu) . \quad (9.39)$$

$$H_{view}(\xi,180) = H_{post}(\xi)\, B(\xi) \sum_{n=-3}^{3} H_{four\text{-}bar}(\xi - n\nu)\, H_{pre}(\xi - n\nu) \cos(n\pi) . \quad (9.40)$$

The equations for A_{bar} and A_{space} now become

$$A_{bar} = W \int_{-\infty}^{\infty} H_{view}\left(\xi, \varphi\right) \cos(2\pi W \xi)\, d\xi . \quad (9.41)$$

$$A_{space} = W \int_{-\infty}^{\infty} H_{view}\left(\xi, \varphi\right) \cos(4\pi W \xi)\, d\xi . \quad (9.42)$$

However, since sampling can cause some unusual images, the presence of the center space must also be checked. Equation (9.43) gives the amplitude at the center of a four-bar pattern:

$$A_{center\ space} = W \int_{-\infty}^{\infty} H_{view}(\xi, \varphi)\, d\xi \ . \tag{9.43}$$

The smallest difference ($A_{bar} - A_{space}$) or ($A_{bar} - A_{center\ space}$) is used in Eq. (9.35) or (9.37) to calculate MRT. If either A_{space} or $A_{center\ space}$ is equal to or greater than A_{bar}, then at least one bar space is not visible.

An example is used to illustrate what occurs during the sampling process. The example imager is a 640 × 480 midwave infrared imager with *f*/3 optics and a 2.5-deg horizontal field of view. The FPA has 20-μm detectors on a 20-μm pitch (100% fill factor). The display is a 6-in high flat panel (square pixels) and is viewed with two eyes from 15 in. The half-sample frequency for this imager is 7.33 cycles per milliradian.

Figure 9.19 shows intensity in object space for an 8.07-cycles-per-milliradian MRT pattern. Figure 9.20 shows both the display intensity and the signal after filtering by the eye. In both figures, the abscissa is angle in milliradians, and the ordinate is intensity. The bar spatial frequency is a factor of 1.1 greater than the sensor's half-sample frequency. The triangles show the sample positions. One sample is at the center, so this is zero sample phase. In Fig. 9.20, the small "wiggles" in the display intensity result from the numerical inverse transform. The intensity across any one display pixel should be uniform. At this sample phase, the four bars are easily seen by the eye.

In Fig. 9.21, the sample positions have moved half of a sample spacing relative to the bar pattern, and sample phase is now 180 deg. Figure 9.22 shows the display intensity and intensity through the eye filters. The bar pattern is now spatially distorted due to the sample phase, and the four bars are not visible. Substantial bar modulation is visible, but not at the frequency of the MRT pattern.

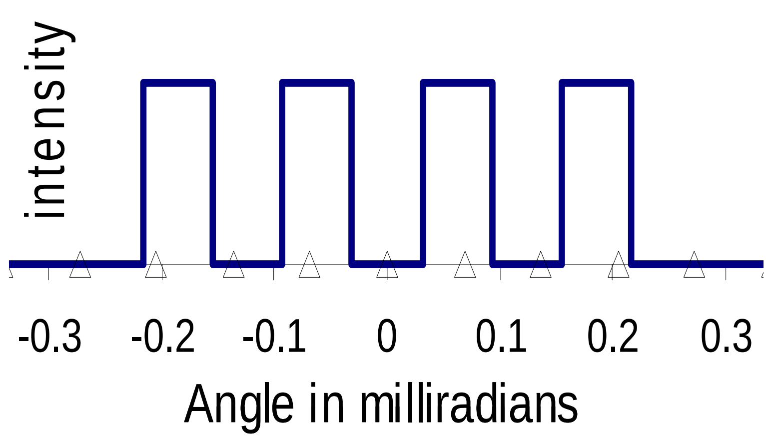


Figure 9.19 Intensity of bar pattern in object space. Sample locations are shown by triangles. This is zero sample phase because one sample is at the exact center of the bar pattern. Bar frequency is 1.1 times the imager's half-sample frequency. The corresponding display and visual cortex signals are shown in Fig. 9.20.

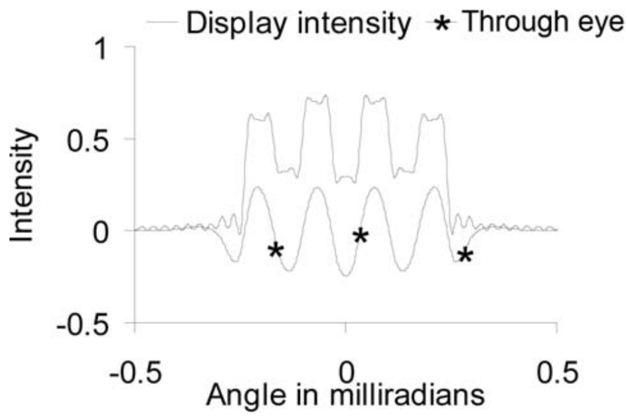


Figure 9.20 The top line shows display intensity of the Fig. 9.19 bar pattern viewed through the example imager. The visual cortex representation of the signal is the curve marked with asterisks (*). At zero sample phase, the bars are easily viewed by an observer.

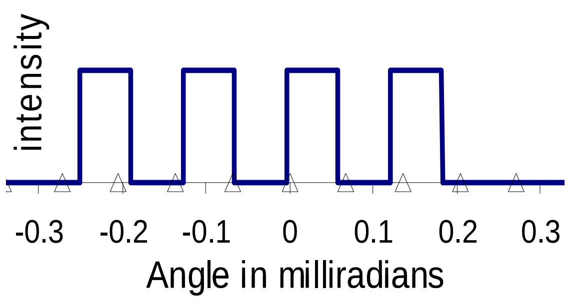


Figure 9.21 Same bar pattern as shown in Fig. 9.19, but the sample locations, marked by triangles, have changed. This is a 180-deg sample phase, meaning that sample locations have moved by half a sample spacing from the Fig. 9.19 locations. The corresponding display and visual cortex signals are shown in Fig. 9.22.

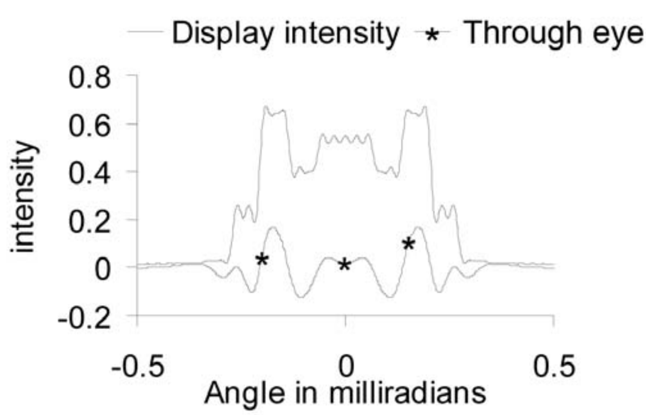


Figure 9.22 Intensity of display signal (top) and signal through eye (*) for an MRT bar pattern sampled as shown in Fig. 9.21. The center of the bar pattern is halfway between sample points. Four bars are not visible.

9.7.3 Improving the minimum resolvable temperature procedure

MRT has proven useful as an indicator of system operation. Broken optics or serious noise problems are quickly detected using the procedure. However, we know from experience that MRT is not a quantitative measurement. The subjective nature of bar-pattern detection is not the only reason for inconsistent data.

In the mid 1980s, 25 production thermal imagers were thoroughly evaluated. The imagers were first-generation common modules fielded by the U.S. Army. The evaluation included MRT performed by experienced laboratory personnel. A wide variation in measured MRT was observed. However, the reason for the variation in MRT could not be identified.

In all cases, the sensor NETD, MTF, and system intensity transfer function were measured. No problems were found. In rare cases, problems with the display were discovered, and the display replaced. The systems were evaluated for electromagnetic interference, and no problem was revealed. Further, the systems were tracked through an extended field exercise, and individual operators were questioned nightly about their experience with the imagers. There appeared to be no correlation between MRT and user opinion about image quality. On some nights imagery was good. On other nights imagery was poor. But the operators reported no difference between individual imagers.

Figure 9.23 shows MRT data from 15 of the 25 systems; data from the remaining 10 systems are lost. The solid line shows predictions of the high-gain model and represents best predicted MRT performance. The dashed line shows predictions for the low-gain model and represents the worst predicted performance. All data lie between the high- and low-gain predictions.

No physical cause could be found for the wide variation in imager MRT shown in Fig. 9.23. There was no difference in performance reported by the operators. And yet, the author and other engineers involved in the evaluation felt certain that an imager problem existed, and that this was proven by the MRT results. Early MRT models were based strictly on signal-to-noise ratio and had no dependence on imager gain and level. The extent to which imager gain and level affected MRT results was not understood.

Avoiding the errors associated with gain and level variation requires changing the typical MRT procedure. The changes require some additional effort. However, the additional measurements involve a procedure that is commonly performed in thermal imager evaluation facilities.

Measuring the system intensity transfer (SIT) function provides information on imager gain state. If imager gain and level are not adjusted during MRT measurement, then the fixed gain SIT must be measured. If gain and level are optimized for each MRT bar pattern, then the maximum imager gain is needed. The operator documents MRT data and SIT for the selected gain state or maximum gain and describes whether gain and level optimization is applicable.

A good fit between model predictions and laboratory MRT indicates that the hardware is operating properly. Furthermore, a match between predictions and

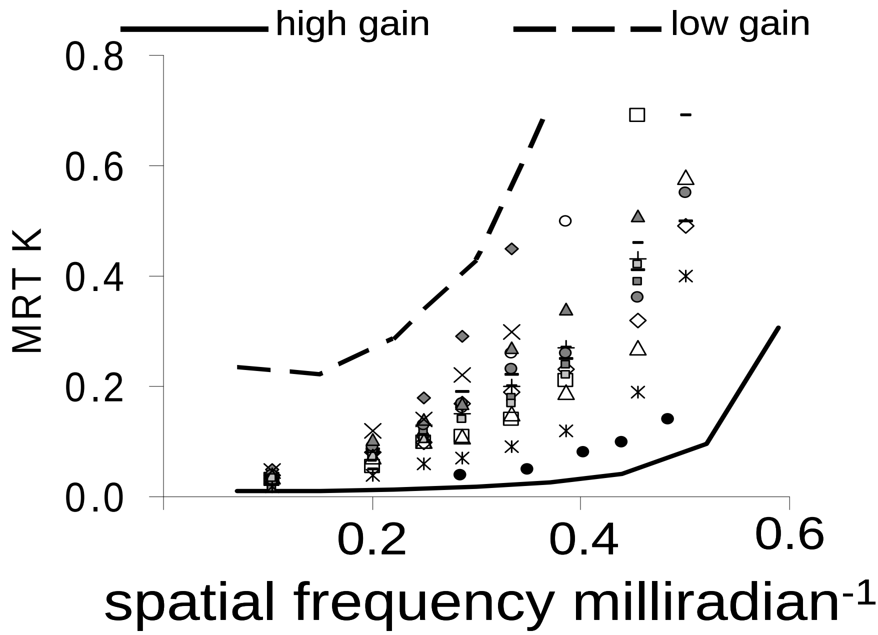


Figure 9.23 MRT versus spatial frequency for 15 production systems. Each symbol represents data for a different imager. The dashed line shows low-gain MRT prediction. The solid line shows predictions for high gain. There is a wide variation in MRT data, but all points fall within the low- and high-gain predictions.

data verifies model inputs. Currently, however, MRT procedure varies from location to location and engineer to engineer. Neither the exact procedure nor the gain state of the imager is recorded with MRT data. Both consistency between experiments and a good match to theory require rethinking current laboratory practices. Selecting a single procedure is not necessary. Recording the procedure used along with the MRT data is necessary. Imager gain state information must also accompany the MRT data.

Bibliography

Driggers, R. G., P. Cox, and T. Edwards, *Introduction to Infrared and Electro-Optical Systems*, Artech House, Boston (1998).

Hopper, G. S., "Forward-Looking Infrared Systems," Chapter 2 in *The Infrared & Electro-Optical Systems Handbook: Vol. 5 Passive Electro-Optical Systems*, S. B. Compana, Ed., SPIE Press, Bellingham, WA, 105–141 (1993).

Hudson, R. D., *Infrared System Engineering*, Wiley, New York (1969).

Land, E. H. and J. J. McCann, "Lightness and Retinex theory," *J. Opt. Soc. Am.* **61**, 1–11 (1971).

Tomasi, C. and R. Manduchi, "Bilateral filtering for gray and color images," *Proc. 1998 IEEE Int. Conf. on Computer Vision*, 836–846 (1998).

Tumblin, J. and H. Rushmeier, "Tone reproduction for realistic images," *IEEE Comp. Graphics & Applications* **13**(6), 42–48 (1993).

Vollmerhausen, R., R. G. Driggers, C. Webb, and T. Edwards, "Staring imager minimum resolvable temperature measurements beyond the sensor half sample rate," *Opt. Eng.* **37**(6), 1763–1769 (1998).

Vollmerhausen, R. H., "Predicting the effect of gain, level, and sampling on minimum resolvable temperature measurements," *Opt. Eng*. **48**(7), 076401 (2009).

Wolfe, W. L. and G. J. Zissis, Eds., *The Infrared Handbook*, The Infrared Information Analysis Center, ERIM, Ann Arbor, MI (1989).

Chapter 10
Imagers of Reflected Light

Imagers of reflected light operate in the spectral band between 0.4 and 2.0 μm. The visible band is from 0.4 to 0.75 μm. The near-infrared (NIR) band extends from 0.75 to about 1.0 μm. However, it is not uncommon to call the whole region between 0.75 and 2.5 μm the near infrared. Light with wavelengths between 1 and 2 μm is called short-wave infrared (SWIR).

Evaluation of reflected-light imagers differs from evaluation of thermal imagers in two important respects. First, the target signature is highly variable. Both the intensity and spectral nature of natural illumination depend on time of day and weather. Also, there is a wide variation in material spectral reflectivities. An object that is highly visible when viewed against one background surface might almost disappear when viewed against a different surface material. The combination of varying illuminations and different background characteristics changes target signal by orders of magnitude.

Photo-electron counting models are generally used in the reflective bands. This is the second difference between reflective and thermal models. In the thermal model, both signal and noise are represented by a joules-per-second flux. Noise is represented by the watts on the detector that generates the same rms amplitude over a 1-sec time period. In the reflective model, the photon flux is integrated for a time period equal to a frame time or an eye integration time. Both models are accurate when properly applied. However, the different treatments lead to different calibration constants.

Section 10.1 discusses factors that affect target-to-background signatures. Section 10.2 develops the system contrast threshold function (CTF_{sys}) for reflected-light imagers. Calculating imager resolution using the TTP metric is described in Section 10.3.

10.1 Calculating Target Set Contrast

Natural light is abundant in the 0.4- to 2.0-μm spectral band. Figure 10.1 shows illumination from sunlight, moonlight, and starlight. Starlight includes airglow. The figure shows illumination through the atmosphere. The moon and sun are both at a 60-deg zenith angle. The visible band is especially bright during the day. The SWIR is the brightest of the three bands on a moonless night.

Depending on the spectral band chosen, night illumination is five to eight orders of magnitude below clear daylight.

Four distinct atmospheric absorption bands are apparent in the illumination spectra. Significant absorption occurs at 0.95, 1.1, 1.4, and 1.9 μm. These absorption bands also affect atmospheric transmission. Transmission over a 1-km horizontal path is shown in Fig. 10.2. In addition to abundant natural illumination, the clear atmosphere is fairly transparent over most of the 0.4–2.0-μm spectral band.

Target and background reflectivities tend to vary with wavelength in this spectral region. Figure 10.3 shows the spectral reflectivity of a foreign-vehicle paint, sand, gravel, a mixed soil, and dead grass. The paint closely matches the gravel and soil out to about 1.2 μm and closely matches the sand beyond 1.2 μm. The paint has very different reflectivity properties from dead grass (the top curve in the figure) over the entire spectral range. The apparent contrast seen by the imager depends on the background as well as on the spectral band chosen.

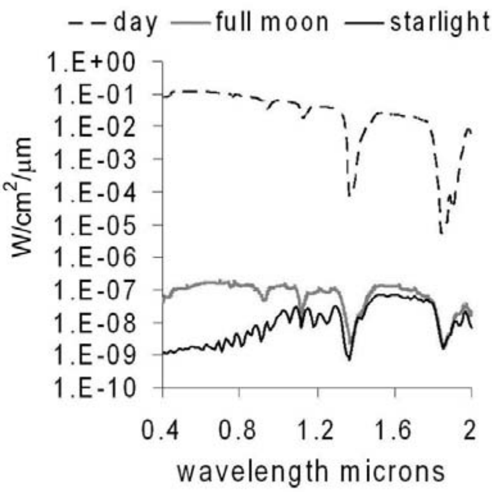


Figure 10.1 Illumination from the sun, moon, and starlight. Most of the "starlight" illumination is actually from air glow.

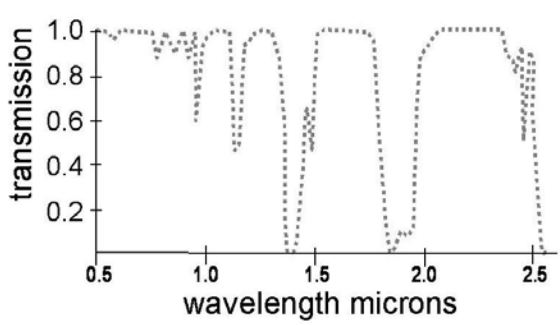


Figure 10.2 Atmospheric transmission over a 1-km horizontal path.

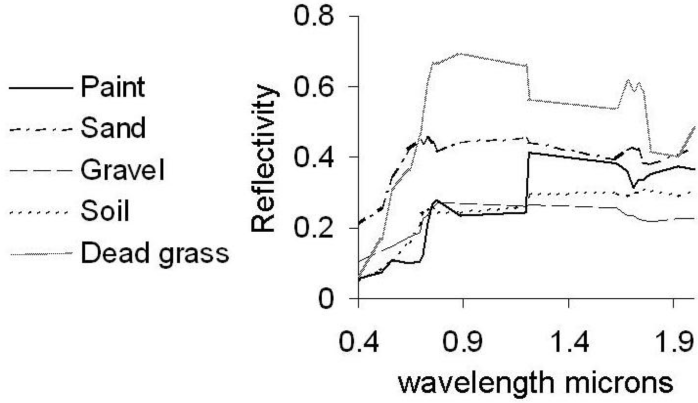


Figure 10.3 Spectral reflectivity of a foreign-vehicle paint compared to several backgrounds. Contrast is good or poor depending on the background material and spectral band selected.

The detector current from a scene element is calculated as follows:

$$\text{photocurrent} = \frac{V_{det} H_{det}}{4\, F\#^2} \int_{-\infty}^{\infty} I(\lambda)\, T(\lambda)\; R(\lambda)\, R_{sp}(\lambda)\; C(\lambda)\, d\lambda\,, \quad (10.1)$$

where

V_{det} = vertical dimension of detector in cm,
H_{det} = horizontal dimension of detector in cm,
λ = wavelength in microns,
F_0 = focal length of objective lens in cm,
$F\#$ = *f*-number: focal length F_0 divided by aperture diameter,
$I(\lambda)$ = illumination in watts cm^{-2} str^{-1} µm^{-1},
$T(\lambda)$ = transmission of atmosphere,
$R(\lambda)$ = spectral reflectance of the scene element,
$R_{sp}(\lambda)$ = detector response in amperes per watt, and
$C(\lambda)$ = objective lens and spectral filter transmission.

Let R_T and R_B represent the Eq. (10.1) spectral integrals for target and background, respectively. E_{av} is the average number of electrons in one second and in a square milliradian. Under bright illumination, the detector dwell might be limited to a fraction F_{gate} of the frame period.

$$E_{av} = \frac{(0.5 \times 10^{-6})\,(R_T + R_B)\, F_0^2\, F_{gate}}{V_{pit}\, H_{pit}\, e^-}, \quad (10.2)$$

where

e^- = charge on an electron (1.6 × 10^{-19} Coulombs per electron),
H_{pit} = horizontal detector pitch in cm,
V_{pit} = vertical detector pitch in cm, and
F_{gate} = fraction of frame period that photo-electrons are integrated.

The ratio $F_0^2/(V_{pit}H_{pit})$ gives the number of photodetectors in a square radian. Multiplying by 1 × 10^{-6} converts square radians to square milliradians. The unit "square radian" rather than steradian may seem strange. Remember, however, that the model treats two dimensions as two one-dimensional calculations.

In the reflective models, target contrast C_{tgt} is given by Eq. (10.3). Modulation contrast T_{con} is given by Eq. (10.4). Note that, unlike the thermal model, modulation contrast is expressed in the scene. In the thermal model, modulation contrast is expressed on the display.

$$C_{tgt} = \left| R_T - R_B \right|. \tag{10.3}$$

$$T_{con} = \frac{\left| R_T - R_B \right|}{R_T + R_B}. \tag{10.4}$$

10.2 System Contrast Threshold Function

In the reflective model, both signal and noise are based on integrating photocurrent for an eye integration time. Display luminance is proportional to t_{eye} E_{av}. Detector noise is proportional to $\sqrt{(t_{eye}E_{av})}$. So eye noise is compared to detector noise for the 0.04-sec eye integration time. The differences between the detectivity model and photo-electron integration model are described in Chapter 8.2. In these equations, κ_{con} is contrast enhancement due to gain and level settings.

$$CTFH_{sys}(\xi) = \frac{CTF(\xi / SMAG)}{M_{dsp}H_{sys}(\xi)}\left(\frac{1}{\kappa_{con}^2} + \frac{\alpha^2 E_{av}\, QH_{hor}\, QV_{hor}}{E_{av}^2}\right)^{1/2}. \tag{10.5}$$

$$CTFV_{sys}(\eta) = \frac{CTF(\eta / SMAG)}{M_{dsp}V_{sys}(\eta)}\left(\frac{1}{\kappa_{con}^2} + \frac{\alpha^2 E_{av}\, QH_{ver}\, QV_{ver}}{E_{av}^2}\right)^{1/2}, \tag{10.6}$$

where

α = 169.6 (a proportionality constant),
ξ = horizontal spatial frequency in $(\text{mrad})^{-1}$,

η = vertical spatial frequency in (mrad)$^{-1}$,
$CTF(\xi/SMAG)$ = naked-eye contrast threshold function,
κ_{con} = contrast enhancement,
$B(\xi \text{ or } \eta)$ = visual cortex bandpass filters; see the appendix,
$H_{eye}(\xi \text{ or } \eta)$ = eyeball MTF; see the appendix,
$H_{elec}(\xi)$ = horizontal electronics MTF,
$V_{elec}(\eta)$ = vertical electronics MTF,
$H_{dsp}(\xi)$ = horizontal display MTF,
$V_{dsp}(\eta)$ = vertical display MTF,
$H_{sys}(\xi)$ = horizontal system MTF,
$V_{sys}(\eta)$ = vertical system MTF,
QH_{hor} = horizontal noise bandwidth for $CTFH_{sys}$ defined by Eq. (10.7),
QV_{hor} = vertical noise bandwidth for $CTFH_{sys}$ defined by Eq. (10.8),
QH_{ver} = horizontal noise bandwidth for $CTFV_{sys}$ defined by Eq. (10.9), and
QV_{ver} = vertical noise bandwidth for $CTFV_{sys}$ defined by Eq. (10.10).

$$QH_{hor} = \int \left| B(\xi) H_{elec}(\xi) H_{dsp}(\xi) H_{eye}\left(\xi / SMAG\right) \right|^2 d\xi . \tag{10.7}$$

$$QV_{hor} = \int \left| V_{elec}(\eta) V_{dsp}(\eta) H_{eye}\left(\eta / SMAG\right) \right|^2 d\eta . \tag{10.8}$$

$$QH_{ver} = \int \left| H_{elec}(\xi) H_{dsp}(\xi) H_{eye}\left(\xi / SMAG\right) \right|^2 d\xi . \tag{10.9}$$

$$QV_{ver} = \int \left| B(\eta) V_{elec}(\eta) V_{dsp}(\eta) H_{eye}\left(\eta / SMAG\right) \right|^2 d\eta . \tag{10.10}$$

Figure 10.4 helps explain the contrast enhancement factor κ_{con}. Scene contrast is depicted in Fig. 10.4(a). Figure 10.4(b) shows the display signal with target contrast equal to A and average display luminance equal to D_{LUM}. In Fig. 10.4(c), gain is used to increase both target contrast and average luminance by B/A. In Fig. 10.4(d), the display brightness control or imager level is used to lower average display luminance back to the original level D_{LUM}. Target contrast increases by κ_{con} equal to B/A.

Imager gain and level can be used to enhance scene contrast. The first term in brackets in Eqs. (10.5) and (10.6) represents eye noise. The effect of eye threshold is mitigated by effective use of the imager gain and level controls. However, as the eye-noise term decreases, the detector noise becomes dominant. Image noise limits contrast enhancement.

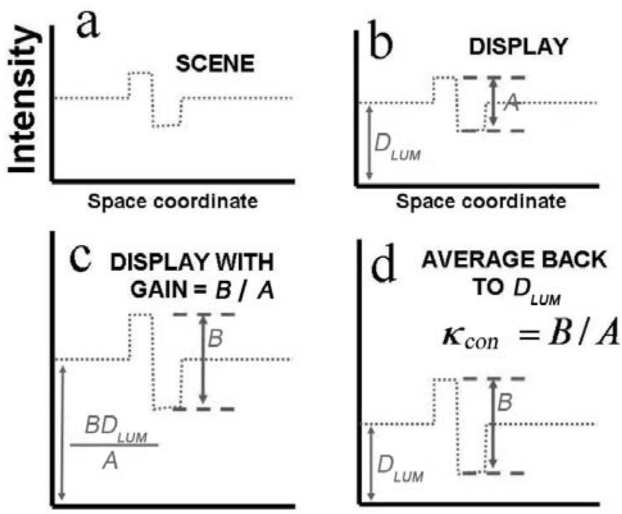


Figure 10.4 Plot (a) represents contrast in the scene. Plot (b) represents the display signal without adjusting gain or level. In plot (c), the display signal is gained by *B*/*A*. Both the contrast and the average luminance increase. In plot (d), the display brightness control lowers average luminance back to D_{LUM}. Target contrast has increased by *B*/*A*.

Equations (10.5) and (10.6) assume ideal shot noise; other noise sources are ignored. This assumption is realistic for most cameras under high-illumination conditions. However, as the light fails, noise sources other than shot noise begin to dominate.

Figure 10.5 illustrates the readout of a CCD imager. Photocharges are clocked down, row by row, until they reach the horizontal shift register. After each row enters the register, it is shifted out at high speed through the video amplifier. In this manner, the imagery collected in parallel at each detector becomes a serial stream. The benefit is a single output line. Generally the format is RS-170 standard video. The penalty paid for the convenience of a single output is noise. The high-speed video amplifier is noisy.

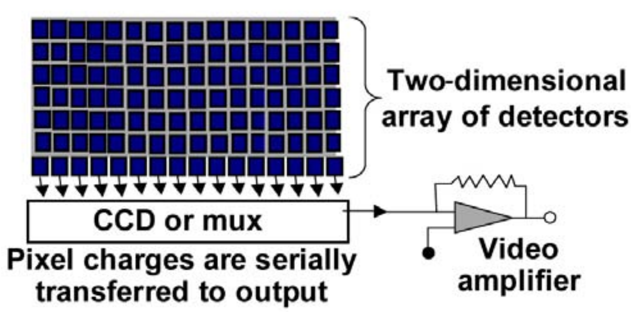


Figure 10.5 Illustration of a CCD. Each row is sequentially shifted down to the multiplexer (mux). The row is then shifted out through the video amplifier. The high-bandwidth video amplifier adds noise to the signal.

The video amplifier noise is typically specified in terms of noise electrons per pixel per field or frame. Although the noise actually arises in the amplifier or readout circuitry, manufacturers provide the equivalent number of noise electrons in order to make calculation of dynamic range and total noise easier.

A second common source of excess noise is dark current. Dark current is often specified as electrons per pixel per frame. Sometimes dark current is specified as current density. For example, the dark current might be specified as 100 μA (microamperes) per square centimeter. In that case, the active detector area and frame time are used to calculate dark electrons per pixel per frame. The noise associated with dark current is the square root of the number of dark current electrons.

A third source of noise is aliasing. The aliasing-as-noise (AAN) model is described in Chapter 8. All noise sources are added in quadrature. The shot, amplifier, and dark current noise in one second and one square milliradian is given by Eq. (10.11):

$$E_{noise}^2 = E_{av} + \frac{T_{CCD}\left(E_{amp}^2 + E_{DC}\right)F_0^2\,10^{-6}}{H_{pit}\,V_{pit}}, \tag{10.11}$$

where

E_{amp} = the amplifier noise in electrons per pixel per frame,
E_{DC} = dark-current electrons per pixel per frame, and
T_{CCD} = fields or frames per second.

Equations (10.12) and (10.13) give threshold vision through the imager. QH_{alias} and QV_{alias} are defined by Eqs. (10.14) through (10.17).

$$CTFH_{sys}(\xi) = \frac{CTF(\xi / SMAG)}{M_{dsp}H_{sys}(\xi)}\left(\frac{1}{\kappa_{con}^2} + \frac{\alpha^2 E_{noise}^2\, QH_{hor}\, QV_{hor}}{E_{av}^2} + \frac{\alpha^2\, T_{con}^2 R_{ng}^2\, QH_{alias}}{L_{tgt}E_{av}^2}\right)^{1/2}. \tag{10.12}$$

$$CTFV_{sys}(\eta) = \frac{CTF(\eta / SMAG)}{M_{dsp}V_{sys}(\xi)}\left(\frac{1}{\kappa_{con}^2} + \frac{\alpha^2 E_{noise}^2\, QH_{ver}\, QV_{ver}}{E_{av}^2} + \frac{\alpha^2\, T_{con}^2 R_{ng}^2\, QV_{alias}}{L_{tgt}E_{av}^2}\right)^{1/2}. \tag{10.13}$$

$$A_H(\xi) = \left[\sum_{n=1}^{\infty} H_{pre}^2(\xi - n\nu) \right]^{1/2} H_{elec}(\xi)\, H_{dsp}(\xi) \,. \tag{10.14}$$

$$A_V(\eta) = \left[\sum_{n=1}^{\infty} V_{pre}^2(\eta - n\gamma) \right]^{1/2} V_{elec}\, V_{dsp} \,, \tag{10.15}$$

where

H_{pre} = horizontal presample MTF,
V_{pre} = vertical presample MTF,
ν = horizontal sample frequency in $(\text{mrad})^{-1}$, and
γ = vertical sample frequency in $(\text{mrad})^{-1}$.

$$QH_{alias}(\xi) = \int \left| A_H(\xi) B(\xi'/\xi)\, H_{eye}\left(\xi'/SMAG\right) \right|^2 d\xi' \,. \tag{10.16}$$

$$QV_{alias}(\eta) = \int \left| A_V(\eta) B(\eta'/\eta)\, H_{eye}\left(\eta'/SMAG\right) \right|^2 d\eta' \,. \tag{10.17}$$

10.2.1 Interlace

Display interlace is used to reduce electronic bandwidth while maintaining a high-resolution image. Figure 10.6 illustrates interlace. The FPA operates at 60 Hz. However, the display operates at a 30-Hz frame rate. The first, third, fifth, and every odd row from the FPA are displayed in the first field. The even rows (two, four, six, etc.) are displayed in the second field. Although interlace does not degrade resolution, the displayed signal-to-noise ratio is affected because half of the available signal from the FPA is discarded.

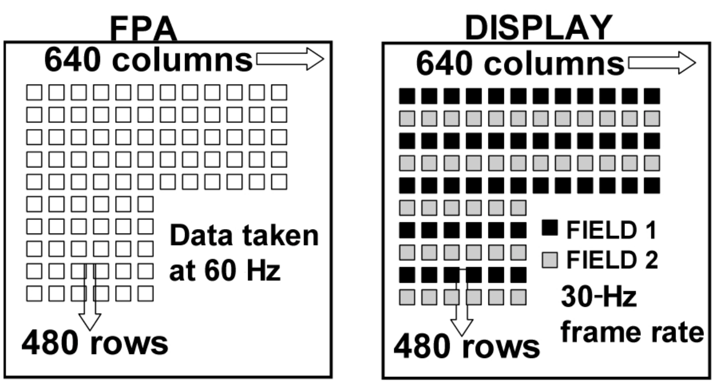


Figure 10.6 Illustration of interlace. Every other FPA row is displayed in every other display field.

Pseudo-interlace is a means for using all of the signal electrons while maintaining the reduced bandwidth benefits of interlace. In the first display field, photo-electrons from pixels in rows one and two are added and presented on display line 1. Pixels from rows three and four are added and presented on display line 3. Pseudo-interlace is illustrated in Fig. 10.7. The process continues, adding odd FPA rows to even FPA rows and displaying on odd display lines. During field 2, FPA rows two and three are added and presented on display line 2. Even FPA rows are added to odd rows and displayed on the even lines.

In Eqs. (10.2) and (10.11), E_{av} is divided by two when interlace is used. We do not reduce E_{av} for pseudo-interlace. Pseudo-interlace uses all of the available signal electrons and therefore maintains image sensitivity. Also, field alignment is properly maintained; samples are in the correct display location. The penalty paid for pseudo-interlace is a decrease in the vertical MTF of the imager. Vertical detector size doubles.

10.2.2 Snapshot and frame integration

When a single frame or snapshot is viewed, the eye does not temporally integrate multiple video frames. Snapshot noise E_{snap} is related to E_{noise} as shown in Eq. (10.18). E_{snap} replaces E_{noise} in Eqs. (10.12) and (10.13).

$$E_{snap} = \sqrt{t_{eye}\, T_{CCD}}\; E_{noise} \; . \tag{10.18}$$

In fact, noise amplitude actually decreases for shorter integration times. However, the signal decreases faster than the noise amplitude. Holding the display luminance constant means that electronic gain increases to overcome the reduced signal. The increased gain makes display noise larger in the snapshot mode.

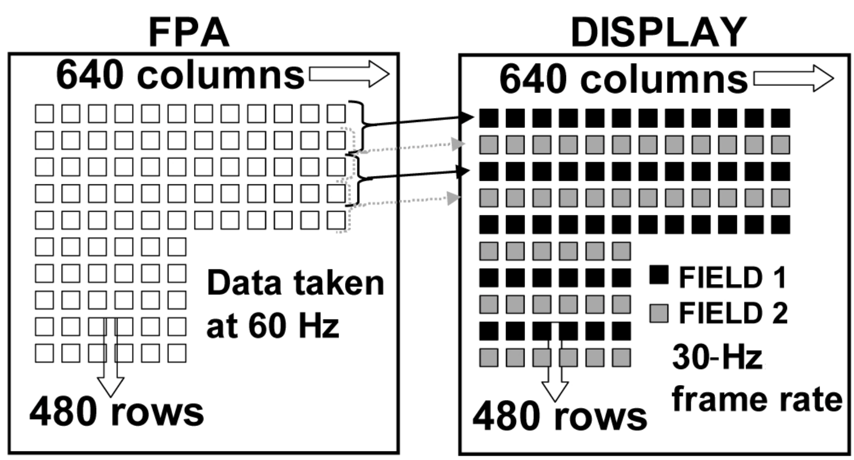


Figure 10.7 Illustration of pseudo-interlace. FPA rows are summed so that sensitivity is maintained. The technique does lead to added vertical blur.

If frame integration is used, then the effect depends on whether the imager is in framing or snapshot mode. If it is in snapshot mode, then Eq. (10.19) gives the benefit of integrating F_{INT} frames:

$$E_{snap} = E_{noise} \sqrt{\frac{t_{eye}\, T_{CCD}}{F_{INT}}}. \tag{10.19}$$

If the imager is in framing mode, then the benefit of frame integration is moderated by the fact that the eye already integrates temporally:

$$E_{frame} = E_{noise} \sqrt{\frac{t_{eye}\, T_{CCD}}{F_{INT} - 1 + t_{eye}\, T_{CCD}}}. \tag{10.20}$$

10.3 Predicting Range Performance

The TTP metric (Φ) at range R_{ng} km is given by Eq. (10.21). This is the separable form of TTP. $\delta[C_{tgt}/CTFH(\xi)]$ is the probability of seeing a sine wave with contrast C_{tgt} at spatial frequency ξ when eye threshold is $CTFH(\xi)$.

$$\Phi = \left[\int \sqrt{\delta\left(\frac{C_{tgt}}{CTFH_{sys}(\xi)}\right)\frac{C_{tgt}}{CTFH_{sys}(\xi)}}\, \frac{d\xi}{R_{ng}} \cdot \int \sqrt{\delta\left(\frac{C_{tgt}}{CTFV_{sys}(\eta)}\right)\frac{C_{tgt}}{CTFV_{sys}(\eta)}}\, \frac{d\eta}{R_{ng}} \right]^{1/2}. \tag{10.21}$$

Equation (10.22) gives the average probability of correctly identifying (PID) members of a target set. Task difficulty for the set is Φ84.

$$PID(\Phi / \Phi 84) = erf(\Phi / \Phi 84) = \frac{2}{\sqrt{\pi}} \int_0^{\Phi/\Phi 84} e^{-t^2}\, dt. \tag{10.22}$$

To predict PID versus range, the following procedure is used: The naked-eye CTF is degraded by imager blur and noise to establish $CTFH_{sys}$ and $CTFV_{sys}$. Φ is found at each range by a numerical integration corresponding to Eq. (10.21). PID is predicted using Φ and Φ84 in Eq. (10.22).

Bibliography

Electro-Optics Handbook, TP-135, 117–120, Burle Industries, Inc., Lancaster, PA (1974).

Driggers, R. G., P. Cox, and T. Edwards, *Introduction to Infrared and Electro-Optics Systems*, Artech House, Boston (1998).

Vollmerhausen, R. H. and T. Maurer, "Night illumination in the visible, NIR, and SWIR spectral bands," *Proc. SPIE* **5076**, 60–69 (2003).

Waldman, G. and J. Wootton, *Electro-Optical Systems Performance Modeling*, Artech House, Boston (1993).

Chapter 11
Computer Programs and Application Data

This chapter presents an overview of the computer programs *Thermal* and *Reflective*. Both programs have help files that describe individual inputs. This chapter does the following:

1. describes the physical models and assumptions implemented in the programs,
2. provides suggestions for modeling the MTF of objective lenses,
3. explains the interface to the detector data described in Chapter 12, and
4. suggests ways to use the programs to evaluate imager performance.

The programs are included for two reasons. First, the programs help the design engineer or systems analyst make imager design trades. The book provides suggestions for modeling optics and displays. The characteristics of available thermal FPAs are directly coupled to the *Thermal* computer program speed performance analysis. The computer programs are designed to provide quick but physically reasonable performance estimates.

The second reason for including the programs is to let the reader experiment with the theory. Blur, noise, and aliasing interact with target size and contrast to establish performance. For a given imager design, the relative importance of each factor is probably not obvious. For example, the dependency of aliasing on blur is clear from the theory. But the impact or importance of aliasing is not clear until the reader runs specific examples. The programs allow the user to vary imager and target characteristics while observing the effect on imager resolution.

Section 11.1 provides suggestions for estimating optical MTF. Section 11.2 provides estimates for the blur associated with the most common display technologies. Section 11.3 discusses the effect of atmosphere on performance and describes modeling assumptions. *Thermal* recalculates detectivity for each background temperature input by the user. The detectivity algorithms used in the program are summarized in Section 11.4. Both *Thermal* and *Reflective* models calculate diffusion MTF as well as spatial MTF for detectors. Detector MTF calculations are also described in Section 11.4. Section 11.5 explains computer program inputs and outputs. Section 11.6 offers suggestions for using the programs to analyze imager performance.

11.1 Optics Modulation Transfer Function

This section provides suggestions for modeling objective lens MTF. Many optics design programs such as CODE V, ZEMAX, and OLSO are available. Some of the numerous books covering optical design are listed in the bibliography. By far, the best procedure for modeling optical MTF is to use these programs in the manner explained in the literature.

In the early stages of system design, however, the output from optical design programs is not available. In the absence of better information, diffraction-limited performance is often assumed. In the MWIR and LWIR, assuming ideal MTF is reasonable for many lenses. The discussion in Section 11.1.1 describes when the ideal can be assumed and what to use for MTF in other cases. Broadband imagers operating in the visible, NIR, and SWIR do not achieve diffraction-limited performance. Section 11.1.2 provides MTF estimates for objective lenses operating in the NIR and SWIR spectral bands.

11.1.1 Thermal imagers

This section provides suggestions for modeling a thermal objective lens. Diffraction-limited performance is sometimes a realistic assumption. The computer model allows the user to easily select other options. This section describes those options and their applicability.

When designing thermal objectives, a number of factors must be considered. Producibility, weight, and volume are all important factors. Athermalization and performance over temperature extremes must be considered. Narcissus limits design options. Due to the high cost of infrared materials and their relatively poor transmission, lens thickness should be minimized. For cryogenic detectors, effective cold shielding is important for reducing detector noise. Most often, the system has multiple fields of view. These design considerations are all important and often require either compromising MTF or increasing the complexity and expense of the optical design.

Figure 11.1 shows MTFs for several production systems versus their diffraction limit. All of these systems except one operate in the MWIR. The abscissa is spatial frequency normalized to the diffraction limit for each optical system. The ordinate is MTF. The labels represent *f*/#, field of view, and whether the MTF is for on-axis (on) or off-axis (off). For example, "*f*/6, 1.5, off" represents an optical system with a 1.5-deg field of view using *f*/6 optics. Field of view is full angle. The "off" means the MTF is for an off-axis field angle of 0.7 of full field. For this example, the field angle is 0.525 deg. The notation "Tmp" means temperature. These MTFs are for a field angle equal to 0.7 of full field and at a temperature extreme of 70 °C.

All of the MWIR data in Figure 11.1 are predicted by optics design programs. The *f*/1.1 LWIR lens MTF is a Gaussian fit to measured data. Degradation due to nonideal spacing or alignment is not included in the MWIR estimates. However, measured data are available at lower spatial frequencies for

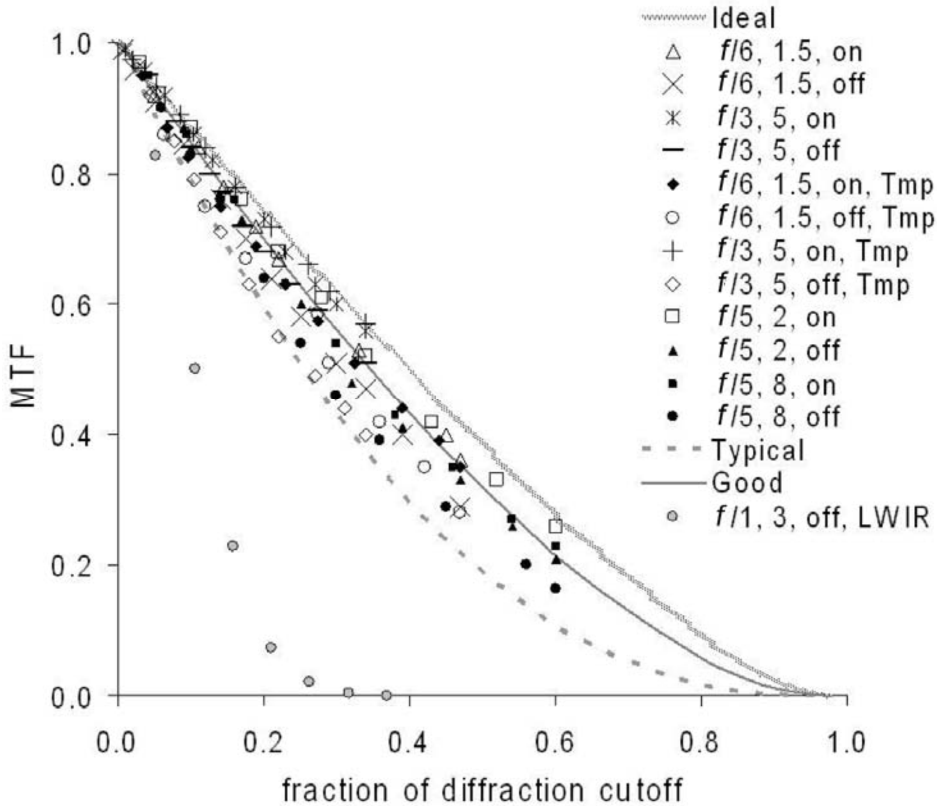


Figure 11.1 MTF for several MWIR objective lenses and one LWIR objective lens. The abscissa is spatial frequency normalized to diffraction cutoff for each lens. The LWIR MTF is a Gaussian fit to measured data. All of the other MTFs are predicted by optics design programs. The labels refer to *f*/#, full field of view in degrees, and on- or off-axis. "Tmp" means that MTFs are for off-axis and at a temperature of 70 °C.

the *f*/6 and *f*/3 lenses, and that measured data supports the MTF predictions. Available data are sparse at higher spatial frequencies due to the habit of not presenting MTF values beyond the half-sample frequency. The half-sample frequency depends on the detector array used. Signal at frequencies beyond "Nyquist" creates aliasing, and these MTF values are important in analyzing system performance. Nonetheless, published data tends to omit this information.

The ideal and good curves in Figure 11.1 bracket on-axis MTF for *f*/3 or higher optics. Off-axis performance and especially degradation due to temperature are bracketed by the good and typical curves. The name typical derives from the idea that off-axis and non-room-temperature environments represent the typical operational usage. For systems with fields of view larger than perhaps 15 deg, the typical MTF curve is probably optimistic.

Figure 11.2 compares the model ideal, good, and typical curves to design examples suggested by Fischer and Tadic-Galeb[1] and by Smith.[2] Design details for the examples are given in Table 11.1. Except for the *f*/0.75 lens, all of these examples use achromatic doublets with a field-flattening lens. Only Optic 6 provides 100% cold shielding of the detector. All designs are for a wide spectral band except for Optic 6; that lens provides poor MTF for the wide 3–5-μm spectral band. These objectives are simpler than designs found in real systems. Production systems provide 100% cold shielding and multiple fields of view by using more lens elements to achieve equivalent performance.

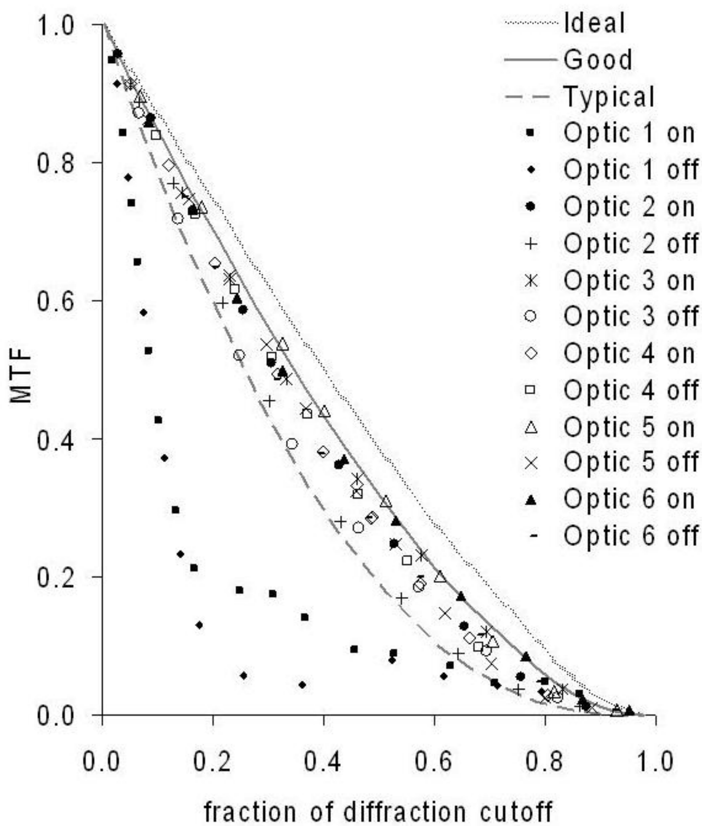


Figure 11.2 Lens design examples compared to model fits. The abscissa is spatial frequency as a fraction of diffraction cutoff for each lens. Ordinate is MTF.

Table 11.1 Design details for objective lenses. MTF for these lenses is shown in Fig. 11.2.

Optic	*f*/#	FOV (degrees)	focal length (cm)	spectral band (μm)	cold shield
1	0.75	6	10	8–12	no
2	1.5	3	16.5	8–12	no
3	1.5	12	10	3–5	no
4	3	2	10	8–12	no
5	3	12	10	3–5	no
6	3	3	10	3.6–4.2	yes

In Figure 11.2, the good and typical model curves bracket all of the MTFs except for the very fast (*f*/0.75) MWIR lens. MTF for high *f*/# and on-axis tends toward the good model. Low *f*/# and off-axis MTF is represented by the typical model curve.

In general, the ideal is easier to achieve in the LWIR than in the MWIR because surface errors are a smaller fraction of wavelength. For most optical designs, the ideal is easier to achieve on axis than off axis. Approaching the ideal is easier for large *f*/#s and smaller fields of view.

The thermal program allows the user to input optics MTF if it is available. Also, the user can select ideal, good, or typical. The MTF used by the program in each case is shown in Figs. 11.1 and 11.2 and given by Eqs. (11.1) through (11.3). Perhaps the best use of the typical, good, and ideal inputs is to run the program once using each option in order to understand the importance of the optical MTF in establishing overall imager resolution.

Ideal MTF is given by Eq. (11.1). The good and typical MTF curves are obtained by raising the diffraction MTF to the 1.2 power and 1.75 power, respectively. This is shown in Eqs. (11.2) and (11.3). Only horizontal MTF is shown; similar formulas are used for vertical MTF.

$$H_{x-diff}(\xi) = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{1000\xi\lambda}{D}\right) - \frac{1000\xi\lambda}{D}\sqrt{1-\left(\frac{1000\xi\lambda}{D}\right)^2}\right]$$
$$\text{for } \frac{1000\xi\lambda}{D} < 1,\ H_{x-diff}(\xi) = 0 \quad \text{for } \frac{1000\xi\lambda}{D} \geq 1, \tag{11.1}$$

where

ξ is horizontal spatial frequency in $(\text{mrad})^{-1}$,
D is aperture diameter in meters,
λ is wavelength in meters, and
$H_{x\text{-}diff}$ is the horizontal diffraction MTF.

$$H_{x-good}(\xi) = \left[H_{x-diff}(\xi)\right]^{1.2}. \tag{11.2}$$

$$H_{x-typical}(\xi) = \left[H_{x-diff}(\xi)\right]^{1.75}. \tag{11.3}$$

11.1.2 Imagers of reflected light

This section provides suggestions for modeling objective lens MTF for reflected-light imagers. The emphasis here and in the reflected-light computer program is on broadband passive imagers operating in the NIR and SWIR that use monochrome displays. The orientation is on long-range target acquisition, security, and mobility imagers.

Optics for passive imagers operating in the visible, NIR, and SWIR do not achieve diffraction-limited performance. Figure 11.3 shows MTF for several production lenses versus diffraction limit. The *f*/1.2 wide field-of-view (WFOV) imager has a 40-deg field of view. The spectral band is 0.6 to 0.9 μm. The *f*/2 SWIR optic has a 9-deg field of view and operates from 0.9 to 1.7 μm. The *f*/5 SWIR is the only reflective optic in this set and has a 2-deg field of view. The *f*/4 SWIR has a 10-deg field of view with a spectral band from 0.9 to 2 μm. As with the thermal examples, data are sparse at higher spatial frequencies due to the habit of not reporting MTF beyond the half-sample frequency.

Figure 11.3 shows the three model inputs available in the computer program: good, typical, and WFOV. The "book typical" curve shown in Fig. 11.3 is taken from Fig. 10.10 of Ref. 1. Only the reflective *f*/5 SWIR optic comes close to the good model and then only on the optical axis and at a fraction of the diffraction cutoff frequency. Off-axis performance for reflective and all of the narrow field-of-view refractive optics is represented by the typical model. For fields of view greater than about 10 deg, the WFOV model is used. These MTF curves are intended to be physically reasonable.

As with the thermal model, optics MTFs can be input by the model user. This is the best approach if the information is available.

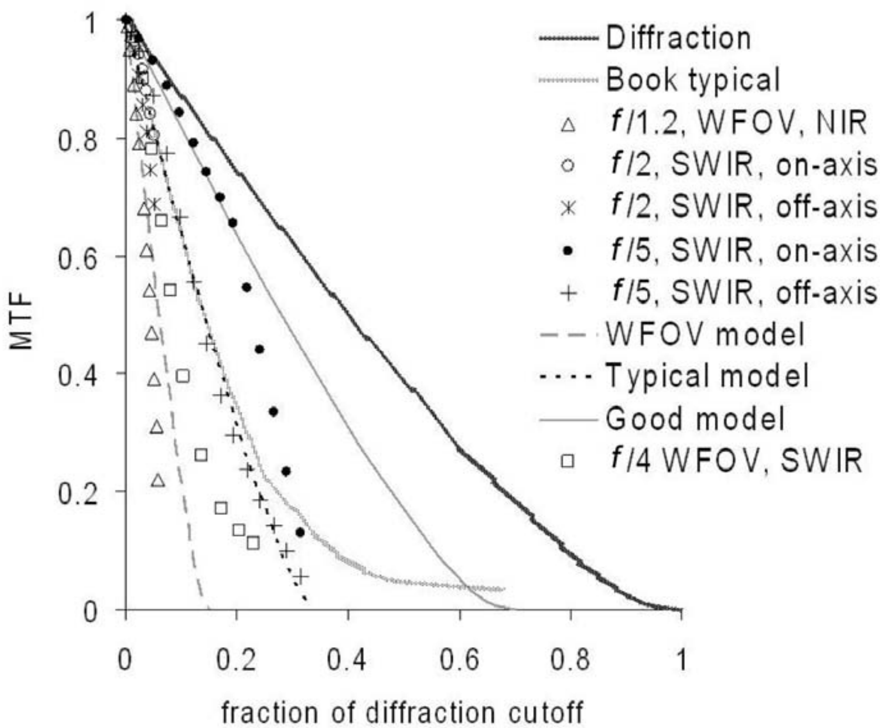


Figure 11.3 MTF of several NIR and SWIR objectives compared to model approximations. The abscissa is spatial frequency normalized to diffraction cutoff for each lens. Ordinate is MTF.

The good, typical, and WFOV model curves are based on diffraction MTF. Equation 11.4 gives the model MTF. The values of Ω are 0.7, 0.35, and 0.15 for good, typical, and WFOV, respectively.

$$H_{x-optic}(\xi,\Omega)=\frac{2}{\pi}\left[\cos^{-1}\left(\frac{1000\xi\lambda}{\Omega D}\right)-\frac{1000\xi\lambda}{\Omega D}\sqrt{1-\left(\frac{1000\xi\lambda}{\Omega D}\right)^{2}}\right]$$
$$\text{for }\frac{1000\xi\lambda}{\Omega D}<1,\ H_{x-optic}(\xi,\Omega)=0\ \text{ for }\frac{1000\xi\lambda}{\Omega D}\geq 1. \tag{11.4}$$

11.2 Display Modulation Transfer Function

This section describes how display MTF is calculated in the computer programs. Four display options are provided. These include high-resolution monochrome cathode ray tube (CRT), color CRT, color liquid crystal display (LCD), and high-resolution monochrome LCD. The color CRT also represents a low-resolution monochrome display. The primary market for the monochrome LCD displays is medical, so these are referred to as medical displays.

11.2.1 Cathode ray tubes

CRT technology is highly developed. Blur size is proportional to display size for diagonal dimensions from 1 to 30 in or more. Modern CRTs present a flat field that is free of line raster and other artifacts.

CRT blur is modeled as Gaussian.[3] Blur size is obtained using the shrinking raster dimension.[4] That is, the lines are spaced just close enough that raster disappears. This assumption provides two performance advantages. First, it maximizes appearance even when the observer is close to the display. Second, it provides the minimum blur size compatible with the flat-field assumption.

For a line pitch s millimeters (mm), spot blur MTF H_{spot} is given by Eq. (11.5). Spatial frequency ξ_{mm} is in cycles per millimeters on the display. Line spacing s is based on display height in millimeters divided by 525 lines for a low-resolution CRT and 960 lines for a high-resolution CRT.

$$H_{spot}(\xi_{mm})=\exp(-5.8s^{2}\xi_{mm}^{2}). \tag{11.5}$$

The user provides the diagonal dimension of the display. The program calculates display height based on a height-to-width ratio of three to four for most displays. For the high-definition television (HDTV) format, a nine-to-sixteen ratio is assumed. The dimension s is then obtained by dividing display height by either 525 or 960, depending on display resolution.

The shrinking-raster assumption is probably optimistic in most cases. The great variety of available CRT displays makes a definitive statement impossible.

However, Figs. 11.4 through 11.6 compare model results to measurement results from the author's experience.

Figure 11.4 compares the shrinking-raster assumption to measurements of a head-up display (HUD) CRT. The HUD uses a 1-in display viewed through a glass combiner and eyepiece. The video format is high resolution at 875 lines. The measurement data are taken through the combiner; this might explain the poor MTF at low spatial frequencies. The sharp drop near zero frequency is likely due to halation. The model line provides a somewhat optimistic but reasonable estimate of the Gaussian blur spot MTF. If the user is aware of additional MTF losses due to halation or contrast loss in the combiner, these can be included as additional post-MTFs.

Figure 11.5 compares model MTF to measurements of a color CRT display. Pixel pitch is 0.38 mm (67 lines per in). The data are represented by Gaussian curve fits. Unfortunately, the original data points are not available. Again, the model is reasonable although somewhat optimistic.

Figure 11.6 compares model to measurements on a high-resolution monochrome medical display. The line pitch in this case is 0.14 mm (180 lines per in). Again, the data are represented by a Gaussian fit. The model is quite optimistic in this case. Apparently, the video lines are closer together than in the shrinking-raster assumption.

The shrinking-raster assumption is optimum in the sense that blur spot size is minimized while maintaining a flat field. The flat field avoids aliasing generated at the display. The user has the option to add post-MTF and compensate for the optimistic display. However, good information on display MTF is generally not publicly available. The model provides a default Gaussian spot that optimizes the display but is physically reasonable.

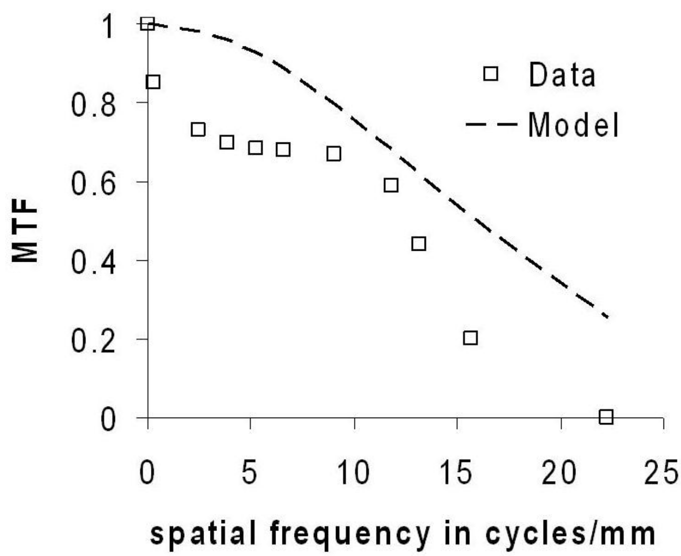


Figure 11.4 Model versus MTF data for a high-resolution HUD display.

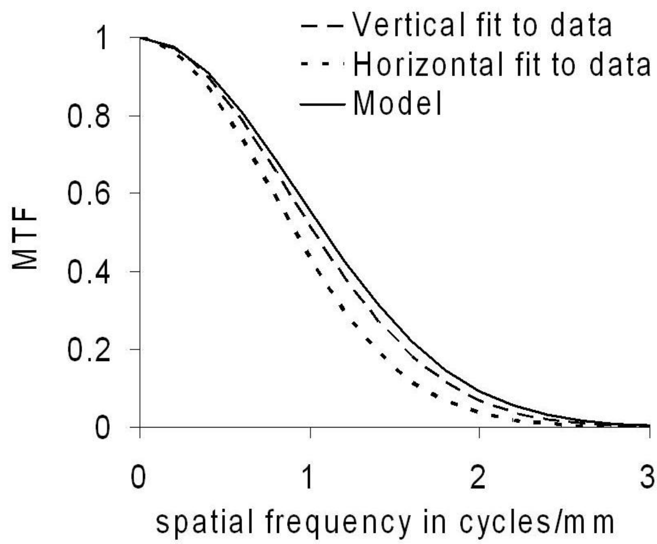


Figure 11.5 Measured MTF of color display compared to model prediction. The measured data are represented by Gaussian curve fits. The model prediction is reasonably close.

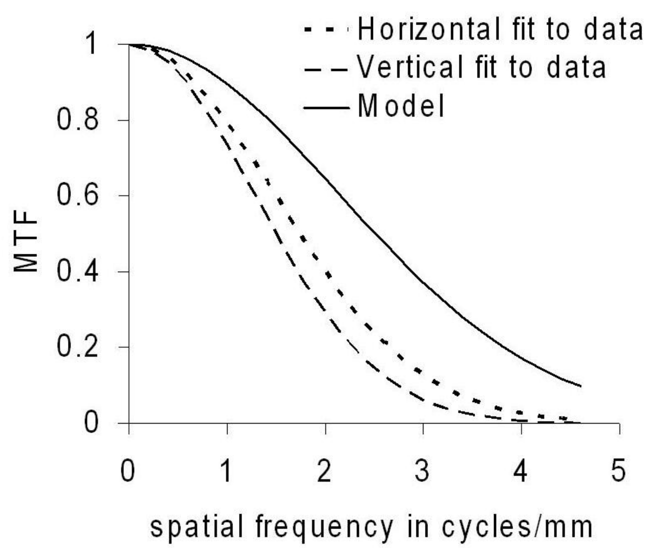


Figure 11.6 Comparison of model MTF to data for a monochrome medical display. The data are represented by Gaussian fits. The model is optimistic.

CRT MTF is also degraded by a horizontal sample and hold circuit. Using H_{pit} for horizontal pitch in mm, sample and hold MTF H_{SH} is given by Eq. (11.6):

$$H_{SH}(\xi_{mm}) = \frac{\sin(\pi\xi_{mm}H_{pit})}{\pi\xi_{mm}H_{pit}}. \tag{11.6}$$

11.2.2 Liquid crystal displays

Figure 11.7 shows a schematic of a color LCD unit cell.[6] This unit cell represents a high-aperture-ratio design. The active areas have a larger fill factor than most LCD displays sold today. The unit cell is small. The blue-, green-, and red-emitting areas are in close proximity. The eye cannot resolve the individual colored areas. Therefore, the cell produces one color, and that color depends on the relative amount of red, green, and blue light. White is produced by the relative intensities shown by the dashed lines in Fig. 11.8. Since the display is used to present a monochrome image, cells produce only white light in the model.

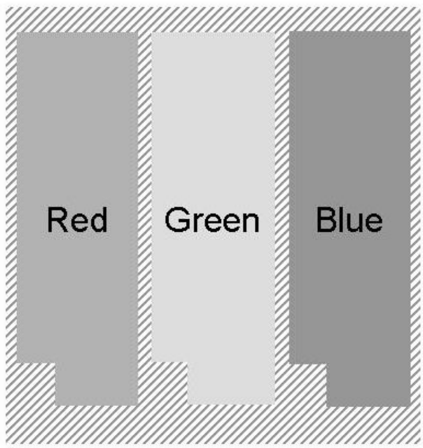


Figure 11.7 Illustration of an LCD unit cell. A high-aperture-ratio cell is shown.

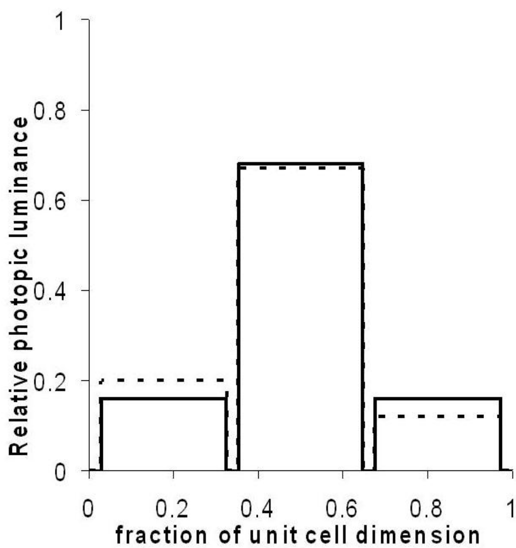


Figure 11.8 The dashed line shows relative luminance for each color cell to produce white light. The solid line shows the intensity profile used in the model.

The model permits only real MTFs. The dashed-line intensity pattern in Fig. 11.8 is not symmetrical; therefore, the MTFs are complex. The relative intensities shown by the solid line are used to model color LCD MTFs. Color LCD MTFs are shown in Fig. 11.9. These MTFs are for a unit cell dimension of 1 mm. The formulas for horizontal MTF H_{LCD} and vertical MTF V_{LCD} for a color LCD are given by Eqs. (11.7) and (11.8). H_{pit} is cell size in millimeters. The width of a color cell is 0.29 H_{pit}. The offset of the blue and red cells from the center is 0.32 H_{pit}. The vertical dimension of the color cells is 0.85 H_{pit}.

$$H_{LCD}(\xi_{mm}) = \frac{\sin(0.29\pi\xi_{mm}H_{pit})}{0.29\pi\xi_{mm}H_{pit}}[0.68 + 0.32\cos(0.64\pi\xi_{mm}H_{pit})]. \tag{11.7}$$

$$V_{LCD}(\xi_{mm}) = \frac{\sin(0.85\pi\xi_{mm}H_{pit})}{0.85\pi\xi_{mm}H_{pit}}. \tag{11.8}$$

The same cell construction is used for medical LCDs except that the color filters are removed. The three color cells are illuminated independently in order to obtain better amplitude quantization. This means that white-light MTF varies with intensity. For the purposes of the model, the three amplitudes are assumed to be equal. The vertical MTF for the medical LCD is given by Eq. (11.8). The horizontal MTF is given by Eq. (11.9). Medical LCD MTF is also shown in Fig. 11.9.

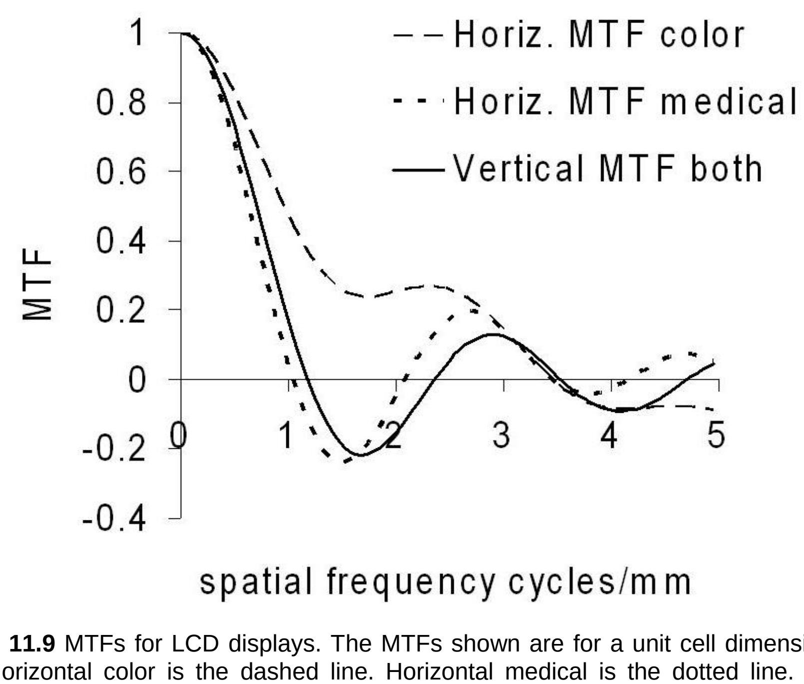


Figure 11.9 MTFs for LCD displays. The MTFs shown are for a unit cell dimension of 1 mm. Horizontal color is the dashed line. Horizontal medical is the dotted line. Vertical MTFs for both types of display are shown by the solid line.

$$H_{LCD}(\xi_{mm}) = \frac{\sin(0.29\pi\xi_{mm}H_{pit})}{0.29\pi\xi_{mm}H_{pit}}[0.33 + 0.67\cos(0.64\pi\xi_{mm}H_{pit})]. \quad (11.9)$$

Willem den Boer provides a clear discussion of available LCD technologies in Ref. 6. He also describes the sources of motion blur and explains how LCD televisions minimize motion blur.

11.2.3 Display interface format

Display interface options are shown in Table 11.2.

Table 11.2 Display interface options.[5]

Description	Standard	Horiz. pixels	Vert. lines	Inter-lace	Frame rate in Hertz
U.S. analog video (RS-170)	NTSC	640	480	Yes	60
High-resolution analog video	RS-343	1280	960	No	50 or 60
HDTV		1280	720	No	60
Computer	VGA series	any	any	No	50 or 60
European analog video	PAL	768	576	Yes	50
European component digital video		720	483	No	50

11.3 Atmospheric Transmission and Turbulence

Expected environment is one important factor in selecting an imager for a task. For example, atmospheric transmission in a high-humidity hot environment is better for MWIR than for LWIR. On the other hand, the longer wavelengths of LWIR improve turbulence MTF. The program provides a selection of atmospheres so that the user can evaluate the impact of atmosphere on imager performance.

The objective of imager analysis is to optimize future performance during a variety of tasks in unknown locations and unpredictable weather. For this reason, the program allows the user to easily change target set characteristics, atmospheres, and backgrounds in order to evaluate the imager under diverse conditions.

Generating a PID-versus-range curve for one atmosphere and for a single target contrast and dimension provides little information about future capability. For example, depending on weather and time of day, the thermal signature for nonexercised targets ranges from 0.1 to more than 10 K. Evaluating imager performance at 1 or 2 K leads to a myopic view of sensor performance.

When predicting contrast transmission, certain assumptions are made consistent with the nature of these models. One assumption is that the target and background are colocated. The target is viewed against local terrain. Distance to the target and to the background are the same. A second assumption is that the average flux seen by the imager does not change with range. User selection of background characteristics establishes background flux.

In reality, path radiance results in range-dependent background flux. This is particularly true in the reflective spectral bands. This means that both imager noise and target contrast are range dependent. Implementing range-dependent flux provides marginally better accuracy in the probability-versus-range curve. Range errors using the constant-flux model are small in the thermal bands. Range errors are largest in the reflective bands and for high sky-to-ground ratios (SGRs).

The objective of these models is to quantify imager resolution. Our view is that implementing the range-dependent model confounds the impact of multiple important factors without significantly improving range prediction accuracy. The simplified model allows the user independent control of both background flux and target signature. The program permits the user to stimulate the imager with a variety of realistic background and target signatures. The exact range at which a target might present the combination of signal and noise is not significant. What is significant is the ability of the imager to manage the diverse situations that will be encountered during years of use. Our goal is to provide the user insight into imager capabilities. The next two sections provide details on implementing the reflective and thermal models.

11.3.1 Atmosphere in the reflective model

Light from the sun, moon, and stars passes through the atmosphere before illuminating the target and background. Absorption bands in the atmosphere act on the light before the light reflects from ground objects. Light absorbed by the atmosphere is not present in the illumination spectra. In the reflective spectral bands, atmospheric absorption has little effect on horizontal path transmission.

Contrast is reduced by aerosol scattering of target signal out of the line of sight. Contrast is also reduced by sunlight, moonlight, or starlight scattered by aerosols into the imager's field of view. See Fig. 11.10 for an illustration. In most scenarios, path radiance resulting from light scattered into the sensor's path is the most serious cause of target-to-background contrast loss.

The atmospheric path can appear brighter at the imager than at the zero-range target and background. This results in substantial loss of contrast. Path radiance is quantified by the SGR. As the atmospheric path lengthens, the path becomes brighter. At some point, the path becomes optically thick. That is, only light from the path is seen, and increasing the path length does not change the path radiance. The SGR is the ratio between the maximum path radiance and the zero-range radiance. SGR does not vary with range because the peak long-range value is used in the ratio. Figure 11.11 shows the effect of SGR on contrast transmission. Table 11.3 gives values of SGR for a variety of terrains and cloud cover.

Equation 11.10 is used to calculate contrast loss for range R_{ng}, Beer's law coefficient β, and target zero-range contrast $C_{tgt\text{-}0}$:

$$C_{TGT} = \frac{C_{TGT-0}}{1 + SGR\left(\beta^{-R_{ng}} - 1\right)}. \tag{11.10}$$

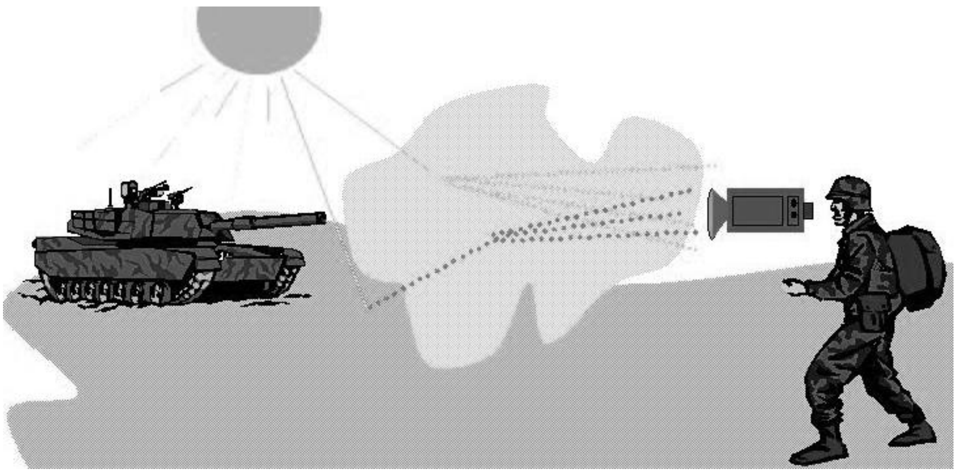


Figure 11.10 Sunlight scattered from atmosphere degrades target-to-background contrast.

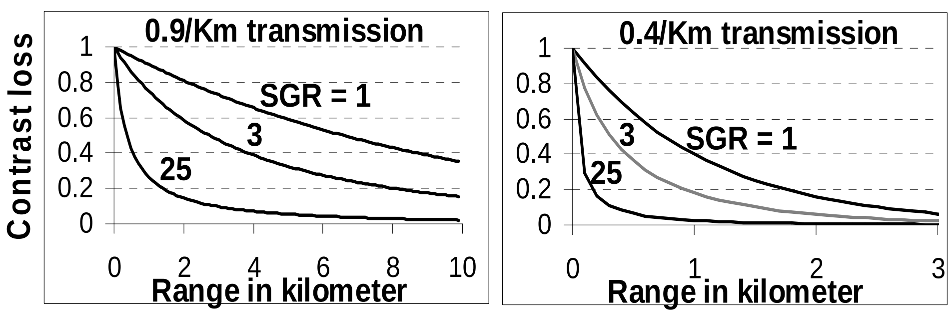


Figure 11.11 Effect of SGR on contrast transmission. The graph on the left shows the effect when the Beer's law transmission is 0.9 per km. The graph on the right shows the effect with 0.4-per-km transmission.

Table 11.3 Typical SGR ratios.

Terrain	Clear	Overcast
Desert	1.4	7
Forest	5	25
Typical	3	

Given knowledge of the path transmission and SGR, Eq. (11.10) is used to obtain target contrast at range. SGR is difficult to estimate, however, and a more empirical approach is used in the computer program.

Visibility is affected by both path transmission and SGR. Visibility is the range at which an observer can just distinguish a high-contrast target. However, visibility range cannot be used directly, because transmission through aerosols improves as wavelength increases. A method is needed to predict the wavelength dependency of contrast loss. The empirical method described by Waldman and Wootton is used in the *Reflective* program.[7] In Eq. (11.11), visibility V and range R_{ng} are both in kilometers. The extinction coefficient is V_{ext}. The center wavelength of the spectral band is λ_{center}.

$$Q = 0.585V^{0.3333} \text{ for } V \le 7 \text{ km.}$$

$$Q = 0.006583V + 1.07 \text{ for } V > 7 \text{ km.}$$

$$V_{ext} = \frac{3.91\left(\dfrac{0.55}{\lambda_{center}}\right)^{Q}}{V}. \tag{11.11}$$

$$\beta = \exp(-V_{ext}).$$

The model user inputs visibility (obviously in the visible band!), and the model calculates β for the imager's spectral band.

11.3.2 Atmosphere in the thermal model

MODTRAN 4 is the *de facto* standard for predicting transmission. However, it is not feasible to include MODTRAN in the program distribution. For that reason, five combinations of standard atmospheres and visibilities are included as transmission tables. The transmissions are calculated for horizontal paths only.

Type 1: U.S. Standard 1976 rural with 23-km visibility
Type 2: Tropical (15 N Latitude) Navy maritime aerosol
Type 3: Sub-Arctic winter with no aerosol
Type 4: Midlatitude winter with urban 5-km visibility
Type 5: Midlatitude summer, desert extinction

The one-wavenumber (cm^{-1}) MODTRAN transmissions are filtered with a triangular slit (40 cm^{-1} full-width half-maximum) and then sampled at 20 cm^{-1}. There are 642 values at each range covering the 2.5- to 13.5-μm spectral band. The data are stored for ranges 0.1, 0.3, 0.5, 1, 3, 7, 10, 20, and 50 km. A logarithmic interpolation at each wavenumber is used to obtain transmissions for intermediate ranges. The RMS errors associated with range interpolation are shown in Fig. 11.12. Except for the three-digit rounding errors, the RMS error goes to zero at stored ranges.

Both the largest absolute error and the largest RMS error occur at 0.2 and 2 km. The errors are associated with Type 5 atmosphere at 0.2 km. The errors occur with the Type 3 atmosphere at 2 km. Figures 11.13 and 11.14 show comparisons of model transmissions to MODTRAN for ranges with the largest errors.

Figure 11.13 compares model transmission for Type 5 atmosphere to MODTRAN. The range is 0.2 km. Figure 11.14 compares Type 3 model transmission to MODTRAN at a 2-km range. The fits are extremely good even at ranges with the largest RMS and absolute errors. The biggest errors occur at spectral locations where the transmissions change dramatically. These locations are circled in the figures.

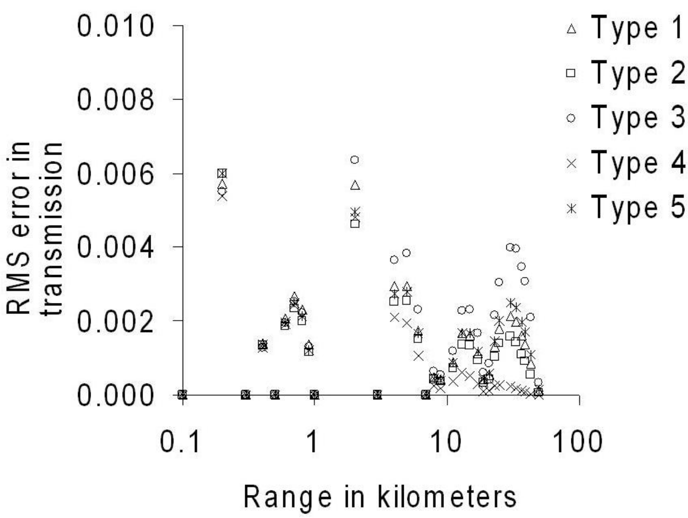


Figure 11.12 RMS error in transmission for various ranges between 0.1 and 50 km.

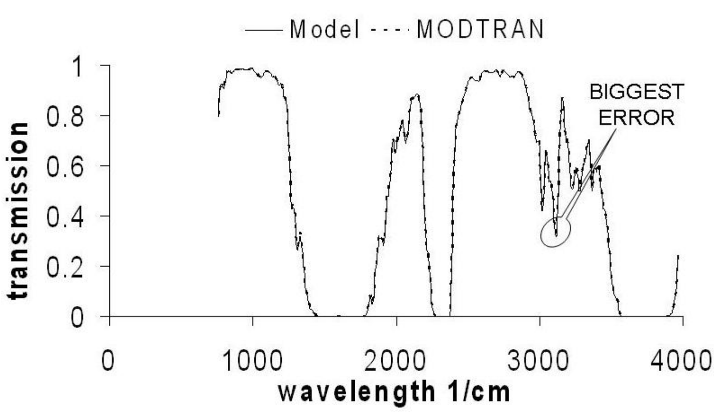


Figure 11.13 Comparison of model transmission to MODTRAN at 0.2 km for Type 5 atmosphere. The biggest error occurs at the circled wavelength.

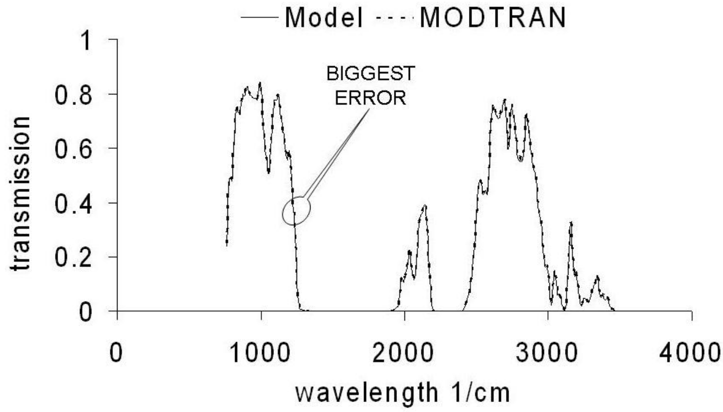


Figure 11.14 Comparison of model transmission to MODTRAN at 2 km for Type 3 atmosphere. The biggest RMS and absolute errors occur at this range. The wavelength with the biggest error is circled.

Model transmissions are virtually the same as equivalently filtered MODTRAN transmissions for these five atmospheres and horizontal paths to 50 km. These atmospheres provide the user with the capability of evaluating the effect of atmosphere on imager performance.

11.3.3 Atmospheric turbulence

All imagers are affected by atmospheric turbulence. In the computer model, the turbulence over a horizontal path can be varied in order to evaluate the susceptibility of the imager to turbulence. Turbulence is highly variable but worst at ground level. For this reason, turbulence has the greatest effect on horizontal paths. Turbulence tends to be low on cloudy days with no wind. Turbulence is low at sunrise and sunset and worst at midday. Hot conditions and certain terrains cause high turbulence.

In the computer model, the user selects no turbulence or low, medium, high, or very high turbulence. The corresponding CN2 values are 0, 1×10^{-14}, 1×10^{-13}, 5×10^{-13}, and 1×10^{-12}, respectively. A discussion of turbulence modeling and turbulence MTF is given by Ref. 8. Kopeika provides algorithms for predicting CN2.[9]

11.4 Detector Calculations

The purpose of this section is to document the algorithms used to calculate detector noise and MTF.

11.4.1 Detector noise

Detector noise is affected by flux from the scene, the enclosure temperature, amplifier noise, dark current, and objective lens transmission and absorption. The program allows the user to change each of these parameters. This section explains how each parameter affects total noise.

The detector flux environment is illustrated in Fig. 11.15. In A at the top, the detector is effectively cold shielded. All of the light hitting the detector is from the scene, with the exception of any light radiated from the lens itself. In B at the bottom of the figure, the cold shield allows light radiated from the enclosure to hit the detector. The flux on the detector depends on enclosure temperature as well as on scene temperature. Cold shielding is not effective, and enclosure temperature affects detector noise.

The physical constants used in the equations are discussed in *The Infrared Handbook*[10] and many other books on radiometry and Planck's blackbody equation. The origin of the constants is not described here. The following parameters, constants, and units are used in the equations:

C1 = 37,418 W μm^4 cm^{-2},
C2 = 14,389 μm K,
C3 = 2E–19 J μm,
$q_e = 1.6 \times 10^{-19}$ C (coulomb) (the charge on an electron),
λ = wavelength in microns,
λ_0 = start wavelength of spectral band,
λ_1 = stop wavelength of spectral band,
H_{pit}, V_{pit} = horizontal and vertical detector pitches in microns,
H_{det}, V_{det} = horizontal and vertical detector sizes in microns,
S_{scene} = electron flux per cm^2 due to scene,
$S_{enclose}$ = electron flux per cm^2 from enclosure,
S_{amp} = equivalent electron flux per cm^2 due to amplifier noise (calculated),
S_{drk} = electrons per cm^2 due to dark current (calculated),
A_{noi} = readout integrated circuit amplifier noise in electrons/pixel/frame (input),
D_{rk} = detector dark current in amperes per cm^2 (input),
σ_{vh} = fraction of flux equal to RMS fixed-pattern noise,
$F\#$ = *f*-number of objective lens,
F_{CS} = *f*-number of cold shield,
T_{CCD} = frame rate in Hertz (either 50 or 60 Hz),
E_{well} = maximum number of electrons in well,
$QE(\lambda)$ = quantum efficiency at wavelength λ,
$\lambda_{pk} = \lambda$ with peak QE,
$D_{\lambda pk}*$ = peak detectivity in Jones normalized per cm^2 and root-Hertz,
T_E = temperature of enclosure in Kelvin,
T_S = scene temperature in Kelvin,
τ = transmission of optics, and
β = absorption in optics.

The equivalent electron flux to generate amplifier noise in a square centimeter of detector area is given by Eq. (11.12):

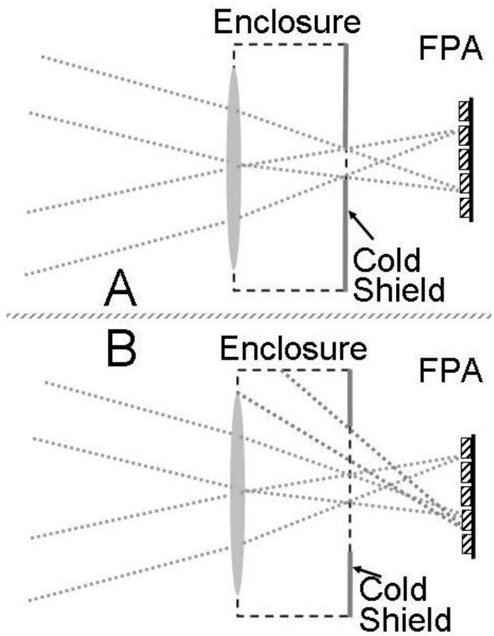


Figure 11.15 Illustration of detector cold shielding. The FPA is to the right of the figure. The cold shield is the barrier in the middle with a hole in it. The lens is to the left. In A (top), thermal flux from the enclosure is blocked by the cold shield. In B (bottom), the detector sees part of the enclosure, and flux from the enclosure increases detector noise.

$$S_{amp} = \frac{1\times 10^{8} A_{noi}^{2} T_{CCD}}{H_{det} V_{det}}. \tag{11.12}$$

The equivalent electron flux for dark current is the dark current per cm^2 converted to electron flux by dividing by q_e:

$$S_{drk} = \frac{D_{rk}}{q_e}. \tag{11.13}$$

The electron flux due to the scene is

$$S_{scene} = \tau \int_{\lambda_1}^{\lambda_2} \frac{\mathrm{C1}\, QE(\lambda)\, d\lambda}{4\mathrm{C3}\lambda^{4}\left[\exp\left(\frac{C2}{\lambda T_S}\right) - 1\right] F\#^{2}}. \tag{11.14}$$

The electron flux due to the enclosure is given by Eq. (11.15):[11]

$$S_{enclose} = \int_{\lambda_1}^{\lambda_2} \frac{C1\, QE(\lambda) d\lambda}{C3\lambda^4 \left[\exp\left(\frac{C2}{\lambda T_E}\right) - 1\right]} \left(\frac{1}{4F_{CS}^2} - \frac{1-\beta}{4F\#^2}\right). \tag{11.15}$$

$D_{\lambda pk}$* is given by Eq. (11.16).

$$D^*_{\lambda pk} = \frac{\lambda_{pk}\, QE(\lambda_{pk})}{C3\left[2(S_{scene} + S_{enclose} + S_{amp} + S_{drk})\right]^{0.5}}. \tag{11.16}$$

11.4.2 Detector modulation transfer function

Detector blur is caused by two distinct physical processes. One blur is associated with the spatial integration of light over the detector active surface. The second blur occurs in planar FPA without reticulated detectors. Planar detectors are subject to diffusion of minority carriers to adjacent pixels. Spatial integration sets the theoretical limit for MTF. However, many detectors also introduce significant blur due to diffusion. Total MTF is the product of spatial and diffusion MTF.

The following definitions are used in the equations:

d_{ff} = distance in centimeters that a photon penetrates into field-free material,
ξ_{cm} = spatial frequency in cycles per centimeters,
L_n = diffusion length in centimeters, and
H_{det} = detector size in microns.

Spatial integration MTF $MTF_{spatial}$ is given by Eq. (11.17). This equation is for horizontal MTF, but a similar equation holds for vertical. The 0.0001 factor converts detector dimension in microns to centimeters.

$$MTF_{spatial} = \frac{\sin(0.0001\pi\xi_{cm} H_{det})}{0.0001\pi\xi_{cm} H_{det}}. \tag{11.17}$$

In Eq. (11.18), L is the spatial-frequency-dependent effective diffusion length. Diffusion MTF $MTF_{diffusion}$ is given by Eq. (11.19).[12]

$$L = \left[\left(\frac{1}{L_n^2} + 4\pi^2\xi_{cm}^2\right)^{-1}\right]^{1/2}. \tag{11.18}$$

$$MTF_{diffusion} = \frac{\exp\left(d_{ff} / L_n\right) + \exp\left(-d_{ff} / L_n\right)}{\exp\left(d_{ff} / L_n\right) + \exp\left(-d_{ff} / L_n\right)}. \tag{11.19}$$

Only the spatial MTF of infrared detectors is generally modeled. This means that the detector MTF calculated in *Thermal* is often more pessimistic than when calculated by other models. An option is provided in the *Detector* window to exclude the diffusion MTF.

11.5 Computer Program Description

Figure 11.16 shows the start-up screen of the *Thermal* program. Clicking on *Lens*, *Focal Plane Array*, *Digital Filter*, *Display*, or *Extra pre- and post-blurs* opens a window to enter parameters associated with the corresponding imager component. Any changes in input parameters made since opening the input file or the last run (whichever occurs later) are recorded in the text window. If the parameters are returned to their original values, the change text disappears.

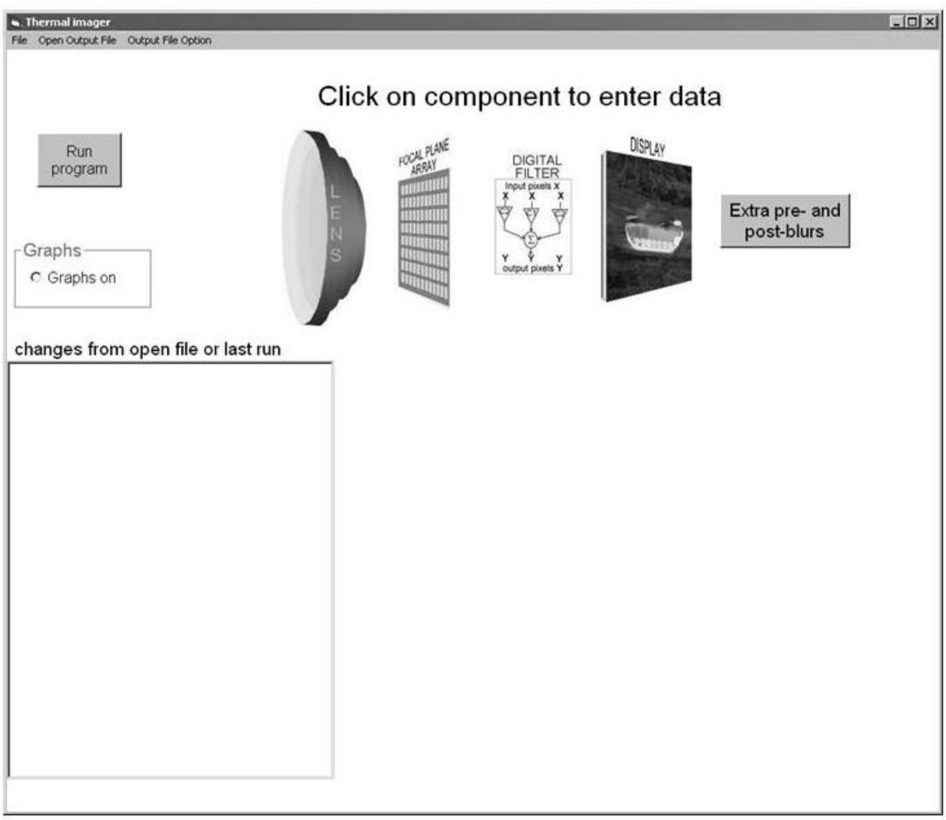


Figure 11.16 Screenshot of the main window of *Thermal*.

The program is run by clicking on the *Run Program* button. Completion of program execution is indicated in two ways. First, the output file is opened automatically. This is shown in Fig. 11.17. Second, the *Graphs* button is enabled. Automatic opening of the output file is disabled by clicking on the *Output File Option* menu and selecting *Open file manually using* "*Open Output File*."

The program starts automatically running the *Thermal.txt* input file. Other input files are loaded or saved by clicking on the *File* menu. The *File* menu also has an exit selection, or the "X" at top right also exits the *Thermal* program.

The main purpose of the output file is to display the parameters calculated in the program. Parameters such as field of view (FOV), system magnification, sample spacing, and detector instantaneous FOV are listed. A short list of selected input parameters is also provided. In addition, the output file provides tables of MTF values.

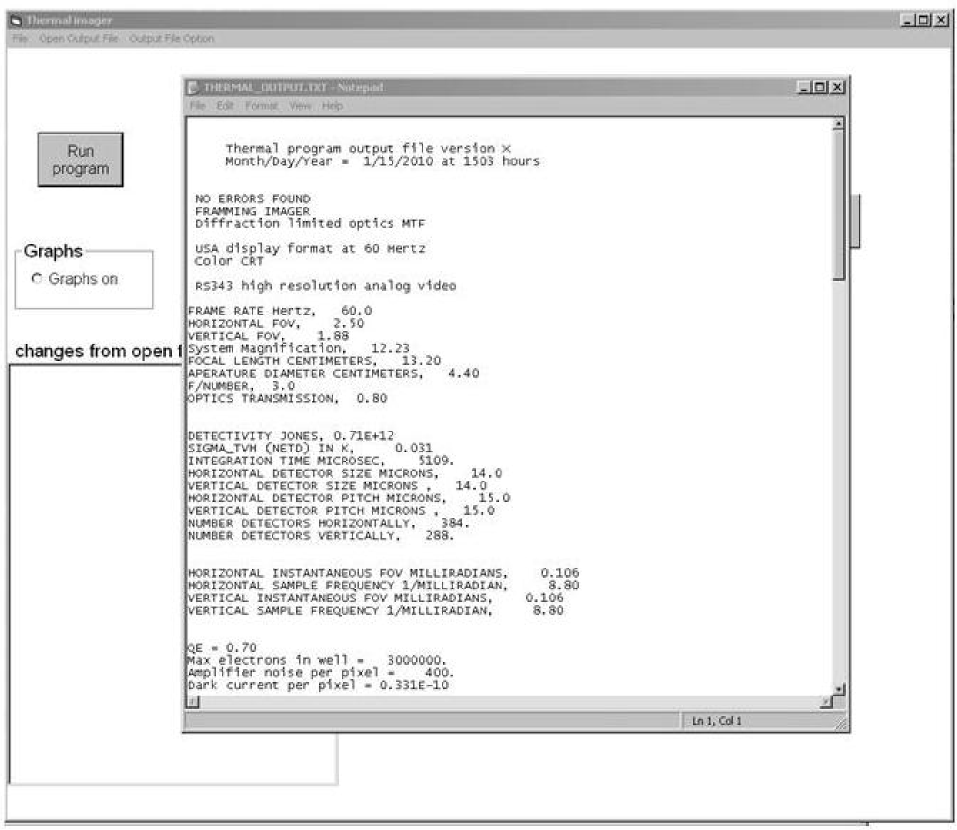


Figure 11.17 After the program executes, the output file is opened automatically and the *Graphs* option is enabled. Automatic opening of the output file is prevented by using the *Output File Option* menu.

Most program information is presented using the *Graphs* windows. Selecting *Graphs* after a run opens the window shown in Fig. 11.18. The pull-down box is used to select one of fourteen graph options:

1. Optics MTF,
2. Detector MTF,
3. Detector spectral response,
4. Digital & interpolate MTF,
5. Display MTF,
6. Horizontal pre, post, & system MTF,
7. Vertical pre-, post-, & system MTF,
8. System contrast threshold function,
9. Plot imager resolution,
10. Range plots,
11. Horizontal variable-gain MRT,
12. Vertical variable-gain MRT,
13. Horizontal aliasing, and
14. Vertical aliasing.

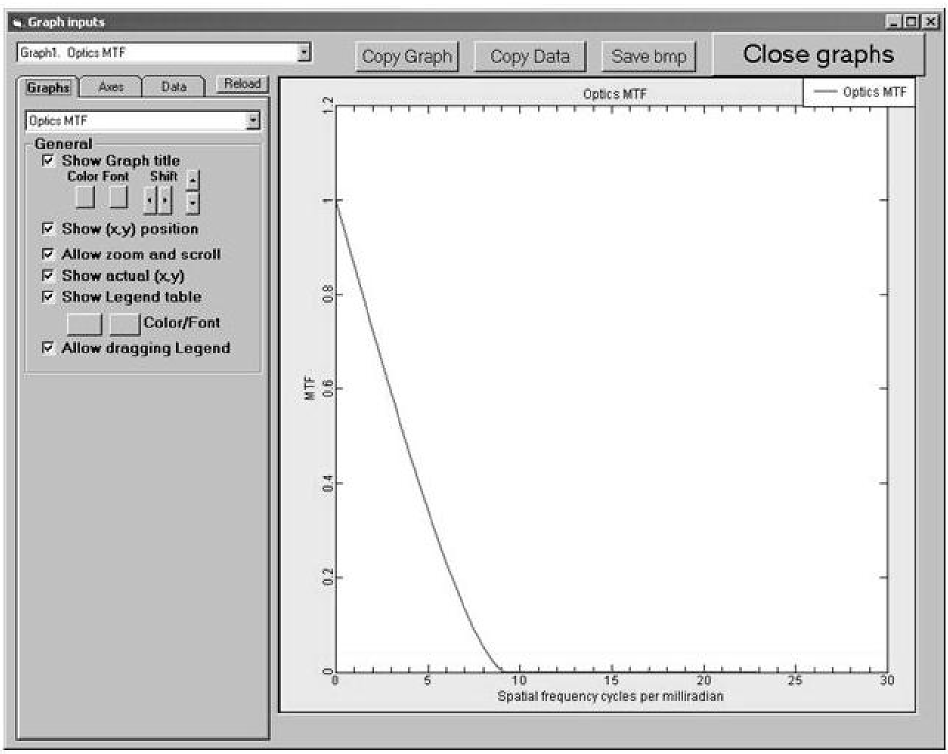


Figure 11.18 The *Graphs* window is the main source of information about the run. After selecting one of fourteen graphs, the user can copy both data and the graph using the Microsoft ClipBoard. The *Copy Graph* and *Copy Data* buttons are at the top of the window.

The model user creates output reports by selecting a graph and then clicking on *Copy Graph* or *Copy Data* and pasting the ClipBoard to Microsoft Excel or Word or another program that accepts the Microsoft ClipBoard.

Clicking on the *Lens* picture opens the input window shown at the left in Fig. 11.19. One of the options *ideal*, *good*, *typical*, or *input MTF array* is selected. The first three options enter MTF automatically. The input MTF array option makes the *Edit* button appear. Clicking on this button opens the window shown to the right in the figure. If an MTF array is entered, all frequencies must be monotonically increasing from top (1) to bottom. All MTFs must be between zero and one. The same MTF is applied horizontally and vertically.

The *OK* and *Cancel* buttons on the optics windows function the same as all other input windows. *OK* enters the data into the run file, and previous values are lost. *Cancel* returns all values to the state before the window was opened. The *Cancel* button on the main optics window (the window to the left in Fig. 11.19) also cancels new array inputs entered in the window at the right in Fig. 11.19.

Clicking on the *Focal Plane Array* picture opens the window shown in Fig. 11.20. Although this window looks complicated, the user can select an FPA from the drop-down lists. Selecting an FPA automatically enters all of the detector parameters. First an array type is selected from the following options:

InSb MWIR,
MCT MWIR,
MCT LWIR, and
QWIP
Uncooled.

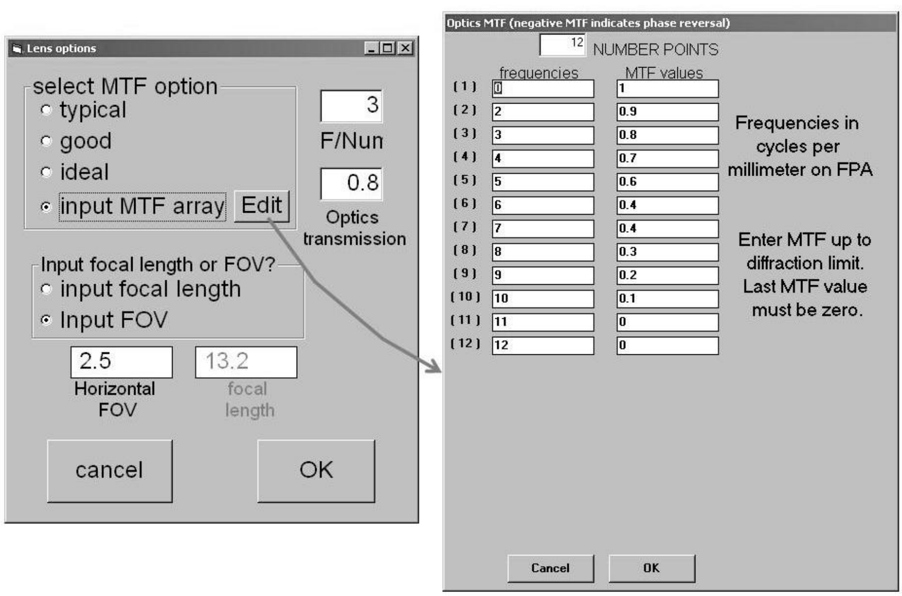


Figure 11.19 The main *Lens* window is on the left, and the array input window on the right. The user has an option of inputting either FOV or focal length (but not both). Whichever is input, the other is calculated by the program.

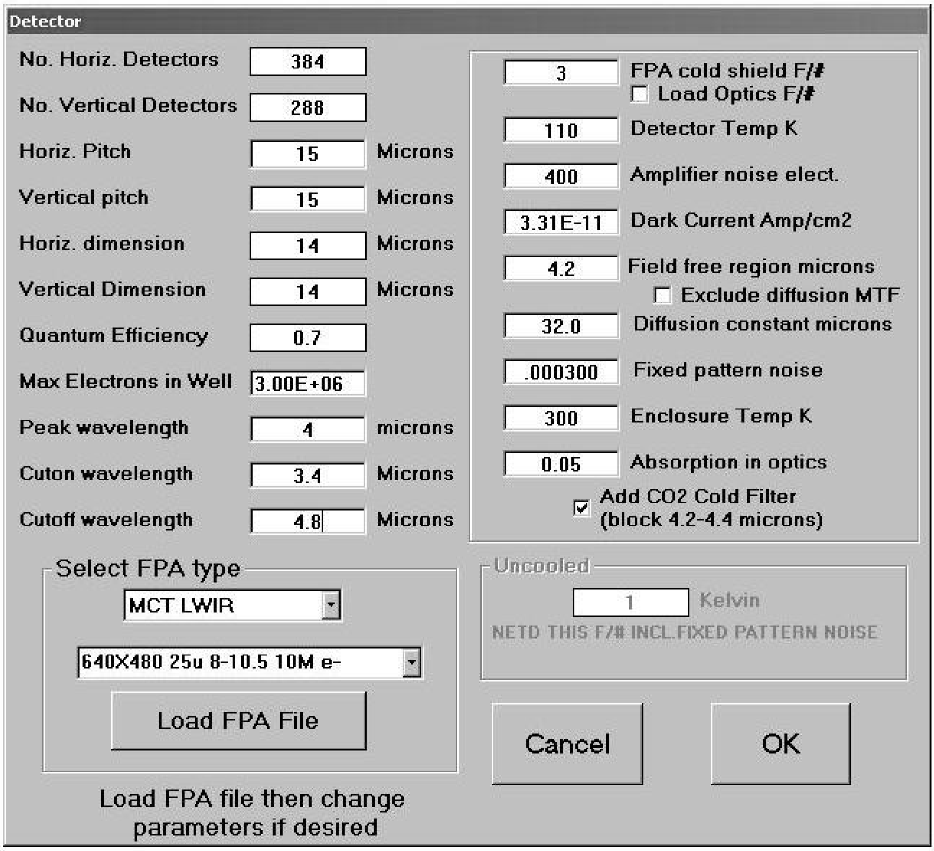


Figure 11.20 Screenshot of the *Detector* window. All input parameters can be input. Alternatively, the *Load FPA File* option is used. The user can also load the FPA file and then change any inputs before storing by clicking the *OK* button.

Then a specific array is selected from the second drop-down list. Clicking on the *Load FPA File* button loads the detector parameters. The user can accept all of the input detector parameters or change any or all of them. When the *OK* button is clicked, the program stores the values in these windows.

The *Uncooled* array type is different from the other detectors in that noise equivalent temperature difference (NETD) is entered. The measured NETD is for *f*/1 optics and optical transmission equal to one. In other words, NETD represents the array and not the current imager. Many of the input parameters disappear when *Uncooled* is selected. The *Thermal* program does not attempt to calculate the effect of environment on uncooled detectors.

The detector window also allows the user to select a carbon dioxide (CO_2) blocking filter. If selected, the program blocks 4.20- to 4.4-μm light from the detector. Since the atmosphere blocks light in this spectral band after only a few feet, allowing this flux on the detector increases noise without helping target signature. The blocking filter helps performance in the MWIR but does represent additional cost.

Another option in the detector window is to exclude diffusion MTF. The input arrays represent our estimate of diffusion parameters. However, this MTF is generally ignored for infrared detectors, and measurements are not generally available. The user might opt for optimism in the MTF predictions versus including an unsubstantiated MTF.

The final option in the detector window is a button to load the optics *f*/# into the cold shield *f*/#. Note that this switch is not stored. If the optics *f*/# is subsequently changed, the cold shield is not changed automatically. This option simply loads the current optics *f*/#.

Clicking on the *Display* picture opens the display window shown in Fig. 11.21. The user selects the type of display and type of interface. The interface selection determines the number of pixels on the display. The number of pixels associated with each interface format is given in Table 11.2 on page 218. The user also fills in viewing distance in centimeters and minimum and average display luminance in foot Lamberts.

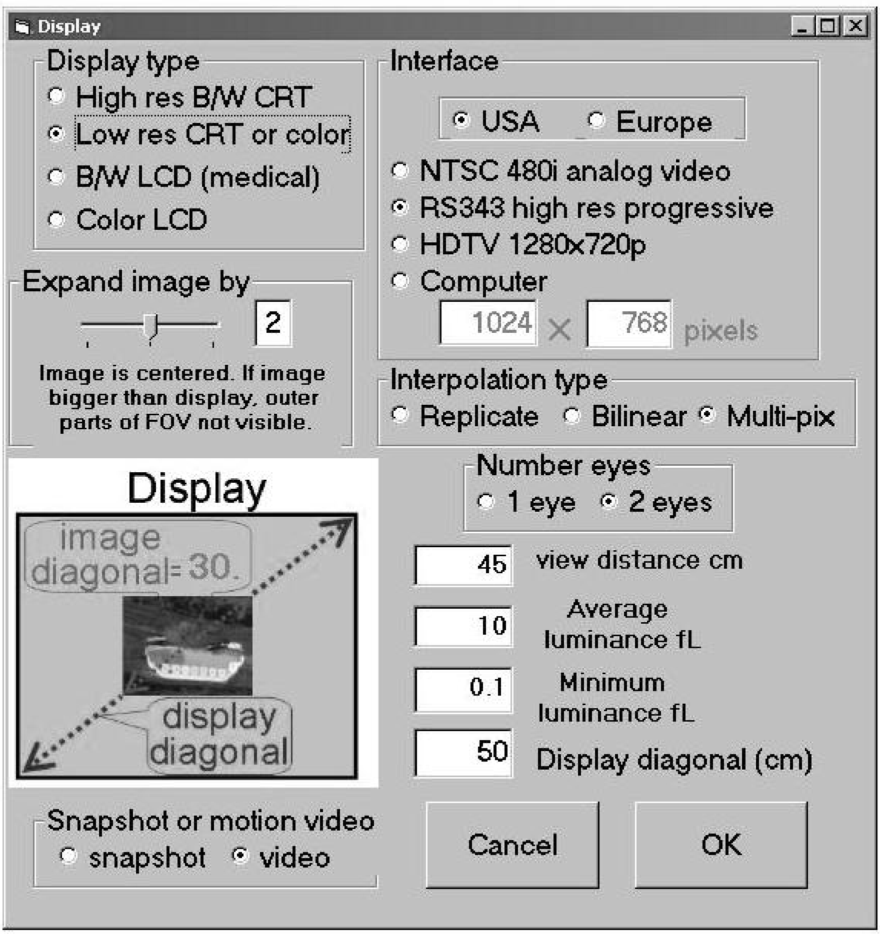


Figure 11.21 The *Display* window.

The display diagonal input is an important entry. This is the physical size of the display and not the sensor portion of the display area. The diagonal size is compared to the total pixels (determined by the interface format) to establish the size of the image. The following example illustrates how the various inputs interact.

In Fig. 11.21, a low-resolution display and RS343 analog video format are selected. Since display diagonal is 50 cm, the display height is 30 cm. A 3:4 aspect ratio is assumed if the display format is not HDTV. HDTV has a 9:16 aspect ratio. The CRT spot size is based on dividing 30 cm by 525 lines. Note that if a high-resolution CRT is selected, then the spot size would be based on 30 cm divided by 960 lines. Spot size is established by the display, not by the interface.

The FPA has 384 × 388 detectors. Since the display has 1280 × 960 pixels, the image diagonal without interpolation is 15 cm. However, as shown in the figure, the image is expanded by two. Therefore, the image diagonal is 30 cm. This value is shown superimposed on the display picture. Image diagonal is recalculated and updated each time the window input parameters are changed. If image size exceeds display size, then only the center portion of the image is displayed.

If the image is expanded by two or three, then the MTF associated with the selected interpolation technique is applied. Interpolation MTF is described in Chapter 4. *Multi-pix* interpolation uses the *6 samples* coefficients from Table 4.1 on page 91.

Clicking on the *Digital Filter* picture opens the window to the left in Fig. 11.22. If the answer to the question "Apply digital filter?" is "no," then no MTF is applied. If the answer is "yes," then the user indicates whether to apply the filter to FPA data or to interpolated data. If interpolation is not used (expand image by 1 in the display window), the FPA-versus-interpolate option makes no difference. To enter digital filter coefficients, the *Edit coefficients* button is clicked. This opens the window shown to the right in Fig. 11.22.

In the coefficient window, the user enters the number of coefficients for horizontal and vertical. Five coefficients are selected in the figure, but up to 20 can be entered. The user also selects whether the filters are *odd* or *even*. Figure 11.23 illustrates how the coefficients are applied for odd and even filters. Figure 11.22 shows *odd* selected for both filters. In Fig. 11.23, the horizontal filter is odd and the vertical filter is even.

Clicking on the *Extra Pre-and Post-MTF* button allows the user to input additional pre- and post-MTFs. Pre-MTFs are prenoise and presample blurs. Post-MTFs are postnoise and postsample blurs. The main window is shown to the left in Fig. 11.24. This window allows the user to independently select additional horizontal or vertical pre-MTFs or additional horizontal or vertical post-MTFs. The MTF is applied only if *Yes* is selected, regardless of the values in the associated array.

To input array values, click on *Edit*. This opens a window similar to the one at the right in Fig. 11.24. The figure shows twenty array inputs, but as few as two can be selected. All frequencies are in object space. The frequencies must be monotonically increasing from top (1) to bottom. Pre-MTFs must be between zero and one (0 ≤ MTF ≤ 1). Pre-MTFs cannot apply gain. The array should span the frequency range to the diffraction limit. Regardless of the frequency span, however, the array must be extended such that the last MTF value is zero.

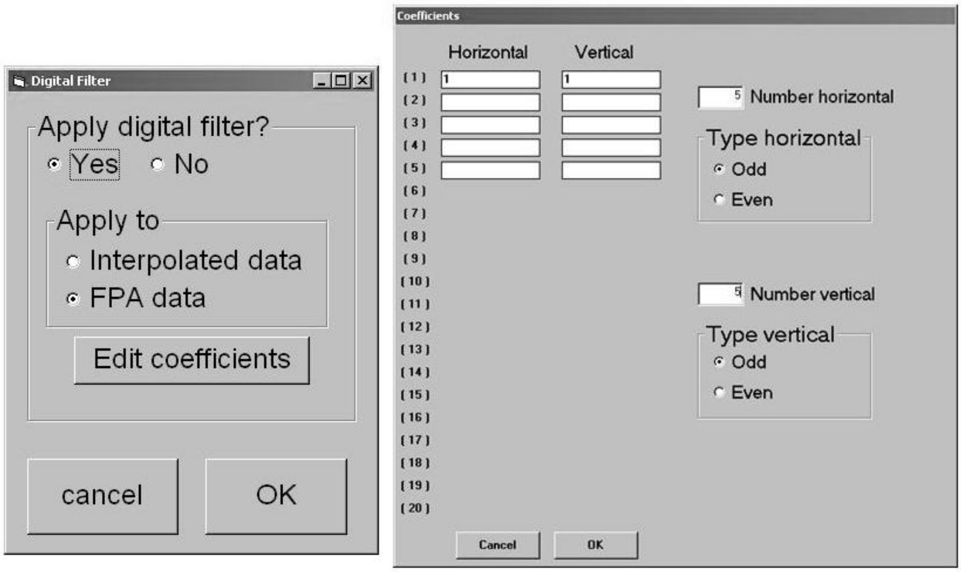


Figure 11.22 The main digital filter window is shown at the left. The window to enter digital coefficients is shown at the right. Up to 20 coefficients can be entered. Different coefficients can be applied horizontally and vertically.

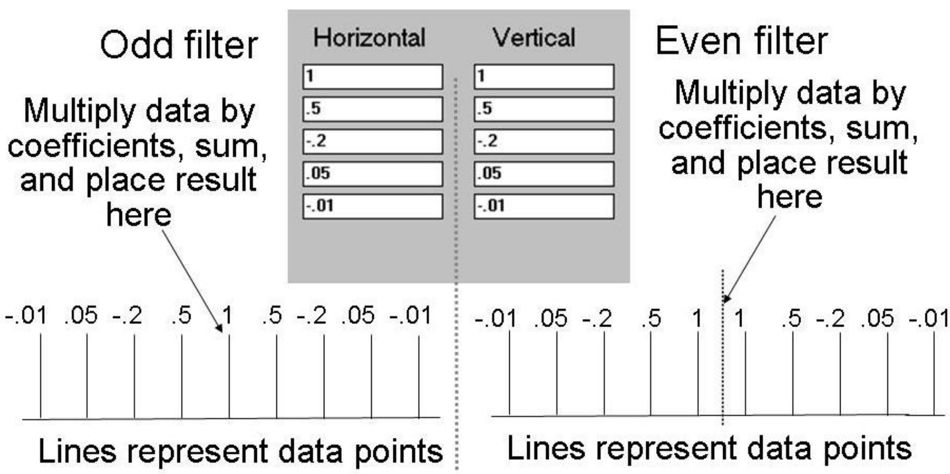


Figure 11.23 The meaning of *odd* and *even* filters is illustrated in this figure. For odd filters, the first coefficient is applied to the center data point. Other coefficients are applied symmetrically to either side. For even filters, a new data point is found by applying the coefficients symmetrically about a point between two original data points.

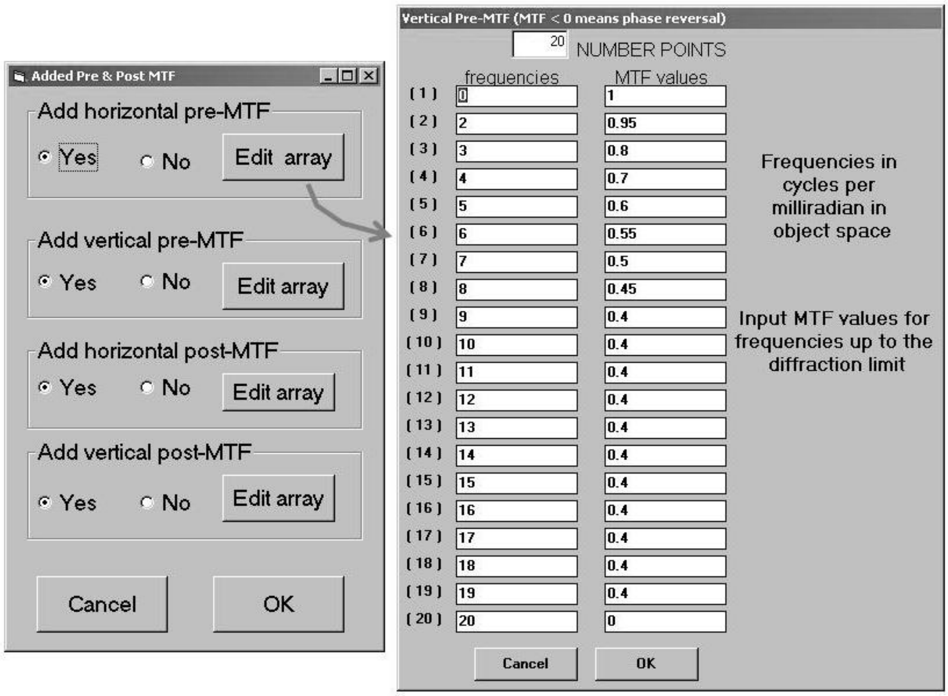


Figure 11.24 This window is used to insert additional MTFs. The MTFs are added only if *Yes* is clicked, regardless of the values in the associated array.

Post-MTF takes any value. Gain can be applied with post-MTF. Again, however, the last MTF value in the input array must be zero. This is because the program cannot extrapolate MTF; the frequency at which MTF goes to zero must be supplied.

11.6 Imager Analysis Using the Programs

Three of the *Graphs* windows are particularly useful for understanding imager behavior. The graphs are selected by clicking on *Graphs* in the main window and using the pull-down menu. The three windows are discussed in the following three sections.

11.6.1 Imager resolution

Figure 11.25 shows the *Plot imager resolution* window. The abscissa is target contrast T_{ebb} in effective blackbody Kelvin. The ordinate is cycles per milliradian resolution (the TTP metric). Imager gain is automatically adjusted as target contrast varies in order to maintain a 0.25 modulation contrast on the display. This contrast would result in saturating some pixels in a complex scene. This contrast setting is used in order to minimize the effect of the display.

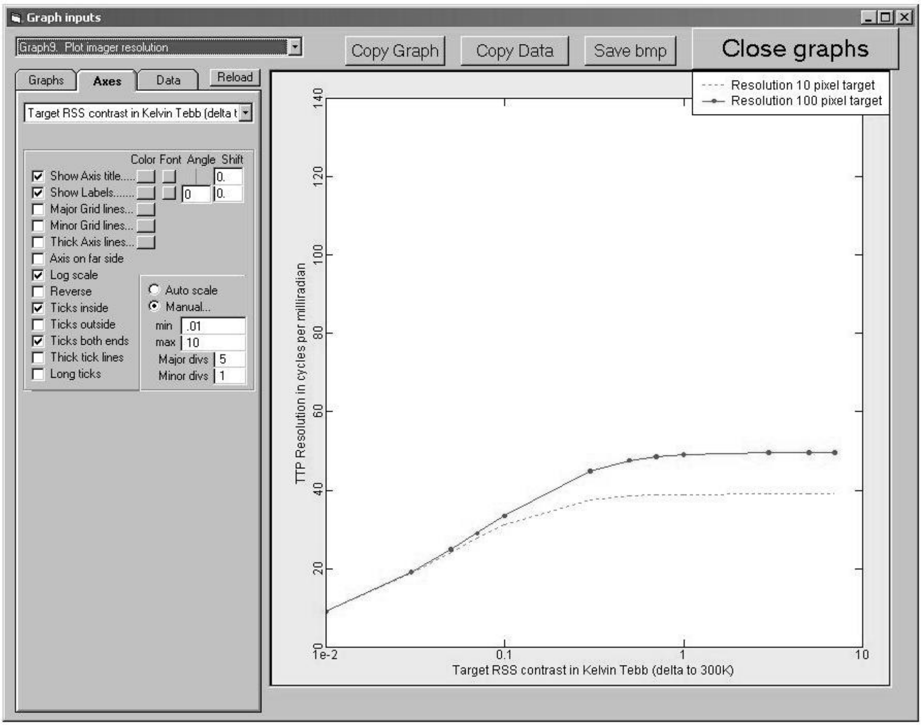


Figure 11.25 Screenshot of the *Plot imager resolution* window.

One curve plots the resolution of a 10-pixel target, and the second curve is for a 100-pixel target. Targets subtending many pixels are not affected by any aliasing in the imager. Targets that subtend only a few pixels are badly corrupted if aliasing is present. The separation between these curves indicates whether and how much aliasing is present.

Good resolution with high target contrast indicates good range performance against hot targets in good weather. The resolution at temperatures near 0.1 K indicates whether the imager performs in cold dreary weather.

11.6.2 System contrast threshold function

The window for *System Contrast Threshold Function* is shown in Fig. 11.26. The abscissa is spatial frequency in object space cycles per milliradian. The ordinate is contrast threshold. Horizontal and vertical are plotted separately. This plot is useful because it shows imager response versus spatial frequency.

The controls at the lower left allow the user to vary target size and contrast. As target contrast is varied, imager gain is adjusted to maintain 0.25 modulation contrast on the display. Decreasing target contrast increases noise because gain increases. Increased noise raises the entire threshold curve. Varying target size allows the user to see the effect of aliasing. When the target size is decreased, threshold increases at frequencies where aliased signal is present. Target contrast and size are noted on the graph and in the copy data.

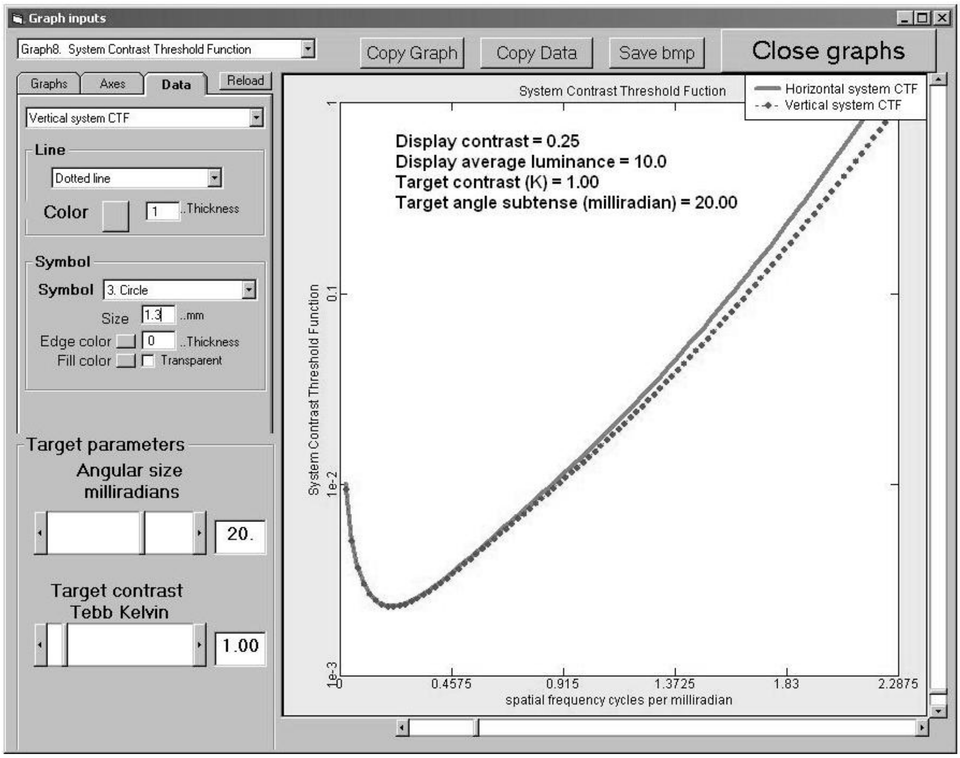


Figure 11.26 Screenshot of the *System Contrast Threshold Function* window.

11.6.3 Range plots

Figure 11.27 shows a PID-versus-range plot. The user adjusts target, background, and atmosphere input parameters using the controls at the lower left. Clicking on the *Calculate PID vs. range* button generates a range plot. An expanded view of the input controls is shown in Fig. 11.28. The notes in the figure explain the function of each control.

Target contrast can be in T_{ebb} or T_{pyro}. The delta temperature is centered at 300 K (T_{ebb}) or at the background temperature (T_{pyro}). These concepts are explained in Chapter 9. When the background is cold, the same value of T_{pyro}. has less energy than T_{ebb}. When the background is hotter than 300 K, T_{pyro} has more energy. Also, the spectral content of the two temperature types is different.

The ground temperature is automatically set when an atmosphere is selected. The ground temperature is set to the air temperature. The user can change the temperature to any value between 240 and 340 K.

Display contrast is given by Eq. (11.20). Imager gain is given by Eq. (11.21). S_{tmp} is scene contrast temperature in Kelvin. S_{tmp} is the scene temperature contrast that raises the display average luminance from black to average. S_{tmp} is not background temperature. The model assumes that the imager has sufficient gain to fill the display with detector noise.

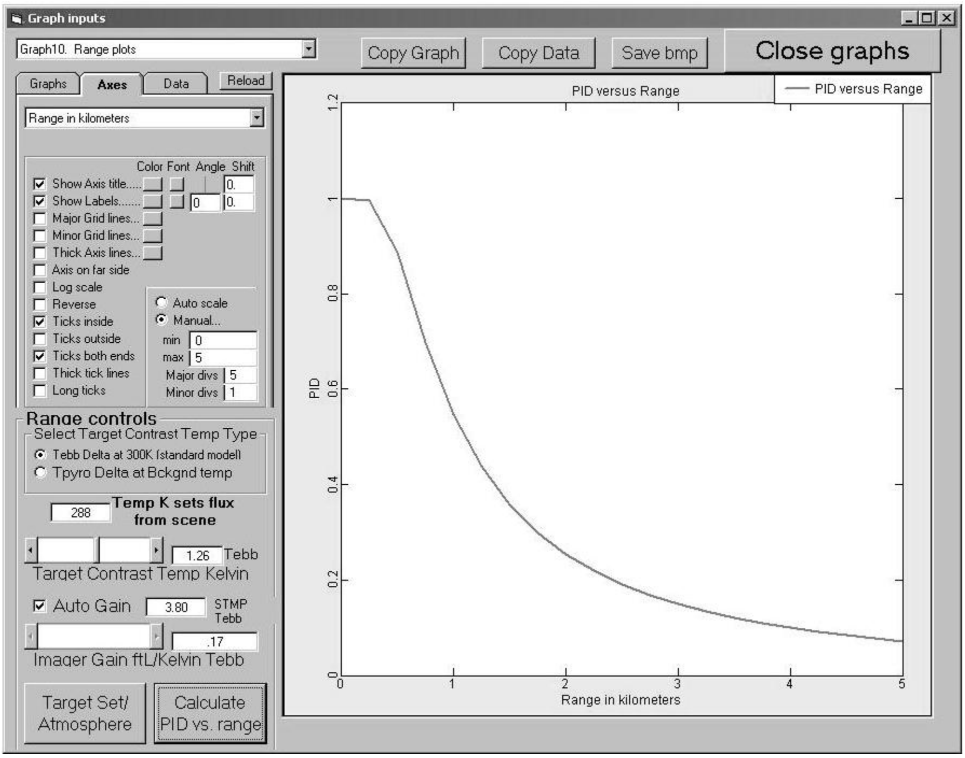


Figure 11.27 Input parameters are adjusted using the controls at the lower left. The range plot is generated by clicking on the *Calculate PID vs. range* button.

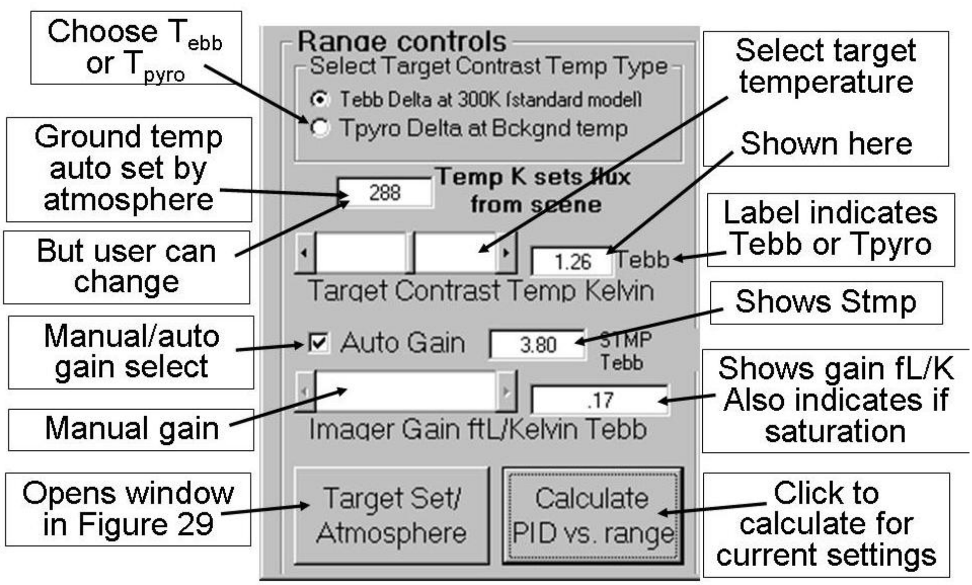


Figure 11.28 Screenshot showing the range controls that appear in the *Graph inputs* window when *Range plots* is selected. The added notes explain what each control does.

$$\text{display contrast} = \frac{\text{target contrast in Kelvin}}{2\,S_{tmp}\ \text{in Kelvin}}. \tag{11.20}$$

$$\text{imager gain} = \frac{\text{display average luminance in fL}}{S_{tmp}\ \text{in Kelvin}}. \tag{11.21}$$

Gain is automatically adjusted to keep S_{tmp} equal to a multiple of three of the target contrast. This maintains displays contrast at 0.17. If *Auto Gain* is turned off, gain is controlled manually by the horizontal slider. If gain is increased such that target contrast is greater than one-third S_{tmp}, the warning "clipping" is displayed in the gain window. If target contrast exceeds S_{tmp}, then the warning "saturation" is displayed. Note that range calculation proceeds, regardless of these warnings.

The 0.17 display contrast setting assumes a complex scene. This setting provides good contrast while avoiding saturation. For cold-soaked scenes, gain can be increased. However, a gain setting in which target contrast exceeds S_{tmp} is physically unrealistic.

When the button *Target Set/Atmosphere* is clicked, the window in Fig. 11.29 opens. This window is used to input target size and task difficulty. Atmosphere and turbulence are also selected using this window.

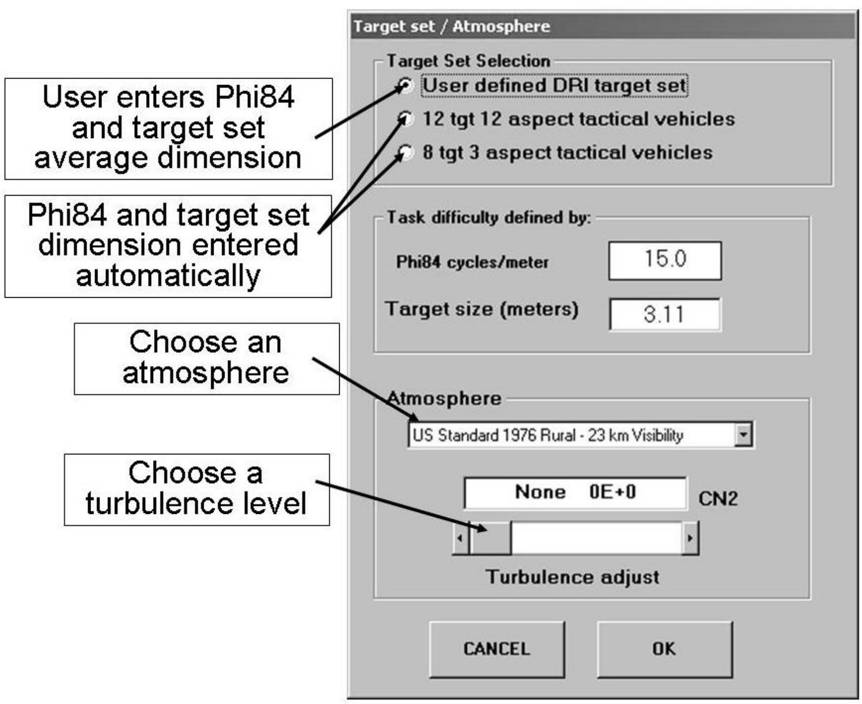


Figure 11.29 Target set size and task difficulty are input in the *Target set / Atmosphere* window. This window is also used to select atmospheres.

When all of the parameters are selected, click on the *Calculate PID vs. range* button located at the bottom left of the *Graphs* window. (See Fig. 11.27). The program calculates PID versus range and plots the results.

References

1. R. E. Fischer and B. Tadic-Galeb, *Optical System Design*, McGraw-Hill, New York (2000).
2. W. J. Smith, *Modern Lens Design*, McGraw-Hill, New York (2005).
3. L. E. Tannas, Jr., *Flat-Panel Displays and CRTs*, p. 175, Van Nostrand Reinhold Co., New York (1985).
4. R. G. Driggers, P. Cox, and T. Edwards, *Introduction to Infrared and Electro-Optical Systems*, p. 257, Artech House, Boston (1999).
5. R. L. Myers, *Display Interfaces, Wiley Series in Display Technology*, p. 132, John Wiley & Sons, Chichester, UK (2002).
6. W. den Boer, *Active Matrix Liquid Crystal Displays*, p. 143, Elsevier, Burlington, MA (2005).
7. G. Waldman and J. Wootten, *Electro-Optical Systems Performance Modeling*, p. 98, Artech House, Boston (1993).
8. K. Buskila, S. Towito, E. Shmuel, R. Levi, N. Kopeika, K. Krapels, R. G Driggers, R. H. Vollmerhausen, and C. E. Halford, "Atmospheric modulation transfer function in the infrared," *Appl. Opt*. **2**, 471–82, 14735966 (2004).
9. N. S. Kopeika, *A System Engineering Approach to Imaging*, p. 463, SPIE Press, Bellingham, WA (1998).
10. W. L. Wolfe and G. J. Zissis, Eds., *The Infrared Handbook*, The Infrared Information Analysis Center, ERIM, Ann Arbor, MI (1993).
11. E. Friedman and J. L. Miller, *Photonics Rules of Thumb: Optics, Electro-Optics, Fiber Optics, and Lasers*, p. 97, McGraw-Hill, New York (2004).
12. J. R. Janesick, *Scientific Charge-Coupled Devices*, SPIE Press, Bellingham, WA (2001).

Bibliography

Caniou. J., *Passive Infrared Detection: Theory and Applications*, Kluwer Academic Publishers, Boston (1999).

Laikin, M., *Lens Design*, CRC Press, New York (2007).

Poynton, C. *Digital Video and HDTV*, Morgan Kaufmann Publishers, Elsevier, New York (2003).

Yeh, P. and C. Gu, *Optics of Liquid Crystal Displays*, Wiley, New York (2010).

Chapter 12
Infrared Focal Plane Arrays

The purpose of this chapter is to provide an understanding of infrared (IR) detector focal plane arrays (FPAs) and to describe the important parameters that contribute to overall system performance. The emphasis here is on building an intuitive view of how IR FPAs work and less on a detailed treatment of the semiconductor physics of the detector materials. There are many excellent books that treat detector physics and materials in much greater depth. A partial list used in preparing this chapter includes Singh,[1] Henini and Razeghi,[2] and Kinch.[3] Section 12.1 is an overview of the basic principles of photon detector/FPA operation. Section 12.2 describes the development of the background-limited photoconductor (BLIP) and flux-based expressions for signal-to-noise ratio. Section 12.3 summarizes the capabilities, limitations, and typical performance parameters for cooled FPAs. Section 12.4 discusses uncooled (bolometer) FPAs.

12.1 Photon Detector Infrared Focal Plane Arrays

12.1.1 Photon detector basic principles

Figure 12.1 shows an abstracted cross-sectional view of a few pixels of a photon detector IR FPA. This particular detector structure is a double-layer hetero-junction (DLHJ) photovoltaic (PV) detector with an n-type absorber region. There are many other detector structures possible, but their general properties are similar. The DLHJ is a widely used structure for HgCdTe FPAs; it is of considerable practical importance and is easy to understand in cross section. The end-to-end signal chain involves a number of steps:

(1) IR radiation (i.e. light) impinges from the top.
(2) The first layer is an antireflection coating that reduces first surface reflection from 20–30% to 2–3%.
(3) The light passes through a transparent but electrically inactive substrate. This detector is “backside” illuminated. Some detectors have no substrate or have the substrate removed.
(4) Photons of sufficient energy to overcome the detector’s bandgap are absorbed in the n-type absorber region generating electron–hole pairs. Holes pass through the semiconductor p–n junction into the p-type

region. This particular detector has been partially delineated into individual cells by an etching process.

(5) A nonconducting coating is applied to the side walls and ends of the detectors to reduce surface noise currents from dangling bonds and other defects (passivation).

(6) A soft indium bump is used to electrically and mechanically couple the detector to the silicon readout integrated circuit (ROIC). Mechanical force is used to "weld" the soft indium to the detector and ROIC. Interspersed epoxy is sometimes used to assist with mechanical bonding.

(7) The silicon ROIC has a small circuit per each pixel (unit cell) to integrate hole photocurrent.

(8) Electronics on the ROIC amplify and multiplex the pixel values and put them on one or more high-speed (megahertz) analog outputs. Additional electronics provide bias voltage to the detectors and clocking signals to set the frame and integration times.

Figure 12.1 underscores the complex interplay of the optical, quantum, electrical, and mechanical elements of an IR FPA. Most IR detectors are cooled to reduce noise, so the additional burdens of operating at cryogenic temperatures and the ability to repeatedly cycle between room and cryogenic temperatures while maintaining specification performance must also be accommodated. While there are many conceivable combinations of detector materials, detector geometries, and ROIC designs, the economic reality is that manufacturers have limited sets of "off the shelf" designs to sell at competitive prices. Unless the buyer is willing to invest significant sums of his own money, it is likely that he

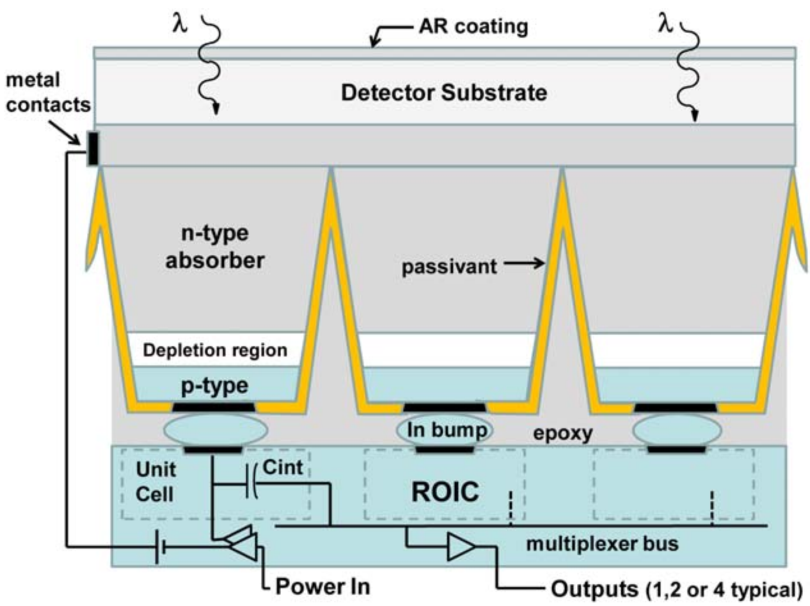


Figure 12.1 Schematic cross-sectional view of an IR FPA. Not all components of the ROIC are shown. Double-layer heterojunction architecture is illustrated. (Not to scale.)

will have to chose from among these different offerings. The good news is that new IR FPAs do come to market on a regular basis as the state-of-the-art in the underlying materials science and device engineering improves. One objective of this chapter is to enable prospective systems engineers to make informed choices about the different available IR FPAs and to understand how the limitations of each individual technology may affect the choice for a particular application.

Figure 12.2 shows how signal is generated in a DLHJ detector at the quantum level. Again, many PV detectors work on similar principles. Generally, the thickness of the n-type layer is engineered so that most or all of the incoming light of the desired frequency is absorbed in either the n-type or p–n junction regions at the specified operating temperature of the detector. The goal is to make the n layer thick enough to absorb most of the incoming light but not too thick as to increase thermal generation of carriers (diffusion dark current). Photons of energy ($E = hc/\lambda$) that exceed the bandgap energy in the n layer generate electron–hole pairs. The photogenerated electrons are in the conduction band, while the holes are in the valance band. An external bias voltage holds the detector (electrically a diode) in slight reverse bias. Holes generated in the n layer within a diffusion length of the junction migrate toward the junction, are swept through it by the internal electric field, then pass through the p layer to the metal contact.

The net result is a flow of photocurrent (from the holes) and an accumulation of electric charge on the integration capacitor, which is in series with the detector and bias voltage source. The flow of signal (in this case the minority carrier holes) to the accumulating capacitor is opposite to the normal electrical path for a forward-biased diode. In the DLHJ the bandgap of the p-type layer is greater than the n-type layer, which helps to substantially reduce the dark current from thermally generated electron–hole pairs in the p-type layer. Generally, the p-type layer is much thinner (or smaller in some detector geometries) than the n-type layer, which also helps to minimize dark current.

Figure 12.3 shows a set of typical detector I–V characteristic curves. Output current is plotted against bias voltage for a detector operated at different temperatures. One important figure of merit that is often quoted is the detector dynamic impedance at zero bias under no illumination, R_0A, given by[4]

$$R_0 A = \left. \frac{dJ_{detector}}{dV} \right|_{V=0} , \tag{12.1}$$

where $J_{detector}$ is the current density (A/cm^2).

Under diffusion-limited conditions, R_0A defines the saturation dark current according to Ref. 5 (see Section 12.1.3 for more detail).

$$J_{diff} = \frac{kT}{e^{-} R_0 A} . \tag{12.2}$$

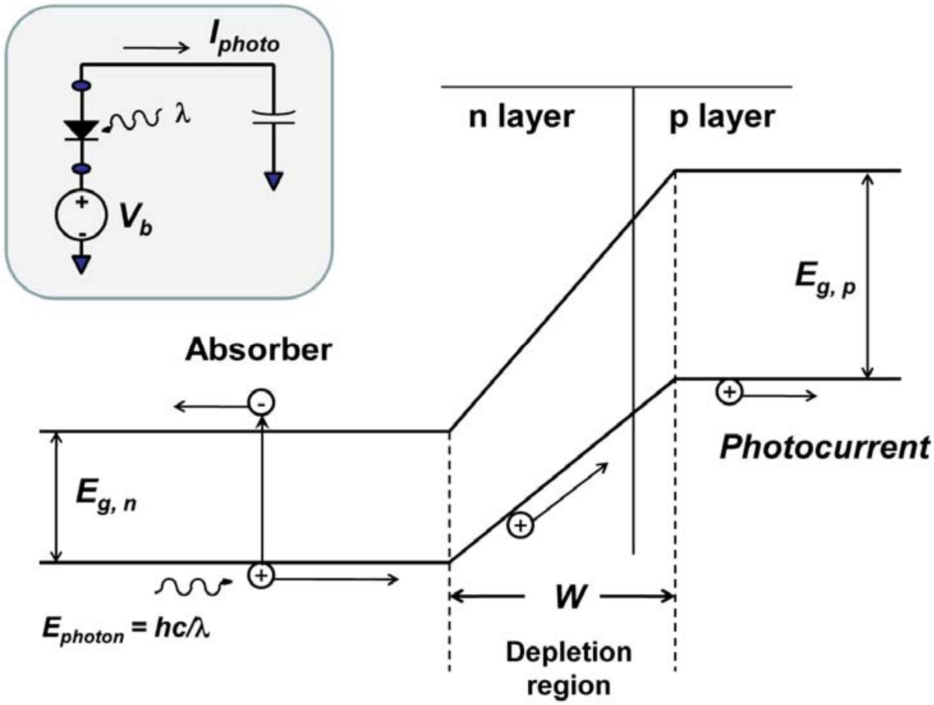


Figure 12.2 Band structure and photogeneration process in an n-on-p DLHJ architecture. Light is absorbed in the n layer. Minority carriers (holes) are responsible for the photocurrent. Note that most of the depletion width is in the n layer and that the heavily doped p layer is much thinner than the n layer. The opposite architecture (p-on-n) is also used for some detectors.

Higher R_0A is preferable because it implies a lower dark current. To generate external photocurrent, PV detectors are normally operated with a small reverse bias (from ~100 mV to ~1 V, depending on the particular detector material and design), which improves the dynamic resistance but also increases the noise current from internal thermal noise processes, as shown in Fig. 12.3. QWIP detectors are photoconductors that are typically operated at a bias of ~ 1–2 V at which the photocurrent is maximized from the multilayer quantum well structure.[2] As discussed in the next section, cooling is the primary method for reducing overall dark current density.

Figure 12.4 shows an idealized I–V curve that includes photocurrent and dark current combined. A slight dark current may exist at zero bias due to surface effects. As can be seen, the normal operating point includes a combination of photocurrent and thermal current. The design goal is to maximize the ratio of photocurrent to dark current. When the detector material has been optimized to improve signal and QE, and to reduce internal noise sources, the only option remaining for better performance is to further cool the detector.

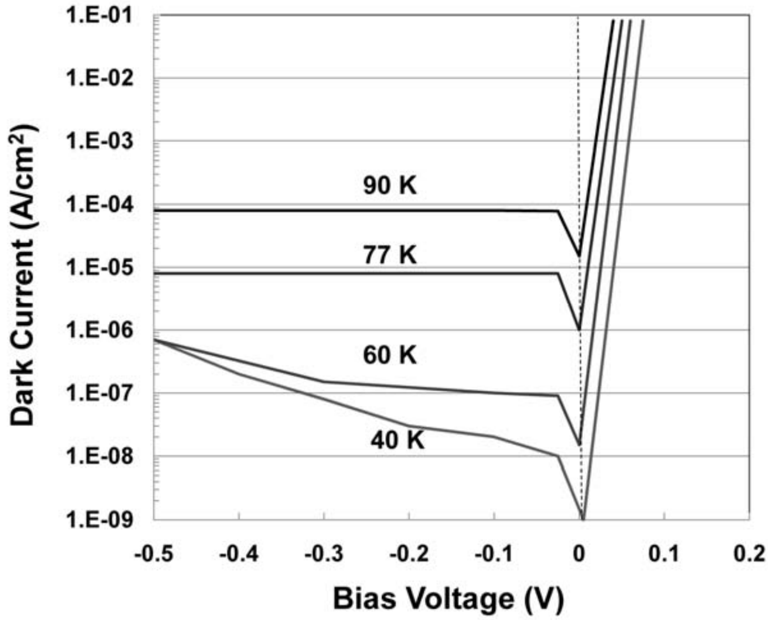


Figure 12.3 Photovoltaic detector I–V curves as a function of bias and detector temperature under no external illumination.[6] The plot shows the total dark current from a HgCdTe detector with a 12-μm cut-off wavelength.

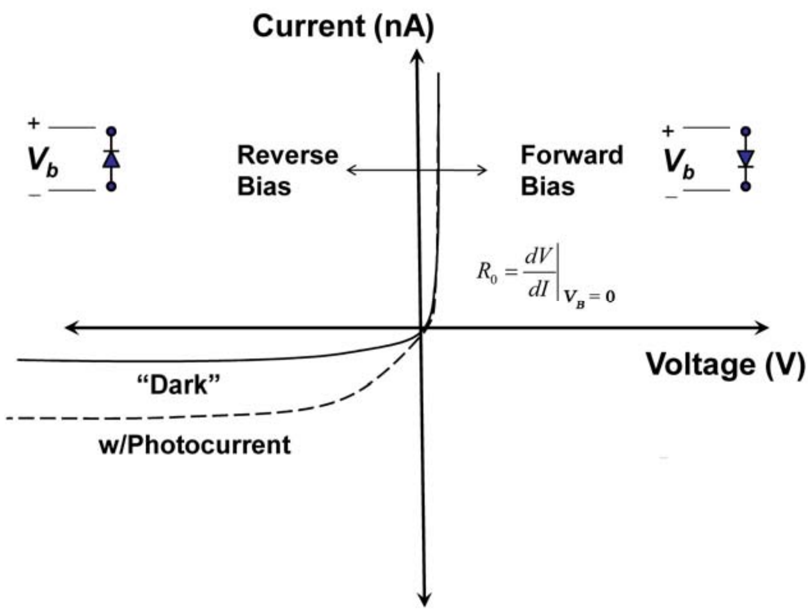


Figure 12.4 Idealized detector I–V curves with and without photocurrent. The solid curve shows dark current only. The dashed curve shows a combination of dark current and photocurrent. Typical operating biases for direct-bandgap detectors such as InSb and HgCdTe detectors are 0.1–0.2-V reverse bias.

12.1.2 Readout integrated circuit

The readout integrated circuit (ROIC) exists to collect the photocurrent from the detector, amplify it to practical levels, and multiplex/format it for transport off the cryogenic space into more conventional signal processors that form the video image. The circuitry associated with each pixel is referred to as a unit cell. Figure 12.5 shows a common unit cell architecture: the direct-injection (DI) circuit. There are at least five common architectures summarized in Refs. 7 and 8. The DI circuit has the advantage of requiring very few components, which allows for maximum possible real estate within the unit cell for integration capacitor, or capacitors if the detector is multiband. DI circuits are often used for high background flux applications, especially in small pixel sizes (25 μm or less). For very low background or larger pixel sizes, buffered direct-injection (BDI) or capacitive transimpedance amplifier (CTIA) circuits offer better bias control and lower cell noise at the expense of lower charge-handling capacity due to decreased unit cell capacitance[7].

Figure 12.5 also shows the representative small signal model for a detector: a current source consisting of photocurrent and dark current, an internal junction capacitance C_d, and a shunt resistance R_d due to surface leakage currents. The MOSFET M_{DI} is used to establish the bias on the detector and to control the timing of photocurrent flow into the unit cell integration capacitor C_{int}. M_{reset} is used to reset the voltage/charge on C_{int} at the start of the integration cycle. For a minority carrier detector, the flow of holes, designated $I_{detector}$, decreases the charge of the capacitor during the integration cycle. M_{sel} is used to select each unit cell for transfer to the multiplexer bus. Additional amplification is applied before the signal is output from the FPA. The injection efficiency is controlled by the ratio of the detector shunt resistance to the MOSFET input resistance. A higher detector shunt resistance and lower MOSFET input resistance maximize photocurrent injected into C_{int}. Primary sources of noise are kTC reset noise and $1/f$ noise in the different MOSFET switches.

The DI circuit is also susceptible to small differences in the turn-on voltage of the DI MOSFET gate altering the detector bias point. Variations in both the MOSFET threshold and individual detector I–V curves lead to fixed pattern noise within the FPA. The bias across the detector V_b is given by Ref 9.

$$V_{det} = V_b - V_{GS} - V_{DI} . \tag{12.3}$$

In the MOSFET subthreshold region, small variations in the threshold voltage affect V_{GS}, altering the detector's bias and thus the output photocurrent for a fixed source current, generally, according to

$$\Delta I_{detector} = \left[I_0 \left(\frac{q}{kT} \right) e^{(qV_b/kT)} + \frac{1}{R_d} \right] \Delta V_b , \tag{12.4}$$

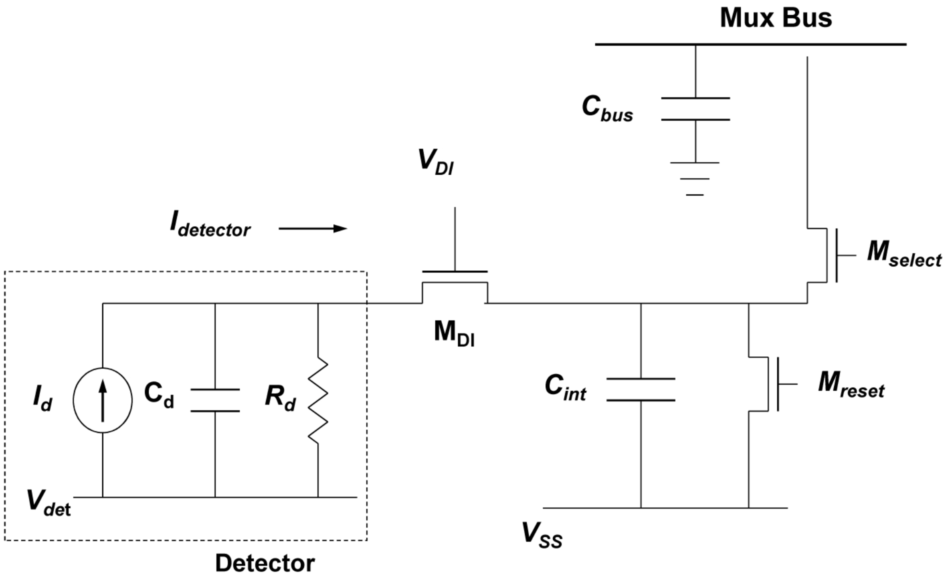


Figure 12.5 Direct-injection ROIC unit cell architecture showing detector small signal model, integration, and pixel select functions.

where I_0 is the detector reverse-bias saturation current. Figure 12.6 illustrates the relationship between the detector I–V and DI MOSFET I–V characteristics. The operating point is the intersection of the two load lines.

As an example for a typical high-background design, a C_{int} of ~ 0.5 pF at 77 K provides a charge storage of ~ 1×10^7 e^- at a typical 3-V ROIC supply voltage.[11] For high-background applications. the overall noise is dominated by background (shot) noise, detector dark current, and nonuniformity. ROIC noise plays only a small part in the overall S/N (signal-to-noise ratio). ROIC noise can be a dominant term for low-background/low-signal applications. Typical ROIC noise for high-background (terrestrial) applications is 500–1000 e^- for integration capacities on the order of ~ 2×10^6 to ~ 2×10^7 e^-.

12.1.3 Photon detector dark current

Detector dark current represents current produced by the detector in the absence of external flux, literally the detector is "in the dark." The sources of dark current are thermally related, and there is a strong thermal dependence in all forms. Fluctuations in dark current account for one source of detector noise, other important contributions being fluctuations in signal current and ROIC noise. Dark currents have been extensively studied in common photon detector materials; there are diode models that enable dark currents to be predicted, depending on material properties, applied bias, and temperature. A comprehensive treatment of PV detectors is provided by Kinch,[3] briefly summarized herein.

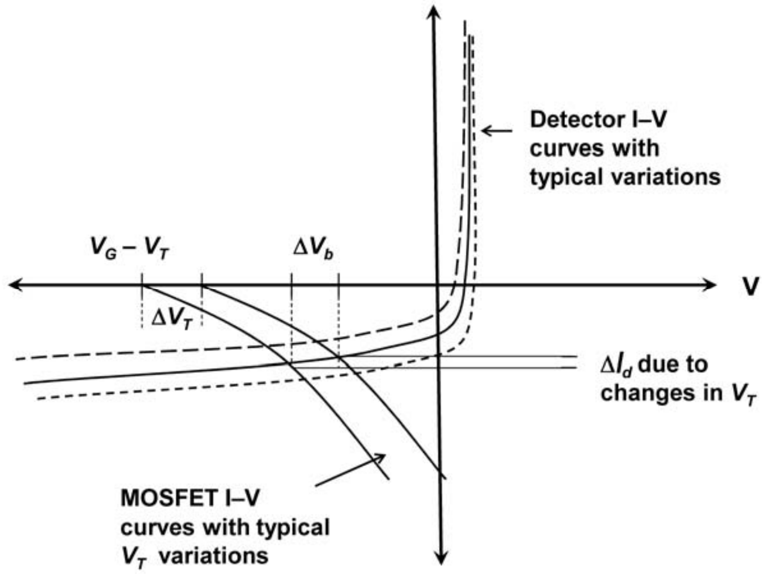


Figure 12.6 Bias determination of a PV detector operated in DI mode. Variations in individual detector I–V characteristics and DI MOSFET threshold voltage affect V_b and I_d, leading to fixed pattern (spatial) noise (after Longo et al.[10]).

The purpose of this section is to provide insight into the major sources of dark current and to illustrate the material and temperature dependencies. Quantum well infrared photodetectors (QWIPs) have similar dark current limitations, but some of the corresponding physical mechanisms and governing equations are different. The primary sources of dark currents are:

- thermally generated diffusion current from both sides of the junction,
- thermally generated current from bandgap states in the depletion region (G–R),
- generation from the diode surfaces,
- tunneling in the depletion region via bandgap states, and
- direct band-to-band tunneling in the depletion region.

For well-passivated detectors at low biases, the first two mechanisms dominate.

The generation and recombination of holes and electrons is responsible for dark current. Three physical mechanisms are most important for typical IR detectors, as shown in Fig. 12.7. *Radiative* (or band-to-band) recombination involves the pairing of a conduction band electron with a valence band hole, producing a photon to conserve energy. *Auger* recombination of electron–hole pairs is enabled by the transfer of the resulting energy to a third particle, shown as a second conduction band electron in Fig. 12.7. *Shockley-Read-Hall* (SRH) recombination involves the existence of a localized state (centers) within the bandgap due to an isolated vacancy, impurity atom, or physical defect. The "trap" state can capture both a conduction electron and a valence hole with a photon being emitted to conserve energy. Each of these processes is reversible, leading to electron–hole pair generation. Under nonequilibrium conditions (such as in the presence of an applied bias) the difference between the generation and recombination rates leads to dark current.[12] Each type of generation-recombination mechanism rate has an associated lifetime τ, with the rate being proportional to the relevant carrier concentration divided by the lifetime. The lifetimes of the various processes add according to

$$\frac{1}{\tau}=\frac{1}{\tau_{radiative}}+\frac{1}{\tau_{Auger}}+\frac{1}{\tau_{SRH}}. \tag{12.5}$$

Therefore, the shortest lifetime dominates. Radiative lifetimes are typically much longer than Auger or SRH and are ignored in the following summary.

Thermal diffusion current arises from thermal recombination mechanisms in the n- and p-absorber regions:

$$J_{diff} = e^{-}n_i^2 t\left(\frac{1}{n\tau_n}+\frac{1}{p\tau_p}\right)\left[e^{\left(\frac{qV_b}{kT}\right)}-1\right], \tag{12.6}$$

where the hole and electron lifetimes $\tau_{n(p)}$ are the lifetimes in the n and p regions, given by Eq.(12.5), t is the thickness of the diode, n_i is the intrinsic carrier concentration, n and p are the doping concentration, V_b is the applied bias, and T is the temperature. n_i is a also function of temperature:

$$n_i=\sqrt{N_C N_V}\,e^{-\left(\frac{E_g}{kT}\right)}, \tag{12.7}$$

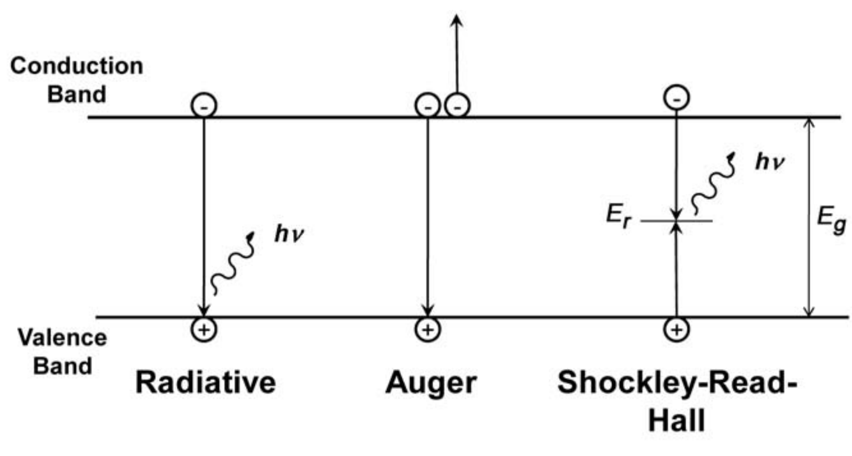


Figure 12.7 Three types of electron–hole recombination mechanisms relevant for IR detectors. (Auger Type 1 shown.)

where N_C and N_V are the conduction band and heavy hole band density of states. For current photon detector technology, either Auger or SRH is the limiting lifetime. From Eqs. (12.6) and (12.7) it can be seen that long carrier lifetimes, lower temperatures, and thinner diodes all contribute to lower diffusion currents. At a typical bias of –0.1 V and at 77 K, the exponential term in Eq. (12.6) is essentially zero, so the current is driven by the prefactor. The downside of reducing the diode layer thickness is a loss of QE. For a heterojunction detector, the diffusion current is dominated by the narrow gap side (typically the n region). Diffusion current from the absorber region represents a sort of theoretical “ideal” dark current.

Depletion current arises from thermal generation in the depletion region via states within the bandgap. Due to the absence of free carriers in the depletion region, SRH recombination dominates. All real semiconductors have energy states within the bandgap that are due to the presence of vacancies, defects, and interstitial atoms. These bandgap states lead to thermal generation recombination within the depletion region via the SRH mechanism. As a simplified treatment, assuming that the SRH level is in the center of the band (at the Fermi level),

$$J_{depl} = \frac{q n_i W}{\tau_{no} + \tau_{po}} \left[e^{\left(\frac{qV_b}{2kT}\right)} - 1 \right]. \tag{12.8}$$

The recombination lifetimes $\tau_{no(po)}$ are given by 1/ $\gamma_{n(p)}$ N_r, where $\gamma_{n(p)}$ are the capture cross sections (cm^{-3} sec^{-1}) for the SRH centers for electrons and holes, and N_r is the density of SRH centers at the Fermi level. n_i is given by Eq. (12.7), and W is the width of the depletion region given by Ref. 1:

$$W = \left[\frac{2\epsilon\epsilon_0}{e} \left(\frac{N_a + N_d}{N_a N_d} \right) \left(V_{bi} + V_b \right) \right]^{1/2}, \tag{12.9}$$

where N_a and N_d are the doping levels and V_{bi} is the built-in potential of the depletion region. N_a, N_d, and V_{bi} are all temperature dependent, so W is a function of the applied bias and temperature. The treatment for an arbitrary SRH center at energy level E_r is similar and is provided by Ref. 3. Lower temperatures and lower bias reduce diffusion currents, as do fewer SRH centers.

Surface current arises from generation recombination on the exposed surfaces/interfaces of the detector and is represented by the general form

$$J_{surf} = \frac{e n_i s}{2}, \tag{12.10}$$

where n_i is the intrinsic carrier concentration [Eq. (12.7)] and s is the surface recombination velocity, a function of the surface trap density, the capture cross section, and the carrier bulk thermal velocity. Surface effects arise from a number of factors including the abrupt termination of the semiconductor crystal, which leads to a large number of localized electrically activated states (traps) and the presence of impurities on surfaces exposed during the semiconductor processing (additional traps). Proper passivation reduces but does not completely eliminate surface currents. Lower temperatures, better *in situ* processing, and effective passivation all serve to reduce surface current. Note that surface current exists at zero bias.

Trap-assisted tunneling and *band-to-band tunneling* can lead to recombination in the depletion region. These effects are not significant at the low biases typically used for passive thermal detectors. Tunneling currents are strongly bias dependent. Figure 12.8 shows modeled dark current components for a HgCdTe detector with a 10.7-μm cutoff at 80 K.[13]

12.2 IR FPA Performance Characterization

In this section we place the primary emphasis on calculating the IR FPA signal-to-noise ratio, which is input into the model as described in Chapter 9. Secondary performance parameters include pixel pitch, fill factor (FF), frame rate/integration time, residual nonuniformity, operating temperature, operability, and spatial and spectral crosstalk (the latter for multiband arrays). The IR detector/FPA literature has, to a fault, many ways to characterize S/N performance. Section 12.2.1 will describe simple background-limited versions of specific detectivity (D^*) and the noise equivalent temperature difference (NETD), which were introduced in Chapter 9. Section 12.2.2 will present a more comprehensive flux-based model for S/N.

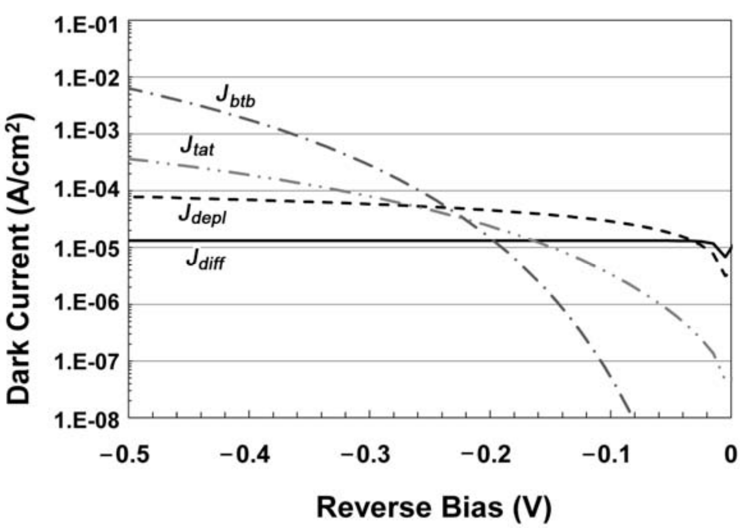


Figure 12.8 Modeled dark current components versus applied bias for a HgCdTe detector diode with a 10.3-μm cutoff.[13] Note that diffusion and depletion (G–R) currents dominate at typical reverse-bias conditions (~ 0.05–0.1 V).

12.2.1 Responsivity and detectivity background-limit performance

The photon detector receives a flux of photons arising from the thermal background and target (approximately equal in many terrestrial cases). The optical system serves to limit the amount of background seen by any one detector element. The resulting flux is Φ_B(photons/sec/cm^2), where B denotes background. Some of the received photons are lost through transmission through the optical system. Some more are lost due to reflection from the front of the detector. The rest enter the detector. Only photons whose energy is greater than the detector bandgap E_g can generate photo-electron–hole pairs, giving a cutoff wavelength λ_c:

$$\lambda_c = \frac{E_g}{hc}, \tag{12.11}$$

where $\lambda < \lambda_c$ contribute to photocurrent. Of these photons, only a certain fraction will be converted into electron–hole pairs that contribute to photocurrent. This fundamental quantity is known as the quantum efficiency η. Generally, the quantum efficiency is dependent on the absorption coefficient of the material α and the thickness of the absorber t

$$\eta \propto 1 - e^{-\alpha t}. \tag{12.12}$$

In principle, a sufficiently thick absorber can yield very high quantum efficiency. Typically, the thickness is a tradeoff between greater absorption and greater dark current, as discussed in Section 12.1.3. In practice, other detector material factors such as short minority carrier lifetime due to the presence of traps and midgap states may limit the QE to less than unity. The reflectivity of the front and back surfaces of the detector also plays a role in overall QE—typically, detectors are antireflection (AR) coated to improve QE. Typical QEs range from a high of 70–80% for HgCdTe and InSb to a low of 2–10% for PtSi and QWP. Quantum efficiency can be a function of wavelength but is often replaced with an average value over the band of interest. In the detectivity model of Chapter 9, fill factor (the ratio of the detector area to the detector pitch) is not included in the QE.

The spectral current responsivity of the detector is the transfer function of optical power into electrical current (A/W). For an ideal photon detector, $R(\lambda)$ is

$$R(\lambda) = \eta e^{-}\left(\frac{\lambda}{hc}\right) \quad (\lambda \le \lambda_c), \tag{12.13}$$

where

e^- = electron charge (1.602×10^{-19} C),
h = Planck's constant (6.626×10^{-34} J sec), and
c = speed of light in vacuum (3.0×10^8 m/sec).

Ideal responsivity for a PV detector is shown in Fig. 12.9. Actual responsivity may slightly or greatly differ from Eq. (12.13). Figure 12.9 also shows for comparison a typical actual responsivity for an FPA. This will be especially true for detectors with narrow-band absorption, e.g., QWIP detectors. For most of the traditional nonbandgap-engineered materials such as HgCdTe, InSb, PtSi, and extrinsic Si at thermal wavelengths (3–12 μm), Eq. (12.13) is typically a good estimate if no additional data is available. Having a correct QE is very important for an accurate model. In some detector literature, QE is separated into different factors, i.e., “absorption QE,” “conversion efficiency,” etc. It is the same case with reflectivity. A gain term is included in photoconductors. The net QE is typically then the product of the different factors. The detector “fill factor” (ratio of active light-sensitive area to total pixel area) may be incorporated. In this model, A_d refers to the detector active area in cm^2, and the FF should not be included in the QE. If the QE values appear to be outside of the typical ranges quoted in the various material sections that follow, then it is highly likely that one or more factors were omitted (or added in the case of FF).

Given the spectral responsivity, a common and useful detector figure of merit is the specific detectivity $D^*(\lambda)$, given by

$$D^*(\lambda) = R(\lambda)\frac{\sqrt{A_d \Delta f}}{I_N}, \tag{12.14}$$

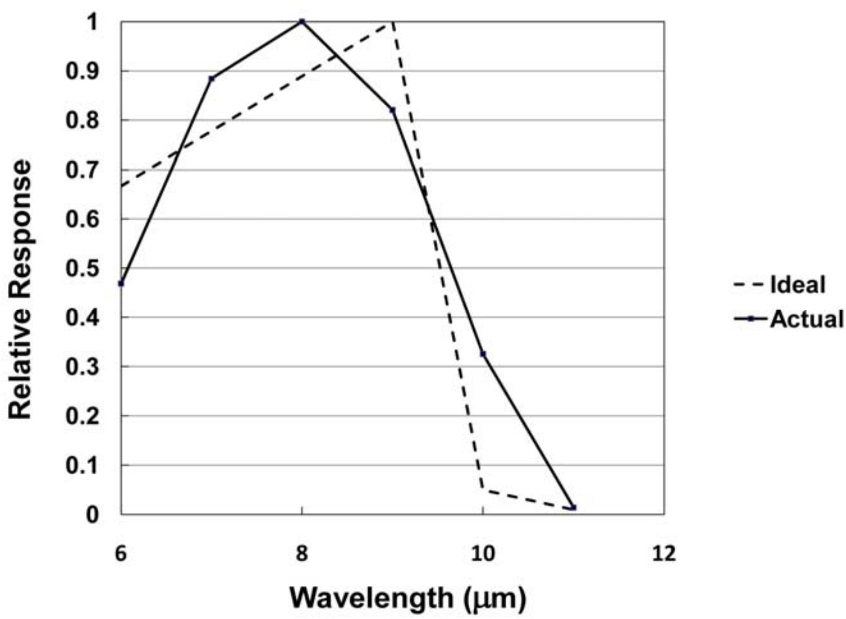


Figure 12.9 Comparison of ideal and actual measured responsivity. The dashed line shows the value predicted by Eq. (12.13), and the solid line shows measured data for a real FPA. Values have been normalized to remove QE differences.

where

A_d is the detector area in cm^2,

Δf is the bandwidth (1/2 t_{int}), and

I_N is the noise current in amperes.

D^* has units of μm cm^2 $Hz^{½}$/W and represents the inverse of the signal required to achieve a S/N of 1 for a particular detector normalized by area and bandwidth. Normalizing out the detector area and the bandwidth (which are specific to particular FPA designs) gives a quantity that is more representative of the fundamental IR detector material. As a figure of merit, D^* also has the satisfying quality that "higher is better." A background-limited infrared photodetector (BLIP) is defined as a detector whose greatest noise source is the shot noise from the thermal background. For a BLIP detector, the background-limited D^* is given by[14]

$$D^*_{BL} = \left(\frac{\lambda}{hc}\right)\sqrt{\frac{\eta}{2\Phi_B}}, \tag{12.15}$$

where η is the QE and Φ_B is the average background flux [Eq.(12.21)]. D^*_{BL} is a good approximation for many detectors, and with sufficient cooling, in principle, all photon detectors should reach background-limited conditions. $D^*_{\lambda peak}$ is often quoted and is the wavelength for which the D^* is maximal. Assuming a detector for which Eq. (12.13) is a good approximation, $\lambda_c = \lambda_{peak}$ and

$$D^*(\lambda) = \frac{\lambda}{\lambda_c} D^*_{\lambda_{peak}} \qquad (\lambda \le \lambda_c). \tag{12.16}$$

If the responsivity differs greatly from Eq. (12.13), then $D^*(\lambda)$ is given by

$$D^*(\lambda) = D^*_{\lambda_{peak}}\, S(\lambda), \tag{12.17}$$

where $S(\lambda)$ is the detector-normalized responsivity (unitless). This is the form used in Chapter 9. $S(\lambda)$ can often be found in detector literature, especially for narrowband detectors such as QWIP detectors. A common plot showing the ideal D^* and typical D^*_{BL} computed for many different IR detector materials can be found in Ref 15.

The noise equivalent temperature difference, which represents the scene temperature difference required to achieve a signal-to-noise ratio of unity is given by[16]

$$NETD = \left(\frac{4}{\pi\, F\#^2 \sqrt{2t_{int} A_d}\, \tau_{opt}}\right) \frac{1}{D^*_{\lambda_{peak}} \int_{\lambda_1}^{\lambda_2} S(\lambda) \frac{\partial W_b(\lambda_{peak}, 300)}{\partial T}\, d\lambda}, \tag{12.18}$$

where

$F\#$ = optics *f*-number,
t_{int} = integration time,
τ_{opt} = average optical transmission,
A_d = detector area (cm^2),
$D^*_{\lambda peak}$ = maximum D^* at wavelength λ_{peak}, and
$\partial W_b\left(\lambda_{peak},300\right)/\partial T$ is the blackbody spectral radiance differential at λ_{peak} and 300 K [Eq. (9.4)].

In general, NETD is an expression of the FPA (or sensor) S/N expressed in terms of the target's effective temperature difference from the background. NETD is a very powerful quantity since it can be easily extended to the entire FPA structure or to the entire sensor if measured data is available. The background-limited approach is typically a good starting point if more detailed information is not available on the FPA of interest. If more information is available, then the flux- based model that follows is more accurate and flexible for accommodating different sensor configurations.

12.2.2 Flux-based signal-to-noise ratio

A more flexible and physically appealing assessment of S/N can be made by accounting for all of the various fluxes (signal and noise) that occur. The goal of this section is to calculate these sources and then incorporate them back into D^* for use in the detectivity model in Chapter 9. The background signal is given by the photocurrent

$$I_S\,(\text{amps}) = \frac{\pi A_d}{4 f_\#^2}\int_{\lambda_1}^{\lambda_2}\tau_{opt}\left(\lambda\right)R\left(\lambda\right)W_b\left(\lambda,300\right)d\lambda\ , \tag{12.19}$$

where

A_d = detector area (cm^2),
$F\#$ = optics *f*-number,
τ_{opt} = optics transmission,
$R(\lambda)$ = spectral current responsivity A/W, and
$W_b(\lambda,300)$ = blackbody spectral radiance at 300 K [Eq. (9.1)].

If Eq. (12.3) is used for the responsivity, then

$$I_S = \eta\Phi_B A_d e^-\ , \tag{12.20}$$

where Φ_B is the background photon flux density (photons/cm^2 sec) falling on the detector:

$$\Phi_B = \left(\frac{\pi}{4 f_{\#}^2}\right)\int_{\lambda_1}^{\lambda_2} \tau(\lambda)\left(\frac{\lambda}{hc}\right) W_b(\lambda, T)\, d\lambda. \tag{12.21}$$

Table 12.1 gives some typical background fluxes at 300 K over thermal wavebands of interest, assuming *f*/1 and that τ_{opt} = 1. Typically, optics transmission is taken as a constant across the waveband.

For an integration time of t_{int}, the resulting signal in electrons n_{source} is simply

$$n_{source} = \frac{I_S\, t_{int}}{e^-}. \tag{12.22}$$

Detector noise is the sum of multiple sources—background noise, thermal noise arising from dark current, ROIC noise, system noise, and nonuniformity. The last term is typically omitted but can be significant for large uncorrected uniformity differences between different detectors in an FPA. Noises are uncorrelated and therefore add in quadrature. The rms system noise $\langle n_{noise} \rangle$ is given by

$$\langle n_{noise} \rangle = \sqrt{\langle n_{shot}^2 \rangle + \langle n_{roic}^2 \rangle + \langle n_{elec}^2 \rangle + \langle n_{pattern}^2 \rangle}, \tag{12.23}$$

where $\langle n_{shot} \rangle$ is the shot noise due to random arrival of signal (background) and dark current.

$$\langle n_{shot}^2 \rangle = \left(\frac{J_{dark} A_d t_{int}}{e^-}\right) + \left(\frac{I_S t_{int}}{e^-}\right), \tag{12.24}$$

where J_{dark} = detector dark current in (A/sec^2).

$\langle n_{roic} \rangle$ represents the ROIC noise in electrons referred back to the detector input. ROIC noise of ~ 500–1000 e^- is typical for high-background FPAs (integration capability on the order of mid-10^6 to ~ 1 × 10^7 e^-). For low-background (e.g., astronomical) FPAs, typical ROIC noise is ~10–100 e^- with (integration capability from 10^4 to 10^5 e^-).

Table 12.1 Photon flux incident on detector at 300-K background (*f*/1, τ_{opt} = 1).

Waveband (μm)	Φ_B (photons/cm^2 sec)
8–12	1.5 × 10^{17}
7.5–9.5	6.3 × 10^{16}
3.4–5.0	3.2 × 10^{15}
3.2–4.2	6.5 × 10^{14}

<n_{elec}> represents off-FPA electronic noise sources. To properly account for such "downstream" noise sources in Eq. (12.23), the sources need to be referred back to the input (i.e., all of the "upstream" gains factored out). This may be difficult in practice, and thermal system performance is usually dominated by the detector and display noise; therefore, unless there is a known issue, this term is typically ignored.

<$n_{pattern}$> represents fixed pattern noise due to uncorrected spatial nonuniformity. Fixed pattern noise (FPN) is also sometimes ignored although it can be a major contributor in FPAs. The difficulty is that the preceding analysis is temporal whereas pattern noise is spatial. A recent trend is for manufacturers to quote residual nonuniformity (fixed pattern noise) as an excess fraction of the temporal NETD. The relative importance of FPN depends on incident flux. FPN is proportional to flux whereas temporal noise is proportional to the square root. An estimate of InSb FPA intrinsic nonuniformity is 0.015 with nonuniformity correction (NUC) providing an additional factor of 50 reduction. The intrinsic nonuniformity of HgCdTe is not as good as that of InSb, but the technology is advancing quickly. QWIP FPAs enjoy good uniformity compared to either InSb or HgCdTe FPAs.

Ignoring electronics and fixed pattern noises, from Eqs. (12.22)–(12.24), the flux-based noise in electrons is

$$\langle n_{noise} \rangle = \sqrt{\left(\frac{J_{dark} A_d t_{int}}{e^-} \right) + \left(\frac{I_S t_{int}}{e^-} \right) + n_{roic}^2} \,. \tag{12.25}$$

Let I_N noise be the noise current in amperes.

$$I_N = \frac{\langle n_{noise} \rangle e^-}{t_{int}} \,. \tag{12.26}$$

At $\lambda = \lambda_{peak}$,

$$D^*_{\lambda_{peak}} = \left(\frac{\lambda_{peak} \eta}{hc} \right) \frac{\sqrt{A_d \Delta f}}{I_N} \tag{12.27}$$

and

$$D^*(\lambda) = D^*_{\lambda_{peak}} S(\lambda) \,. \tag{12.28}$$

The simplest way to incorporate flux into the thermal model is to calculate I_S and I_N and then use Eq. (12.27) to calculate the effective $D^*_{\lambda peak}$. $D^*_{\lambda peak}$ can then be incorporated into Eq. (9.10) and subsequent equations. Typically, NETD is calculated for a frame's integration time t_{int} but can also be calculated for the half-well fill time, in which case

$$t_{int}\,(\text{half fill}) = \frac{C_{roic}}{2\,n_{source}}, \tag{12.29}$$

where C_{roic} is the ROIC unit cell integration capacity expressed in electrons.

Sometimes only an NETD is available for a particular FPA or camera core (FPA plus camera electronics). This is typical for thermal (uncooled) detectors. In this case, $D^*_{\lambda peak}$ can be calculated as the inverse of Eqs. (9.10) and (9.11):

$$D^*_{\lambda_{peak}} = \frac{4F\#^2}{\pi\tau_{opt}\sqrt{2A_d t_{int}}\,NETD}\left[\int_{\lambda_1}^{\lambda_2} S(\lambda)\frac{\partial W_b(\lambda, 300\,\mathrm{K})}{\partial T}\,d\lambda\right]^{-1}, \tag{12.30}$$

where it is assumed that the measured NETD was quoted on a frame basis (t_{int}), and $F\#$ is the *f*-number of optics. For an FPA level measurement, it would be sensible to set $\tau_{opt} = 1$. $F\#$ would be given by the cold shield. If $S(\lambda)$ is not known, a reasonable approximation would be λ/λ_{peak}, $\lambda \leq \lambda_{peak}$ for a photon detector and 1 for a thermal (uncooled) detector.

12.3 Commonly Available Photon Detector FPAs

For terrestrial imaging, the commonly available FPAs are InSb, QWIP, and HgCdTe. Older PtSi detectors may still be available. In the very near future, new III–V type-2 strained-layer superlattices may be available. Extrinsic silicon, used for low-background (space and astronomy) applications, is not typically used for terrestrial imaging. In this section we will give some basic information on InSb, QWIP, and HgCdTe FPAs.

12.3.1 Indium antimonide (InSb) detectors

InSb (indium antimonide) is a well-established detector material. InSb is normally grown as an n-type bulk crystal. After being sliced and polished into wafers, individual detectors are fabricated by ion implantation of a p-type area to form the p–n junction. InSb has also been grown epitaxially. InSb was one of the first detector materials to be fabricated into FPAs.

InSb detector advantages include the following characteristics:

(1) a well-characterized and mature material system with low defects,
(2) excellent uniformity,
(3) ease of fabrication of large FPAs (limited by crystal boule diameter), and
(4) high quantum efficiency and excellent performance at 77 K.

InSb detector disadvantages include the following characteristics:

(1) fixed bandgap limits maximum cutoff wavelength to ~ 5 μm (MWIR only),

(2) excess G–R noise requires cooling to ~ 80 K to achieve BLIP operation, and
(3) fixed bandgap means InSb is not suitable for multiband/multicolor.

High-quality InSb FPAs and cameras are widely available from manufacturers in the U.S., France, and Israel. First-generation InSb FPAs were primarily half-TV or full-TV format (~ 320 × 240 and ~ 640 × 480, respectively) with pixel sizes from 50–25 μm. Since 2004, a new, second generation of FPAs has appeared with much smaller pixel sizes (20–12 μm), enabling high definition and larger formats (1 K × 1 K to 2 K × 2 K). A QE of ~ 80% across the MWIR band (3.4–5.0 μm) is typical with a thinned detector and antireflection coating. At 77 K, reported dark currents average 3–5 × 10^{-8} A/cm^2. Dark current is dominated by G–R noise, which has limited InSb's potential for higher-temperature operations.[3]

Figure 12.10 shows measured and modeled InSb dark currents from several sources. Table 12.2 provides a partial list of available InSb FPAs. For 15-μm reticulated detectors, FF limits QE to ~ 50–60% unless microlenses are incorporated into the detector,[17] in which case the tradeoff for reticulation is improved MTF as discussed below. At 12 μm this effect would presumably be more pronounced. Currently available InSb FPAs do not incorporate microlenses; however, this may change in the near future.

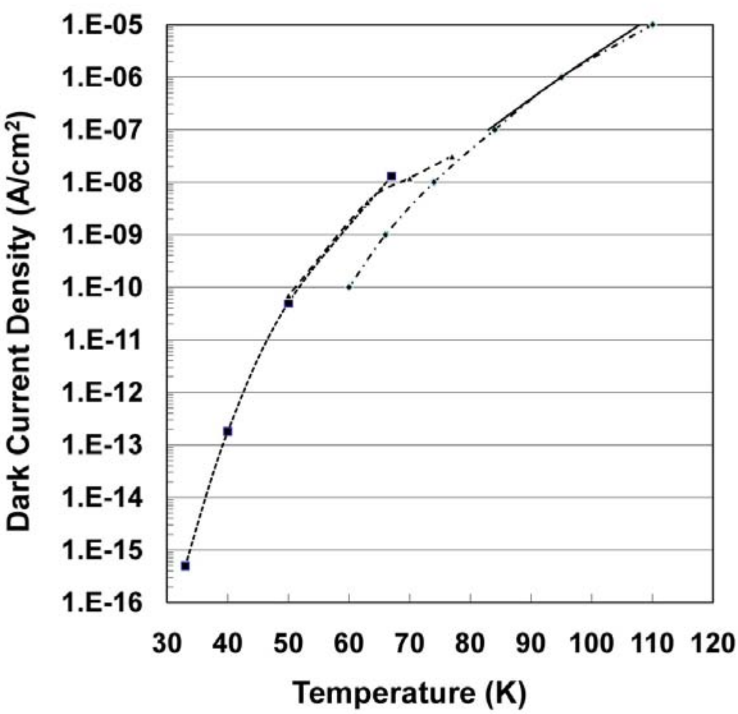


Figure 12.10 Measured and modeled dark current for InSb photdiodes from several sources. Dashed lines are measured data and the solid line is a theoretical model.

Table 12.2 Representative InSb FPAs.

Manufacturer	Format	Spectral Range (μm)	Detector Pitch (μm)	ROIC Capacity (e^-)	NETD (mK)	Notes
Lockheed Martin Santa Barbara Focal Plane Array[20]	640 × 512	< 1–5.3	20	7.0 M		
	320 × 256		30	20.0 M		
	640 × 512		20	8.0 M		
	1280 × 1024		12	2.0 M		
	1024 × 1024		25	10.0 M		
L3 Cincinnati Electronics[17,21]	640 × 512	3–5.5	28	13.0 M	30 (*f*/4)	QE > 0.8
	640 × 512	3–5.5	20	7.0M	<30 (*f*/4)	QE> 0.8, reticulated
	1280 × 1024	3–5.5	15			QE> 0.8, reticulated
Raytheon Vision Systems[22,23]	256 × 256	3–5 (< 1–5)	30			QE × FF > 0.55
	512 × 512	3–5	25	6.0 M		
	640 × 480	3–5	20	18.0 M	20 (16 ms, *f*/6)	QE × FF = 0.9, N_{roic} = 500–900 e^-
	1024 × 1024	3–5	20	17.6 M	11 (*f*/2, 60 Hz)	QE × FF =0.9
SCD (Israel)[24,25,26]	320 × 250	3.6–4.9	30	11.0 M	25 (all at 50%)	*f*/1.5, *f*/3, *f*/3.6, *f*/4
	320 × 256	3.6–4.9	20	13.0 M	20 (well fill)	*f*/1.5–*f*/5.5
	480 × 384		20	13.0 M	20	
	640 × 512	3.6–4.9 (1–5.4)	25	11.2 M	20	*f*/2.9, *f*/4.1, N_{roic} = 350 e^-
	640 × 512	3.6–4.9 (1–5.4)	20	9.5 M	20	*f*/3.1, *f*/5.3, N_{roic} = 300 e^-
	640 × 512	3.6–4.9	15	6.5 M	22	*f*/4
	1280 × 1024	3.6–4.9	15	6.0 M	22	*f*/4.3

InSb FPAs that use a planar detector without reticulation are subject to lateral diffusion of minority carriers to adjacent pixels, leading to crosstalk. The resulting MTF is over and above the usual sinc function associated with the detector physical dimensions.[18] The diffusion MTF for a backside-illuminated planar detector is given by[19]

$$MTF_{diff}\left(f\right) \approx \frac{\cosh\left(L_D / L_0\right)}{\cosh\left(L_D / L\right)}, \tag{12.31}$$

where

$$L^{-2} = L_0^{-2} + \left(2\pi f\right)^2. \tag{12.32}$$

The spatial frequency is f (cy/mm), L_0 is the diffusion length, and L_D is the field free depth in the absorber. In Ref. 18, a simplified version of Eq. (12.31) was fit to experimental data, producing a good fit for bulk InSb. Typical values for InSb at 77 K of L_0 ~ 30 μm and L_D ~ 5–7 μm provide a good fit to the MTF discussed in Ref. 18, where it is also reported that this effect is not seen in reticulated InSb detectors. This MTF should apply to other backside-illuminated infrared detectors; however, the values will depend on detector design details.

12.3.2 Quantum well infrared photoconductor (QWIP) detectors

QWIP detectors are formed by alternating layers of strain-matched GaAs/AlGaAs, such that a series of potential barriers are formed in the conduction band at the boundaries of the alternating layers (Fig. 12.11). QWIP detectors have been made practical in the last decade due in part to rapid advances in the low-cost fabrication of complex III–V-compound semiconductors as opto-electronic components for the telecommunication industry. The potential barriers shown in Fig. 12.11 form a series of quantum mechanical wells with the allowed energy states within the well calculable from the width and depth of the well. At cryogenic temperatures, electrons provided by dopant atoms reside in the lowest energy (ground) state of the well. The most advanced QWIP design has an excited state close to, or directly at, the top of the well. A photon of correct wavelength can excite the ground-state electron to the higher state in the well, and an applied electric field (the bias) enables the electron to sweep or tunnel out of the well into the conduction band (Fig. 12.11).

Under constant illumination, a photocurrent flows through the detector. It should be clear that unlike InSb and HgCdTe, QWIPs are inherently narrowband devices. The center of the band is tuned to the ground-state–excited-state energy difference. The spectral range is determined by the width of the ground- and excited-state energy levels. Heavily doped thicker cap layers above and below the quantum well structure are used as electrical contacts to the device. Although QWIPs are photoconductors, for modeling purposes, they behave sufficiently similar to PV devices to be treated the in the same manner, apart from accounting for the photoconductive gain in QE.

Sources of dark current include generation-recombination noise from thermal excitation and decay of carriers in the well, as well as Johnson noise arising from the series resistance of the QWIP.[26] At typical operating temperatures, G–R noise is dominated from thermionic emission of electrons over the potential barriers in the conduction band shown in Fig. 12.11. Overall photon absorption can be improved using multiple well layers—the tradeoff is that more layers increase the probability of recombination and increase dark current. The practical upper limit appears to be about 50 layers.

It is important to recognize that the quantum wells exist in one dimension only, which is 90 deg to the natural orientation for incoming light for a focal plane array. To achieve absorption, the light must be reoriented so that its Poynting vector is in the direction parallel to the face of the detector, typically accomplished by placing an optical diffraction grating on top of the detector (shown schematically in Fig. 12.11). Only the incident light with its E field polarized in the plane of the well can be absorbed.[28] These two characteristics serve to limit the QE of a QWIP detector compared to the intrinsic direct-bandgap detectors such as InSb and HgCdTe.

QWIP detector advantages include the following characteristics:

(1) uses mature low-cost semiconductor material and fabrication techniques derived from telecommunications,
(2) detector material has inherently high uniformity, and
(3) uses flexible material: tailorable band gap (MWIR and LWIR) and multicolor/multiwavelength devices.

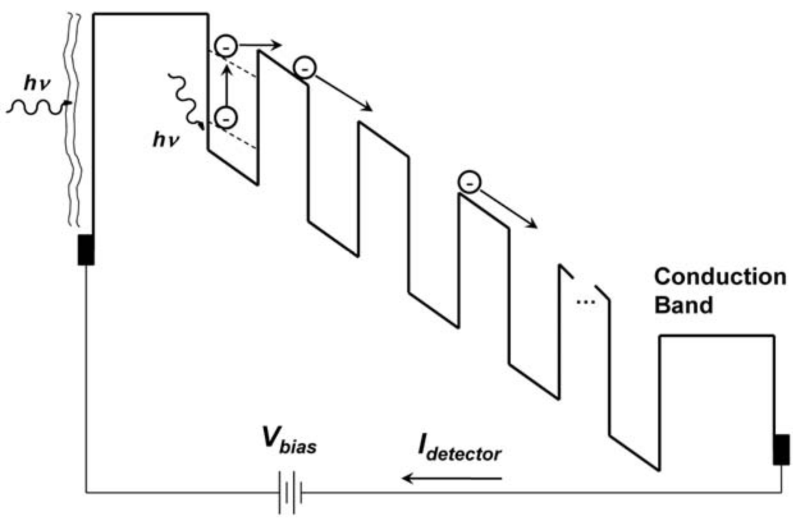


Figure 12.11 Schematic energy diagram of a QWIP detector.[27] Only the conduction band is shown. Dopant atoms provide the ground-state electrons. A scattering grating or some other mechanism is required to reorient the incoming photons parallel to the wells.

QWIP detector disadvantages include the following characteristics:

(1) low QE and low photoconductive gain,
(2) narrow spectral band, and
(3) limited potential for low background or operation above ~ 65–70 K at LWIR due to excess dark current from thermionic emission.

QWIP FPAs and cameras are available from manufacturers in the U.S., France, Sweden, and Israel in a variety of formats up to 1000 × 1000, but typically have a 320 × 240 or 640 × 480 format. The most advanced QWIPs available are purported to have QEs of about 6%, although 2% is typical. One potential source of confusion in the QWIP literature is splitting the QE into multiple terms, such as internal quantum efficiency, conversion efficiency, photoconductive gain, etc. As of this writing, any production QWIP peak QE above ~ 6% should be examined carefully; perhaps one of the terms is not being included. QWIPs are reticulated, so if no FF is given, one can be estimated assuming ~ 2-μm trench width. For a 25-μm pixel, the FF would be on the order of 85%.

Figure 12.12 shows some typical QWIP detector I–V curves and illustrates the bias dependence in dark current. MWIR QWIPs have been reported in the literature, although their current production status is uncertain. There is a general inverse relationship between the peak responsivity and spectral width; however, due to the possible variations in well design, number of wells, and bias condition, it is not possible to give general rules for QE, J_{dark}, and the spectral band. It is not so much that the absolute performance is better for different recipes but rather that the various parameters can be traded off against each other. Table 12.3 gives some examples of LWIR QWIPs that illustrate typical configurations.

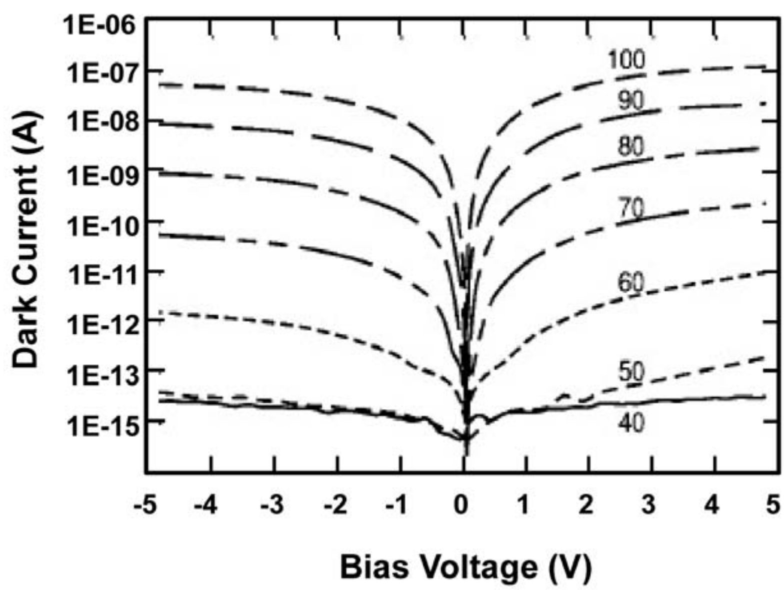


Figure 12.12 Typical QWIP detector I–V curves as a function of temperature and bias.[27]

Table 12.3 Representative QWIP FPA parameters.

QWIP FPA Parameters	QWP Tech[30,31]	Thales Catherine XP[29,32]	JPL[28]
Array Size	1024 × 1024	384 × 288 w/4:1 dither	640 × 486
Nominal Operating Temperature (K)	65 K	75 K	70 K
Pixel Pitch (μm)	19.5 × 19.5μm	25 × 25 μm	25 × 25 μm
Fill Factor	81%	85%	51%
Peak Responsivity (A/W)	0.45 A/W	0.6 A/W	0.13 A/W
Estimated Peak QE	6.6%		3.3%
Peak Wavelength Approximate FWHM	8.4 μm 0.75 μm	8.6 μm 0.9 μm	8.3 μm 1.0 μm
Relative Responsivity @ Wavelength	.05 7 μm .20 7.5 μm .50 8 μm 1.0 8.5 μm .50 8.75 μm .08 9.5 μm	.08 7 μm .40 8 μm 1.0 8.6 μm .52 9 μm .05 9.5 μm	0.09 6.5 μm 0.19 7 μm 0.46 7.5 μm 0.85 8 μm 1.0 8.3 μm 0.27 9 μm 0.06 9.5 μm
Dark Current (A/cm^2) @ stated bias & temp.	3.3e–6 @ –2.0 V, 65 K	3e–4 @ –1.5 V, 75 K	3.1e–6 @ –2.0 V, 70 K
Frame Rate	60 Hz	100 Hz	60 Hz
Integration Time	3.5–4.0 msec	4 msec	
ROIC Storage Capacity	6.7e6 e$^-$	18.5e6 e$^-$	9e6 e$^-$
ROIC Noise	1000 e$^-$	850 e$^-$	500 e$^-$

Figure 12.13 shows a typical relative (spectral) response for a QWIP tuned to a center wavelength of 8.3 μm.[29] MWIR response curves are qualitatively similar. It should be clear from Fig. 12.13 that QWIP detectors have a much narrower spectral response than have InSb or HgCdTe detectors.

Most QWIPs offered as cameras have *f*-numbers on the order of *f*/2–*f*/2.5, with typical integration times on the order of 3–4 msec. Reported D^* values are in the range of ~1×10^{11} cm^2 Hz$^{1/2}$μm/W at *f*/2–*f*/2.5. Typical operating temperatures are 60–70 K, with 65 K being fairly common.

12.3.3 Mercury cadmium telluride (HgCdTe) detectors

The II–VI tertiary alloy of mercury, cadmium, and tellurium is a very powerful and flexible material system for IR detectors. HgCdTe is an intrinsic direct-bandgap semiconductor. Adjusting the ratio of mercury to cadmium ($Hg_xCd_{1-x}Te$) during growth produces a variable bandgap with cutoff wavelengths from the visible (< 1 μm) out to the very long-wave IR (> 14 μm).

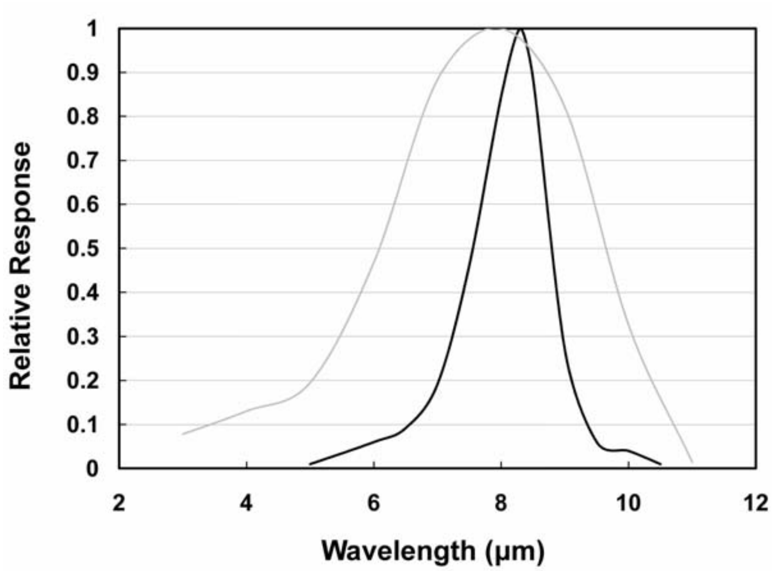


Figure 12.13 Typical relative response curve for a QWIP detector tuned to 8.3 μm. For comparison, the dashed line shows an HgCdTe detector with a nominal cutoff wavelength of ~ 9 μm.

The resulting detectors are notable for high quantum efficiency and near theoretical (diffusion-limited) noise characteristics over the thermal wavebands (MWIR and LWIR). High-performance multiband/multiwavelength detector FPAs have been developed and produced. HgCdTe for modern IR FPAs is grown using liquid phase epitaxy, chemical vapor deposition, and molecular beam epitaxy. High-quality HgCdTe FPAs are available from manufacturers in the U.S., France, and the UK.

HgCdTe has been the material of choice for LWIR sensors since the 1970s. Operation at 77 K is typical in the LWIR. A recent trend is to use very compact MWIR sensors operating at elevated temperatures (120 K or more), where BLIP-limited operation is still possible due to HgCdTe's low dark current at MWIR wavelengths. Lower-cost detectors grown on commercial substrates such as Si and GaAs are beginning to come to market as of this writing. There are several important limitations to HgCdTe. Uniformity and operability are typically lower than for III–V materials, especially for InSb. Detector material costs are higher than for the III–V materials (QWIP and InSb) due to the specialized foundries required for growth. Producibility has been a historic challenge, although recent trends suggest more stable manufacturing processes at the major suppliers.

HgdTe detector advantages include the following characteristics:

(1) high QE and low dark current,
(2) detector sensitivity from NIR to very LWIR possible by changing material composition,
(3) high-temperature operation in MWIR demonstrated (140 K), and
(4) suitability for multiband or multicolor wavelengths.

HgCdTe detector disadvantages include the following characteristics:

(1) lower uniformity and operability than III–V detector materials, and
(2) high cost.

Table 12.4 lists some representative HgCdTe IR FPAs. In some cases QE and FF were separately estimated from a specified QE × FF, and dark current was estimated from similar published data. Manufacturers should be consulted for more detailed information. HgCdTe possesses a long minority carrier lifetime (~ 1 μsec) and a lack of excess noise mechanisms at cutoff wavelengths less than 12 μm (e.g., ~ 0.124 eV).[32] HgCdTe dark current is dominated by diffusion current at 77 K. Diode I–V characteristics are typically excellent, as shown in Fig. 12.14. Typical reverse bias for MWIR and LWIR detectors is about 0.1 V.

Figure 12.15 is a plot of Tennant's "Rule 07," an up-to-date empirical fit to HgCdTe dark current at reverse bias as a function of temperature and cutoff wavelengths for HgCdTe detectors grown on CdTe or CdZnTe substrates.[34] Rule 07 can be used to make an estimate of the dark current given changes in operating temperature or cutoff wavelength down to a $\lambda_c T$ product of ~ 400. Typical QEs with antireflection coating are 70–80%. For detectors operating at ~ 100 mV of reverse bias, typical dark currents are ~ 1×10^{-5} A/cm^2 at 10.5 μm in the LWIR at 77 K and ~ 1×10^{-8} A/cm^2 at 5 μm in the MWIR at 120 K. At 80 K, MWIR dark currents are very low (< 1×10^{-10}). Fill factors depend on the device geometry. For mesa-type detectors (Fig. 12.1), a trench width of ~ 2 μm is typical.

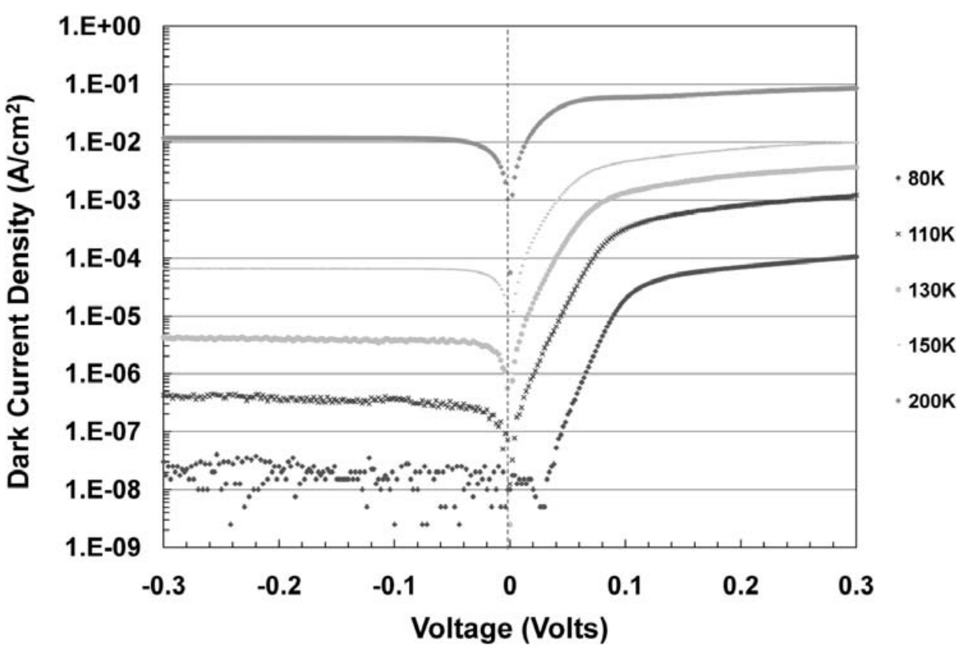
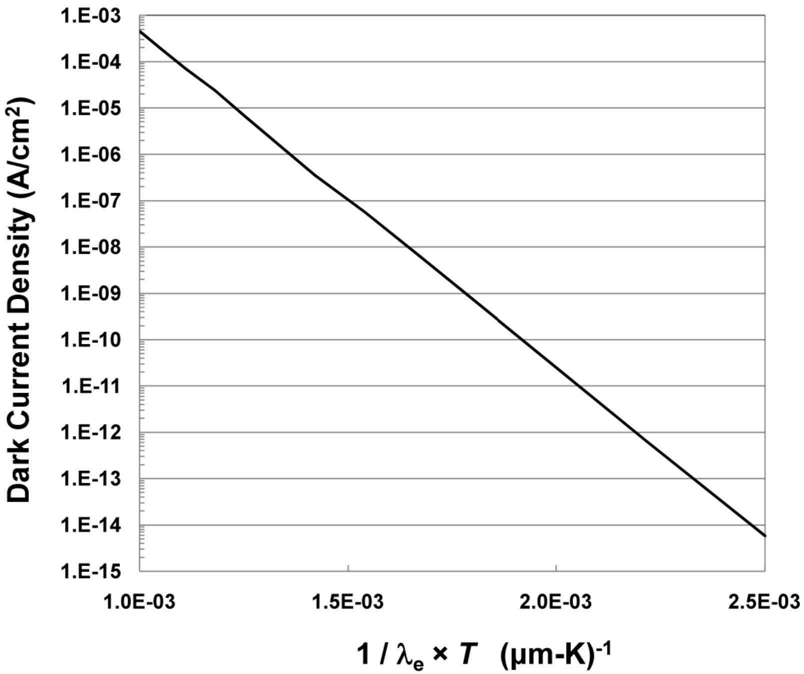


Figure 12.14 Typical HgCdTe detector diode I–V curves that show diffusion-limited characteristics in reverse bias over a wide range of temperatures.[40]

Table 12.4 Representative HgCdTe FPA parameters.

Manufacturer	Array Size	Pitch	Waveband	Operating Temp.	Estimated Dark Current (A/cm^2)	ROIC Noise (e^-)	Integration Capacity	QE	Fill Factor
DRS[3,36]	256 × 256	40 μm	3–5	77 K	1e–9			0.8	0.75
	256 × 256	40 μm	8–10.5	77 K	1e–5			0.73	0.75
	640 × 480	25 μm	3–5	77 K	1e–9			0.8	0.75
	640 × 480	25 μm	8–10.5	77 K	1e–5			0.8	0.75
Sofradir[35,37,38]	320 × 256	30 μm	8–9.5 (11)	77K (70K)	2e–6 (2e–5)	1000	12 e6	0.8	0.7
	320 × 256	30 μm	3.7–4.8	110 K	1e–7	1000	12 e6		
	384 × 288	15 μm	3.4–4.8	110 K	1e–7	400	3.4 e6	0.7	0.9
	384 × 288	25 μm	7.7–9.5	77 K		400	4.8 e6		
	640 × 512	20 μm	3.7–4.8	85 K		400	4.8 e6		
	640 × 512	15 μm	3.7–4.8	91 K		420	6.5 e6	0.7	0.9
	1280 × 1024	15 μm	3.7–4.8			400	4.2 e6	.7	0.9
Selex[39]	320 × 256	30 μm	8–9.4	77 K	1e–3			0.7	0.75
	640 × 512	24 μm	8–9.4	77 K	1e–3			0.7	0.84
(HgCdTe on GaAs)	640 × 512	16 μm	8–9.4	77 K	1e–3			0.7	



$$J = J_0 e^{C(1.24q/k\lambda_e T)}$$

$$\lambda_e = \lambda_{cut-off} \text{ for } \lambda_{cut-off} \geq \lambda_{threshold}$$

$$\lambda_e = \lambda_{cut-off} \Big/ \left[1 - \left(\frac{\lambda_{scale}}{\lambda_{cut-off}} - \frac{\lambda_{scale}}{\lambda_{threshold}}\right)^{Pwr}\right]$$

$$\text{for } \lambda_{cut-off} < \lambda_{threshold}$$

Where:

J_0 =	8367.000019
Pwr =	0.544071282
C =	-1.162972237
λ_{scale} =	0.200847413
$\lambda_{threshold}$ =	4.635136423

Figure 12.15 Tennant's "Rule 07" for estimating HdCdTe dark current density for HgCdTe detectors grown on CdTe or CdZnTe substrates. Tennant et al. report their Rule 07 to be an accurate estimate of dark currents for high-quality HgCdTe photodiodes within a factor of 2 over the range of cutoff wavelengths and operating temperatures shown.[34]

For planar detectors with small central junctions, FF will be influenced by the minority carrier diffusion length. For planar pixels on the order of 20–40 μm, a FF of 70–80% is typical. At 90% FF, a 15-μm pixel pitch has been reported.[35] It is notable that the implanted junction detectors tend to have the lowest dark current, which may be desirable in very low-background conditions.

12.4 Uncooled Detectors

12.4.1 Introduction

Uncooled detectors represent an exciting breakthrough in the size, weight, cost, and power consumption of thermal imagers. Thermal imagers using uncooled FPAs have become the *de facto* standard for many application areas where these constraints are paramount. Uncooled detectors work on a different physical principal than do cooled photon detectors. Photon detectors sense external radiation by conversion of photons to electron (or hole) current and integration, i.e., by counting. Uncooled detectors sense external radiation through a physical change in the detector material, typically resistance, as a function of power absorbed by the detector.

While the cooled photon detector's ROIC, under continuous bias, accumulates current in an integration capacitor for a frame integration time, the uncooled detector is allowed to absorb radiation (heat up) for a frame time, and the ROIC uses a continuous bias or *bias pulse* to sense the detector resistivity for that frame. Since there is no counting of current during the frame-integration time, the uncooled detector does not require operation at cryogenic temperatures to reduce the flow of dark current to manageable levels for the ROICs integration capacitor. This is not to say that uncooled detectors outperform cooled detectors (the opposite is true) or operate at the BLIP limit, merely that cooling is not *required* to achieve a practical device.

The current state-of-the art in uncooled detectors is the thin-film-suspended microbridge bolometer. See Fig. 12.16 for a typical single-pixel uncooled microbolometer geometry. A thin film [typically 400–800 Å (angstroms)] of active material, typically vanadium oxide (VO_x) or amorphous silicon (α-Si), serves as the active material. The active material is deposited on a silicon layer and coated with an absorber, typically SiN. A typical detector thickness is on the order of 0.1–0.2 μm.[41] Both VO_x and α-Si have a significant change in resistance around room temperature, defined by the temperature coefficient of resistance (TCR). Narrow legs provide an electrical (and a heat-conduction) path to the ROIC. The legs may be either on the sides of the detector or tucked underneath to maximize the FF of the active area. Advances in semiconductor photolithography have made practical the creation of the necessary small features, and both VO_x and α-Si are compatible with standard silicon CMOS (complementary metal-oxide semiconductor) processes, making uncooled detectors inherently less expensive than either the III–V or II–VI semiconductors used in photon detectors. The detector is enclosed in a vacuum dewar in order to minimize heat conduction through air.

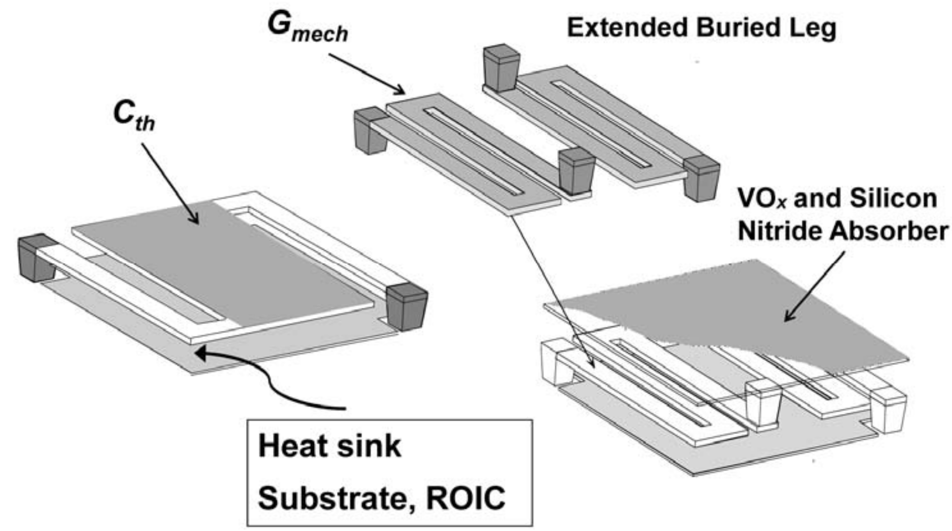


Figure 12.16 Example of an uncooled microbolometer pixel showing a two-level structure. One-level structures with legs on the side also are in use.

The principal figure of merit for an uncooled detector is the TCR, defined by

$$\alpha = \frac{1}{R}\frac{dR}{dT}, \tag{12.33}$$

where the heat capacity of the detector C_{th} and the thermal conductance of the bolometer G_{mech} are typically dominated by the conductance of the legs. Typical practical values of α are ~ 2.5% for VO_x and ~ 3–4% for α-Si. For state-of-the-art 25-μm pixels, a typical C_{th} ~ 0.5 nJ/K and G_{mech} ~ 50 nW/K.[42] Other important parameters are the detector FF and the detector emissivity. While it might be tempting to think that performance would be dominated by TCR, in reality C_{th} and G_{mech} play equal roles as will be demonstrated below. The total conductance of the bolometer is the sum of the leg conductance and the bolometer's direct radiation, denoted G:

$$G = 4\varsigma\sigma T_d^3 + G_{mech}, \tag{12.34}$$

where ς is a parameter that accounts for emissivity of the detector, the waveband (if less than the full blackbody waveband), and the detector effective area accounting for both sides and any reflective surfaces on the substrate. A typical value of ς might be 0.7–0.8.

Unlike photon detectors, whose response can be considered instantaneous for practical purposes, the combination of C_{th} and G define a characteristic response time of the uncooled detector with time constant τ

$$\tau = \frac{C_{th}}{G}. \tag{12.35}$$

Thus, the maximum bandwidth B, independent of frame rate, is given by

$$B = \frac{1}{2\tau}. \tag{12.36}$$

It is typical to see uncooled detectors operating at 60 Hz with time constants on the order of 10–15 msec. The equivalent for a photon detector would be an integration time on the order of 10–15 msec, which is extremely high. While uncooled detectors may achieve reasonable S/N, it needs to be understood that this is useful only for the case of limited target-to-sensor movement.

12.4.2 Signal-to-noise ratio and performance limits

The signal associated with a small change in scene ΔT follows the treatment in presented in Chapter 9.

$$W_S = \frac{\eta}{4F\#^2} H_{pit} V_{pit} \frac{dP_S(300\,\mathrm{K})}{dT_S} \Delta T, \tag{12.37}$$

where

$$\frac{dP_S(300\,\mathrm{K})}{dT_S} = \int_{\lambda_1}^{\lambda_2} \frac{dW_b(300\,\mathrm{K})}{dT} dT \tag{12.38}$$

and $F\#$ is the f-number of optics. dP_S/dT_S is equivalent to δ_{scene} in Eq. (9.12). dP_S/dT_S is 1.98×10^{-4} W cm^{-2}K^{-1} for the 8- to 12-μm spectral band and 2.11×10^{-5} W cm^{-2}K^{-1} for the MWIR spectral band. η includes the optical transmission loss, fraction of incident radiation absorbed (analogous to QE), and the FF fraction, analogous to Eq. (9.14).

Noise consists of contributions from multiple sources, including

- detector fluctuations in temperature from heat exchange with the background,
- Johnson noise from the applied bias signal,
- $1/f$ noise, and
- ROIC noise.

Elementary treatments of uncooled detectors consider only the temperature fluctuation noise and can lead to very optimistic appraisals of the potential for uncooled performance. The *NVESD Handbook for Uncooled Bolometer Detector Technology*[42] provides a detailed model for uncooled detector performance, which accounts for the primary noise mechanisms under both pulse and continuous bias. The simplified prediction of NETD from that model follows.

$$\Delta T_{NETD} \approx \gamma \, \Delta T_{TF} \sqrt{1 + \frac{2}{\alpha^2 T_{sub} \Delta T_r} + \frac{C_{th}}{k T_{sub}^2} \frac{\chi^2}{\alpha^2} B_{1/f} + \ldots}, \tag{12.39}$$

where

$$\Delta T_{TF} = \sqrt{\frac{k T_{sub}^2}{Cth}}. \tag{12.40}$$

$$\gamma = \frac{4 \varsigma \sigma T_d^3 + G_{mech}}{\frac{\eta}{4\, F\#^2} H_{pit} V_{pit} \frac{dP_S(300\,\mathrm{K})}{dT_S}}, \tag{12.41}$$

where ΔT_{TF} is the temperature fluctuation noise of the detector, *F#* is the *f*-number of optics, and γ is a function of the conductance and the input signal. The second term in the radical in Eq. (12.39) represents the Johnson noise associated with the applied bias raising the temperature of the detector ($T_{sub}\Delta T_r$), and the third term represents the 1/*f* noise. Additional terms not shown would include the ROIC noise. Some of the tradeoffs are apparent. For example, to decrease the thermal time constant requires either raising *G* (fatter legs) or decreasing *C* (thinner detectors). Either method raises the NETD. A smaller detector also raises the NETD. Moving to smaller pixels while maintaining NETD and time constant requires a reduction in the bias, 1/*f*, and ROIC noise terms.

Looking at only the temperature fluctuation noise for an ideal 25-μm pixel and the state-of-the-art parameters discussed above leads to a theoretical ideal (*f*/1, $\eta = 1$) LWIR NETD of ~ 11 mK, which is impressive, but unrealistic. When a *theoretical design goal* (Johnson noise from the bias, 1/*f* noise, and ROIC noise all equal to the temperature fluctuation noise) is assumed (still optimistic), then the NETD is 23 mK. When a *more realistic* absorption efficiency (still at *f*/1) of 50% is included, the NETD is 46 mK.

Compare this to a typical measured InSb or HgCdTe NETD at *f*/3–*f*/4 of 15–20 mK and it is clear that, while uncooled detectors have usable capability under limited circumstances (*f*/1, 60 Hz, LWIR, limited target-to-sensor movement), they cannot directly compete with cooled photon detectors when the sole metric is performance. Recall that dP_S/dT_S drops ~ 10× in the MWIR relative to the LWIR, making MWIR operation for imaging virtually impractical. However, uncooled sensors are significantly less expensive, making their use attractive when the above constraints can be satisfied. They are also lighter, consume less power, and are smaller and more reliable than their cooled counterparts, all of which may be deciding factors in certain applications.

To summarize,

Uncooled detector advantages include the following characteristics:

(1) lowest-cost imaging thermal sensor available,
(2) low power and small size and weight, and
(3) high reliability.

Uncooled detector disadvantages include the following characteristics:

(1) low performance (typically requires *f*/1 optics),
(2) slow time constant, and
(3) operation practically limited to LWIR.

12.4.3 Typical uncooled detectors

Typical uncooled detectors vary less than their cooled counterparts. FPAs and camera cores are commonly available from the U.S. and France. Manufacturers in other countries such as Japan and China may come to market soon. The current state-of-the-art is based on ~ 25 × 25-μm detectors. Obsolete 35- to 50-μm detectors may still be available. A new generation of 17- to18-μm devices is imminent as of this writing but not currently available. Formats are standard TV and half-TV (640 × 480, 320 × 240, or European equivalents). High-definition formats will be available with the new 17- to18-μm devices. Frame rates are a typical 60 Hz, with *f*/1 operation assumed in quoting performance. NETDs range from 40–80 mK, depending on time constant, pixel size, and other factors. Current time constants range from ~ 7–15 msec. Table 12.5 lists some commonly available uncooled FPAs.

Due to the strong dependence of detailed fabrication parameters on ultimate performance, there is no simple analog to the flux-based model described in

Table 12.5 Representative uncooled focal plane arrays.

Manufacturer	Format	Detector Pitch (μm)	Time Constant	NETD (mK)	Notes
DRS VO_x[43]	320 × 240 640 × 480	25 25		40 40	
RVS VO_x[22]	640×512/480 320 × 240	25 25	< 16 msec < 16 msec	< 50 < 50	NETD @ *f*/1, 30 Hz NETD @ *f*/1, 30 Hz
BAE VO_x[44]	640 × 480 320 × 240	28 28		< 100 < 100	
Sofradir (ULIS) α-Si[37,45]	160 × 120 384 × 288 640 × 480 640 × 480 10224 × 768	25 25 25 17 17	7 msec 7 msec 7 msec 10 msec 10 msec	70 70 70 < 60 < 60	 @ *f*/1, 300 K, 30 Hz

Section 12.2. Uncooled detector performance is universally summarized in the literature by NETD. For the purposes of this model, it is sufficient to calculate an effective D^*_{peak} using Eq. (12.30), assuming that $S(\lambda) = 1$, $f/1$, $t_{int} = \tau$, and A_d = detector pitch (i.e., FF = 1). This effective D^*_{peak} can be used in generating the SN_{det} and Γ_{det} [Eqs. (9.13) and (9.15)].

References

1. J. Singh, *Semiconductor Optoelectronics: Physics and Technology*, McGraw Hill, New York (1995).
2. M. Henini and M. Razeghi, *Handbook of Infrared Detection Technologies*, Elsevier, Oxford (2002).
3. M. A. Kinch, *Fundamentals of Infrared Detector Materials*, SPIE Press, Bellingham, WA (2007).
4. T. J. de Lyon, R. D. Rajavel, J. A. Roth, and J. E. Jensen, in *Handbook of Infrared Detection Technologies*, M. Henini and M. Razeghi, Eds., p. 337, Elsevier, Oxford (2002).
5. S. K. Singh, V. Gopal, and R. H. Mera, "Relationship between deep levels and R_0A product in HgCdTe diodes," *Opto-Electronics Review* **9**, 385–390 (2001).
6. P. Norton, "Status of infrared detectors," *Proc. SPIE* **3379**, 102–114 (1998).
7. C. C. Hseih, C. Y. Wu, F. W. Jih, and T. P. Sun, "Focal plane arrays and CMOS readout techniques of infrared imaging systems," *IEEE Trans. on Circuits and Systems for Video Technology* **7**, 594–605 (1997).
8. B. Pain and E. R. Fossum, *A Review of Infrared Readout Electronics for Space Science Sensors*, JPL Report (1993).
9. W. Wang, Y. Fan, Q. Guo, and J. Liu, "Nonuniformity analyses of IRFPA with DI readout circuit," *Proc SPIE* **7383**, 738344 (2009).
10. J. T. Longo, D. T. Cheung, M. A. Andrews, C. C. Wang, and J. M Tracy, "Infrared focal planes in intrinsic semiconductors," *IEEE Trans. on Electron. Devices* **ED-25**, 213–233 (1978).
11. FLIR Indigo website: www.corebyindigo.com.
12. B. Van Zeghbroeck, *Principles of Semiconductor Devices*, Ch. 2, http://ecee.colorado.edu/~bart/book/contents.htm, 2007.
13. C. Billman, personal communication (2010).
14. A. Rogalski, *Infrared Detectors*, p. 17, Gordon and Breach, Amsterdam (2000).
15. A. Rogalski, "Narrow-gap semiconductor photodiodes," *Proc. SPIE* **3287**, 2–12 (1998).
16. R. G. Driggers, P. Cox, and T. Edwards, *Introduction to Infrared Electro-Optical Systems*, p. 289, Artech House, Boston, (1999).

17. R. Rawe, C. Martin, M. Garter, D. Endres, B. Fischer, M. Davis, J. Devitt, and M. Greiner, "Novel high fill-factor, small pitch, reticulated InSb IR FPA design," *Proc SPIE* **5783**, 899–906 (2005).
18. M. Davis, M. Greiner, J. Sanders, and J. Wimmers, "Resolution issues in InSb focal plane array system design," *Proc. SPIE* **3379,** 288–299 (1998).
19. D. F. Barbe, "Imaging devices using the charge-coupled concept," *Proc. IEEE* **63**, 38–67 (1975).
20. Lockheed Martin Santa Barbara Focal Plane Array website: www.sbfp.com.
21. L3 Cincinnati Electronics website: www.cinele.com.
22. Raytheon Vision Systems website: www.raytheon.com.
23. A. W. Hoffman, E. Corrales, P. J. Love, J. Rosbeck, M. Merrill, A. Fowler, and C. McMurtry, "2K × 2K InSb for astronomy," *Proc. SPIE* **5499**, 59–67 (2004).
24. SCD USA website: www.scdusa.com.
25. J. O. Schlesinger, Z. Calahorra, E. Uri, O. Shick, T. Fishmann, I. Shtrichman, E. Sinbar, V. Nahum, E. Kahanov, B. Shlomovich, S. Hasson, N. Fishler, D. Chen, and T. Markovitz, "Pelican: SCD's 640 × 512/15 μm pitch InSb detector," *Proc. SPIE* **6542**, 654231 (2007).
26. O. Nesher, I. Pivnik, E. Ilan, Z. Calahorra, A. Koifman, I. Vaserman, J. O. Schleisenger, R. Gazit, and I. Hirsh, "High resolution 1280×1024, 15 μm pitch compact InSb IR detector with on-chip ADC," *Proc. SPIE* **7298**, 72983K (2009).
27. J. Jiang and M. Razeghi, "GaInSb(P) Based QWIPs on GaAs, InP, and Si Substrates for Focal Plane Arrays", in *Handbook of Infrared Detection Technologies*, M. Henini and M. Razeghi, Eds., p. 132, Elsevier, Oxford (2002).
28. S. D. Gunapala, S. V. Bandera, J. K. Liu, C. J. Hill, S. B. Rafol, J. M. Mumolo, J. T. Trinh, M. Z. Tidrow, and P. D. LeVan, "Multicolor megapixel QWIP focal plane arrays for remote sensing instruments," *Proc. SPIE* **6308**, 63080P (2006).
29. E. Costard and A. Nedelcu, X. Marcadet, E. Belhaire, P. Bois, "QWIP development status at Thales Research and Technology," *Proc. SPIE* **6206**, 62060D (2006).
30. A. Goldberg and K. K. Choi, "Recent progress in the application of large-format and multispectral QWIP IR FPAs," *Proc. SPIE* **5406**, 624–638 (2004).
31. A. Goldberg, E. Cho, and B. McQuiston, "Laboratory and field performance of a megapixel QWIP focal plane array," *Proc. SPIE* **5783**, 755–766 (2005).
32. E. Costard and Ph. DuBois, "Thales long-wave QWIP thermal imagers," *Infrared Physics and Technology* **50**, 260–269 (2007).

33. R. G. Pratt, J. Hewett, P. Capper, C. L. Jones, and M. J. Quelch, "Minority carrier lifetime in n-type Bridgman $Hg_{1-x}Cd_xTe$," *J. Applied Physics*, **54**(9), 5152–5157 (1983).
34. W. E. Tennant, D. Lee, M. Zandian, E. Piquette, and M. Carmody, "MBE HgCdTe technology: A very general solution to IR detection, described by 'Rule 07,' a very convenient heuristic," *J. Electronic Materials* **37**(9), 1406–1410 (2008).
35. F. Pistone, S. Verdet, P.Tribolet, and G. Destefanis, " Progress in MCT large staring array," *Proc. SPIE* **6295**, 62950K (2006).
36. M. A. Kinch, "HDVIP FPA technology at DRS Infrared Technologies," *Proc. SPIE* **4369**, 566–578 (2001).
37. Sofradir website: www.sofradir-ce.com.
38. A. Manissadjian, P. Tribolet, G. Destefanis, and E. DeBorniol, "Long-wave HgCdTe staring arrays at Sofradir: from 9 µm to 13+ µm cut-offs for high performance applications," *Proc. SPIE* **5783**, 231–242 (2005).
39. C. L. Jones, L. G. Hipwood, C. J. Shaw, J. P. Price, R. A. Catchpole, M. Ordish, C. D. Maxey, H. W. Lau, R. C. Mistry, M. C. Wilson, A. D. Parsons, J. Gillespie, L. Baggaley, and M. Wallis, "High-performance MW and LW IR FPAs made from HgCdTe grown by MOVPE," *Proc. SPIE* **6206**, 356–367 (2006).
40. K. Dang, personal communication (2009).
41. D. Lohrmann, personal communications (2010).
42. J. Brown, P. Howard, D. Lohrmann, W. Mason, K. McCormack, and J. Miller, *NVESD Working Handbook for Uncooled Bolometer Detector Technology, Theory and Practice*, U.S. Army NVESD Report, Draft, p. 14 (2010).
43. DRS website: www.drsinfrared.com.
44. BAE USA website: www.baesystems.com.
45. C. Trouilleau, B. Fieque, F. Giner, D. Pochic, A. Durand, P. Robert, S. Cortial, M. Vialain, J. L. Tissot, and J. J. Yon, "High-performance uncooled amorphous silicon TEC less XGA IRFPA with 17-µm pixel-pitch," *Proc. SPIE* **7298**, 72980Q (2009).

Appendix
Observer Vision Model

To accurately model visual performance, the impact of imager blur and noise on human vision must be quantified. Human vision is treated as a "black box" that is characterized by its signal transfer response and detection thresholds. This appendix describes an engineering model of observer vision.[1] It also provides the psychophysical parameters used to characterize human vision.

A.1 Contrast Threshold Function

Contrast threshold function (CTF) and its inverse, the contrast sensitivity function (CSF), quantify the spatial frequency response of human vision. A sine-wave pattern is presented to an observer, and a response is solicited as to whether the sine wave is visible. In Fig. A.1, the observer is viewing a sine-wave pattern. While holding constant the average luminance to the eye, the contrast of the bar pattern is lowered until the pattern is no longer visible to the observer. That is, the dark bars are lightened and the light bars darkened, holding the average constant, until the bar-space-bar pattern disappears. A decrease in contrast from left to right is shown at the top right in the figure. The goal of the experiment is to measure the amplitude of the sine wave that is almost not visible to the observer.

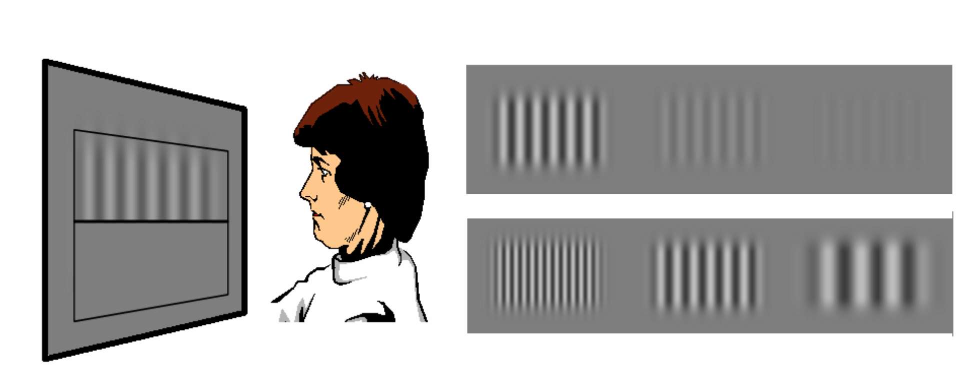


Figure A.1 Experimental setup for measuring CTF/CSF. Top right shows variation in contrast. Bottom right shows variation in spatial frequency.

Although experimental practice varies, one procedure is described here in order to fully explain the CTF concept. CTF data is sometimes taken using two alternative forced-choice (2afc) experiments. In these experiments, the observer is shown one blank field and one with the sine wave. The observer must choose which field has the sine wave. These experiments measure the sine-wave threshold where the observer chooses correctly half the time, independent of chance. That is, the 2afc experiment provides the threshold that yields a 0.75 probability of correct choice. The procedure is repeated for various bar spacings—that is, for various spatial frequencies. See the bottom right of Fig. A.1 for an illustration of spatial frequency; high spatial frequency is at the left, and lower spatial frequency is to the right. The curve of threshold contrast versus spatial frequency for each display luminance is called the CTF at that luminance.

Note that contrast threshold is the sine-wave amplitude at which the observer is correct half of the time, independent of chance. There is a finite probability of seeing the sine wave at reduced contrast, and there is some chance of not seeing the sine wave at contrasts above threshold. The function $\delta(C/\text{CTF})$ describes the probability of seeing a signal with contrast C when eye threshold is CTF. The function δ is available from published data.

A.2 Engineering Model of the Eye

An engineering model of the eye and visual cortex is shown in Fig. A.2. This figure shows the MTF associated with the eyeball and visual cortex. The figure also shows the points where noise is injected into the visual signal.

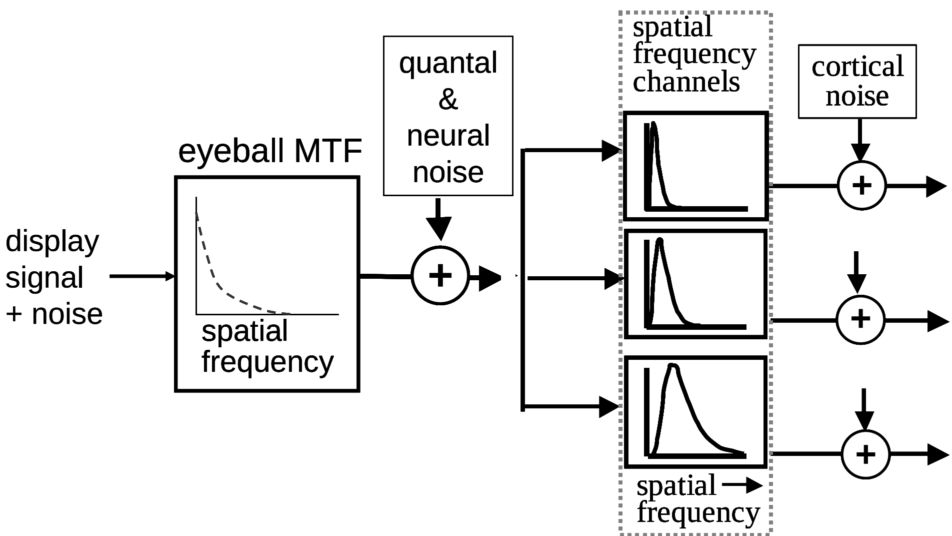


Figure A.2 Engineering model of the eye showing the spatial filters and noise sources acting on the display signal and noise.

Based on the experiments of Nagaraja[2] and others,[3–6] the effect of noise is explained by assuming that the brain is taking the RSS (square root of the sum of squares) of display noise and some internal eye noise. Furthermore, for display luminance above the de-Vries-Rose law region and for foveated and fixated targets, the RSS of eye and display noises occurs in the visual cortex.[6]

In Eq. (A.1), n_{eye} is cortical noise, and σ is display noise filtered by the eyeball and visual cortex bandpass MTF. σ is also appropriately scaled in amplitude. Using Weber's law, eye noise is proportional to display luminance. The parameter α is an empirically established calibration factor. Once α is established through experiments, all of the parameters in Eq. (A.2) are known or measurable, and CTF_{sys} is calculated.

$$CTF_{sys}^2(\xi) = CTF^2(\xi)\left(\frac{n_{eye}^2(\xi)+\sigma^2(\xi)}{n_{eye}^2(\xi)}\right). \tag{A.1}$$

$$CTF_{sys}^2(\xi) = CTF^2(\xi)\left(1+\frac{\alpha^2\sigma^2(\xi)}{L^2}\right). \tag{A.2}$$

Although α is established empirically, the same value is used consistently for all types of imagers and to predict target acquisition probabilities and bar-pattern thresholds. It is true that psychophysical data such as eyeball MTF and naked-eye CTF vary between observers. Adjusting model predictions based on known observer characteristics is, of course, sensible. However, fitting model predictions to data based on assumed variations in the observer obscures all model errors.

Model calculation starts with measured naked-eye thresholds and then estimates the threshold elevation that results from adding imager blur and noise. The target-acquisition models use a numerical approximation to measured naked-eye CTF provided by Barten.[7] The Barten numerical fit is selected based on Beaton's comparison of several numerical approximations to experimental data.[8] The Barten numerical approximation to naked-eye CTF data is given by Eqs. (A.3) through (A.5):

$$CTF(\xi) = \left[a\xi e^{-b\xi}\sqrt{1+0.06e^{b\xi}}\right]^{-1}. \tag{A.3}$$

$$a = 540\left(1+\frac{0.2}{L}\right)^{-0.2} \Bigg/ \left(1+\frac{12}{w^2\left(1+5.8\,\xi\right)^2}\right). \tag{A.4}$$

$$b = 5.24\left(1+\frac{29.2}{L}\right)^{0.15} \tag{A.5}$$

The independent variable is luminance L in fL. The parameter w is display size in degrees. We do not follow Barten in varying w in this formula. Barten varies w in an attempt to predict the effect of display size on image quality. One problem with Barten's method is that the variation in CTF with field of view is based on the data of Carlson.[9] That data is based on an adapting luminance field that is equal to display size; i.e., when the display size is small, the observer is looking at a small bright spot on a dark surrounding field.

Another fundamental problem with varying w is that the resulting CTF approximation does not represent the sine-wave threshold response. In the imager resolution model, CTF is strictly used to represent the measured threshold response of the eye. The intent is to represent vision in frequency space. In the context of the resolution model, degrading CTF because only a few sine-wave cycles are presented to the observer represents an experimental error. A display size of 15 deg presents a sufficient number of cycles to accurately measure sine-wave threshold.[1] The Barten numerical fit with w equal to 15 deg provides the appropriate CTF values. The display field-of-view size is fixed at 15 deg in order to represent Fourier domain eye thresholds. Reference 1 compares Eqs. (A.3)–(A.5) (with w equal to 15 deg) to experimental CTF data.[16–25] The fit is excellent.

Eyeball MTF is also needed to predict the effect of noise on threshold. Formulas to predict eyeball MTF are taken from Stefanik's distillation of the data in Overington.[10,11] Total eyeball MTF is predicted by multiplying optics, retina, and tremor MTF. Optical MTF depends on pupil diameter. Pupil diameter versus light level is given in Table A.1. For each pupil diameter, the parameters i_0 and f_0 are given by Table A.2. Equation (A.6) gives the optics MTF. The total eyeball MTF is therefore the product of optics, retina, and tremor MTF. The numerical eyeball MTF compares favorably with experimental data.[11,26–29]

Table A.1 Pupil diameter in millimeters (mm) versus light level.

diameter	7.0	6.2	5.6	4.9	4.2	3.6	3.0	2.5
log fL	–4	–3	–2	–1	0	1	2	3

Table A.2 Parameters for optics MTF.

Pupil diameter (mm)	1.5	2.0	2.4	3.0	3.8	4.9	5.8	6.6
f_0	36	39	35	32	25	15	11	8
i_0	0.9	0.8	0.8	0.77	0.75	0.72	0.69	0.66

$$MTF_{optics} = \exp\left[-(\xi / f_0)^{i_0}\right]. \tag{A.6}$$

$$MTF_{retina} = \exp(-0.375\,\xi^{1.21}) . \tag{A.7}$$

$$MTF_{tremor} = \exp(-0.4441\,\xi^2) . \tag{A.8}$$

The visual cortex bandpass filters $B(\xi)$ are taken from Barten,[12] who created a numerical fit for the visual cortex filters by using psychophysical data. See Eq. (A.9). Again, this is a numerical fit and not a theoretical result. In Eq. (A.9), ξ is the frequency of the sine-wave grating, and ξ' is a dummy variable used to integrate over noise bandwidth.

Reference 1 compares Eq. (A.10) CTF [using the Eq. (A.9) visual cortex filters] to experimental data.[30,31] The data are widely scattered and show inconsistent behavior between subjects. An absolute match between the observer model and predictions is not expected, as the measured data are too scattered. However, the observer model does predict the effect of nonwhite noise on CSF. In all cases, the data trends are predicted by the model. The model accurately predicts the frequencies at which nonwhite noise affects CSF. Onset and cessation of CSF degradation is correctly predicted.

$$B(\xi') = \exp\left\{-2.2\left[\log(\xi' / \xi)\right]^2\right\} . \tag{A.9}$$

Equation (A.10) predicts CTF_{sys} for horizontal gratings viewed through an imager. A similar formula is used for vertical gratings.

$$CTF_{sys}(\xi) = \frac{CTF(\xi / SMAG)}{H_{sys}(\xi)}\left(1 + \frac{\alpha^2\sigma^2(\xi)}{L^2}\right)^{1/2}, \tag{A.10}$$

where

α = 169.6 root-Hz,
SMAG = system magnification,
σ = noise affecting threshold at grating frequency ξ,
$\rho(\xi,\eta)$ = noise spectral density in fL sec$^{1/2}$ mrad,
$H_{sys}(\xi)$ = system MTF from scene through display,
$H_{eye}(\xi)$ = eyeball MTF,
$B(\xi)$ = filters in the visual cortex, and
$D(\xi)$ = MTF of display blur.

$$\sigma^2(\xi) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|B(\xi'/\xi)\, D(\xi')H_{eye}(\xi')\right|^2 \left|D(\eta)H_{eye}(\eta)\right|^2 \rho^2(\xi,\eta)\, d\xi' d\eta \,. \quad \text{(A.11)}$$

The observer model pertains to cone vision of foveated targets. Cones mediate vision down to about 0.01 fL. An observer at night might set his display as low as 0.1 fL. This low display luminance is a compromise between maintaining dark adaptation and effective viewing of display information. Therefore, the display luminance levels of interest here vary from about 0.1 fL to several hundred fL. However, there is little variation in visual thresholds above about 100 fL.

The visual system filters temporally as well as spatially. However, an explicit adjustment for the variation in temporal integration of the eye is not included in Eq. (A.10). We do not know *a priori* whether variations in temporal integration alter the relationship between luminance, display noise, and cortical noise. Although increasing temporal integration certainly applies additional filtering to display noise, it also increases the gain applied to both display signal and noise.

Furthermore, no clear evidence exists that cone temporal integration is light-level dependent. Cone temporal integration does not change with a variation of photopic light level down to about 10 fL.[13] At photopic light levels, adaptation is by pigment bleaching.[14] Data is not available on cone temporal integration below 10 fL. It is certainly true, however, that most of the observed increase in temporal integration at low luminance is due to the rod system. Rods begin to come out of saturation at 10 fL. Furthermore, regardless of what is actually happening physiologically, the match between the Eq. (A.10) model and image intensifier data suggests that a temporal adjustment is not needed in Eq. (A.10).[15]

This does not mean, however, that the eye treats static and dynamic noise equally. This is a change in the nature of display noise, not a change in the visual system. For a nonframing imager, noise is integrated for a dwell time, not for the cone integration time. For a framing imager with frame rate F_R sec^{-1}, single frame noise is $\sqrt{(0.04F_R)}$ more effective than dynamic noise at masking signal. The value 0.04 sec for cone integration is taken from Ref. 13.

The match between model and data is excellent considering the experimental errors.[1] The model predicts the impact of display luminance on CSF. The model also predicts the relative impact of noise at one frequency on the CSF at a different frequency. In all cases, the model either accurately predicts the data or at least predicts the trends in observed behavior.

References

1. R. H. Vollmerhausen, “Representing the observer in electro-optical target acquisition models,” *Opt. Exp.* **17**, 17253–17268 (2009).
2. N. S. Nagaraja, “Effect of luminance noise on contrast thresholds,” *J. Opt. Soc. Am.* **54**, 950–955 (1964).

3. D. G. Pelli, "Effects of Visual Noise," Doctoral dissertation, Physiological Laboratory, Churchill College, Cambridge University (1981).
4. G. E. Legge, D. Kersten, and A. E. Burgess, "Contrast discrimination in noise," *J. Opt. Soc. Am*. **4**, 391–404 (1987).
5. D. G. Pelli, "Why use noise?" *J. Opt. Soc. Am. A* **16**(3), 647– 653 (1999).
6. M. Raghavan, "Sources of Visual Noise," Ph.D. dissertation, Syracuse University, NY (1989).
7. P. G. J. Barten, "Formula for the contrast sensitivity of the human eye," *Proc. SPIE* **5294**, 231–238 (2004).
8. R. J. Beaton and W. W. Farley, "Comparative study of the MTFA, ICS, and SQRI image quality metrics for visual display systems," Report AL-TR-1992-0001, DTIC ADA252116, Armstrong Lab., Air Force Systems Command, Wright-Patterson AFB, OH (1991).
9. C. R. Carlson, "Sine-wave threshold contrast-sensitivity function: dependence on display size," *RCA Rev.* **43**, 675–683 (1982).
10. R. J. Stefanik, "Performance modeling for image intensifier systems," Report NV-93-14, Night Vision and Electronic-Sensors Directorate, U.S. Army Research, Development, and Engineering Command, Fort Belvoir, VA (1993).
11. I. Overington, *Vision and Acquisition*, Chapters 1, 2, and 4, Crane, Russak & Co. (1976).
12. P. G .J. Barten, *Contrast Sensitivity of the Human Eye and Its Effect on Image Quality*, SPIE Press, Bellingham, WA (1999).
13. R. A. Moses and W. M. Hart, "The temporal responsiveness of vision," in *Adler's Physiology of the Eye: Clinical Application*, Mosby Publishing, London (1987).
14. H. Davson, *Physiology of the Eye*, 5th ed., pp. 221 & 271, Macmillan Academic and Professional Ltd. (1990).
15. R. H. Vollmerhausen, "Predicting the effect of gain, level, and sampling on minimum resolvable temperature measurements," *Opt. Eng*. **48**, 076401 (2009).
16. A. van Meeteren and J. J. Voss, "Resolution and contrast sensitivity at low luminance levels," *Vision Research* **12**, 825–833 (1972).
17. J. J. DePalma and E. M. Lowry, "Sine wave response of the visual system. II. Sine wave and square wave contrast sensitivity," *J. Opt. Soc. Am*. **52**, 328–335 (1962).
18. A. Watanabe, T. Mori, S. Nagata, and K. Hiwatashi, "Spatial sine wave response of the human visual system," *Vision Research* **9**, 1245–1263 (1968).
19. A. S. Patel, "Spatial resolution by the human visual system. The effect of mean retinal luminance." *J. Opt. Soc. Am*. **56,** 689–694 (1966).
20. F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physio.* **197**, 551–566 (1968).
21. Z.-L. Lu and B. A. Dosher, "Characterizing the spatial-frequency sensitivity of perceptual templates," *J. Opt. Soc. Am*. **18**, 2041–2053 (2001).

22. K. R. Boss and J. E. Lincoln, *Engineering Data Compendium: Human Perception and Performance*, Vol. 1, Harry G. Armstrong Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio (1988).
23. V. Virsu and J. Rovamo, "Visual resolution, contrast sensitivity, and the cortical magnification factor," *Experimental Brain Research* **37**, 475–494 (1979).
24. F. W. Campbell and R. W. Gubisch, "Optical quality of the human eye," *J. Physio*. **186**, 558–578 (1966).
25. C. R. Carlson, "Sine-wave threshold contrast-sensitivity function: dependence on display size," *RCA Rev.* **43**, 675–683 (1982).
26. F. L. Van Nes and M. A. Bouman, "Spatial modulation transfer in the human eye," *J. Opt. Soc. Am*. **57**, 401–406 (1967).
27. J. Rovamo, H. Kukkonen, and J. Mustonen, "Foveal optical modulation transfer function of the human eye at various pupil sizes," *J. Opt. Soc. Am. A* **15,** 2504–2513 (1998).
28. A. van Meeteren, "Calculations on the optical modulation transfer function of the human eye for white light," *Optica Acta* **21**, 395–412 (1974).
29. P. Artal and R. Navarro, "Monochromatic modulation transfer function of the human eye for different pupil diameters: an analytic expression," *J. Opt. Soc. Am*. **11**, 246–249 (1994).
30. A. van Meeteren, and J. Valeton, "Effects of pictorial noise interfering with visual detection," *J. Opt. Soc. Am. A* **5**, 438–444 (1988).
31. C. F. Stromeyer and B. Julesz, "Spatial frequency masking in vision: critical bands and spread of masking," *J. Opt. Soc. Am*. **62**, 1221–1232 (1972).

Index

Q

R

S

T

Richard H. Vollmerhausen currently consults in the areas of electro-optical systems analysis and modeling. He retired as head of the Model Development Branch at the U.S. Army's Night Vision Lab (NVL). During his tenure at NVL, the branch developed and validated a target acquisition metric to replace the traditional Johnson criteria. His experience includes project engineer and EO systems analyst for numerous Army weapon systems. His previous work included designing air-to-air missile seekers for the Navy and working as an instrumentation engineer for Douglas Aircraft on the Saturn/Apollo program. Mr. Vollmerhausen is the author of two books on electro-optical systems analysis and has published numerous journal and symposium papers.

Donald A. Reago, Jr. has 24 years experience in electro-optics gained through his work at the Army Night Vision Lab (now NVESD). His background includes development on infrared systems, component development technology, and signal processing/multisensor fusion for automatic target recognition. He currently serves as the Deputy Director for Technology and Countermine at NVESD.

Ronald G. Driggers received a doctorate in electrical engineering from the University of Memphis in 1990. Dr. Driggers has 23 years of electro-optics experience and has worked or consulted for Lockheed Martin (Orlando), SAIC, EOIR Measurements, Amtec Corporation, Joint Precision Strike Demonstration Project Office, Redstone Technical Test Center, and Army Research Laboratory. He was recently appointed to the Senior Executive Service as the Superintendent of the Optical Sciences Division at the Naval Research Laboratory. Previously, he was the Director of the Modeling and Simulation Division at the U.S. Army's Night Vision and Electronic Sensors Directorate (NVESD). Dr. Driggers is the author of four books on infrared and electro-optics systems and has published over 100 research papers. He was Editor-in-Chief of the *Encyclopedia of Optical Engineering* (Taylor and Francis). He was selected as the 2002 Army Materiel Command's Engineer of the Year, the 2001 CERDEC Technical Employee of the Year, and the 2001 NVESD Technical Employee of the Year. He is a U.S. Naval Reserve Officer and was selected as the 2001 Naval Engineering Duty Officer of the Year (William Kastner Award). He is also a Fellow of SPIE, the Optical Society of America, and the Military Sensing Symposium. In January 2010, Dr. Driggers took over duties as Editor-in-Chief of SPIE's flagship journal, *Optical Engineering*.

Richard H. Vollmerhausen, Donald A. Reago, Jr., Ronald G. Driggers, *Analysis and Evaluation of Sampled Imaging Systems*, SPIE Press, Bellingham, WA (2010).

If the above button does not work, the supplemental files for this eBook are also available for download by copying and pasting the following link to a browser:

http://spie.org/Samples/Pressbook_Supplemental/TT87_sup.zip