PHOTONICS RULES OF THUMB

Third Edition

John Lester Miller, Ed Friedman, Jack Sanders-Reed,
Katie Schwertz, and Brian McComas

SPIE PRESS
Bellingham, Washington USA

Contents

Chapter 1
Astronomy

Introduction

This chapter contains a selection of rules specifically involving the intersection of astronomy and electro-optics (EO). Sensors frequently look upward, so astronomical objects often define the background for many systems. Moreover, many sensors are specifically designed to detect heavenly bodies, so astronomical relationships define the targets for many sensors.

Over the past few hundred years, astronomy has driven advances in optics, and then in photonics and optics in merged systems for astronomy. These disciplines have been as interwoven as DNA strands. Frequently, key discoveries in astronomy are impossible until photonic technology develops to a new level. Conversely, photonic development has often been funded and refined by the astronomical sciences as well as the military.

Military interests have been an important source of new technology that has furthered the application of EO in astronomy. The authors contend that one of the most important contributions of the Strategic Defense Initiative (SDI) and its subsequent versions was the advancement of certain photonic technologies that currently benefit astronomers, e.g., space-based telescopes, adaptive optics, synthetic guide stars, large and sensitive focal planes, advanced materials for astronomical telescopes, new methods of image stabilization, and advanced computers and algorithms for interpreting images distorted by atmospheric effects.

Recent developments in the size and technology of ground-based telescopes allow them to compete with space-based observation capabilities (except in spectral regions where the atmosphere strongly absorbs or scatters). The dual Keck 10-m telescopes represent an amazing electro-optical engineering achievement that is being surpassed in size by several telescopes. Some initial compromises in the field of view (FOV) over which atmospheric correction can be obtained are expected, but work continues to overcome these issues.[1] By employing segmented lightweight mirrors and lightweight structures, and by adjusting the mirrors in real time, many of the past notions and operating paradigms of both ground-based and space-based telescopes have been discarded. This new emphasis on novel technology applied to Earth- and space-based telescopes represents a major addition to the astronomical community's toolbox and a shift in the electro-optical and astronomical communities' perceptions.

In the near future, these high-technology telescopes, coupled with advanced precision instruments, will provide astronomers with new tools to make new and wondrous discoveries. There is no inherent reason why the technologies used in ground telescopes

It is likely an accident that the peak of the Sun's radiation is well matched to a major transmission window of the atmosphere. On the other hand, it is no accident that the peak performance of the human vision system is well matched to the solar radiation that reaches the ground. Evolution of visual systems has ensured that performance is best around 555 nm. Due to the absorption properties of the atmosphere, the Sun deviates significantly from blackbody properties when seen from the ground.

References

1. E. A. Gurtovenko and V. A. Sheminova, "Formation depths of Fraunhofer lines," arxiv.org/pdf/1505.00975.pdf (5 May 2015).
2. E. Hecht, *Optics*, 2nd Edition, p. 234, Addison Wesley (1987).
3. Wikipedia, "Fraunhofer lines," en.wikipedia.org/wiki/Fraunhofer_lines (2020).

Number of Stars as a Function of Wavelength

For a given sensitivity at visible wavelengths and beyond, the longer the wavelength is, the fewer stars can be sensed. The falloff in the number of stars is approximated by the following:

$$\#S\lambda_2 \approx \#S\lambda_1 \times 10^{-0.4R},$$

where $\#S\lambda_2$ is the number of stars at wavelength λ_2 (λ_2 larger than λ_1) at a given irradiance, R is the ratio of one wavelength to another (λ_2/λ_1), and $\#S\lambda_1$ is the number of stars at wavelength λ_1.

Discussion

This rule is based on curve-fitting empirical data by the authors. It is useful for separate narrow bands, from about 0.7–15.0 microns. Generally, this provides accuracy to within a factor of 2. Most stars are like our Sun and radiate most of their energy in the visible part of the spectrum. As the wavelength increases, there are fewer stars, because the Planck function is dropping for the stars that peak in the visible, and fewer stars have peak radiant output at longer wavelengths.

A Simple Model of Stellar Populations

The total number of stars above a given visual magnitude M_V can be estimated from a numerical model derived from measured data:[1]

$$\#S = 11.84 \times 10^{0.4204 M_V},$$

where #S is the approximate number over the whole sky.

Discussion

This simple rule is accurate to within a factor of 3 between magnitudes 0 and 18.5. It provides a good match—no worse than a factor of 5 for most magnitudes. A reminder,

where the equation specifically notes the approximate blackbody temperature of the Sun (5900 K). The notation $L^{bb}(\lambda, 5900\ \mathrm{K})$ describes the blackbody radiance of the Moon as a function of wavelength λ for a temperature of 5900 K. R_m is the reflectivity of the Moon, which has typical values of 0.1 in the visible wavelengths, 0.3 for 3–6 microns, and 0.02 for 7–15 microns.[1] Ω is the solid angle formed by the Moon when viewed from Earth. In the infrared (IR), the Moon's thermal emission must also be considered as it can be the dominant source of photons in some bands.

Discussion

The Moon is an important (and sometimes dominant) source of radiation in the night sky. Its signature includes radiation ranging from the visible to the IR, so all types of sensors must be designed to tolerate its presence. Many sensors (such as image intensifiers, low-light-level cameras, and modern consumer cameras) can exploit this light.

The light seen when viewing the Moon is the superposition of emitted radiation, reflection of solar radiation, and emission from the atmosphere:

$$L_{\mathrm{moon}}(\lambda) = \tau(\lambda)\left[L_{\mathrm{reflected}}(\lambda) + L_{\mathrm{emitted}}(\lambda)\right] + L_{\mathrm{atm}}(\lambda),$$

where τ is the transmission of the atmosphere, L_{emitted} is the radiance of the Moon, and L_{atm} is the radiance of the atmospheric radiance of Earth's atmosphere at the lunar bands in question.

The IR signature from the full Moon is defined by its apparent blackbody temperature of 390 K. Anyone using the following equation should note that the actual temperature of the Moon depends on the time elapsed since the location being imaged was last illuminated by the Sun. This can result in a substantial difference from the following equation, but it is good enough if you don't know the details of the lunar ephemeris:

$$L_{\mathrm{emitted}} = \varepsilon(\lambda) L^{\mathrm{bb}}(\lambda,\ 390\ \mathrm{K}).$$

Nighttime temperatures are around 95 K[2]. The spectral emissivity $\varepsilon(\lambda)$ in the equation above can be estimated by using the reflectivity numbers quoted previously, remembering that $1 - R = \varepsilon$.

As a result of changes in the distance from Earth to the Moon, the solid angle of the Moon seen from Earth is

$$\Omega = 6.8 \times 10^{-5}\ \mathrm{sr}\ [\text{with variation from } 5.7 \times 10^{-5} \text{ to } 7.5 \times 10^{-5}].$$

It is an accident of nature that the angular size of the Moon is close to the angular size of the Sun. Under rare conditions, the Moon's position in its orbit results in a lunar angular size that almost exactly matches the angular size of the Sun. If an eclipse occurs under these conditions, the entire disk of the Sun is blocked, and it is possible to view the solar corona. If the Moon is slightly farther away from Earth during an eclipse, then the entire disk of the Sun is not covered, and an annular eclipse occurs.

References

1. J. Shaw, "Modeling infrared lunar radiance," *Opt. Eng.* **38**(10), 1763–1764 (1999).

Chapter 2
Atmospherics

Introduction

It is hard to imagine a subject more complex, and yet more useful, than the study of the propagation of light in the atmosphere. Because of its importance in a wide variety of human enterprises, considerable attention has been paid to this topic for several centuries. Initially, the effort was dedicated to learning how the apparent size and shape of distant objects depend on the properties of the atmosphere. The advent of modern science provided new tools for measuring and theorizing about the optical properties of the atmosphere. Much of the motivation for progress was related to the military, astronomy, and navigation at sea. Modern applications are still important, and the prospect of laser communication links has provided continued impetus for progress.

Maturation of the field of spectroscopy led to a formal understanding of the absorption spectra of significant atmospheric species and their variation with altitude. Computer models[1–3] that include virtually all that is known about the absorption and scattering properties of atmospheric constituents have been assembled and can provide very complete descriptions of transmission as a function of wavelength with a spectral resolution of about 1 cm^{-1}. This is equivalent to a wavelength resolution of 0.1 nm at a wavelength of 1 micron. In addition to the current chapter, the reader may wish to review the Degraded Visual Environments chapter, which provides atmospheric absorption curves and discussion of the effects of various particulate obscurants.

In addition to gradually refining our understanding of atmospheric absorption by considering the combined effect of the constituents, there is also a rather complete and elaborate theory of scattering in the atmosphere. These results have been extended greatly by computer modeling, particularly in the field of multiple scattering and Monte Carlo methods. For suspended particulates of known optical properties, reliable estimates of scattering properties for both plane and spherical waves can be obtained for conditions in which the optical thickness is not too large. Gustav Mie (1868–1957) was particularly influential, as he was the first to use Maxwell's equations to compute the scattering properties of small spheres suspended in a medium of another index of refraction. A number of references suggest that Mie was not the first to solve the problem but was the first to publish the results. His work, along with that of Debye, is now generally called

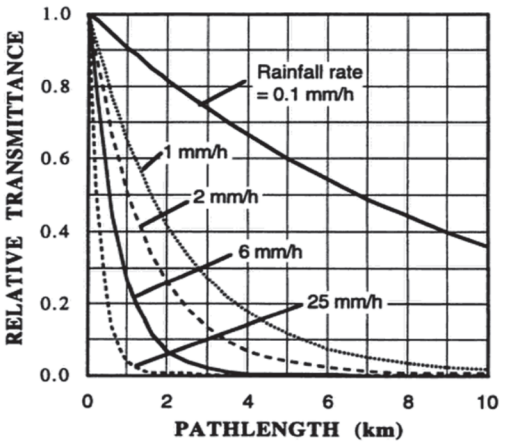

Figure 2.2 Transmittance of 10.6-μm IR radiation in rain.

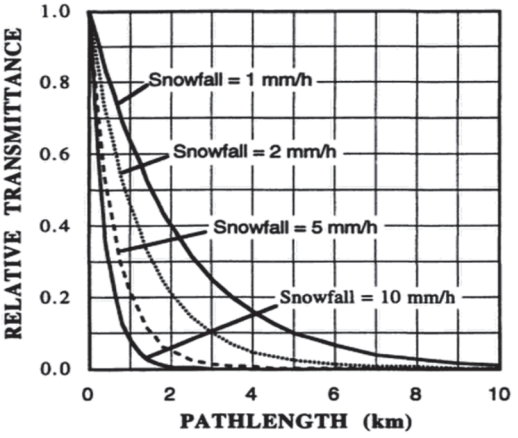

Figure 2.3 Transmittance of 10.6-μm IR radiation in snow.

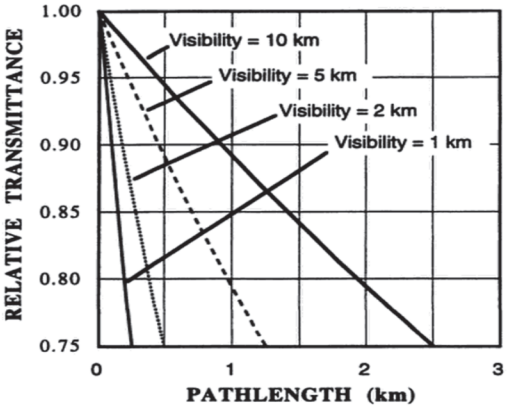

Figure 2.4 Transmittance of 10.6-μm IR radiation in dust with different visibilities reported by a person based on visual sighting of distant objects.

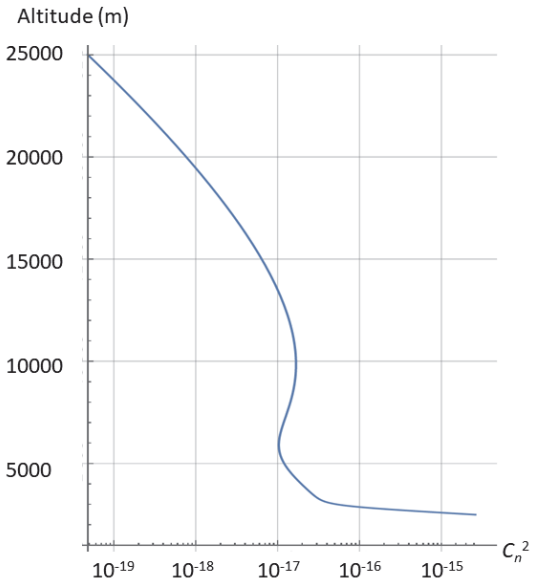

Figure 2.6 An example graph of the C_n^2 for the HV 5/7 model. Altitude is in meters.

$$C_n^2(h) = 8.2\times10^{-26}\left(\frac{h}{1000}\right)^{10} W^2 e^{-h/1000} + 2.7\times10^{-16} e^{-h/1500} + 1.7\times10^{-14} e^{-h/100},$$

where h is the height in meters, and W is the wind correlating factor, which is selected as 21 for the HV 5/7 model.[8] Another version of HV is provided by[8]

$$C_n^2(h) = A\left[2.2\times10^{-23}\left(\frac{V}{V_o}\right)^2 h^{10} e^{-h} + 10^{-16} e^{-h/1.5} + 1.7\times10^{-14} e^{-h/0.1}\right],$$

where A is approximately unity, V is the upper-atmosphere wind speed, and V_o is the mean upper-atmosphere wind speed.

In many cases, the C_n^2 value can be crudely approximated as simply 1×10^{-14} during the night and 2×10^{-14} during the day. The R384 database[3] has a minimum value of 7.11×10^{-19} and a maximum value of 1.7×10^{-13} for ground-based, horizontal measurements. The R384 average is almost exactly 1×10^{-14}.

Another model is the CLEAR 1,[7] which pertains to the New Mexico desert. It was obtained by averaging and statistically interpolating a number of radiosonde observation measurements obtained over a large number of meteorological conditions. The lower limit of 1.23 km of ground elevation is related to the altitude of New Mexico, where the measurements took place:

$$C_n^2(h) = 10^{-17.025-4.3507h+0.8141h^2} \text{ for } 1.23 < h \le 2.13 \text{ km},$$

$$C_n^2(h) = 10^{-16.2897+0.0335h-0.01341h^2} \text{ for } 2.13 < h \le 10.34 \text{ km},$$

$$C_n^2(h) = 10^{-17.0577-0.0449h-0.00051h^2+0.6181\exp(-0.5(h-15.5617)/12.0173)} \text{ for } 10.34 < h \le 30 \text{ km}.$$

that result in a meaningful value of E_T. For example, we expect α to be in km^{-1}, so V must be in km. Other units can be used so long as the product of the distance and attenuation coefficient is unitless. For E_T to be in W/m^2, the units of V in the denominator must be in meters. The equation above derives from a type of beam spread model, coupled with an attenuation term.

Reference 6 describes a measure of pilots viewing down a runway as the visual threshold E_t(lux), which is approximated by

$$\log E_t = -5.7 + 0.64 \log B,$$

where B is the background luminance (cd/m^{-2}). A lower threshold on E_t is set at 6.8×10^{-6} lux, the night limit for background luminance.

Hudson[3] shows that the effect of rain on visual range and scattering coefficient σ can be estimated from

$$\sigma_{rain} = 1.25 \times 10^{-6} R / r^3,$$

where R is the rainfall rate (cm/sec), and r is the radius of the drop (cm).

Alternatively, Ref. 1 gives the scattering coefficient of rainfall as

$$\sigma_{rain} = 0.248 f^{0.67},$$

where f is the rainfall rate in mm/hr.

Oakley and B. Satherley[5] provide some insight into the effect of aerosols into scattering in the atmosphere. They point out that, for a uniform distribution of particles of concentration D and radius a, the scattering coefficient is

$$\sigma_{rain} = D \pi a^2 Q_{sc},$$

where Q_{sc} is the Mie scattering coefficient, which is a complicated function[9,10] of the ratio $(2\pi a)/\lambda$.

As a increases, either by considering scattering at shorter wavelengths or by increasing the aerosol size, Q_{sc} becomes 2. The result is that for a large particle size or short wavelength, the particles have a scattering cross-section twice their geometric size.

For non-monochromatic radiation, the exponential relationship mentioned earlier in this rule may not apply. For example, we include an example using MWIR data for the band from 3–5 microns.[10] The typical exponential decay in transmission with distance is not observed. For a band as large as the one shown in Fig. 2.8, this effect results in a more complicated relationship between intensity and distance.

For a more extensive discussion of atmospheric effects on visibility, refer to the chapter on Degraded Visual Environments.

References

1. *Burle Electro-Optics Handbook,* Burle Industries, Lancaster, PA, p. 87, available at ftp://ftp.bartol.udel.edu/jmc/pmt/rca_eo_hb.pdf (2020).
2. en.wikipedia.org/wiki/Visibility (2020).
3. R. Hudson, *Infrared Systems Engineering,* John Wiley & Sons, pp. 161–165 (1969).

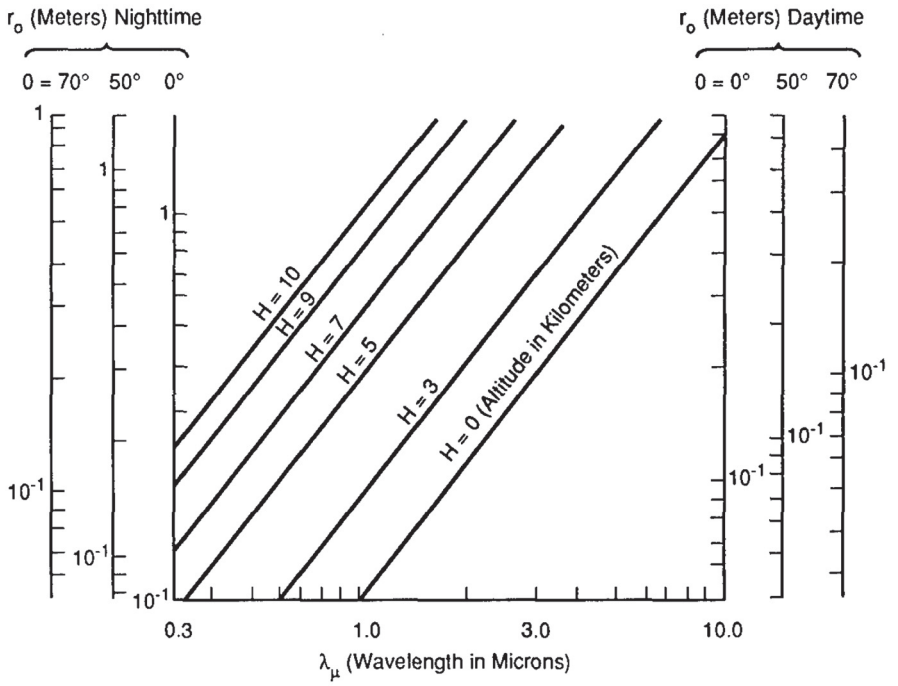

Figure 2.10 Coherence length r_0 as a function of wavelength at various altitudes for day and night conditions.[1]

Phase Error Estimation

The maximum phase error induced by the atmosphere can be written as

$$maximum\ phase\ error = 0.57k\sqrt{LC_n^2 D^{5/3}}.$$

Here, L is the distance through which the aberrated wave propagates, C_n^2 is the atmospheric structure function, D is the aperture diameter, and k is the wave propagation constant $2\pi/\lambda$.

Discussion

The stroke of actuators in a deformable mirror system must be able to accommodate the phase errors induced by the atmosphere and aberrations in the telescope. The rule shown here compares with a discourse by Tyson[2] in which he shows the phase effects of various optical aberration terms induced by the atmosphere. Some algebra shows that what is shown above (as well as the terms described by Tyson) can be put in a form that includes the ratio of the aperture diameter and the Fried parameter r_0. We can compare the various results using the following equation:

$$\sigma^2 = n\left(\frac{D}{r_0}\right)^{5/3}.$$

Discussion

Not surprisingly, the performance of an adaptive optics system depends on the quality and number of aberration modes detected and corrected. Winocur[1] has data on the correction coefficients for the first 21 Zernike modes and a formula for the higher-order modes.[1,5]

Table 2.8 shows how the number of corrective modes improves the resulting Strehl ratio (SR) performance. The shaded cell entries are the values of $-\sigma_M^2$ for two different parameters: (1) the levels of aberration correction measured as the number of aberration modes M that are corrected, and (2) the relative size of the aperture to the Fried parameter D/r_0. M is provided in the left column. In each of the shaded cells,

$$\sigma_M^2 = weighting\ coefficient \cdot \left(\frac{D}{r_0}\right)^{5/3}.$$

Table 2.8 Parameters used to determine the aberrations induced by the atmosphere. Noll[5] provides the weighting coefficients.

Modes Corrected M	Weighting Coefficient $n(M)$	Zernike Aberration Names	Ratio of the Aperture to Fried Parameter (D/r_0)		
			1	5	10
1	1.0299	Piston	0.36	0.00	0.00
2	0.5820	1D Tilt x	0.56	0.00	0.00
3	0.1340	2D Tilt y	0.87	0.14	0.00
4	0.1110	Defocus	0.89	0.20	0.01
5	0.0880	Astigmatism and defocus	0.92	0.28	0.02
6	0.0648	Astigmatism and defocus	0.94	0.39	0.05
7	0.0587	Coma and x-tilt	0.94	0.42	0.07
8	0.0525	Coma and y-tilt	0.95	0.46	0.09
9	0.0463	Coma	0.95	0.51	0.12
10	0.0401	Coma	0.96	0.56	0.16
11	0.0377	Third-order spherical	0.96	0.58	0.17
12	0.0352		0.97	0.60	0.20
13	0.0328		0.97	0.62	0.22
14	0.0304		0.97	0.64	0.24
15	0.0279		0.97	0.67	0.27
16	0.0267		0.97	0.68	0.29
17	0.0255		0.97	0.69	0.31
18	0.0243		0.98	0.70	0.32
19	0.0232		0.98	0.71	0.34
20	0.0220		0.98	0.72	0.36
21	0.0208		0.98	0.74	0.38

Adaptive Optics Influence Function

Actuation of an element of an adaptive optics array will influence neighboring actuators. The magnitude of this effect (which is characterized by a single number, the mirror fitting parameter), will determine the wavefront correction performance of the array, as described here. Optimal performance is achieved by statistical methods.

Discussion

The wavefront variance σ^2 resulting from the use of adaptive optics depends on the atmosphere through which the light is propagating (through the Fried parameter r_0), and d is the spacing of the actuators and the influence function according to

$$\sigma^2_{fit\ for\ minimum\ error} = \alpha_i \left(d / r_0 \right)^{5/3} = \sigma^2_{fit/\min}.$$

We can find the Strehl ratio reduction for the mirror fitting error according to Table 2.9.

Table 2.9 Mirror fitting parameter α_i.[1]

0.229	Gaussian
0.284	Pyramid
0.399	Itek mirror
1.297	Piston only
0.319	Tyson (Gaussian)
0.141	Greenwood and Fried
0.134	Noll
0.130	Fried

For example, a Gaussian fitting function (Tyson's value of 0.319) applied to adaptive optics with 3-cm actuator spacing in an atmosphere where r_0 = 10 cm results in a Strehl ratio of

$$SR = \exp\left(-\sigma^2_{fit/\min} \right) = \exp\left[-0.319 \left(3/10 \right)^{5/3} \right] = 0.87.$$

The structure of the equations shows that increasing the density of actuators (smaller d) and/or better atmospheric conditions (larger r_0) increases the Strehl ratio.

Reference

1. S. R. Robinson, Ed., *Emerging Systems and Technologies*, Vol. 8 of *The Infrared and Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, Eds., Infrared Information Analysis Center and SPIE Press (1993).

Shack–Hartmann Noise

The variance (square of the standard deviation) in the wavefront resulting from non-perfect sensing in a Shack–Hartmann sensor can be expressed (in radians^2) as[1]

References

1. R. Byren, "Laser Rangefinders," p. 103, in *Active Electro-Optical Systems*, C. Fox, Ed., Vol. 6 of *The Infrared and Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, Eds., Infrared Information Analysis Center and SPIE Press (1993).
2. J. Isterling, "Electro-Optical Propagation thorough highly aberrant media," p. 43, Ph.D. Dissertation, University of Adelaide, digital.library.adelaide.edu/au/dspace/bitstream/2440/64292/8/02whol.pdf (2020).

Free-Space Link Margins

The atmosphere has a distinct impact on the ability of terrestrial laser communications. The data in Table 2.10 for wavelengths of 1550 nm indicates the relative impact of different conditions.

Discussion

Atmospheric absorption, scatter, and scintillation will all decrease the SNR and, if bad enough, will eliminate the ability of an electro-optical system to detect a target or send information to another location.

Table 2.10 provides some guidelines for the link margins suitable for various weather types. As in all weather conditions, the real world can exhibit both spatial and temporal variation, so some care must be taken when using the numbers in the table.

The practitioner is encouraged to get the local weather statistics for his link to determine the link margin needed for a give locale. Obviously, Adelaide and Tucson will need a lower margin for a given reliability than Seattle or Halifax.

Visible wavelengths perform slightly worse, and the long-wave infrared (LWIR) slightly better. Carlson[1] gives the first five entries; the last entry is from the author.

Other rules in this chapter, the Appendix, and the DVE chapter provide additional information about how attenuation varies with rainfall or snowfall rates.

Reference

1. R. Carlson, "Reliability and Availability in Free Space Optical Systems," *Optics in Information Systems* **12**(2) (2001).

Table 2.10 Suitable link margins.

Weather Condition	Required Link Margin (dB/km)
Urban haze	0.5
Typical rainfall	3
Heavy rainfall	6
Typical snow, heavy downpour, or light fog	10
White-out snowfall or fog	20
Heavy-to-severe fog	30–120

Chapter 3
Acquisition, Tracking, and Pointing

Introduction

The fundamental purpose of an *acquisition, tracking, and pointing* (ATP) system is to keep a moving target in the FOV of a sensor, allowing the development of a time history of the object. As such, ATP is a system-level problem involving sensors, algorithms to detect targets, algorithms to assign multiple detections to multiple active target tracks (detect-to-track assignment), control loops to drive a pointing system (typically gimbals) to maintain the FOV centered on the target, and finally the gimbals themselves. Successful development of an ATP system requires expertise in optics, sensors, algorithms, real-time processing, control loops, and gimbal systems. Note that this is very different from an *intelligence, surveillance, and reconnaissance* (ISR) system, which simply gathers imagery of an area with functions such as detection, recognition, and identification (DRI) performed at least quasi-independently (although rules for DRI are included in this chapter). While functions such as ISR or "scene understanding" can benefit from temporal data, they are not fundamentally time-based functions, whereas ATP at its core is a temporal process. While ATP usually involves a pointing system, a special case involves the use of a fixed line of sight (LOS) sensor collecting a video sequence. In this case, the target is detected and tracked as its image moves across the image (focal plane).

ATP as a field of study has its roots in radar processing, and in fact, many detection and tracking algorithms were first developed in the radar community and later adapted to electro-optical (visible) and infrared (EO/IR) systems. Radar systems usually emit a scanning beam, which has implications for data rates and coverage rates compared to today's staring array focal planes, which provide faster coverage and high data rates. Early EO/IR systems commonly used scanning techniques with one or a few pixels scanned across the scene to build an image. With these early systems, the adaptation of radar detection algorithms was relatively straightforward due to the similarity between a scanning radar beam and a (passive) scanning EO/IR sensor. As EO/IR technology evolved to produce high-pixel-count staring array focal plane arrays (FPAs), while many of the radar detection principals are still applicable, the algorithms have needed to be re-cast to more accurately represent the simultaneous collection of millions of pixels (as opposed to the sequential collection of a scanning system).

As discussed in the "Tracker vs. Detection" rule in this chapter, we often separate detection and tracking into two independent functions, but they are ultimately a single

number of hits [a constant false alarm rate (CFAR) process], which should correspond to the maximum number of hits that can be processed in a given timeframe.

The above M out of N test is one form of a multiple hypothesis tracker (MHT), widely regarded as the "best" detect-to-track assignment algorithm.[3] An MHT is distinguished by the fact that all detections are assigned to all tracks with a probability based on how well the particular detection matches a given track. In principle, all possible trajectories are evaluated, but in practice, this rapidly leads to an unmanageable number of combinations, and hence various techniques are used to reduce the number of possible trajectories.

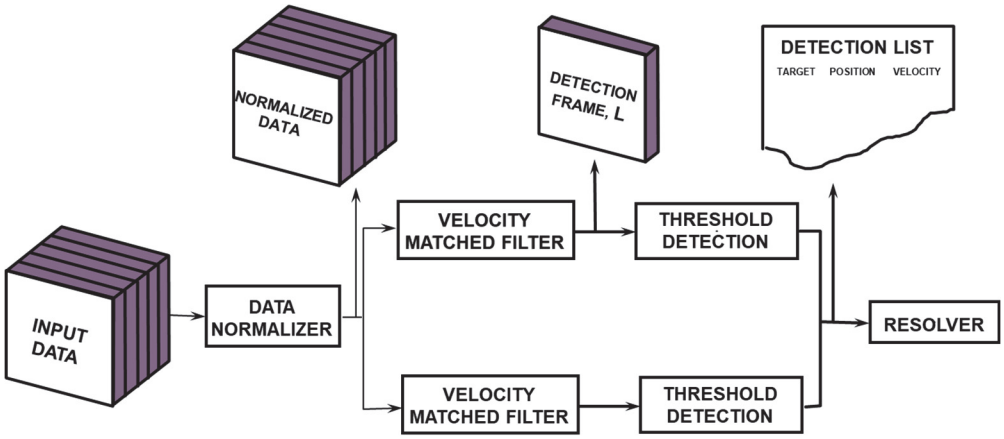

Figure 3.1 Maximum likelihood detection using an assumed velocity filter bank. Reprinted with permission from *IEEE Trans AES* **34**(3) (1998).

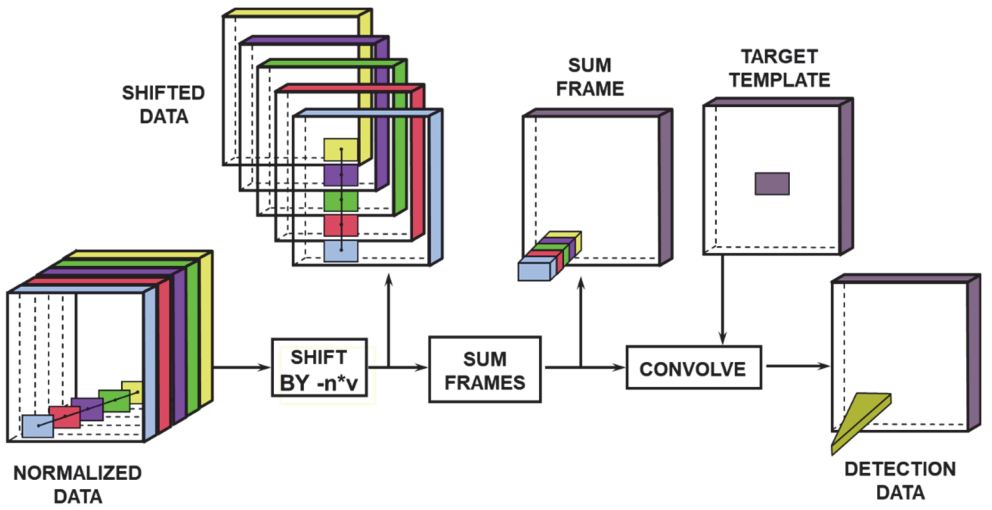

Figure 3.2 Simultaneous detection and tracking summing target energy along an assumed velocity trajectory.

3. D. Wilmot et al., "Warning Systems," p. 61, in *Countermeasure Systems*, D. Pollock, Ed., Vol. 7 of *The Infrared and Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, Eds., Infrared Information Analysis Center and SPIE Press (1993).

Psychometric Function

The psychometric function (Fig. 3.4) is well matched by a Weibull function as follows:[1]

$$P(x) = 1 - (1-\gamma)2^{-\left[\frac{x}{\alpha}\right]^{\beta}},$$

where P is the fraction of correct responses, x is the stimulus strength, β is a parameter that determines the steepness of the curve, γ is the guess rate (0.5), and α is the stimulus strength at which 75 percent of the responses are correct.

Discussion

Bijl and Valeton[1] point out that "the probability of a correct response (in a detection task) increases with stimulus strength (that is, contrast). If the task is impossible because the contrast is too low, the probability of a correct response is 50 percent (guess rate), and if the task is easy, the observer score will be 100-percent correct. The relationship between stimulus strength and probability of a correct response is called the *psychometric function*."

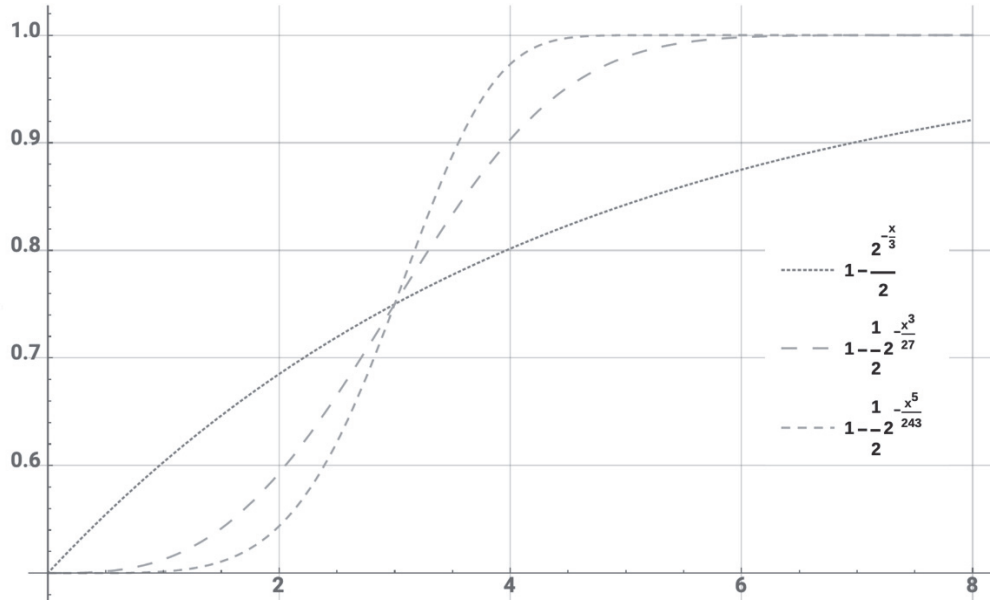

Figure 3.4 Example of the psychometric function. The fraction of correct responses gradually increases with stimulus strength (i.e., contrast) from 50 percent (representing chance) to 100 percent. The threshold is defined as the stimulus strength at which the observer scores 75 percent correct. The threshold is independent of the decision criterion of the observer.

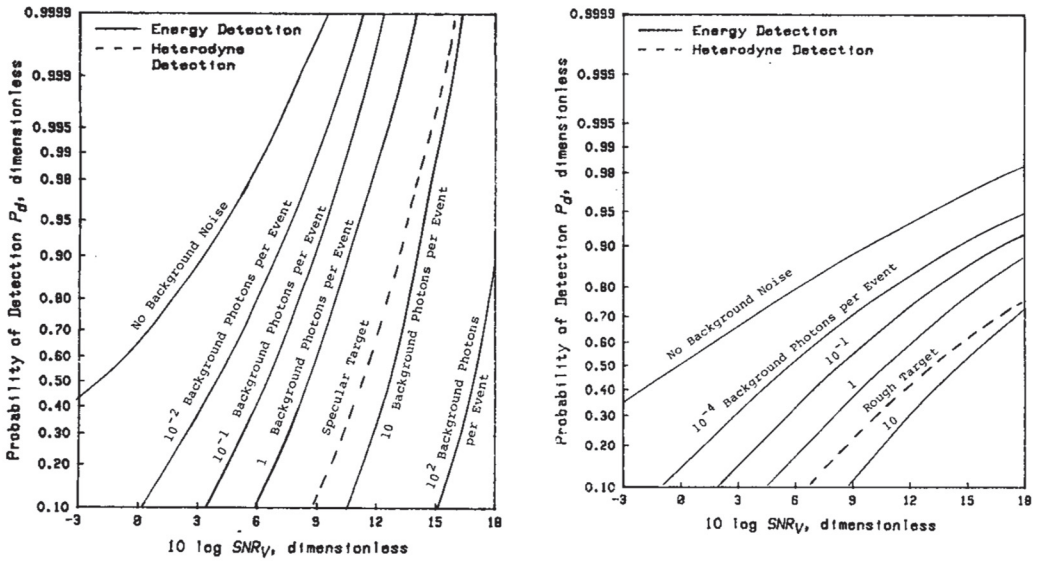

Figure 3.8 SNR and P_d for active sensors for two classes of surface roughness presuming a FAR of 10^{-6}. The left panel describes performance for specular or well-resolved rough targets. The right panel shows performance against rough targets.[5]

References

1. *Burle Electro-Optics Handbook*, Burle Industries, p. 111, ftp://ftp.bartol.udel.edu/jmc/pmt/rca_eo_hb.pdf (1974).
2. G. Kamerman, "Laser Radar," p. 45, in *Active Electro-Optical Systems*, C. Fox, Ed., Vol. 6 of *The Infrared and Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, Eds., Infrared Information Analysis Center and SPIE Press (1993).
3. K. Seyrafi and S. Hovanessian, *Introduction to Electro-Optical Imaging and Tracking Systems,* Artech House, pp. 147–157 (1993).
4. W. Press et al., *Numerical Recipes*, Cambridge University Press, p. 164 (1986).
5. "Military Handbook Quantitative Description of Obscuration Factors for Electro-Optical and Millimeter Wave Systems," DOD-HDBK-178(ER), Section 5, p. 10 (1986).

Limits of Position Estimation

When measuring the location of a point target using a quad cell (which is analogous to using 4 pixels on an FPA to sample the PSF distribution), the standard deviation of the angular position error σ can be estimated using

$$\sigma = \frac{PSF}{2SNR}.$$

The standard USAF resolution test pattern[13] shows the typical resolution pattern and the alternating dark and white lines mentioned above (Fig. 3.10).

When discussing DRI and the number of cycles across a non-symmetric object such as a person, the "critical" dimension to use is usually the smaller dimension (the width of the person, not their height).

Resolution as described above will result in proper performance by 50 percent of the observers asked to perform the observational task under nominal conditions. More detail on the expected number of line pairs on a specific target is contained in Table 3.2, from Johnson.[1] If you are unsure about the number of cycles required and need to do a calculation, use the nominal values in Table 3.3 as a guide.

The Johnson criteria were developed by a military analyst viewing single still imagery to identify military forces. In this case, the "trained observer" is indeed well trained, familiar with the type of imagery (backgrounds, resolution, etc.), and looking for well-known objects. As a result, many practitioners tend to round up slightly, say, to 1 full cycle for detection and 4 cycles for recognition. Orientation does not seem to be commonly used in current literature or practice, hence the reference to DRI instead of DORI.

The Johnson criteria were developed at a time when photographic film was the standard for surveillance applications. As a result, degradation factors common to modern digital electronic systems, including sensor pixelization, display resolution limits, and other factors, must be compensated for prior to applying the Johnson criteria.

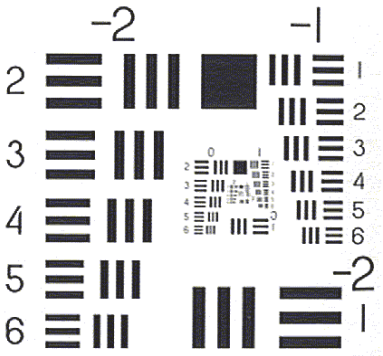

Figure 3.10 USAF resolution test target.

Table 3.2 More detail on the expected number of line pairs on a target.[2]

Target	Detection	Orientation	Recognition	Identification
Truck	0.9	1.25	4.5	8.0
M48 tank	0.75	1.2	3.5	7.0
Stalin tank	0.75	1.2	3.3	6.0
Centurion tank	0.75	1.2	3.5	6.0
Half-track	1.0	1.5	4.0	5.0
Jeep	1.2	1.5	4.5	5.5
Command car	1.2	1.5	4.3	5.5
Soldier	1.5	1.8	3.8	8.0
105 Howitzer	1.0	1.5	4.8	6.0

National Image Interpretability Rating Scale

NIIRS Rating Level	Examples of Exploitation Tasks (Visible)	Examples of Exploitation Tasks (Multispectral)	Examples of Exploitation Tasks (Infrared)
0	Interpretability of the imagery is precluded by obscuration, degradation, or very poor resolution.	Interpretability of the imagery is precluded by obscuration, noise, poor registration, degradation, or very poor resolution.	Interpretability of the imagery is precluded by obscuration, noise, degradation, or very poor resolution.
1 (>9 m GSD)	Distinguish between major land use classes (urban, forest, water, etc); detect a medium-sized port facility; distinguish between taxiways and runways.	Distinguish between urban and rural areas; identify a large wetland (>100 acres); delineate coastal shoreline.	Distinguish between runways and traffic ways; detect large areas (>1 km^2) of marsh or swamp.
2 (4.5–9.0 m GSD)	Detect large buildings (hospitals, factories); detect military training areas.	Detect multilane highways; detect strip mining; delineate extent of cultivated land.	Detect large aircraft and large buildings; distinguish between naval and commercial port facilities.
3 (2.5–4.5 m GSD)	Detect individual houses in residential neighborhoods; detect trains on tracks but not individual cars; detect a helipad; identify a large surface ship in port by type.	Detect vegetation/soil moisture differences along a linear feature; identify golf courses; detect reservoir depletion.	Distinguish between large and small aircraft, distinguish between freighters and tankers of 200 m or more in length; identify individual thermally active flues running between the boiler hall and smoke stacks at a thermal power plant.
4 (1.2–2.5 m GSD)	Identify farm buildings as barns, silos, or residences; identify by general type, tracked vehicles, and field artillery; identify large fighters by type.	Distinguish between two-lane improved and unimproved roads; detect small boats (3–5 m) in open water.	Identify the wing configurations of small fighter aircraft; detect a 50-m^2 electrical transformer yard in an urban area.
5 (0.75–1.20 m GSD)	Detect large animals in grasslands; identify a radar as vehicle mounted or trailer mounted.	Detect an automobile in a parking lot; detect disruptive or deceptive use of paints or coatings on buildings at a ground-forces installation.	Distinguish between single-tail and twin-tail fighters; identify outdoor tennis courts.
6 (0.40–0.75 m GSD)	Identify individual telephone/electric poles in residential neighborhoods; or the spare tire on a medium-sized truck.	Detect a foot trail through tall grass; recently installed minefields in ground-forces deployment areas; or navigational channel markers and mooring buoys in water.	Distinguish between thermally active tanks and APCs, or between a 2-rail and 4-rail launcher; identify thermally active engine vents atop diesel locomotives.
7 (0.2–0.4 m GSD)	Identify individual railroad ties; or fitments and fairings on a fighter-sized aircraft.	Detect small marine mammals on sand or gravel beaches; distinguish crops in large trucks; detect underwater pier footings.	Identify automobiles as sedans or station wagons; identify antenna dishes on a radio relay tower.

Chapter 4
Backgrounds

Introduction

A typical problem with most electro-optical sensors is that backgrounds can be complex and can include influences that reduce the chances of discerning targets that are superimposed on such backgrounds. The backgrounds against which targets are viewed are often as bright as, and sometimes brighter than, the targets themselves. To detect targets, the sensed background noise (or clutter) levels must be different from the sensed target levels and able to be processed so as to provide acceptable false detection rates. In addition, the background may include spatial variation that includes a size distribution that matches that of the target. In general, it is desirable for any sensor to possess the capability of detecting targets in a variety of environments and backgrounds.

Borrowing from the radar community, the spatial and temporal amplitude variation in the background is usually referred to as *clutter*. Clutter is a structured background phenomenon that is not spatially or temporally constant. It is not a characteristic that allows the signals to be combined in a root-sum-squared (RSS) fashion; rather, the signals must be simply added to other noise sources. From elementary statistics, we know that adding noise sources creates a total that is larger than would be the case if the terms were added in an RSS manner. Clutter requires more sophisticated techniques of image processing, as described below.

The sources of clutter for the chosen band pass may include large weather fronts, Sun glints, clouds and cloud edges, variations in water content, lakes, certain bright ground sources, and variances in emissivity, reflectivity, and temperature. For viewing from space, the altitude of the clutter sources will depend on the wavelength of operation. Ground sources will only be detected by a space-based sensor at wavelengths that penetrate the atmosphere. The same sources are present for most terrestrial or down-looking airborne sensors.

Examples of how clutter prevents detection are easily observed. A poppy seed dropped on a white piece of paper is easy to find. The same seed dropped on a Jackson Pollock painting is almost impossible to detect. In addition to the effect induced by clutter, detection systems must also deal with a background effect that is analogous to trying to detect the seed on a gray piece of paper.

Spatial, spectral, morphological, and temporal band pass filtering are commonly used techniques in contemporary sensor systems to reduce the effects of unwanted

References

1. W. Wolfe and G. Zissis, *The Infrared Handbook*, ERIM, pp. 3-11 to 3-12 (1978).
2. D. Kryskowski and G. Suits, "Natural Sources," p. 144, in *Sources of Radiation*, G. Zissis, Ed., Vol. 1 of *The Infrared and Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, Eds., Infrared Information Analysis Center and SPIE Press (1993).
3. J. Lekner and M. C. Dorf, "Why some things are darker when wet?" *Appl. Opt.* **27**(7), 1278–1280 (1988).

Effective Sky Temperature

The temperature of the sky can be estimated using the following equations:[1]

$$T_{sky} = \sqrt[4]{\varepsilon_{sky}}\, T_{ambient},$$

where T_{sky} is the radiative temperature of the sky if it were assumed to be a blackbody, ϵ_{sky} is the emissivity of the sky = 0.787 + 0.764 ln(T_{dp}/273)F_{cloud}, T_{db} is the absolute dry bulb temperature in K, T_{dp} is the dew point temperature in K, and F_{cloud} is the cloud cover factor defined as

$$F_{cloud} = 1 + 0.0224N - 0.0035N^2 + 0.00028N^3,$$

where N is the opaque sky cover in tenths. If observers provide no cloud cover data, assume 50 percent.

Discussion

The equations convey that

1. A measurement of sky irradiance quickly provides an estimate of the effective sky temperature.
2. The cloud cover and the dew point dictate the emissivity.
3. If cloud cover is zero, F_{cloud} = 1, and ϵ_{sky} = 0.787 + 0.764 ln(T_{dp}/273).
4. An example is helpful: assuming N = 0 and T_{db} = 20°C (293 K), then

$$\epsilon_{sky} = 0.787 + 0.764 \ln(293/273) = 0.841.$$

References

1. ciks.cbt.nist.gov/bentz/nistir6551/node5.html (2019).
2. G. N. Walton, *Thermal Analysis Research Program-Reference Manual,* NBSIR 83-2655, U.S. Department of Commerce (March 1983, updated 1985).
3. R. Zarr, *Analytical Study of Residential Buildings with Reflective Roofs,* NISTIR 6228, U.S. Department of Commerce (October 1998).
4. S. Idso, "A Set of Equations for Full Spectrum and 8- to 14-μm and 10.5- to 12.5-μm Thermal Radiation from Cloudless Skies," *Water Resources Res.* **17**(2), 295–304 (1981).

Chapter 5
Cost and Economics

Introduction

Some of the most popular rules of previous editions regarded estimating or scaling costs. We collected such rules into this chapter, as well as other economic and production rules.

Photonics is commercially and militarily ubiquitous, found in our cellphones, laptops, DVD players, and automobiles. Photonics is an enabler and critical technology in almost every aspect of modern life, from semiconductor manufacturing to medicine, from the houses we inhabit to the vehicles we drive.

The economics of photonics is a study to itself without a lot published. The photonics market is approximately $500 billion annually, employs about 300,000 people,[1] and is likely to experience healthy growth on the order of 8 percent (some sectors could reach as high as 20 percent) per year well into the mid-2020s.[2] Photonics impacts about 8-trillion-dollar commercial markets and the entire global defense, intelligence, security, and law enforcement markets. Photonics companies (exhibiting at SPIE and OSA) are approximately 88-percent private and only 12-percent public.[3]

There are some potentially disruptive technologies and markets that can substantially reduce the cost of photonics components and systems. At the time of writing, key technologies include quantum-dot FPAs, molded optics, graphene applications, and nanophononics. Moreover, the adaptation of photonics into the massive global vehicle market may result in drastically falling prices for the full spectrum of cameras, lidars, and other photonic components used in automobiles. Also manufacturing processes continue to reduce production costs.

These details are all approximately correct at the time of publication. Costs and economics are a roller coaster that can change direction dramatically in a short period of time, e.g., the cost fluctuations of germanium in the past due to nonphotonics influences.

References

1. www.swissphotonics.net/libraries.files/Photonics21_Strategic_Roadmap_2014-2020.pdf, p. 13 (2019).
2. www.nap.edu/read/13491/chapter/4#30 (2019).
3. ec.europa.eu/enterprise/sectors/ict/files/kets/ photonics_final_en.pdf (2019).

Moore's Law

1. The capacity (or number of transistors per unit area and cost) of an integrated circuit (IC or chip) of a given size doubles approximately every 12 to 24 months.[1,2]
2. This is often paraphrased to read, "The power of a microprocessor (or memory on a chip) doubles every 18 months."

Discussion

The more common paraphrase implies a simple mathematical equation that can relate a cost or price drop every 18 months to a given (constant) processing power as (Fig. 5.1)

$$C_y = \frac{C_i}{2^{2y/3}},$$

where C_y is the cost for a given amount of processing at year y, C_i is the initial cost (at year zero), and y is the difference in years between i and y (initial and projected).

"Moore's observation transformed computing from a rare and expensive venture into a pervasive and affordable necessity. All of the modern computing technology we know and enjoy sprang from the foundation laid by Moore's law."[1]

In 1965, Gordon Moore,[2] founder of Fairchild Semiconductors and Intel, extrapolated from the capability of early chips that the number of transistors on a chip would approximately double each year for the next ten years (Fig. 5.2). Max Roser has a graph of processor power over time that can be found in Ref. 3.

Moore revised this law to say it would double every two years, but with other improvements, it seemed like performance or capacity would double every 18 months. Moore later stated,[3] "The original prediction was to look at 10 years, which I thought was a stretch. The fact that something similar is going on for 50 years is truly amazing. But someday it has to stop. No exponential like this goes on forever."

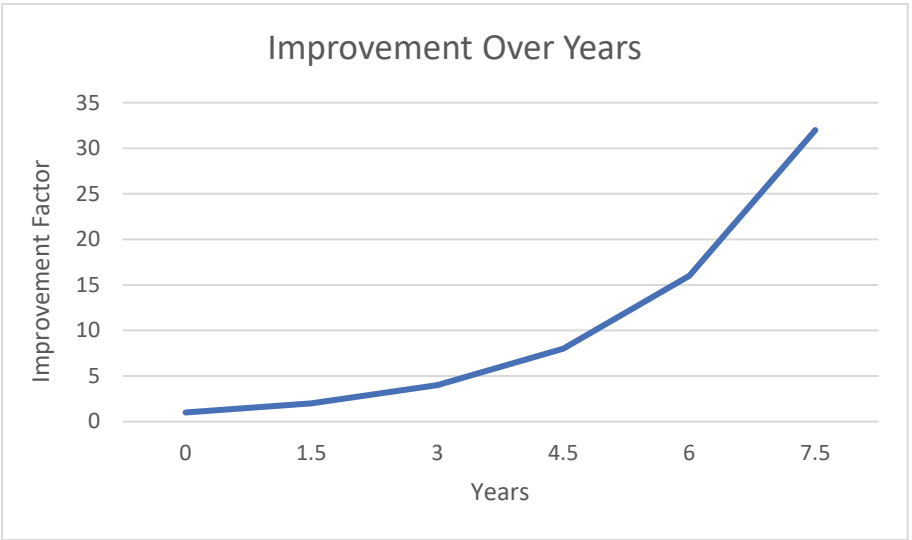

Figure 5.1 Improvement over time given constant costs for the 18-month Moore's law.

Additional References

- www.mooreslaw.org (2019).
- spectrum.ieee.org/semiconductors/processors/the-multiple-lives-of-moores-law (2019).

Metcalfe's Law

A network's "value" is the square of the number of interconnections, and it increases to the square of the number of interconnections.[1]

Discussion

This is a fundamental insight developed by George Gilder and Robert Metcalfe (co-inventor of Ethernet)[1] to value telecommunications networks. However, this insight has far-reaching value for estimating the effectiveness of diverse items including neural networks, animal brains, conferences, meetings, personal networks (such as LinkedIn, Zoom, and Facebook), and even traffic flow and cars on the road. If a car can communicate with another car, that is valuable and worthwhile, but if every vehicle on the road could communicate with each other car, it would be far more productive and valuable. The first fax machine was useless, but as they proliferated, they because immensely useful and valuable. Then, when their numbers declined, they lost value and usefulness. The last fax machine will be useless and of no value.

"Metcalfe's law is related to the fact that the number of unique possible connections in a network of n nodes can be expressed mathematically as the triangular number $n(n-1)/2$, which is asymptotically proportional to n^2."[2]

This is important for photonic practitioners to understand. Increasingly, our cameras and instruments are being integrated in a network of other like cameras (and unlike phenomenological modalities). Photonic sensors are being integrated with other modalities and multiple phenomenologies (e.g., hyperspectral, radar, sonar, etc.) to provide increased insight and intelligence to tasks.

"Metcalfe's law states that a network's impact is the square of the number of nodes in the network. For example, if a network has 10 nodes, its inherent value is 100 (10*10). The end nodes can be computers, servers, and/or connecting users."[3]

This is related to Engelbart's law and two others not in this book (Reed's and Zipf's):[4] "Zipf's law says that if we order some large collection by size or popularity, the second element in the collection will be about half the measure of the first one...The k^{th}-ranked item will measure about $1/k$ of the first one."

As Sundeimeier[5] points out, this applies to meetings and conferences:

> "While at one of the author's (JLM) previous company we lead an engineering conference of approximately 100 engineers from various EOIR related companies which were not co-located. To quantify the value of the conference we could use the following hypothetical:
>
> Average Engineer's annual cost (salary, benefits, travel, etc): $250,000

Cost Reduction Techniques

To successfully reduce costs, a program/company must establish a dedicated team, gather factual savings, focus on near-term gains, and define rewards.

Discussion

Cost-efficient designs are essential in today's environment. Most cost-reduction efforts fall short. For any business to remain viable, they must focus on delivering a cost-conscious solution to their buyer. Whether the customer is an electronics consumer or government entity is irrelevant; the demand for affordable and competitive solutions is paramount.

Unfortunately, many programs or companies focused on cost reduction deliver less than anticipated results or fail completely. The main cause is because lowering costs sounds straightforward, but most people have never had to lead a cost-reduction team and therefore do not fundamentally appreciate the potential roadblocks or do not provide the leadership, resources, and/or authority to make the necessary changes.

However, there are things that can be done to maximize the likelihood of success, and although they seem obvious, they are mandatory to achieve desired results:

1. **Establish a dedicated team**: More often than not, the cost reduction team is also the development team, the production team, or a group of individuals who were selected arbitrarily to help reduce cost. Unfortunately, these groups' primary roles and responsibilities will always trump the affordability efforts, resulting in missed timelines and opportunities to capitalize on cost reduction opportunities.
2. **Gather factual savings**: Cost reduction ideas typically start with the people most familiar with the product putting together a list of efforts and the potential savings; this method is fraught with pitfalls as savings values were based on estimates and set unrealistic targets. To mitigate this issue, it is suggested that savings be solidified through formal activities such as soliciting documentation through the supply chain to ask suppliers to state that if the defined efforts are executed, then the savings expected will be as quoted.
3. **Focus on near-term gains**: It is easy for a team to get discouraged without seeing the results of their work. It is paramount that the easier, "low hanging fruit" opportunities be driven to closure as rapidly as possible so that the team sees the results of their work and does not get distracted
4. **Define rewards**: Identify what the team will receive if certain milestones are met and what it means for the company as a whole (e.g., a program is not canceled, and co-workers get to keep their jobs); consider offering incentives for measurable cost reduction to reinforce the buy-in and deliverables.

Considering and addressing the four aforementioned areas will greatly increase a business's likelihood of success and may ultimately develop a culture of continual cost reduction and incremental improvement.

Reference

1. Private communications with Darrick Buban (2019).

Table 5.13 Incremental impact of tolerances[2] for numbers as low as 100 (for easy-to-meet tolerances) and higher (as tolerances are tightened). For example, a thickness tolerance of 0.0125 mm doubles the cost involved.

Variable	➔ More Difficult ➔ ➔				
Diameter (mm)	±0.10	±0.05	±0.025	±0.0125	±0.0075
	100	100	102	105	125
Thickness (mm)	±0.20	±0.10	±0.05	±0.025	±0.0125
	100	103	115	140	200
Stain	<2	2	3	4	5
	100	103	110	140	175
Cosmetics (Scr-Dig)	80-50	60-40	40-20	20-10	10-5
	100	100	120	150	250
Test (fringes)	5-2	3-1	2-½	1-¼	½-1/8
	100	105	125	175	250
Wedge (arc minutes)	3	2	1	½	¼
	100	105	110	125	150
Doublets (arc minutes)	6	3	2	1	< ½
	100	105	110	150	200
Aspect Ratio	<10:1	15:1	20:1	30:1	50:1
	100	120	175	250	350
Delivery Time (weeks)	8	6	4	2	1
	100	100	130	170	200

This collection of cost rules identifies how each step in the production process adds to cost with details on the influence of requirements impacts. A summary of the equations (and other material) is provided in Ref. 2.

References

1. A. Ahmad, Ed., *Optomechanical Engineering Handbook*, CRC Press, Figure 1.11 and subsequent (1999).
2. wp.optics.arizona.edu/optomech/wp-content/uploads/sites/53/2016/10/Plympton.pdf (2019).
3. R. R. Willey, "The Impact of Tight Tolerances And Other Factors On The Cost Of Optical Components," *Proc. SPIE* **0518** (1985).

Additional Reference

- wp.optics.arizona.edu/optomech/wp-content/uploads/sites/53/2016/10/Tolerancing-Optical-Systems-E.Milby-2009.pdf (2019).

Stahl Segmented Cost Rule

Stahl[1] provides the following rule for estimating the cost of ground optical telescopes:

$$Ground\ OTA\ Cost\ \$\sim SF\ D^{1.8}\lambda^{-0.5}e^{-0.04(Y-2000)},$$

where D is the primary mirror diameter, λ is the wavelength diffraction-limited performance, Y is the year of development (less 2000) for reduction in technology cost over

Chapter 7
Focal Plane Arrays

Introduction

A detector is a universal component of the electro-optical system. It is a transducer that changes the energy in the form of light into electrical impulses. Detectors come in many shapes and sizes, from simple quad-cells that can be used to detect the position of the image, to line arrays used as scanning systems, and finally to 2D arrays used in digital cameras and cellphone cameras. These 2D arrays are also called *focal plane arrays* (FPAs). Traditionally, detector arrays have held the dubious position of being the system sensitivity limiter, resolution limiter, and cost driver. However, advances in the technology are making the FPA less of a performance and cost concern. A million-dollar forward-looking infrared (FLIR) system may contain an FPA that costs $5,000 or less. Visible sensors are ubiquitous thanks to the cellphone explosion and uses in automobiles. These high-density FPAs cost a few dollars and provide years of reliable service using low-cost, easily processed silicon (nature's gift to humanity) as the detector.

Historical Review of Detectors

To understand modern detectors, a quick review of the history of imaging systems is in order. Willoughby Smith first reported the photoconductive effect in 1873 while studying selenium crystals. This—and subsequent investigations of what would become electronic detectors for the UV, visible, and IR—was an important line of work, because film relying on silver halide (which dominated image detection for 100 years) does not respond well in the ultraviolet or light beyond about 1.2 μm. Additionally, film does not lend itself to the electrical digitization needed by modern processors, displays, and electronic communication.

Albert Einstein won a Nobel prize in physics for explaining the photoelectric effect. Einstein's equation, first published in 1905, explained the energy of the free electron from the material and is given by

$$eV_0 = h\nu - \phi,$$

where eV_0 is the energy of the electron, h is the Planck's constant, ν is the frequency of the light, and ϕ is the work function (or the minimum energy required to free the electron). The work function is a characteristic value of the material. This demonstrates the quantum

Table 7.2 Wavelengths of common detector materials.

Material	Typical Useful Spectral Region (μm)	Notes
CdS	0.30 to 0.55	Rarely used today.
GaAs	0.6 to 1.8	Linear arrays are commercially available. Doping with phosphorus extends the cut-off wavelength.
GaAs QWIP	2 to 20	Tunable at time of manufacture, limited in spectral bandwidth, and suitable for dense arrays and low-cost production.
Ge:xx	2 to ~100	Doped Ge has long been an IR detector, with one to a few elements per array. Ge:Hg can respond as low as 2 μm, whereas Ge:Ga can respond at 100 μm at 3 K.
HgCdTe	2 to 22	This material is tunable at time of manufacture. Arrays up to 12 μm are commonly available.
InGaAs	0.8 to 1.7	Cut-off can be extended to ~2.6 μm by adding phosphor. Usually, only thermoelectric cooling is required.
InSb	1.0 to 5.5	Some versions have response into the visible wavelength spectrum.
$LiTaO_3$	5 to 50	Pyroelectric materials with 2D arrays are available.
PbS	1 to 3	Usually photoconductive, so 2D arrays are rare, although many linear arrays are made.
PbSe	2 to 5	2D arrays are rare, but many linear arrays are in production.
Pt:Si	1 to ~5	Pt:Si is highly uniform and producible in large formats but offers low quantum efficiency at MWIR wavelengths. Although a popular FPA material in the 1980 and 1990s, it has become archaic.
Si	0.3 to ~1.0	Red-enhanced and blue-enhanced versions are available; it lends itself to IC manufacture.
Si:X	0.3 to 26.0	Doping silicon allows detection into the LWIR; requires cooling well below liquid nitrogen temperatures.
Micro-bolometer	8 to 15	Via micromachining, silicon can be made into a tiny bolometer, lending to dense arrays. Currently available with vanadium oxide or amorphous silicon coatings.
SLS	4 to 14	Manufactured from strained layers of III-V materials such as InAs and InGaSb. These new IR detectors have electronic and optical properties similar to (maybe, in the near future, superior to) those of HgCdTe.

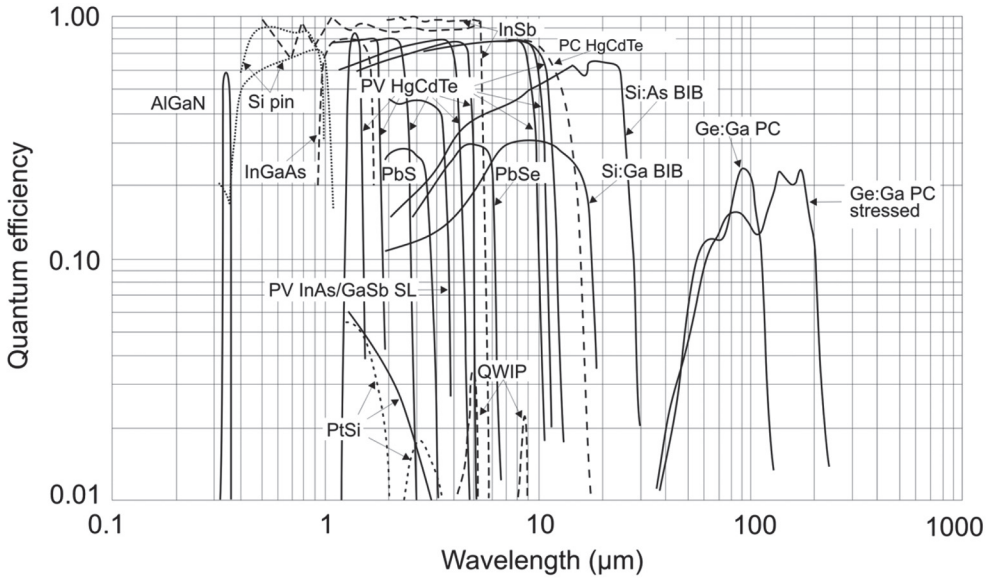

Figure 7.3 Quantum efficiency of various detectors. Reprinted from Rogalski et al.,[1] p. 17.

References

1. A. Rogalski, M. Kopytko, and P. Martyniuk, *Antimonide-Based Infrared Detectors: A New Perspective*, SPIE Press (2018).
2. R. C. Jones, "Performance of detectors for visible and infrared radiation," in *Advances in Electronics*, Vol. 5, L. Morton, Ed., Academic Press, pp. 27–30 (1952).
3. B. Saleh and M. Teich, *Fundamentals of Photonics*, John Wiley & Sons, p. 649 (1991).

Responsivity and Quantum Efficiency

An optical detector's responsivity (in amps per watt) is equal to its quantum efficiency divided by 1.24 times the wavelength in microns, or

$$R = \frac{QE}{1.24}\lambda.$$

Discussion

Responsivity is a measure of a detector's output signal (defined in amps, volts, or electrons per second) to a given input radiant signal, regardless of the noise. Higher responsivities are always better than lower responsivities.

This rule occurs because of the definition of quantum efficiency. It is defined as the number of electrons generated per second per incident photon/second on the active area

Shot Noise Rule

If the photocurrent from a photodiode is sufficient to produce a 50-mV voltage drop across the load resistor at room temperature, the shot noise equals the Johnson noise at 300 K.

Discussion

Shot noise is an important noise source for modern infrared photodiodes. At tactical background levels, it is frequently the dominant noise source. The above result can be proved by investigating the comparison of Johnson noise (caused by the random motion of carriers within a detector, usually thermal in nature) and shot noise (the result of the statistics of the photon-to-electron conversion process in the detector). The rms Johnson noise voltage is expressed as

$$\sqrt{4kTR\Delta f},$$

where k is the Boltzmann's constant (8.6×10^{-5} eV/K), T is the temperature in K, R is the resistance, and Δf is the noise bandwidth. The noise bandwidth is generally $1/2t_i$, where t_i is the integration time. Shot noise (current) is expressed as

$$\sqrt{2eI\Delta f},$$

where I is the average current in the detector, and e is the charge of the electron. The shot noise voltage is the product of the noise current and the resistance defined above. That is, the shot noise voltage is

$$R\sqrt{I2e\Delta f}.$$

If we now take the ratio of the Johnson noise to the shot noise and cancel some terms, we obtain

$$\sqrt{\frac{2kT}{RIe}}.$$

We can compute this ratio for $T \sim 300$ K, and the voltage drop $RI \sim 50$ mV:

$$\sqrt{\frac{2\times 8.6\times 10^{-5}\,\text{eV/K}\times 300\ \text{K}}{50\times 10^{-3}\,\text{V}\times e}} = 1.02.$$

Using these values, we find the ratio to be nearly unity, which proves the assertion in the rule. To get a ratio of precisely 1, the shot noise voltage drop needs to be 51.6 mV.

References

1. P. Hobbs, *Building Electro-Optical Systems: Making it All Work*, John Wiley, p. 117 (2000).
2. W. Sloan, "Detector-Associated Electronics," in *The Infrared Handbook*, W. Wolfe and G. Zissis, Eds., ERIM, pp. 16-4 to 16-6 (1978).

Chapter 9
Lasers

Introduction

Revolutions in photonics are occasional and important. Since the earliest photonics developments (TV, photomultipliers, and others in the 1920s), most of this discipline's history has involved slow evolution punctuated with occasional, dramatic technological improvements. Like the invention of the lens, film, and electronic sensors, lasers caused a revolution throughout the discipline. And they provided the impetus for advancements in many areas of photonics.

Lasers have unique diagnostic capabilities essential for evaluating effluents from exhausts, detecting gravitational waves, diagnosing optical defects in the human eye, and thousands of other applications. Laser systems provide high-quality, stable-alignment reference sources and allow lens and mirror quality testing using interferometry. Lasers have led to the development of several applications that would be unthinkable using conventional light sources. Fields such as optical communications, laser eye surgery, and active tracking would be impossible without the unique features of lasers.

Lasers, in the form of laser pointers, CDs, DVDs, and laser speed guns (among others), have been one of the few photonics advancements to become a part of the lexicon of the average citizen. Their enormous brightness and spectral purity have changed many parts of the electro-optics environment and have provided, after about 50 years of development, new advancements in consumer electronics can be created only with high-performance laser lithography.

As a result of lasers' widespread application, researchers have invested heavily in understanding the characteristics of the beams they produce and the interaction of those beams with various types of targets and detectors. A full understanding of the application of lasers requires new insight into the way electromagnetic waves propagate in the atmosphere. The close relationship of laser light propagation and the medium in which they travel requires this chapter (and others) to include a mixture of rules. As a result, rules are presented here and in the chapters on Atmospherics and Astronomy pertaining to the properties of the beams as they might propagate in vacuum as well as how they interact with the propagation medium.

The laser field has been the source of many interesting stories about how science and industry do business. For instance, when the first optical laser was developed, a press conference was held, pictures were taken, stories were written, and predictions were made.

$$\frac{PA}{\lambda^2 BQ^2}.$$

This rule is related to the antenna theorem,[3] which states that $A\Omega \approx \lambda^2$. More detail on this powerful expression is provided in the rule "Etendue." This can be illustrated simply by noting that for a diffraction-limited beam, the solid angle that is obtained Ω is

$$\pi\left(\frac{1.22\lambda}{D}\right)^2.$$

The area of the aperture is $(\pi D^2)/4$, so that the product of these two terms is $3.67\lambda^2$.

As an example of laser brightness, consider a HeNe laser operating at 0.6328 micron and producing 1 milliwatt, emitted through a 1-mm aperture. Using the first equation of this rule, we get

$$\frac{PA}{\lambda^2} = \frac{10^{-3} \cdot 7.8 \times 10^{-7}}{\left(0.6328 \times 10^{-6}\right)^2} = 1.94 \times 10^3 \text{ W/sr.}$$

A comprehensive list of definitions appropriate to laser brightness appears in Shukla.[5]

References

1. G. Fowles, *Introduction to Modern Optics,* Dover Publications, p. 223 (1975).
2. G. Golnik, "Directed Energy Systems," p. 451, in *Emerging Systems and Technologies*, S. R. Robinson, Ed., Vol. 8 of *The Infrared and Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, Eds., Infrared Information Analysis Center and SPIE Press (1993).
3. A. Siegman, *Lasers,* University Science Books, p. 672 (1986).
4. T. Ross and W. Latham, "Appropriate Measures and Consistent Standard for High Energy Laser," *J. Directed Energy* **2**, 22–58 (summer 2006).
5. P. Shukla, "Understanding laser beam brightness: A review and new prospective in material processing," *Optics and Laser Technol*. **75**, 40–51 (2015).

Laser Beam Quality

Beam quality is defined as

$$BQ = \exp\left(\frac{1}{2}(2\pi WFE)^2\right) = \frac{1}{\sqrt{SR}},$$

where *SR* is the Strehl ratio (unitless), and *WFE* is the rms wavefront error. This equation is only valid for wavefront errors less than about 1/5 of a wave.

Chapter 13
Optical Design and Analysis

Introduction

Developments in optics have been linked to specific engineering applications. Optics of antiquity, until about 1700, existed mainly as an aid to vision or for a tiny part of the population, or astronomy or military applications. From about 1600 to World War II, the main impetus for optical development was to develop better instruments for astronomy, window glass, and other industrial applications. Many major developments during this era were in some way geared to astronomy (e.g., the Foucault knife-edge test, interferometers, new telescopes, and so on). Military needs dominated optics development from World War II to the 1990s. In the 1990s, the military remained a driving force, but its (and NASA's) technologies were commercialized, leading to an explosion in applications in consumer products, high-bandwidth communications (fiber optics). and science. More recently, the driving force is smartphone cameras, self-driving vehicles, bionics, and robotics.

The science of optics is ancient. Archeological findings involving the Phoenicians suggest that powered lenses were made over 3000 years ago. All the ancient cultures studied light and its interaction with matter. In the ancient world, Aristotle, Plato, Euclid, Pythagoras, and Democritus all wrote extensively about vision and optics between approximately 550 and 250 BCE. Seneca (4 BCE to 45 CE) was the first to write about observing light divided into colors by a prism. To these early investigators, the world was full of rules of thumb and principles explained by the thought process alone. Frequently, this resulted in poorly made or poorly understood observations. Among the incorrect theories was that vision resulted from "ocular beams" emitted from the eye. This theory was finally rejected by al-Kindi (801–873 CE) and al-Haytham (965–1039 CE).

No one knows who invented spectacles. One of the earliest paintings of a person wearing glasses is attributed to Crivelli's painting of Hugues de St. Cher in 1352; however, spectacles and lenses were known to glassmakers for several centuries prior to this period.

Likewise, much controversy surrounds the inventor of the first telescope, although it probably occurred around 1600 by Lippershey (who applied for a patent in 1608), Adriaanzoon, Jansen, or someone else. Refractive telescopes were already being sold as toys and navigation aids when Galileo and others turned them to the heavens for astronomy. Prophetically, Galileo remarked that the science of astronomy would improve with further observations from better telescopes. The microscope was invented about the same time, with almost as much controversy concerning the original inventor.

For angles other than 45°, the lateral deviation can be calculated by[1]

$$\Delta y = t\sin\theta\left[1-\sqrt{\frac{1-\sin^2\theta}{n^2-\sin^2\theta}}\right],$$

where θ is the angle of incidence. If small angles are being used, the above equation reduces to the following, with θ in radians:

$$\Delta y = \frac{t\theta(n-1)}{n}.$$

Note that even though it is not captured in these equations, inserting a tilted PPP into a converging or diverging beam will result in an array of aberration impacts. The material properties of the PPP and the system geometry, including the index and Abbe number of the PPP, the thickness of the PPP, the angle of tilt, and the *f*/# of the system, control the magnitude of the aberrations. In general, there will be more impact inserting a PPP into the beam for faster systems (i.e., an *f*/1 system will be more affected than an *f*/4 system). Table 13.2 provides equations for calculating the amount of each third-order aberration introduced into the system[2] when a PPP is introduced into a converging beam at angle θ (in radians) and the rays are entering the PPP at angle *u* to the optical axis (see Fig. 13.3).

In the table, *V* is the Abbe number of the PPP, and *u* is the angle of the ray (in radians) with respect to the optical axis. Inserting a PPP at 45° into a converging beam can severely degrade an image; the astigmatism introduced according to the equation above is approximately a quarter of the thickness of the PPP. This can be avoided by adding a wedge to the PPP, or the astigmatism can be cancelled out by inserting another PPP (with the same material properties and thickness) tilted at 90° relative to the first plate.[2]

Table 13.2 Equations for calculating aberrations from a PPP.

Aberration	Equation
Spherical	$\frac{tu^2(n^2-1)}{2n^3}$
Coma	$\frac{tu^2\theta(n^2-1)}{2n^3}$
Astigmatism	$\frac{-t\theta^2(n^2-1)}{n^3}$
Lateral color	$\frac{t\theta(n-1)}{n^2V}$
Longitudinal color	$\frac{t(n-1)}{n^2V}$

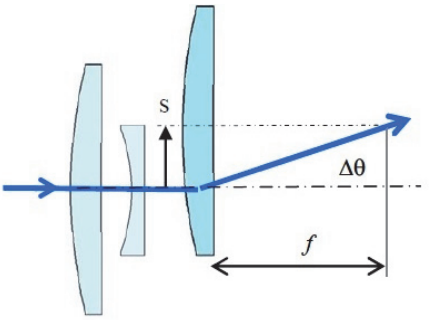

Figure 13.5 Lateral decenter of a lens.[1]

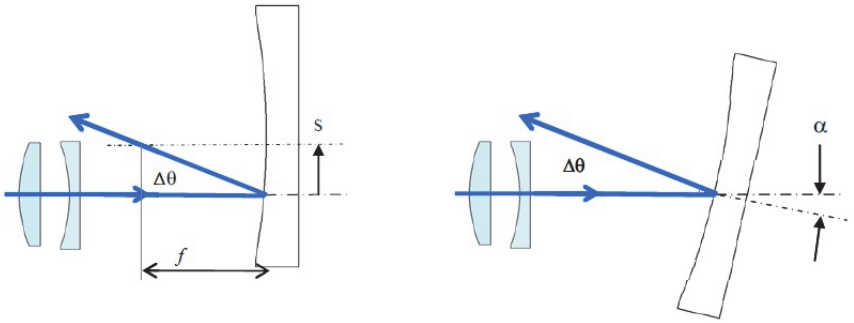

Figure 13.6 Lateral decenter (left) and tilt (right) of a powered mirror.[1]

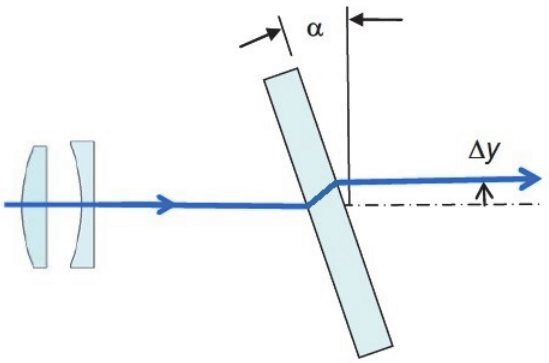

Figure 13.7 Tilt of a plane parallel plate.[1]

Discussion

The table provides image motion and LOS changes for individual elements. For an optical system, the results for Δθ can be used to estimate the image shift for the system ε by

$$\varepsilon = (f\#) \cdot D_i \cdot \Delta\theta_i - \frac{NA_i}{NA} \Delta y_i,$$

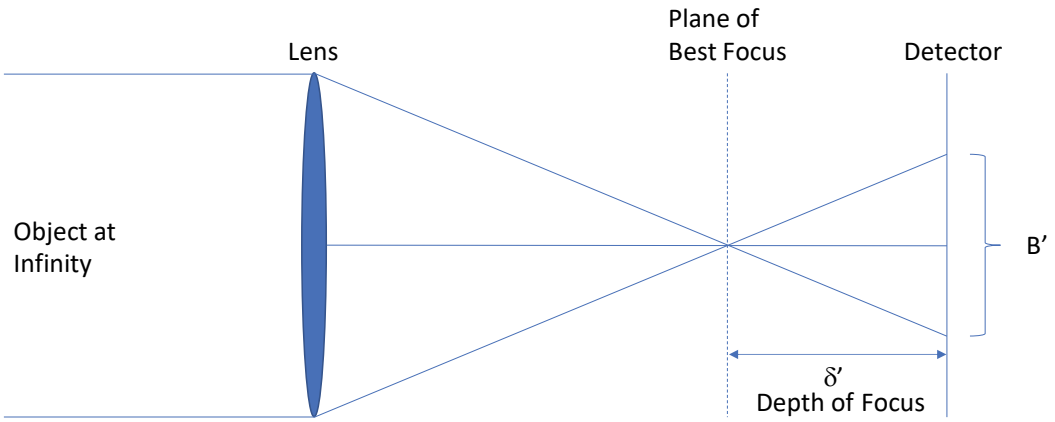

Figure 13.8 Depth of focus.

Solving for δ and assuming the depth of field is small compared to the distance D, then this equation reduces to

$$\delta = \frac{\beta D^2}{A}.$$

For the image side, the above relationship can be rewritten as

$$\delta' = \frac{\beta D'^2}{A},$$

and for a telescope focused at infinity, the distance to the image plane D' equals the focal length of the system F. Also, by making use of the definition of the f-number, we derive the original rule provided above:

$$\delta' = \frac{\beta F^2}{A} = F\beta f\# = B' f\#.$$

Added angular blur from defocus seriously degrades system performance when it exceeds the in-focus angular blurring from other aberrations and detector IFOV. The above relationship allows one to estimate the amount of defocus from a system (focused at infinity) attempting to image an object not at infinity. Note that the depth of field toward the optical system is smaller than that away from the system.

Reference

1. W. Smith, *Modern Optical Engineering,* 3rd ed., McGraw Hill, pp. 155–156 (2000).

Hyperfocal Distance

The hyperfocal distance (in meters) can be approximated as

4. Private communications with George Spencer (1995).
5. Private communications with Max Amom (1995).
6. W. Smith, "Optical Systems," in the *Handbook of Military Infrared Technology*, W. Wolfe, Ed., Office of the Naval Research, pp. 427–429 (1965).
7. G. Holst, *Electro-Optical Imaging System Performance*, JCD Publishing, pp. 459–460 (1995).
8. R. Hudson, *Infrared Systems Engineering,* John Wiley, p. 180 (1969).
9. O. von Voss, "Stanley Kubrick's Legendary f/0.7 ZEISS Lens Explored," Cinema5D.com (July 2016).
10. www.cinema5d.com/stanley-kubricks-legendary-f0-7-zeiss-lens-explored (2020).

f-Number for Circular Obscured Apertures

The standard definition of *f*/# is generally inappropriate when an obscuration is present. In such cases, use the effective *f*/# ($f\#_{effective}$) defined as[1,2]

$$f\#_{effective} = \frac{effective\ focal\ length\ of\ the\ overall\ system}{diameter\ of\ primary\ mirror} \sqrt{\frac{1}{\left(1-\left(\frac{D_o}{D_p}\right)^2\right)}},$$

where D_o is the effective diameter of the obscuration, and D_p is the diameter of the primary mirror or other defining entrance aperture.

Discussion

This rule can be derived from the following argument: The effective *f*/# equals the ratio of the effective focal length and the effective aperture diameter (which is just the actual aperture area converted to a circular aspect):

$$f\#_{effective} = \frac{f_{effective}}{D_{effective}} = \frac{f_e}{\sqrt{\frac{4}{\pi}A_e}} = \frac{f_e}{\sqrt{\frac{4}{\pi}\left(\frac{\pi}{4}\left(D_p^2 - D_o^2\right)\right)}},$$

which is equivalent to the equation in the rule. In the equation, A_e is the effective area of the entrance aperture. This relationship is useful for determining the *f*/# or effective focal length that should be used and for estimating the impact of a central obscuration.

There is a special-case simplification of the above rule that applies to many telescopes (especially for visible and IR astronomy). If the diameter of the central obscuration is small as compared to the diameter of the aperture, then

$$\frac{1}{\sqrt{1-\varepsilon^2}} \approx 1 + \frac{\varepsilon^2}{2},$$

where ε is the obscuration diameter divided by the aperture diameter (D_o/D_p).

1. Lower-order (Seidel) aberrations have ratios from √3 to √32;
2. Crinkly wavefronts (more or less random) have a ratio of 2√2; and
3. Surfaces with high spatial frequencies, such as diamond-turned surfaces, have lower ratios around 2–3.[3]

This rule seems to hold for many nonstandard aberrations such as circumferential grooves in the wavefront (which vary sinusoidally with radius), grooves varying co-sinusoidally, grooves with a square-wave variation with radius, and a two-level zone plate.

In each of the above, the PV-to-rms ratio is either 2 or √2 when there are many periods from center to edge. For higher-order Zernikes, the ratio seems to tend toward 8. Truly random wavefront errors are arguably described by PV ratios between 2 and 8.

Often, the PV figure error is available by inspecting an interferogram. Conversion to the rms error is more tedious than using this rule, as it requires some computation of a 2D integral. However, the effort may be worth it, as the rms wavefront error is useful in computing other quality measures (such as the Strehl ratio) so long as it is not too large. If the result is a very large PV-to-rms ratio, it is often the case that there are noisy data points in the measurement (e.g., dust on the surface). Figure 13.36 illustrates how this data is presented in typical optics lab measurements.

Some specific cases are shown in Table 13.9.

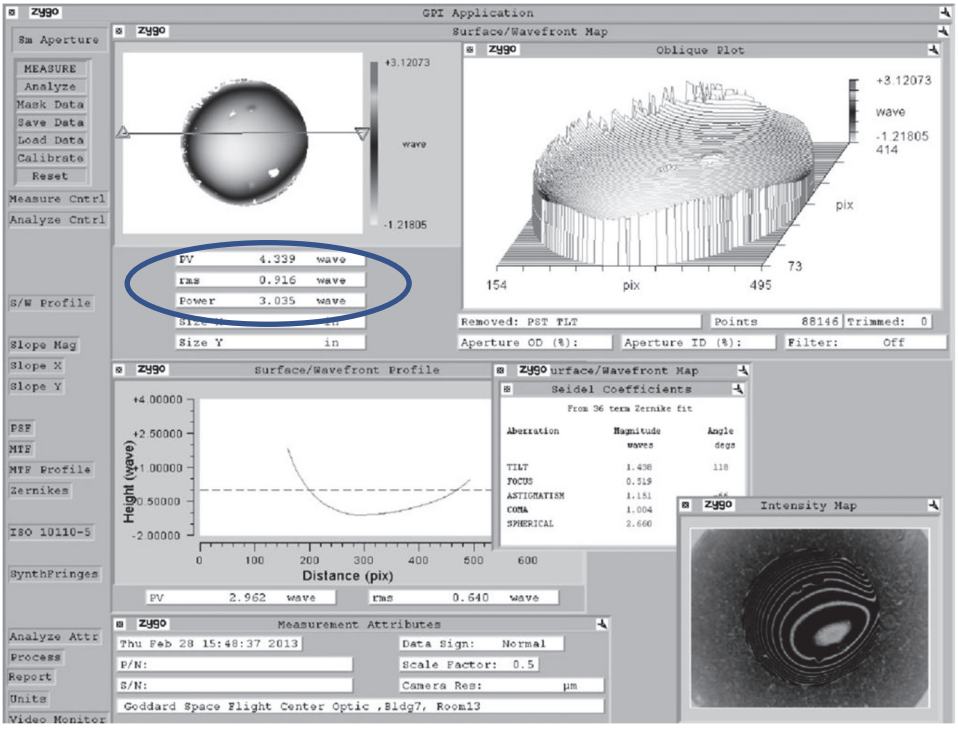

Figure 13.36 The output of a typical interferometric analysis system. Key metrics include peak to valley (PV) and root mean square (rms), as highlighted by the oval.

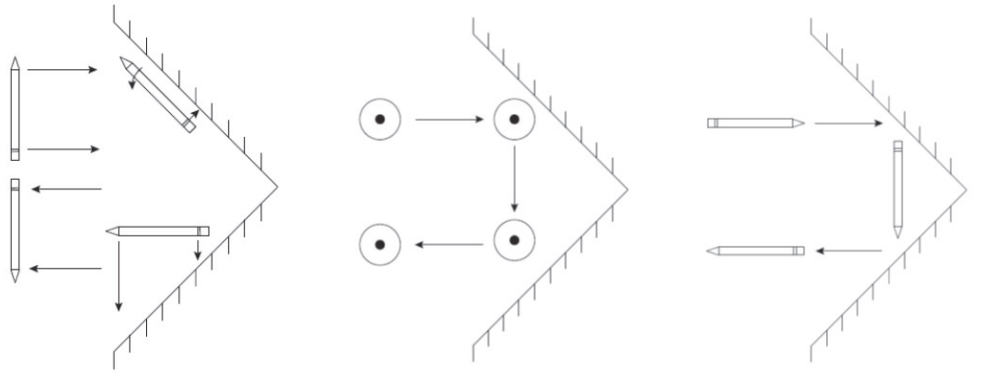

Figure 13.54 Example visualization of a pencil "bouncing" off two successive mirrors. The left diagram shows the pencil traveling as the *x* axis and results in a 180° image rotation. The next two diagrams show the pencil traveling as the *y* axis (middle) and *z* axis (right), which result in no change in the image orientation.

Discussion

This is a quick mental technique that can be used to determine the orientation of an object after it experiences a reflection without need for complicated analysis. By repeating this exercise through each successive reflection in a system, and in each perpendicular axis, the entire image orientation can be determined.

Reference

1. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill (2000).

Thermal Gradients in a Primary Mirror

A constant thickness plane or shallow spherically curved mirror is distorted by an axial steady state thermal gradient. The distortion manifests itself primarily as a change in the radius of curvature according to[1]

$$\frac{1}{R_o} - \frac{1}{R} = \frac{\alpha q}{K},$$

where R_o is the original radius of curvature, R is the new radius of curvature, α is the thermal coefficient of expansion, K is the thermal conductivity, and q is the heat flux per unit area.

Discussion

One of the authors [Friedman] encountered this issue when working in the field of underwater optics. Laser light sources projected through windows into a water-filled test tank heated the center of the window and complicated the curvature induced by the water pressure.

Chapter 16
Radiometry

Introduction

Radiometry is the study of the creation, transport, and absorption of electromagnetic radiation, and the wavelength-dependent properties of these processes. The term is also often used to include the detection and determination of the quantity, quality, and effects of such radiation. The term *photometry* describes these phenomena for the visible portion of the spectrum only. Photometry and its terms and dimensions are a result of normalizing (or attempting to normalize) the measurement of light to the response of the human eye.

William Herschel (1738–1822) not only discovered infrared radiation but also attempted to draw the first distribution of thermal energy as a function of wavelength, and thus he can be considered the father of radiometry. Johann Lambert (1728–1777) noted that the amount of radiated (and in some cases reflected) energy in a solid angle is proportional to the cosine of the angle between the emitter and receiver. Incidentally, Lambert also proved that π is an irrational number and introduced the hyperbolic functions sinh and cosh. A few decades later, Gustav Kirchoff (1824–1887) discovered that the emissivity of a surface equals its absorptivity and that the total of reflection, absorption, and transmission of a material always equals 1. Later, Austrian physicist Josef Stefan (1835–1893) determined the total radiant exitance from a source from all wavelengths to be equal to the emissivity multiplied by a constant (the Stefan–Boltzmann constant) times its temperature raised to the fourth power. In 1866, Langley used a crude bolometer to study the radiation of carbon at different temperatures. In 1886, Michelson employed Maxwell's laws to develop crude blackbody laws. Additionally, two famous rules of thumb (or useful approximations) that led to the Planck function were described by Wien and Rayleigh. Jeans found (and repaired) a numerical error in Rayleigh's equation, so it is now known as the Rayleigh–Jeans law. Others, such as Lummer and Pringsheim, made important pre-Planckian additions to blackbody theory.

The main architect of modern radiometry was Max Karl Ernst Ludwig Planck (1858–1947). Planck began his scientific career under the influence of Rudolf Clausius (developer of the second law of thermodynamics), giving him a strong background in thermal physics. Planck is most noted for describing the blackbody radiation in a simple equation and for developing the quantum theory of energy (which states that energy is not infinitely divisible but exists in units whose energy is defined by the frequency). He noted that the Rayleigh–Jeans approximation agreed well with experimentation for long

Peak Wavelength of Wien's Displacement Law

The peak wavelength (in microns) of a blackbody is approximately 3000 divided by the temperature in kelvin. That is,

$$\lambda_{peak} = \frac{3000\ \mu\text{m/K}}{T}.$$

Discussion

This result derives from taking the derivative of the Planck function and setting it equal to zero. This requires a numerical solution to a transcendental equation. According to Planck's law, a blackbody will have an energy distribution with a unique peak in wavelength. For a blackbody, this peak is solely determined by the temperature and equals 2898/*T*. This assumes the emitter is indeed a blackbody and not a spectral emitter.

Hudson[1] points out that about 25 percent of the total energy lies at wavelengths shorter than the peak and about 75 percent of the energy lies at wavelengths longer than the peak. Additionally, Hudson gives the following shortcuts:

- To calculate wavelengths where the energy is half of the peak (half power or at the 3-dB points), divide 1780 by the temperature in kelvin for the lower end and 5270 for the higher end.
- You will then find
  - 4 percent of the energy lies at wavelengths shorter than the first half-power point.
  - 67 percent of the energy lies between the half points.
  - 29 percent of the energy lies at wavelengths longer than the longest half-power point.[1]

Of course, there are other ways to describe blackbodies, and each has its own version of Wien's law. For example, the maximum occurs for photon emission from a blackbody when the product of wavelength and temperature equals 3670 micron-K.

Reference

1. R. Hudson, *Infrared Systems Engineering,* John Wiley, pp. 58–59 (1969).

Choice of Waveband

The long-wave infrared (LWIR) band from 8–12 microns provides the best combination of target (self-emission) radiance and minimal solar illumination effects (shadows, glare, solar blinding) that is particularly attractive for machine vision and automation. The mid-wave infrared (MWIR) band (3–5 microns) provides the best target signal for hot objects such as engines or thrusting vehicles.

Incorrectly Sizing Radiometric Areas

To properly determine the source or detector area, the imaging characteristics of the optical system need to be considered. Incorrectly identifying the source or detector area will result in incorrect area identification for throughput calculations.[1,2]

	Situation	Restriction	Figure
1	Sensor FOV is smaller than the target	Use the imaged area of the source	16.12
2	Full target size is smaller than the full focal plane	Use the illuminated area of the detector	16.13
3	Angular size of the target is smaller than the IFOV of a detector	Use the total area of the source and not the pixel area	16.14

Discussion

A natural mistake in radiometry is the misidentification of the areas in the throughput equation. There are two areas in the throughput equation $A\Omega$ (see the rule "Etendue" in this chapter). One is for the source, and the other is for the receiver. The areas can typically be switched; however, this cannot be done without considering the optical system. Typically, it is the physical size of the source or the detector that can be used; however, there are three exceptions to this rule.

First, if the source is too large to be imaged onto the detector, as shown in Fig. 16.12, the area of the source cannot be used. The imaged area must be used to get the correct answer. Second, if the source does not completely fill the detector, then only the illuminated portion of the detector should be used in the throughput calculation (see Fig. 16.13).[1] Finally, the third case occurs when the source is smaller than a single pixel on the detector.[2] The object can be considered a sub-pixel object or a point source. For a sub-pixel target, the calculation must use the target area and the solid angle of the sensor aperture, as seen in Fig. 16.14. Additionally, the optical system's PSF should be used to determine the amount of light falling onto a single pixel.

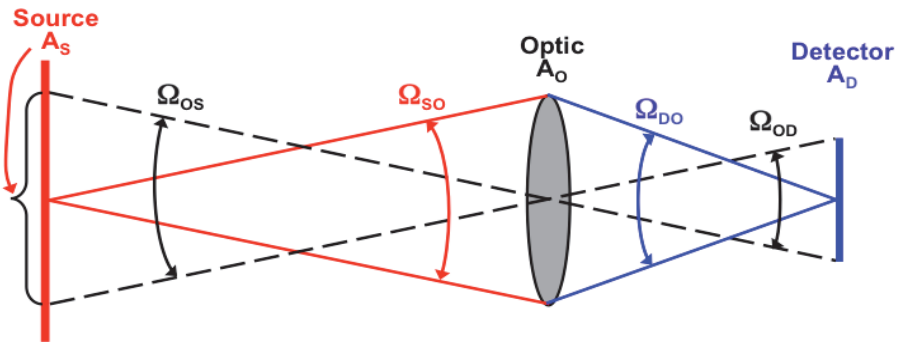

Figure 16.12 Example of a super-sized source.

Chapter 18
Target Phenomenology

Introduction

Generally, the properties of targets and their signatures, such as are summarized in this chapter, fall into the domain of the military designer. However, increasingly, many of these rules apply to non-military segments of the photonics market, such as machine vision, self-driving vehicles, security cameras, paramilitary organizations, search and rescue, homeland defense systems, environmental monitoring, general surveillance, remote diagnostics, remote sensing, and industrial security.

This chapter provides a brief look into the shortcut characterizations that were largely developed to assess the signatures of various potential targets for typical EO systems. Regardless of their heritage, several of these rules are applicable in the generic business of assessing what a sensor might be able to detect, recognize, or identify.

Although most of these rules were developed for the infrared spectrum, they illustrate important principles that may be applied (albeit with caution) to other parts of the spectrum, including UV, visible, millimeter-wave, and imaging radar.

Often, targets of interest to the EO sensor designer consist of the metal body and frame containing some kind of engine (such as an airplane or car). Such a target can be detected by sensing the metal hardbody (e.g., the roof of the car), the hot engine compartment (the heat dissipated from under the hood), or the spectral engine emission (e.g., the hot CO_2 coming out of the tailpipe). The emission of hot gases and particles is generally called a *plume*. Although all internal combustion engines produce significant plumes, those of jet engines and rockets draw the most attention.

Rocket and jet plumes have long been of interest to EO designers, as these provide bright signature-to-background ratios. Much of the early work was done in the 1950s and 1960s in remote plume diagnostics by EO instruments for jet and rocket engine development. Additional effort was expended in the 1960s on large and small rocket plume signatures, thanks to the space and arms races. The signatures of tactical air-to-air and surface-to-air missiles were investigated in the hope of providing effective countermeasures. Plume investigations of large rockets continued in support of early warning efforts. Maturation of this study and the perceived need were formalized regarding missile defense.

As expected, there has been considerable interest in the reflectivity of target surfaces, because this determines the amount of laser tracker radiation that can be expected

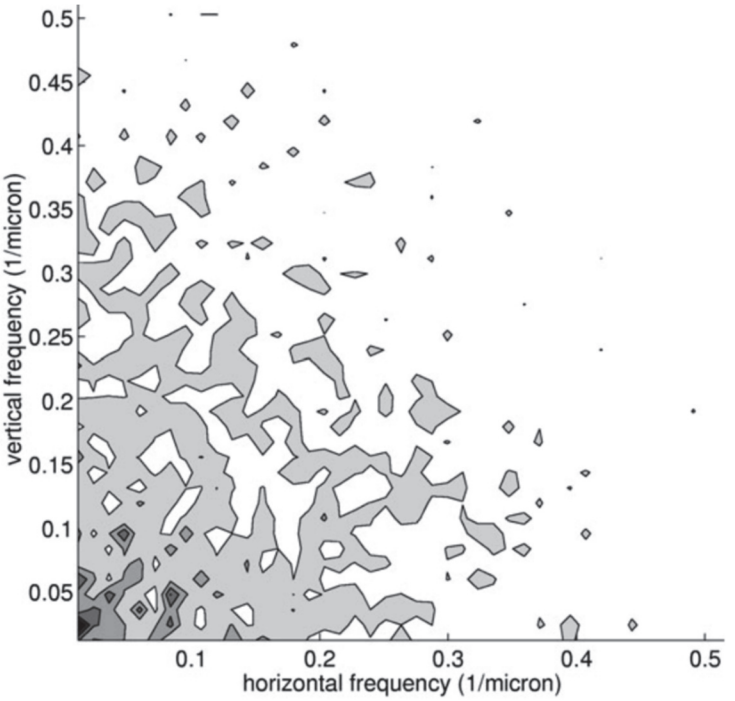

Figure 18.2 A discrete Fourier transform of Fig. 18.1. The strength of the darkness denotes an increase in power. This shows that most of the structure is concentrated at a scale of 10 microns or larger, which indicates that the surface should be black at those wavelengths (which it is). Image reprinted courtesy of The Research Triangle Institute.

The above is useful only for a quick estimate of an object's emissivity when little is known about it. The appendix has a table of emissivities for common materials. All of these tables are approximate infrared emissivities and should be used with caution, as emissivity varies with temperature, wavelength, and surface roughness. There are several emissivity libraries on the internet.

Solar Reflection Always Adds to the Signature

The Sun is bright. When present, solar reflection always adds signature to a target, regardless of the bandpass. Specific impacts depend on the conditions, sensor, and band pass.

Discussion

Imaging systems detect reflected light, emitted light, or both. Increases in signature from solar reflection are usually great for wavelengths less than about 3 microns and inconsequential beyond about 5 microns (see the associated rule in the chapter on Backgrounds). However, beyond 5 microns, the target may absorb solar radiation, causing an increase in temperature. Solar reflection is usually measurable but not significant between these two wavebands.

Table 18.1 Radiant intensity as a function of band pass that does not consider heating but does consider solar reflection.

	0.4–0.6-µm band	3–5-µm band	8–12-µm band
Thermal radiant intensity (W/sr)	Essentially zero	1.5	28
Solar reflection (W/sr)	35	2.2	0.7
Total (W/sr)	35	3.7	29
Percent of signature contribution by the Sun	100%	60%	2%

References

1. P. Jacobs, *Thermal Infrared Characterization of Ground Targets and Backgrounds*, SPIE Press, pp. 34–37 (1996).
2. rredc.nrel.gov/solar//spectra/am1.5 (2020).

Lambertian vs. Specular

No surface is perfectly Lambertian or specular. Generally, a surface is considered Lambertian if its bidirectional reflectance distribution function (BRDF) peak reflection at angles near normal incidence is less than one order of magnitude above the average. Conversely, a surface is generally considered specular if its peak is four orders (or more) of magnitude above its average.

Discussion

A perfectly Lambertian surface emits equally over 2π steradians. A perfectly specular surface emits in an infinitesimally small angle determined by Snell's law. Nothing is perfect in nature, so surfaces behave like a combination of Lambertian and specular. Note that these definitions do not have anything to do with total reflectance. Surfaces exist that are very low in reflectance yet are still very specular, and vice versa.

For example, gloss black paint has an overall reflectance that is quite low, yet it is specular. An automobile with a high-quality black finish will provide a very nice reflective image. Paper is the opposite: it is designed to be of high reflectivity but is also intended to be Lambertian (diffuse). Most readers find that the glossy paper used in some books and magazines is annoying, because the reflection of the light source can be distracting. The same is true of some computer screens. The characteristics desired in a target, background, or hardware surface treatment must be carefully analyzed and should be determined by statistical ray traces.

Figure 18.3 illustrates the difference between a notional specular (or mirror-like) surface and a notional Lambertian surface. If an incident beam encounters the surfaces at a 45° angle from normal, the Lambertian surface will have about the same level of reflectance at all observing angles. The specular surface will generally have a lower reflectance at all angles except near the Snell reflection angle, where it will be many orders of magnitude greater.

on the surface morphology, which can be a function of temperature. Generally, the surface should be rough and cone-like at a scale greater than the wavelength to trap the photons. When cooled, most surfaces contract. A given structure that is very "black" at room temperature because of its surface morphology will contract when cooled to cryogenic temperatures. This surface morphology change can result in the surface being more specular and/or reflective at the wavelength of interest.

Note that some BRDF measurements include cosine corrections, while others do not. Additionally, caution is advised as the angles are often defined differently.

References

1. F. Nicodemus, et al., *Geometric Considerations and Nomenclature for Reflectance*, NBS (now NIST) Monograph **160** (1977).
2. J. Stover, *Optical Scattering*, SPIE Press, pp. 19–22 (1995).
3. J. Conant and M. LeCompte, "Signature Prediction and Modelling," pp. 318–321, in *Electro-Optical Systems Design, Analysis, and Testing,* M. Dudzik, Ed., Vol. 4 of *The Infrared and Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, Eds., Infrared Information Analysis Center and SPIE Press (1993).
4. W. Wolfe, "Radiation Theory," Ch. 1 in *The Infrared Handbook*, W. Wolfe and G. Zissis, Eds., ERIM, pp. 1-30 to 1-31 (1978).
5. W. Wolfe, "Optical Materials," Ch. 7 in *The Infrared Handbook*, W. Wolfe and G. Zissis, Eds., ERIM, pp. 7-78 to 7-79 (1978).

Hagen–Rubens Relationship for the Infrared Reflectivity of Metals

The reflectivity R of a metal at a particular wavelength is related to its DC conductivity by[1]

$$R \approx 1 - 2\sqrt{\frac{\omega}{\sigma}},$$

where ω is the frequency of the radiation, and σ is the DC conductivity of the metal.

Discussion

This estimation assumes that the emissivity is measured perpendicular to the surface plane, that the wavelength is not shorter than 10 microns, and that the material is a good conductor.

The amount of light reflection of a given (solid) object depends on the exterior surface material, roughness, and surface coating. The reflectivity of a metal is related to the complex index of refraction of the surface coating. This can be estimated for a given wavelength by the metal's absolute magnetic permeability and electrical conductivity.

Note that the equation in the rule may appear as follows if given in SI (international) units:[2]

$$R \approx 1 - 2\sqrt{\frac{2\varepsilon_0\omega}{\sigma}},$$