A GUIDE TO THE
Use and Calibration of Detector Array Equipment

A GUIDE TO THE

Use and Calibration of Detector Array Equipment

Gordon R. Hopkinson
Teresa M. Goodman
Stuart R. Prince

Published in cooperation with Sira Technology Ltd. and NPL (UK)

sira NPL
National Physical Laboratory

SPIE PRESS
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Hopkinson, Gordon R.
A guide to the use and calibration of detector array equipment / Gordon R. Hopkinson , Teresa M. Goodman, Stuart R. Prince.
p. cm.
Includes bibliographical references and index.
ISBN 0-8194-5532-6 (alk. paper)
1. Detectors--Calibration. I. Goodman, Teresa M. II. Prince, Stuart R. III. Title.

TA165.H67 2004
681'.25--dc22 2004016376

Published by

SPIE—The International Society for Optical Engineering
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: +1 360 676 3290
Fax: +1 360 647 1445
Email: spie@spie.org
Web: http://spie.org

Printed in the United States of America.
PDF ISBN: 9781510607927

CONTENTS

INTRODUCTION

The original version of this user guide appeared in March 2000 and resulted from a program of work sponsored by the National Measurement System Policy Unit (NMSPU) of the U.K. Department of Trade and Industry (DTI). The aim was to benefit industry by identifying the factors influencing the accuracy of measurements obtained with detector array equipment, by developing calibration techniques, and by establishing best practice procedures. It was written primarily for users of instrumentation rather than for detector specialists, and produced as a booklet mainly for small-scale distribution within the U.K. Hopefully, however, the guide has enough detail that even these "experts" will find it useful.

The reception to the original version was extremely favorable and we were therefore encouraged to make it more widely available. In revising the guide for a more worldwide readership, we have tried to leave the majority of the text intact, but to remove those parts that were felt to be overly U.K.-centric.

We also took the opportunity to bring the guide up to date by adding brief sections on electron multiplying (EM) CCDs for low-light-level imaging and on lag effects in CMOS active pixel sensors (APS). The list of references has also been revised and significantly extended. The original guide contained many references to information available via the Web. These have been checked and updated where necessary, and it is hoped that the book will carry on being an up-to-date reference guide for users of detector array instrumentation for a considerable time to come.

Solid state imaging arrays have many advantages as sensors of optical radiation compared with discrete, single-element detectors. For example, the large number of elements gives fine spatial sampling when imaging and a reduction in the measurement time when performing spectrometry. On the other hand, the large number of pixels and the complexity of array structures lead to special issues for measurement and calibration; issues which, until now, have not been addressed in any systematic way. These can arise both for the developer of array-based instrumentation (who wishes to specify, measure, and predict detector performance for his application and to provide an adequate calibration) and for the user (who may wish to assess and maintain the accuracy of his overall system).

At the start of the original study a survey was made, by questionnaire and selected interviews, of U.K. users and manufacturers/developers. This was undertaken with the help of Atkins Management Consultants, with the objective of determining equipment usage and calibration requirements and obtaining views on the preferred content and format of the user guide. Key parameters and calibration methods were then assessed in detail and measurements on typical systems performed. This work culminated in the preparation of this guide, which contains both a review of the existing literature and a large amount of new experimental data (obtained during the course of the work). The main emphasis has been on UV, visible, and near-infrared systems that use silicon detector technology, but an attempt has also been made to address the issues arising in thermal imaging with infrared detector arrays.

An outline of the guide is presented in the table on the following page. In understanding requirements for instrument calibration, it is important to realize that the detector array is not simply a black box that translates optical to electrical signals, but is a complex assemblage in which a variety of physical process are taking place (for example, optical absorption, charge transport, and noise/dark current generation). Even in detectors for the visible region there are several detector forms, each with its own characteristics; and in the infrared the variety of array technologies multiplies rapidly. Hence, the guide starts (in Chapter 1) with an overview of detector technologies and key performance parameters. It is hoped that Sec. 1.2, which deals with the detailed definition and measurement of detector parameters will be of interest, not only to detector manufacturers and instrument developers, but also to users—first, because many of the techniques described can be used at instrument level (although more detailed guidance for users of array based systems can be found in Chapter 4); and second, since some insight is provided into the dependence of the results on measurement conditions. Chapter 1 also discusses measurement units, which are often a considerable source of confusion.

Chapter 2 gives a brief introduction to the various types of instrumentation that use detector arrays and the requirements for calibration. Chapter 3 discusses in some detail the influence of operating conditions on detector performance. It is these changes in performance which often result in the need for instrument calibration.

Further information on calibration techniques is given in Chapter 4, where there is a special emphasis on the calibration of instruments for spectrometry. Chapter 5 reviews the use of common items of calibration equipment, such as light sources, filters, and ceramic tiles, and gives guidance on the selection of components. Chapter 6 briefly reports on international efforts to harmonize measurement standards.

Each chapter is provided with a list of literature references, which the reader is encouraged to consult for further information. Sources of information of more general interest are included in Appendix A. Throughout the guide, references are made to information available on the Web. Though these resources are somewhat ephemeral (websites and their content may change), the wealth of easily accessible information considerably eases the task of assessing technology trends and sharing standards information. Appendix A also includes a list of abbreviations, and a glossary of terms can be found in Appendix B. Appendix C summarizes the recommended practices for spectrometer and imager calibration in a Quick Guide.

Gordon Hopkinson
Teresa Goodman
Stuart Prince
August 2004

Outline of the User Guide

Interest	Chapter /Section
General	
Reference guide	Appendix A
Glossary	Appendix B
Detector arrays	
Types of detector array (CCDs, CIDs, APS, PDAs, etc.)	1.1
Key parameters and their measurement	1.2
Influence of operating conditions on array performance and instrument calibration	3
Instrumentation	
Types of instruments and their calibration requirements	2
Calibration	
Imagers (including color and thermal imagers)	4.1, 2.1
Spectrometers (including Raman and imaging spectrometers)	4.2, 4.3, 4.4
Equipment (e.g., sources and filters)	5
International standards	6
Quick Guide	Appendix C

CHAPTER 1

DETECTOR ARRAYS

In order to understand the performance limits and calibration requirements of an instrument it is necessary to have an understanding of the nature of the detector array, its key performance parameters, and the way these are defined and measured. These are all considered in this chapter, which serves to introduce many of the detector terms that are discussed later in the guide. At first sight there is a confusing multitude of array types, parameters and definitions, and an extensive published literature; however, it is hoped that the discussion below can act as a basic summary and a pointer to references for further reading, should these be required.

1.1 Types of Detector Arrays

This discussion is concerned exclusively with instruments that use detector arrays fabricated using semiconductor materials (sometimes termed solid-state detector arrays). Such arrays have two parts:

1. A mosaic of closely spaced detector elements that convert the incident electromagnetic radiation into electrical signals,
2. A readout circuit that relays and multiplexes the electrical signal from each detector element (or pixel) to a small number of output amplifiers (usually one or two).

Although both functions may (as in a visible CCD) be realized with the same basic structure, on a single piece of material (die) this is not necessarily the case. Hence, it is important when considering the properties of the different types of solid-state arrays that the various architectures are distinguished.

Regarding the detector elements, Fig. 1.1-1 illustrates the wide variety of semiconductor materials that can be used and their wavelength ranges and typical operating temperatures. Most are *photon* detectors in that they respond only to those photons that have an energy greater than the band gap of the semiconductor material. The detector response is proportional to the rate of absorption of these photons. There is another class of *thermal* detectors, which includes the pyroelectric and microbolometer arrays used for near–room temperature (usually termed *uncooled*) sensing in the infrared. In these the detector measures the total amount of heat that is absorbed, regardless of the number of photons.

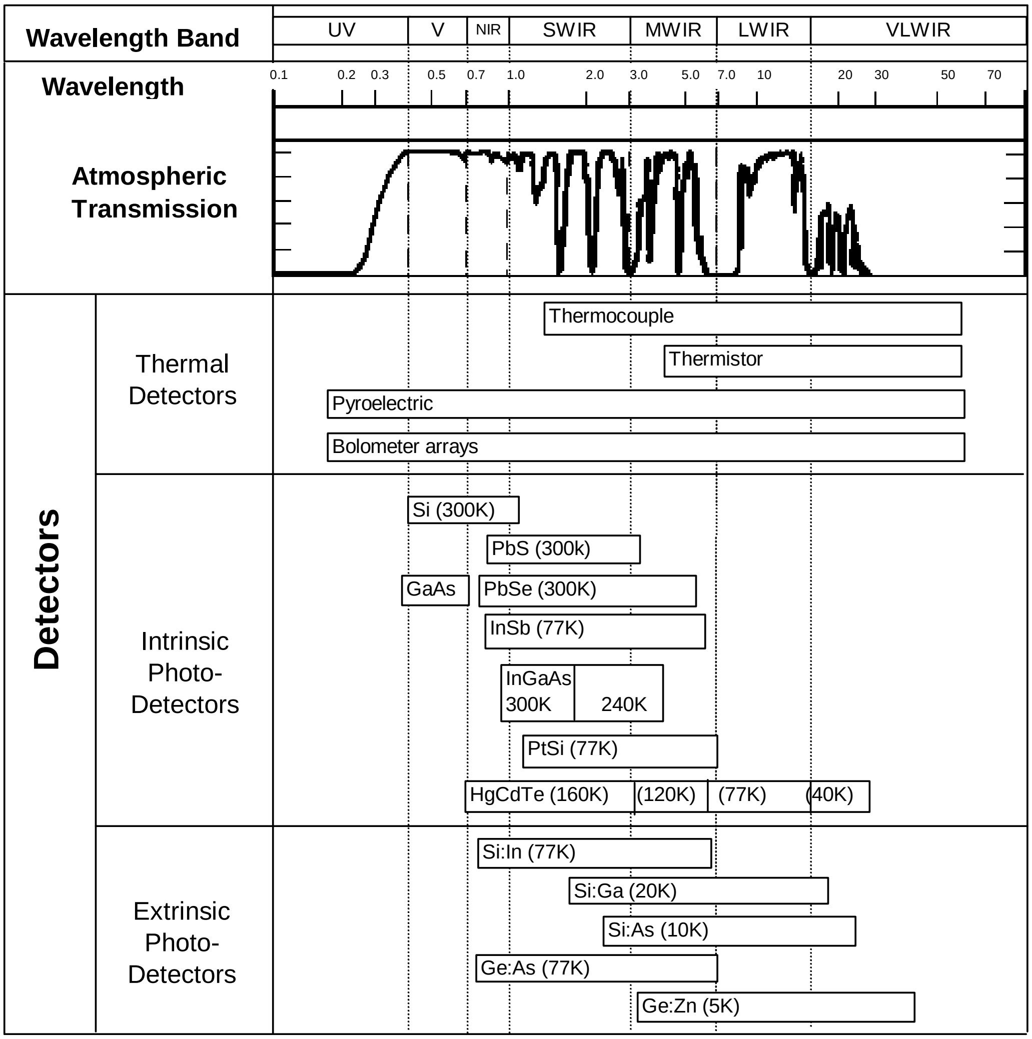


Figure 1.1-1 Semiconductor detector materials.

As well as different detector materials, there are also a variety of detector constructions. Photon detector materials can be used in three basic types of structures:

1. photoconductive (PC) mode detectors, where the absorbed photons cause a change in conductivity. When a constant bias is placed across the detector, the modulation in the current through the semiconductor is proportional to the change in irradiance.
2. photovoltaic (PV) mode detectors, where the absorbed photons create a change in voltage at a potential barrier either produced by
   - a p-n junction or
   - induced by electric fields with a metal-insulator-silicon (MIS) structure.

3. photoemissive detectors, such as silicide Schottky barrier devices. In these, photons are absorbed in the silicide electrode and a small fraction of the excited carriers are emitted over the Schottky barrier into the semiconductor. The charge generated at the electrode is then transferred to a readout structure.

There are also a number of different designs for readout electronics. Among these are

CCD (charge-coupled device) readouts
MOSFET switch readouts (including CMOS circuits)
CID (charge injection device) readouts
CIM (charge imaging matrix) readouts.

Many of the above detector types and readout circuits have been developed for imaging in the infrared. For wavelengths less than 1100 nm (1.1 μm) the choice is largely restricted to silicon detectors operating in a photovoltaic mode (either as photodiodes or as MOS capacitors). Only these will be considered in detail in the remainder of this guide, since the chief interest is in imaging and spectrometry applications in the ultraviolet (UV), visible, and near-infrared (NIR). Occasional reference will, however, be made to calibration issues for IR detectors, many of which will be similar to those for silicon detectors (see, in particular, Sec. 4.1.2). For more information on the basic principles of IR arrays the reader is referred to the following general reviews:

Harnly et al., "Solid-state array detectors for analytical spectrometry,"[1]
Scribner et al., "Infrared focal plane array technology,"[2]
Norton, "Infrared image sensors,"[3]
Rogalski, "New trends in semiconductor infrared detectors,"[4]
Rogalski, "Infrared photon detectors,"[5]
Crowe et al., Chap. 4 in *The Infrared & Electro-Optical Systems Handbook (Vol. 3)*.[6]

Since CMOS and MOS circuits are commonly used for readout of IR arrays, these are briefly reviewed in Sec. 1.1.6.

1.1.1 Types of Silicon Detector Arrays

As mentioned above, this guide is primarily concerned with silicon detectors. Photons are absorbed via the photoelectric effect and create electron-hole pairs. Since a photon has an energy hc/λ where h is Planck's constant, there is a limit on the longest wavelength that can be detected. This is given by

$$\lambda_{\text{cutoff}}\left(\mu\text{m}\right) = \frac{1.24}{\text{bandgap}\left(\text{eV}\right)}. \qquad (1.1.1\text{-}1)$$

Silicon has a bandgap of 1.1 eV and is sensitive up to roughly 1.1 μm. There is roll-off in response from about 0.8 μm upwards because the detector is thin (about 20 μm) and the absorption length becomes large. At short wavelengths (blue and UV) there is also a limit because the absorption length becomes small and photons are absorbed in surface layers and do not get through to the silicon. In the visible region, one electron-hole pair is created for each photon absorbed. The holes diffuse to the substrate and recombine, the electrons are collected within a pixel and give rise to the signal charge. The basic response of silicon detectors is shown in Fig. 1.1.1-1. At x-ray and far-UV wavelengths, each photon has enough energy to create several electron-hole pairs (in fact, one for every 3.6 eV of photon energy). So, a 20-keV x-ray (e.g., from a medical x-ray tube) will produce about 6000 electron-hole pairs by direct interaction with the silicon.

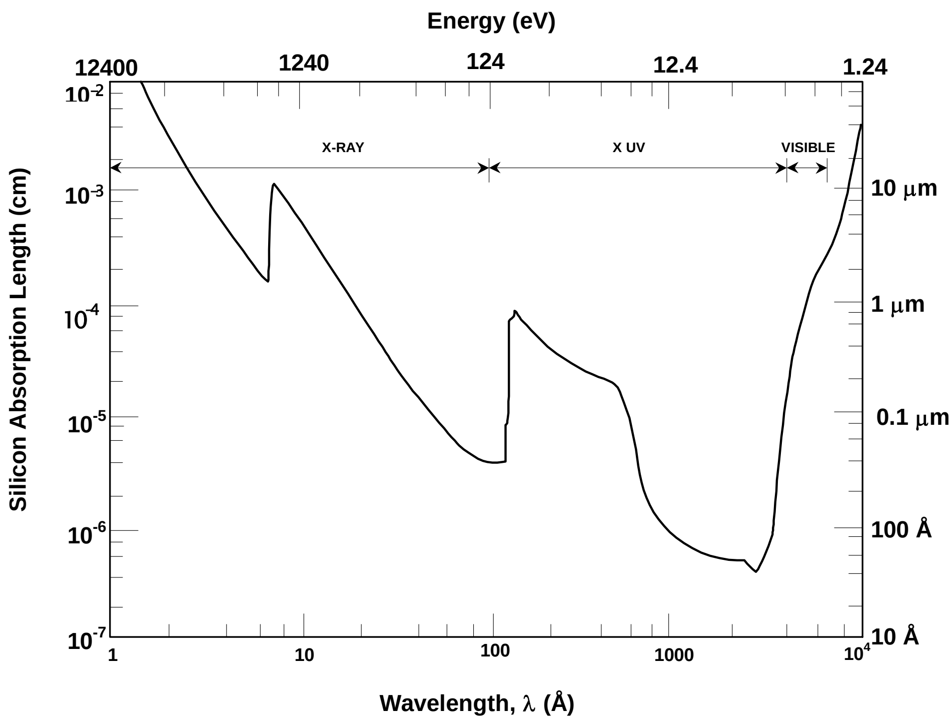


Figure 1.1.1-1 Wavelength response of silicon.

Silicon detectors operate in the photovoltaic mode and there are two basic detector types; the photodiode and the metal oxide silicon (MOS) capacitor. These are illustrated in Fig. 1.1.1-2. Photodiodes have only a small part of the surface covered by the contact electrode. The MOS capacitor has a semi-absorbing gate electrode and tends to have a reduced blue response when front-illuminated (c.f. Fig. 1.1.1-3). An alternative is to thin the detector and illuminate from the back. With front illumination the blue response can be improved if polysilicon, the standard electrode material, is replaced with indium tin oxide (ITO) or if "windows" are created in the electrode structure (as with *open-phase* and *virtual-phase* devices).[7] Blue/UV response can also be boosted by using a phosphor coating (e.g., lumogen). For x-ray sensitivity, a phosphor such as

CsI or gadolinium oxysulphide (Gadox) can be used as an alternative to the direct interaction discussed above.

In some applications (for example, imaging), the red response of the detector (out to 1.1 μm) is unwanted (because images are blurred due to chromatic aberrations in the lens). Hence, some commercial CCD cameras use detectors with a thin (~2 μm) active region (implemented via a special drain structure) to "kill" the red response. Other applications (for example, Raman spectroscopy) need good IR response and sometimes use a thick (e.g. 50 μm) active region with low doping, so that the depletion region is also thick and spatial resolution is not degraded (these are often called *deep-depletion* devices).

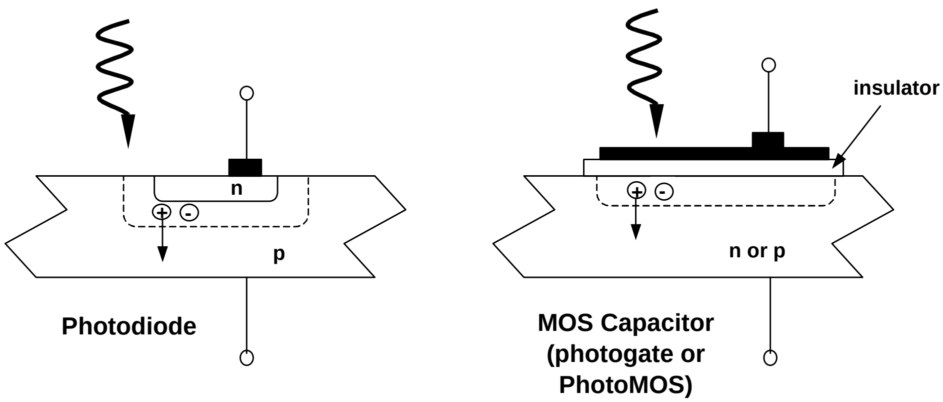


Figure 1.1.1-2 Photodiode and MOS capacitor structures.

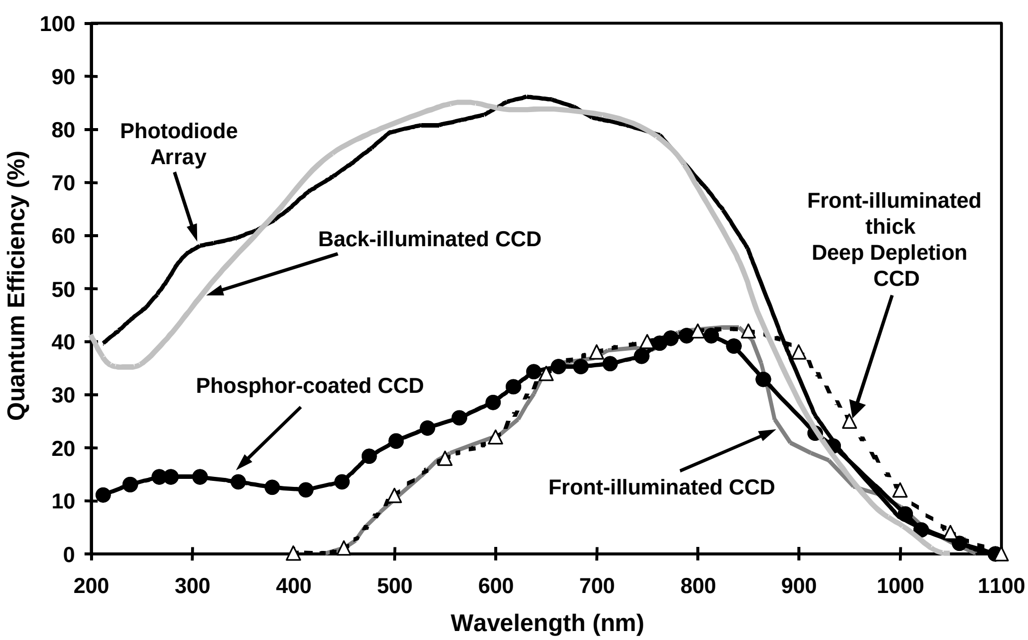


Figure 1.1.1-3 Quantum efficiency for typical CCDs and photodiodes.

As discussed in Sec. 1.1, there are also two basic types of readout for silicon sensors: the charge-coupled device (CCD) and the MOS X-Y addressed switch array. In a CCD the photo-generated charge is moved across the chip towards an output amplifier by applying a clocking sequence to the CCD electrodes. A typical sequence for a three-phase device is illustrated in Fig. 1.1.1-4. CCDs can be manufactured with two, three, or four phases. Two-phase devices are the easiest to clock but need additional implants during manufacture to define the direction of charge transfer. The following simple sequence illustrates the basic principle of operation, but it should be noted that CCD manufacturers will define particular waveforms that are optimized for their own devices.

When an electrode is pulsed high, a potential well is created underneath it. At point A in the clock waveform (left of Fig. 1.1.1-4) charge collects under phases 1 and 2. When phase 1 goes low (point B) the charge is squeezed under phase 2, and when phase 3 goes high it is shared between phases 2 and 3. The sequence carries on, and at the transition F to A we arrive back at the starting point of the clock sequence, at which point it can be seen that charge has moved sideways by one pixel.

There are four basic types of readout architectures for a CCD: frame transfer (FT), full frame (FF), interline transfer (ILT or IL), and frame interline transfer (FIT). These are shown in Fig. 1.1.1-5.

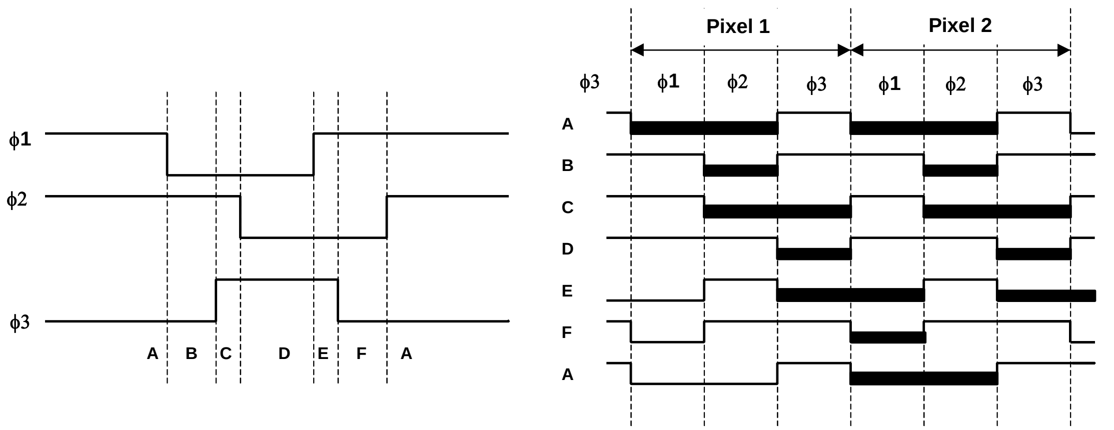


Figure 1.1.1-4 Typical CCD clocking sequence for a three-phase device.

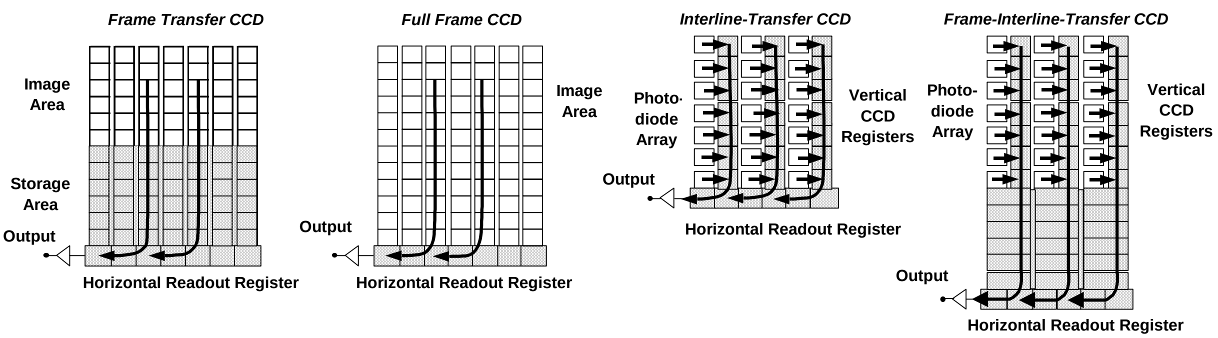


Figure 1.1.1-5 CCD architectures. The directions of charge transfer are shown arrowed. The shaded areas are shielded from light by a metallized layer on the chip.

FF CCDs have the simplest architecture and are the easiest to make and operate. They consist of a parallel CCD shift register, a serial CCD shift register, and a signal-sensing output amplifier. Images are optically projected onto the parallel array, which acts as the image plane. The resulting rows of scene information are then shifted in a parallel fashion to the serial register, which shifts each row of information to the output as a serial stream of data. The process repeats until all rows are transferred off the chip. Because the parallel register is used for both scene detection and readout, a mechanical shutter or synchronized strobe illumination must be used to preserve scene integrity. The simplicity of the FF design yields CCD imagers with the highest resolution (for a given chip size)

The FT CCD architecture is similar to that of the FF CCD, except that FT CCDs have a separate, identical parallel register, called a storage region, which is not light sensitive. The idea is to quickly shift a captured scene from the image array to the storage array. Readout off chip from the storage register is then performed as described in the FF device, while the storage array is integrating the next frame. The advantage of this architecture is that a continuous or shutterless/strobeless operation is achieved, resulting in faster frame rates. Performance is compromised, however, because integration is still occurring during the image dump to the storage array, which results in image smear (see Sec. 3.1.16). Because twice the silicon area is required to implement this architecture, FT CCDs have lower resolutions and higher costs than FF CCDs.

ILT CCDs address the shortcomings of FT devices by separating the photo-detecting and readout functions with isolated photosensitive regions between lines of nonsensitive or light-shielded parallel readout CCDs. After integrating a scene, the signals collected in the pixels are simultaneously transferred to the light-shielded parallel CCD. Transfer to the output is then carried out in much the same way as in FF and FT CCDs. During readout, the next frame is being integrated, thus achieving a continuous operation and a higher frame rate. Because of the architecture, the image smear that occurs in FT CCDs during readout is significantly reduced. It is not entirely eliminated, however, because photons can migrate sideways from the photodiode sensor to the vertical CCD shift register either by scattering or by multiple reflections in the dielectric layer under the gate electrode structure, a process known as *light-piping*.[8] The FIT CCD was invented to reduce the smear that is still present in ILT CCDs.

The FIT CCD has a storage region (as in the FT) CCD so that charge does not need to stay in the vertical CCD registers during serial readout of a line. Hence, it will pick up less smear from the scene being imaged.

The major disadvantages of ILT and FIT CCD architectures arise from the complexity of the devices, which leads to higher unit costs and lower sensitivity. Lower sensitivity occurs because less photosensitive area (i.e., a reduced aperture) is present at each pixel because of the associated light-shielded readout CCD. Furthermore, quantization, or sampling errors, are greater because of the reduced aperture (hence, aliasing effects are more pronounced). Lastly, some ILT and FIT architectures using photodiodes suffer image lag as a consequence of charge transfer from the photodiode to the CCD.

With the frame transfer and full frame devices it is quite common to have output amplifiers at each end of the serial readout register so as to reduce the readout time. There is often the choice to use just one amplifier (and have just one video channel) or to use both. There is a penalty in the additional circuitry (and pin-out) and so in cost—low-cost cameras will usually have a single video output. A variant of the frame transfer array is to have two storage regions, one at the top and one at the bottom of the image area. Half of the image is transferred up into the top storage region and half to the bottom. This results in a faster readout (two serial registers are used) and half the smear (since the time taken to transfer into the storage region is halved). Again, there is a penalty in system complexity and cost.

X-Y addressed MOS and CMOS arrays are shown in Fig. 1.1.1-6. The vertical (row) register is used to connect a whole row of detectors and the horizontal (column) register is used to connect each pixel in that row, in turn, to the output amplifier. This type of readout has the advantage of random access (not all the pixels need to be read out) but the disadvantage of high readout noise (because of the high capacitance of the readout lines) unless an amplifier is incorporated into each pixel as in Fig. 1.1.1-6(b)). Also, part of the detector area is taken up by the switching transistors, so the fill factor is low.

In summary, we have two detector types (photodiodes and MOS capacitors) and two basic types of readouts (CCD and X-Y addressed). The combinations of these are shown in Fig. 1.1.1-7. The abbreviations are defined as follows:

CCD	charge-coupled device
CID	charge injection device
CCPD	charge-coupled photodiode array—a linear photodiode array with a linear CCD alongside (sometimes on both sides) for readout
ILT CCD	interline transfer CCD
FIT CCD	frame interline transfer CCD
FT CCD	frame transfer (including full frame) CCDs

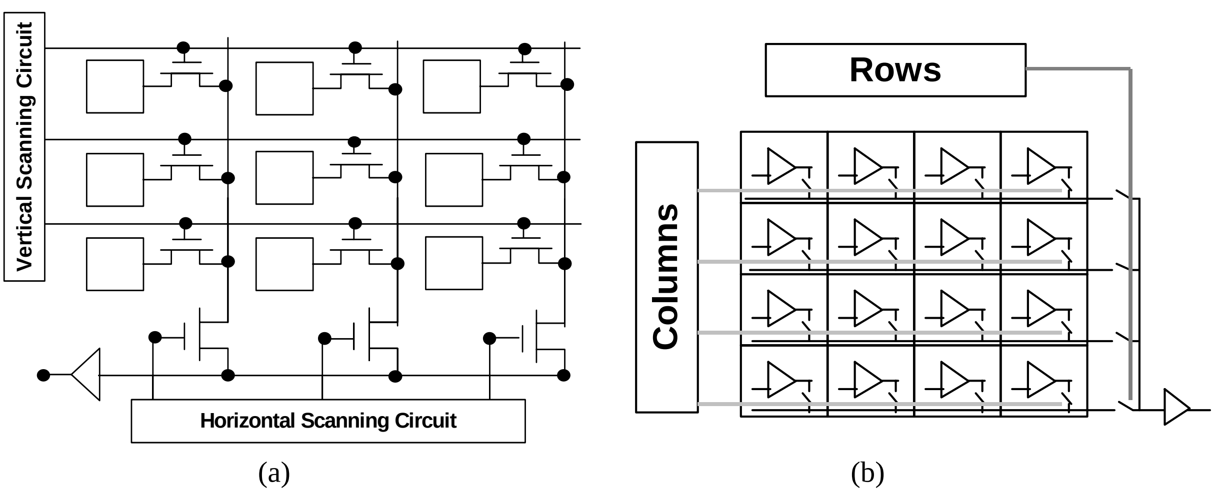


Figure 1.1.1-6 Readout structure for (a) MOS and (b) CMOS-APS X-Y addressed arrays.

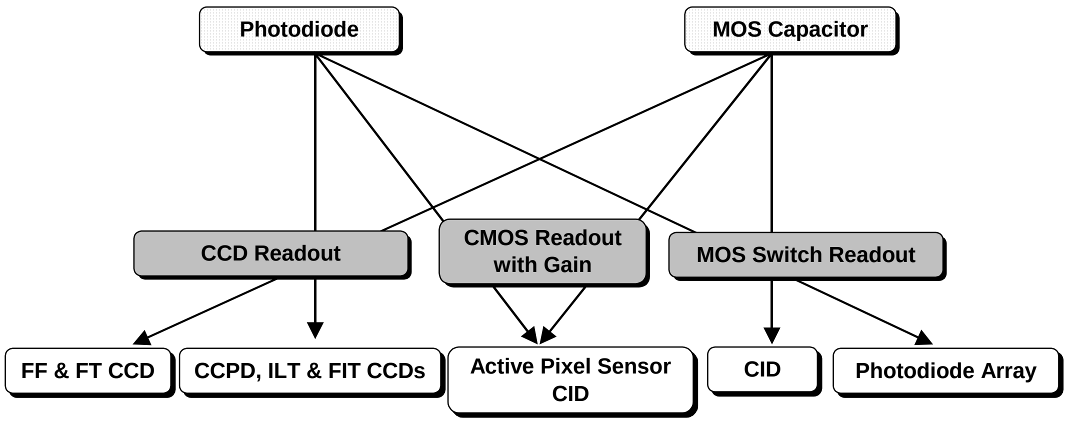


Figure 1.1.1-7 The combinations of detector element and readout structure that are possible with silicon sensor arrays.

The performance of these detectors is briefly summarized below. The performance parameters are considered in more detail in Sec. 2.2.

CCD detector: MOS capacitor and CCD readout. This has good noise performance (can be as low as a few electrons), high fill factor, modest full well capacity (a few $\times 10^5$ electrons) but poor blue response (unless using a special electrode structure, phosphor coating, or back illumination).

Charge injection device (CID): MOS capacitor and X-Y readout. This has the possibility for random access but high readout noise due to the clock line capacitance (and low output voltage). However, there is the possibility for nondestructive readout (in which case noise can be reduced by averaging readouts).

Photodiode array (PDA): photodiode and X-Y readout. This has good blue response and high full well capacity ($>10^6$ electrons) but high readout noise. Hence, photodiode arrays are usually used for high-light-level applications.

Charge-coupled photodiode array (CCPD): photodiode and CCD readout. In linear arrays (as used in linescan cameras), it is possible to have the best of both worlds and have a photodiode detector (for good blue response and high full well capacity) and CCD readout register alongside (for low noise). Many 2D interline transfer and frame-interline transfer CCDs in fact use photodiodes as the detector elements. Also, many so-called photodiode arrays (PDAs) will have a CCPD structure and in common use the term PDA usually applies to a linear array with either a MOS (true PDA) or a CCPD structure. The term LPDA or linear PDA is also common.

The active pixel sensor (APS): photodiode or MOS capacitor with CMOS X-Y readout and gain. The active pixel sensor (APS) can have either photodiodes or MOS capacitors (photogates) as the detector elements and uses a CMOS switch array for readout with a gain transistor (MOSFET) within each pixel.[9] This improves the noise performance since the detector is isolated from the capacitance of the address lines, but leads to increases in fixed pattern noise because of variations in the offset voltage of each gain transistor. Full well capacity is similar to that for a CCD, and readout noise is somewhat higher.

The CMOS readout technology reduces cost, simplifies drive requirements (only a few TTL clocks are needed) and allows for additional circuitry on-chip, for example, digitization (on-chip ADC), edge detection, thresholding, motion detection and spatial filtering. Random addressing is also possible in some devices. APS devices are ideally suited to many low-cost imaging and machine vision applications and the performance can in some circumstances rival that of scientific CCDs (i.e., CCDs used for scientific rather than imaging applications).

Until recently most active pixel sensors were *monolithic*. That is, both the detector elements (e.g., photodiodes) and the CMOS switching circuits were manufactured on the same silicon die. These devices are sometimes called MAPS (monolithic active pixel sensors). Over the past year or so, hybrid silicon APS devices have become commercially available and these provide an alternative for visible imaging.[10] In hybrid devices the detector and readout circuits are formed on separate dies that are indium- or solder-bump-bonded together. This is a common situation for infrared arrays where the detectors are necessarily made of semiconductors other than silicon. However, though expensive, the advantages in using hybrid technology in the visible include 100% fill factor and compatibility with back illumination. Also, the readout circuit can be more complex, giving advantages in readout noise.

1.1.2 Color Imaging

The most common method of achieving 2D color imaging in high-resolution (e.g., broadcast-quality) systems is to use three CCDs (one for each primary color) and to split the light between them by means of a prism. For more commercial applications, it is common to use either stripe or mosaic color filter arrays (CFAs) placed directly on to the detector array (Fig. 1.1.2-1). Further information is given in Refs. 8 and 11. Reference 12 gives an interesting discussion on the trade-offs between the different CFA patterns.

Stripe filters are usually used in conjunction with frame transfer CCDs. For example, each pixel in a column being covered by a red, green, or blue filter (in fact, there are twice as many green pixels as either red or blue, because the eye is more sensitive to green light and because the coding for TV color signals has a luminance signal, which is sum of all the colors but weighted towards the green). There is an advantage in using the complementary colors cyan and yellow since then only two (overlapping) filters are needed (the overlapped combination of the two gives green). There is another advantage in that a green signal is present in each pixel and can be used to derive a high-spatial-frequency component in the image ("high-frequency green").

For interline transfer CCDs the number of horizontal pixels is at a premium (since the vertical shift registers also need to be accommodated). Hence, it is usual to use a matrix (or mosaic) of filters arranged in, for example, 2×2 pixel blocks. In this way, the horizontal resolution is half that of a monochrome imager (as opposed to a third for a stripe filter), but the vertical resolution is also degraded by a factor of 2. Again, complementary color filters can be employed. Video processing ICs (ASICs) are used with mosaic CCDs to perform the arithmetic for deriving the RGB data necessary for display. There are several possibilities for color imaging with a linear array, depending on the trade-off between spatial resolution and data rate, as shown in Fig. 1.1.2-2 (after Ref. 8).

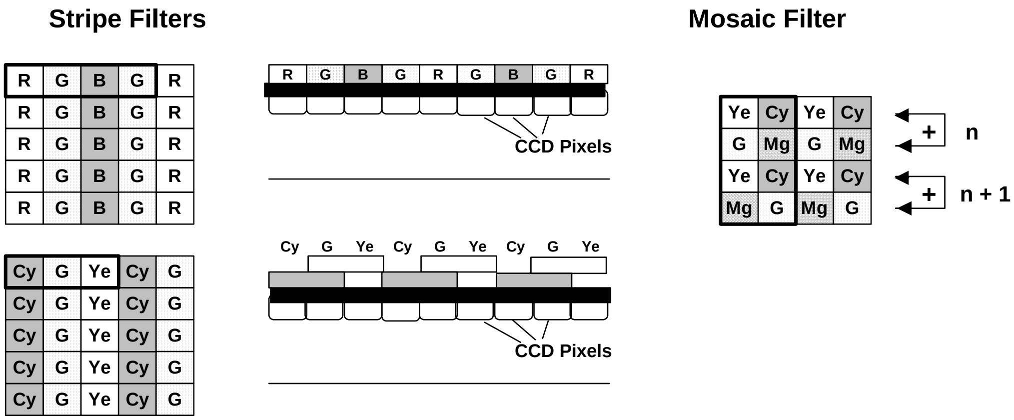


Figure 1.1.2-1 Examples of stripe and mosaic filters. For the mosaic filter shown, two vertically adjacent colors are always mixed together; which two depends on the interlace field (*n* or *n*+1).

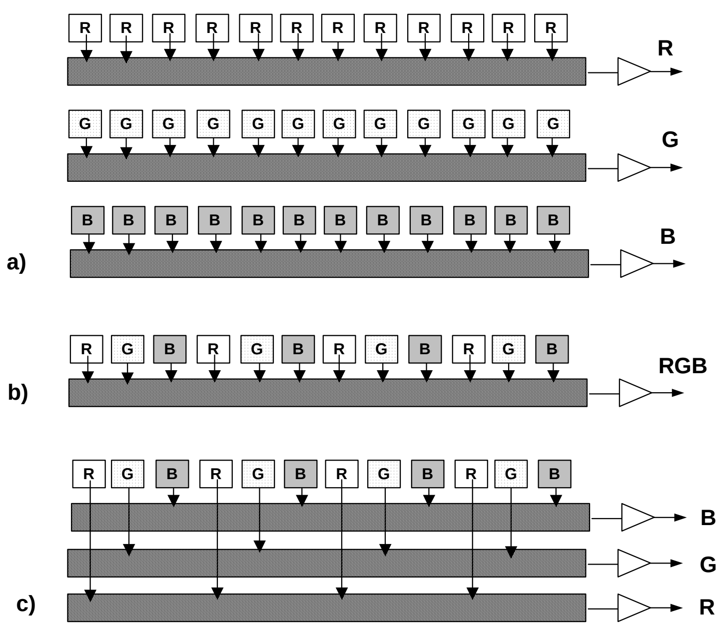


Figure 1.1.2-2 Example configurations for color imaging with linear arrays.

1.1.3 Intensified CCDs

A typical intensified CCD camera converts the incident light into electrons, which are then intensified and passed either directly (electron bombardment mode) or via a phosphor screen to the CCD. A schematic diagram of a typical device is shown below. If a phosphor is used, the light can be coupled to the CCD either via a lens or a fiber-optic plug (Fig. 1.1.3-1(a)).

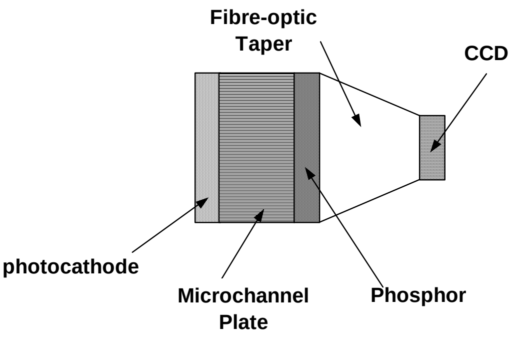


Figure 1.1.3-1(a) Schematic of an intensified CCD.

The electron gain is usually implemented with a microchannel plate or MCP. This is a bundle of glass fibers with pore size of a few microns. A high voltage (several hundred volts) is applied across the plate and electron multiplication takes place at the surface of the pores down which the electrons are channeled. Typically the intensification factor is ~1000, and good-quality imaging can be obtained for illuminance levels greater than 10^{-3} lux (i.e., starlight). An advantage is that the voltage to the intensifier can be gated so as to obtain short exposures. A disadvantage of all intensified cameras is that they can be permanently damaged by over-exposure to light.

1.1.3.1 Electron Multiplying CCDs

Electron multiplying (EM) CCDs have recently been developed to achieve the low light level performance of intensified CCDs, without the need for a separate intensifier.[13, 14] The EM CCD uses impact ionization to provide high gain before the signal is read out through the output amplifier. Usually the CCD is arranged to have a charge multiplication register, which is an extension to the normal readout register, as shown in Fig. 1.1.3-1(b). Two types of EM CCDs are commercially available at present: L3Vision™ devices from e2v Technologies and IMPACTRON™ devices from Texas Instruments. In both cases the amount of multiplication can be adjusted by varying the amplitude of the voltage pulses in the multiplication register.

Since the multiplication process results in increased noise (by a factor of $\sqrt{2}$ for high gains) the CCD will usually be operated in a *photon counting mode*, where the signal due to a single photon is boosted to a level much higher than the readout noise, and it is also arranged that no more than one photon is detected in each pixel during the integration time. Since the multiplication process will also amplify the dark current, it is usually necessary to cool the CCD to low temperatures.

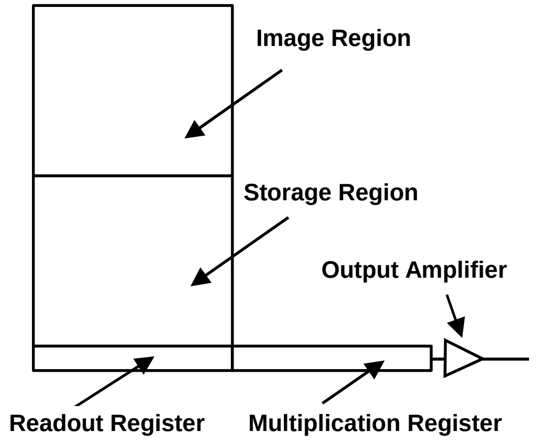


Figure 1.1.3-1(b) Schematic of an EM CCD.

Such devices are finding application in areas such as such as night-time surveillance or the detection of fluorescent and luminescent markers in life sciences. An advantage is that the EM CCD can be used both for low-light-level imaging and also at high light levels (which would "burn out" an intensifier). However, the EM CCD cannot replace the intensified CCD for applications that require fast gating of the exposure time.

1.1.4 Specialized CCD Readout Modes

In this section we discuss the various ways that detector arrays can be clocked in order to control performance. In general it will be necessary to calibrate the instrument/array (at least for response and dark current nonuniformity) for each clocking condition.

1.1.4.1 Windowing and Binning

With some CIDs and APS devices it is possible to either randomly access and read out any specific pixel or to only read out one (or perhaps more) selected rectangular regions of interest. With photodiode arrays (which are usually linear devices) and CCDs it is necessary to clock out every pixel. However, not every pixel has to be read out at the same rate. Specific regions of interest can be readout and digitized at a slow rate (to give low noise) while unwanted portions of the image can be readout (i.e., *dumped*) at high speed. This process, known as *windowing* (or sometimes, with linear arrays, *skipping*) is illustrated in Fig. 1.1.4-1. Some CCDs have a special "dump drain" that runs in parallel with the readout register. Operation of a "dump gate" transfers charge directly into this drain so that it does not have to be cleared out by horizontal readouts along the readout register. This gives a large increase in the rate that unwanted lines in the image can be dumped and is particularly useful for imaging spectrometers where only some (spectral) lines (or bands) are of interest (the direction of the spectrometer slit being aligned with the CCD line direction), or in spectrometers where several spectra are collected simultaneously (sometimes termed *multi-track* spectrometers), the slit being aligned perpendicular to the CCD line direction.

As well as windowing it is possible to achieve variable spatial resolution with a CCD by *binning* pixels together on-chip. Binning in the vertical (parallel) direction is achieved by transferring several lines into the readout register without readouts (shifts along the readout register) in between. Binning in the horizontal direction is achieved by clocking several pixels into the output node without reset pulses in between. Most modern CCDs are designed so that the readout register has three or four times the full well capacity of the image/storage region pixels so that several lines can be binned before charge starts to overflow (bloom) back into the storage region. Likewise, the output amplifier is usually designed to be linear over a range approximately 4 to 10 times the image area full-well capacity so that binned images can be accommodated. Large binning factors (e.g., greater than 3×3 or sometimes 4×4 pixels) will not, however, usually be obtainable with bright images (with signals close to image area full well). It should be noted that the binning process itself is noiseless and so gives an improvement in signal-to-noise ratio (SNR) at the expense of loss of spatial resolution. Vertical binning is often used in nonimaging spectrometers to improve SNR since the along-slit (vertical) direction does not contain useful information.

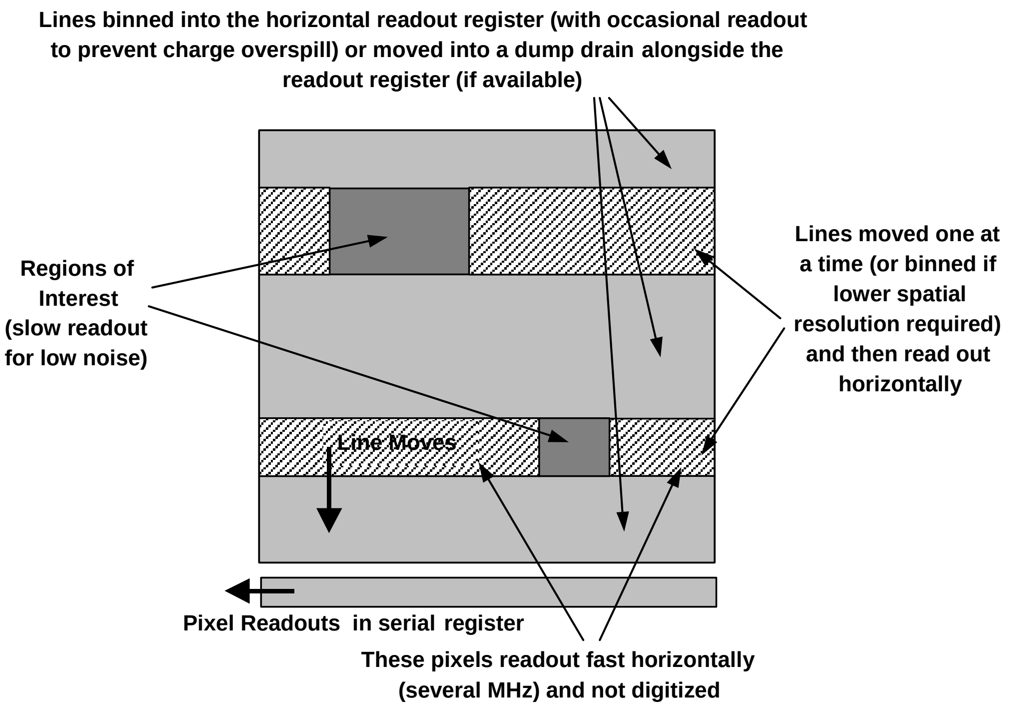


Figure 1.1.4-1 Schematic diagram of windowed readout with an area CCD.

For horizontal binning the usual sequence is to first reset the output amplifier, then to transfer in the pixels to be binned and finally to measure the output signal. However, since scientific CCDs have typical pixel readout times of 1 to 5 µs, the delay between the reset and readout pulses can be quite long so that the two samples are no longer correlated in time and low-frequency (correlated) noise is no longer cancelled out. Hence, the readout noise is increased.

Some low-noise CCDs have a *summing well,* which allows the temporary storage of charge from binned pixels before it is transferred into the output node. This ensures that the reset and read pulses can be close together so that noise performance is not degraded.

With windowed or binned readout in a CCD there should not be any change in fixed pattern noise, although the video offset voltage may change (due to the change in capacitative feedthrough of the readout clocks). A change in binning will (by definition) cause a change in dark current and response nonuniformity.

1.1.4.2 Kinetics Mode

Kinetics mode is a form of windowed operation with a CCD (c.f. Sec. 2.1.4.1) where most of the CCD (apart from the first few lines) are masked from light. The spectrum from the polychromator falls on the unmasked lines but is then rapidly shifted to the shielded area for storage prior to readout. By repeating the process, a time sequence of spectra can be stored on the CCD and then read out at slow speed.

1.1.4.3 Variable Integration Time

In many applications it is desirable to increase the dynamic range by varying the time for which the detector integrates charge from the signal. A simple way to change the integration time of any array is to change the master clock frequency (from which all the drive waveforms are derived). However, this has the considerable disadvantage that the readout speed (and hence the bandwidth, noise, and general optimization) of the off-chip electronics is changed. Instead, it is arranged that the readout speed is kept constant and the time for which the pixels integrate charge (when the clocks are idle)—known as the integration time (or the stare time)—is varied.

There are no constraints on the possible integration time of a FF CCD equipped with an external shutter other than the shutter open/close time (which will prevent uniform exposures of short duration) and the generation of thermal dark current (which limits the longest exposure that is possible). Shutter open/close times will depend on the size of the shutter but will typically be ~10 ms. However, with a FT CCD (or an ILT CCD) it is not often possible to have an exposure time less than the time needed to read out the image. There are two possible methods to produce a shorter integration time. Either the image region charge can be clocked backwards into a drain structure at the top of the array for part of the time that storage region charge is being read out (a procedure that tends to lead to clock breakthrough and a large fixed pattern noise), or the image region charge can be transferred into a drain structure (associated with each pixel) for part of the readout time. The latter process (which requires special CCD architectures) is termed *gated antiblooming, charge reset,* or *electronic shuttering*[8] and will not usually cause any side-effects to the imaging performance (i.e., offsets or nonuniformity).

With an X-Y addressed array the reset pulse to each line of pixels can usually be controlled, and this allows the integration time to be varied (in units of the line read time). Such changes may result in differences in the fixed pattern noise (dark signal nonuniformity), depending on the type of device used (because of changes in clock breakthrough).

An advantage of intensified arrays (Sec. 2.1.3) is that the voltage to the intensifier can be gated so as to achieve very short integration times.

1.1.4.4 Time Delay and Integration (TDI)

Time delay and integration is a way of using a 2D array to improve the sensitivity for 1D applications (such as scanning inspection systems, airborne reconnaissance, or spectroscopy). Usually there are a large number of pixels in the horizontal (line) direction and a relatively small number of lines (*n*). The scene is scanned in the column direction at the same rate (but in the opposite direction) as lines are transferred down the CCD (Fig. 1.1.4-2). In this way, the integration time is increased by a factor equal to *n* without any sacrifice in data rate. Smearing is avoided by careful synchronization between the movement of the scene and the shifting of charge within the CCD, and there is no limit to the scan length that can be used. Since several lines of data are being added, it must be ensured that the total signal is within the full well capacity of the array. The trade-offs involved in the design of TDI arrays are discussed in detail in Ref. 15.

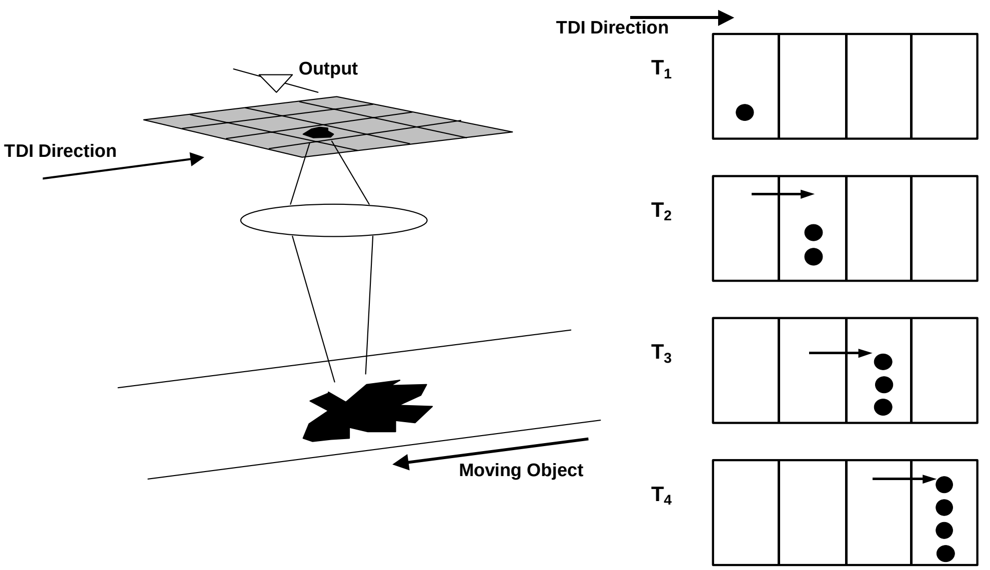


Figure 1.1.4-2 Illustration of the principle of TDI.

1.1.4.5 Imaging: Interlaced and Progressive Scan

In imaging applications, the most common scanning modes are interlaced and progressive scanning. Interlacing is a technique to increase the vertical resolution without increasing the signal bandwidth. The field frequency on a TV monitor is 50 Hz for the CCIR standard (Europe) and 60 Hz for the EIA standard (U.S. and the Far East). In interlaced scanning, alternate fields are shifted vertically by half a pixel. Each TV image (called a frame) is formed from the combination of the two (odd and even) fields, as shown in Fig. 1.1.4-3. With FT CCDs the interlacing is straightforward to achieve by changing the electrode phase under which charge is integrated. For (frame-) interline transfer devices the detector needs to have twice the number of photodiodes. In the most common architecture (called field-integration mode interlacing), the signals from the photodiodes are added in pairs within the vertical shift registers. Alternate fields have a pixel added to the one above or below. With X-Y addressed devices (including CIDs), interlacing is achieved in a similar manner—with the number of photosensitive elements being doubled.

Interlaced scan will produce distorted images of moving objects because of the staggered timing of the two fields (as shown in Fig. 1.1.4-4).

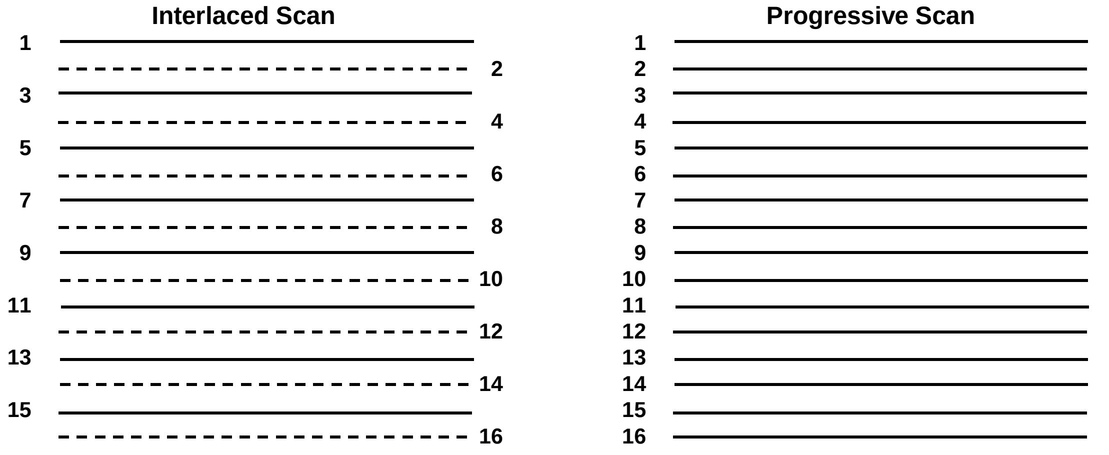


Figure 1.1.4-3 Illustration of interlaced and progressive modes of scanning.

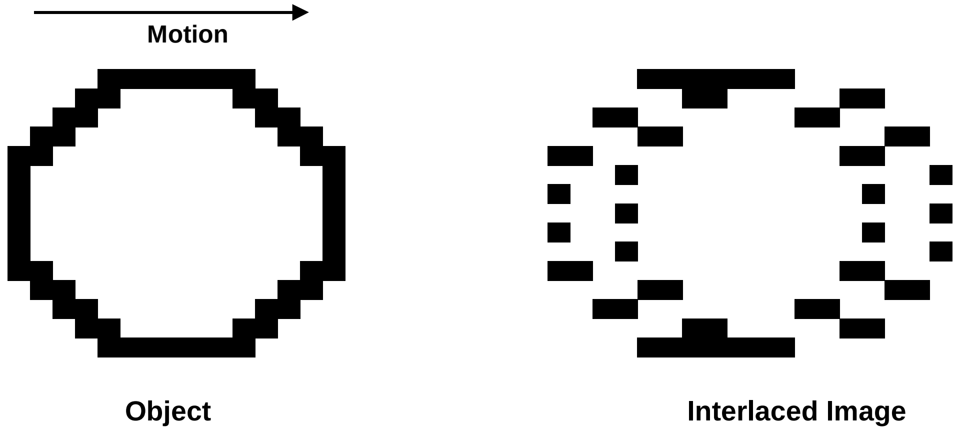


Figure 1.1.4-4 Appearance of a moving image with interlaced scan.

In machine vision, display on computer monitors and for high-speed applications, a different technique known as *progressive scan* is used. In this method, each field is identical (thereby removing flicker effects) but the number of lines in each field is doubled. In (frame-) interline transfer devices we already have double the number of detector elements, and only the vertical shift registers need to be modified (to transfer twice the number of charge packets). In FT CCDs, the number of lines needs to be doubled. Some cameras have the facility to switch between interlaced and progressive scan modes. “Progressive scan” basically means noninterlaced imaging, which will be the normal mode for scientific imaging and spectrometry applications.

1.1.4.6 Dark Charge Suppression

The theory of dark current generation in semiconductor devices is discussed in Ref. 16. In a silicon detector array thermal dark charge is generated predominantly within the depletion region or at a depleted surface. In these regions the free carrier concentration is low, and emission processes (i.e., generation) dominate over capture processes (recombination). The generation rate is greatest when there are states near the middle of the bandgap (for example, impurity atoms in the bulk of the silicon depletion region or interface traps at the silicon/silicon dioxide surface). In detector materials other than silicon (for example, those used in the IR) there are other generation mechanisms, for example, diffusion from regions outside the depletion region and tunneling processes.

The component that is generated by interface traps (surface dark current) tends to be large in silicon devices (~1 nA/cm^2 at room temperature; c.f. Sec. 1.2.5 for a discussion of dark charge units). However, surface dark current can be reduced by biasing the surface into inversion so that holes from the substrate or from p-doped regions (e.g. CCD channel stops) migrate to the Si/SiO_2 interface and fill the interface traps. This is achieved by decreasing the electrode voltages relative to the substrate voltage. In normal operation, one of the clock phases (the barrier phase) must be left high so as to form a potential well in which to collect signal charge, and thus the silicon under this phase cannot be inverted. However, there are two ways in which nearly the whole surface can be inverted (with as much as a factor 100 suppression of surface dark current):

1. Using a doped implant under one electrode phase to define a potential well even when all the clock voltages are reduced. CCDs with this type of implant are known as inverted-mode operation (IMO) or multi-phase pinned (MPP) devices. The latter name comes from the fact that when inverted, the surface potential is no longer controlled by the electrode voltage and is pinned to the potential of the substrate. Inverted-mode operation can also be achieved with some types of photodiode arrays. IMO devices tend to have a slower clock rate (for parallel transfers) and a lower full well capacity, though the latter can be improved in so-called advanced inverted mode or advanced MPP designs.
2. Using a clocking technique known as *dither clocking*, whereby the barrier phase that has to be kept high during integration (to define the potential well) is periodically swapped with one of the inverted phases. If the holes are trapped for longer than the time between switches of the barrier phase, then inversion can be achieved. The

theory of this process is discussed in Refs. 17 and 18. Dither clocking also gives a measure of antiblooming performance and when used specifically for this purpose is sometimes called *clocked antiblooming* (see below). Note that during readout the storage region is continually clocked and so automatically gives dither clocking. The efficiency of dither clocking in reducing surface dark charge improves as the temperature is decreased (since the hole trapping time increases). Drawbacks to dither clocking are an increase in power consumption and a tendency to “pump” trapping states and so produce black/white pairs of pixels.[19] This trap pumping effect can be reduced (sometimes eliminated) by careful choice of clock waveform levels.

1.1.4.7 Clocked Antiblooming

If a CCD is biased so that at saturation the charge comes into contact with interface traps at the surface before it spills over into adjacent pixels (a condition known as surface full well), then if the surface is kept populated with holes (by dither clocking, as explained previously) these will recombine with any excess photo-charge produced by over-exposure. This will therefore prevent charge from being (a) smeared due to trapping and emission by interface traps (giving poor CTE), and (b) overflowing into adjacent pixels. Note that charge can only be “soaked up” by holes at the rate that they can be supplied from the channel stops, so there is a linear relation between the over-exposure that can be tolerated and the frequency of dither clocking. Note that IMO (MPP) operation on its own does not give antiblooming properties (though dithered IMO mode gives good dark charge suppression and antiblooming performance).

1.1.4.8 CCD Erasure

For low-light-level applications with long exposure times, it is sometimes advantageous to clear a CCD of charge (e.g., accumulated in between exposures) by a fast clear-out process. This is easily achieved if the CCD has a special dump drain structure (in which all charge clocked into it will recombine). Another method is simply to clock the charge out in the normal way but at a higher rate and without digitization (a process known as *flushing*). Another technique, which works on some CCDs, is to operate the storage region clocks in inversion and to bias the readout register clocks at a fixed bias but not inverted, the idea being that the readout register then acts as a drain (in which charge recombination occurs).

1.1.5 Resetting of Output Amplifiers and Pixels

Figure 1.1.5-1 shows a typical CCD output circuit and Fig. 1.1.5-2 a typical output waveform. For every pixel, charge is clocked from the readout register through the output gate (OG) onto the capacitance of the output amplifier (where it is converted to a voltage signal). After every pixel the capacitance is discharged (reset) by a pulse (Rϕ) to the reset transistor. This produces a reset feedthrough pulse followed by a reset (or reference) level. Then, charge for that pixel is clocked into the output node (to give the video level). Usually the measured “signal” is the difference between the reset and video levels (taking this difference is known as correlated double sampling). Resetting the capacitance, C, introduces a noise proportional to kTC, where k

is Boltzmann's constant and *T* the absolute temperature. Not surprisingly, this noise source is called *reset* or *kTC noise*. A feature of the correlated double sampling process is that reset noise (which tends to dominate other noise sources at low signal levels) is cancelled. Note that in a CCD all the pixels integrate photo-generated charge (stare) at the same time.

Because of differences in the way reset occurs, X-Y switched arrays will not normally give an output waveform with reset and video levels for every pixel. The reset is performed either for the whole array or for a whole line or column prior to the integration. For example, Fig. 1.1.5-3 shows an example of an active pixel sensor with resetting one line at a time. In this architecture all the pixels stare for the same duration (the time needed to readout the array), but the start of an integration starts after each line reset, and so a line has its integration period delayed by one line read time from the pixels in the line before; hence, the pixels do not all stare at the same time. Imagers that can have pixels staring at the same time are sometimes known as *snapshot* imagers.

In an X-Y addressed array a reset level will normally appear on the output and is followed by a string of pixel video levels for each line (or for a whole frame if reset is once per frame).

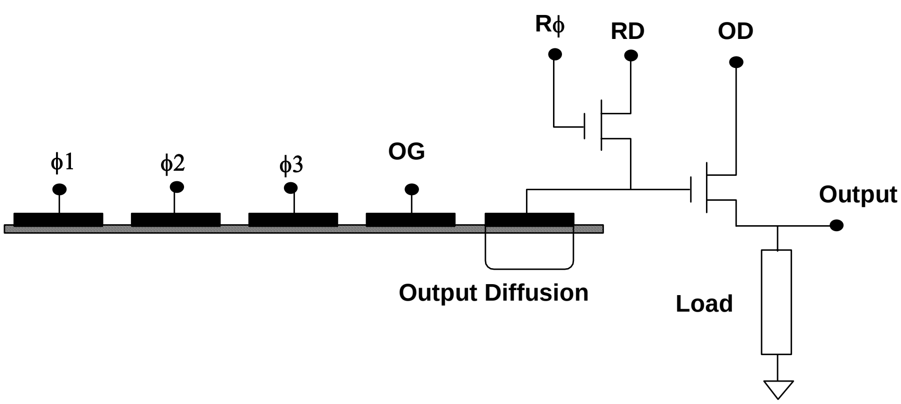


Figure 1.1.5-3 A typical floating diffusion CCD output stage.

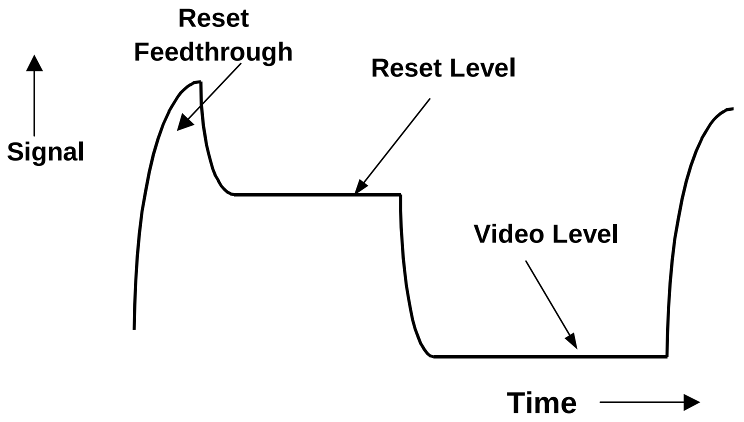


Figure 1.1.5-2 CCD output waveform. In correlated double sampling (CDS) the difference is taken between the reset and video levels.

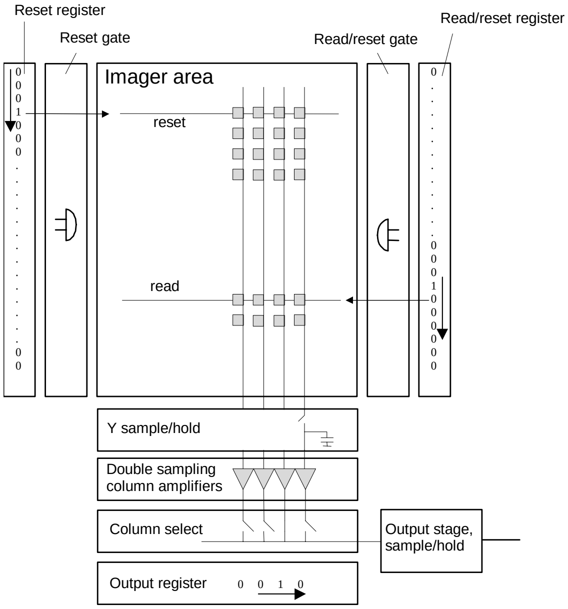


Figure 1.1.5-3 Architecture of a line addressed active pixel sensor.

Because the reset and video levels are far apart in time, the two measurements are no longer correlated in terms of low-frequency noise (though the kTC noise due to resetting the pixel will be correlated). Because of the need to hold the reset level for a long period of time (which requires additional sample/hold circuits) this form of double sampling is usually replaced in line reset arrays by measurement of a pixel video level followed by a measurement of the level after the next reset (the two levels occur close together in time). The kTC noise is not cancelled in this process, but low-frequency noise is reduced.

1.1.6 Readout of IR Arrays

It was mentioned in Sec. 1.1 that there are several semiconductor technologies that can be used for detecting infrared illumination. There are also several types of CMOS or MOS readout circuits that can be used in addition to the CCD (surface-channel rather than buried-channel CCDs are used in the IR because of the higher charge-handling capacity). CMOS circuits have been reviewed by Fossum and Pain[20] and also by Bluzer and Jensen,[21] Kozlowski et al.,[22] Kozlowski,[23] and Vampola (Chap. 5 of Ref. 6). The basic circuit types are

source follower per detector (SFD)
direct injection (DI)
buffered direct injection (BDI)
gate modulation input (GMI)
cascode amplifier per detector (CAD)
capacitative transimpedance amplifier (CTIA).

These various readout circuits give differing performance in terms of readout noise, power consumption, charge handling, bandwidth, fixed pattern noise, and pixel area. These parameters can therefore be traded off for a particular application. Their basic properties are summarized below.

Source follower per detector (SFD)

- simple circuit, so can be used in high-density arrays
- low power consumption
- high 1/*f* noise
- high fixed pattern noise (due to threshold voltage nonuniformities)
- disadvantage of no control of diode bias
- used in U.S. for a wide range of MCT, Si:As blocked impurity band and InSb arrays.

Direct injection (DI)

- simple circuit, so can be used in high-density arrays
- low power consumption
- injection efficiency low for low-resistance LWIR diodes and for low backgrounds
- large integration capacitance implies low voltage sensitivity and increased noise due to downstream circuits
- commonly used for high background applications.

Buffered direct injection (BDI)

- amplifier inserted in the DI circuit to improve injection efficiency, so allowing low background operation—at the expense of a more complex circuit
- good control of detector bias
- power consumption increased due to amplifier
- used mainly in linear arrays where additional circuit size can be easily accommodated; some 2D arrays have also used BDI.

Gate modulation input (GMI)

- bias voltage developed across a load resistor used to modulate the gate voltage of an output transistor
- large charge-to-voltage conversion gain leads to high charge detection sensitivity and reduced input-referred noise levels
- hence, can have a large integration capacitance (high charge capacity) and still achieve good noise performance
- current gain self-adjusts, depending on background flux—can be used for background suppression

- disadvantage in that pixel-to-pixel current gain variations cause large fixed pattern noise, also variations in linearity.

Capacitive transimpedance amplifier (CTIA)

- each pixel has an inverting amplifier with a capacitor in the feedback loop
- complex circuit, so difficult to accommodate with high density arrays but possible with state-of-the art CMOS processes.
- good control of detector bias (input to amplifier is a virtual ground)
- can have small integration capacitance and very good noise performance
- high power consumption.

Cascode amplifier per detector (CAD)

- like the GMI circuit, charge-to-voltage conversion gain is increased; otherwise properties similar to the SFD circuit
- complex circuit, so difficult to accommodate with high-density arrays, but possible with state-of-the art CMOS processes.

1.2. Key Parameters and Their Measurement

In this section we consider the definition and measurement of the key parameters for detector arrays, with special emphasis on silicon devices (CCDs, CIDs, and PDAs). These are discussed at the device level, and the measurements will be of interest for array manufacturers and equipment suppliers. This section should also be of interest to users since the performance of the detector array often dominates the overall performance of a complete instrument and largely determines its specification. Also some measurement techniques (e.g., linearity, response nonuniformity, and MTF) can be directly applicable to instrument-level calibrations; users should also refer to Chapter 5 for further information on these topics.

The definition and measurement of array parameters can be a confusing subject because of the lack of standardization. A good discussion of the issues is given in Ref. 24. On the one hand, detector arrays can be viewed as an electronic component (just as a microprocessor or a memory chip), and performance can be specified in the manufacturer's data sheet. However, detector arrays have features that make them somewhat unique:

- they are analog devices with a high dynamic range and a large number of detector elements, each with its own response
- they are time consuming and expensive to test
- there are a large number of performance parameters
- performance depends on operating conditions (temperature, clocking, illumination).

An *array manufacturer/vendor* will test devices under a particular set of conditions (for example, at room temperature and at a specific readout rate). Also, he will use his own test methods developed in-house, which will be dedicated to specific test equipment. This equipment is usually expensive to set up and validate, and so cannot be easily changed.

A *user* will usually want to operate the detector array (in his equipment) at a temperature and readout rate different to that used by the manufacturer, he will have specific illumination conditions, and may be particularly concerned with some parameters and not with others. For example, a spectroscopist measuring transient emission phenomena may be interested in fast readout (at the expense of higher noise) and need a high full well capacity so that emission lines do not saturate the detector. In contrast, a Raman spectroscopist may have very low signals and need a low-noise CCD with low dark current/low operating temperature (so that long integration times can be used). The number and type of defects (bright pixels, dark pixels or pixels with traps) that can be tolerated will also vary with application. In some cases defects can be removed by image processing software, but in others (particularly in real-time systems) such defects may trigger false events or identifications.

Bridging the gap between the two is the *system integrator/equipment supplier*, who must take a detector with the manufacturer's specification and yet guarantee to the user that the complete instrument will meet its overall specification. This job may be made easier by knowledge of the performance of previous systems and if the array manufacturer can tailor his testing to the application, or is willing to extrapolate the detector performance to the end-user conditions. Otherwise the system integrator may need to perform his own selection testing, which may be a costly and/or high-risk undertaking. Whichever approach is adopted, all sides are encouraged to gain a good understanding of the way in which performance parameters change with conditions, and it is hoped that this will be helped by the information given in this guide. In view of the wide range in applications (from consumer imaging cameras, through machine vision, to high-performance scientific/spectroscopy systems), it is unlikely that there will be a significant standardization in manufacturers' specifications in the near future.

The following discussion is based on the definitions and test methods described in a specification[25] adopted by the European Space Agency (Space Components Co-ordination group) following advice from an ad hoc working group of European CCD manufacturers and users within the space community; however, the definitions have been adapted and extended to make them more widely applicable.

The definitions given in the following sections should be interpreted in relation to the typical CCD output waveforms shown in Sec. 1.1.17. The extension to other sensor types (such as photodiode arrays) should be straightforward in most cases. A CCD will normally have prescan and postscan pixels in the readout register. These do not contain any charge (apart from a negligible amount of readout register dark current) and are useful for determination of a signal baseline (though clock feedthrough can produce an offset in this baseline). Also, the CCD will sometimes have image-area pixels that are shielded from light (usually along an edge) so that they act as a reference for measuring average dark signal.

1.2.1 Units

At the array level the manufacturer/developer will be most interested in measuring signals in electrons (one electron is created for each detected photon for wavelengths longer than about 300 nm). However, the user will be more concerned in the signal produced for a given amount of light. A particular source of confusion is that specifications for general imaging applications are often given in photometric units (e.g., lux), whereas scientific imaging and spectrometry applications use radiometric units (e.g., W/m^2 or photons/pixel). Section 1.3 discusses the relationship between the two types of units.

Another complication is that output signals are sometimes quoted in volts (since the output is a voltage waveform) and sometimes in the number of electrons (n) generated in each pixel during the integration time. Since the charge generated (q) is the number of electrons multiplied by the electronic charge, e, (= 1.602 10^{-19} coulombs), the two measurements are related by the capacitance (C) of the output node ($q = ne = C$ V). Unfortunately, manufacturers do not always quote the value of C (or, equivalently, the charge-to-voltage conversion factor CVF = e/C V/electron), and so conversion between the two sets of units cannot always be achieved. Typically, however, the CVF is in the range of 0.5–10 μV/electron. The signal in electrons/pixel/integration time is sometimes multiplied by the number of pixels and converted to a current for the whole device (usually in nA) —this is the current that (on average) needs to be supplied to the reset drain of the device to balance the total photo-generated current. Conversion from nA to the average signal in electrons/pixel/integration time requires a knowledge of the number of pixels illuminated and the integration time (recall that the definition of current, in amperes, is the charge flowing, in coulombs, per second). See also Sec. 1.4 for a discussion on signal sizing.

1.2.2 Electrical Test Methods

Reference 25 contains definitions and measurement techniques for the following CCD parameters (and similar definitions will apply to other types of detector arrays):

leakage current on input gates
insulation leakage current between pins
power supply current
DC output level
amplitude of reset feedthrough
output impedance
electrode capacitance
output signal waveform features (such as settling time).

These parameters are mainly of interest for instrument designers and do not usually affect instrument stability and so will not be discussed in this guide; the reader is referred to Ref. 25 for more information.

1.2.3 Measurements on Dark Images

Measurements with the detector array in darkness are important in defining the baseline that needs to be subtracted from the signal obtained during illumination so that the photo-generated component (the signal of interest to the user) can be obtained. The dark signal can contain several components, some that may be the same for each pixel and some that may not:

Thermal dark charge. This is produced by generation of electron-hole pairs within the depletion region (bulk generation) or at a depleted surface (surface generation). The generation rate increases rapidly with temperature (c.f. Sec. 3.1.7). The thermal dark charge is
- proportional to integration time
- varies strongly with temperature
- has an average value and a nonuniformity (pixel-to-pixel variation).

Offset level. Produced by electrical feedthrough of clock signals (especially in a CCD)
- independent of integration time
- uniform across the array
- can vary with temperature (but not as strongly as thermal dark charge).

Fixed pattern noise. This is produced by electronic effects such as crosstalk with the array. It is usually negligible for CCDs (but can be important for X-Y addressed arrays).
- usually independent of integration time
- is nonuniform (pixel-to-pixel variations)
- usually independent of temperature.

Spurious charge. Though rare, in some CCDs operated in inverted mode, charge can be generated by impact ionization by the electric fields induced by the clock voltage swing.[26, 27] This spurious dark signal is
- independent of integration time
- varies slowly with temperature (not in the same way as thermal dark charge, the variation being due to the temperature dependence of the impact ionization coefficient).
- may be somewhat nonuniform.

Amplifier glow. In some CCDs operated at low light levels, the biases applied to the output amplifier can cause light emission. The effect is only present in some CCDs (and is becoming rarer as designs improve). Where it does occur it can be reduced by applying the amplifier bias only during readout (not during long integrations). It is essentially
- independent of temperature
- proportional to integration time
- nonuniform (apparent only for pixels close to the output amplifier).

Spurious charge and amplifier flow occur rarely and can be considered as image anomalies. The other parameters are discussed further below.

1.2.4 Offset Level

Definition:
Offset voltage is the difference between the reset and signal levels for pre- or postscan elements of the output waveform; i.e. those that do not contain signal charge or thermal dark current. It arises from the clock feedthrough of readout register clocks onto the output waveform (c.f. Sec. 3.1.2), but will also contain dark charge (usually negligible) generated in the readout register.

This parameter can be important for system-level performance, but is difficult to specify for an individual CCD since it is influenced by the shape of the readout clock waveform (and the temperature). Hence, it is not normally quoted by detector manufacturers, but has to be measured at system level.

In some (rare) cases, temporal variations in clock waveforms can cause the offset level to vary across a line.

Measurement:
If the pre- or postscan pixels are digitized for the system, then the offset level can be measured directly (either as a digital number or a voltage) averaged over several pixel values. If they are not digitized, then only the dark level (which is the offset plus the dark signal) can be measured. This may not be important for the user since it is the dark level that is ultimately of interest; but it is useful for predictive purposes to separate the two components.

1.2.5 Average Dark Signal

Definition:
Dark signal is the output signal (offset subtracted) in the absence of any illumination, under specified conditions (temperature, integration time bias).

For a FT CCD the signal in a dark image varies with line number because of the progressively more time that charge spends in the storage regions before readout. The average dark signal in the first few lines read out (excluding pixels subject to edge effects) gives the average dark signal from the image region. The average dark signal in the last few lines read out (again excluding pixels subject to edge effects) is the sum of the average dark signals in image and storage regions. (Fig. 1.2.5-1).

The units of dark signal are usually electrons/pixel/s or nA/cm^2. Strictly speaking, the latter is a unit of dark current density, calculated as follows:

$$\underset{(\mathrm{nA/cm^2})}{\text{dark current density}} = \underset{(\mathrm{electrons/pixels/s})}{\text{dark signal}} \times e \times \frac{10^9}{A} \tag{1.2.5-1}$$

where e is the electronic charge (1.602×10^{-19}C) and A is the pixel area (cm^2).

Measurement:
Straightforward derivation from analysis of dark images. Related parameters are the temperature coefficients of the image and storage region dark signals and the inversion voltage (at which the surface becomes inverted and dark signal falls steeply).

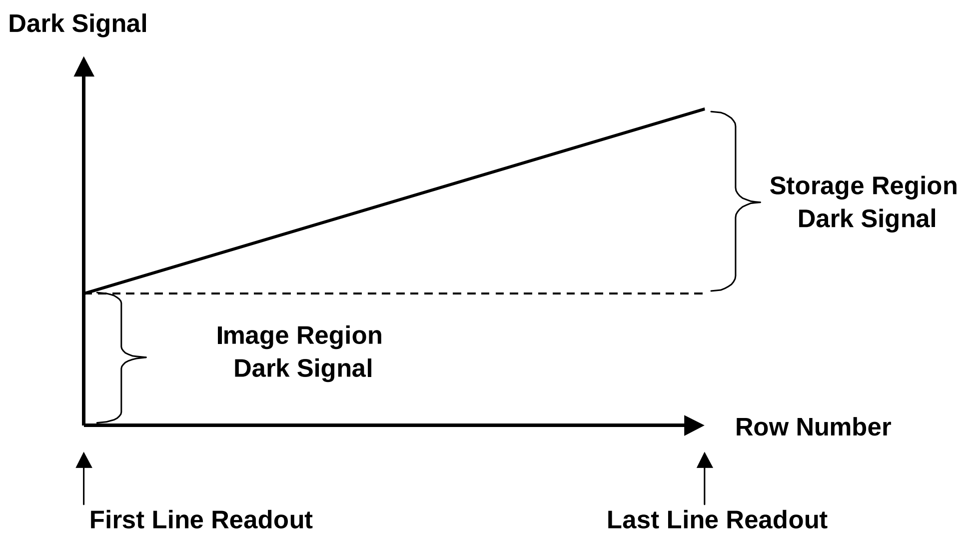


Figure 1.2.5-1 Measurement of image and storage region average dark charge for a FT CCD.

1.2.6 Dark Signal Nonuniformity (DSNU)

Definition:
The dark charge signal will differ from one pixel to another. The DSNU is the nonuniformity in the dark image, after allowing for any slope in the average dark charge (as arises, for example, in FT CCDs as discussed previously). In this guide we define the DSNU as the dark signal nonuniformity that is not fixed, but varies with temperature and integration time. It arises principally from thermal dark current, though for "defect" pixels (white spots and columns) there will be a component that comes from spurious charge injection. Nonuniformity that is fixed we term *fixed pattern noise* (see Sec. 1.2.7).

There are several ways to specify DSNU. If we define for pixel i

$$\mathrm{DSNU}_i = (Vs_i - V_a)/V_a\,,$$

where V_a = average signal in darkness, and Vs_i = pixel signal in darkness then any of the following can be specified (depending on the device and the application):

- peak-to-peak DSNU
- maximum and minimum, max (DSNU_i) and min (DSNU_i)
- standard deviation of the DSNU
- defect pixels (spikes and dips) beyond a level $[-a; +a]$
- for linear devices, a uniformity plot can be produced

- for area devices, a map of DSNU defects and/or a histogram of pixel values or of cumulative dark signal (percent of pixels having a dark signal higher than a given value).

It should be noted that in devices with low average dark current (operated in inverted or MPP mode), the distribution of dark pixel values is non-Gaussian (there are dark current spikes due to “hot” pixels that have an anomalously high value). In such cases the standard deviation of the DSNU is not a good measure.

Measurement:
By analysis of dark images. Averaging of several frames may be needed in order to remove effects of temporal noise and/or contamination due to pick-up.

1.2.7 Fixed Pattern Noise (FPN)

Definition:
The term “fixed pattern noise” suggests a random nature, but this need not be the case. It is a type of DSNU; the term “noise” refers to spatial noise. It is defined in various ways in the literature and in manufacturers’ data sheets; however, the most common definition (and the one recommended in this guide) is that it is the dark signal nonuniformity arising from electronic sources (i.e., other than thermal generation of dark current). This fixed pattern noise can arise either from clock breakthrough (which is synchronized to the readout pattern and therefore fixed) or from offset variations in row, column or pixel amplifiers/switches (in X-Y addressed arrays such as active pixel sensors, APSs). In a CCD the electronic fixed pattern noise is usually negligible, and there is only the DSNU due to thermal dark current and a uniform offset level (we exclude pick-up since it is not synchronized, i.e., fixed). In X-Y addressed arrays the fixed pattern noise can dominate over the DSNU due to thermal dark current.

In common usage, the difference between DSNU and fixed pattern noise is somewhat blurred. In CCDs the electronic fixed pattern is small and the term fixed pattern noise is rarely used. In X-Y addressed arrays the fixed pattern noise is sometimes defined in terms of the total dark signal nonuniformity (including both electronic and thermal components), but, as mentioned previously, it is to be encouraged that the term refers only to the electronic component—so that the thermal DSNU has to be separately specified. In practice it may be difficult to separate the two components. To do this requires measurement of the nonuniformity at several temperatures or integration times: the electronic component will tend not to change (at least over a small range). An example of the constant nature of electronic fixed pattern noise is given in Fig. 1.2.7-1. As with DSNU, fixed pattern noise can be specified in a variety of ways (peak-to-peak, max/min, rms, histograms, etc.).

In some literature the fixed pattern noise is defined as including both nonuniformity in pixel offsets (electronic and thermal) and responsivity variations (defined in this guide as photo-response nonuniformity or PRNU, c.f. Sec. 1.2.8). For individual uncalibrated images this may have some justification, since both give the appearance of spatial noise; however, the two quantities (DSNU/fixed pattern noise and PRNU) have quite different effects (PRNU depends on signal level) and are derived from different calibration measurements. Derivation of offset nonuniformity comes from dark images and PRNU from uniformly illuminated images (flat fields). Hence, the practice of using fixed pattern noise to include both is to be discouraged.

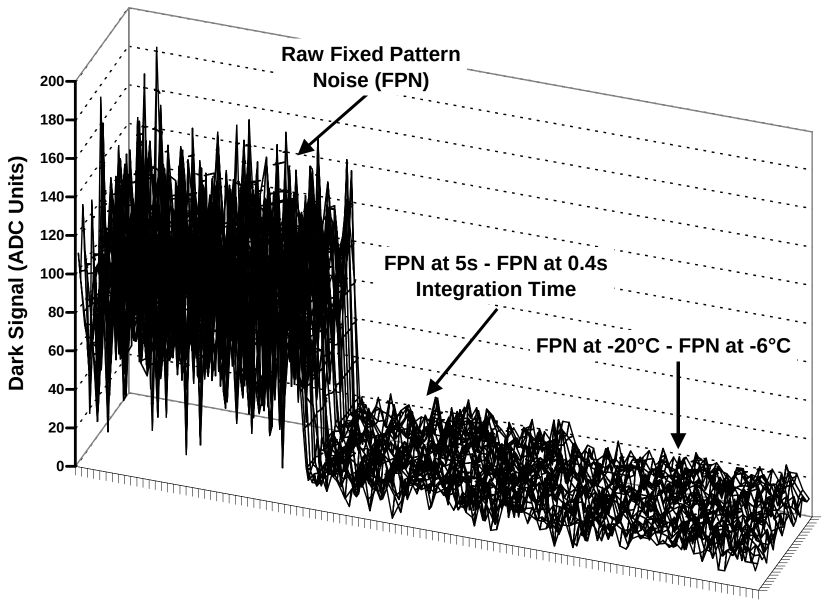


Figure 1.2.7-1 Fixed pattern noise for a CMOS active pixel sensor (left side of the plot). For the temperatures and integration times used, the dark charge was negligible; though at higher temperatures and integration times the thermal dark charge nonuniformity changes as expected. At the right side of the plot are the differences between images obtained for two integration times and two temperatures. Although the fixed pattern noise in a single image is large, the differences are much smaller (the residual nonuniformity is due to thermal dark charge, temporal noise, and small changes in electronic fixed pattern noise). The conclusion is that, in this case, the electronic fixed pattern noise is essentially independent of integration time and temperature.

Measurement:
The same as for DSNU, but analysis of images for the component of nonuniformity that does not change with temperature or integration time is also needed.

1.2.8 Photo-Response Nonuniformity (PRNU)

Definition:
For a uniform illumination, the photo-response nonuniformity (PRNU) is the difference in percentage between the signal in each pixel and the average signal of the total photosensitive area (excluding the extreme edge pixels in some cases). As discussed for DSNU in Sec. 1.2.6, PRNU can be specified in various ways:

- peak-to-peak PRNU
- maximum and minimum, max ($PRNU_i$) and min ($PRNU_i$)

- standard deviation of the PRNU
- defect pixels (spikes and dips) beyond *a* level [$-a$; $+a$]
- for linear devices, a uniformity plot can be produced
- for area devices. a map of PRNU defects and/or a histogram of pixel values or of cumulative dark signal (percent of pixels having a dark signal greater than a given value).

PRNU is usually obtained from the whole chip area, but it can also be obtained from calculation on small groups of pixels (with the PRNU being taken as the maximum of the local PRNU values for all the pixel groups).

Measurement:
By analysis of uniform bright field images (*flat fields*) after subtraction of dark signal. Either the dark image has to be for the same integration time or it has to be verified that the dark signal is dominated by either time-independent or time-dependent (proportional) components. In the latter case, the dark signal can be scaled to the same integration time as the bright field exposure. Note that if traps are present (c.f. Sec. 3.1.4), then measurement may be needed at several illumination levels. Also, the measured PRNU will depend on the wavelength and may depend on the cone angle of the light beam used to form the uniform illumination (Sec. 3.1.4).

Uniform illumination can be achieved in various ways:

- point source (e.g., pinhole) at a large distance
- diffuser (but the beam needs to be projected from a distance or the cone angle on the array will be large)
- integrating sphere
- uniform translation of the illumination beam (from a pinhole, diffuser or integrating sphere) over the plane of the array (or in-plane translation of the array itself). This can be useful for improving the uniformity of illumination for large arrays. Performing deliberate (known) shifts of the illumination or the array by ~1 pixel has been suggested[28, 29] as a means to improve flat-fielding in situations (such as astronomy) where the uniformity of the flat field cannot be guaranteed. However, the mathematical techniques needed to process the images are complex and not yet fully developed (at least for the 2D case).
- imaging of a near-uniform target (reflection or transmission).

The last method (imaging of a target) is convenient for camera systems but will lead to errors (even if defocused) due to nonuniformities in the target. Healey and Kondepudy[30] have suggested an improved technique involving taking images with n_1 configurations (obtained by varying the target and source positions), $n_1 = 20$ in their experiments. For each image, the light level falling on the array was found by averaging pixels in an $m \times m$ pixel window centered on a given pixel, where m is large enough that responsivity variations are averaged out, but small enough that variation in illumination and surface reflectance (or transmission for a transmission target) are small over the window (the window was 9×9 pixels in their experiments). The reader is referred to their paper for more details. Note that these authors use the term fixed pattern noise instead of the recommended term PRNU for responsivity variations.

1.2.8.1 Spectral Photo-Response Nonuniformity

A quantity related to the PRNU is the spectral photo-response nonuniformity (SPRNU). This is related to the spectral variation in the responsivity and can be measured by taking the ratios of the PRNUs at selected wavelengths or narrow wavebands to the PRNU for a broad waveband (encompassing all the selected wavelengths or wavebands). The ratios can be weighted according to the spectral shape of the broadband illumination.

1.2.9 Charge-to-Voltage Conversion Factor (CVF) and Conversion Gain

Definition:
The charge-to-voltage conversion factor (CVF) is the ratio between the output voltage for a particular pixel (or the average voltage for a group of pixels) and the number of electrons transferred to the output stage for that pixel (or the average for a group of pixels). Usually CVF is expressed in μV/electron, and this is an important parameter at chip level. At system level the output voltage will usually be amplified and digitized. The measured signal in electrons is given by

$$\underset{(\text{electrons})}{\text{signal}} = \underset{(\text{ADU})}{\text{signal}} \times \underset{(\mu\text{V/ADU})}{\text{system gain}} \;/ \underset{(\mu\text{V/electrons})}{\text{CVF}} . \tag{1.2.9-1}$$

The system gain is purely related to the gain of the off-chip electronics and the characteristics of the ADC, and hence can be either calculated knowing the component values or measured electrically. However, it may be that neither is a practical possibility. The parameter system gain / CVF (in electrons/ADU) can often be measured as an alternative. Indeed, for the user this is the more useful parameter. This parameter is often termed the *conversion gain*.

In this guide we make no fundamental distinction between the two parameters: CVF in μV/electron or conversion gain in ADU/electron (or the reciprocal quantity in electrons/ADU). For the user, conversion gain is the parameter of interest for the instrument. For the manufacturer/developer, the system gain (in μV/ADU) will usually be known so that measurements in either μV/electron or electrons/ADU can be made and readily interrelated.

Measurement:
Several measurement methods are possible, naturally giving results either in μV/electron or electrons/ADU.

Reset Drain Current
The device is uniformly illuminated and the average current in the bias supply to the reset drain (I_{RD}) is measured. Since this current will fluctuate during the readout cycle, it is recommended to operate in a continuously clocked (rather than frame transfer) mode if possible.

$$CVF\ (\mu V/e) = \frac{V_a}{I_{RD} t_{\mathrm{int}}} eN \times 10^{15}, \tag{1.2.9-2}$$

where V_a is the average output voltage from the CCD (in μV), I_{RD} is measured in nA, t_{int} is the integration time (ms), e is the electron charge (1.602×10^{-19}C) and N is the number of useful pixels (i.e., those which generate a signal on illumination).

For a digitized signal, V_a is replaced by the average signal in ADU, and we have the conversion gain measured in (ADU/electron) and its reciprocal in electrons/ADU.

Since I_{RD} is small and can be influenced by leakage currents, it is recommended to use a differential method:

$$CVF = \frac{\Delta V_a}{\Delta I_{RD}} eN \times 10^{15}. \tag{1.2.9-3}$$

To avoid risk of electrical overstress when connecting a picoammeter to the reset drain line, I_{RD} can be monitored by measuring the voltage across a suitable resistor.

X-ray Method

X-ray illumination of a silicon detector will cause near-spherical charge clouds of ~1 μm diameter to be generated via the photoelectric effect. This charge will then be transported by drift or diffusion and be collected in the depletion region of a pixel (or group of pixels).[31] The total charge produced in each x-ray absorption event is proportional to the x-ray energy, with an average of one electron-hole pair being produced for every 3.6 eV of energy at 25°C. Thus, the event size (in ADU) can be directly related to the number of electrons in the charge packet.

A source that is often used for x-ray calibration is the radioactive isotope Fe^{55}. This gives characteristic x-rays of energy 5.9 keV, giving events of ~1600 electrons. However, these x-rays do not have enough energy to pass through the glass window of the array, and devices have to be delidded. A more convenient (though surprisingly little used) isotope is Cd^{109}. This decays to ^{47}Ag, having characteristic x-ray emissions at 22 and 25 keV that will readily pass though the detector window (and also through a few millimeters of external optics). To be more precise, there are several x-ray wavelengths produced:[32]

Description	Energy (keV)
$K\alpha_2$	21.9903
$K\alpha_1$	22.1629
Average	**22.0766**
$K\beta_3$	24.9115
$K\beta_1$	24.9424
$K\beta_2$	25.4564
$K\beta_5$	25.145
$K\beta_4$	25.512
Average	**25.1935**

Hence, at 25°C the signals from the Kα and Kβ lines are 6130 ±50 electrons and 7000 ±80 electrons, respectively, where we have made an estimate of the errors due to calculating the average x-ray energy and converting to the average number of electron-hole pairs.

The quantum efficiency for silicon arrays (the proportion of x-rays absorbed within the active region of the detector) will be ~1%, but the event rate does not need to be high (<< 1/pixel/integration time). A drawback of using Cd^{109} is the short half-life (450 days); however, a 3mCi (111 MBq) source will usually give ample count rates for at least five years and can be readily purchased[33] (though, as with any radioactive source, the appropriate licensing regulations need to be observed).

A complication of the x-ray method is that the charge is only confined to single pixels (or split between adjacent pixels) if it is generated within the depletion region. For events generated in the field-free region below the depletion layer, charge is collected by diffusion (Fig. 3.1.1-1), a process that leads to spreading over several pixels (these events are usually termed "partial").

A typical histogram of pixel values from a Cd^{109} CCD room-temperature illumination is shown in Fig. 1.2.9-1. Most pixels receive no x-ray events and give a "zero" peak corresponding to the dark signal nonuniformity. There is a flat region due to partial and split events, and two peaks corresponding to Kα and Kβ x-rays absorbed in the depletion region (the amplitude of these peaks will depend on the size of the depletion relative to the total volume). In this case the peaks appear at 84.3 and 95.3 ADU above the center of the zero peak, giving a calibration of 73.1 electrons/ADU (see figure inset).

For accurate work the temperature dependence of the average energy per electron-hole pair has to be taken into account. This has been discussed by Fraser et al.,[34] who give a variation of roughly 0.01%/K, hence the value changes from 3.6 eV at room temperature to 3.65 eV at 170 K.

Note that x-ray illumination is a good way to check the overall imaging performance of a system and can be used to measure CTE (c.f. Sec. 1.2.13).

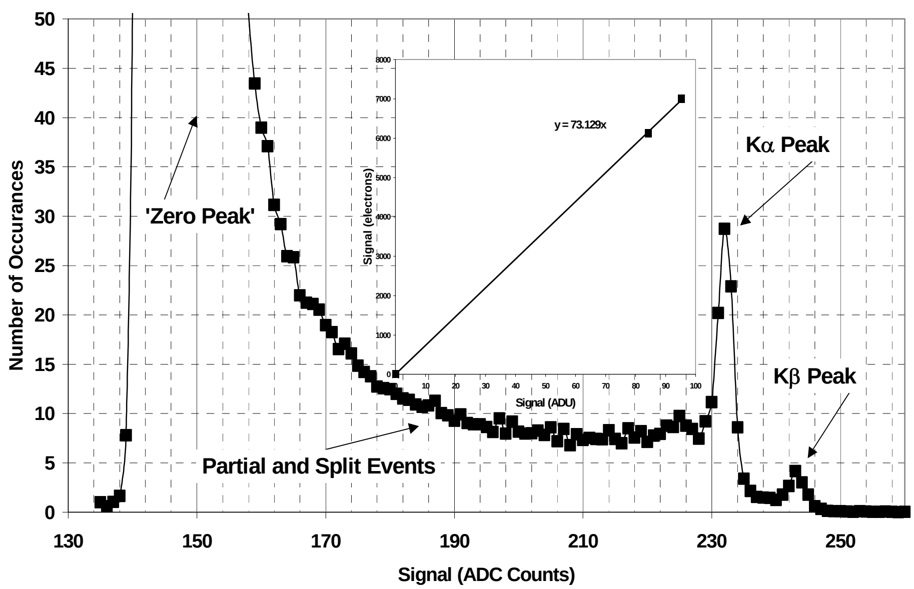


Figure 1.2.9-1 Histogram of pixel values from a Cd^{109} illumination of a CCD at room temperature.

Mean-Variance Method

This method (Ref. 26, see also Ref. 35) relies on the fact that the noise in an image increases as the signal increases because of the presence of shot noise. If the signal in a pixel is measured as S (ADU) and the conversion gain is K (ADU/electron), then S is related to the number of signal electrons (N) by

$$S = K \times N + \text{readout noise.} \tag{1.2.9-4}$$

Since the readout noise will have a mean of zero, the mean signal, $\langle S \rangle$, is given by

$$\langle S \rangle = K \times \langle N \rangle \tag{1.2.9-5}$$

and the noise variance by

$$\text{noise variance} = K^2 \sigma_N^2 + \sigma_R^2, \tag{1.2.9-6}$$

where σ_R is the rms readout noise and σ_N^2 is the shot noise variance (in electrons), which for Poisson statistics, is equal to $\langle N \rangle$:

$$\sigma_\mathrm{N}^2 = \langle N \rangle = \langle S \rangle / K, \tag{1.2.9-7}$$

from (1.2.9-2); hence,

$$\text{noise variance} = K \langle S \rangle + \sigma_\mathrm{R}^2, \tag{1.2.9-8}$$

and a plot of noise variance against average signal is a straight line of slope K and intercept equal to the readout noise variance.

If over-scanned pixels can be digitized so as to determine the electronic offset (the output, in ADU, for pixels containing no signal or dark charge), then the plot can be obtained by taking pairs of flat-field images for the same illumination conditions *with the offset subtracted* (FF1 and FF2), and obtaining the mean and difference images ([FF1 + FF2] and [FF1-FF2]):

$$\begin{aligned} \text{mean signal} &= \langle S \rangle = \langle [\mathrm{FF1} + \mathrm{FF2}] \rangle / 2 \\ \text{shot noise variance} &= (\text{variance}[\mathrm{FF1}-\mathrm{FF2}]) / 2. \end{aligned} \tag{1.2.9-9}$$

K is found from the slope of this plot (or the intercept, if a log-log plot is used).

Alternatively, K can be found using pairs of flat-field (FF) and dark (DK) images *of the same duration*:

$$K = \frac{\left(\left[\langle \mathrm{FF1} \rangle + \langle \mathrm{FF2} \rangle\right] - \left[\langle \mathrm{DK1} \rangle + \langle \mathrm{DK2} \rangle\right]\right)/2}{\left[\text{variance}\left(\mathrm{FF1} - \mathrm{FF2}\right) - \text{variance}\left(\mathrm{DK1} - \mathrm{DK2}\right)\right]/2} \tag{1.2.9-10}$$

and values of K for different pairs can be averaged.

Figure 1.2.9-2 shows data obtained for a CCD in an imaging spectrometer that was obtained in a slightly different way, with near-uniform illumination of the CCD. Since signals could be binned in the vertical direction before readout, a special clocking pattern was devised which first read out a single line, then two binned together, then three binned, and so on, up to a binning of 19 lines. In this way, a CCD "image" contained a range of signal sizes (alternatively, a nonuniform CCD illumination could have been used). A sequence of such images were obtained and the variance obtained for selected pixels (one in each line). This method allows fast, automatic collection of data but has the drawback of requiring stability of the illumination source during the duration of the measurements (though, in this case, the time needed was only of the order of minutes).

Although the mean-variance method is straightforward to implement with most instruments, large errors can result if there is significant noise due to pick-up. Also, there must be no variation in the spatial content between the pairs of flat-field images or any significant nonlinearity (a condition that is usually met with CCDs as long as the illumination level is below saturation).

If the sensor has a nonlinear response (as is sometimes the case, for example, in CMOS active pixel sensors), then the CVF will depend on the signal level. In fact, the CVF can be defined in two ways, either as an *integral* value (the output voltage or ADC value divided by the signal level in electrons) or as a *differential* value (given by the instantaneous slope of the output versus signal electrons curve). Also, while the off-chip electronics noise will remain constant, the sensor noise may not (because of the nonlinearity). Modifications to the method for a nonlinear sensor have been discussed by Pain and Hancock,[36] and these authors describe a method for deriving both the CVF (or conversion gain) and the noise as a function of signal.

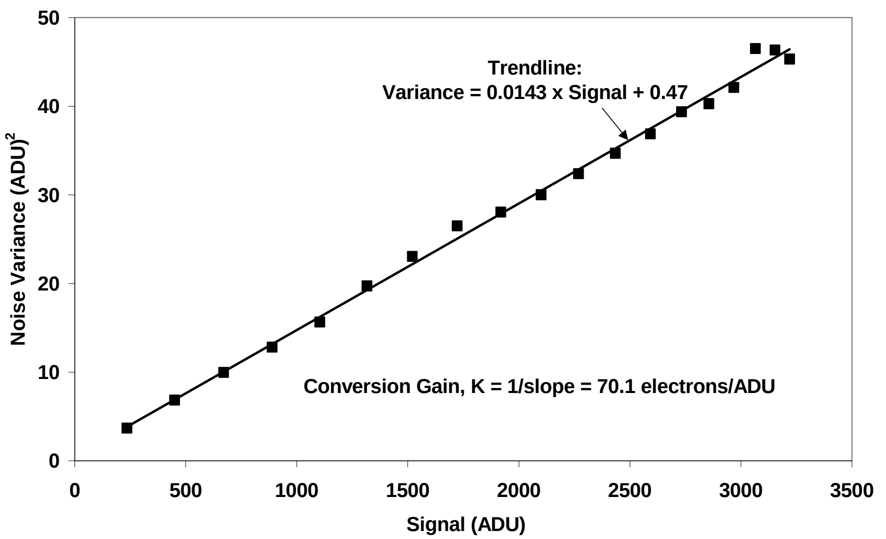


Figure 1.2.9-2 Noise variance versus signal data for a 2-CCD used in an imaging spectrometer. Varying signal levels were obtained by binning along a column.

1.2.10 Temporal Noise

Definition:
The temporal noise on the output waveform arises from the following sources:

- noise associated with the output amplifier
- reset (kTC) noise associated with the resetting of the output stage
- shot noise on any signal (arising either from illumination, spurious charge injection or thermal dark current)
- noise due to external (off-chip) electronics

In a well-designed measurement system, component (4) should be small, but it is unlikely to be zero.

Measurement:
Methods of measuring noise can be categorized as follows:

a. If the array design allows access to the output amplifier connections (e.g., as for a single stage amplifier), then this can be connected to a spectrum analyzer. The reset drain is biased permanently on, and the amplifier is biased as for normal operation. This method is usually only feasible for the array manufacturer.

b. Forming the rms of N successive pixel values X_i for a single pixel, the rms noise is given by

$$\text{rms noise} = \frac{1}{G}\left[\left(\frac{1}{N-1}\right)\sum_{i=1}^{N}\left(X_i - \overline{X_{ij}}\right)^2\right]^{1/2}, \qquad (1.2.10\text{-}1)$$

where G is the gain of the off-chip electronics. For accurate results the number of measurements, N, should be at least 1000, preferably grouped into four lots of 250 so that repeatability (and stability) can be assessed.

The measurement should be made for pre- (or post-) scan pixels if possible and otherwise for pixels in darkness (in which case the noise will include shot noise on the dark charge). In the latter case it is advisable to check the noise for several pixel locations in the array. For IR devices the noise can vary greatly from one pixel to another, and so noise is often measured for every pixel and can be displayed as a matrix of rms values or as a histogram.

c. If the noise on successive pixels in a row or column is not correlated (i.e., if a correlated double sampling amplifier is used to remove low-frequency noise), then noise can be measured as in the mean-variance CVF method discussed above. If two successive acquisitions of a group of pixels are made (X_1, X_2....X_N and $X_1', X_2'...X_N'$), then

$$\text{rms noise} = \frac{1}{\text{G}}\left[\frac{\left(X_1 - X_1'\right)^2 + \left(X_2 - X_2'\right)^2 + ...\left(X_N - X_N'\right)^2}{2N}\right], \qquad (1.2.10\text{-}2)$$

where G is the gain of the off-chip electronics.

In cases (b) and (c) the measured noise will include a contribution from the off-chip electronics (they will add in quadrature). The electronics noise can be separately assessed by replacing the array with a resistor of value approximately equal to the output impedance of the array (typically a few hundred ohms) placed between the reset drain (VRD) and the output source (VOS) connections. This resistor is necessary to ensure that the off-chip electronics is provided with a DC signal of the correct voltage and impedance.

1.2.11 Linearity

Definition:
The array response will depart from linearity at high signal levels due to saturation of the diodes or MOS capacitors and with some types of detectors (but not usually visible CCDs or PDAs,) nonlinearities can also be seen for small signals.

The linearity error (LE) is defined as the discrepancy between the output signal and the ideal (straight line) behavior. The maximum allowed linearity error will usually also define the useful saturation level or full well capacity, since the largest departure from linearity will generally occur at saturation, when either charge comes into contact with interface traps and CTE is degraded (surface full well) or charge overflows into neighboring pixels (bloomed full well).

A problem with linearity error is that it can be specified either as a percentage of the signal at full well (the commonest case) or as a percentage of the signal level at which it is measured (tending to give a higher value of LE for small signals). The latter definition is to be preferred even if a higher LE has to be allowed for small signals because of measurement errors.

Measurement:
Measurement of linearity error involves obtaining output data for a range of known exposure levels. Either uniform illumination of a group of pixels (or the whole array) or spot illumination of a single pixel can be used; however, the two methods will often give different results near saturation. Uniform illumination will tend to give smaller linearity error close to full well, because interface traps are kept filled by preceding charge packets. Hence, if a high full well capacity is needed for spot illumination conditions, then a measurement with this type of illumination should be specified. The exposure can be varied in several ways:

1. Using neutral density filters or liquid absorption standards (see Secs. 4.2.5 and 5.2.1). It is difficult to achieve high levels of accuracy using this technique unless care is taken with alignment and cleanliness of the filters, and avoidance of spurious reflections. It is emphasized that the filter transmittance will vary with wavelength and the wavelength of calibration and use should be as close as possible. Metal film filters tend to have less wavelength variation but suffer more from spurious reflections. A stable light source is also needed (e.g., lamp or LED) and for accurate work this may require use of an optical feedback circuit. For low-light-level illumination, a “beta light” tritium source[37] can give suitably stable illumination. (The stability is determined by the half life of tritium, which is 12.3 years, giving a stability of 0.016% per day.)

2. Using a pulsed LED. LEDs can be pulsed on and off at high speed. The exposure can be controlled by varying the number of pulses received during the exposure time. This is a simple and accurate technique; however, a custom-made circuit will usually be needed to control the pulses and to synchronize with the array readout. Stabilization of LED current and temperature may be required for accurate work but will not generally be needed for short-duration checks. A suitable LED source can be constructed in the laboratory (see Ref. 38), or commercial units are available (see, for example, Ref. 39).
3. Changing the integration time. This generally gives accurate measurements, provided that the light source is stable (see above). This method relies on the user having the ability to change the array clocking waveforms so as to vary the integration time, and may become impractical if many long exposures are needed.
4. Using a good flat-field illumination and operating the array in a continuous readout (or TDI mode)—if the array electronics allows this mode of operation.
5. Techniques that rely on "flux superposition"; for example, multiple apertures or sources. Also, the inverse square law fall-off in intensity from a point source can be used. These methods are more often used for instrument-level calibrations (particularly for spectroradiometers) and are discussed in more detail in Sec. 4.2.5.2.

Figure 1.2.11-1 shows linearity measurements at three wavelengths for a CCD instrument, obtained by varying the integration time. It can be seen that over most of the range the linearity error (in this case expressed as a percentage of full well) is less than 0.2% and does not depend on the wavelength of illumination (though there is a slight trend for the linearity error to increase for large signals in the red). If the linearity error is expressed as a percentage of signal, then it is still less than 0.5%, except for the smallest signals (where measurement errors tend to dominate).

It is usually assumed that for most of their range, the linearity of CCDs is better than a few percent and that measurements tend to be limited by experimental effects (e.g., source stability) rather than instrument response. However, linearity is an important check on the overall "health" of an instrument, and is also important for the interpretation of the output data. Hence, frequent linearity checks are recommended. In systems that use pixel binning (Sec. 1.1.4.1) to increase the dynamic range, it is important to check the linearity over the full range (see also the following section on full well capacity). CMOS active pixel sensors (APS) can be more nonlinear than CCDs because of changes in the photodiode capacitance with signal. Hence, it can be especially important to check linearity with these devices.

It will not usually be necessary to perform linearity checks at more than one wavelength (or waveband), but for accurate work the user should satisfy himself that there is no wavelength dependence to the linearity error.

1.2.12 Full Well Capacity (Saturation)

Full well capacity is often defined as the point at which linearity error exceeds a defined limit (usually ~5%) and is measured as a by-product of linearity measurements. However, it can be checked directly by examination of smear into over-scanned pixels or into areas of the image that have low signal (achieved, for example, by imaging a suitable target).

For a CCD, full well capacity can be defined separately for the image region, the readout register and the output amplifier, the latter two quantities being important when pixel binning is used.

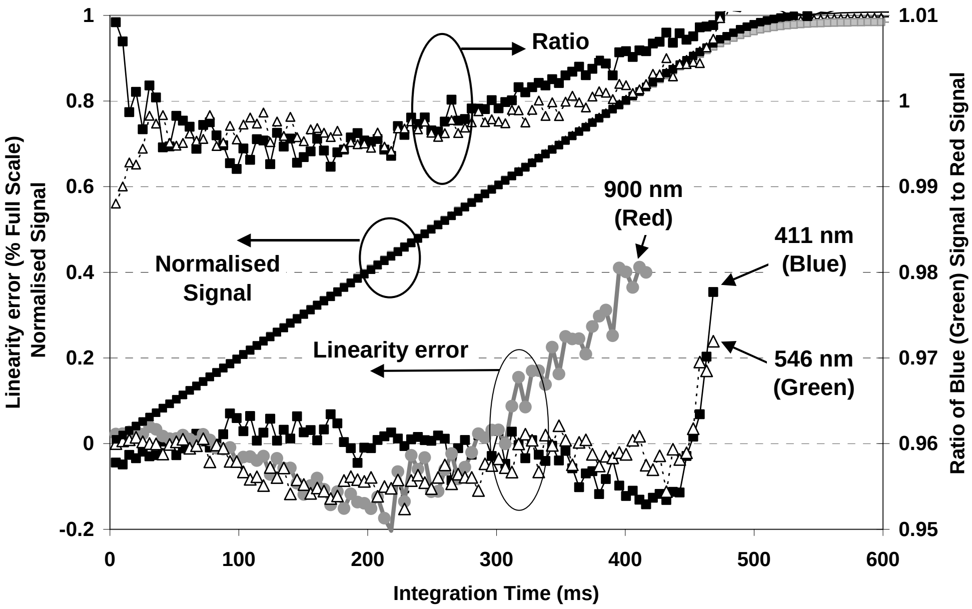


Figure 1.2.11-1 Linearity measurements at three wavelengths for a CCD instrument, obtained by varying the integration time. The signals were normalized to have unity saturation level.

1.2.13 Charge Transfer Efficiency (CTE) of a CCD

Definition:

The charge transfer efficiency (CTE) of a CCD, denoted by ε, is usually defined as the proportion of charge that is transferred from *one pixel to the next*, so that after N pixels the fraction that is correctly transferred (and not trapped or deferred into trailing pixels) is $(1–\varepsilon)^N$. Note that CTE is sometimes defined as the proportion of charge transferred from *one electrode phase to the next*, but this definition is to be discouraged.

In an area CCD the CTE can be defined for horizontal (serial) transfer (HCTE) and for vertical (parallel) transfer (VCTE).

Measurement:

There are several methods for measuring CTE, but they can be grouped into two categories:

1. methods based on measuring the charge lost from a charge packet, as a function of the number of transfers;

2. methods based on measuring the charge deferred into trailing pixels, as a function of the number of transfers.

In general, charge lost during a transfer may be deferred over several pixels so that the two types of measurements will not necessarily give the same results in all circumstances. Although CTE is often quoted as a constant parameter (having a single value), it will usually depend on signal and background charge level, as well as on clocking speeds and temperature. Not all measurement methods can give CTE as a function of signal.

Area CCDs can also exhibit smear due to the frame transfer process (Sec. 3.1.16), and this has to be subtracted before the CTE is calculated.

EPER Method

The extended pixel edge response (EPER) method[24] relies on uniform optical illumination and over-scanning in either the horizontal or the vertical direction (depending on whether measuring HCTE or VCTE), which can often be done for CCD camera systems. Ideally, these over-scan or "extended" pixels will not contain any signal; but if the CTE is not perfect, then a charge will be spread out (or "deferred") over several trailing pixels. The CTE is estimated as follows:

$$\mathrm{CTE} = \frac{\text{total deferred charge (summed over all trailing pixels with significant signal)}}{\text{flat field signal level (in each pixel)} \times \text{number of pixels in a column (or line)}}\ . \qquad (1.2.13\text{-}1)$$

Although the EPER technique can be readily implemented for a range of signal levels, with many CCD instruments it relies on the deferred charge not being spread over too many pixels (or measurement errors will result) and on there not being any edge structures, such as masked columns or lines that might affect the detection of deferred charge. An average of several images can be taken so as to reduce noise and improve measurement accuracy.

Test Targets

Test targets such as grid or bar patterns can be used to give a quick qualitative assessment of CTE, especially to detect any problems at low light levels. However, the accuracy of measurements of CTE will be affected by the finite image blurring produced by optical effects.

First Pixel Response (FPR) Method

The FPR method is a technique to measure total charge loss using uniform illumination. It relies on there being a split between electrode structures (e.g., the split between image and storage region electrodes in a FT CCD) so that a "knife-edge" can be electronically created in the image. The difference between the first line of pixels read out after the knife-edge and subsequent flat-field pixels gives the charge lost, from which the CTE can be calculated knowing the flat-field signal level and the number of transfers.

The method is usually straightforward to implement for vertical CTE measurement with area CCDs, but it does require special clocking sequences and so will not be generally applicable to instrument-level measurements by users. Measurement of horizontal CTE can be performed on CCDs with a split readout register (with an amplifier at both ends). Further details have been given by Janesick,[26] Gregory et al.,[40] and Hopkinson.[27] As with the EPER technique, averages of several images can be obtained so as to improve accuracy.

Spot Illumination

CTE can be measured by comparing the brightness of spot signals obtained from different regions of the CCD (for which differing numbers of transfers are needed before readout) or from the charge deferred to trailing pixels. If the CCD clocking signals can be varied, then an improved technique[41] is to transfer the spot image into the shielded storage region of a frame transfer CCD (or to close an external shutter or use a pulsed LED) and then to repeatedly move the charge up and down (or from side to side in the readout register, in the case of horizontal CTE) so as to increase the number of charge transfers and eliminate the need to produce multiple spot images.

X-ray Methods

As described in Sec. 1.2.9, x-rays from a Fe^{55} or Cd^{109} (or similar) source can be used to produce point charge "events" in the CCD. CTE can be measured by plotting the size of the x-ray peak (not counting partial or split events) as a function of line (or column number). Such a plot is sometimes termed the *stacked response*.[26] Although a single x-ray source does not give CTE as a function of signal, the method gives a simple check on the imaging performance of an instrument and can be used to identify traps or local regions of poor CTE.

Electrical Injection

Some (but not many) CCDs have an injection gate at the end of the horizontal register, which can be used to inject a pulse of charge. If a train of pulses is injected, then the difference in readout signal between the first and the last pulse gives a measure of the charge loss. This is often termed the *periodic pulse technique*.

1.2.14 Modulation Transfer Function (MTF)

Definition:

The quality of an imaging system can be characterized by the response to point illumination (the point spread function, PSF) or by the Fourier transform of the PSF, the optical transfer function (OTF). The magnitude of the OTF is the modulation transfer function (MTF). The phase gives the phase transfer function (PTF), which in most imaging systems is small and can be neglected.

MTF is a measure of the spatial resolution of the image produced by CCD system. If an image is formed of a sinusoidal grating and the maximum and minimum intensities in the image are I_{max} and I_{min}, then

$$MTF = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad (1.2.14\text{-}1)$$

at the particular spatial frequency of the imaged grating. See also Sec. 3.1.13. If a square-wave target (rather than a sinusoidal pattern) is used, then the measured quantity is the contrast transfer function (CTF).

With a sampled imaging system (such as any array-based instrument) the form of the image depends on the position of the image relative to the sampling grid. Hence, the system is not shift-invariant and the simple theory of image formation does not apply. In this case, the MTF cannot be defined in the normal way. This is has been discussed in detail by Williams,[42] who has defined the MTF of a sampled system as "the modulus of the Fourier transform of the one-dimensional responsivity across the effective sampling aperture" (i.e., of the pixel, including crosstalk to other pixels); see also Ref. 43.

Because crosstalk between pixels depends on the absorption depth in the detector, MTF is wavelength dependent. If the pixels are non-square, or in interline transfer arrays (where columns are separated by vertical CCD registers), the MTF will have different values in the horizontal and vertical directions. Also, the MTF for moving objects will be affected by the integrating nature of the detector array.

Measurement:
The MTF can be measured in several ways: by scanning a slit image across the array, by imaging a knife edge, or by use of sine-wave or random targets (either directly or using holographic or laser speckle techniques). Some methods are more suited to spot checks at a single spatial frequency or wavelength. Some can be used for instruments (with associated optics) as well as for arrays. The methods have recently been reviewed by Williams[42] and are briefly discussed below. Readers wishing to perform MTF measurements on arrays or instruments are encouraged to consult the references for further information. The scanned and tilted slit methods are briefly described in ISO standard 15529:1999, "Optical transfer function – principles of modular transfer function (MTF) of sampled imaging systems."

Slit Image (scanned)
When using slit images, the slit can be scanned across the array so as to produce a line spread function (LSF) which can be Fourier transformed to form the MTF. A good-quality optical projection system is needed (whose MTF is known) as well as a scanning system. The method tends to exhibit poor signal to noise ratio, particularly when measuring arrays with small pixel size—since the slit width needs to be reduced (this also leads to diffraction effects). Hence, the other techniques, discussed below, are to be preferred, though the method is useful if the MTF needs to be characterized for small regions of the device (i.e., for a few pixels rather than as an average over many pixels). The effect of the finite slit width can be allowed for by dividing the measured MTF by the Fourier transform of the slit profile.

Knife Edge (Tilted)
Use of the knife edge technique has been discussed in detail by Tzannes and Mooney.[44] The advantages with this method are that no scanning system is needed, signals tend to be high and (with a 2D array) the number of samples can be increased by tilting the knife edge so that several profiles can be obtained with one image, each with a slightly different phase with respect to the pixel positions. The steps in the process (as described by the authors) are as follows:

1. A single image (or average of several images) containing a slightly skewed horizontal or vertical edge is obtained.
2. For each line of pixels in the image an estimate of the position of the edge is calculated (to subpixel accuracy). This is done by a linear interpolation of the pixel on either side of the intensity midpoint, which is the mean of the average pixel value at the top of the step and the average pixel value at the bottom of the step.
3. A straight line fit though the edge-position estimates from (2) is obtained.
4. Knowing the tilt of the edge with respect to the pixel array (from step (3)), all the scans are shifted to a single reference location so that the (interpolated) edge points in all of the scans are aligned. Superimposing all the scan lines creates an edge spread function (ESF) with high spatial resolution.
5. To improve accuracy, the ESF can be fitted by an analytical function (the authors suggested a summation of three Fermi functions, each of the form $F(x) = a \big/ \left\{ \exp\left[(x-b)/c \right] - 1 \right\}$, where a, b, and c are constants).
6. The ESF is differentiated with respect to the scan direction to give the LSF (this is straightforward if the analytical fit is used).
7. The LSF is sampled at a high enough sampling rate to avoid aliasing in the discrete Fourier transform (DFT).
8. The DFT of the sampled LSF is computed; the magnitude of the DFT is an estimate of a 1D profile though the center of the 2D system MTF (in the direction perpendicular to the knife edge).

Slit Image (tilted)

Instead of using a tilted knife edge, it is also possible to use a tilted slit. The averaged line spread function (LSF) is obtained by averaging the LSF for each row (or column) after shifting each LSF to a reference location (as for the tilted knife edge, above). As for the scanned slit method, the effect of the finite slit width can be allowed for by dividing the measured MTF by the Fourier transform of the slit profile.

Imaging of Sine-wave, Holographic or Random Patterns

Sine-wave test targets (see, for example, Ref. 45) can be imaged and the MTF calculated directly from Eq. (1.2.16-1). Results are needed for several phases with respect to the pixel array (encompassing both the minimum and maximum MTF curves) and can be obtained by tilting the target. The MTF should be taken as midway between the maximum and minimum values. A disadvantage of the sine-wave target is that a given target only provides information at one spatial frequency; however, the technique is simple and easy to use for spot checks of the MTF. An alternative is to employ the fringe pattern from two interfering laser beams,[46,47,48] though these cannot be used if there are any intervening optics (e.g., camera lenses). The spatial frequency of the pattern can then be changed by altering the angle of the mirror in one of the laser beams, but a major disadvantage is that it is difficult to perform measurements at several wavelengths. Another disadvantage is that multiple reflections can cause spurious interference patterns and artifacts in the image.

To overcome the disadvantage of needing a separate sine-wave target for each spatial frequency, Sensiper et al.[49] have reported the use of laser speckle. In this, an integrating sphere and aperture are used to produce a speckle pattern on the array whose features scale with the aperture dimension, the laser wavelength, and the distance between the aperture and the array. The nature of the speckle pattern ensures that its phase is random across the array, and since the input power spectrum is known (it is proportional to the autocorrelation of the aperture intensity function) and the output spectrum can be measured, the MTF can be found from the ratio (as a function of aperture to array distance). Ducharme and Boreman[50] have reported an improved technique that reduces the laser power needed. This uses a holographic target. Such targets are now commercially available (for example, see Ref. 45). Playback of the hologram involves illumination by a divergent laser beam (formed using a plano-concave lens). Speckle images are recorded for various distances between the array and the real image formed by the hologram of the aperture used to produce it. For the highest diffraction efficiency, the reilluminating laser beam should be incident on the hologram at the same angle as used for the recording. Images are recorded for several array to image distances and the power spectrum of the speckle pattern determined in each case. The peak of the side band and the central zero frequency are measured and the ratio is $0.5 \times (\mathrm{MTF})^2$.

A method that is straightforward to use at any wavelength with CCD cameras as well as bare arrays has been described by Daniels et al.[51] This uses a random test pattern that has a flat power spectrum and a uniform light source rather than a laser. As with the holographic and laser fringe methods, the whole of the image plane can be measured at one time. In the authors' description, a transparent target was used and methods are given for generating targets for the visible and IR. Reflective targets on semi-matte photographic paper are also commercially available.[45] The MTF is simply the square root of the measured power spectral density.

Guerineau et al.[52] have reported a method of determining the MTF of mid-IR arrays that uses a canted periodic target.

As an alternative to using sine-wave targets, rectangular (bar) targets are often used to yield the contrast transfer function (CTF). As with sine-wave targets for MTF, results for a range of image positions relative to the pixel grid are needed to average out the phase variation of the CTF.[53] In fact, it is possible to use square-wave targets to derive MTF. Since a square-wave target is composed of a sum of sinusoidal harmonic elements, an inversion of this relationship will provide the sinusoidal response (i.e., the MTF) in terms of a sum of square-wave harmonic responses. This technique (originally suggested by Coltman[54]) has been discussed by Dutton et al.[55] and can give good results, but relies on high spatial frequencies being attenuated (or bandlimited) by the MTF of the optics used.

Issues related to the MTF characterization of TDI arrays have been discussed by Lomheim et al.[56]

1.2.15 Responsivity (and Spectral Responsivity)

Definition:
Responsivity, R, is usually defined as the ratio of useful signal voltage (i.e., excluding dark signal) to exposure (in $\mathrm{J/m^2}$) for a given waveband of illumination. An alternative is to define R

in terms of the total photo-current generated in the array, in which case R is specified in A/W.

Measurement:
Determination of the output voltage is straightforward (either directly, using a calibrated oscilloscope) or via the digitized signal (knowing the conversion gain in V/ADU). The current can be measured by determining the current in the reset drain bias supply line.

The exposure is the irradiance multiplied by the integration time (which will be known). Measurement of the exposure requires either a calibrated narrowband source (for example, broadband source, waveband filter, and integrating sphere combination) positioned at a defined distance from the array (see Sec. 4.2.2 for more information on responsivity calibration using sources) or an uncalibrated narrowband source and a calibrated reference detector (for example, a large-area silicon photodiode). In the latter case, measurements are first made with the source illuminating the array, then the array is replaced with the calibrated detector (placed at exactly the same location). The reference detector will give the irradiance directly in W/m^2 or similar units. The area of the detector will usually be known (or can be measured) so that irradiance (on the reference detector) can be converted to radiant flux in watts.

There are two variations on the experimental method. Either the array is uniformly illuminated with a measured irradiance in W/m^2 or a small group (spot) of pixels is illuminated (for example, with a optic–optic probe placed close to the array surface, or a using a spot projection system) and the total radiant flux is measured (with the reference detector) in watts. The *average* signal for N pixels enclosing the spot is then determined; the area of the array illuminated being N × pixel area (pixel area is usually quoted by the manufacturer). The illuminated area is used to convert the radiant flux (on the array) to the irradiance in J/m^2. In the latter (spot projection) method, which is often used in conjunction with a monochromator (for wavelength selection), it may be that the spot has a larger area than that of the reference detector, in which case an integrating sphere attachment for the reference detector will be needed (and the reference detector should be calibrated with the sphere in place).

An example reference detector is the UDTS370,[57] which can be supplied with various detector heads, each with a calibration EPROM module for a range of wavelengths. Such detectors can be independently calibrated by standards institutes such as NIST or NPL (see Sec. 7.1.2) and need recalibration only every three years or so (although annual calibration may be needed for accurate work).

Note that an absolute responsivity determination is needed at only one wavelength. If the responsivity is needed at several wavelengths (the spectral responsivity), then these can be measured relative to the absolute responsivity measured at a single wavelength. As an example, a tungsten lamp and monochromator may be used to illuminate the array at known wavelengths. The calibrated reference detector is needed to measure the relative output of the monochromator (so as to correct for the spectral response of the lamp and the efficiency of the grating), but it does not need to be positioned in the same place as the array except for the single absolute measurement (as long as it is not moved during the course of the measurements).

The most difficult aspect of responsivity measurements is usually the matching of signal levels for the calibrated detector and the array, especially at the extremes of the spectral response (in the red and blue/UV). The calibrated detector, being a large-area photodiode, is more suited to high illumination levels at which the array may saturate. Hence, an accurate method of reducing the exposure for the array (e.g., calibrated neutral density filters or changes in exposure time) may be needed. If necessary, the spectral response can be measured in a piece-wise fashion, with several changes in illumination (as long as the illuminations can be related to each other).

Methods for the calibration of the responsivity of array-based instruments for spectrometry are given in Sec. 4.2.2.

1.2.16 Quantum Efficiency (QE)

Definition:
The quantum efficiency at a given wavelength is the ratio between the number of electrons generated in the semiconductor (corresponding to useful signal, i.e., excluding dark signal) and the number of incident photons.

Measurement:
QE can be related to the responsivity, *R*, (discussed earlier):

$$QE = \frac{hcR}{CVF\lambda} \times 10^{15}, \tag{1.2.16-1}$$

where

h = Planck's constant = 6.626×10^{-34} J s,
c = velocity of light = 3.0×10^{8} m/s,
R (V/(J/m^2) = responsivity,
CVF (μV/e) = charge-to-voltage conversion factor (Sec. 1.2.9), and
λ (nm) = center wavelength of the spectral band used for measurement.

If *R* is specified in A/W, then

$$QE = \frac{hcR}{e\lambda} \times 10^{9}, \tag{1.2.16-2}$$

where e = electronic charge = 1.602×10^{-19} C.

1.2.17 Lag

Definition:
Lag is the effect carried forward from one image frame to subsequent frames. The effects are most noticeable for sharp transitions in image brightness. Depending on the array design, there can be effects either for bright-to-dark or dark-to-light transitions (or both). The effect can be seen in some types of (frame-) interline transfer or linear CCDs and also in some CMOS active pixel sensors. (See also Sec. 3.1.12).

Measurement:
Specific clocking waveforms are needed to synchronize array readout with exposure from a pulsed source (e.g., an LED).

1.2.18 Crosstalk

Definition:
Crosstalk is the amount of signal generated within one pixel that is detected in adjacent pixels and is usually specified as a percentage (see also Sec. 3.1.11).

Measurement:

Spot Method
If a spot of light (of diameter that is less than pixel dimension) is scanned across several pixels (in the vertical or horizontal direction), then the crosstalk can be obtained as in Fig. 1.2.18-1:

$$\text{crosstalk (\%)} = \frac{\text{area A} + \text{area B}}{\text{area A} + \text{area B} + \text{area C}} \times 100. \quad (1.2.18\text{-}1)$$

Spot illumination can be achieved using a pinhole (e.g.,10 μm diameter) mounted 160 mm (the mechanical tube length) from the shoulder of a suitable microscope objective (e.g., × 10 or × 20). A long working distance objective may be needed if the array is mounted behind a thick optical window (e.g., in a cryostat). As an alternative, an edge can be scanned across the array and the pixel profile obtained by differentiation.

Alternative Method
The illumination is arranged so that

- signal (S_0) is measured for a pixel whose geometrical aperture is illuminated but the rest is in darkness

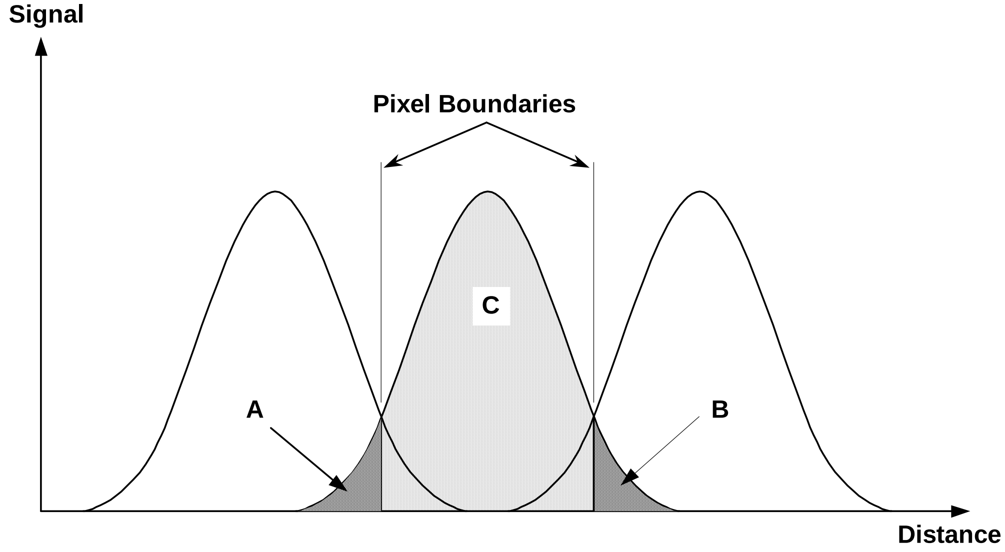


Figure 1.2.18-1 Definition of crosstalk from pixel scan measurements.

- signal (S_1) is measured for the same pixel when its geometrical aperture is in darkness but the rest is under illumination (at the same level)

$$\text{crosstalk}(\%) = S_1/(S_0 + S_1) \,. \tag{1.2.18-2}$$

1.2.19 Geometrical Measurements

Errors in the geometry of the pixel grid are discussed in Sec. 3.1.17, where reference is given to one of the few measurement techniques. Alignment of the array within its package (and of the package within an instrument) is straightforward by conventional microscope inspection. Flatness of the detector array, where important, is usually measured by the manufacturer by interferometric techniques (usually to an accuracy of a few tenths of 1 μm).

1.3 Photometric Versus Radiometric Units

Two sets of units can be used in specifying illumination levels, intensities, etc. Radiometric units relate to the amount of energy that is received by a detector, whereas photometric units relate to the perceived response of the human eye, and are therefore weighted accordingly. The various quantities and their units are given in Table 1.3.1.

Each of the radiometric quantities can also be expressed in terms of their distribution with wavelength; spectral radiant intensity, for example, is the radiant intensity as a function of wavelength and has the units W/sr/nm.

The conversion factor between photometric and radiometric units, K, (in lm/W) is

Table 1.3.1 Radiometric quantities and their units.

Photometric Quantity	Photometric Unit	Radiometric Quantity	Radiometric Unit
Luminous Flux	lumen (lm) = cd · sr	Radiant Flux	W = J/s
Luminous Intensity	candela (cd)	Radiant Intensity	W/sr
Illuminance	1m/m^2 = lux	Irradiance	W/m^2
Luminance	cd/m^2	Radiance	W/m^2/sr
Quantity of Light	1m · s	Radiant Energy	J
Luminous Exposure	lux · s	Radiant Exposure	J/m^2
Exitance	1m/m^2	Emittance	W/m^2

$$K = 683 \frac{\int \phi_{e,\lambda}(\lambda) V(\lambda) d\lambda}{\int \phi_{e,\lambda}(\lambda) d\lambda} , \tag{1.3-1}$$

where $V(\lambda)$ is the photopic spectral luminous efficiency function that defines the response of the average human eye at normal lighting levels, and $\phi_{e,\lambda}(\lambda)$ is the spectral radiant quantity. *K* is known as the luminous efficacy of the radiation. In this equation there are no limits to the integration (that is, the integration is over all wavelengths).

In most cases, the theoretical luminous efficacy of a particular source of optical radiation is of relatively little interest. What is more important is to be able to relate a measurement made in radiometric units to a corresponding value in photometric units. For example, the absolute spectral responsivity of a silicon photodiode, $R(\lambda)$, is frequently given in terms of A/W, whereas it may be of more interest to know its responsivity for a particular source in terms of A/lm. A responsivity correction factor, K_{resp}, can be determined as follows:

$$K_{resp} = 683 \frac{\int_{waveband} \phi_{e,\lambda}(\lambda) V(\lambda)\, d\lambda}{\int_{waveband} \phi_{e,\lambda}(\lambda)\, d\lambda}, \tag{1.3-2}$$

where the waveband is set by the spectral sensitivity range of the array or by the source.

The fundamental quantity for a detector array is the amount of charge detected, usually expressed as a number of electrons. Charge can be converted to output current (say in nA) if the exposure time is known. The detected charge can also be related to the incident exposure (in J/m^2), provided that the relative spectral output of the source, the quantum efficiency (as a function of wavelength), and the pixel area are known. Knowing the eye response, $V(\lambda)$, and the spectral output of the source, the conversion factor, *K*, can also be calculated and hence the responsivity, in nA/lux, and the saturation illumination, in lux.

Figure 1.3-1 shows the eye response and the normalized quantum efficiency of a typical front-illuminated CCD. It is seen that they are not well matched—the CCD has much higher response in the red. This leads to variations in responsivity (in nA/lux) and saturation illuminance depending on the source used—as shown in the Table 1.3.2.

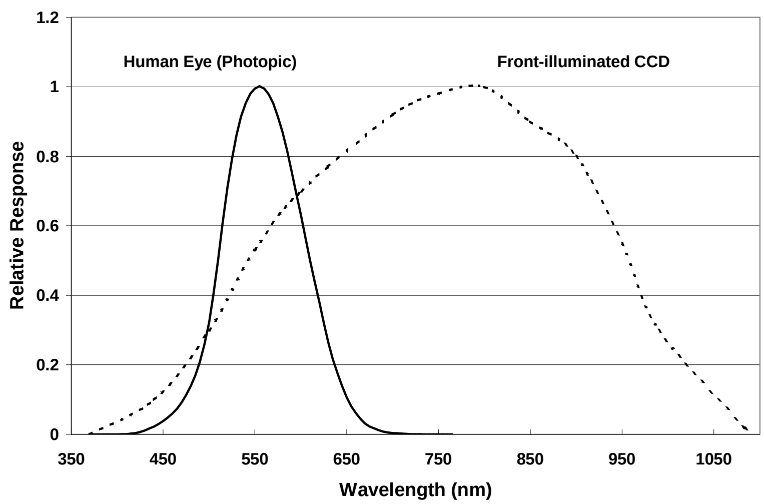


Figure 1.3-1 The photopic (light-adapted) response of the human eye, and the response of a typical front - illuminated CCD.

Table 1.3.2 Values of the conversion factor, K, for various light sources, and also the responsivity and saturation illuminance of a typical CCD camera when illuminated by these sources.

	Conversion Factor, K (lm/W)	**Responsivity (nA/lux)**	**Saturation Illuminance *at the sensor*, at TV rate (lux)**
Daylight, no clouds	140	-	-
3000 K	-	75	2.7
3000 K, IR filter	-	22	9.1
2850 K	16	85	2.4
2850 K, IR filter	350	23	8.7
Red LED	60		
Green 555 nm	683 *	5	40

* by definition, at 555 nm: 1 W = 683 lumens, 1 W/m^2 = 683 lux

Typical levels of illuminance for a variety of lighting conditions are given below.

Lighting	**Illuminance (lux)**
full daylight	10^4
overcast sky, TV studio/shop window	10^3
office/room	~ 250
streetlights	~ 10
twilight	1 – 10
full moon	10^{-1}
starlight	10^{-3}

1.4 Output Signal Calculations

In the visible region of the spectrum, each photon that is absorbed by the silicon creates an electron that is stored in the detector element (MOS capacitor or photodiode). The fraction of incident photons that are absorbed is the quantum efficiency (QE). Usually the charge that is accumulated during an exposure is converted to a voltage at the output amplifier. A typical capacitance for the output stage is 0.15 pF, corresponding to a CVF of 1 μV/electron. Signals can be specified in either volts or the equivalent number of electrons.

The output signal per pixel in volts is calculated as

$$\text{Signal (V)} = (CVF \times 10^{-6}) \int_{waveband} QE(\lambda) T(\lambda)\, E(\lambda)\, \frac{\lambda}{hc} t_{int} \frac{A}{4F^2} d\lambda\,, \qquad (1.4\text{-}1)$$

where $T(\lambda)$ is the optics transmission, $E(\lambda)$ is the scene irradiance (in W/m^2/nm), hc/λ is the energy of each photon (h is Planck's constant and c the speed of light), F the F-number of the lens, A the pixel area, and t_{int} the integration time for the array.

References

[1] J. M. Harnly and R. E. Fields, "Solid-state array detectors for analytical spectrometry," *Appl. Spectrosc*., vol. 51(9), pp. 334A-351A (1997).

[2] D. A. Scribner, M. R. Kruer and J. Killiany, "Infrared focal plane array technology," *Proc. IEEE*, vol. 79(1), pp. 66-85 (1991).

[3] P. R. Norton, "Infrared image sensors," *Opt. Eng*., vol. 30(11), pp. 1649-1663 (1991).

[4] A. Rogalski, "New trends in semiconductor infrared detectors," *Opt. Eng*., vol. 33(5), pp. 1395-1411 (1994).

[5] A. Rogalski, *Infrared Photon Detect*ors, (SPIE Optical Engineering Press, Bellingham, WA, 1995).

[6] W. D. Rogatto, ed., *The Infrared and Electro-Optical Systems Handbook, vol. 3: Electro-Optical Components* (Infrared Information Analysis Center, Ann Arbor, MI & SPIE Optical Engineering Press, Bellingham, 1993).

[7] D. J. Burt, "Extending the performance limits of CCD image sensors," *GEC J. of Res*., vol. 12(3), pp. 130-140 (1995).

[8] A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices* (Kluwer, Dordrecht, 1995).

[9] E. R. Fossum, "CMOS image sensors: electronic camera on a chip," *IEEE Trans. Elec. Dev*., vol. 44(10), pp. 1689-1698 (1997).

[10] J. Janesick, "Charge coupled CMOS and hybrid detector arrays," *Proc. SPIE*, vol. 5167, pp. 1-18 (2004); see also, in the same volume: G. Atlas and M. V. Wadsworth, "Hybrid imaging: a quantum leap in scientific imaging," pp. 121-126; and Y. Bai, S. G. Bernd, J. R. Hosack, M. C. Farris, J. T. Montroy and J. Bajaj, "Hybrid CMOS focal plane array with extended UV and NIR response for space applications," pp. 83-93. On-line information available from http://www.rsc.rockwell.com/html/cmos.html.

[11] K. A. Parulski, "Color filters and processing alternatives for one-chip cameras," *IEEE Trans. Elec. Dev*., vol. 32(8), pp. 1381-1389 (1985).

[12] B. Dierickx, "Some of the best RGB patterns," available on http://www.fillfactory.com/htm/technology/htm/rgbfaq.htm.

[13] P. Jerram et al., "The LLLCCD: low light level imaging without the need for an intensifier," *Proc. SPIE*, vol. 4306, pp. 178-186 (2001).

[14] J. Hynacek, "Impactron - a new solid state image intensifier," *IEEE Trans. Elec. Dev*., vol. 48(10), pp. 2238-2241 (2001).

[15] H.-S. Wong, Y. L. Yao and E. S. Schlig, "TDI charge-coupled devices: design and applications," *IBM J. Res. Develop.*, vol. 36(1), pp. 83-106 (1992).

[16] S. M. Sze, *Physics of Semiconductor Devices, 2nd Ed.*, (Wiley-Interscience, New York, 1981).

[17] B. Burke and S. A. Gajar, "Dynamic suppression of interface-state dark current in buried channel CCDs," *IEEE Trans. Elec. Dev.*, vol. 38(2), pp. 285-290 (1991).

[18] W. J. Toren and J. Bisschop, "Complete characterisation of dark current in frame transfer image sensors," *Philips J. Res.*, vol. 48(3), pp. 207-231 (1994).

[19] M. Blouke, D. Heitmann, F. Yang and J. Janesick, "Traps and deferred charge in CCDs," in *Instrumentation in Astronomy: Present and Future*, L. B. Robinson, ed., p. 462 (Springer-Verlag, New York, 1988).

[20] E. R. Fossum and B. Pain, "Infrared readout electronics for space science sensors: state-of-the-art and future directions," *Proc. SPIE*, vol. 2020, pp. 262-285 (1993).

[21] N. Bluzer and A. S. Jensen, "Current readout of infrared detectors," *Opt. Eng.*, vol. 26(3), pp. 241-248 (1987).

[22] L. J. Kozlowski, W. V. McLevige, S. A. Cabelli, A. H. B. Vanderwyck, D. E. Cooper, E. R. Blazejewski, K. Vural and W.E. Tennant, "Attainment of high sensitivity at elevated operating temperatures with staring hybrid HgCdTe-on-Sapphire focal plane arrays," *Opt. Eng.*, vol. 33(3), pp. 704-715 (1994).

[23] L. J. Kozlowski, "Low noise capacitive transimpedance amplifier performance vs. alternative IR detector interface schemes in submicron CMOS," *Proc. SPIE*, vol. 2745, pp. 2-11 (1996).

[24] C. T. Troy, "How camera makers cope with a confused CCD market," *Photon. Spect.*, February 1996, pp. 74-86.

[25] ESA/SCC Basic Specification No. 25000, "Electro-optical Test Methods for Charge Coupled Devices," Issue 1, July 1993.

[26] J. Janesick, *Scientific Charge-Coupled Devices* (SPIE Press, Bellingham, WA, 2001).

[27] G. R. Hopkinson, "Proton damage effects in p-channel CCDs," *IEEE Trans. Nucl. Sci.*, vol. 46(6), pp. 1790-1796 (1999).

[28] F. Wang, X.-J. Shen, X.-H. Sun, W. Dong and B. Zhou, "A new method of measuring the CCD device photoresponse nonuniformity," *Proc. SPIE*, vol. 3553, pp. 295-296 (1998).

[29] W. J. Wild, "Reconstructing CCD flat fields using non-uniform background illumination sources," *Proc. SPIE*, vol. 3355, pp. 713-720 (1998).

[30] G. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Trans. Pattern Anal. Mach. Intellig.*, vol. 16(3), pp. 267-276 (1994).

[31] G. R. Hopkinson, "Analytic modelling of charge diffusion in charge-coupled device imagers," *Opt. Eng.*, vol. 26(8), pp. 766-722 (1987).

[32] *Handbook of Chemistry and Physics, 85th Ed.* (CRC Press, Cleveland, Ohio).

[33] See, for example, AEA Technology QSA, Radiation Sources for Industrial Gauging and Analytical Instrumentation.

[34] G. W. Fraser, A. F. Abbey, A. Holland, K. M. McCarthy, A. Owens and A. Wells, "The x-ray energy response of silicon: (A) theory, *Nucl. Instrum. Meth. Phys. Res. A*, vol. 350 (1-2), pp. 368-378 (1994).

[35] B. Stark, B. Noelting, H. Jahn and K. Andert, "Method of determining the electron number in charge-coupled measurement devices," *Opt. Eng.*, vol. 31(4), pp. 852-856 (1992).

[36] B. Pain and B. Hancock, "Accurate estimation of conversion gain and quantum efficiency in CMOS imagers," *Proc. SPIE*, vol. 5017, pp. 94-103 (2003).

[37] Saunders-Roe Ltd., http://www.betalight.com.

[38] S. Tulloch, "Design and use of a light source for linearity measurement in CCD cameras," RGO Technical Note 111, Aug. 1997, available online from http://www.ing.iac.es/Astronomy/observing/manuals/ps/tech_notes/tn111.pdf.

[39] Gamma Scientific, http://www.gamma-sci.com/rs-5-lightsource.htm.

[40] J. A. Gregory, B. E. Burke and M. J. Cooper, "Solid State Research," Quarterly Technical report (ESC-TR-92-149), Lincoln laboratory, MIT, 31 Jan 1993, pp. 48-50.

[41] I. H Hopkins, G. R. Hopkinson and B. Johlander, "Proton-induced charge transfer degradation in CCDs for near-room temperature applications," *IEEE Trans. Nucl. Sci.*, vol. 41(6), pp. 1984-1990 (1994).

[42] T. L. Williams, *The Optical Transfer Function of Imaging Systems*, ch.10, "MTF of Sampled Imaging Systems" (IOP Publishing, Bristol, 1999).

[43] J. C. Feltz and M. A. Karim, "Modulation transfer function of charge-coupled devices," *Appl. Opt.*, vol. 29(5), pp. 717-722 (1990).

[44] T. M. Tzannes and J. M. Mooney, "Measurement of the modulation transfer function of infrared cameras," *Opt. Eng.*, vol. 34(6), pp. 1808-1817 (1995).

[45] Sine Patterns LLC, Pittsford NY, http://www.sinepatterns.com.

[46] M. Marchywka and D. G. Socker, “Modulation transfer function measurement technique for small-pixel detectors,” *Appl. Opt*., vol. 31(34), pp. 7198-7213 (1992).

[47] J. E. Grievenkamp and A. E. Lowman, “Modulation transfer function measurement technique for small-pixel detectors,” *Appl. Opt*., vol. 31(34), pp. 7198-7213 (1992).

[48] M. Willemin, N. Blanc, G. K. Lang, S. Lauxtermann, P. Schwider, P. Seitz and M. Wany, “Optical characterization methods for solid-state image sensors,” *Opt. Las. Engineer*., vol. 36 (2), pp. 185-194 (2001).

[49] M. Sensiper, G. D. Boreman, A. D. Ducharme and R. Snyder, “Use of narrowband laser speckle for MTF characterisation of CCDs,” *Proc. SPIE*, vol. 1683, pp.144-151 (1992).

[50] A. D. Ducharme and G. D. Boreman, “Holographic elements for modulation transfer function testing of detector arrays,” *Opt. Eng*., vol. 34(8), pp. 2455-2458 (1995).

[51] A. Daniels, G. D. Boreman, A. D. Duchame and E. Sapir, “Random transparency targets for modulation transfer function measurement in the visible and infrared regions,” *Opt. Eng*., vol. 34(3), pp. 860-868 (1995).

[52] N. Guerineau, J. Primot, M. Tauvy and M. Caes, “Modulation transfer function measurement of an infrared focal plane array by use of the self-imaging property of a canted target,” *Appl. Opt.*, vol. 38(4), pp. 631-637 (1999).

[53] S.-Y. Lin and W.-H. Chan, “MTF and CTF measuring system for digital still cameras,” *Proc. SPIE*, vol. 3019, pp. 50-57 (1997).

[54] J. W. Coltman, “The specification of imaging properties by response to a sine wave input,” *J. Opt. Soc. Amer*., vol. 44, pp. 468-471 (1954).

[55] T. Dutton, T. Lomheim and M. Nelson, “Survey and comparison of focal plane MTF measurement techniques,” *Proc. SPIE*, vol. 4486, pp. 219-246 (2002).

[56] T. S. Lomheim et al., “Electro-optical hardware considerations in measuring the imaging capability of scanned time-delay-and-integrate charge-coupled imagers,” *Opt. Eng*., vol. 28(8), pp. 911-927 (1990).

[57] UDT Instruments, http://www.udtinstruments.com.

CHAPTER 2

DETECTOR ARRAY EQUIPMENT AND REQUIREMENTS FOR CALIBRATION

In this chapter we consider applications of detector arrays in instrumentation and the impact on calibration requirements. This cannot be an exhaustive review, because the uses of CCDs and photodiode arrays is extremely diverse. However, there are particular classes of applications that have features in common and for which a generic set of calibration requirements can be derived.

The information that comes from an array-based instrument relates to the

1. signal in each pixel—calibration involves an assessment of how faithfully the signal in a pixel (or group of pixels) matches the irradiance falling on the array.
2. sharpness of an image (i.e., the spatial resolution)—calibration involves an assessment of how crosstalk between pixels and image artifacts affect the appearance of images.
3. geometric properties of the image—calibration involves an assessment of distortions in the image caused by deviations from an ideal, uniformly spaced grid of pixels.
4. time at which signals are detected.

In a general sense, the term “image” can be used to describe the ensemble of all the pixel signals from the array, whatever the application. However, we will use the term “imaging” to refer to systems with an imaging, or camera, lens through which objects are viewed, to distinguish these from spectrometry applications where the image is usually of the entrance slit of a spectrometer. Spectrometry applications are generally the most demanding in terms of calibration requirements, and this is reflected in the greater attention to this subject in Sec. 4.2 compared with imaging applications (in Sec. 4.1).

The fourth item in the list above will be important for transient phenomena, but does not usually present any particular calibration problem since array detectors will normally be driven by clocking signals derived from a stable master oscillator. Timing information (as long as it is accessible) will usually be adequate. If absolute time-stamping of images is required, then a possibility is to use the IRIGB signal from a GPS (global positioning system) satellite receiver. This carries absolute time information that is accurate to better than a microsecond. Note that, in general, array-based instruments cannot be instantly triggered by external events, since the exposure and readout follow a sequence defined by the array clocking. If asynchronous external triggering is required, then a special mode is normally used whereby the array is left “in the dark" in an idle state (no clocks operating) or is kept in a continuous flushing state (Sec. 1.1.4.8) prior to the trigger that initiates an exposure and readout sequence. An exception to this general rule are some types of CMOS APS cameras that operate in a continuous current monitoring rather than charge integration mode. These cameras can be triggered for readout at any time. Calibration of timing information will not be discussed further in this guide.

2.1 Imaging Applications

There are a host of applications that use detector arrays for general imaging. Many of these are demanding in terms of detector performance (e.g., low light level, high speed, UV/x-ray imaging), but relatively few require calibration of the radiometric or geometric performance.

As examples of imaging applications of silicon detector arrays that have requirements for calibration, we will briefly discuss the following broad areas:

machine vision
photometric imaging
color measurement
Earth remote sensing
fluorescence microscopy
photogrammetry
thermal imaging

In most cases, the main calibration requirements are to measure dark field and bright field uniformity (dark and flat fielding) and to check linearity and dynamic range (noise and full well capacity). In low-light-level applications the stability of the dark signal will be important as will cosmic ray removal for long exposures. An absolute calibration of responsivity (perhaps as a function of wavelength) will be needed in some cases. Applications such as photogrammetry, which require good geometric accuracy, will sometimes entail calibration of pixel positions.

2.1.1 Machine Vision

Apart from general 2D imaging applications, the most common use of arrays is for line scanning inspection of moving objects (e.g., on a production line) using linear or TDI arrays (TDI is discussed in Sec. 2.1.4.4). The term "web" inspection is often used and applies to objects such as metal, paper, textiles, plastic, or wooden sheets. Reference 1 discusses some of the calibration requirements. Because the primary aim is to detect small defects in materials, good imaging resolution (good MTF and a large number of pixels) is important, as is the calibration of photo-response nonuniformity. In common with many industrial applications, there is often an additional requirement for camera systems to be interchangeable—implying identical calibrations (particularly for responsivity).

2.1.2 Photometric/Radiometric Imaging

Photometric imaging is the use of a two-dimensional CCD camera to determine the brightness of objects. Calibrated digital cameras are now becoming available for this purpose. Applications include inspection of airfield and street lighting systems, automotive, avionics and flat-panel displays, and calibration of light sources. For the latter application, Ashdown and Rykowski[2] have recently reported on the use of data compression algorithms to make the recording of the large number of CCD images required to characterize the near-field output of light sources a practical possibility. Other applications make use of intensified arrays (Sec. 3.2). In this case the output of the camera can be adjusted by varying the voltage applied to the intensifier, and the relation between voltage and detector output has to be calibrated.[3]

Array detectors, with their large number of pixels, nonuniformities, aliasing effects, subpixel response and dependence on operating conditions, are (though excellent in other respects) not naturally suited to absolute radiometric or photometric imaging applications. However, provided that the image does not have detail at a comparable scale with the pixel dimensions, and if the operating conditions are stable, then calibration for photometric use is practicable (through dark and flat fielding, responsivity, and linearity measurement).

Since photometric cameras give an output in photometric units, they must have the spectral sensitivity of the detector array corrected so as to match that of the CIE photopic spectral luminous efficiency function. Filters are used to reduce the spectral bandwidth (CCDs and PDAs are sensitive to longer wavelengths than the human eye). Calibrations usually involve use of a variable, spatially uniform luminance source (such as an integrating sphere at various source distances), with a calibration needed for each lens aperture and focus setting (see, for example, Ref. 4). Also, the dark current and response nonuniformity are calibrated for each pixel. Regular recalibrations of these instruments are necessary to maintain photometric accuracy. The frequency of recalibration will depend on the accuracy required, but will generally be at least annual. The PRNU can only be calibrated by the manufacturer for certain source/filter combinations. Hence, it is likely that scenes with a strong spectral content or involving narrow-band sources (e.g. LEDs) may require calibration by the user.

Photometric cameras will be most accurate when the average signal from a group of pixels is measured. When there is important scene information at or below the pixel level (i.e., when there are significant pixel–pixel variations in irradiance), then it is likely that intrapixel nonuniformities and residual PRNU and DSNU errors (e.g., from traps) will affect the photometric accuracy.

2.1.3 Color Measurement

Instruments used to measure the color of objects require calibration of the relative responses of the three color outputs (red/green/blue) from a color camera (Sec. 1.1.2). This calibration is often incorporated in a look-up table (LUT), which translates the CCD signals into a color value (e.g., a point in tristimulus color space). Because crosstalk between pixels will spread signals from one color pixel to another, there is usually a mixing of color information. Hence, a matrix transformation is sometimes preferred. Since each pixel has its own responsivity, there ideally needs to be a separate matrix (or separate LUT) for each pixel; otherwise, pixel-to-pixel color errors will result. However, this is rarely implemented in practice. Calibration is often performed by imaging a standard color checker (chart or tiles) and adjusting the values in the matrix or

LUTs until the correct color signal from the camera is produced. Color calibration cannot be considered in detail within the scope of this guide; however, Sec. 4.1.1.1 gives a brief discussion.

2.1.4 Earth Remote Sensing

In Earth remote sensing applications it is common for images to be taken at several wavelengths. This is termed *multispectral* or *hyperspectral imaging*, depending on the number of measured wavelengths. Imaging with array detectors is usually performed in *pushbroom* mode, where the array is scanned over the surface of the Earth by the motion of the satellite (or aircraft). Either linear (or TDI) arrays can be used (in conjunction with color filters), or a 2D array can be used in an imaging spectrometer. Since the scene will usually contain high-frequency spatial information, aliasing errors (c.f. Sec. 3.1.14) can be important. Radiometric accuracy is also important for determination of color information and albedo values (albedo is the fraction of sunlight reflected from a material surface). Calibration of camera systems can be performed either in the laboratory or in-flight, using reflectance data from known target areas on the ground. Laboratory measurements often involve imaging of diffuse reflecting surfaces (and comparison with a standard surface, such as Spectralon). References 5 and 6 give examples of laboratory and in-flight calibrations of airborne multispectral CCD cameras. Calibration of imaging spectrometers is discussed in Secs. 2.2.6 and 4.4.

2.1.5 Fluorescence Microscopy

Many microscope applications use a CCD for image recording and subjective analysis. Fluorescence microscopy is an example of when information can be used quantitatively. Fluorescence occurs when an atom or molecule absorbs light and emits at another wavelength. In biological applications, slight variations in the surrounding environment can cause large changes in the fluorescence from *reporter* molecules or *fluorophores*, and these can be used as sensitive probes for the measurement of changes in pH, ion concentration, nucleic acids, and proteins.

A cooled, back-illuminated CCD and long exposures are usually needed because light levels are low (partly because overexposure to light can cause photobleaching of fluorophores). A common approach is to uniformly illuminate a sample with light from a xenon arc lamp that has been filtered or passed through a monochromator so as to select a particular set of excitation wavelengths (often only two wavelengths are needed). Ratios of the fluorescence images obtained with these excitation wavelengths are then obtained so as to determine biological information (such as concentrations). Good linearity, high dynamic range, high spatial resolution and calibrated DSNU and PRNU are required. Ideally the PRNU should be obtained for flat-field images obtained at the useful fluorescence emission waveband.

A related field is fluorescence in-situ hybridization (FISH), a technique that is used, for example, in gene mapping. In this technique, “probe” nucleotide sequences are synthesized so as to incorporate fluorescent molecules. The probes are then hybridized with the denatured (strands separated) target DNA. Fluorescent emission then reveals the relative position of the probes on the gene. High spatial resolution and low noise are important in this application.

A technique that circumvents the common problem of bleaching of the fluorescent material is fluorescence lifetime imaging spectroscopy (FLIM). The excitation is modulated and phase shift of the detected signal relative to the excitation is measured.[7] Since the lifetimes are short, an intensified CCD is used (Sec. 3.2).

2.1.6 Photogrammetry, Profilometry, and Noncontact Metrology

There are many applications in photogrammetry, profilometry and noncontact metrology that require accurate calibration of image distortions. Examples include architectural photogrammetry, accident recording, building monitoring, measurement of motion or deformation in structures and inspection tasks within sterile or nuclear environments, as well as numerous medical and industrial applications. Calibration is usually performed at system level and includes the combination of detector array, optics, and camera/framegrabber electronics. Some target tracking and scientific applications (e.g., astrometry) also require good geometric accuracy.

2.1.7 Thermal Imaging

In the infrared region of the spectrum the brightness of an image is often used to determine either the reflectance properties or the temperature of an object (depending on the wavelength range used). Detector nonlinearity and response nonuniformity are usually significantly worse than for visible arrays and require frequent calibration and/or compensation since they are sensitive to operating voltage and thermal drifts. In addition, images are usually of low contrast, making the effect of response nonuniformity more severe. Calibration of infrared arrays and imaging systems is really a subject in its own right and will only be discussed briefly in this guide (see Sec. 4.1.2).

2.2 Spectrometry Applications

In this section we discuss some of the areas of spectrometry that use detector arrays and the associated performance issues and calibration requirements. The discussion is inevitably somewhat generalized, but it is to be hoped that the reader will gain a flavor of the range of spectrometry applications that use detector arrays. Calibration is particularly important for spectrometric systems; since quantitative results are usually required, some form of calibration for either wavelength scale or signal size (or both) is almost universal. While this section gives a brief overview, detailed guidance on calibration methods and standards for spectrometry are given in Secs. 4.2 to 4.4.

An excellent review of the design and use of the various types of CCD-based spectrometers has been given in the book *Charge-Transfer Devices in Spectroscopy*, edited by Sweedler, Ratzlaff, and Denton.[8] Ferraro et al.[9] have given a review of instruments for UV-VIS, NIR and Raman absorbance spectroscopy, including a detailed list of manufacturers.

Detector array systems can offer a number of advantages over scanning systems. For example:

- All wavelengths in the spectral range are sampled simultaneously, enabling data to be collected very rapidly (an important consideration in applications such as process

control).
- Simultaneous wavelength sampling also makes these systems well suited to measurements on pulsed or time-varying sources.
- The absence of moving parts means instruments can be more stable, reproducible and rugged, and usually also smaller and therefore more portable.
- It is usually possible to integrate the signal over a period of time, which can help reduce the effect of noise in some measurements.

As in imagers, noise, dark signal, linearity, and response nonuniformity are important parameters. However, there are special issues for spectrometric instruments that do not arise for imagers. For example, response nonuniformity is not always easy to measure because the array is often "buried" inside the instrument and not easy to illuminate uniformly. Also, the issue of wavelength calibration only arises in spectrometers, the accuracy of which depends on the methods used (e.g., software algorithms) and the interplay of factors such as pixel size, spectral line shape, and response nonuniformity (both inter- and intrapixel). In a traditional scanning instrument a sine-bar is generally used to linearize the relationship between scan distance and wavelength; calibration at just a few points can give a good assessment of how well this has been done. In array-based instruments, the spectrum is spread across the evenly spaced pixels of the array. Software algorithms are used to link the individual elements of the array with their corresponding wavelength. The relationship is not straightforward, but depends on how well the spectrum is focused at the exit plane, the angle at which the grating has been set, the regularity of the size and spacing of the pixels in the array, etc. Due to these problems, the accuracy of the wavelength scale depends critically on the number of points at which the system is calibrated and the algorithm calculated. For array-based systems, therefore, it is necessary to check the accuracy of the instrumental wavelength scale by calibration at many points spanning the entire wavelength range of interest.

The problems of in-system stray light are also considerably more acute with array-based spectrometers than with traditional scanning systems, and in the blue and UV spectral regions in particular, often dominate all other sources of measurement uncertainty. In-system stray light is due to scattering and reflections within the spectrometer, and levels are often high in array-based systems because

- unlike most high-quality traditional spectrometers, which use double monochromators, array systems use a single monochromator to disperse radiation across the array;
- the small size of most array-based systems, and the need to spread radiation across the whole array rather than a narrow exit slit, makes it difficult to make effective use of baffles in the monochromator;
- radiation can be reflected off the array, onto the walls and/or components within the monochromator, and then back onto the array;
- most array spectrometers use silicon detectors, which have high responsivity in the red and near-infrared and much lower response in the blue and UV spectral regions. This makes them highly sensitive to any stray radiation from the longer wavelengths,

which is also the region of the highest emission from most commonly used sources, such as tungsten and xenon lamps or daylight.

2.2.1 General Spectroscopy

There are a host of instruments that use detector arrays for general, low-accuracy spectrometry in, for example, spectrophotometry, colorimetry, medical diagnostics, and environmental and process analysis. The trend is for miniaturization of the design and no moving parts, so that instruments are portable (hand-held) and inexpensive. Usually a small detector array (sometimes even a few tens of elements) is used and a wide spectral range is covered—so that resolution is low. Nevertheless, wavelength accuracy and low stray light performance is often required. A optic–optic bundle or a single optic is usually used to input light into the spectrometer. Examples of general purpose miniature spectrometers are given in Refs. 10, 11, 12, and 13.

2.2.2 Absorbance and Reflectance Spectrometry

One of the main applications of absorbance spectrometry is the detection and analysis of chemical samples, in the pharmaceutical industry, for example. The samples are usually liquid or gas and require containment in a sample holder—though reflectance probes can be used in the study of solid samples. Often the components of a solution need to be separated, for example, by gas chromatography (GC) or high-performance liquid chromatography (HPLC) (see, for example, Chap. 6 of Ref. 6 for an introductory discussion and Ref. 14 for details on method development). Reflectance spectrometry is widely used in the printing, dyeing, and paint industries for process control and color matching. A large proportion of the array-based instruments used operate in the UV and visible (UV-VIS), but the use of the NIR region is becoming more popular. (Silicon CCDs and PDAs have useful response out to just below 1.1 μm, and InGaAs detectors can be conveniently used out to 1.7 μm, with extension to ~ 2.5 μm). Though written some time ago, Ref. 15 gives a good introduction to the use of photodiode arrays for UV-VIS spectrometry, and Ref. 16 discusses some of the main issues in single- and multicomponent quantitation.

Absorbance and reflectance spectrometry can usually be divided into the two categories—continuous wave (CW) and transient—and these can each be split into low and high absorbance/reflectance cases. In the high-absorbance/low-reflectance case, very little illumination reaches the array and performance tends to be limited by *sensitivity* (the lowest signal level that can be used before noise dominates) and by *stray light* performance. However, the wavelength reproducibility of array-based instruments often allows measurements to be made on the sloping edge of an absorbance/reflectance peak rather than at the peak itself (where signal has a minimum). In the low-absorbance/high-reflectance case, the user is measuring small fluctuations on a large continuum signal, and high signal levels (and high detector *full well capacity*) are often needed if the measurements are not to be degraded by shot noise on the signal.

It is common for measurements to be made with and without the sample in place so that the ratio gives the absorbance/reflectance spectrum, without any need for calibration of the responsivity (provided that the source is stable). However, both the reference and the sample spectra must first be corrected for the dark signal, and linearity and wavelength accuracy must be assessed.

Transient measurements can often be made with conventional systems with temporal resolution down to a few milliseconds, but if higher temporal resolution is needed, then techniques such as kinetics mode (c.f. Sec. 1.1.4), gated intensified CCDs or pulsed sources are needed. With pulsed sources there can be a problem with shot-to-shot reproducibility (in source intensity), and use is sometimes made of a 2D CCD or two photodiode arrays (end-to-end or side by side, depending on the configuration) to record both the reference and the sample beam at the same time.

Optic–optic probes are becoming common attachments to absorption and reflection spectrometers and are useful for sampling from existing containers, viscous samples, and samples that may be hard to remove after measurement. Another important trend is improvements in analysis software, which can now incorporate correction for scattering from the sample as well as wavelength and intensity calibration. A commonly used analysis package is GRAMS/32 (Ref. 17 gives details of the algorithms available).

In analytical chemistry, the development of a validated model for deriving chemical compositions from the multitude of lines in an absorbance spectrum is an important task and forms part of procedures for good laboratory practice (GLP). Commercial spectrometers for use in analytical chemistry are often supplied with software to aid the validation process (see, for example, Ref. 18). Several algorithms such as principal components regression (PCR), partial least squares (PLS), piecewise direct standardization (PDS), and genetic regression (GR) have been developed for chemical analysis. These techniques can improve the precision of measurements in the presence of wavelength shifts and use of multiple instruments, but their specialized application and the degree of mathematics involved puts them outside the scope of this guide. The reader is referred to Refs. 17, 19, 20, and 21 and the references therein for further information. In particular, Ref. 19 gives detailed guidelines for multivariate quantitative analysis that, although specified for the IR, could also be applicable for UV-VIS spectrometry.

2.2.3 Emission Spectrometry

Emission spectrometry covers measurements of the spectral output of optical radiation (e.g. the sun, fluorescent lamps, LEDs) and spectroscopic measurements in which a sample is excited to produce a characteristic spectrum from which information on the sample can be obtained. In atomic emission spectroscopy (AES) and optical emission spectroscopy (OES), a plasma is excited by radio frequency (RF), glow discharge (GD), or spark. Fluorescence spectroscopy (Chap. 8 of Ref. 6) and Raman spectroscopy (Chap. 7 of Ref. 6) also use excitation of a sample to produce characteristic emission spectra, though resolution tends to be lower (particularly in florescence spectroscopy, where 0.2 nm is often sufficient). However, light levels can be low so that integration times are long. Dark signal subtraction and removal of cosmic ray artifacts then become important.

Applications of emission spectrometry include monitoring of solar UV, lamp manufacture, metallurgy, medicine, forensic science, and water analysis. Similar instruments are also used for plasma diagnostics (e.g., in the semiconductor industry). The wavelength, intensity and line width of the emission lines are all important parameters. In general, a large wavelength range and high resolution are needed so as to measure a range of lines at high precision (good linearity and stray light performance are also needed). Sullivan and Quimby[22] discussed a special PDA spectrometer that could be translated along a flat focal plane (produced by a holographic grating), but most modern instruments use echelle spectrometers with 2D CCD or CID arrays [c.f. Chap. 10 of Ref. 6]. Since emission lines are produced with a range of intensities, the random access capabilities of the CID are useful (in effect, different parts of the array can have different integration times), but CCDs can be used with a range of exposures, provided that antiblooming protection is available (to prevent saturated lines from contaminating other parts of the image).

2.2.4 Vibrational Spectrometry

Vibrational spectrometry is used to characterize materials by way of the vibrational spectra of the constituent molecules. There are three basic techniques: (1) NIR spectroscopy, (2) Mid-IR spectroscopy, (3) Raman spectroscopy.

The first two are absorption techniques, whereas Raman is a scattering technique where the sample is irradiated with monochromatic laser light. Due to the different excitation conditions in each case, the same molecular vibration will lead to very different responses. The trade-offs between these techniques include consideration of intensities of the absorption bands (which decrease from the mid-IR to the visible), sample preparation (not needed for Raman and NIR), and water content (water is a weak Raman scatterer but a strong infrared absorber, hence Raman spectroscopy is the best technique for studying aqueous solutions).

An excellent review of the fields of application has been given by Workman,[23] and Coates[24] has reviewed types of instrumentation.

Mid-IR and NIR are forms of absorption spectroscopy and, as such, have been discussed earlier. Raman spectroscopy is discussed in the next section.

2.2.5 Multichannel Raman Spectroscopy

In Raman spectroscopy the sample is illuminated by a monochromatic laser beam and molecules are raised to an excited energy level. Molecules may return to the ground state by elastic scattering, thereby emitting the Rayleigh line that has the same wavelength as the excitation. If the excitation returns by inelastic scattering, the emitted line (the Stokes line) has a higher wavelength, the difference corresponding to the vibrational energy. An anti-Stokes line is also produced but is not usually used as it has lower intensity.

It is important to note that the intensity of the Stokes line is extremely weak, being $\sim 10^{-4}$ the intensity of the Rayleigh line and $\sim 10^{-8}$ times the source intensity. Also, the Raman intensity varies *inversely* as the *fourth* power of the excitation wavelength. Although the NIR region is preferred because fluorescence is lower, the Raman line itself is weaker.

The use of Raman spectroscopy has increased significantly over the past few years due to the availability of low-noise CCD systems, solid-state lasers, holographic gratings and laser rejection (notch) filters, and advances in optic–optic couplings and software. These changes have led to the introduction of more efficient, rugged, and user-friendly systems for use not only in the laboratory but also for online monitoring of industrial and environmental processes. Reference 25 gives a typical example from the chemical industry. Uses in medicine (for example, for in-situ measurements) and in surface-enhanced Raman spectroscopy (SERS) are being developed. Hybrids of Fourier-transform (FT) Raman and multichannel Raman are also being investigated.[26] References 27 and 28 discuss issues for Raman spectrometers used for online chemical analysis (using optic–optic probes).

The NIR is being exploited through the use of Ti:sapphire and diode lasers, and requires the use of CCDs tailored to the 700–1100 nm region (for example, showing reduced fringing effects, c.f. Sec. 3.1.4.2, or having a thick active volume, using deep-depletion CCDs). This region has the advantage of giving less background scatter than the visible. The UV is also exploited through the use of back-illuminated CCDs (fluorescence scattering is again less than in the visible). Pulsed Raman instruments—for example, for coherent anti-Stokes Raman spectroscopy (CARS) and time-resolved vibrational spectroscopy—usually use gated intensified CCDs.

In resonant Raman spectroscopy a rejection filter is not always available for every laser wavelength, and use is still made of triple spectrometers for the suppression of stray light. These have low throughput, however, and accentuate the need for low-noise CCDs.

Particular issues for Raman spectrometers are the need for wavelength calibration relative to the laser wavelength, preserving the same illumination geometry for both samples and calibration, and checking for dark level shifts and cosmic-ray events during long exposures. These issues are discussed in Sec. 4.3.

2.2.6 Imaging Spectroscopy

Imaging spectroscopy uses 2D detector arrays to form an image with 1D spatial information in one (usually the line) direction and spectral information in the other. A full 2D spatial image can be formed by scanning the scene across the entrance slit. The images that are collected can be built up to form a "data-cube" as shown in Fig. 2.2.6-1. Though commonly used in airborne and space-based reconnaissance and Earth sensing, many ground-based applications and low-cost systems are emerging. For example, use of a prism/grating/prism combination allows an in-line spectrometer design that can be "bolted" onto existing solid-state cameras.[29]

Since imaging spectroscopy provides both spatial and spectral information, both aspects require consideration during calibration. It is possible for the spatial calibration to vary with wavelength and for the wavelength calibration to vary with position in the image. This is discussed further in Sec. 4.4.

In recent years there has been a significant growth in the use of arrays in imaging spectroscopy, for example, in laser confocal Raman[30] and Fourier transform in the IR (FT-IR)[31,32,33] spectroscopic imaging microscopy. In the confocal Raman technique the laser beam is spread out in one direction to form a line source which is then scanned across the target (the confocal design allows light scattered from the volume above or below the focal plane to be attenuated, thus giving optical sectioning). Raman spectroscopy can also be performed by spreading out the laser beam to globally illuminate the target and then selecting the wavelengths which fall onto the detector (either using a monochromator or with liquid crystal tunable filters

(LCTFs)). At present, most systems use CCD detectors and calibration requirements are similar to those for CCD cameras and spectrometers.

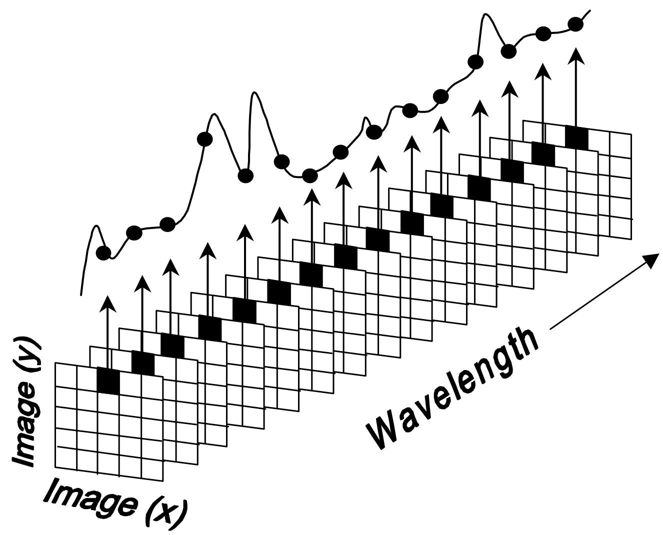


Figure 2.2.6-1 Spectral images can be thought of as a data cube having one spectral and two spatial dimensions. Spectra can be constructed for each individual pixel.

FT-IR spectroscopy conventionally uses single-element detectors, but there is a growing trend for FT-IR spectrometers to be coupled to infrared focal plane arrays. This has many applications in chemical analysis of microscopic volumes. A conventional step-scan Michelson interferometer is coupled to an IR microscope. A 2D image is formed on the detector array for each step (mirror position) of the interferometer. For each pixel in the image, the sequence of signal values (one for each interferometer step) is Fourier transformed to obtain the spectrum. FT-IR is normally carried out in the mid- or long-wave IR, and detector materials such as InSb or HgCdTe are commonly used (see Sec. 1.1). Spectroscopic imaging has also been performed in the IR using acousto-optical tunable filters (AOTFs).[34]

In the infrared, detector arrays usually show large response nonuniformities that require frequent calibration, and the issues are similar to those arising with thermal imagers (Sec. 3.1.2).

2.3 Other Applications

There are many applications of solid-state detector arrays that fall into none of the above categories. Examples are document scanners[35] and bar code readers[36]. Although ruggedness, low noise, and high dynamic range are often important, these instruments will rarely need to be calibrated.

References

[1] J. Laitinen and I. Moring, "Method for evaluation of imaging in automated visual web inspection," *Opt. Eng.*, vol. 36(8), pp. 2184-2196 (1997).

[2] I. Ashdown and R. Rykowski, "Making near-field photometry practical," 1997 IESNA annual conference, Seattle, WA, available from http://www.radimg.com.

[3] C. C. Ross and R. A. Sturz, "High speed radiometric imaging with a gated, intensified, digitally-controlled camera," *Proc. SPIE*, vol. 2869, pp. 386-394 (1997).

[4] Iqcam® Photometric Image Analysis System, http://www.iqcam.com.

[5] A. Edirisinghe, J. P. Louis and G. E. Chapman, "Radiometric calibration of multispectral airborne video systems," *Int. J. Rem. Sens.*, vol. 20(14), pp. 2855-2870 (1999).

[6] C. M. U. Neale and B. G. Crowther, "An airborne multispectral video/radiometer remote sensing system: development and calibration," *Rem. Sens. Environ.*, vol. 49(3), pp. 187-194 (1994).

[7] L. K. van Geest, F. R. Boddeke, P. W. van Dijk, A. F. Kamp, C. J. R. van der Oor and K. W. J. Stoop, "System for fluorescence lifetime imaging spectroscopy," *Proc. SPIE*, vol. 3605, pp. 55-64 (1999).

[8] J. V. Sweedler, K. L. Ratzlaff and M. B. Denton, eds., *Charge-Transfer Devices in Spectroscopy* (VCH Publishers Inc., New York, 1994).

[9] J. R. Ferraro, K. Martin and R. J. Jarnutowski, "Five year summary of commercial instrumentation for absorption spectroscopy (1992-1996)," *Spectroscopy*, vol. 12(8), pp. 18-47 (1997).

[10] MicroParts, Dortmund, Germany, http://www.microparts.de.

[11] J. T. Brownrigg, "Design and performance of a miniature dual-beam diode-array spectrometer," *Spectroscopy*, vol. 10(9), pp. 39-44 (1995).

[12] Ocean Optics Europe, Eerbeck, The Netherlands, http://www.oceanoptics.com.

[13] Carl Zeiss GmbH, http://www.zeiss.de/de/spectral/home_e.nsf.

[14] L. R. Snyder, J. J. Kirkland and J. L. Glajch, *Practical HPLC Method Development, 2nd Ed.* (John Wiley & Sons, 1997).

[15] D. G. Jones, "Photodiode array detectors in UV-VIS spectroscopy: part I," *Anal. Chem*., vol. 57(9), pp. 1057A-1073A (1985). And D. G. Jones, "Photodiode array detectors in UV-VIS spectroscopy: part II," *Anal. Chem*., vol. 57(11), pp. 1207A-1214A (1985).

[16] B. T. Owen, "Advances in UV-VIS spectrscopy: single and multicomponent quantitation," *Intl. Lab.*, pp. 24-29 (June 1987).

[17] Galactic Industries Corporation, Salem, NH, http://www.galactic.com.

[18] Good Laboratory Practice with a UV-Visible Spectroscopy System, Hewlett Packard Publication Number 12-5963-5615E (1995), also available on Hewlett Packard CD-ROM *Explore the World of UV-Visible Spectroscopy* ('98 Edition).

[19] ASTM Practice E1655-97el, "Standard practices for infrared multivariate, quantitative analysis," (ASTM, Conshohocken PA, USA).

[20] D. Ozdemir, M. Moseley and R. Williams, "Hybrid calibration models: an alternative to calibration transfer,"*Appl. Spectrosc*., vol. 52(4), pp. 599-603 (1998).

[21] D. Ozdemir, M. Moseley and R. Williams, "Effect of wavelength drift on single- and multi-instrument calibration using genetic regression," *Appl. Spectrosc*., vol. 52(9), pp. 1203-1209 (1998).

[22] J. J. Sullivan and B. D. Quimby, "Characterization of a computerized photodiode array spectrometer for gas chromatography – atomic emission spectrometry," *Anal. Chem.*, vol. 62, pp. 1034-1043 (1990).

[23] J. J. Workman Jr., "Review of process and non-invasive near-infrared and infrared spectroscopy," *Appl. Spectrosc. Rev*., vol. 34(1&2), pp. 1-89 (1999).

[24] J. Coates, "Vibrational spectroscopy: instrumentation for infrared and Raman spectroscopy," *Appl. Spectrosc. Rev*., vol. 33(4), pp. 267-425 (1998).

[25] N. Everall, H. Owen and J. Slater, "Performance analysis of an integrated process Raman analyzer using a multiplexed transmission holographic grating, CCD detection, and confocal fiber-optical sampling," *Appl. Spectrosc*., vol. 49(5), pp. 610-615 (1995).

[26] J. Zhuao and R. L. McCreery, "Multichannel FT-Raman spectroscopy: noise analysis and performance assessment," *Appl. Spectrosc*., vol. 51(11), pp. 1687-1697 (1997).

[27] T. J. Vickers and C. K. Mann, "On-line monitoring by Raman spectroscopy: instrument control and calibration," *Proc. SPIE,* vol. 2367, pp. 219-227 (1995).

[28] J. M. Tedesco and K. L. Davis, "Calibration of dispersive Raman process analyzers," *Proc. SPIE,* vol. 3537, pp. 200-212 (1998).

[29] Spectral Imaging Ltd., http://www.specim.fi.

[30] P. Dhamelincourt, J. Barbillat and M. Delhaye, "Laser confocal Raman microspectrometry," *Spectrosc. Eur.*, vol. 5/2, pp. 16-26 (1993).

[31] J. M. Chalmers, N. J. Everall, K. Hewitson, M. A. Chesters, M. Pearson, A. Grady and B. Ruzicka, "Fourier transform infrared microscopy: some advances in techniques for characterisation and structure-property elucidations of industrial material," *Analyst*, vol. 123, pp. 579-586 (1998).

[32] L. Kidder, I. W. Levin and E. N. Lewis, "Infrared spectroscopic imaging microscopy: applications to biological systems," *Proc. 11th Intl. Conf. on Fourier Transform Spectroscopy*, AIP, CP430 (1998).

[33] N. A. Wright, R. A. Crocombe, D. L. Drapcho and W. J. McCarthy, "The design and performance of a mid-infrared FT-IR spectroscopic imaging system," *Proc. 11th Intl. Conf. on Fourier Transform Spectroscopy*, AIP, CP430 (1998).

[34] E. N. Lewis and I. R. Levin, "Real-time, mid-infrared spectroscopic imaging microscopy using indium antimonide focal-plane array detection," *Appl. Spectrosc*., vol. 49(5), pp. 672-678 (1995).

[35] X. Feng, J. T. Newell and R. L. Triplett, "Noise measurement technique for document scanners," *Proc. SPIE*, vol. 2654, pp. 317-324 (1996).

[36] D. Tsi, E. Marom, J. Katz and J. Swartz, "System analysis of CCD-based bar code readers," *Appl. Opt.*, vol. 32(19), pp. 3504-3512 (1993).

CHAPTER 3

INFLUENCE OF ARRAY PARAMETERS ON INSTRUMENT CALIBRATION

One of the advantages of solid-state detector arrays is that they are intrinsically stable, both geometrically and in terms of the basic properties of the electronic building blocks (photodiodes or photogates, shift registers, amplifiers, etc.) of which they are constructed. Their stability arises because doping densities, carrier mobilities, device dimensions, etc., are not normally expected to change with time. However, it will be seen in the following discussion that several of the key performance parameters depend on operating conditions (e.g., temperature, applied voltages, clocking rates, and illumination conditions) so that unless these conditions are well defined, the device performance can vary. For this reason there is an important need for frequent calibration of array instrumentation (and for data correction) when radiometric or wavelength accuracy is required. It will also be seen that the most important parameters for the majority of applications in the UV, visible, and near-infrared can be calibrated by a simple arrangement of stable light source and optical filters, thus making the provision of calibration artifacts feasible and relatively straightforward.

3.1 Key Array Parameters

As a very general rule (though there will be some exceptions) we can identify the following *system-level* parameters as being important for users of detector array equipment:

Spectrometry
- response (including quantum efficiency, detector and off-chip electronics gain)
- response nonuniformity
- wavelength accuracy
- linearity
- stability (including stability against temperature changes)
- stray light

Imaging
- response (as above)
- response nonuniformity
- linearity
- geometric stability

As we saw in Sec. 1.2, the list of *detector-related* parameters is larger, but not all of them will affect performance at system level. A reasonably complete list of these detector parameters is given below, with a discussion in the following sections, though those having a minor influence on instrument calibration are considered only briefly. The parameters of special importance for intensified arrays are discussed in a separate section.

For some factors it may be important to calibrate routinely (sometimes even for every measurement set). For others it may only be necessary to carry out periodic checks to ensure that the system is performing to specification. The remaining parameters are important for system design (at the manufacturing level) but are unlikely to be important for system calibration.

In this guide we will only consider changes in array performance caused by changes in normal operating conditions. Degradation can also be produced by operation in hazardous environments (e.g., radiation, high-power illumination, or high temperature), but this will not be considered here.

Parameter	Importance to system performance	Dependence on operating conditions
quantum efficiency[1]	**important** - *affects absolute radiometric accuracy*	depends on temperature
on-chip amplifier response[1]	**important** - *affects absolute radiometric accuracy*	amplifier gain and offset can be affected by temperature and applied voltages
off-chip electronics response[1]	**important** - *affects absolute radiometric accuracy*	can drift with time due, e.g., to temperature and voltage changes
response nonuniformity (inter pixel)	**important** - *affects radiometric accuracy and can affect wavelength calibration accuracy in spectrometers*	varies with wavelength and can depend on illumination conditions—can get fringing effects in the NIR with thinned CCDs
response nonuniformity (intrapixel)	**can affect** wavelength calibration accuracy in spectrometers and radiometric accuracy *if* there are high spatial frequency features in the image	varies with wavelength—worse nonuniformity when illumination passes through nonuniform surface layers (e.g., in a front-illuminated CCD)
stray light	**important** - *affects radiometric accuracy*	depends on wavelength and illumination conditions
dark current and dark current nonuniformity	**important** - *affects radiometric accuracy*	varies strongly with temperature—Dark current in individual pixels can sometimes be seen to change with time
nonlinearity - of the array - of the off-chip electronics (amplifiers and ADC)	**important** - *affects radiometric accuracy*	usually stable

Parameter	Importance to system performance	Dependence on operating conditions
temporal noise	**important to check**	can change with temperature. may be influenced by external electromagnetic interference
crosstalk between pixels - due to the array architecture - due to external electronics	**can be important to check** *– affects spatial resolution and color error in color cameras*	important at design level but usually stable
smear	**can be important** - *for imaging systems*	important at design level but usually stable
lag	**important** - *for some linear and IL CCDs*	important at design level but usually stable
charge transfer efficiency in CCDs	any checks would be part of overall crosstalk or MTF tests—Radiometric accuracy can be affected if CTE is poor	important at design level but usually stable
aliasing	**can be important** - *for imaging systems*	depends on spatial frequencies in scene
geometrical stability	usually not important to calibrate	can change slightly with temperature
cosmic ray effects	**can be important** - *for long exposures*	location of the array (altitude and surrounding materials) will have a moderate effect
electrical parameters - power consumption - capacitances - output impedance - leakage currents	-	only important at design level
[1] At instrument level these parameters would normally be grouped (multiplied) together as part of an overall system responsivity.		

3.1.1 Quantum Efficiency (QE)

The quantum efficiency of a detector is the proportion of incident photons that are absorbed. Figure 3.1.1-1 shows the processes that take place. For a front-illuminated detector, photons are not normally reflected at the back surface. Because the bandgap of silicon varies with temperature, the absorption coefficient (α) and hence the QE also vary. The magnitude of the effect will depend on the detector thickness and the number of multiple reflections taking place. Figure 3.1.1-2 shows results from a model[*] that assumes a single pass through a detector of

[*] This model uses the expression: $\alpha(\lambda) = [(84.732/\lambda) - 76.417]^2$ from Ref. * (α in cm^{-1}, λ in μm), and an expression for the temperature shift from Ref. *.

25 µm thickness. It can be seen that temperature changes can give changes in QE down to ~800 nm and that at 1000 nm the changes can be 1% per degree Celcius.

Changes in absorption coefficient cannot cause QE variations much below 800 nm. However, changes can be produced at shorter wavelengths due to variations in diffusion length, reflectivity or surface recombination velocity, or, occasionally, due to instabilities in the surface structure. For example, phosphor-coated devices can sometimes show degradation of the coating at temperatures appreciably above room temperature or in humid conditions.[1] Some older types of back-illuminated CCDs rely on UV flooding or *flash gates* to improve blue/UV response.[2,3]

These back surface processes tend to give an unstable QE, but more modern devices usually rely on ion implantation and laser annealing to give a stable back-surface treatment and so have a stable QE for blue photons absorbed close to the back surface.

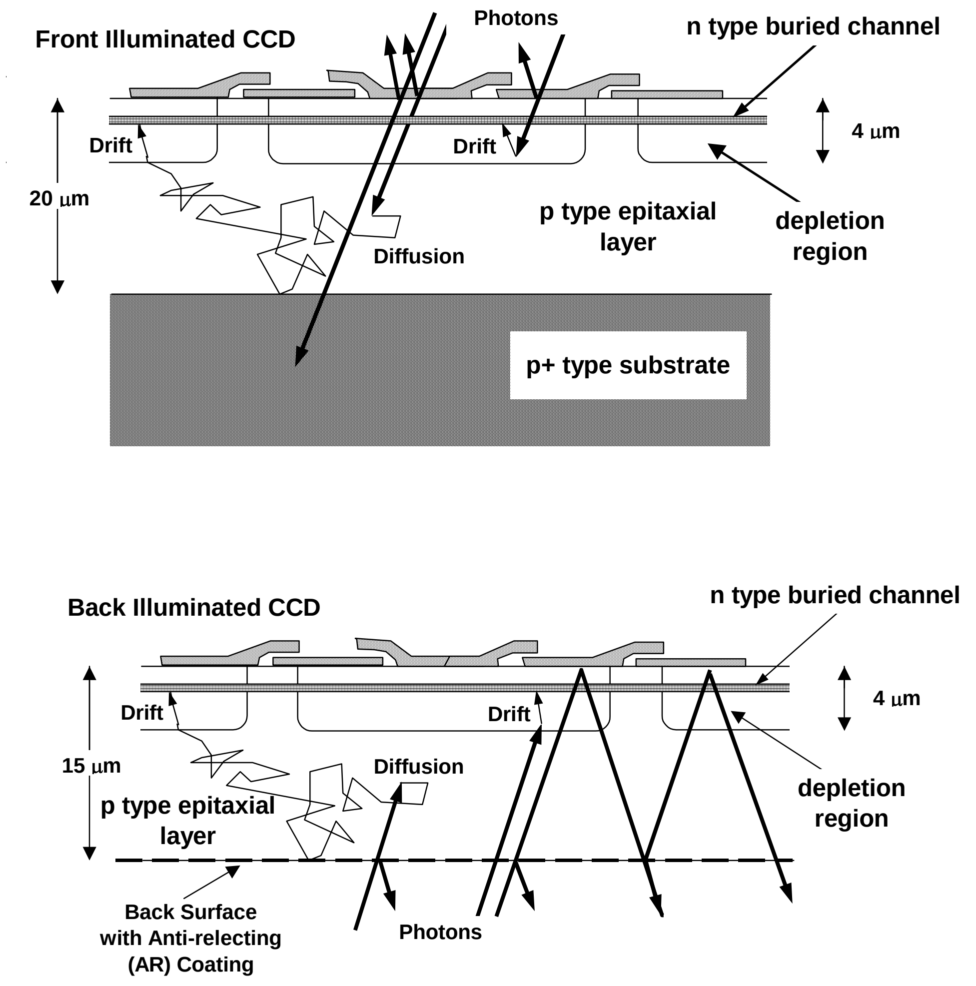


Figure 3.1.1-1 Schematic of the processes taking place when a CCD is illuminated by photons.

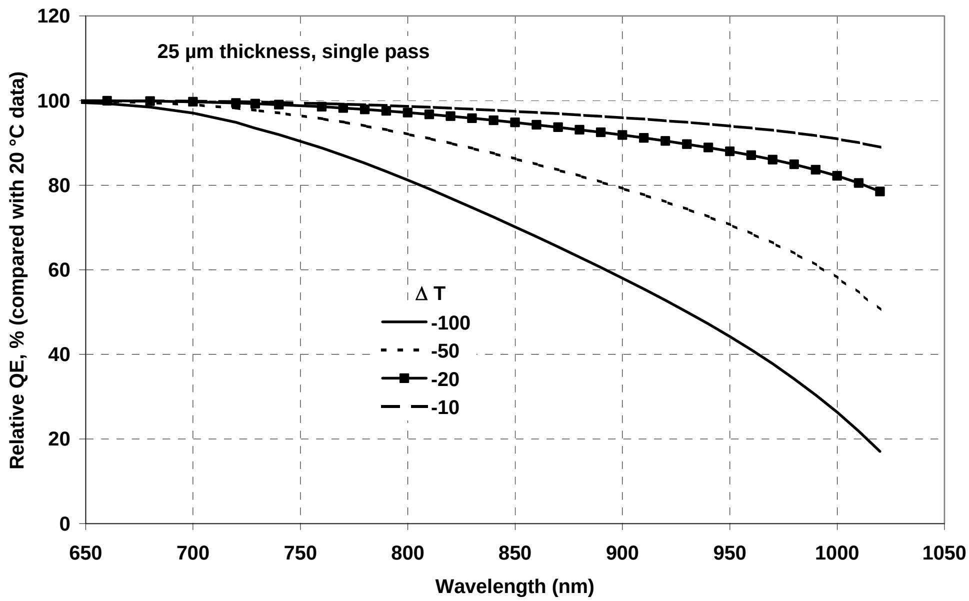


Figure 3.1.1-2 Variation in QE as a function of temperature for a CCD, assuming a single pass (no internal reflections) and a CCD thickness of 25 µm.

The absorption length increases with wavelength, and so red photons penetrate deeper into the device than green or blue. When a visible photon is absorbed it generates an electron-hole pair. (In the UV and x-ray regions a photon has enough energy to create more than one pair). The holes are repelled from the depletion region and eventually recombine. The electrons are quickly collected if the absorption takes place in the depletion region (because of the high electric field), but electrons generated in the field-free region below the depletion layer are collected by diffusion and can recombine before being collected (a process characterized by the diffusion length, *L*). They can also travel sideways to be collected in an adjacent pixel (thereby causing crosstalk). Photons can also be reflected from the front surface of a front-illuminated CCD (or a photodiode array) and at both surfaces with a back-illuminated CCD. Note that temperature-induced changes in quantum efficiency will usually be uniform across the array (with the possible exception of phosphor-coated devices).

A hysteresis effect termed residual bulk image has been reported for some types of CCDs operated at low temperatures;[3] it arises due to trapping of signal at trapping sites in the dielectric layers either at the front or the back surface. However, this is not expected to be important for most applications. A similar effect has been seen in single-element photodiodes[4] and may also occur in photodiode arrays (Hattenburg and Shumaker[5] have reported wavelength-dependent nonlinearities at the 1% level for a PDA). The effect would only occur for photons absorbed deep in the device (i.e., in the red or NIR for front-illuminated arrays). Since it is a trapping effect, once the traps are filled (for example, by a large signal), then further trapping cannot occur. This can lead to response nonlinearities since only small signals will lose charge due to trapping (see also Secs. 1.2.11 and 4.2.5).

If the detector array is operated below 0°C there is the possibility that contaminants (from dewar outgassing) can condense onto the array surface in a thin film. This will cause interference effects and will cause a wavelength-dependent change in responsivity. This is discussed further in Sec. 3.1.4.2.

In summary, the responsivity of silicon array-based instruments should be checked if the operating temperature is variable, particularly if the detector is used at wavelengths close to 1000 nm, and particular care needs to be taken with cooled CCDs so that the surface does not become contaminated. With other detector materials, similar changes in responsivity can be caused for photon energies close to the bandgap. With some devices the blue/UV sensitivity can be variable.

3.1.2 On-Chip Amplifier Response

In a solid-state detector array, the photo-induced current is either directly sensed and converted to a voltage by the output amplifier (as in the case of some photodiode arrays and CMOS sensors) or is integrated within the pixel to form a signal charge that is later converted to a voltage when transferred to the capacitance of the output amplifier. The capacitances involved are essentially fixed by the architecture of the device and are stable in time (though in some cases they will vary slightly with signal size, thus producing a nonlinearity). However, the gain and the offset of the amplifier can depend on temperature and the applied bias voltages. The sensitivity to these conditions is dependent on the amplifier design and the voltages applied. Shifts in offset voltage can be significant (of the order of 100 μV/K) and shifts in gain ~ 0.1%/K.

Shifts in offset voltage can also be caused by changes in the clocking waveform. Figure 3.1.2-1 shows the output waveform of a CCD. An offset can be produced by capacitative feedthrough of the readout register (R) clock onto the video waveform. The particular phase concerned is the one associated with the electrode nearest the output (when this phase goes low, charge is transferred into the output node). Thus, any change in the R clock waveform caused, for example, by changes in the clock frequency or in pixel binning will cause a change in offset. Changes of the order of 100 μV can result from this.

3.1.3 Off-Chip Electronics Gain

Stability of the off-chip signal processing chain (amplifiers' sample and hold and analog-to-digital converters, ADCs) can vary with time and operating temperature, either directly or via changes in the amplifier bandwidth. However, the changes can be expected to be small, of the order of a few parts per million per degree Celcius temperature change.

When CCD video cameras with an analog output are used in conjunction with a framegrabber, then thermal effects can alter the synchronization between the camera and the framegrabber clocks. This can lead to geometrical distortions in the image, which are noticeable when geometrical accuracy is important (e.g., in photogrammetric applications); and warm-up effects have been reported.[6, 7] There is a trend to use cameras either with a digital output (ADC in the camera itself and not on the framegrabber) or with pixel clock outputs (for synchronization of an external ADC); hence, these problems are less likely to occur in the future.

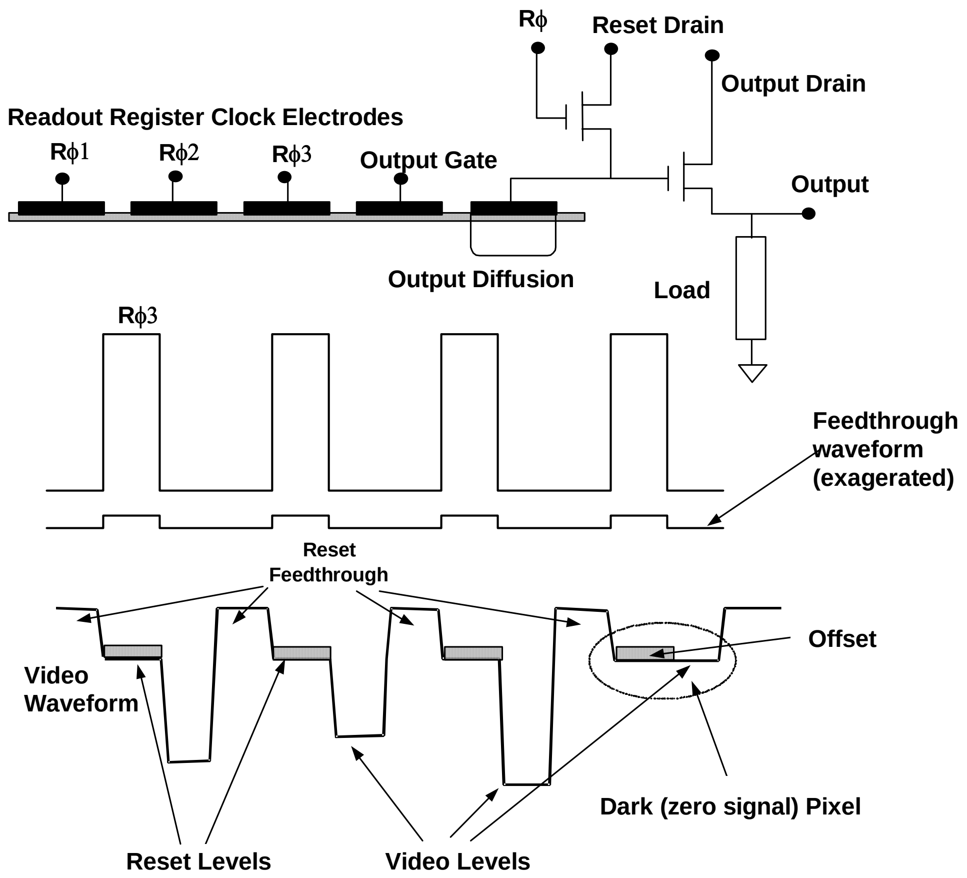


Figure 3.1.2-1 An offset voltage can be produced in a CCD readout register by feedthrough of the RΦ3 clock. The signal measured by the off-chip electronics is the difference between the reset and video levels. The addition of the offset results in a dark pixel with no true signal presenting a spurious signal (which can be positive or negative, depending on the polarity of the feedthrough waveform) to the off-chip electronics.

3.1.4 Response Nonuniformity (Interpixel)

The responsivity to illumination will vary from pixel to pixel, for example, due to differences in pixel area or the thickness of surface layers. These factors can usually be expected to be stable in time since the architecture of the device will not change. Such responsivity variations are usually of the order of a few percent for CCDs and can be <1% for photodiode arrays. (Reference 8 gives example data for a 1024-element PDA; in this case, the authors comment that the pixel responsivities needed to be calibrated every few months, particularly when the array was less than a year old.) Note that these small pixel-to-pixel variations in responsivity due to pixel size or surface structure can depend (in a CCD) on the electrode phases used for integration. In a four-phase CCD an improvement in response nonuniformity can be gained by integrating under three phases compared with one phase. However, this is at the expense of vertical resolution. (With three-phase integration the signal is blurred over two pixels in the vertical direction.)

Larger deviations can be produced in some pixels by spurious deposits formed during manufacture (e.g., stains or metallization deposits) or by dust either on the surface of the array or on the window (inside or outside surfaces). The effects of dust particles will depend on the cone angle of the illumination. A narrow cone angle gives a more pronounced shadow and greater contrast between the defect pixel (or pixels) and its (their) neighbors. Although array manufacturers take appreciable care to reduce particulate contamination, this will occasionally occur. Adherence of dust to the outside of the detector window will be more common but will tend to give an out-of-focus, low-contrast region of lower response (often in the form of a halo). The effect of cone angle on the appearance of dust is illustrated in the images shown below (Fig. 3.1.4-1).

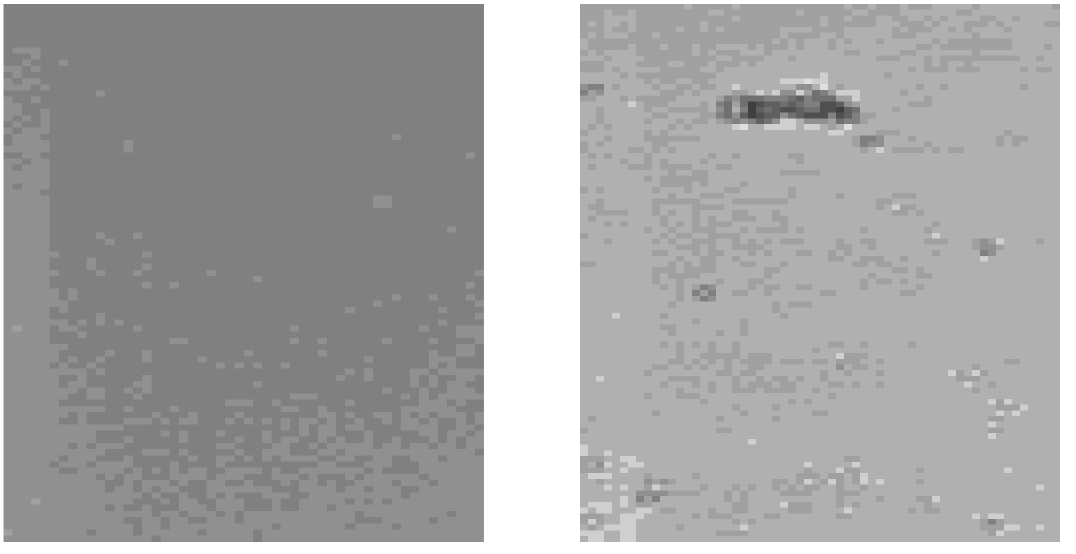


Figure 3.1.4-1 Flat-field CCD images obtained with a large cone angle using a diffuser (left) and with near parallel illumination (right). Shadows from dust on the array window are only noticeable for small cone angles.

3.1.4.1 Traps

Dark pixel defects can also be produced in CCDs by charge traps. In an area CCD, charge has to be transferred down the columns to the readout register and then along the readout register to the output node. For good performance this transfer has to be nearly perfect. If there is an obstruction in a column that impedes charge transfer, then a partial dark column results, as shown in Fig. 3.1.4-2. Lesser obstructions lead to defects called *traps*. These can be caused by defects in the silicon used to fabricate the CCD or by defects introduced during manufacturing; occasionally a design fault in the layout of the masks used in manufacturing leads to a "potential pocket" that can cause trapping, but this is uncommon with present day devices.

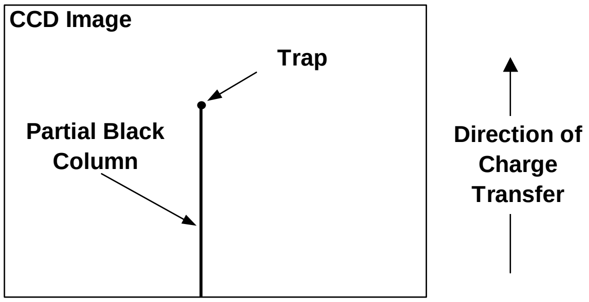


Figure 3.1.4-2 Generation of a dark column response defect by a large trap or column obstruction.

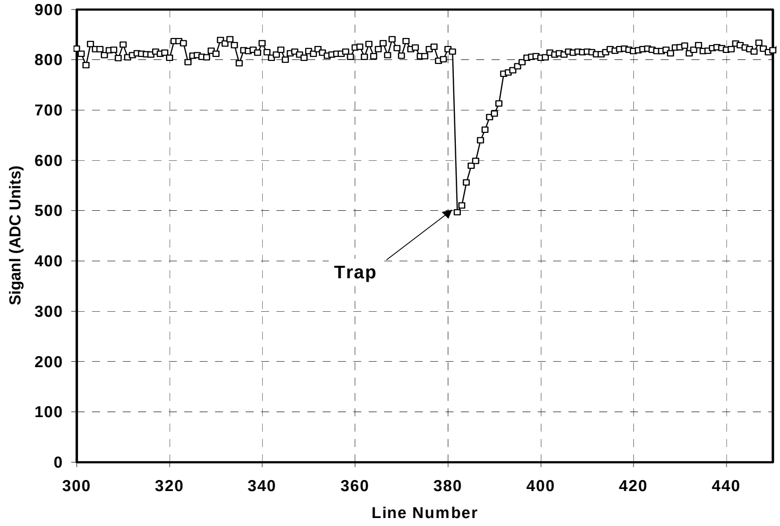


Figure 3.1.4-3 Vertical slice (in the column direction) through a high-level trap.

There are two basic types of traps that can be classified as low-level or high-level. Low-level trap defects absorb a small amount of the charge that passes through the pixel.[9] At low signal levels, not enough charge is transferred by each pixel to fill the trap, and the defect appears as a string of dark pixels following the trap. However, once the charge from several pixels has added up to fill the trap, then charge can pass through successfully and the remainder of the column gives a good image. As the signal level increases, the number of pixels that are affected by the trap decreases, and at high light levels many traps disappear. It can occasionally happen that a trap that gives a dark trail at low signal levels gives a bright one (rather than no trail at all) at high levels.

The other type of trap (the high-level trap) shows the opposite behavior and only appears at high light levels. These high-level traps are the equivalent of a constriction in the charge path through a pixel. Small amounts of charge can pass through these traps unhindered, but signals above a certain level cannot be transferred. Figure 3.1.4-3 shows a vertical slice (in the column direction) through such a trap. In this case, signals below 500 ADC units (= 30,000 electrons) show no effects, but at higher signals a string of dark pixels occurs as for the low-level trap. At higher signals still, the trap can sometimes become filled (as for the low-level trap) and disappear again.

Although trap defects are rare in high-grade CCDs, a large-area device may well have two or three. Clearly, their effect depends on the signal level, so, in effect, there is a signal-dependent response nonuniformity which can at times be difficult to detect. Even in a prestigious project, such as the CCD imaging camera on NASA's *Galileo Jupiter* spacecraft, it was found by detailed calibration after launch that there were traps present in the serial register that had been previously undetected.[10] (It has to be said, though, that in that particular case, the CCDs used were not state-of-the-art by today's standards, and the traps caused only a small effect on image quality).

Some, but not all, CCD manufacturers now include traps (however, usually only high-level ones) as pixel defects in their grading of cosmetic image quality.

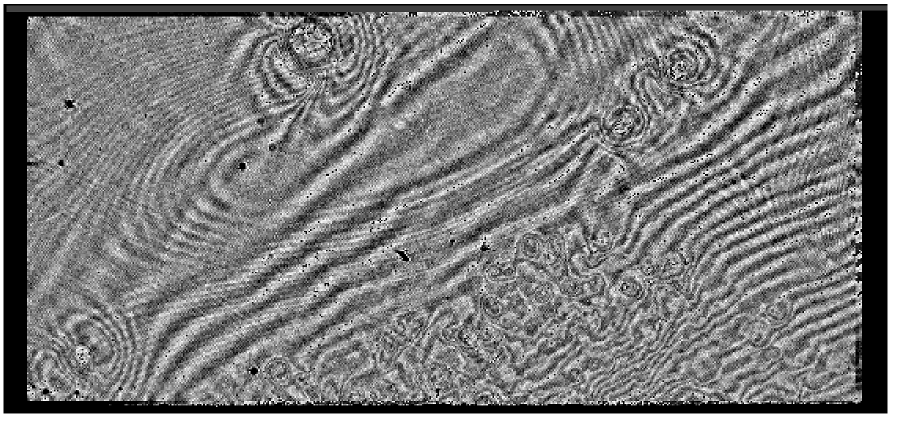
3.1.4.2 Fringe Effects

There is an effect that occurs with thinned back-illuminated detector arrays when illuminated in the red or near-infrared. At these wavelengths the absorption depth is comparable with the detector thickness (typically 10–15 μm) and the silicon effectively becomes semi-transparent. The front and back surfaces are nearly parallel. An optical cavity is formed, and multiple reflections between the two surfaces produce interference fringes. Fringing effects are classified in this guide as an interpixel phenomenon since there are variations between pixels, but fine-scale fluctuations will also cause effects within a pixel (and cause changes in the intrapixel response). Interference effects can also occur in front-illuminated CCDs and photodiode arrays due to multiple reflections in the gate oxide layer. However, these are less severe, being uniform across the array and, in effect, can be considered as producing ripples on the spectral response curve of a typical pixel. In systems where the CCD is cooled to low temperatures there is the risk that contaminants can condense onto the CCD surface to form a thin film. This will cause changes to the interference pattern; an example (for a front-illuminated diode array) is given at the end of this section.

For a back-illuminated CCD of 16 μm thickness, the optical thickness (t_{opt}) is ~60 μm, since the refractive index is ~ 3.7.[11] There will be constructive interference when

$$n\lambda = 2\, t_{opt} = 120\ \mu\text{m}. \tag{3.1.4-1}$$

For example, there will be fringe maxima at 800 nm (n = 150) and at 805 nm (n = 149), and the interference pattern will repeat at intervals of about 5 nm (the fringe spacing = $\lambda^2/2\, t_{opt}$). This causes two types of instrumental effects: one spatial, the other spectral. The spatial effect occurs when imaging with near monochromatic light. The thickness of the CCD will not be constant over the whole device. A change in mechanical thickness of 1/16 λ (or 0.05 μm at 800 nm) will change the interference from constructive to destructive. Figure 3.1.4-4 shows a flat-field image for a typical thinned CCD illuminated by a tungsten lamp at 900 nm; the bandwidth was ~10 nm. Also shown is a horizontal slice across the image, showing modulation of ~5% (a fairly typical value, though the fringes would have higher contrast if the bandwidth was smaller). In many imaging applications this effect will not cause problems. Firstly, because multiple reflection only occurs at long wavelengths (where the absorption coefficient is comparable with the detector thickness), and secondly, because the spectral bandwidth will often be large enough (>10 nm) to span several fringe cycles.



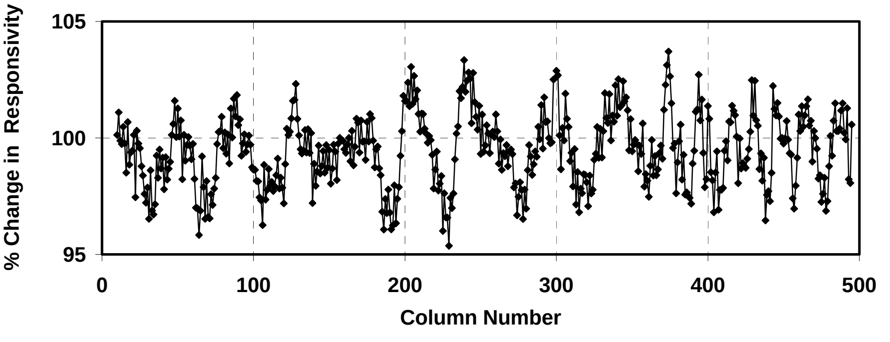


Figure 3.1.4-4 A flat-field image for a typical thinned CCD illuminated by a tungsten lamp at 900 nm; the bandwidth was ~10 nm.

In contrast, spectroscopic instruments will usually have a narrow bandpass (<10 nm) falling on a given pixel. Even if the detector were to have a perfectly uniform thickness, the variation in wavelength across the array would produce fringes with a period of 5 nm (in the example given above). This is the second instrumental effect (spectral effect). In a spectrometer, both spatial and spectral effects will occur together. Jorden et al.[12] give a good example of fringe effects in the ISIS grating spectrograph for the William Herschel Telescope (WHT). They found fringe amplitudes ranging from 2% at 600 nm to 60% at 900 nm with their thinned CCD (see also Ref. 13). Another example is given below. Figure 3.1.4-5 shows results from a thinned back-illuminated CCD used in an imaging spectrometer. The plot shows the output from pixels in the across slit (dispersion direction) as the input illumination to the imaging spectrometer is scanned in wavelength (the input illumination was itself from a monochromator). The asymmetric peaks arise from fringing effects as demonstrated in Fig. 3.1.4-6, which shows the normalized output for summed pixels. In this case the modulation is roughly 20% and the fringe spacing is about 6 nm (the CCD was ~15 μm thick). In general, these fringing effects can be expected to cause wavelength errors in the wavelength calibration of the spectrometer, although Fig. 3.1.4-7 shows that in this case the wavelength errors are small. The plot was obtained by finding the center of gravity of the center pixel and the three pixels on each side of it (in the dispersion direction) using the algorithm

$$\text{Peak position (pixels)} = \frac{-3\times A - 2\times B - C + 0\times D + 1\times E + 2\times F + 3\times G}{A + B + C + D + E + F + G}, \tag{3.1.4-2}$$

where A, B…G are the signals in each pixel. The calculation was performed for each input center wavelength (derived from the input monochromator grating position) for three center pixels (each with signal D in the above equation) labeled 1, 2, and 3 in Fig. 3.1.4-5. As expected, the wavelength corresponding to the center of a pixel shifts by exactly the same amount when the center pixel is changed from 1 to 2 to 3. It can be seen that the plots are essentially straight lines, and the wavelength error is negligible in this case. However, if the center wavelength positions were obtained by a quadratic fit to the three brightest pixels (the four brightest if the peak fell close to the boundary between two pixels), then much-larger wavelength errors were produced. This is probably a worst case for wavelength errors since the linewidth was of the same order as the pixel dimension and using a quadratic fit is not optimal, but it illustrates the impact of center wavelength algorithm on the wavelength calibration error on this type of spectrometer.

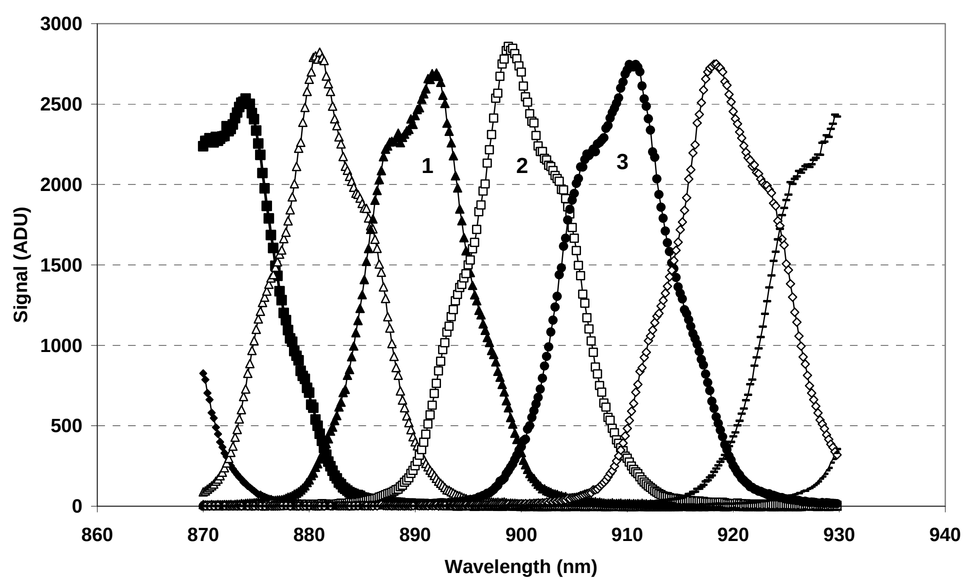


Figure 3.1.4-5 Results from a thinned back-illuminated CCD used in an imaging spectrometer. The plot shows the output from pixels in the across slit (dispersion direction) as the input illumination to the imaging spectrometer is scanned in wavelength.

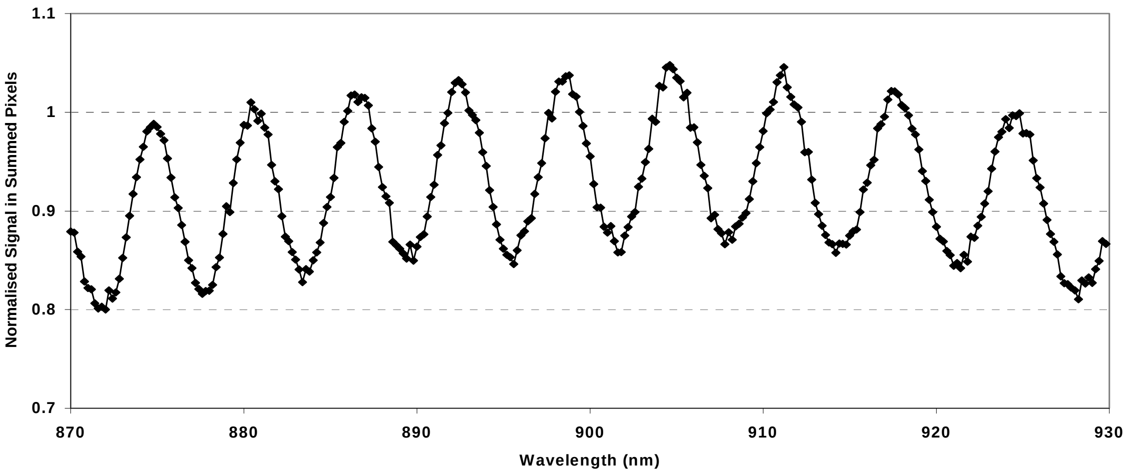


Figure 3.1.4-6 Normalized output for summed pixels (derived from the data of Fig. 3.1.4-5).

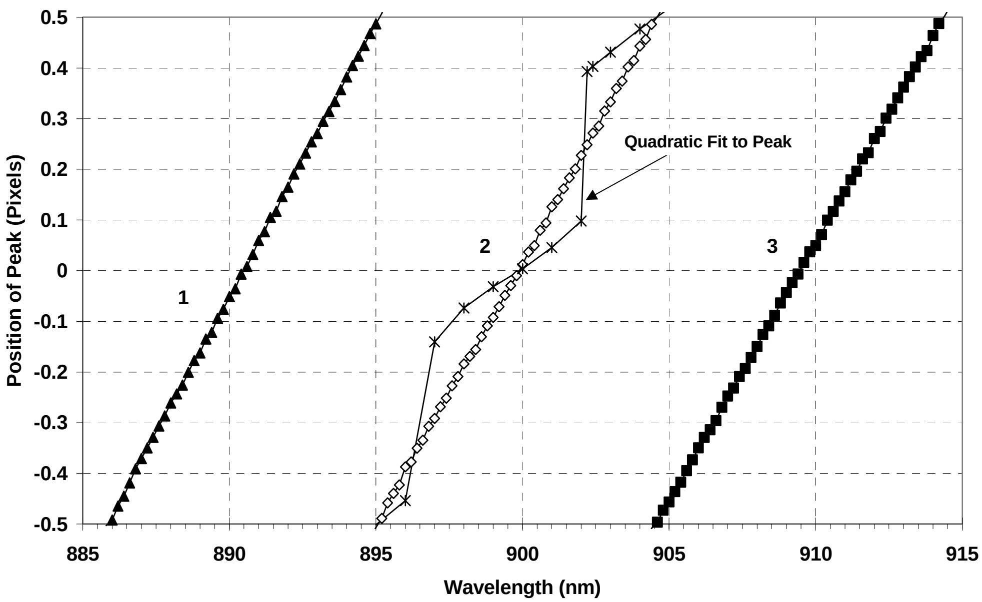


Figure 3.1.4-7 Center wavelength position derived from the imaging spectrometer (one pixel corresponds to roughly 9.5 nm) derived using the center-of-gravity method versus the input wavelength (for the three pixels labeled 1, 2, and 3 in Fig. 3.1.4-5 taken as the center pixel).

Wavelength calibration of spectrometers is discussed in detail in Sec. 4.2.1. Normally the calibration is performed using a fixed wavelength feature, such as an emission or absorption line. Fringing (and other) nonuniformity effects may then not be immediately apparent—though they will often still be present and can ultimately limit the accuracy of wavelength calibration.

Although the fringe pattern is often found to be stable in time, changes in optical configuration (e.g. grating shifts) or thermal effects will result in variations. It is therefore imperative to perform frequent flat-field calibrations when fringing effects are important. The normal procedure is to calibrate the responsivity against a standard continuum source (such as a tungsten lamp). If the CCD is used in a mode where pixels are binned together (as is often the case in the along-slit direction, to improve sensitivity), then a calibration will be needed for each binning configuration (the binning will tend to average out the spatial effect over several pixels).

Although it can be seen from the above that the fringing phenomenon (sometimes called etaloning) can be pronounced in the NIR, there are ways of reducing the effect. For example, optimization of the antireflective (AR) coating on the back of the CCD will reduce the reflection from that surface. Often the AR coating cannot be made perfect for near-IR wavelengths since good quantum efficiency is needed in the UV, but there is scope for tailoring the AR coating for specific applications. Use of a thicker (deep-depletion) device will also reduce the effect. Finally, there are other (proprietary) modifications to the internal device structure that are being developed by CCD manufacturers. Hence, it is likely that spectroscopic detector arrays will be less prone to fringing effects in the future. It might be thought that using a low F-number (large cone angle) might reduce the effect due to variations in the path length through the CCD for the arrays traveling at different angles. However, the high refractive index of silicon (~4) refracts the converging beam to one that is nearly parallel, thus reducing the path length differences to much less than a wavelength.

Mount et al.[13] discuss an example where outgassing from materials inside a vacuum dewar resulted in a thin film being deposited on the passivating surface of a photodiode array. This caused changes in the fringe pattern arising from the cavity formed by the front surface of the array and the back surface of the detector window and resulted in drifts in the array responsivity (as a function of wavelength). With a CCD temperature of –84°C the amplitude modulation of the fringe pattern was initially observed to drift at a rate of several percent per 0.5 hr interval. This rate decreased with time (as the rate of outgassing decreased) and was reduced at higher temperatures. The solution to the problem was to use a tilted window (the window itself being wedged so as to eliminate interference between its front and back surfaces).

3.1.4.3 Microlens Arrays

A way of improving sensitivity when the fill factor of the sensor array is small is to use a microlens array (which can be formed on top of the sensor as part of the manufacturing process). For example, many ILT (and FIT) CCDs and some CMOS sensors incorporate such arrays. However, this leads to a response nonuniformity that depends on the cone angle of the illumination. This has been discussed by Theuwissen (Sec. 7.2.1 of Ref. 14) and is illustrated in Fig. 3.1.4-5. For off-axis pixels, light is no longer fully focused onto the sensitive area of the pixel, the effect being worse for low F-numbers. This vignetting effect has been experimentally demonstrated by Penkthman[15] and Putnam et al.[16] in ILT CCD cameras. In the latter case the photodiodes were narrower than they were tall, and so the effect was more pronounced in the horizontal direction. Note that the amount of vignetting will also depend on the aperture of the lens.

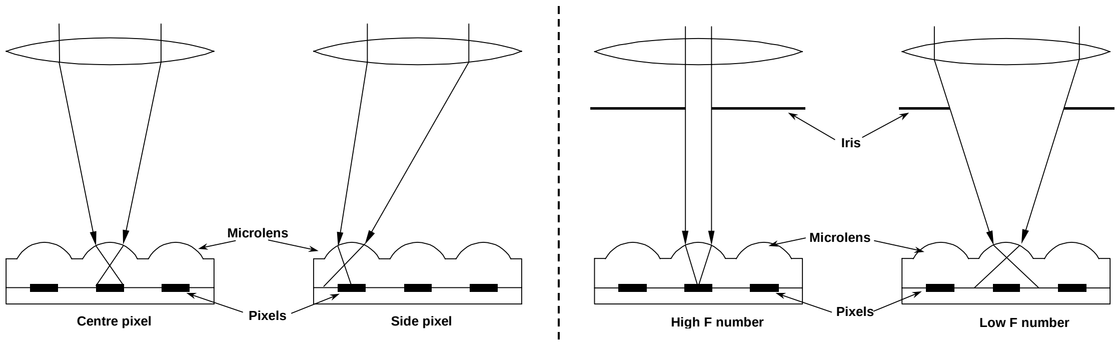


Figure 3.1.4-5 Illustration of the effect of illumination conditions on the sensitivity of detectors incorporating a microlens array (after Ref. 16).

3.1.5 Response Nonuniformity (Intrapixel)

Ideally, each pixel in an imaging array will have a flat “top-hat” response. In practice, this is often far from the case, particularly for arrays that have nonuniform surface structures such as front-illuminated CCDs. Photons will be preferentially absorbed in those surface regions that are most opaque (for example, the region where electrodes overlap), leading to a nonuniform response. Hence, both the signal in an individual pixel and the total signal (summed over adjacent pixels) can vary depending on the position of the image (or spectral line) relative to the pixel.

The response will also vary with wavelength. Long-wavelength photons (with a higher absorption length that can penetrate deeper into the device) will be less affected by the surface nonuniformities but will produce charge that tends to spread sideways in the field-free region below the depletion region (Fig. 3.1.1-1), and so produces more crosstalk between adjacent pixels. The effect of absorption in the electrode structure will be most important at short wavelengths. The effect is particularly important for open-electrode (OE) CCDs used in the UV. Figures 3.1.5-1 to 3.1.5-5 illustrate the various effects caused by absorption in the surface layers of a front-illuminated CCD. Back-illuminated CCDs will show a wavelength-dependent crosstalk but without the strong modulation in response across the pixel (since there is no transmission through the electrode structure). There will, however, be some modulation due to the geometry of the depletion region.[17] For diode arrays, the pixel response is often quoted as having a trapezoidal shape;[18] but, as for a front-illuminated CCD, surface structure will affect response in the blue and increased crosstalk is apparent in the red. CCDs and PDAs used for scientific imaging in the visible will normally have negligible (or at least small) dead spaces between pixels. This is not the case for some commercial CCD cameras (e.g., those using interline transfer arrays) or in the infrared (where technological restrictions often result in a reduced fill factor). Pronounced dead spaces between pixels will clearly have an effect on the relative response for the signals focused on the center and the edge of a pixel. As discussed in Sec. 3.1.4.3 above, the use of microlenses can lead to response nonuniformities.

Subpixel response (sometimes termed the *pixel aperture response*) has been discussed by Jorden et al.,[17] Kavaldjiev et al.,[19] and Lind et al.[20]. The data presented here is from a contract report prepared for the European Space Agency.[21] Figures 3.1.5-1 and 3.1.5-2 show scans across the CCD electrodes for four- and three-phase devices in the green (550 nm). The modulation in the pixel response is ~50% at this wavelength. For the four-phase device the pixel structure is symmetrical, so that a spot image (or a spectral line) falling on the center of a pixel gives the same (summed) signal as one split between two pixels. This is not the case with a three-phase device. Figure 3.1.5-2 also shows that, as expected, the total signal does not depend on which electrodes are used for integration (n.b., the phases used in the figure: either $\phi 1$ or $\phi 1$ plus 2, correspond to the two interlace fields in a TV image). Figure 3.1.5-3 explains how the pixel profiles arise from the overlapped electrode structure, and Fig. 3.1.5-5 shows that there can also be changes in response for scans across the channel stops at short wavelengths. At 850 nm most of the photons are absorbed below the electrode structure and so the response profiles are smooth Gaussian-like curves. The three-phase device still showed a lower summed response when the signal is split between two pixels, whereas the four-phase device showed an essentially flat response at this wavelength. Note that intrapixel response is usually very similar for all pixels in an array.

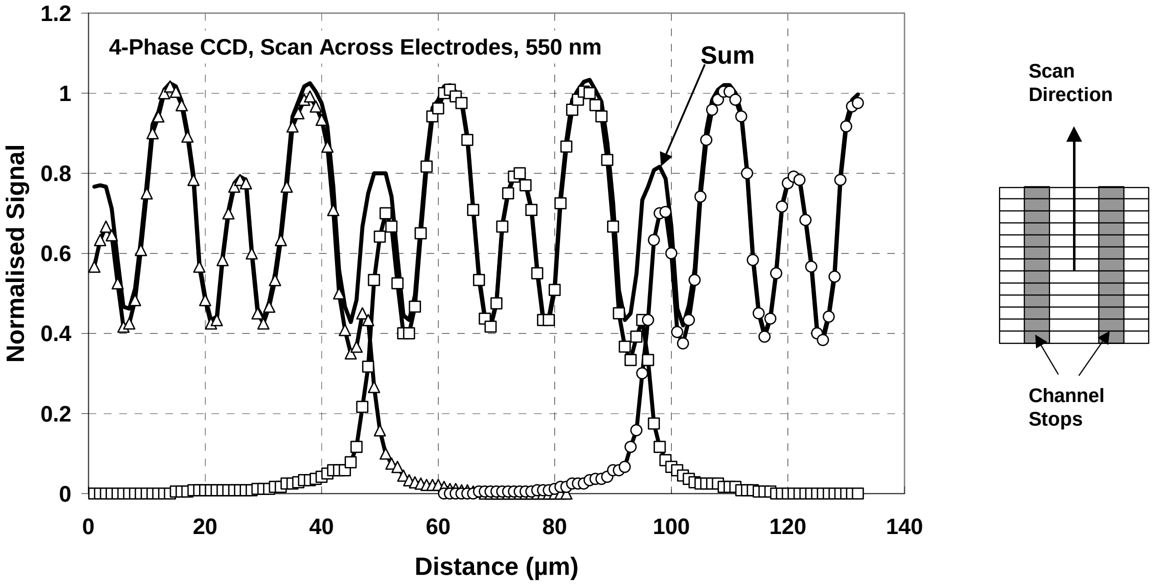


Figure 3.1.5-1 Across-electrode (column direction) scan for a four-phase CCD at 550 nm. The sum of signals in adjacent pixels is shown by the thick line. In this case the electrode structure is symmetrical and the summed signal is the same for spot illumination centered on a pixel as it is for centered on the boundary between two pixels.

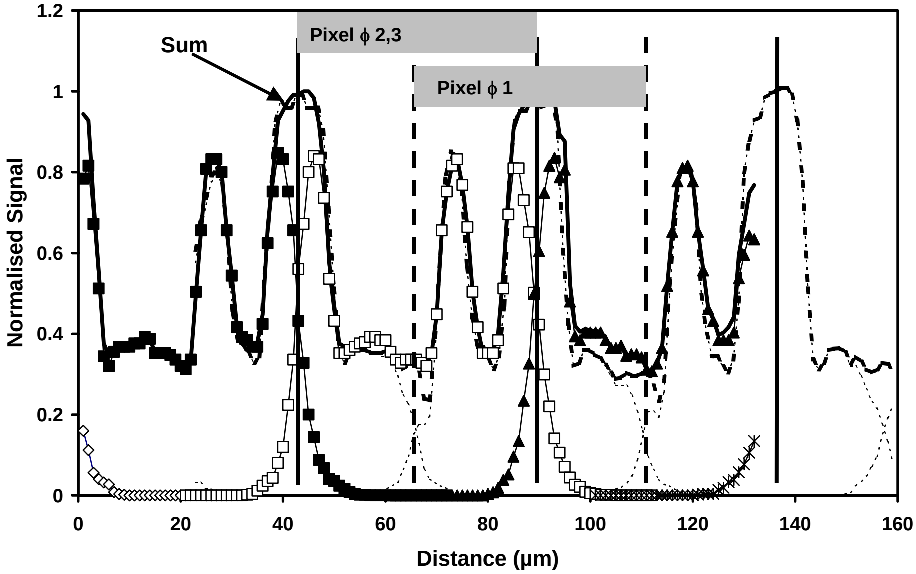


Figure 3.1.5-2 Across-electrode (column direction) scan for a three-phase CCD at 550 nm. The sum of signals in adjacent pixels is shown in thick. In this case the electrode structure is not symmetrical and the summed signal is not the same for spot illumination centered on a pixel as it is when centered on the boundary between two pixels. Full lines are for integration under phases 2 and 3 and dotted lines for phase 1 high during integration. The effective pixel boundaries are shown as thick vertical lines. The summed response is the same (within measurement error) in the two cases.

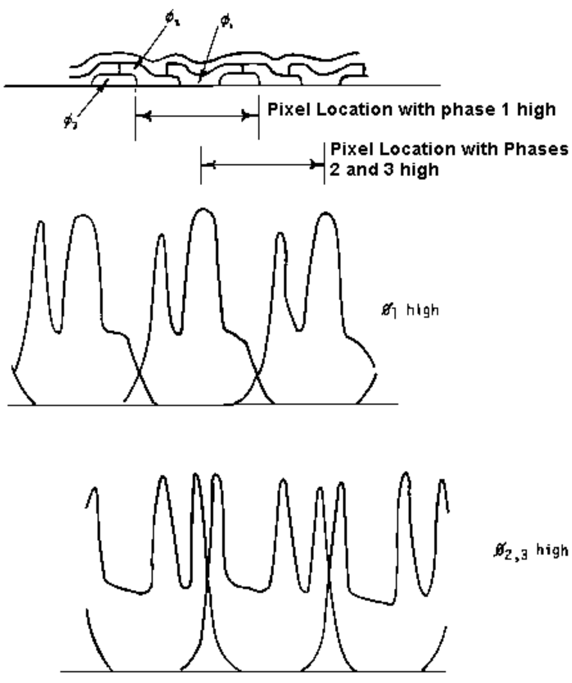


Figure 3.1.5-3 Across-electrode scans at 550 nm for a three-phase front-illuminated CCD for phase 1 high during integration and for phases 2 and three high during integration, and an idealized diagram of the electrode structure. More light is absorbed in the regions where the electrodes overlap.

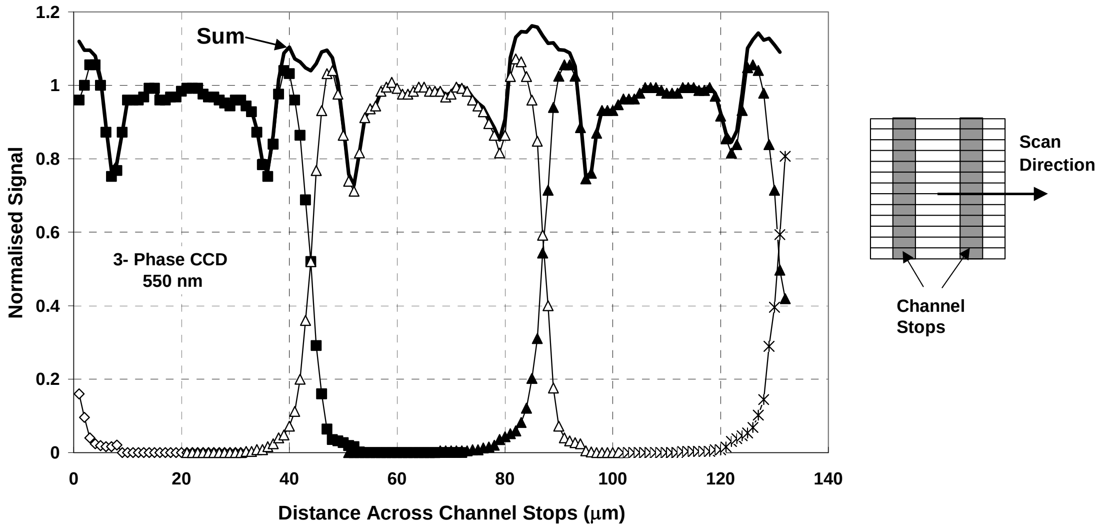


Figure 3.1.5-4 Scan across the channel stops of a three-phase front-illuminated CCD at 550 nm.

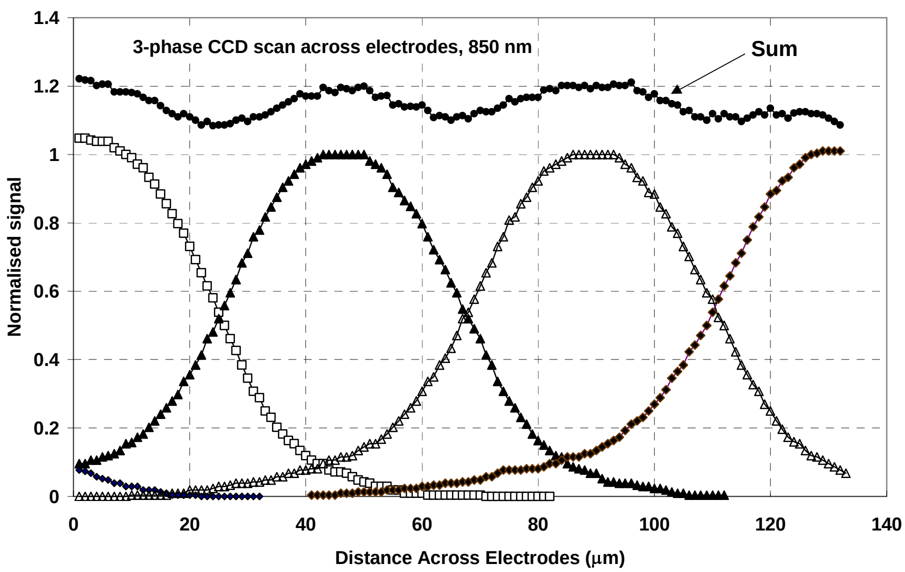


Figure 3.1.5-5 Scan across the electrodes of a three-phase front-illuminated CCD at 850 nm. The scan across the channel stops had a similar appearance. For the four-phase device the summed scan was flat at this wavelength.

The presence of nonuniformity within a pixel has two effects:

1. The "center of gravity" of the pixel response is not necessarily identical to the geometric center of the pixel, and is wavelength dependent.
2. Features in the image that have finer detail than the pixel dimension, *L*, (i.e., have higher spatial frequency than *I*/*L*) cannot be resolved spatially but will interact with

the pixel profile to give an overall signal that depends on the exact registration of the pixel with the fine structure in the image. This will result in calibration errors.

In most spectroscopic applications the entrance slit width is such that the bandwidth of spectral features (either emission or absorption lines) used for wavelength calibration are broader than the resolution (the bandpass associated with each pixel or the spectral sampling distance)—typically this is a few tenths of 1 nm. These effects will then not be of major importance, likewise for most imaging applications. However, the presence of response nonuniformity within each pixel should always be borne in mind when accurate calibrations are being considered. Unfortunately, manufacturers will rarely be able to supply details of pixel response since this is not normally measured.

3.1.6 Stray Light

There are particular problems with stray light with array-based instruments because of the possibility of light reflected off the surface of the array being directed back to the detector by reflection off either the array window or nearby optical components. The effect tends to be more noticeable in back-illuminated CCDs and in photodiode arrays since the refractive index of silicon is large (~4). With front-illuminated CCDs the reflection from the electrodes on the front surface is usually less.

An antireflection coating on the back surface will reduce the reflection at the design wavelength, λ (where the thickness is $\lambda/4$ and reflection is minimized), but for other wavelengths the reflection coefficient can be large.

The stray light will depend on the cone angle and on the curvature of nearby reflecting surfaces. In a poorly designed optical system the stray light can be focused back onto the array to give a ghost image. The effect on the measurement will be most pronounced when the stray light is of a wavelength for which the detector responsivity is large and it falls on a part of the image (or spectrum, in the case of a spectrometer) where the signal is normally low. For these reasons the measurement of stray light is an important part of the calibration procedure for array-based instruments, and is discussed in detail for spectrometers in Sec. 4.2.4.

3.1.7 Dark Signal and Dark Signal Nonuniformity

As implied by the name, the dark signal is the signal measured in each pixel of the array when the instrument is in darkness. This signal can arise from several causes (see also Sec. 1.2.3):

electronic offsets, usually termed fixed pattern noise
thermally generated dark current
white pixels, or columns
clock-induced dark current

Fixed pattern noise is caused by electrical feedthrough of clocking or switching signals onto the output video waveform. It is usually negligible in CCDs but can be noticeable in CMOS detectors, charge injection devices, and some photodiode arrays. In a sense, the Rϕ3 feedthrough that causes an offset voltage in a CCD (see Sec. 1.2) can be thought of as a fixed pattern noise; however, it is the same for every pixel. Usually the fixed pattern noise will give an uneven response for entire columns or rows (or both). Some photodiode arrays use a different readout register and output amplifier for odd- and even-numbered diodes. Any difference in offset in the two amplifiers will give rise to an "odd/even" effect in the image.

For silicon devices the thermal dark signal (D) varies with absolute temperature (T) as[3]

$$D = \text{constant}\, T^{\alpha} \exp\left(\frac{-Eg}{2kT}\right), \tag{3.1.7-1}$$

where the bandgap $Eg = 1.1157 - 7.021\times10^{-4}T^2/1108 + T$, and k is Bolzmann's constant (8.602×10^{-5} eV). The exponent α is usually quoted as 1.5, but in fact its value depends on the source of the dark current, being 2 for bulk dark current and 3 for surface dark current. A simpler expression, which gives a good approximation to Eq. (3.1.7-1), is

$$\text{dark signal} = \text{constant}\ \exp(-E_{act}/kT)\ , \tag{3.1.7-2}$$

where the activation energy E_{act} ~ 0.63 eV. At room temperature the dark current doubles for every ~8°C increase in temperature, and at −40°C for every 5°C increase. For other detector materials, for example, those used in the infrared (HgCdTe, Insb, InGaAs, etc.), the change in dark signal with temperature can be even larger.

Because of the large change with temperature it is important to check the dark signal for each set of measurements, particularly as the thermal dark signal varies from pixel to pixel. Ideally, the same integration time should be used for dark images as for those under illumination (see Sec. 1.2 for a further discussion of the effect of integration time and Sec. 4.2.6 for methods for measuring dark signal in spectrometers).

Even if a detector array in an instrument is cooled (e.g., using a thermoelectric cooler) this is not necessarily a guarantee that the array temperature is stable. For example, a thermoelectric cooler may be run at constant current (or voltage) so as to give a constant ΔT relative to the heat sink temperature, which may itself vary depending on the method of cooling it (usually via forced air or circulating water). Also, the heat dissipated in the CCD will vary with the clock sequence being used. Fast clocking (for example, to dump unwanted parts of the image) can result in extra heat loads on the order of 100 mW in an area CCD. This can be sufficient to change the CCD temperature by of the order of 1°C.

The dark signal nonuniformity, expressed as a percentage of the average value, is highest for low dark current silicon devices, where the surface component of the dark current is suppressed (c.f. Sec. 1.1.4.5) and only bulk dark current generation occurs.

This bulk dark current arises mainly from individual crystal defects and varies widely from pixel to pixel. Those pixels having the largest dark signal are often termed dark current spikes or "hot" pixels. The dark current nonuniformity will usually vary with temperature because the larger thermal dark signals arise from a mechanism called field-enhanced emission.[22,23] In this, the bulk dark current is enhanced because the crystalline defect occurs in a region of high electric field. A characteristic of this effect is a reduced value of the activation energy (see Eq. (3.1.7-2)) from a value of ~0.63 eV to values as low as 0.3 eV. Figure 3.1.7-1 shows the effect of this change in E_{act} on the way the dark current changes with temperature. With a lower E_{act} the dark current does not reduce as quickly with temperature, so that dark current spikes "cool out" slowly and the dark signal nonuniformity is relatively more pronounced at lower temperatures. This is illustrated in Fig. 3.1.7-2. Since the signals from hot pixels can depend on the electric field, they can also be influenced by signal level (i.e., as a potential well fills with charge, the electric field changes).

On rare occasions it has been observed that the dark current from a hot pixel switches between two (sometimes more) discrete values in a random fashion[24*] as illustrated in Fig. 3.1.7-3. In this case, a dark current map taken at one instant may not be valid at later times. Fortunately this random telegraph signal behavior is very infrequent and is likely to affect only one or two pixels in a large-area array, and most arrays will have no such pixels. Flickering pixels (essentially the same effect) have been reported in CCDs used for spectroscopy.[25]

Bright pixels can also be produced in a CCD by nonthermal mechanisms such as short circuits between clock electrodes. These short circuits can be caused by small pinholes or regions of dielectric breakdown in oxide of the layers that insulate the polysilicon electrodes. Such defects are particularly bright, giving signals close to the full well capacity of the device. A good cosmetic quality device will usually have no such defects (having only the smaller dark current defects). If the defect is particularly bright, or the integration is long, then the signal may be large enough to spread down a column, thus giving a "white column" defect.

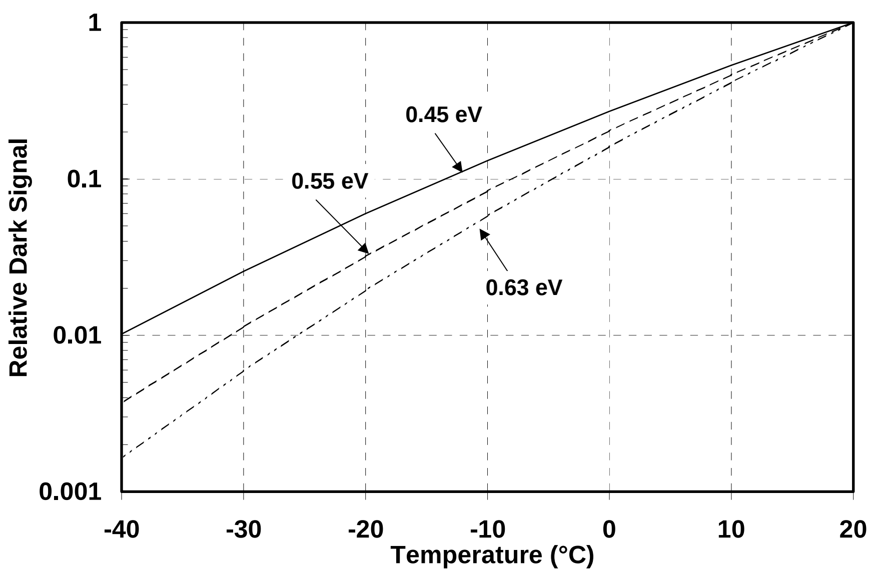


Figure 3.1.7-1 The effect of dark current activation energy on the variation of dark current with temperature. Most pixels will have an activation energy of about 0.63 eV, but most "hot" pixels will have a lower value.

* Although Ref. 22 is primarily concerned with random telegraph signals in irradiated detectors (which can be a common effect), such signals can occasionally be observed in unirradiated devices.

A low-contrast partial white column is produced when charge is injected from a bright white defect into all the pixels that are transferred through it. The partial white column will become a full white column if the defect occurs in the storage region of the CCD.

Note that these white defects tend not to decrease upon cooling (since they arise from electrical injection effects rather than thermal generation). Also, their location can vary with the particular electrode phases that are used for integration. (Hence, a manufacturer's defect map taken with integration under clock phase 1, for example, may not give the same white defects as a map for integration under phases 2 and 3).

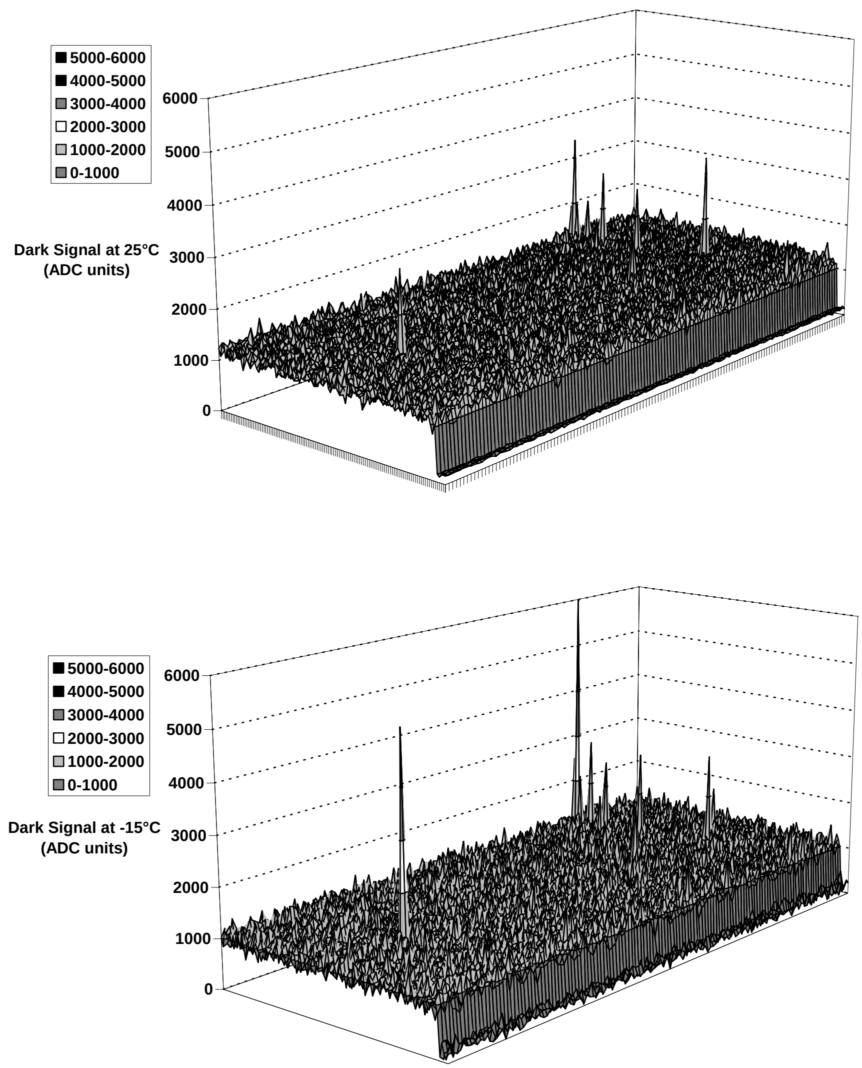


Figure 3.1.7-2 Dark signal nonuniformity for a CCD at 25°C and –15°C. The integration time and gain are such that the average dark signal is the same in both cases, but the spikes are relatively more pronounced at the lower temperature.

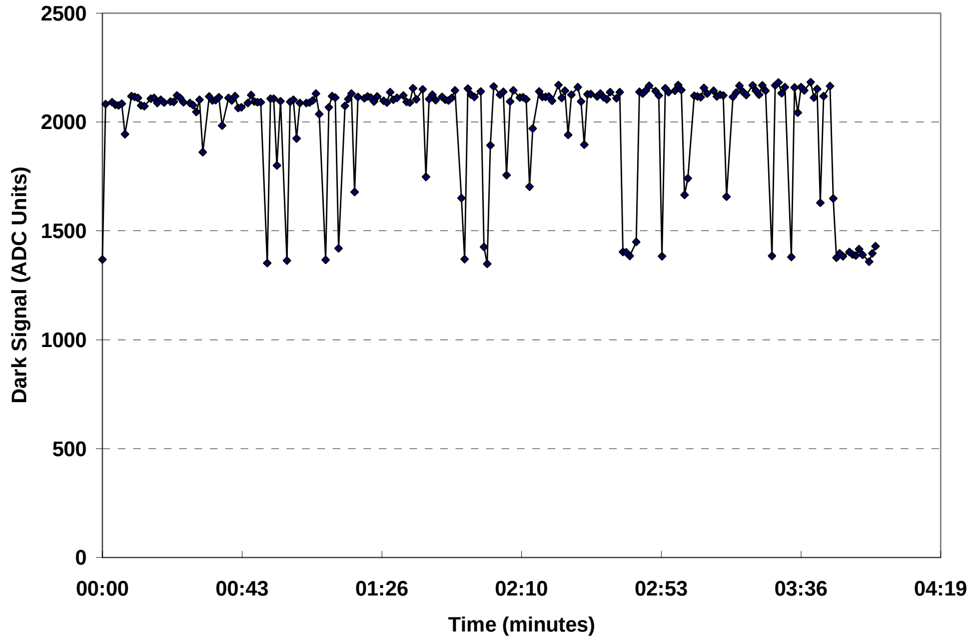


Figure 3.1.7-3 On rare occasions a hot pixel can have a dark signal that switches randomly in time, like a random telegraph signal. The time constant can vary between several seconds and several minutes at room temperature.

Finally, another source of (usually uniform) dark signal in CCDs is clock-induced impact ionization. In some types of CCDs operated in inversion and with fast clock edges or large clock swings, the electric fields at the surface can be high enough for holes to be accelerated enough to cause impact multiplication and spurious injection of charge. When this occurs, the associated dark signal will not depend on integration time or temperature, but only on the image/storage area clock drive waveform and the number of clock pulses. This phenomenon is rare, however, and can usually be ignored for current commercially available devices.

In summary, it can be said that dark signal average level and nonuniformity depends very much on operating conditions (particularly temperature) and should be characterized for each set of instrument readings.

3.1.8 Nonlinearity

Nonlinearity can arise either in the detector array or in the off-chip electronics. Usually any nonlinearity is most pronounced at the extremes of the sensitivity range (i.e., either for small or for large signals). CCDs and photodiode arrays will usually be linear for small signals, however some types of CMOS detectors can be nonlinear in this region. In some cases a nonlinearity (e.g. a logarithmic response) will be deliberately introduced by the array manufacturer to extend the dynamic range, but these types of arrays will rarely be used for applications requiring any radiometric accuracy.

Otherwise, the main instance of detector nonlinearity will occur for high signals, close to the full well capacity (or saturation level). The response will become nonlinear either when charge spills over into adjacent pixels (an effect termed *blooming*) or when charge starts to interact with trapping states at the silicon surface.[26] Which condition occurs first (bloomed full well or surface full well) depends on the voltages applied to the device. For a CCD in the surface full well case, the nonlinearity will depend on the nature of the illumination since this determines the occupancy of the surface traps. Uniform illumination will tend to keep the surface traps full and extend the linearity range (increase the full well capacity). On the other hand, for spot illumination much of the array surface is in darkness and surface traps are emptied. This tends to reduce the full well capacity. For this reason it is recommended that nonlinearity is checked using illumination that is similar to that used in the application, though it is recognized that in many cases it will only be feasible to use quasi-uniform illumination. Methods to determine array and system linearity are discussed in detail in Secs. 1.2.11 and 4.2.5.

An example of an instrument-level linearity measurement for a CCD-based imaging spectrometer is given in Sec. 1.2.11. With CCDs it is usual to expect a linearity of ~ ± 0.1% (of full scale) over the dynamic range, with no wavelength effects.

In a CCD there is the possibility to bin pixels together, either vertically in the readout register, or horizontally in the output amplifier (see Sec. 1.1.4.1). Hence, the full well capacity of the readout register and nonlinearities in the on-chip amplifier are also important. Usually the readout register pixels have several times the area of the image area pixels so as to accommodate some binning (full well capacity is proportional to pixel area). The point at which the output amplifier becomes nonlinear depends on the bias voltages that are applied, but will usually be of the order of 1 V. This can be a problem for arrays that have a high CVF—a feature that is otherwise desirable since it reduces the readout noise. This is because a 1 V output swing will correspond to only 10^5 electrons if the CVF is 10 μV/electron, but to 10^6 electrons if the CVF is 1 μV/electron. For this reason, high-sensitivity arrays sometimes have a switchable CVF so as to allow for high signal illumination.

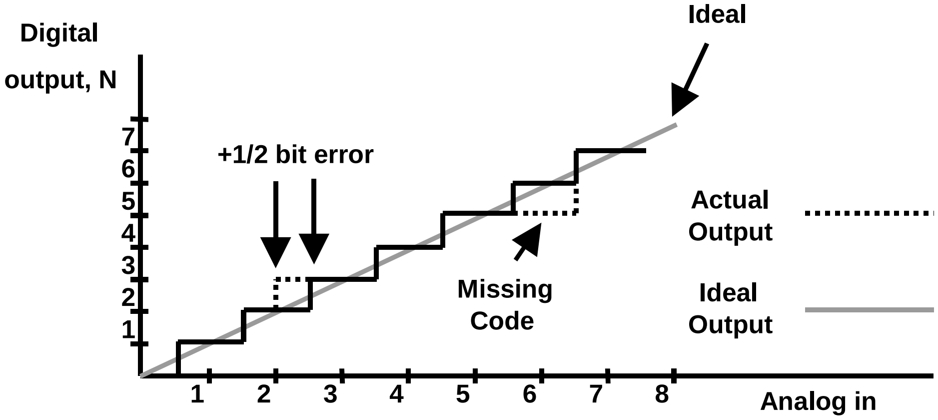
Nonlinearity can also arise in the off-chip electronics, for example, external amplifiers and the ADC circuitry. For an ADC there are two parameters: the integral nonlinearity and the differential nonlinearity (as illustrated in Fig. 3.1.9-1).

To summarize, the linearity of the detector system will normally be stable in time, but it can depend on illumination conditions and on applied voltages, as well as the amount of pixel binning. The linearity is a good check on the general health of a system and the absence of any fault conditions. It is also an important parameter for the user because of its effect on the measured data. Hence, linearity is a useful parameter to check on a regular basis.

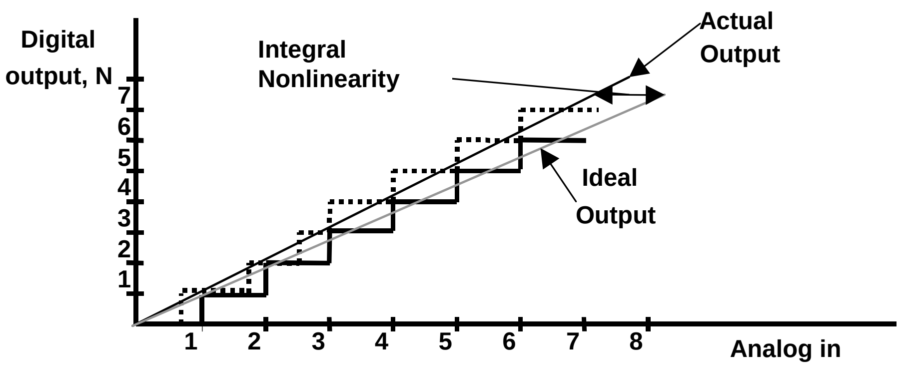
3.1.9 Full Well Capacity

This parameter has been discussed in the previous section. In its own right, the full well capacity determines the dynamic range of the system:

$$\text{dynamic range} = \frac{\text{full well capacity}}{\text{noise}} . \tag{3.1.9-1}$$



Differential nonlinearity



Integral nonlinearity

Figure 3.1.9-1 Illustration of differential and integral nonlinearity in an ADC. Differential nonlinearity (DNL) is defined for code i as $(w_i-w)/w$, where w_i is the width of the i^{th} code (ADC bin) and w is the average width. Hence, if there is a missing code the DNL is –1.

3.1.10 Noise

In the context of an array-based instrument, the noise parameter usually refers to temporal noise in the image data. If samples of the output of a given pixel are repeatedly measured over a period of time and the rms value is taken, then this is the noise (see Sec. 1.2.10). There is also spatial noise due to pixels having a different dark signal or responsivity, but this is not normally referred to as noise. Here we reserve the term to mean temporal noise only.

There will be a temporal noise component that varies as the square root of the signal (termed *Poisson noise*), but for pixels that have no signal the noise comes solely from the readout process. This readout noise will normally be stable, but it can vary if the CCD temperature changes by a large amount (tens of degrees) or if there is a change in the bias voltages applied to the output amplifier. The readout noise will also be influenced by any electromagnetic interference (or “pick-up”). Imaging arrays are normally sensitive enough that a small amount of pick-up will usually be present, if only from the various clock signals that are used to drive the array or to exchange information with the control and signal processing electronics (and host computer). However, this pick-up should be small compared with the intrinsic readout noise of the array in a well-designed system.

The readout noise will have an effect on the minimum detectable signal and hence on the dynamic range, so this should be periodically checked, particularly to make sure that there is no degradation due to electromagnetic interference or changes in grounding.

3.1.11 Crosstalk Between Pixels

Crosstalk between pixels is a measure of the signal generated (by optical absorption) in a given pixel which is spread to neighboring pixels by processes of charge collection, the readout or the off-chip signal handling (see also Sec. 1.2.18).

For example, the charge initially generated within the boundaries of a pixel may spread sideways by diffusion and be collected in a neighboring pixel (see Fig. 3.1.1-1). The process of charge transfer in a CCD may smear charge into following pixels (either in the vertical or horizontal directions), and the finite bandwidth of the electronic signal processing circuits can result in a crosstalk from one pixel to that in the next pixel readout (in the line direction).

In addition, if an array has several outputs, each fed to a separate off-chip signal processing chain, then signal capacitive coupling between circuit board tracks or components or current loading (if the circuits share the same power supply) can result in a crosstalk between the video signals; a pixel read out from one output affects the pixel on the other output(s) that is read out at the same time. Crosstalk is expressed as a fraction of the signal in a given pixel since it will usually scale with this signal. An exception is the charge transfer efficiency (CTE) of a CCD. Since this is defined as the charge that is left behind after each pixel transfer *divided by the signal*, it might be expected to be independent of signal size. However, when caused by trapping effects the CTE can depend on signal size as well as on the amount of background charge (e.g., from signal or dark current). CTE also differs from the other crosstalk phenomena in that the effect on an image depends on position. The signal in a pixel that is located far from the output and thus has to undergo many pixel transfers will suffer from increased smearing due to imperfect charge transfer efficiency, whereas charge diffusion and off-chip electronics crosstalk will be independent of pixel location. For this reason, CTE is usually distinguished from other sources of crosstalk and is often listed as a separate parameter (see Sec. 1.2.13 and Sec. 3.1.11.1 below).

Crosstalk (from any of the above mechanisms) is important to specify, but the mechanisms are usually stable in time and should only require infrequent performance checks on image quality or spectral resolution. The effects can, however, depend on the nature of the signal and so cause difficulties in calibration.

It should be noted that any change in crosstalk will also affect response and dark signal nonuniformity.

3.1.11.1 CTE

In the case where the charge transfer efficiency of a CCD is dominated by trapping effects, the image quality and photometric accuracy can be affected. This will not usually happen in normal use (although it can be important in radiation environments); however, it can occasionally happen that a particular CCD suffers from poor CTE due to problems in manufacture. In these cases the CTE is usually dependent on both signal and the level of background charge, as well as the operating temperature and the clocking speed (since all these factors influence the trapping and detrapping of charge). The traps that affect the global CTE are usually single electron traps (in contrast to the larger traps that occur in localized defect pixels as described in Sec. 3.1.4.1). They are usually characterized by a capture time τ_c and an emission time τ_e. The capture time depends on the signal (the higher the signal, the shorter the capture time), but the emission time depends only on the temperature and the nature of the trap (usually an impurity atom or a combination of an impurity, or a dopant atom with a lattice vacancy). If large amounts of charge become trapped and are released only after many intervening pixels are read out, then photometric accuracy will be affected. In such situations it will usually be necessary to develop an empirical model of the effect before the response can be calibrated. For a further discussion the reader is referred to Ref. 27 and the references therein. Although this reference deals specifically with CCDs used in *proton radiation environments*, the discussion on CTE is also applicable to the rare cases where *as-manufactured* CCD image quality is dominated by charge trapping.

3.1.12 Lag

3.1.12.1 Lag in CCDs

Lag is an effect seen in some types of linear CCD arrays (with photodiode detector elements), some photodiode arrays and some ILT (and FIT) CCDs. It is due to residual charge left behind after transfer from a photodiode to a CCD register. This process has to take place via a transfer gate. The shorter the open period of the transfer gate, the greater the lag effect that will become noticeable when images vary in time—a preceding image becomes partially imprinted onto succeeding ones. The amount of lag depends on the signal level and is usually reduced when a background charge (or “fat zero”) is present. The amount of lag also depends on the array architecture. In some designs only the excess charge that is generated above an equilibrium level is transferred from the photodiode to the transfer gate (and then to the CCD readout register). This equilibrium level is affected by subthreshold conduction[28] whereby, over long timescales, some charge can “leak” over the transfer gate. The signal readout then depends on the previous signal—hence, lag effects and nonlinearities result. If the photodiode has a *pinned* design, then the potential levels can be designed so that transfer is nearly complete.[29] (Pinned diodes have an additional doped layer near the surface which “pins” the surface to a particular potential, independent of signal). So, pinned photodiode arrays tend to have a much reduced lag. Lag is not present in FF or FT CCDs, apart from trapping effects and residual bulk image effects, as discussed in Sec. 3.1.1. Figure 3.1.12-1 shows data from a linear CCD with unpinned photodiodes. It can be seen that the lag is significant at low signal levels. The data was obtained using a flashing LED, synchronized to the array readout.

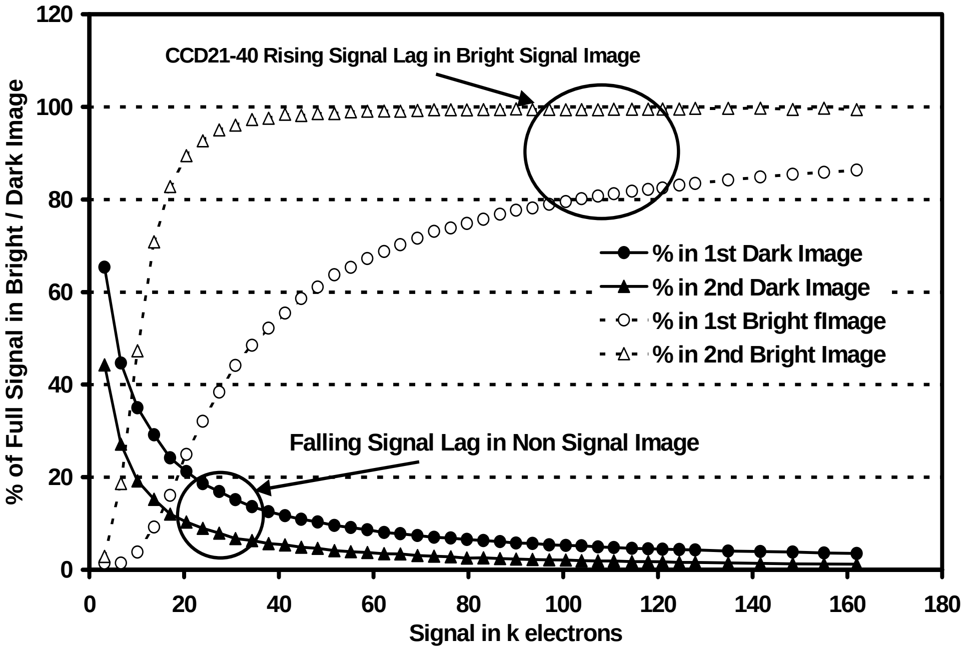


Figure 3.1.12-1 Image lag for a linear CCD with unpinned photodiodes, both for light-to-dark and dark-to-light transitions.

3.1.12.2 Lag in CMOS Active Pixel Sensors

Lag in CMOS active pixel sensors arises from a similar cause to that in a CCD—namely, subthreshold conduction.[30,31] It is most prevalent in three-transistor (3-T) photodiode designs where the reset transistor operates in the subthreshold region at the end of the reset period. This can happen because the reset transistor and the analog supply rail both have the same voltage (e.g. 5 V). However, the reset transistor will have a finite threshold voltage (typically about 1 V), and when the voltage across it falls below that level the transistor will start to cut off (and operate in a region known as "subthreshold"). In this case the reset is "soft," and the photodiode and the reset drain do not reach thermal equilibrium; conduction still occurs across the gate, but at a slow rate.[28] If the signal level on a photodiode suddenly changes, then it can take a while (several frames) for the equilibrium to be re-established. If the light level suddenly decreases, then we have *discharging lag,* seen as comet tails on bright parts of the image or as residual ghost images. If the light level suddenly increases, then we have *charging lag,* where parts of the image are darker than they should be (resulting in dead zones).

A by-product of soft reset is that the readout noise is often reduced (kTC noise is usually reduced by a factor √2).

There are ways of ameliorating the lag problem with CMOS APSs. For example, use of a four-transistor circuit (within the pixel) can result in very low lag.

3.1.13 Resolution and MTF

The spatial resolution in an image is predominantly determined by the number of lines read out (in TV mode devices, determined by the TV standard), by the number of pixels in a line (and the pixel dimensions), and by the crosstalk between pixels (Sec. 3.1.11).

In imaging camera systems, vertical resolution is often expressed in TV lines. In systems with a monitor, a parameter known as the Kell factor (= 0.7) is sometimes used to allow for the average perceived resolution (as judged by the eye).

A quantity related to resolution is the modulation transfer function (MTF), which is a measure of the contrast in an image. If an image is formed of a sinusoidal grating and the maximum and minimum intensities in the image are $I_{\max}$ and $I_{\min}$, then

$$MTF = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{3.1.13-1}$$

at the particular spatial frequency of the imaged grating. If a square-wave target (rather than a sinusoidal pattern) is used, then the measured quantity is the contrast transfer function (CTF).

The MTF of any imaging system will fall off at high spatial frequencies and, for a detector array system, will fall to zero at 1/(pixel width). On the other hand, the sampling frequency f_s is 1/(pixel pitch), and the highest frequency that can be reliably sampled in an image (without the aliasing effects discussed in the next section) is the Nyquist frequency, f_N = 1/(2× pixel pitch). A difficulty with sampled systems (such as detector arrays) is that the MTF will depend on the phase of the signal relative to the pixel positions. This can produce disturbing effects with moving images.[32]

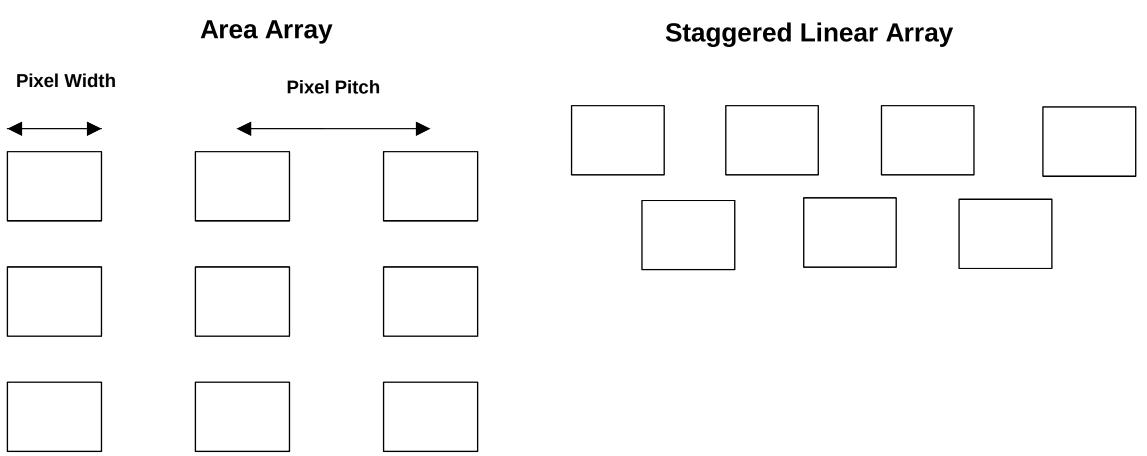


Figure 3.1.13-1 Pixel width and pitch.

The MTF can also be found by measuring the response to slit illumination (i.e., the line spread function) and taking the Fourier transform (other methods are discussed in Sec. 2.2.14). An important property of the MTF of a system is that it is made up of the product of the separate MTFs of the array and the lens:

$$MTF_{\text{system}} = MTF_{\text{detector}} \times MTF_{\text{lens}} . \qquad (3.1.13\text{-}2)$$

3.1.14 Aliasing

For a given pixel pitch, MTF is higher for noncontiguous pixels. However, this is often not particularly useful because of aliasing. This is the beating effect caused by spatial frequencies in the image that are higher than the Nyquist frequency, $f_N = 1/(2\times\text{pixel pitch})$, which interfere with lower frequencies in the image (c.f. Figs. 3.1.14-1 and 3.1.14-2). If aliasing is to be avoided then the MTF at high spatial frequencies has to be reduced (either by reducing the MTF of the detector or of the optical system). Aliasing effects are worse when there are gaps between pixels, for example, with ILT CCDs or CCDs with antiblooming structures. Staggered arrays (Fig. 3.1.13-1) suffer least from aliasing effects.

It is possible to reduce aliasing effects using an optical anti-aliasing (birefringent) filter, by digital processing, or by dithering the position of the array. However, the high-frequency information in the scene (above the Nyquist limit) can never be uniquely recovered, though it can be estimated and aliasing effects can be reduced.[33]

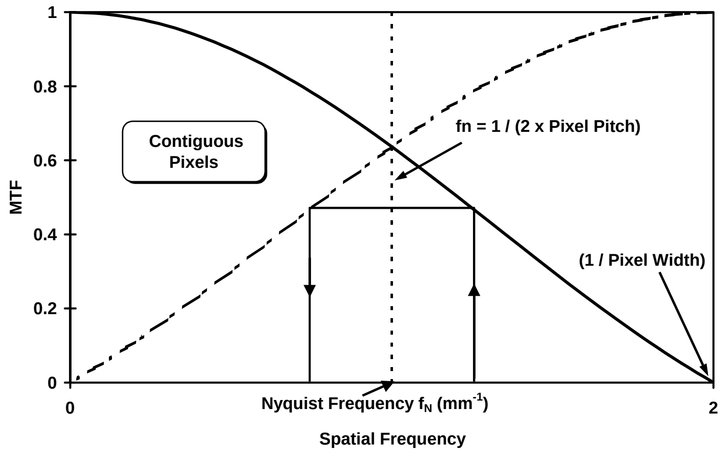


Figure 3.1.14-1 MTF and Nyquist frequency.

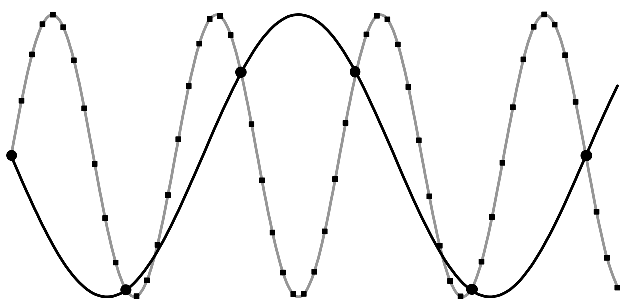


Figure 3.1.14-2 Illustration of aliasing effects at high spatial frequencies. The gray curve is the real signal, but when undersampled is interpreted as being represented by the waveform in black.

3.1.15 Cosmic Ray Effects

Cosmic ray interactions in the upper atmosphere cause showers of particles, predominantly muons that produce transient signals in solid-state array detectors.[34] The muon flux will also produce a flux of electrons and protons by nuclear reactions in local materials. The overall flux will vary with altitude and geomagnetic latitude, but a flux of a few particles/cm^2/minute is typical at sea level. The events tend to be bunched together in time (presumably resulting from the same primary cosmic ray interaction) so that several events can occur in one image (with none in the images preceding or following it). Another source of background events is radioactive emission from nearby materials such as concrete, optical glasses, thin film coatings, and the array package itself (usually ceramic).

When a particle passes through the array it produces electron-hole pairs along its track by ionization. Minimum ionizing particles deposit 80 electrons per micron of track on average. For a 25-µm-thick detector the average signal is therefore 1600 electrons. Note, however, that the distribution of event sizes is non-Gaussian, and events two or three times the average value are usually possible. Figure 3.1.15-1 shows dark CCD images (the exposure time was several minutes and the CCD was cooled to −90°C). Cosmic ray tracks are usually straight; however, the tracks of Compton electrons, produced by radioactive decay, are curved (and are often termed "worms"[35]).

It is to be expected that long-exposure images will occasionally be contaminated by cosmic ray events. To detect these it is usual to compare two identical exposures. If parts of the image have little spatial information (for example, the along-slit direction in a spectrometer), then these also can be used for detecting localized cosmic ray events. Some Raman spectrometers (which often rely on long CCD exposures to detect weak spectral lines) now have built-in software for cosmic ray rejection. Astronomers often take long CCD exposures and are well used to "cleaning" images of these effects. Reference 36, which is available on the Web, gives information on algorithms for processing CCD images. Although these algorithms may not be directly appropriate to most users (they function as part of the general astronomy package

IRAF), they illustrate the general principles involved in cases where the maximum amount of information needs to be extracted from a contaminated image.

Figure 3.1.15-1 Cosmic ray tracks in two CCD images.

Kay and Sadler[37] suggest a simple technique for removing cosmic ray events. If two otherwise identical images are available, then a third—called "lesser" image—can be constructed by comparing the brightness of each pixel in the first image with the corresponding pixel brightness in the second image and storing the lesser of the two. This will automatically remove spurious high signal events. The journal article calculates the effect on mean value and noise—which is usually small.

3.1.16 Smear

Unless an instrument is fitted with a mechanical shutter, light falling onto a detector array while it is being read out will cause spurious signal to be generated, known as *smear*. If the scene contains a bright highlight, then the smear will appear as a low-contrast bright vertical column running through the highlight. The cause depends on the type of detector, though the effects tend to be similar. In a FT CCD light falls directly on the pixels in the image area while they are being read out. The effect depends on the ratio of the integration time to the time to perform a frame transfer:

smear/pixel = (average signal in the image/pixel) × (frame transfer time) / (integration time).

In ILT (and FT) CCDs and in most X-Y addressed arrays, smear is generated by

- diffusion of photo-generated electrons from the photodiode into the vertical CCD register
- multiple reflection ("light-piping") of photons so that they generate electrons directly in the vertical CCD registers

Only the charge injection device (CID) is fully smear-free.[14]

In all types of CCDs, smear will be generated in an image both from the readout of the previous image before the integration takes place ($smear_{before}$) and from readout of the current image after integration (*smear*$_{after}$). Hence, smear appears both above and below an image highlight.

The *smear*$_{before}$ component can be removed in CCDs with gated antiblooming (charge reset). To remove the *smear*$_{after}$ component requires the last line (at least) of the image area to be shielded from light (e.g., by an extension of the storage region shield). The signal in this "smear line" will contain dark charge and the smear signal from the previous image. This can be used as a correction (either automatically in the camera, or in the signal processing), provided that the image does not contain moving objects. If the scene is moving, then a special CCD architecture that gathers several smear lines is needed.[14]

A method for correcting smear by image data processing has been discussed by Powell et al.[38]

3.1.17 Geometrical Stability

Although the fabrication of solid-state detector arrays is performed by photolithographic processes that have good geometric accuracy and stability, errors can arise in the definition of the pixel geometry as follows:

> *Lithographic mask positioning inaccuracy*. Several masks are used in the fabrication process. These are positioned with respect to reference marks on the silicon wafer. The positioning of the masks determine the dimensions of individual pixels, that is, their center and size. Typically the masks can be positioned to better than 0.1 µm.
>
> *Optical distortion of the lithographic system.* The optical system used to image the mask will introduce distortion that is apparent as deviation of the actual center of pixels from the nominal center. The worst case absolute distortion at any point within the field used in the lithographic process is usually about 0.25 µm, and the worst case distortion between the centers of any two pixels is < 0.4 µm.
>
> *Sub-array stitch inaccuracy*. Large CCD arrays are usually formed by a step-and-repeat process, whereby sub-arrays are "stitched" together. Errors in pixel position can result from misalignment in the stitch process, but these should be less than ~ 0.2 µm.

These geometrical errors will be fixed at manufacture and will not vary with time thereafter, although it is possible to get changes in flatness due to distortions of the package (during temperature variations). Effects are usually small. Early back-thinned devices often had the active detector area as a thin membrane that was prone to geometrical distortions. Modern devices usually have the electrode surface mounted on a substrate and stability is as good as for front-illuminated devices.

Geometrical errors may need to be assessed (perhaps even calibrated) for accurate work such as photogrammetry or astrometry but need not be considered for most applications. In photogrammetry it is common to be able to measure to accuracies ~0.01 pixels,[39] which is comparable to the errors given above. Shaklan et al.[40] have reported a way of directly measuring pixel registration errors using a two-beam interference pattern (sine wave) imaged onto the CCD. They found a step-and-repeat error of 0.033 pixels and other errors of the order of 0.01–0.03 pixels; however, this was for an early-generation CCD and performance is expected to be better with most current devices.

3.1.18 Image Anomalies

There is a general class of effects that for want of a better term we will call image anomalies. These are effects (many of which have already been mentioned above) that produce artifacts in images, arising either in the detector array itself or in the detector/off-chip electronics combination. The effects are not fundamental (i.e., they will not occur in a well-designed detector/system) and will be specific to the particular system; also, they will usually only be important in high-performance instruments. They are, however, effects that should be watched for. Examples are

Residual images. As mentioned in Sec. 3.1.1, with some CCDs operated at low temperatures ($< -50°C$) charge can become trapped within the dielectric surface layer (either at the front Si/SiO_2 interface with a front-illuminated CCD or at the back surface of a thinned back-illuminated device). This leads to charge trapping and a weak persistence of a bright image. The length of time that the persistence lasts will depend strongly on the temperature. It should be noted, however, that only some CCDs suffer from these effects. In many cases there is no detectable residual image effect.

Ghost images. Ghost images can be formed by reflections off the surface of the detector array onto the detector window, filters, and field flatteners. The usual remedy is to employ antireflection coatings and tilting of optical components relative to the array. Another cause of ghost images in some back-illuminated CCDs is reflection off the substrate material used to support the thinned detector. In most cases reflection at this surface is suppressed by use of appropriate coatings.

Traps. As discussed in Sec. 3.1.4.1, CCD traps can result in dark or bright trails behind a defect pixel, depending on the nature of the trap and the signal level. Some traps are only visible in faint images and some only in bright ones.

CTE problems. Section 3.1.11.1 discussed the occurrence of poor charge transfer efficiency (CTE) in CCDs. Poor CTE will result in a smearing (i.e., loss of spatial resolution) in an image. If the CCD is operated at low temperature, then the charge that is trapped may be released only slowly and is spread over many pixels. In this case the smearing may not be discernible, but there will be a loss in the signal from the original pixel and this will affect photometric/radiometric accuracy, particularly for pointlike images. The effect will depend on position in the image, being most severe for pixels furthest from the output amplifier, which therefore require the greatest number of transfers before readout.

Crosstalk. In systems with more than one signal channel (either with more than one detector array or with an array with more than one output), crosstalk can occur not only between adjacent pixels (as discussed in Secs. 3.1.1, 3.1.5, and 3.1.11) but also between the signal channels. For example, a bright image in one detector channel may give rise to a faint ghost image in another channel for those pixels that are read out at the same time. These crosstalk ghost images (not to be confused with the optical ghosts discussed above) can be either positive (whiter than average) or negative (darker than average), depending on the electronic crosstalk effect responsible. Examples of electronic effects that can affect more than one channel at once are power supply or voltage reference loading or transients on ground planes. For the ultimate performance in multichannel systems, the drive and signal processing electronics for each channel should be separated as far as possible (though there is a trade-off against complexity and cost).

Contamination. Contamination of the surface of array detectors can be a problem, particularly for low-temperature operation in the UV. In such cases the array is normally mounted inside an evacuated enclosure. Contaminants from vacuum pump oil or from outgassing of electronic components or printed circuit boards will tend to collect on the coldest surface, which is normally the detector itself. This can cause a loss in quantum efficiency in the UV (where the absorption length is low enough that even a thin surface film can affect the response). The remedy is usually to warm the array to room temperature for several hours so as to evaporate contaminants.

Stray light and filter leaks. Stray light can be a particular problem with array detectors because of their large area and wide spectral response. If the region being investigated is at an extreme of the response (in the blue/UV or in the red), then the quantum efficiency is low and signals tend to be small. Hence, any stray light of other wavelengths (where the sensitivity is higher) will give rise to spurious signal. Also, when using narrowband filters it should be checked that the out-of-band transmission is negligible for the whole bandwidth of the detector. This is not always the case; for example, some UV filters allow some transmission in the red ("red leaks").

Vignetting. Systems using large area detectors can be susceptible to vignetting, where the optical aperture is not sufficient to allow uniform illumination of the array. There is then a fall-off in illumination at the edges of the field of view. When this occurs the sensitivity will clearly depend on the detector position and the illumination conditions. Any change in these will affect the flat-field response.

Pick-up. Since array-based instruments have to operate using clocking waveforms to readout the detector and to operate the off-chip electronics (clamps, sample and holds, ADCs, etc.), it is inevitable that there will be some capacitative pick-up of these signals onto the video waveform. Also, the small signals used (the noise level of the detector is typically a few μV) makes these instruments susceptible to elecromagnetic interference. In a well-designed system these effects will usually be comparable with the noise level;

however, this may not always be the case. Pick-up will usually have the appearance of bar or herringbone patterns in the image. Electromagnetic interference can occur in short bursts (e.g., from switch on/off of nearby equipment), in which case horizontal glitches may appear in the image.

Spurious dark signal. As discussed in Sec. 2.2.3, clock-induced dark charge and amplifier glow can be produced in some CCDs.

3.1.19 Electrical Parameters

The following parameters are only of interest for instrument design and are not expected to change with operating conditions; hence, they do not impact on calibration procedures:

power supply current(s)
output impedance (of the on-chip amplifier)
electrode capacitances
leakage currents (between pins and across the gate isolations)
DC output voltage
reset clock feedthrough

3.2 Intensified Arrays

In this section we briefly discuss parameters important to the use of PDAs and CCDs coupled to a photocathode/microchannel plate (MCP)/phosphor combination. Intensified cameras are classified in terms of their generation:

GenI	Multi-alkali photocathode, electrons focused on to a phosphor screen, heavy and bulky
GenII	Same as GenI but with an MCP to improve gain, size, and weight
GenII+	Same as GenII but with the photocathode optimized for the near-infrared
GenIII	Same as GenII but with a GaAs photocathode to enhance near-infrared response (at the expense of blue response)

Further information on intensified arrays is given in Ref. 41 (see also Sec. 1.1.3). The main issues are

1. Geometric distortions, for example due to nonuniform spacing of the pores of the MCP. As with the array itself, these distortions will be fixed at manufacture.
2. Response nonuniformity. As well as array variations, there will be low spatial frequency variations (typically 10%) across the photocathode and in MCP gain. There can also be localized hot spots in the MCP response caused by damaged or contaminated MCP channels. These can give a time-varying response.
3. In photon counting instruments, nonlinearities can be produced at high signal levels due to pulse-pile-up effects. For example, at a signal level of 0.2 electrons/pixel/CCD readout time, the nonlinearity will be ~ 1%.[41]

4. Nonlinearities can also be produced in the MCP at high illumination levels if the current through the channels reaches the replenishment limit (set by the resistance of the MCP and the potential difference across it.[42]
5. The photocathode will have a characteristic dark charge generation rate, usually described as the equivalent background illumination or EBI. The effect can be minimized by operating in gated mode where the intensifier is only "on" for the short periods when signal is integrated. In continuous mode it may be necessary to cool the photocathode. A 20°C reduction in temperature gives a decrease in EBI by a factor 10–20 for UV photocathodes and of 50 to 100 for NIR-enhanced ones.
6. The persistence of the phosphor screen should also be taken into account for critical photometric applications with a time-varying signal.[43]

3.3 Conclusions

In this chapter, the variations in array parameters due to changes in operating conditions have been described. Changes in operating temperature and applied voltages can be expected to cause changes in dark signal (and dark signal nonuniformity), offset voltage, and quantum efficiency (especially at long wavelengths). In addition, parameters such as response nonuniformity, linearity and stray light will vary with illumination conditions. Noise, and hence dynamic range can be altered if there is a change to the level of electromagnetic interference (for example, if grounding arrangements change).

With only a few exceptions, silicon-based arrays for detection in the UV, visible, and near-infrared are linear devices. Hence, a *two-point* calibration of each pixel is sufficient. This means that the response of a pixel can be characterized by just two parameters: an offset (i.e., dark signal) O_j and a responsivity R_j . For illuminations $P1_j$ and $P2_j$, the signals $S1_j$ and $S2_j$ are

$$S1_j = O_j + R_j * P1_j , \tag{3.3-1}$$

$$S2_j = O_j + R_j * P2_j . \tag{3.3-2}$$

This gives two equations in two unknowns, which can be solved for O_j and R_j. These need to be determined separately for each pixel. Usually one of the illuminations will be "darkness," in which case the corresponding P_j will be zero. In that case, the usual formula for deriving the true signal from the raw data is

$$\text{Corrected signal} = \frac{(\text{Raw} - \text{Dark})}{(\text{Flat field} - \text{Dark})} \times (\text{mean value of Flat field} - \text{Dark}) . \tag{3.3-3}$$

It is advisable to calibrate O_j for each set of measurements (since the offset and dark current depend strongly on temperature). This can be simply accomplished by recording and subtracting the signal in darkness for each pixel (preferably averaged over several frames so as to reduce noise). The responsivities R_j and most other parameters will require only periodic checks.

To measure responsivity, linearity, full well, and stray light, it will normally be necessary to illuminate the whole array using a stable light source and filters of known transmittance (see Secs. 4.2 and 5.2).

The process of correcting for pixel responsivity variations is known as *flat fielding*. At instrument level this will correct for the array response (which will be wavelength dependent), imperfections in the optics, dust, and any pattern in the light source. Hence, it is important that the illumination conditions are as close as possible to those encountered during normal use. Poor flat fielding can be worse than having no correction. Note that to achieve good signal-to-noise during the flat-field calibration, the illumination level should be high (roughly half-saturation) but not so high as to saturate any pixels. Averaging of image frames may be needed if the light level is low (for example, at the extremes of the wavelength scale).

Measurement of crosstalk, CTE, and geometric stability will require more specialized equipment, such as optics for spot projection, but these parameters are rarely required to be measured by users.

It has been mentioned that array full-well capacity can depend on the method of illumination and that spot illumination tends to be more influenced by surface traps and to be given a reduced value relative to uniform illumination. This will be important for some applications, and users will need to bear this in mind; however, it will not generally be feasible for users to measure linearity/full-well capacity at system level with by means other than uniform illumination. Spot illumination measurements are more suited to calibrations (by the manufacturer) of the arrays themselves.

References

[1] M. A. Damento, A. A. Barcellos, and W. V. Schempp, "Stability of lumogen films on CCDs," *Proc. SPIE*, vol. 1762, pp. 204-210 (1995).

[2] J. Janesick, T. Elliot, T. Daud, J. McCarthy and M. Blouke, "Backside charging of the CCD," *Proc. SPIE*, vol. 570, pp. 46-79 (1985).

[3] J. Janesick, "CCD transfer method–standard for absolute performance of CCDs and digital camera systems," *Proc. SPIE*, vol. 3019, pp. 70-102 (1997).

[4] A. R. Scaefer, E. F. Zalewski and J. Geist, "Silicon detector nonlinearity and related effects," *Appl. Opt.*, vol. 22(8), pp. 1232-1236 (1983).

[5] A. T. Hattenburg and J. B. Shumaker, "Linearity study of a diode array radiometer," *Appl. Opt.*, vol. 23(19), pp. 3257-3258 (1984).

[6] S. Robson, T. A. Clarke and J. Chen, "The suitability of the Pulnix TM6CN camera for photogrammetric measurement," *Proc. SPIE*, vol. 2067, pp. 66-77 (1993).

[7] A. Cumani, A. Guidicci, P. Grattoni and G. Pettiti, "Characterisation of commercial solid-state TV cameras for non-contact metrology," *Proc. SPIE*, vol. 2067, pp. 115-122 (1993).

[8] J. M. C. Plane and C.-F. Nien, "Differential optical absorption spectrometer for measuring atmospheric trace gases," *Rev. Sci. Instrum.*, vol. 63(3), pp. 1867-1876 (1992).

[9] F. H. Yang, M. M. Blouke, D. L. Heidtmann, B. Corrie, L. D. Ridley and H. H. Marsh, "Trap characteristics in large area CCD image sensors," *Proc. SPIE*, vol. 1071, pp. 213-228 (1989).

[10] K. P. Klaasen, M. J. Belton, H. H. Breneman, A. S. McEwen, M. E. Davies, R. J. Sullivan, C. R. Chapman, G. Neukum, C. M. Heffernan, A. P. Harch, J. M. Kaufman, W. J. Merline, L. R. Gaddis, W. F. Cunningham, P. Helfenstein and T. R. Colvin, "Inflight performance characteristics, calibration, and utilization of the Galileo solid-state imaging camera," *Opt. Eng.*, vol. 36(11), pp. 3001-3027 (1997).

[11] D. E. Groom, S. E. Holland. M. E. Levi, N. P. Palaio, S. Perlmutter, R. J. Stover and M. Wei, "Quantum efficiency of a back-illuminated CCD imager: an optical approach," *Proc. SPIE,* vol. 3649, pp. 80-90 (1999).

[12] P. R. Jorden, A. P. Oates, S. M. Tulloch, D. J. Ives and X. Gao, "New generation CCDs-applications and performance results," *Proc. SPIE*, vol. 3355, pp. 457-468 (1998).

[13] G H. Mount, R. W. Saunders and J. W. Brault, "Interference effects in reticon photodiode array detectors," *Appl. Opt.*, vol. 31(7), pp. 851-858 (1992).

[14] A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices* (Kluwer, Dordrecht, 1995).

[15] J. A. Penkethman, "Calibrations and idiosyncasies of micro-lensed CCD cameras," *Proc. SPIE*, vol. 3779, pp. 241-249 (1999).

[16] G. Putnam, S. Kelly, S. Wang, W. Davis, E. Nelson and D. Carpenter, "Photography with an 11-megapixel 35-mm format CCD," *Proc. SPIE*, vol. 5017, pp. 371-383 (2003).

[17] P. R. Jorden, J.-M. Deltorn and A. P. Oates, "The non-uniformity of CCDs and the effects of spatial undersampling," *Proc. SPIE*, vol. 2198, pp. 836-850 (1994).

[18] R. K. Winge, V. A. Fassel and D. E. Eckels, "Spectral line-diode registry effects with photodiode array detectors," *Appl. Spectrosc.*, vol. 40(4), pp. 461-464 (1986).

[19] D. Kavaldjev and Z. Ninkov, "Subpixel sensitivity map for a charge-coupled device sensor," *Opt. Eng*., vol. 37(3), pp. 948-954 (1998).

[20] T. A. Lind, R. K. Reich, W. H. McGonagle and B. B. Kosicki, "Intrapixel response test system for multispectral characterisation," *Proc. SPIE,* vol. 3649, pp. 232-238 (1999).

[21] ESA Contract Report on contract 5254/82, "Evaluation of two-dimensional GEC and Thomson-CSF imaging arrays" (Sira Ltd, 1986).

[22] J. Van Der Spiegel and G. J. Declerck, "Characterization of dark current non-uniformities in charge-coupled devices," *Solid State Electronics*, vol. 27(2), pp. 147-154 (1984).

[23] J. R. Srour and R. A. Hartmann, "Enhanced displacement damage effectiveness in irradiated silicon devices," *IEEE Trans. Nucl. Sci.*, vol. NS-36(6), pp. 1825-1830 (1989).

[24] G. R. Hopkinson, "Proton irradiation induced RTS in CCDs," *Proc. 14th Int. Conf. on Noise in Physical Systems and 1/f Fluctuations*," C. Claeys and E. Simeon, eds., pp. 218-223 (World Scientific, Singapore, 1997).

[25] H. Prytherch, "The characterisation of a CCD and camera system for atomic emission spectroscopy," *Proc. SPIE*, vol. 2416, pp. 25-84 (1995).

[26] J. Janesick, T. Elliott, R. Winzenread, J. Pinter and R. Dyck, "Sandbox CCDs," *Proc. SPIE*, vol. 2415, pp. 2-42 (1995).

[27] G. R. Hopkinson, C. J. Dale and P. W. Marshall, "Proton effects in CCDs," *IEEE Trans. Nucl. Sci.*, vol. 43(2) pp. 614-627 (1996).

[28] T. E. Hendrickson, "A simplified model for subpinchoff conduction in depletion-mode IGFETs," *IEEE Trans. Elec. Dev.*, vol. 25(4), pp. 435-441 (1978).

[29] E. C. Fox et al., "A high speed linear CCD sensor with pinned photodiode photosite for low lag and low noise imaging," *Proc. SPIE*, vol. 3301, pp. 17-26 (1998).

[30] H. Tian, B. Fowler and A. El Gamal, "Analysis of temporal noise in CMOS photodiode active pixel sensor," *IEEE J. Solid State Cir.*, vol. 36, no. 1, pp. 92-101 (2001).

[31] E. R. Fossum, "Charge transfer noise and lag in CMOS active pixel sensors," *2003 IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, Elmau, Germany (2003).

[32] D. J. Bradley, C. J. Baddiley and P. N. J. Dennis, "The modulation transfer function of focal plane array systems," *Proc. SPIE*, vol. 807, pp. 33-41 (1987).

[33] A. H. Lettington and Q. H. Hong, "An anti-aliasing algorithm for focal plane arrays," *Proc. SPIE*, vol. 2470, pp. 134-140 (1995).

[34] J. F. Ziegler and W. A. Lanford, "The effect of sea level cosmic rays on electronic devices," *J. Appl. Phys.*, vol. 52(6), pp. 4305-4312 (1981).

[35] A. R. Smith et al., "Radiation events in astronomical CCD images," *Proc. SPIE*, vol. 4669, pp. 172-183 (2002).

[36] L. A. Wells and D. J. Bell, "Cleaning images of bad pixels and cosmic rays using IRAF," http://iraf.noao.edu/iraf/web document clean.ps.Z (1994).

[37] L. Kay and D. A. Sadler, "A method for processing CCD images to remove cosmic rays and other randomly positioned spurious events – theory and experiment," *Meas. Sci. Technol.*, vol. 2(6), pp. 532-535 (1991).

[38] K. Powell, D. Chana, D. Fish and C. Thompson, "Restoration and frequency analysis of smeared images," *Appl. Opt*., vol. 38(8), pp. 1343-1347 (1999).

[39] J. M. Raynor and P. Seitz, "The technology and practical problems of pixel-synchronous CCD data acquisition for optical metrology applications," *Proc. SPIE*, vol. 1395 (part1), pp. 96-103 (1990).

[40] S. Shaklan, M. C. Sharman and S. H. Pravdo, "High-precision measurement of pixel positions in a charge-coupled device," *Opt. Eng.*, vol. 34(29), pp. 6672-6681 (1995).

[41] Y. Talmi, "Intensified array detectors," in *Charge-Transfer Devices in Spectroscopy*, J. V. Sweedler, K. L. Ratzlaff and M. B. Denton, eds. (VCH Publishers Inc, 1994).

[42] R. A. Antcliff, M. E. Hillard and O. Jarrett, Jr., "Intensified silicon photodiode array detector linearity: application to coherent anti-stokes Raman spectroscopy," *Appl. Opt*., vol. 23(14), pp. 2369-2375 (1984).

[43] B. R. Sandel, D. F. Collins and A. L. Broadfoot, "Effect of phosphor persistence on photometry with image intensifiers and integrating readout devices," *Appl. Opt*., vol. 25(20), pp. 3697-3704 (1986).

CHAPTER 4

CALIBRATION TECHNIQUES FOR INSTRUMENTS

4.1 Imaging Instruments

4.1.1 Silicon-based Array Systems

Calibration for most visible imaging systems involves a combination of dark and flat-field measurements (to derive the gain and offset of each pixel, as discussed in Sec. 3.3), perhaps with additional measurements of spectral response and MTF. The techniques involved are essentially the same as for individual arrays, and the reader is referred to Sec. 1.2. Not all the techniques will be suitable for a complete system, however, depending on accessibility to the appropriate electronic signals and whether the collection optics are removable. In cases where a complete system (including optics) requires calibration, it should be borne in mind that re-calibration (especially for absolute intensity response) may be needed for each configuration of the optics (e.g., after focus, aperture, or filter changes).

In the field of digital photography (i.e., electronic still picture cameras) it is common to use test charts both for resolution measurements and for determination of the relation between digital output and scene luminances (for electronic cameras the log illuminance is used and the latter relationship is termed the optoelectronic conversion function, or OECF). The following ISO standards are applicable for digital cameras:

ISO 12231:1997	Photography—Electronic still picture cameras—Terminology
ISO 12233:2000	Photography—Electronic still picture cameras—Resolution measurements
ISO 14524:1999	Photography—Electronic still picture cameras—Methods for measuring opto-electronic conversion functions (OECFs)
ISO 15739:2003	Photography—Electronic still picture cameras—Noise measurements
ISO 15529:1999	Optics and optical instruments—Optical transfer function—Principles of measurement of modular transfer function (MTF) of sampled imaging systems

In commercial CCD cameras the horizontal spatial resolution can be degraded by too low a bandwidth in the video signal processing electronics. This bandwidth is often chosen for good noise performance and a pleasing image appearance rather than good (spatial) impulse response.

The effect may be observable in MTF measurements (which tend to show an initial rise rather than a fall-off with spatial frequency) but overshoot and ringing effects (Fig. 4.1.1-1) can often be better assessed by imaging of spot or bar targets.

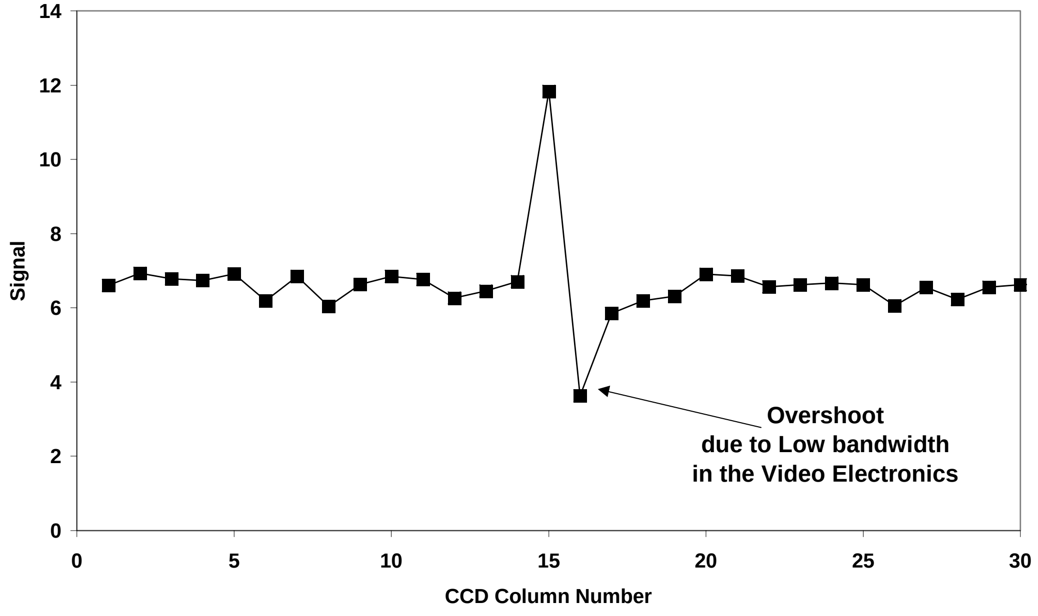


Figure 4.1.1-1 Effect of low video bandwidth and "overshoot" on the appearance of spot or line illumination. The plot shows a scan along a CCD line.

Stray light can be important for radiometric applications, especially if there are bright sources outside the nominal field of view. Assessment can be made using bright point sources or lasers.

Calibration for geometrical accuracy (e.g. for photogrammetry) is a somewhat specialized activity and usually involves imaging of special line or dot matrix targets. There are many examples in the literature, for instance, Shortis et al.[1] and Cumani et al.[2] mention targets for visible cameras, whilst Amin et al.[3] report on a target for calibrating thermal imagers (see also Ref. 4) and Wang et al.[5] on a target for X-ray cameras using intensified CCDs. In a camera system that employs digitization in the framegrabber, image distortions can arise due to errors in synchronization (see, for example, Ref. 6) and this can be separately assessed using video pattern generators (in place of the camera). Luo et al.[7] have also discussed the measurement of line jitter in CCD images. However, jitter and synchronization problems can be avoided in digital cameras, where the digitization is performed in the camera itself. Warm-up effects have been observed in the geometric performance of both cameras and frame grabbers. Recently, Habib and Morgan[8] have described a simple (though computer intensive) automatic method for distortion calibration using straight line features in the image.

A calibration technique for optical profiling applications has recently been described by Asundi and Wensen[9]—though in these applications the distortions are more likely to be produced by the optical system than in the array itself.

4.1.1.1 Color Imagers

With color imagers it is usually the relative color values (the chromaticity) rather than the absolute brightness that is important.[10] For example, for a camera with red/green/blue (RGB) outputs, it is the values of R/(R+G+B) and G/(R+G+B) that are of concern.

To improve the color rendition of array-based cameras it is necessary to perform both white balance and color correction.[11] White balance is the operation whereby separate gains are applied to the R, G, and B pixels so that a reference white in the scene gives equal response in the those pixels. Color correction is then applied to transform from the native color space of the camera to the color space of the display standard. Since crosstalk errors can result in photons of one color spreading to pixels of another color (particularly for long- wavelength photons absorbed deep in the device), it is generally necessary to use a 3×3 color correction matrix, C, and an offset (r_{offset}, g_{offset}, b_{offset}) to convert the white-balanced and interpolated RGB outputs (R_{cam}, G_{cam}, B_{cam}) to the correct output values (R_0, G_0, B_0):

$$\begin{bmatrix} R_0 \\ G_0 \\ B_0 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \times \begin{bmatrix} R_{cam} \\ G_{cam} \\ B_{cam} \end{bmatrix} + \begin{bmatrix} r_{offset} \\ g_{offset} \\ b_{offset} \end{bmatrix} \qquad (4.1.1\text{-}1)$$

The matrix coefficients can be found by imaging colored tiles or color checker charts. The method used by Blanksby and Loinaz[11] was to use a gradient descent algorithm to find the matrix coefficients that minimized the color error, defined as the root sum square of the differences between the measured (after correction) and the true tristimulus color values. In that case they used 24 color tiles.

Reference 11 gives an example of a CMOS active pixel sensor (APS) where the color error had a particularly large value of 7.7, compared with a value of 4.8 for a CCD camera. The authors used the CIE L*u*v* color space (although other color spaces can be used). An error of 1 unit in the CIE L*u*v* color space gave a just noticeable difference. The large error for the APS device was due to the low quantum efficiency in the blue and a large crosstalk (the blue pixels had a substantial response for wavelengths 500–750 nm). This was confirmed by separate measurements of the average quantum efficiency of each color pixel, obtained using a monochromator.

As pointed out by Putnam et al.[12], color errors can arise when light hits the array at a shallow angle. It can than, for example, pass through a green filter but become absorbed in an adjacent blue or red photodiode. Hence, color error can be a function of angle of incidence (and hence lens configuration).

4.1.1.2 Recommendations

1. Flat fielding (for correction of response nonuniformity) should be performed with illumination conditions (wavelength and cone angle) as close as possible to those in normal use and at an illumination level bright enough to give good SNR but not high

enough to give saturation. Flat fielding at several illumination levels may be needed to detect traps.

2. Dark images should, if possible, have the same exposure time as the images to be corrected. A check should be made on dark charge stability (flickering pixels).
3. Linearity measurements should follow the recommendations given in Sec. 4.2.5.4 (see also Sec. 1.2.11)
4. Other measurements can be made using those methods from Sec. 1.2 that are most applicable. Note that a typical imager will have its optical performance influenced by the camera lens and any band-limiting filters.
5. Warm-up effects can influence both the camera electronics (e.g., gain) and the geometric stability. For accurate work this may require investigation by the user and the derivation of a suitable start-up procedure.
6. Checks should be made for spurious image artifacts (c.f. Sec. 3.1.18) and electronics effects (e.g., pick-up or bandwidth effects).

4.1.2 Thermal Imagers

As mentioned earlier, the testing and calibration of thermal imagers is a specialized subject and cannot be exhaustively discussed within the scope of this guide. The following is a brief discussion of some of the main issues. The reader is referred to Ref. 13 for a general treatment of testing at the camera level. Nonuniformity and nonlinearity issues are discussed in Refs. 14, 15, 16, 17, 18 and19.

4.1.2.1 Nonuniformity

The nonuniformity, both in pixel responsivity and offset (i.e., dark current and fixed pattern noise) tends to be worse for long-wavelength detectors than it is in the visible. (This is because of the nature of the detector materials technology and also because IR images tend to have a high background component and low contrast.) This nonuniformity can also be bias dependent. We consider a particular example as an illustration. Infrared photodiodes are often coupled to the (usually silicon) readout circuit by a direct injection gate (see Sec. 1.1.6). The injection efficiency (the fraction of charge that is transferred from the diode to the readout circuit) depends critically on the bias applied to the gate (the direct injection gate voltage, DIG). This is because this voltage determines the bias voltage (V) across the photodiode and hence its slope resistance (R). A change in resistance will affect the impedance matching to the output circuit and hence the injection efficiency:

$$\text{Injection efficiency, } \eta = \frac{g_m}{g_m + 1/R}, \tag{4.1.2-1}$$

where $1/g_m$ is the input impedance of the readout circuit, given by

$$g_m = \left(\frac{e}{nKT}\right)\left(\eta I_D + \frac{V}{R}\right), \tag{4.1.2-2}$$

where I_D is the diode current (dark current plus photo-current). Hence, η is reduced for small currents (signals) and low-resistance diodes. The dependence on diode current also leads to nonlinearities, as discussed below.

Nonuniformity can be measured by viewing a blackbody source. Measurements need to be made for at least two source temperatures in order to determine both the gain and the offset for each pixel. A true dark field is not usually possible because, first, IR emission from the dewar window and fore-optics will produce some illumination, and secondly, if a cold target is used the resultant signals are likely to be small enough that the response will be nonlinear.

Figure 4.1.2-1 shows a typical uncorrected flat-field image for a 64 × 64 LWIR mercury cadmium telluride (MCT) photodiode array showing the large nonuniformity. Figure 4.1.2-2 shows histograms for several blackbody temperatures, with the response to a 30°C blackbody subtracted. This subtracts offset variations but not responsivity variations, and so the histograms are broadened as the blackbody temperature is increased. To get reasonable images these responsivity variations have to be measured and corrected for each pixel. For the temperature range used in this test the diode response was linear, as illustrated in Fig. 4.1.2-3 (though this will not always be the case; see below). However, the slope of the plot will depend on the DIG voltage.

Figure 4.1.2-4 shows the effect of only a few millivolt change in the DIG voltage for a group of pixels selected for fairly uniform response. Pixels that have a high responsivity at one voltage tend to still have a high responsivity at another DIG voltage; however, the absolute response changes and there is also a scatter in the points, showing that the pattern of the nonuniformity changes slightly with DIG voltage. It will also change slightly with a change in array operating temperature. Figure 4.1.2-5 shows the change in average absolute responsivity with DIG voltage for this array.

Since pixel responsivity and offset can be sensitive to operating conditions, especially for LWIR detectors, frequent flat-field calibration is recommended.

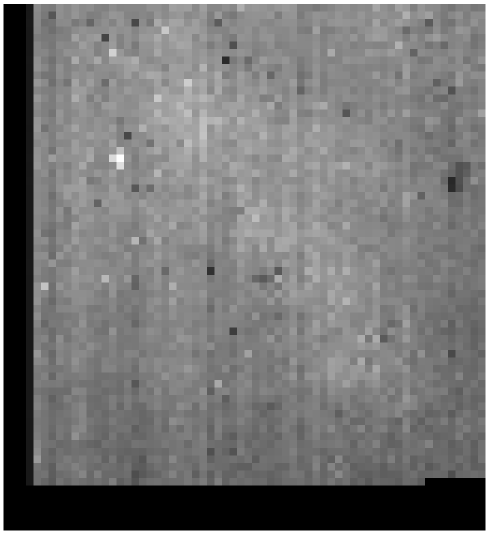


Figure 4.1.2-1 Uncorrected flat-field image from a 64 x 64 LWIR MCT photodiode array.

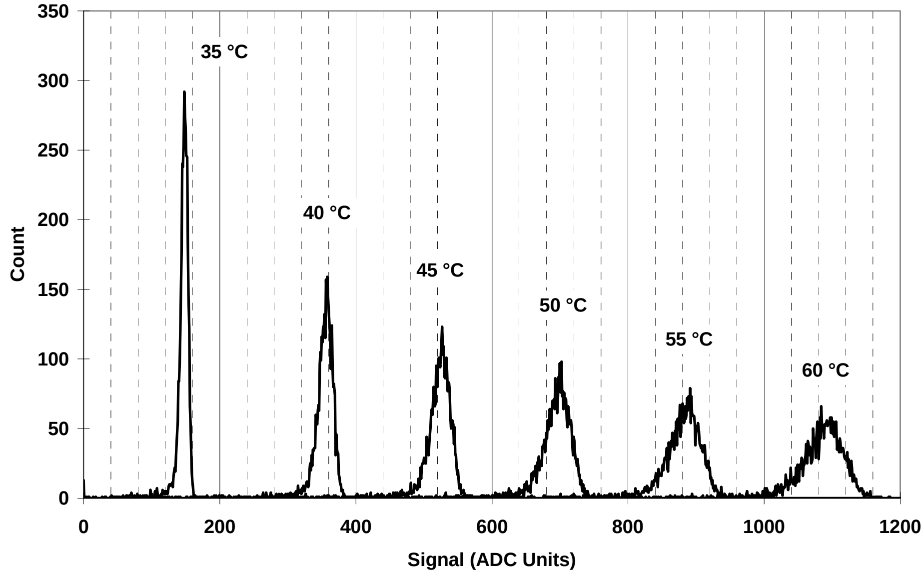


Figure 4.1.2-2 Histograms for several blackbody temperatures for selected photodiodes, with the response to a 30°C blackbody subtracted (128 x 128 LWIR MCT array).

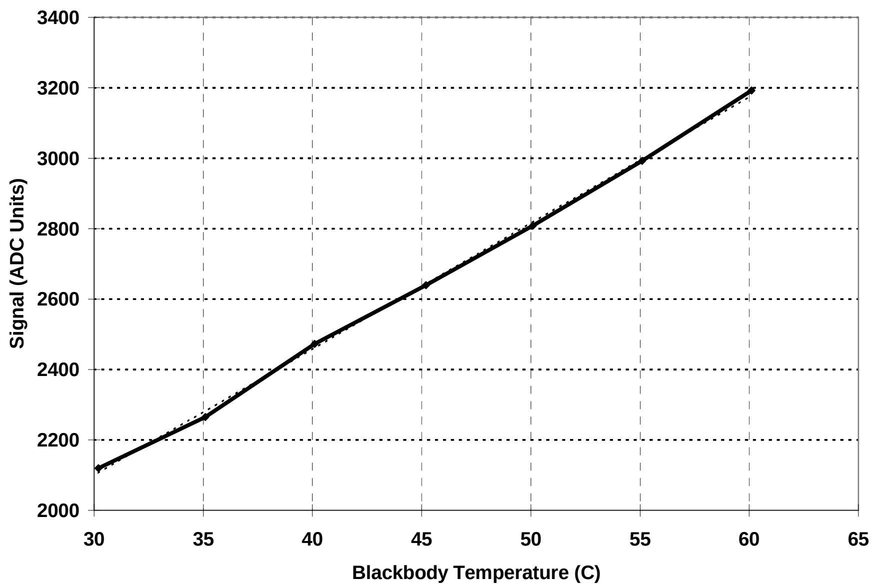


Figure 4.1.2-3 Average signal versus blackbody temperature (from the data of Fig. 4.1.2-2).

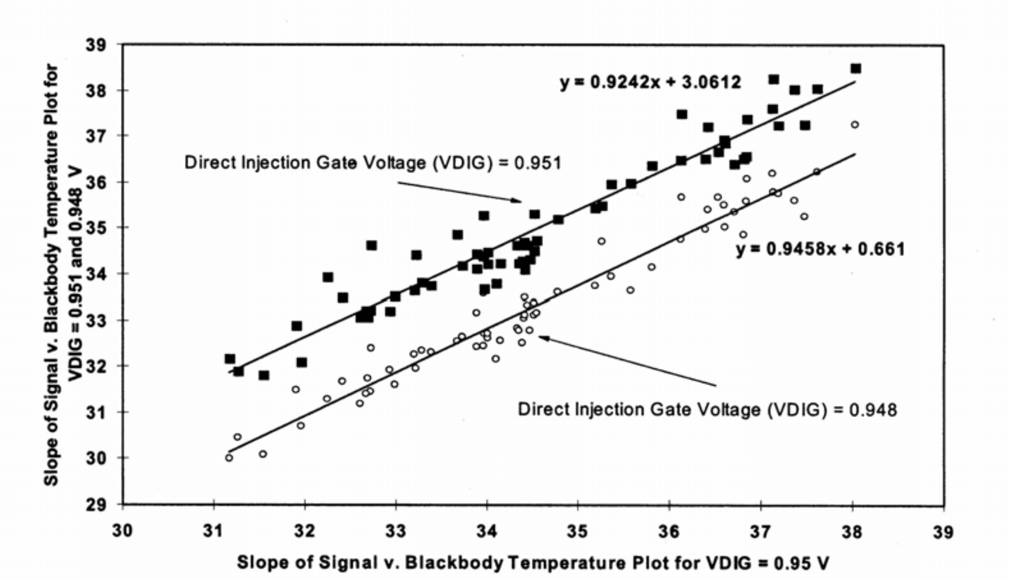


Figure 4.1.2-4 Change in average absolute responsivity with DIG voltage for an array with LWIR MCT photodiodes and direct injection (CMOS readout circuit).

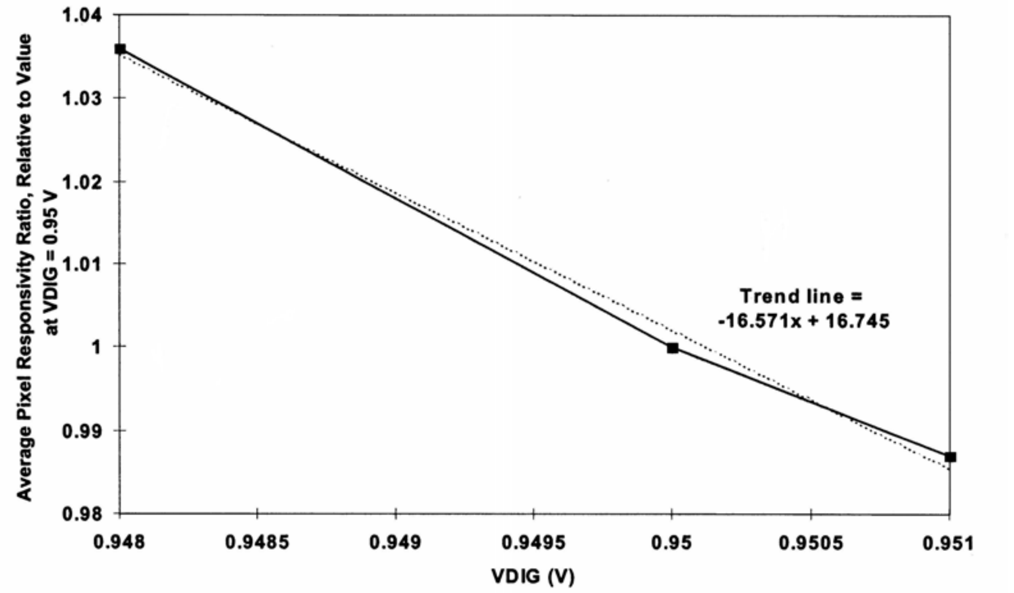


Figure 4.1.2-5 Change in average absolute responsivity with DIG voltage from the data of Fig. 4.1.2-4.

4.1.2.2 Nonlinearity

Linearity for IR detectors can be measured in several ways:

- Varying blackbody source temperature (as above)
- Varying aperture
- Varying integration (stare time)

Because the injection efficiency for a direct injection circuit depends on the signal, there is an inherent nonlinearity that depends on the diode resistance. This is illustrated in Fig. 4.1.2-6, which is an example from a particular instrument simulation model.

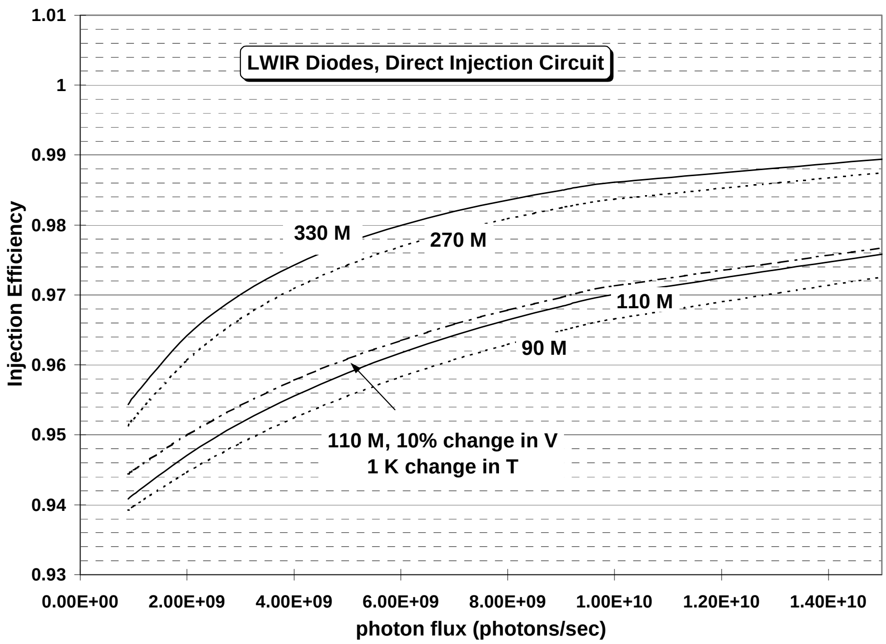


Figure 4.1.2-6 Example of injection efficiency versus photon flux for a particular instrument using LWIR MCT photodiodes. The plots are labeled with the diode resistance values, R.

There are several other effects that should be considered. Firstly, there is a dependence of R on the photon flux. This is believed to be because, for LWIR diodes in particular, R is limited by avalanche multiplication processes.[20] These are sensitive to applied voltage and lead to a resistance term dependent on the flux.

Another effect occurs in the SWIR, where diode currents are low (because of the low photon flux). This leads to high values for the input impedance of the multiplexer [Eq. (4.1.2-2)]. This is not a problem for injection efficiency since the diode resistance is also high. But the high resistance in combination with the diode capacitance leads to a long RC time constant compared with the stare time, and this reduces the injection efficiency and makes it dependent on operating conditions.

For a typical minimum current of 1.5×10^6 electrons/pixel/s we have an input impedance of $5.4 \times 10^{10}\ \Omega$ from Eq. (4.1.2-2). A typical diode capacitance is 0.1 pF, giving a time constant of ~5 ms. Bluzer and Jensen[21] give a frequency dependent $\eta(f)$ as

$$\eta(f) = \frac{g_m R}{1 + g_m R} \times \frac{1}{1 + RCf}, \qquad (4.1.2\text{-}3)$$

where f is the operating frequency. This leads to an injection efficiency ~ 0.8 for 50 Hz operation.

Finally, the linearity of the output amplifier of the readout circuit might be expected to change with operating point (and hence on bias and temperature). However, all pixels read out through a given amplifier would be affected, so this will not cause pixel–pixel variations in linearity.

4.2 Instruments for Spectrometry

This broad title describes a diverse range of instruments and techniques, but these generally fall into two categories—spectroradiometry and spectrophotometry. Since calibration techniques vary, depending on the class of instrument, this classification is discussed further below.

Spectroradiometry covers the measurement of spectral radiance, spectral irradiance, spectral radiant intensity and geometrically total spectral radiant flux (usually referred to as spectral total flux). These measurements quantify the optical radiation emitted from a source, or incident on a surface, in terms of the distribution of energy or power with wavelength, under defined geometric conditions. Spectral irradiance, for example, is the spectral distribution of the radiation incident on a surface, per unit area of that surface. A spectroradiometer normally consists of input optics, a device for splitting the radiation into its constituent wavelengths (often using a grating or prism—but band-pass filters may also be used) and a suitable detector system. If the wavelength splitting component has an exit slit, then it is usually termed a *monochromator* and a *polychromator* if it does not (note: array systems will not have an exit slit and so will employ a polychromator, a term which will be used in the remainder of the guide). The input optics are used to collect the radiation and pass it to the polychromator, and their design will vary according to the geometry of the measured quantity. For example, for spectral radiance measurements a focusing mirror is generally used to image the source onto the entrance to the polychromator; whereas in the case of spectral irradiance measurements, the use of a small integrating sphere or diffusing plaque is typical. The design of the input optics can have a significant effect on the performance of a spectroradiometer and is therefore discussed further in Sec. 4.2.2 (responsivity calibration) and in Refs. 22, 23 and 24. A schematic of an array-based spectroradiometer for irradiance measurements is shown in Fig. 4.2-1.

Spectrophotometry is the measurement of the proportion of incident optical radiation that is transmitted, reflected, or absorbed by an artifact, as a function of wavelength and under defined geometric conditions. A spectrophotometer typically comprises a source, illuminating optics, sample holder, collection optics, a polychromator, and a detector system. The illuminating and collection optics are used to achieve the required measurement geometry; e.g., for 0/45 reflectance measurements, the sample is irradiated at normal incidence and the radiation reflected at 45° to the normal is collected. The polychromator may be placed either before the illuminating or after the collection optics, although for a diode array-based system it will always be in the latter position. A schematic of a typical array detector-based spectrophotometer for regular transmittance measurements is shown in Fig. 4.2-2.

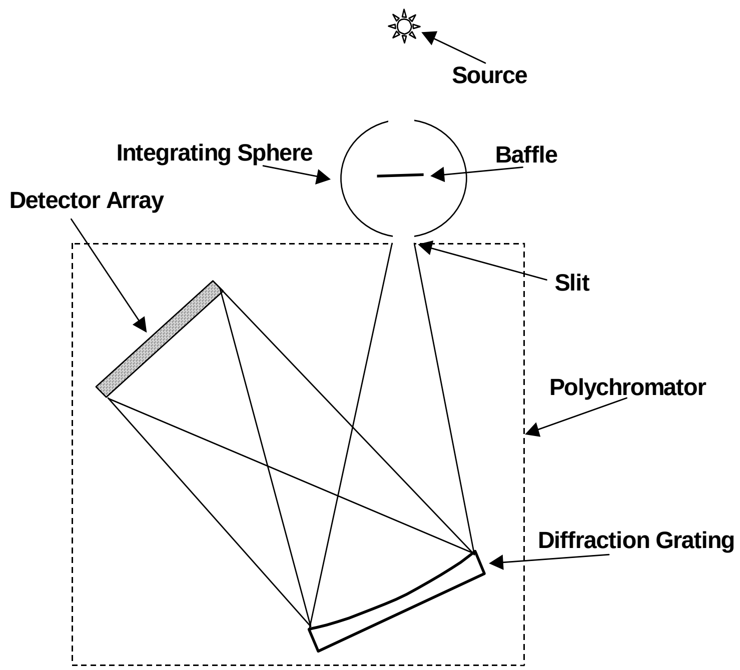


Figure 4.2-1 Schematic of an array-based spectroradiometer for irradiance measurements.

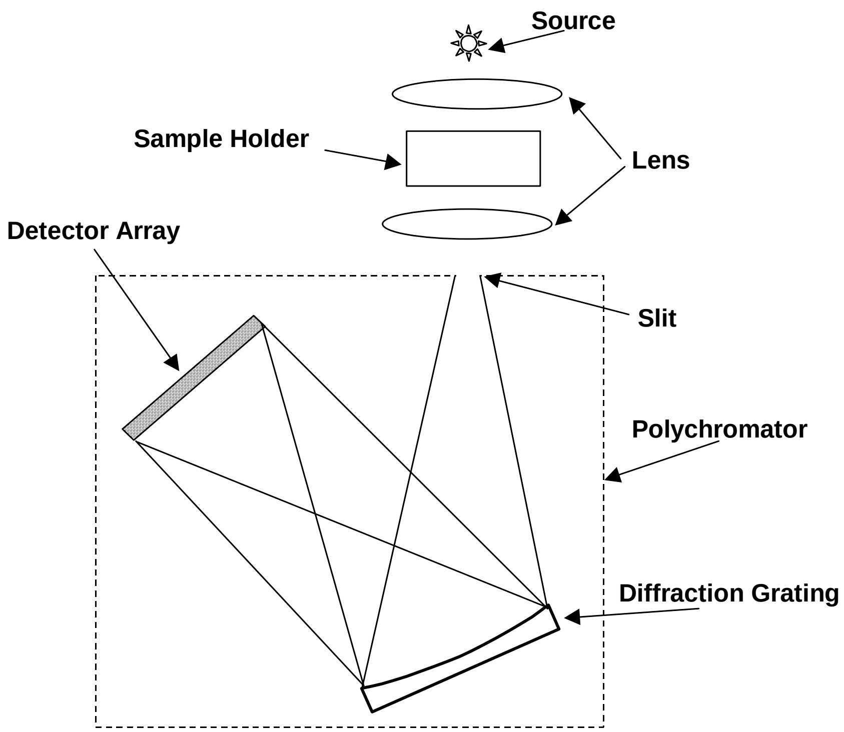


Figure 4.2-2 Schematic of a typical detector-array-based spectrophotometer for transmittance measurements.

The following sections describe potential sources of error, transfer standards and recommended methods for the calibration of key aspects of the performance of array-based spectrometers. Many of the examples given refer to systems using a linear photodiode array, but the principles can be directly applied to systems using CCDs or 2D detector arrays. Some of the techniques and standards described are suitable for use with all types of spectrometers, but more generally the recommended approach for a spectroradiometer and a spectrophotometer will differ, and this is highlighted as appropriate.

4.2.1 Definitions

The following definitions are used in this section. Although alternative definitions are possible, those given below are in most common usage. Note that in the case of resolution the term "spectral sampling interval" is to be preferred. Also, if the dispersion of the spectrometer is nonlinear, then the bandwidth and resolution (expressed in nanometers) will vary with pixel number.

Slit function	The relative spectral transmittance of a polychromator (monochromator) for a given spectral adjustment and a given setting of its slit width(s)
Bandwidth	Also termed *instrument bandwidth* or *spectral width*. It is the width (usually the full width at half maximum, or FWHM) of the line spread function. This function is the convolution of the slit function and the pixel aperture response. Bandwidth is measured in wavelength units (usually nanometers).
Resolution	Also termed *pixel pitch* or *spectral sampling interval*. It is the spacing (in wavelength units) between samples of the spectrum. Note that the effective spectral pixel width will be less than the pixel pitch if there are dead spaces between pixels, and will be larger than the pixel pitch when there is crosstalk between pixels. Although the term resolution is common in spectroradiometry, use of spectral sampling interval is preferred.
Spectral range	The full wavelength range over which a spectrometer can measure optical radiation. For a fixed grating (or prism) this will be the number of pixels in a line multiplied by the bandwidth, but will be greater than this if the grating is moveable.

4.2.2 Wavelength Calibration

4.2.2.1 Introduction

For any spectrometer it is usually essential to know how well the instrumental wavelength reading agrees with the true wavelength, i.e., to calibrate the wavelength scale. Before discussing how this is done, it is first necessary to consider the characteristics of the spectrometer in a little more detail. In a "perfect" spectrometer, it would be possible to make measurements at an infinite number of wavelengths using infinitely narrow wavebands. Of course this is not possible, and measurements are actually made across finite spectral bands. As discussed above, the width of each spectral band is the bandwidth and is determined by the attributes of the instrument, in particular the entrance slit width and the width of the pixels. For a given wavelength setting λ_s, the detected signal (the line spread function) is a function both of λ_s and of the wavelength λ of the incident radiation. The shape of this function $T(\lambda,\lambda_s)$ is the convolution of the slit function of the polychromator and the aperture response of the pixels. In an aberration-free instrument with a plane grating, the slit function is rectangular since there is no exit slit (with a curved holographic grating the slit function tends to be more Gaussian). If the array response is also rectangular, then the line spread function is trapezoidal in shape, centered about λ_s and falling to zero on either side (as shown in Fig. 4.2.2-1). If the entrance slit and the pixel width are the same size, the line spread function will be triangular. In a practical instrument the corners of the function will be rounded to some extent by the effects of diffraction, scattered radiation and the non-rectangular shape of the pixel response. The width $\Delta\lambda$ of the line spread function at the half amplitude points is known as the bandwidth of the spectrometer for the wavelength λ_s, and in a plane grating spectrometer will usually be reasonably constant across the spectrum for a given combination of slit width and grating pitch.

Changing the slit width will change the line spread function. For an array system, the pixel size is fixed and will therefore set a lower limit to the bandwidth. The bandwidth can be increased, however, by increasing the entrance slit width.

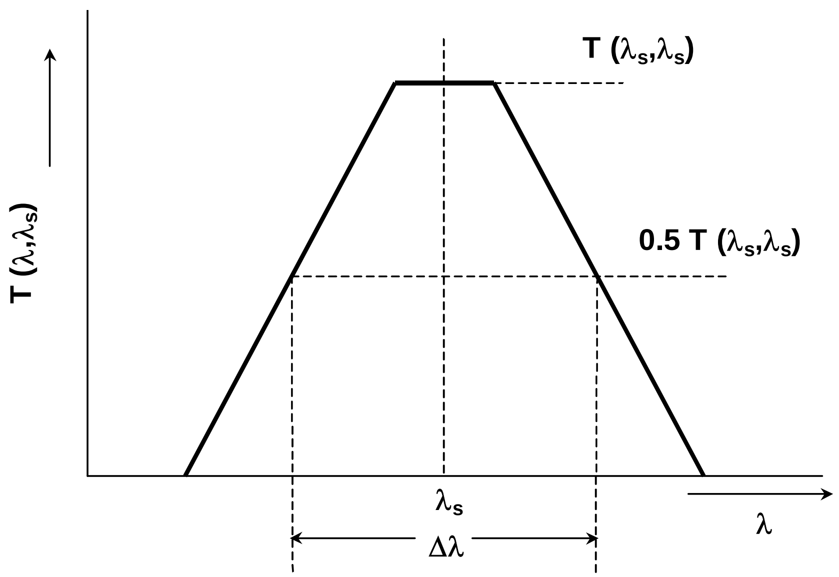


Figure 4.2.2-1 Line spread function for a plane grating spectrometer with an ideal rectangular pixel response.

If a narrow spectral feature (e.g., the emission line from a low-pressure discharge lamp) is used to illuminate the polychromator, then examination of the line spread function allows the magnitude of any errors in the wavelength scale of the instrument to be determined. Suppose the true wavelength of the spectral feature is λ_T and that the center of the measured line-shape (taken as midway between the points at which the signal falls to half its maximum value) is λ_S. Then, the error in the wavelength scale is $\lambda_S - \lambda_T$. This is sometimes referred to as the FWHM method of determining bandwidth and wavelength.

There are a number of features of array-based systems that can make calibration of the wavelength scale particularly difficult compared to the calibration of conventional scanning systems. In scanning instruments a sine-bar is generally used to linearize the relationship between scan distance and wavelength; calibration at a few points can give a good assessment of how well this has been done. In array-based instruments, the spectrum is spread across the (evenly spaced) pixels of the array. Software algorithms are used to link the individual elements of the array with their corresponding wavelength. The relationship is not straightforward but depends on how well the spectrum is focused at the exit plane, the angle at which the grating has been set (this is often adjustable, to allow a single unit to measure across several spectral regions), the regularity of the size and spacing of the pixels in the array, etc. Because of these problems, the accuracy of the algorithm, and hence the accuracy of the wavelength scale, depends critically on the number of points at which the system is calibrated and the algorithm calculated. Experience with a number of systems has shown that large errors (even as high as several nanometers) can be introduced if, for example, only three strong lines from a mercury discharge lamp are used for calibration of a spectroradiometer in the visible spectral region. Use of a large number of emission lines covering the entire working wavelength range of the instrument, on the other hand, has been found to yield wavelength scale errors that are approximately within the resolution (pixel size) of the array instrument, as shown in Fig. 4.2.2-2. The sources used in the multiple case illustrated were mercury, neon and argon lamps, and a helium-neon laser; the resolution of the instrument (the spectral sampling interval) was 0.5 nm and the bandwidth used was approximately 1 nm. The main peak wavelengths used in the calibration are summarized in Table 4.2.2-1.

Table 4.2.2-1 Wavelengths of main emission peaks of some commonly used wavelength calibration sources (and used in obtaining Fig. 4.2.2-2).

Mercury	Neon	Argon	HeNe laser
253.65 nm	585.25 nm	696.54 nm	632.8 nm
296.73 nm	594.48 nm	706.72 nm	
302.15 nm	614.31 nm	738.4 nm	
404.66 nm	621.73 nm	750.39 nm	
435.84 nm	626.65 nm	772.38 nm	
546.07 nm	638.3 nm		
	659.9 nm		
	692.95 nm		

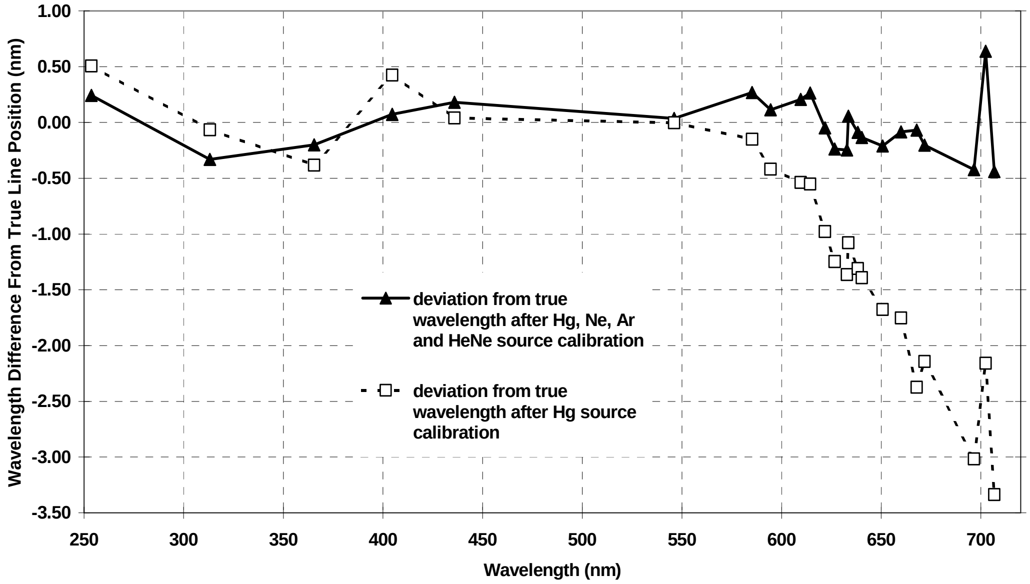


Figure 4.2.2-2 Wavelength calibration errors for a 1024-element diode array spectroradiometer for two calibration scenarios.

Wavelength-scale error can also vary with the type of fitting algorithm applied to the calibration points in order to define the wavelength scale. The most commonly used are quadratic, polynomial, and cubic spline, and experience has shown that the latter generally gives the optimum fit and hence the smallest wavelength-scale errors (see Refs. 25 and 26). In some cases, however, a fit using trigonometric functions can give an improved standard error of estimate,[26] and Lindrum and Nickel[27] and Diem et al.[28] have reported good results using trigonometric functions in the case where the exact grating equation is known. Tests using a range of different fitting algorithms (if the system software supports this) may aid the user to determine the best fit and lowest wavelength-scale error, rather than relying on the default fit contained in the software.

Having found a relation between pixel number and wavelength, the calculation of the resolution (the wavelength range spanned by one pixel; in effect, the dispersion) as a function of pixel number is straightforward. This is often a valuable parameter for assessing system performance.

The results obtained in a wavelength calibration will also be influenced by the sampling nature of the array output. Instead of the continuous slit function obtained with a scanning instrument, one effectively records a histogram, similar to that shown in Fig. 4.2.2-3, and the central wavelength can then only be determined by interpolation between the points on this histogram (or some other arithmetic manipulation). The number of points available for this calculation (which is determined by the relationship between the bandwidth and the pixel size) and the interpolation method chosen can clearly have a significant influence on the results obtained.

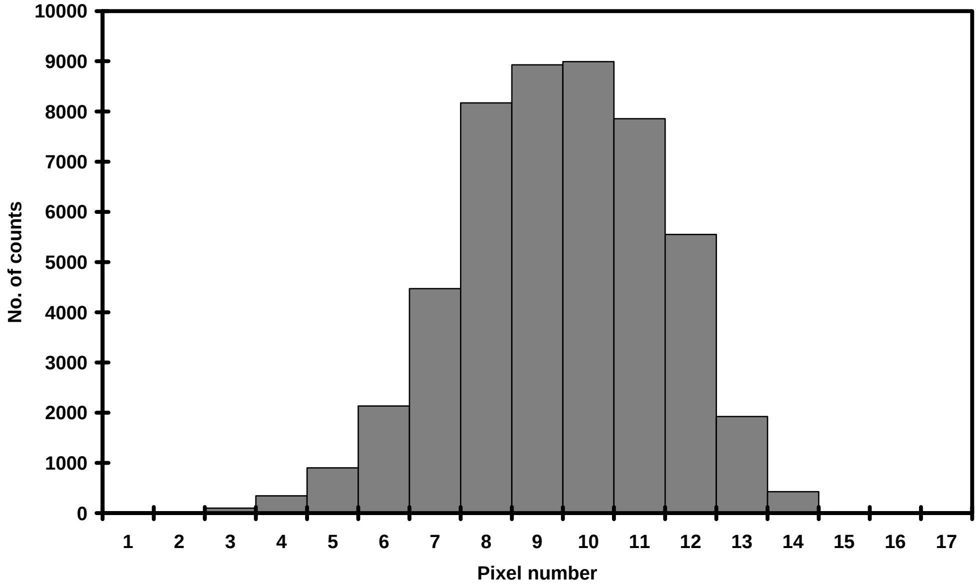


Figure 4.2.2-3 Histogram recorded by a PDA spectroradiometer during wavelength calibration.

Optimal results will usually be obtained by at least five data points for the calculation (see, for example, Ref. 25 for a discussion of wavelength interpolation errors). In other words, the bandwidth should be adjusted (by adjusting the entrance slit width) so that it is several times larger than the pixel width. This will also tend to minimize the effects of pixel response nonuniformity in the array (typically a few percent pixel response nonuniformity will give < 0.05 pixel error in line centroid position if there are four or more pixels within the bandwidth). However, it is equally necessary to ensure that the bandwidth is not too wide— this will make it more difficult to determine the true shape of the line spread function and also increase the influence of noise on the measurements and reduce the ability to resolve closely spaced spectral features. Whatever the bandwidth chosen for the wavelength calibration, this should not be changed when using the system for measurements. Conversely, if a particular bandwidth (or entrance slit width) is necessary when using the system for measurements, then this same bandwidth should be used when calibrating the system.

Use of the FWHM method on linearly interpolated data points (as discussed above) is generally acceptable but is not recommended for high-accuracy work—a cubic spline fit applied to the peak prior to determination of the FWHM value gives better results. Several other approaches have been described by Sullivan and Quimby,[29] who also give references to previous studies:

- Parobolic (quadratic) fit to three or more pixels spanning the top of a peak (this method was extended by Tseng et al.[30] who interpolated the data by adding zeros to the Fourier transform (which was apodized to force the line shape to a Gaussian), the peak was then found by fitting a polynomial.
- Median method, the line center is taken as the fractional pixel location that bisects the base-line-corrected area.
- Center of gravity method (described below).

Sullivan and Quimby found that for their spectrometer, the center of gravity method was superior to both the median algorithm and a parabolic fit (see Ref. 31 for other results obtained using this method). The line center is defined as the quotient of the sum of the base-line-corrected pixel signals (S_i), weighted by the pixel numbers (i), and the line area:

$$\text{center of gravity} = \frac{\Sigma i S_i}{\Sigma S_i}\,. \tag{4.2.2-1}$$

This algorithm is also straightforward to use when there are only a few samples across the spectral line.

Berlot and Locascio[32] have discussed the particular case where there are dead spaces between pixels and few samples within the bandwidth; however, they assumed somewhat idealized rectangular pixel response functions.

Most of the above methods (but not those fitting to the center of the peak) rely on making an allowance for the baseline level. This is usually straightforward by linear, multilinear, polynomial or functional fitting to clear regions of the baseline (depending on the nature of the background spectrum); but if the spectrum is crowded and clear regions hard to find, then this can give errors.

Figure 4.2.2-4 shows an example of the FWHM method applied to a calibration using the 546.07-nm emission line from a mercury discharge lamp, using straight-line interpolation to join the data points. The amplitude of the maximum signal is first determined (line *a*), and the points at which a line drawn horizontally at half of this amplitude (line *b*) intersect with the interpolated data are obtained. The difference between these points is the bandwidth, and the point midway between them is taken as the calibration wavelength (line *c*). In this example, the true wavelength is 546.07 nm and the calibration wavelength is 546.1 nm; i.e., there is no significant wavelength error. Note that if the peak signal (545.6 nm) were taken as the calibration wavelength, then a wavelength error of 0.5 nm would erroneously be determined. Applying the center of gravity method to the results shown in Fig. 4.2.2-4 gives a value of ~546.0 nm for the wavelength of the 546.07-nm emission line, i.e., a slightly greater error than the FWHM method.

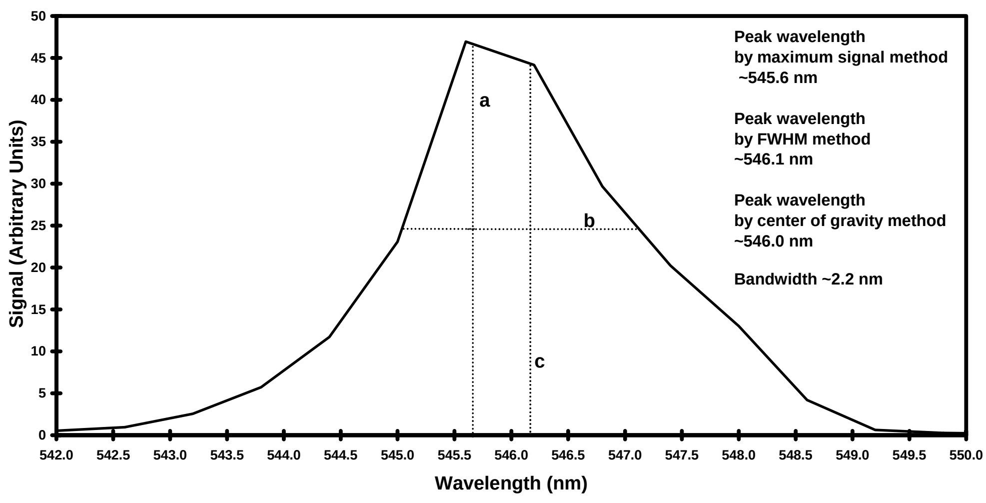


Figure 4.2.2-4 Examples of wavelength calibration using the 546.07-nm peak from a mercury discharge lamp.

Figure 4.2.2-5 shows the effect of changing entrance slits from 0.025 mm to 0.05 mm or 0.12 mm on an array-based spectroradiometer with a pixel resolution of 0.5 nm. The change in bandwidth can be clearly seen, but additionally a small shift in center wavelength is apparent. This illustrates the importance of using the same bandwidth for calibration and for measurement.

As mentioned above and in Chapter 3, there may be instances where wavelength accuracy is influenced by response nonuniformity (both pixel-to-pixel and within a pixel). For example, some individual defect pixels can show low responsivity and thinned CCDs can suffer from fringing effects in the NIR. These effects can be checked for by examination of the line-spread function (Fig. 4.2-3). If the bandwidth contains several pixels, then this will tend to have a flat top and fluctuations in this will indicate pixel-to-pixel variations in response. The responsivity profile from a continuum source (c.f. Sec. 4.2.2) will also show fluctuations if nonuniformity is present. If an external monochromator can be used to provide an input to the spectrometer under test, then this input can be scanned in wavelength to produce plots as in Figs. 3.1.4-5 and 3.1.4-6, which can reveal both inter-and intrapixel non-uniformity. If any of the above checks reveal the presence of bad defect pixels, then these parts of the array should be avoided when performing wavelength calibration (by omitting the corresponding lines or shifting them to "good" pixels by rotating the grating).

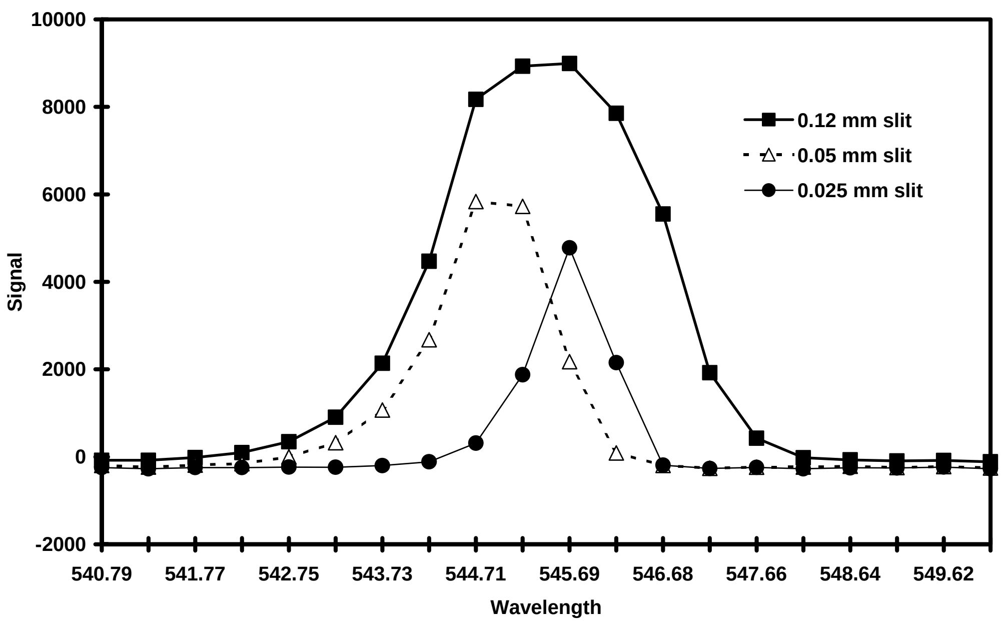


Figure 4.2.2-5 Wavelength calibration of a PDA spectroradiometer using the 546.07-nm mercury line and three different entrance slit widths.

4.2.2.2 Calibration Methods and Artifacts

A range of low-pressure gas discharge lamps (mercury, neon, argon, cadmium, etc.) are available and by careful selection it is usually possible to use a combination of these to provide a number of calibration lines spanning the entire wavelength range of interest. For example, for the calibration of a spectroradiometer operating in the visible region, i.e., 380 nm to 780 nm, a mercury lamp providing lines at 404.7 nm, 435.8, 546.1 nm, 576.96 nm, and 579.07 nm could be used with additional lines at higher wavelengths provided by either a neon or argon discharge lamp (refer to Table 4.2.2-1 for the major peaks). It should be noted that unlike the majority of scanning systems, many array-based systems incorporate fixed second-order filters and, in this case, the use of second-order peaks (e.g., 730.0 nm from second-order 365.0 nm) is therefore often not possible.

The use of more than one source to perform a wavelength calibration may not always be feasible, depending on how the system control software has been written. In such cases it may be possible to use an alternative method, which exploits the ability of most array-based systems to acquire data rapidly. The procedure is to monitor a mercury discharge lamp, or similar, during its warm-up period. Immediately after switch-on, emission lines caused by neon are visible, which quickly disappear as equilibrium is reached inside the lamp, after which the mercury emission lines predominate (see Fig. 4.2.2-6). The neon lines seen during the warm-up period can be used to provide additional reference points for the wavelength calibration.

Use of a discharge lamp or lamps for the wavelength calibration of a spectroradiometer is comparatively straightforward, since the lamp can generally be placed in the test source position without difficulty. However, it is important to ensure that the input optics are overfilled (this should guarantee that the polychromator optics and entrance slit are correctly illuminated) and that an appropriate entrance slit width is used. As discussed earlier, the slit width will affect the bandwidth of the spectroradiometer and hence the number of array elements over which the radiation from the emission line is spread. The optimum arrangement is for several elements to be irradiated, since this allows the line spread function to be resolved and an interpolation of the peak wavelength to be made better than the resolution provided by each array element as shown in Fig. 4.2.2-4. In this case the pixel size/resolution is 0.6 nm and the bandwidth is ~2.2 nm; i.e., about eight data points are contained within the line spread function.

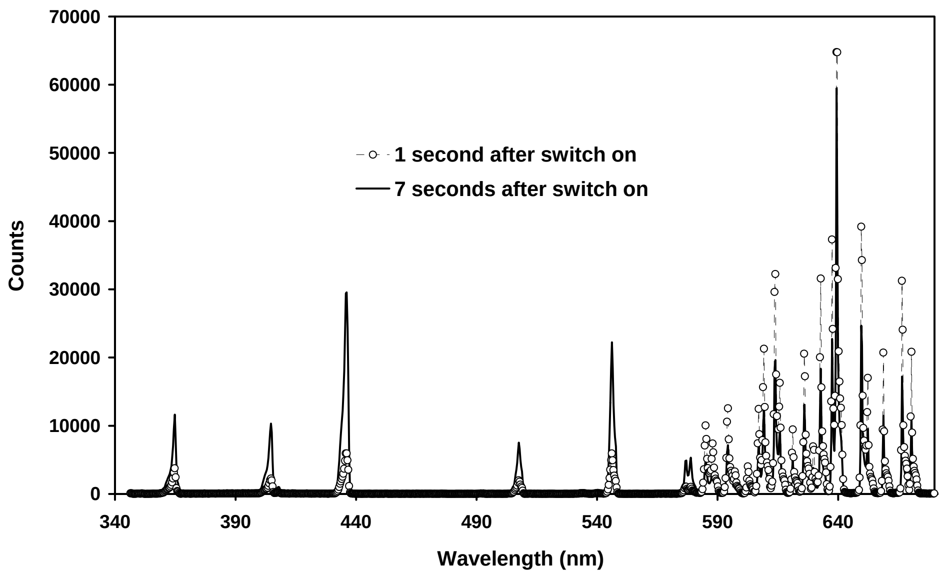


Figure 4.2.2-6 Warm-up of a mercury discharge lamp, monitored using the kinetic function of a 1024 x 512 pixel CCD spectroradiometer.

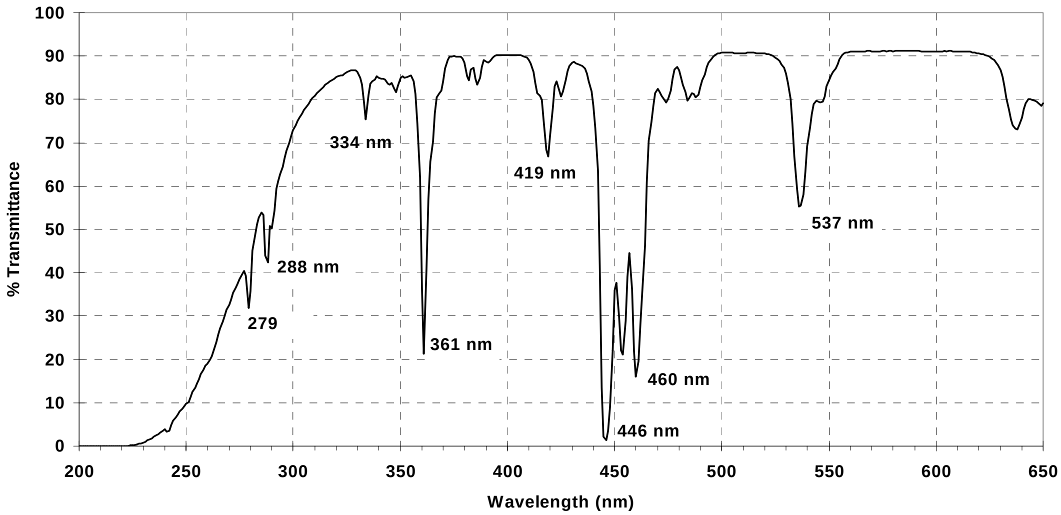


Figure 4.2.2-7 Typical plot of transmittance for a holmium oxide doped filter glass, showing nominal filter peaks.

If it is not possible to place the wavelength calibration lamp in the normal test lamp position (e.g., because the signal levels obtained are too low), then direct illumination of the entrance slit may be used, but with caution. It is essential that the lamp is positioned so that the slit is evenly illuminated and the polychromator optics are overfilled (if this condition is not satisfied, then the wavelength calibration will vary with the lamp position). The lamp must also be situated far enough away from the polychromator to minimize scatter from the slit edges, as discussed in Ref. 33.

Some spectrophotometers have built-in emission line sources (in addition to the broad- band source or sources used for measurements on artifacts) that can be used for wavelength calibration in the same way as described earlier. Any suitable narrowband features superimposed on the continuum output of the broadband source can also be used; e.g., many deuterium lamps have reasonably strong emission lines at 486 nm and 656 nm, as discussed in Ref. 34. In spectrophotometers that do not incorporate a selection of built-in emission line sources, it may be possible to insert them temporarily for the purposes of wavelength calibration (e.g., in the sample compartment[32]).

For many spectrophotometers, however, wavelength calibration using emission line sources is not an option and alternative techniques are necessary, based around the use of reference materials (filter glasses, solutions, ceramic tiles, etc.). It is important to note that in contrast to the emission lines from a low-pressure discharge lamp that always occur at well-documented wavelengths, the precise location and form of the wavelength features of reference materials vary from one artifact to another, and traceable calibration (preferably by an approved laboratory) is therefore essential.

Filters or solutions containing compounds of rare-earth elements exhibit finely structured absorption bands that can be used for wavelength calibration (examples of the use of such filter glasses and solutions is discussed in Ref. 34 and Refs. 35, 36, and 37). Oxides of holmium and didymium (a mixture of praseodymium and neodymium) are the two most commonly used materials and can be incorporated in solutions, filter glass or ceramic tiles. Other rare- earth salts and oxides, as well as solutions of potassium dichromate and benzoic acid are also used.[36] A typical transmittance plot for a holmium oxide filter glass is shown in Fig. 4.2.2-7; the most commonly used calibration features are highlighted. Holmium oxide in solution provides an even larger number of absorption features, as described in Ref. 37 (those at low wavelengths are not apparent in holmium oxide filter glasses due to absorption by the glass itself). The width of the absorption bands of holmium and didymium calibration artifacts can be a limiting factor with high-accuracy work, and an alternative is to use a crystalline material supplied by McCrone Scientific Ltd., which has considerably narrower bands.[38] A typical absorption spectrum for a McCrone filter is shown in Fig. 4.2.2-8.

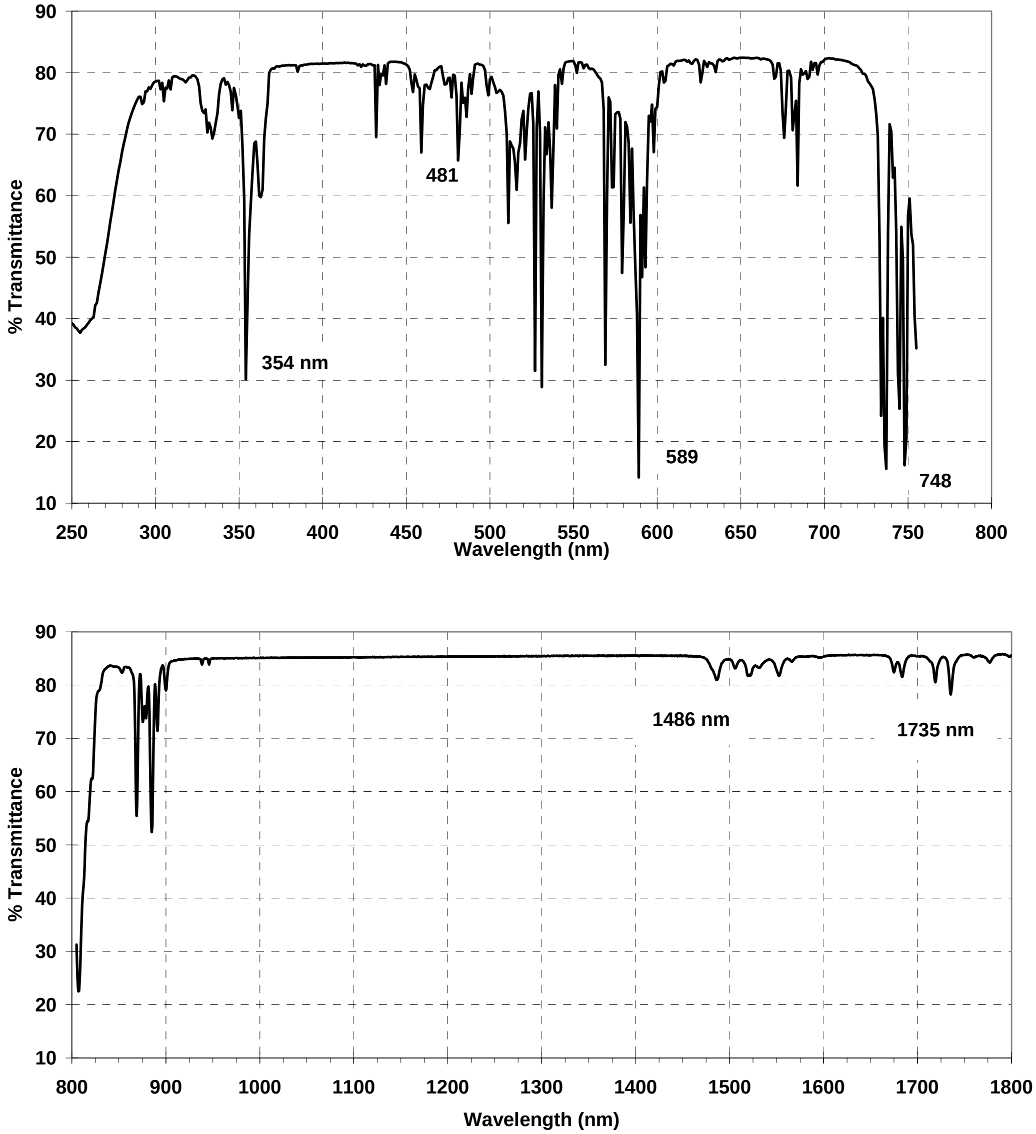


Figure 4.2.2-8 Typical transmittance trace of a McCrone filter showing nominal calibration wavelengths.

It is important to note that the measured wavelength of the absorption features for all such filters and solutions is strongly influenced by the bandwidth at which measurements are made. Calibration laboratories will often provide data for several different bandwidths, enabling users to select the most appropriate for their particular application. Alternatively, the user can request calibration at a specified bandwidth.

These standards should be stable over a period of several years after calibration (Ref. 37 quotes 10 years, but 5 years would be more typical) providing they are stored carefully and in accordance with the calibrating laboratory's guidelines. Studies on the ruggedness of the holmium oxide solution have been conducted[39] by varying temperature of use, purity of holmium oxide, and solution concentration. Although the absolute transmittance/absorption of the standard changed, the peak wavelength values remained within quoted uncertainties.

Another possibility, suggested by Schlemmer and Machler,[40] is to use a thin air gap between two glass plates to give an interference spectrum that gives evenly spaced features throughout the spectrum, whose wavelength can be determined if an approximate calibration (from a few known lines) is available. The order *m* of an extremum (fringe maximum or minimum) closest to one of the calibration lines is determined by

$$m = \frac{\lambda_{m+1}}{\lambda_m - \lambda_{m+1}}, \tag{4.2.2-1}$$

where λ_{m+1} is the wavelength of extremum $m+1$, while its wavelength λ_m is assumed to be true. This value of m is then rounded to the exact integer or half-integer and the wavelengths of the nth extremum can be found from

$$\lambda_n = \frac{m}{n}\lambda_m. \tag{4.2.2-2}$$

In the case of spectrophotometers designed for reflectance measurements, it is often not possible to incorporate either a spectral line source or a filter into the system. The optimum approach to wavelength calibration in this situation is to use a suitable, calibrated, reflectance standard. Ceramic tiles that have been doped with holmium oxide are available (see Ref. 41); a typical reflectance spectrum is shown in Fig. 4.2.2-9. Some broadening and shifting of the absorption features is apparent when compared with those seen in a holmium oxide filter or solution, probably due to the high-temperature processing of the ceramic. Materials in which a range of rare-earth oxides have been sintered into a fluorocarbon matrix have also been produced (described in Ref. 41). One of the most recent of these is a mixture of high concentrations of three rare-earth metal oxides (holmium, erbium, and dysprosium oxides) in a fluorocarbon matrix processed under high pressure but low temperature, which shows little evidence of broadening or shifting of the absorption features. As in the case of filters and solutions, traceable calibration of all such reflectance standards, at an appropriate bandwidth, is essential before they can be used for wavelength calibration.

In situations where a higher degree of uncertainty in the wavelength calibration can be tolerated, calibrated colored ceramic tiles of the type that are widely used to check the performance of reflectance spectrophotometers may be used for wavelength calibration. These tiles have smoothly varying reflectance profiles with no localized absorption features. They can be used to assess wavelength accuracy in two ways:

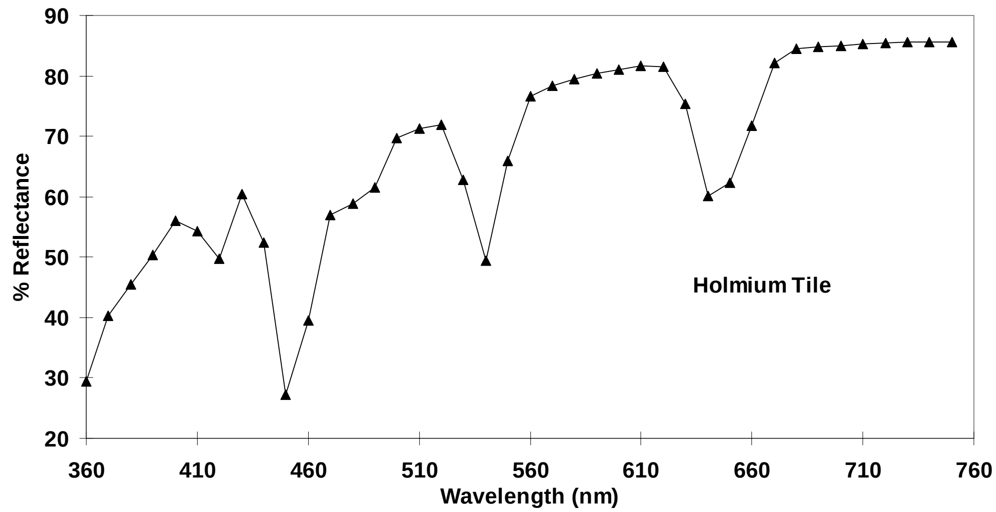


Figure 4.2.2-9 Typical reflectance plot for a ceramic tile doped with holmium oxide, obtained using a 512-pixel spectrophotometer and binning into 10-nm bands.

1. The reflectance curves for two ceramic tiles of complimentary color will typically slope in opposite directions to each other and, if plotted on the same graph, will cross over at a particular wavelength. By comparing the position of this crossover point as measured using the spectrophotometer under test with that obtained by plotting the calibrated reflectance values, the wavelength error in the former can be determined. An example of the results from such a comparison is given in Fig. 4.2.2-10; in this case the agreement is within 0.3 nm. Such crossover pairs ideally need to have steep but smooth slopes, exhibit negligible thermochromism (see Ref. 42), and be available in a wide range of colors so as to provide crossover points across the whole of the visible spectrum. All of these requirements have so far been difficult to achieve, but research is ongoing (at NPL).
2. For any calibrated tile, it is possible to plot the calibrated reflectance values against wavelength, and from this plot to read off the wavelength associated with any particular reflectance value. If the calibrated values are compared with the results obtained from the spectrophotometer under investigation, it is possible to make an assessment of the approximate magnitude of any wavelength errors. However, the uncertainty associated with such an assessment is relatively high and is further increased by factors such as thermochromism of the tiles and the performance of the spectrophotometer in other aspects such as linearity. This method is therefore only suitable for the identification of gross wavelength errors.

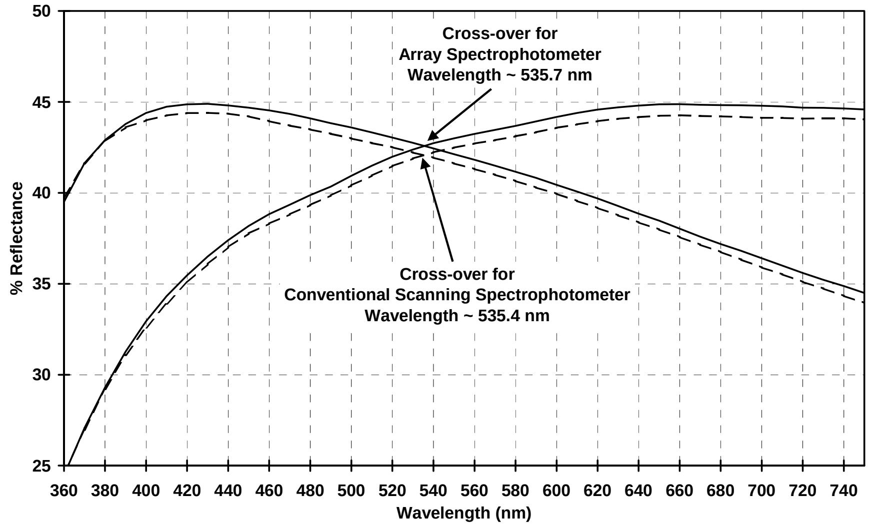


Figure 4.2.2-10 Use of cross-over tiles for wavelength calibration.

4.2.2.3 Recommendations

1. A number of calibration features (i.e., emission lines, transmittance or reflectance peaks) spanning the entire wavelength range of interest should be used.
2. A cubic spline fitting algorithm generally gives an acceptable fit to the wavelength scale; but, if possible, several algorithms should be tested and the best selected on the basis of the results obtained.
3. Determination of the position of the peak wavelength by use of the FWHM or center of gravity methods should give acceptable results in most cases.
4. Where possible, the entrance slit width used for the wavelength calibration should be sufficient that each calibration feature covers several elements in the array. Furthermore, the bandwidth used during routine measurements using the system should be matched to that used for the wavelength calibration or vice versa.
5. Where possible, the array output should be checked for nonuniformities. It may be necessary to exclude regions near bad pixels from the wavelength calibration.
6. Wavelength calibration should be performed using emission lines from low- pressure discharge lamps wherever possible. Where this is not feasible (as is often the case with spectrophotometers), glass filters or solutions providing relatively narrow absorption features should be used. For reflectance spectrophotometers, ceramic tiles or sintered materials, which have been doped with rare-earth metal oxides, may be used.
7. The emission lines from a low-pressure discharge lamp always occur at well-documented wavelengths; calibration of such lamps is therefore not necessary.
8. When using a low-pressure discharge lamp for wavelength calibration, it is essential that the spectrometer optics, including the polychromator, are overfilled and

uniformly illuminated. Ideally, the lamp is placed at the normal test lamp position; direct illumination of the slit should only be used if absolutely essential; and if this approach is taken, it is crucial to ensure that the polychromator optics are overfilled.

9. Filters, solutions or reflectance standards must be traceably calibrated and care taken to maintain cleanliness. The calibration bandwidth should be the same as that for the user's system.

4.2.3 Responsivity

4.2.3.1 Introduction

As discussed in Sec. 1.2.15, responsivity is the relationship between the output from a detector and the input optical signal, and is usually expressed in A/W, V/(J/m^2), ADU/(J/m^2), etc. Absolute spectral responsivity is the relationship between the output from a detector and the input optical signal at a particular wavelength. Relative spectral responsivity is the variation of responsivity with wavelength, normalized at a specified wavelength, and is by definition dimensionless.

For a spectrometer the responsivity will depend not only on the properties of the photosensitive element(s) but also on the characteristics of the polychromator and any other optical components, including the input optics (if any). A calibration of the spectral responsivity of the complete system can provide a useful diagnostic check (e.g., to identify any bad pixels) and will also be required whenever it is necessary to know the relative amounts of radiation present in individual wavebands across the spectrum or the absolute amount of radiation in one or more wavebands. Routine spectral responsivity calibration is not necessary if the system is only to be used to locate the position of spectral emission features (e.g., in many spectroscopic applications) or to perform ratios between measurements made with and without a sample in the beam (i.e. spectrophotometric measurements). Thus, in practice it is usually only spectroradiometer systems that require a spectral responsivity calibration to be performed on a regular basis, and the remainder of this section will therefore deal exclusively with these. Note that in this context we take spectroradiometry to include those spectroscopic measurements on sources where it is necessary to know the relative intensities of spectral features as well as their wavelengths (and this will include many atomic emission, fluorescence, and Raman spectroscopy applications).

Spectroradiometers should ideally be calibrated for spectral responsivity on each occasion of use; in this way a history of the system reproducibility and drift is built up. If the magnitude of the system drift is negligible compared to the acceptable error level for the task at hand, then a less-frequent calibration regime can be invoked. Since most array spectroradiometers have few, if any, moving parts, they tend to be more reproducible than many conventional systems. Less-frequent calibrations can therefore usually be accepted; but for a new system daily or weekly checks may initially be appropriate until acceptable performance has been proved.

The measurement system may also experience short-term drift during a batch of measurements due, for example, to changes in temperature. This is most easily monitored by measuring the spectrum of a stable source at the start, during and at the end of the measurement schedule, which will give an indication of the short-term system stability or repeatability. The source used for this monitoring should be stable within known limits, but does not necessarily have to be calibrated.

4.2.3.2 Input Optics

The purpose of the input optics is to direct optical radiation from the source into the spectrometer and thence onto the diffraction grating (or prism), and to ensure that the radiation follows the same path through the spectrometer and suffers the same attenuation, regardless of the size, shape, polarization, etc., of the source being measured. The latter requirement arises because

- the transmission of a polychromator is dependent on the way in which it is illuminated;
- polychromators typically show a strong polarization dependence (as much as 40%);
- in an array system, the responsivity of the pixels is often nonuniform.

Thus, the system responsivity will vary depending on how it is illuminated, and constant responsivity can only be achieved if the illumination conditions are also constant.

Diffusing input optics are often used when the array detector system is configured to measure spectral irradiance or spectral radiant intensity. These two quantities are essentially the same as far as the input optics required are concerned. The spectral radiant intensity is the spectral radiant flux per steradian emitted by the source in a specified direction. The spectral irradiance at a surface is the spectral radiant flux per unit area incident upon that surface. Knowing the distance between the source and the surface being irradiated enables the conversion of one quantity to the other.

$$\text{Spectral radiant intensity: } I_{e,\lambda} = \frac{d\phi_e}{d\Omega} \quad \text{W sr}^{-1}\text{ nm}^{-1}, \qquad (4.2.3\text{-}1)$$

where $\phi_{e,\lambda}$ is spectral radiant flux and Ω is a solid angle.

$$\text{Spectral irradiance: } E_{e,\lambda} = \frac{I_{e,\lambda}\cos\theta}{d^2} \quad \text{W m}^{-2}\text{ nm}^{-1}, \qquad (4.2.3\text{-}2)$$

where θ is the angle between the normal to the surface and the line joining the source to the surface, and d is the distance between the source and the irradiated surface.

Due to the small size of array detector systems and the need for portability, the most suitable diffusing optic is an integrating sphere. This should ensure that optical radiation from a reference and test artifact follows the same optical path through the system and that all radiation entering the polychromator will have been subjected to at least two reflections within the integrating sphere, thereby depolarizing it. Care needs to be taken that the radiation from the sphere does not grossly overfill the polychromator, since this will cause significant spread of the radiation around the edges of the optical components within the polychromator and thus increase the stray light. Use of a diffusing plaque (e.g., a matte ceramic tile) is also possible.

Imaging input optics are normally used when the system is set up to measure radiance. Spectral radiance for a given line of sight is defined as the spectral radiant flux per unit projected area of the source in the given direction, emitted into unit solid angle. Equivalently, it is the spectral radiant intensity per unit projected area of the source in the given direction

$$L_{e,\lambda} = \frac{dI_{e,\lambda}}{dA \cdot \cos\theta} \text{ W sr}^{-1}\text{ m}^{-2}\text{ nm}^{-1}, \tag{4.2.3-3}$$

where θ is the angle between the normal to the surface and the direction being viewed, $I_{e,\lambda}$ is the radiant intensity, and A is the area.

The imaging optics usually take the form of lenses or mirrors. Lenses suffer from chromatic aberration and are therefore generally only used for short spectral ranges. The use of mirrors, on the other hand, can introduce astigmatism due to off-axis imaging, and this can introduce errors when measuring nonuniform sources. Radiance-measuring spectroradiometers often have an eyepiece to allow accurate alignment of the measuring system with the source. The optical imaging system must be provided with an aperture to isolate the required area of the source, which is frequently the spectrometer entrance slit.

For radiance sources that show nonuniformities within the imaged area or are strongly polarized, it may be necessary to use the imaging optics in conjunction with a diffuser arrangement. In this case, the selected area of the source is imaged onto the entrance to the integrating sphere or diffusing plaque, and the latter then irradiates the spectrometer, as described previously.

4.2.3.3 Calibration Methods and Artifacts

In terms of spectral responsivity calibration, array systems pose no particular difficulties beyond those encountered when calibrating conventional spectrometer systems. Thus, the artifacts already available as spectral responsivity transfer standards will generally also be suitable for use with array systems. For example, a stable and reproducible light source provides an excellent calibration standard for determining the absolute or relative spectral responsivity and long-term reproducibility of a spectroradiometer, regardless of whether this is a scanning or detector array type.

Further details on the requirements for and selection of continuum sources for responsivity calibration are given in Sec. 5.1.1.

4.2.3.4 An Example of a Spectral Responsivity Calibration

Consider an array-based spectroradiometer system that is to be used for spectral irradiance measurements on a group of three test sources. It is assumed that a suitable spectral irradiance transfer standard has been selected following the guidance given in Sec. 5.1.1. An example of the measurement schedule might be:

$$\text{Standard1} - \text{Test1} - \text{Test2} - \text{Test3} - \text{Standard} - \text{Test3} - \text{Test2} - \text{Test1} - \text{Standard1} \qquad (4.2.3\text{-}4)$$

The system should be used to measure the output of the spectral irradiance transfer standards in exactly the same way as they were calibrated. For each transfer standard, the quotient of the calibrated irradiance value for a given waveband by the system "counts" recorded by the pixel corresponding to that waveband gives the system responsivity factor for that waveband:

$$\text{Responsivity} = \frac{[\text{Standard}]_{\text{irr}}}{[\text{Standard}]_{\text{counts}}}, \qquad (4.2.3\text{-}5)$$

where $[\text{Standard}]_{\text{irr}}$ is the calibrated irradiance value for the transfer standard and $[\text{Standard}]_{\text{counts}}$ is the number of system counts for the transfer standard. This is used to convert the system counts for the test source into spectral irradiance, for that particular waveband, as follows :

$$[\text{Test}]_{\text{irr}} = \text{Responsivity} \times [\text{Test}]_{\text{counts}}, \qquad (4.2.3\text{-}6)$$

where $[\text{Test}]_{\text{irr}}$ is the calibrated irradiance value for the test source, and $[\text{Test}]_{\text{counts}}$ is the number of system counts for the test source.

In practice, most array-based spectroradiometers will only allow the user to calculate the spectral responsivity of the system from a single calibration measurement. The software will then use this spectral responsivity to calculate the spectral irradiance for all subsequent measurements. In this case, repeat measurements using the transfer standard(s) can be used to assess system stability. It is recommended that at least one such repeat measurement should be made, preferably at the end of a sequence of test lamp measurements. At least two measurements should be made for each test lamp—these can then be averaged to give the final result for that particular lamp and also provide valuable information on the measurement repeatability.

Note: The user usually specifies the conditions (with advice from the calibrating laboratory) under which the transfer standard is calibrated; e.g., for an irradiance standard source the following would usually be specified:

alignment
distance
warm-up time
operating current or correlated color temperature
polarity
(wavelength) sampling interval
bandwidth

These parameters should be quoted explicitly in the certificate of calibration from the calibration laboratory, thus providing unambiguous instructions to any user of the transfer standard.

Note that use of the calibration reference standard at the beginning and end of a series of measurements as described above allows the short-term drift or repeatability of the system to be monitored and noted, which is an important consideration when calculating the overall uncertainty of the measurement.

4.2.3.5 Recommendations

1. Aside from an initial or diagnostic check on an instrument, spectral responsivity calibration is generally only necessary for spectroradiometer systems or other applications where it is necessary to measure the relative amounts of radiation present at each individual waveband across the spectrum or the absolute amount of radiation at one or more wavebands.
2. Suitable input optics are essential. For irradiance measurements or radiance measurements on nonuniform or polarized sources, a diffuser arrangement should be used. For radiance measurements on diffuse sources, direct imaging of the source onto the entrance slit is acceptable.
3. A group of suitably calibrated sources should be selected to perform the calibration; these standard sources should be stable, reproducible, rugged, and user friendly, and the calibration should be traceable to national standards.
4. Wherever possible, the selected calibration sources should be similar in terms of size, shape, and spectral distribution to the types of source to be measured. Where this is not possible or appropriate, specially designed and calibrated tungsten lamps (for the visible and near-infrared) or deuterium lamps (for the UV) are usually a good choice.
5. The short-term reproducibility (repeatability) of the system should be assessed by measuring the spectrum of a stable lamp (not necessarily calibrated) at the start, during, and at the end of a batch of measurements to give an indication of system drift.
6. The spectral responsivity of the system should be calibrated on each occasion of use until a sufficiently long history of long-term reproducibility exists. After this, the system need only be calibrated often enough to stay within the required error limits.

4.2.4 Stray Light

4.2.4.1 Introduction

Stray light is a general term used in spectrometry to describe the level of optical radiation measured at a given wavelength position that reaches the detector by other than the direct path from the artifact under evaluation to the detector, or is of a different wavelength to that associated with the wavelength position. There are two basic types:

1. external or homochromic stray light—caused by, for example, scattering around the edges of a filter or reflections off walls and ceilings;
2. In-system or heterochromic stray light—caused by scattering and reflections within the spectrometer.

External stray light is a potential problem in all optical radiation measurements and is no more, or less, acute with diode array systems than with traditional systems. It can be minimized by the use of appropriate baffles and by coating surfaces with nonreflecting paint; general principles are discussed in Refs. 43 and 44, and methods for assessing levels of external stray light are presented in Sec. 4.2.6.2.

The problem of in-system stray light, on the other hand, is significantly more acute with array systems than with more traditional approaches to spectrometry, for the following reasons:

- Array systems usually use a single polychromator to disperse the radiation across the array, whereas the majority of high-quality traditional spectrometers use a double polychromator (which has much better stray light rejection) and mechanically scan through the spectrum by rotating the polychromator gratings.
- Radiation can be reflected (or inter-reflected) off the array, onto the walls and/or other components within the polychromator and then back onto the array. Reflections from detectors placed after an exit slit (as in scanning systems) are much less likely to re-enter the polychromator and be re-reflected onto the detector again.
- The attraction of many array systems lies in their small size and portability; it is much more difficult to make effective use of baffles in a physically small system.
- Radiation must be spread across the whole array, making the placement of baffles within the polychromator more difficult than in the situation where only the narrow exit slit is to be irradiated.
- Most array spectrometers use silicon detectors, which have high responsivity in the red and near infrared (c.f. Fig. 2.1-4) and much lower response in the blue and ultraviolet spectral regions. This makes them highly sensitive to any stray radiation in the red and near infrared, which is also the region of highest emission from many commonly used sources, such as tungsten and xenon lamps or daylight.

As a result of these problems, in-system stray light is a major factor influencing the performance of array systems for spectrometry and, in the blue region in particular, this can dominate all other sources of measurement uncertainty. Accordingly, some instruments include *stray light filters*, either fitted permanently in front of the array or available for insertion by the user. In general it is easier to manufacture a filter with a sharp transmission cut-on (at short wavelengths) than it is to have a sharp cut-off at long wavelengths. Hence, the majority of stray light filters tend to transmit radiation only above a certain wavelength and are therefore suitable for reducing, or even eliminating, stray light arising from wavelengths below the cut-on wavelength. It should be noted, however, that it is the longer wavelength region that is often more of a problem, since not only is the responsivity of silicon higher in the red than in the blue (making most systems very susceptible to stray radiation at longer wavelengths) but also many sources (for example, tungsten lamps and daylight) emit proportionately more radiation at longer wavelengths. Recent filter developments may address this problem. For example, dielectric short-pass filters with a reasonably sharp cut-off wavelength are available for a limited selection of transmittance ranges, although care must be taken with these since they are not absorptive but act by reflecting back the unwanted radiation (potentially contributing further to the stray light); and also they do not have a flat transmittance over the pass band but tend to show ripples at the ~ 5% level. Simple colored glass filters or heat-reflecting ("cold") mirrors can also be used to limit the amount of red/NIR radiation in some cases. Wedge interference filters are also starting to become available that can be placed directly in front of the array. These can be used to give a narrow bandpass that varies across the array (in the dispersion direction) and are very effective in limiting stray light, in principle giving the level of performance previously only achievable with a double monochromator system.

If stray light filters are supplied, it is strongly recommended that these be used. It should be noted that the second-order filters often incorporated into array-based systems are actually a special form of stray light filter and must be used to prevent, for example, second-order 400-nm radiation from being measured at 800 nm.

Beyond the use of stray light filters as appropriate, there is generally nothing the user can do to improve the stray light performance of a system. Furthermore, the amount of stray light at any selected wavelength will vary according to the spectral distribution of the radiation entering the polychromator, thus making stray light errors almost impossible to correct for numerically. If an instrument shows significant levels of stray light during the tests described below, the optimum approach is to restrict its use to those regions and/or types of source for which acceptably low stray light is seen. Alternatively, in situations where only a limited spectral region is of interest, it may be possible to use a broadband interference filter, a second polychromator with a wide bandwidth (as described in Ref. 42), or similar, to reduce the level of “unwanted” radiation entering the system. If neither of these approaches is possible, an allowance for potential levels of stray light should be made in the uncertainty budget for the instrument, taking into account the spectral distribution of the incident radiation.

4.2.4.2 Calibration Methods and Artifacts

The most common method by which to assess in-system stray light is to irradiate the instrument with a monochromatic source (typically a laser, although line sources or even an auxiliary monochromator are sometimes used, as described in Ref. 45) and measure the signal recorded at wavelengths away from the source wavelength. This signal, expressed as a percentage of the peak of the source signal, is taken to be representative of the in-system stray light. Even with scanning systems this generally gives an artificially low value; with array systems the results can be almost meaningless.

A better approach is to use a series of edge filters or notch filters, and this is highly recommended for array systems. An edge filter transmits radiation only above a certain wavelength, whereas a notch filter transmits radiation over a wide wavelength range and only absorbs over a narrow wavelength range. The important feature of both types of filter is that for a certain wavelength range the transmission is near zero. The procedure for using the filters is as follows:

1. Using a broadband continuum source (e.g. a tungsten lamp) and no filter, make a measurement covering the entire wavelength range of interest. In the case of a spectroradiometer, the results $X(\lambda)$ will be in units of W m^{-2} nm^{-1} or similar; whereas for a spectrophotometer they will be expressed by percent transmission or similar.

2. Repeat the measurement with the edge or notch filter placed between the source and the entrance slit. (The optimum placement of the filter is discussed in more detail on page 146.) Record the result, $Y(\lambda)$.
3. Determine the ratio between the two measurements, $Z(\lambda) = Y(\lambda) / X(\lambda)$. (Note that $Z(\lambda)$ is essentially the filter transmittance as measured using the system under test).
4. For an edge filter, $Z(\lambda)$ should be zero for wavelengths below the cut-on wavelength of the filter. For a notch filter, the value of $Z(\lambda)$ at the notch wavelength should be zero. Any nonzero signal in these regions is likely to indicate the presence of in-system stray light, although some care has to be taken in interpreting the results, as described below.
5. Repeat the above for a sufficient number of filters to enable an assessment of stray light performance across the whole region of interest to be made—performance at one particular wavelength cannot be taken as representative of the whole.

When analyzing the results from stage 4, it is important to remember that the stray light at a given wavelength depends not only on the properties of the instrument itself, but also on the broadband source used (since this determines the relative proportions of “wanted” and “unwanted” radiation at each wavelength). The most appropriate choice of source will depend on the application. For a spectrophotometer, the instrument will generally automatically select the source for a given wavelength range, and this is then clearly the source that should be used for the stray light assessment in that wavelength range. With a spectroradiometer, however, the situation is not so clear-cut. The source chosen for the stray light tests should be selected to be similar to the types of source that are to be measured, but also needs to have good stability over the period of the tests and adequate spectral output throughout the spectral region in which measurements are made. In general, the optimum choice is therefore a tungsten lamp for measurements in the visible or a deuterium lamp for measurements in the UV. For measurements in the long-wave UV region (about 300 to 400 nm) where signal levels from both sources are usually fairly similar, the recommended choice would be a tungsten lamp if the test sources have strong emission in the visible (e.g., daylight) or a deuterium lamp for sources which emit more strongly in the UV (e.g., lamps used for water sterilization).

Edge filters with a range of cut-on wavelengths are widely available at a reasonable cost, and they are therefore often the preferred choice for stray light evaluation. An example of stray light measurements on a spectroradiometer using a series of edge filters is shown in Figs. 4.2.4-1 to 4.2.4-4. Here, three filters (GG495, OG2 and RG665) have been used, which cut-on at about 450 nm, 500 nm, and 600 nm, respectively (note that the cut-on wavelength should be taken to be just below any tail region). Two different broad-band sources have been used: a tungsten lamp (Figs. 4.2.4-1 and -2) and a deuterium lamp (Figs. 4.2.4-3 and -4). From the results it is apparent that this system shows significant stray light with the tungsten lamp at wavelengths below 450 nm and that either (a) its use with such sources should be restricted to longer wavelengths, or (b) the uncertainty budget for measurements below 450 nm should allow for an appropriate contribution from stray light. The low signal level below about 300 nm with the tungsten lamp means that a meaningful assessment of stray light levels in this region cannot be made with this source—the deuterium lamp provides better information. Note the significant difference in the measured stray light in the 300 to 400 nm region using the two sources; at 350 nm, for example, the assessment using the tungsten lamp yields a value of about 15%, whereas with the deuterium lamp the value is zero. This indicates that it is radiation from the visible region of the spectrum that is responsible for the out-of-band radiation at this wavelength.

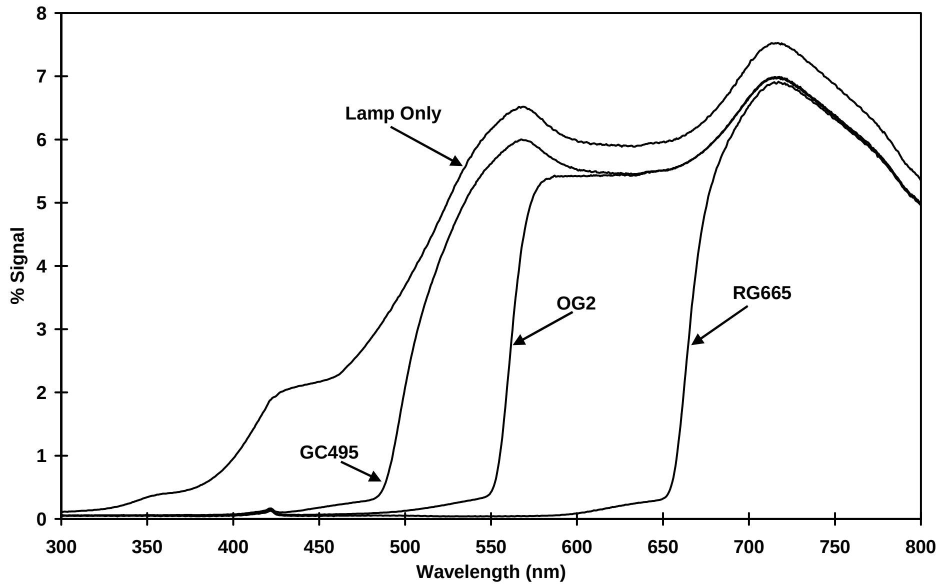


Figure 4.2.4-1 Signal recorded from 1024-pixel spectroradiometer (with integrating sphere) using a tungsten lamp and edge filters.

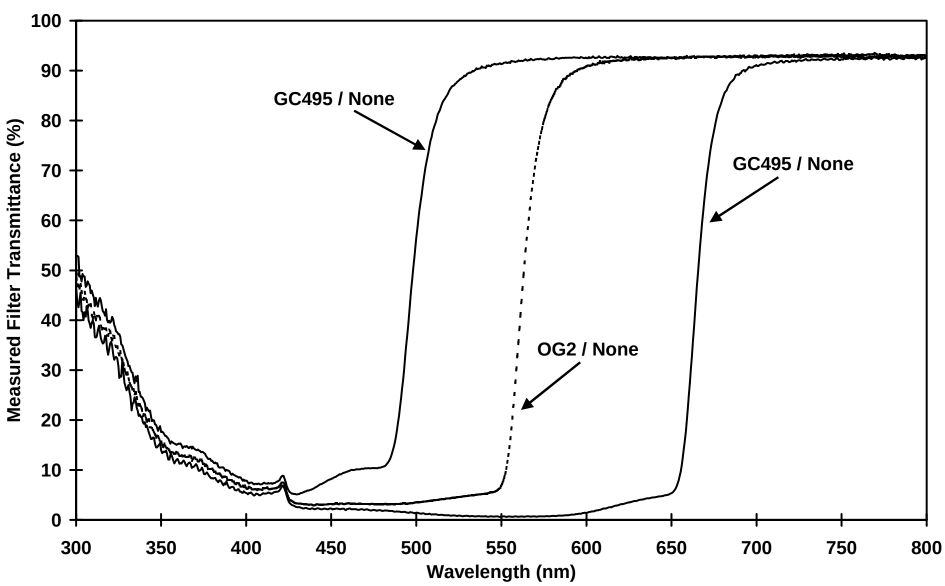


Figure 4.2.4-2 Measured transmittance of edge filters, calculated from results in Fig. 4.2.4-1.

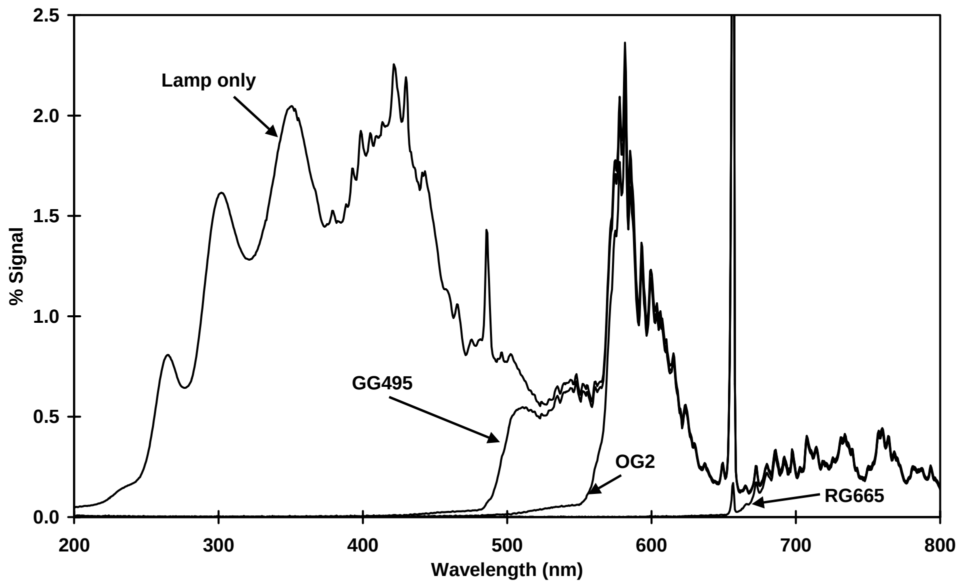


Figure 4.2.4-3 Signal recorded from spectroradiometer using a deuterium lamp and edge filters.

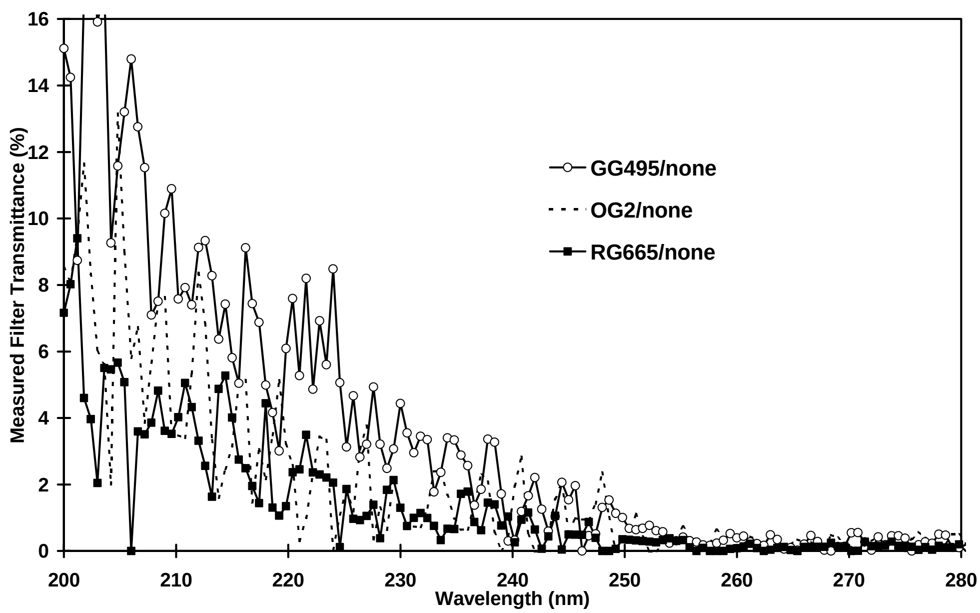
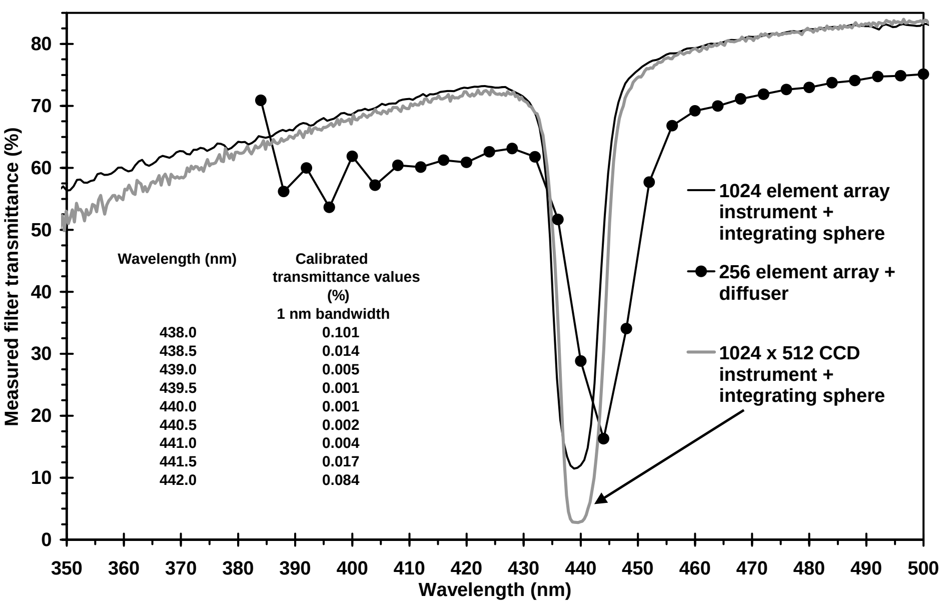


Figure 4.2.4-4 Measured transmittance of edge filters, calculated from results in Fig. 4.2.4-3.

An example of the use of a notch filter is given in Fig. 4.2.4-5; in this case, the notch wavelength is about 440 nm. Notch filters can offer a more precise stray light determination than edge filters, particularly if the transmittance of the filter has been calibrated, but they are relatively expensive. In addition, each filter provides information on stray light at only one wavelength, and a large number of filters must therefore be used in order to assess performance throughout the spectral region of interest. Figure 4.2.4-5 shows the notch filter transmittance values of the trough measured using a high-quality scanning spectrophotometer (i.e., the calibrated values) and also shows the results of measurements using this filter on three different array spectroradiometers. When analyzing such results, it should be noted that the bandwidth used can significantly influence the measured minimum transmittance, and the bandwidth used for the tests should therefore be matched to that of the calibration, or vice versa. In the example illustrated, the minimum transmittance has been calibrated for various bandwidth settings (the 1-nm bandwidth data is tabulated in Fig. 4.2.4-5); e.g. with a 1-nm, 2-nm, and 8-nm bandwidth the minimum transmittance is 0.001%, 0.004%, and 15.2%, respectively. For the three-array systems assessed, the minimum measured transmittance varies from about 3% (the bandwidth for this system is about 2 nm) to about 16% (system bandwidth of about 8 nm). It is apparent that even allowing for bandwidth effects, in all three systems significant out-of-band radiation is reaching the array detector. It is also interesting to note the apparent shift in the notch wavelength on one of the systems, which is due to an uncorrected wavelength error; hence, notch filters can provide another wavelength scale check point, *provided a traceable transmittance calibration of the filter has been performed*.

As mentioned earlier, the correct placement of an edge or notch filter for stray light tests is between the source and the entrance slit. With a spectroradiometer this normally presents no problem; there will usually be sufficient room for the filter either immediately in front of the integrating sphere (if there is one), or close to the lamp or the entrance slit.

With a transmittance spectrophotometer, the simplest approach is to place the filter in the normal sample position. However, the sample holders in many such instruments are designed only for samples in cuvette form and will not accept standard format (glass) edge or notch filters.



Wavelength (nm)	Calibrated transmittance values (%) 1 nm bandwidth
438.0	0.101
438.5	0.014
439.0	0.005
439.5	0.001
440.0	0.001
440.5	0.002
441.0	0.004
441.5	0.017
442.0	0.084

Figure 4.2.4-5 Example of stray light check on three spectroradiometers using a notch filter.

In this case, an alternative is to use liquid edge filters, as described in Refs. 46, 47, and 48. A detailed method for measuring the *stray radiant power ratio* using these filters is described in Ref. 49, which also gives the chemical composition of a range of liquid filters suitable for stray light evaluation throughout the optical spectral region. A similar but simplified method is discussed in Ref. 50.

The design of reflectance spectrophotometers can make it very difficult to perform stray light tests—it is usually not easy to gain access to the source, so that it is not possible to place an edge or notch filter in this position, and the geometry of the sample compartment and collection optics generally also leaves little room for the insertion of a filter. However, with care it is normally possible to find a suitable location; e.g., most systems for diffuse reflectance measurements include an integrating sphere as part of the collection optics, and in these it is often possible to mount an edge filter against the sphere exit port, between the sphere and the detector. Some systems incorporate internal edge filters (usually intended to remove lower wavelength radiation which may cause fluorescence) that can be used for stray light checks. However, care is needed when using these, since in some cases the software is designed to return zero readings below the cut-on wavelength of the filter regardless of the actual signal recorded. Figure 4.2.4-6 shows the results of measurements on a reflectance spectrophotometer using a series of edge filters mounted against the sphere exit port. In this example there is evidence of stray light at short wavelengths when using the GG475 filter.

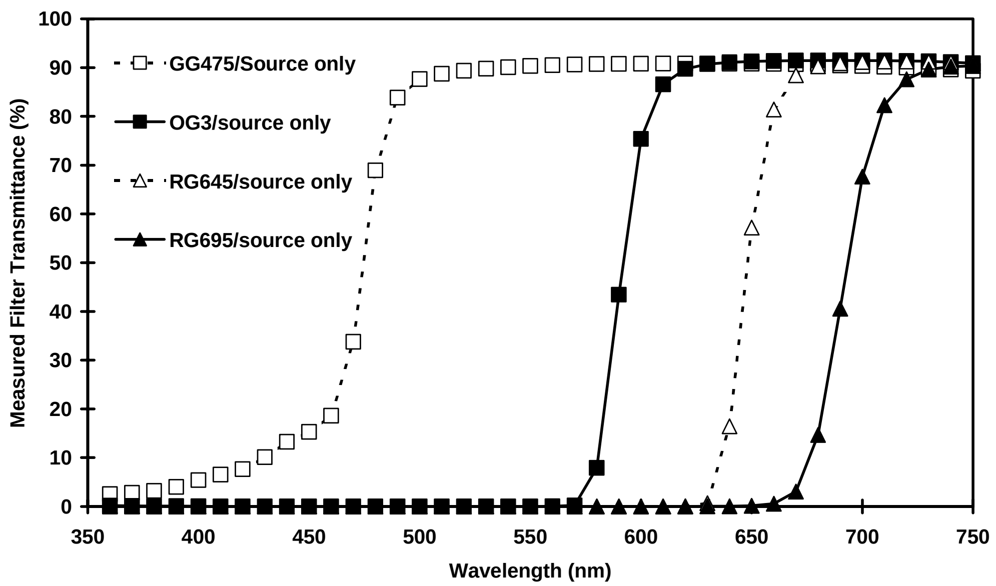


Figure 4.2.4-6 Example of stray light checks on a spectrophotometer using an edge filter.

4.2.4.3 Recommendations

1. A series of edge or notch filters should be used spanning the wavelength range of interest.

2. Calibration of the transmittance profile of edge filters is not necessary. Notch filters should ideally be calibrated, although for many purposes it is sufficient to assume zero transmittance at the notch wavelength (this may give a slight over-estimate of the level of stray light).
3. When using calibrated notch filters, the calibration bandwidth and the measurement bandwidth should be the same.
4. The filters should be placed in a convenient position between the source and the entrance slit. This may not always be easy, but is almost always possible.
5. The source used for stray light checks will affect the results obtained. For spectrophotometry the source will depend on the wavelength and will usually be selected by the instrument software. For a spectroradiometer, a tungsten or deuterium lamp is usually the best choice, with the final selection being determined by the wavelength range and the types of test source to be measured.
6. Stray light errors are almost impossible to correct. If an instrument shows significant levels of stray light during the above tests, the optimal approach is to restrict its use to those regions and/or sources for which acceptably low stray light is seen. The use of a stray light filter may be possible in some cases. If neither approach is feasible, an allowance for potential levels of stray light should be made in the uncertainty budget for the instrument.

4.2.5 Linearity

4.2.5.1 Introduction

A system is linear if the output varies in direct proportion to the input or, put another way, if the responsivity of the system is constant as the input is varied. If a system is nonlinear, a straightforward relationship between the measured signal and the input quantity does not exist and measurements become more difficult to interpret. Nonlinearities may arise due to the characteristics of the individual pixels in the array or because of imperfections in the amplifiers or other electronics, so it is important to check the performance of the system as a whole. In most cases the system will be linear within a certain range of operating conditions, but will become nonlinear if these conditions are exceeded.

Manufacturers frequently state that their systems are “linear” without further qualification of this statement and often any linearity checks that are performed relate only to one element, such as the electronics, rather than the complete system. For example, a known electrical signal may be applied to the detector electronics and a record made of the number of counts produced. The deviation from a straight line fit of this data is then quoted as the linearity of the system; in fact, it is the linearity of the detector electronics alone and the system itself may show very different linearity characteristics.

Although the need to assess linearity is common to all spectrometers, some of the features of array systems make this assessment particularly difficult:

- With most array spectrometers the signal is integrated over a period of time and then processed before being presented to the user. It can be difficult to gain access to the raw (unprocessed) data and determine the actual measured signal levels.
- Some detectors behave nonlinearly at high incident power levels. For example, silicon photodiodes operating in the near-infrared region can show an output that increases more rapidly than expected with increasing incident power level, as described in Ref. 51, and saturation effects (a fall-off in response) can be seen at high signal levels at all wavelengths with most detectors. Linearity tests on a conventional spectrometer can usually allow the user to set an upper acceptable signal level at any particular wavelength, below which the system will be linear. With an array spectrometer, however, the same integrated signal may be obtained, for example, by measurement of incident power P for a time T or power $2P$ for time $0.5T$. It is not possible to set an integrated signal level below which the system will behave linearly without also specifying the integration time for which this applies. Consequently, it is necessary to assess performance for a range of integration times and at varying incident power levels in order to determine safe operating limits.
- When measuring time-varying signals, the peak power can similarly affect the linearity, but this peak power is often difficult to determine because of the signal integration. The problem is particularly acute when measuring a signal that varies over orders of magnitude in a short space of time, such as the radiation produced during one pulse of a flash gun, but it can also be a factor when measuring other devices, e.g., a fluorescent lamp.
- In many array spectrometer systems the integration time is set automatically by the software and cannot be controlled by the user. In this case, it may be impossible for the incident power level to be determined and the results of linearity checks may therefore be especially difficult to interpret.

The linearity tests described below are intended primarily to determine whether the system behaves linearly within a defined range of operating conditions, but they may also yield information about the input–output characteristics of the system that can sometimes be used to correct for nonlinear behavior. This approach is not recommended with array spectrometers because of the problems associated with gaining access to the unprocessed output signal.

4.2.5.2 Calibration methods and artifacts

A wide range of techniques have been developed for the assessment of the linearity of detectors and traditional types of spectrometer (see also Sec. 1.2.11). These methods are generally also applicable to array spectrometers, but some are especially suited to the investigation of particular aspects of array systems (such as the effect of integration time), while for others additional precautions and care in interpretation of the results may be necessary. This section will therefore concentrate on those issues relating to linearity measurements on array spectrometers; the methods themselves will be described in outline only, and readers are referred to the many texts and papers on this subject for further details (see, for example, Refs. 42, 43, 52, and 53 for excellent summaries of the most widely used techniques).

Flux superposition methods

The most fundamental method by which to determine the linearity of a system is by flux superposition. There are many variations on this theme, but the basic principle is the same for each: if two fluxes A and B are incident on a detector first separately and then together, and the outputs Y_A, Y_B, and Y_{A+B} are recorded for each condition, then for a device that is linear between the minimum and maximum of these three signals, the following relationship will hold:

$$\text{Linearity factor} = Y_{A+B} / (Y_A + Y_B) = 1. \qquad (4.2.5\text{-}1)$$

Alternatively, the nonlinearity between these limits is given by

$$\text{Nonlinearity (\%)} = 100 \times [Y_{A+B} / (Y_A + Y_B) - 1]. \qquad (4.2.5\text{-}2)$$

(A positive value for the nonlinearity means the output increases more rapidly than expected as the input increases.) Some of the most widely used manifestations of this technique are

a) The *double aperture* method (see Refs. 42 and 54, for example). Here two apertures are placed adjacent to each other between the source and detector. The two apertures can be closed by means of shutters, enabling the signals for aperture A only, aperture B only, and both apertures together to be recorded. Generally the two apertures are of equal size, in which case this method allows the linearity over a 2:1 step to be assessed. The addition of stepped attenuators, such as neutral density filters with transmittances of 50%, 25%, 12.5%, etc., allows cascading of these 2:1 steps to cover a wide range of signal levels.
b) The *multiple aperture* method (see Refs. 55 and 56). A series of apertures are used, which can be opened singly or in combination. This allows a wide range of signal levels to be assessed without the need for additional elements such as neutral filters, but has the disadvantage of being considerably more complex to implement. A slight variant on this idea is a single aperture with a variable shutter, which can be used to generate apertures of various sizes, as described in Ref. 57.
c) The *multiple source* method (e.g., Ref. 58). Here the two fluxes are provided by two separate sources, each with a shutter. As with the double aperture method, the fluxes from the two sources are usually designed to be of similar magnitude and filters or similar, which is necessary if linearity over a wide range of levels is to be assessed.
d) The *beam conjoining* method (see Refs. 45 and 59). The radiation from the source is divided into two using a beam splitter, and each beam is subjected to variable attenuation using neutral density filters and apertures before being recombined. Shutters are also included so that radiation from one beam, the other, or both, is incident on the detector. Thus, the principle is very similar to the double aperture method, but the ability to vary independently the radiation in each beam allows investigation over a wider range of levels.

It will be clear from the above that the investigation of the linearity of a system by use of a superposition method can be a complex and time-consuming affair. Thus, although this is the most fundamental and accurate approach, its use is generally restricted to the most demanding applications, such as calibration of reference artifacts. Furthermore, the small size of most array spectrophotometers, and the consequent difficulty in placing a double aperture, etc., at a suitable position in the system, means that superposition methods normally cannot be used with these systems. This additional limitation does not apply to instruments for measuring sources, however, and Fig. 4.2.5-1 shows the results of an investigation of the linearity of an array spectroradiometer using the double aperture method.

In this case, the linearity of the system has been checked for 2:1 steps for three different integration times (0.1 s, 1 s, and 10 s) across the 400 to 880 nm wavelength range, using the same incident power levels in each case. The results show the system is nonlinear for the 10 s integration time, and further examination of the data reveals that this is due to saturation of the readout unit. Note also that the noise on the signal at the 0.1 s integration time appears as noise on the measured nonlinearity.

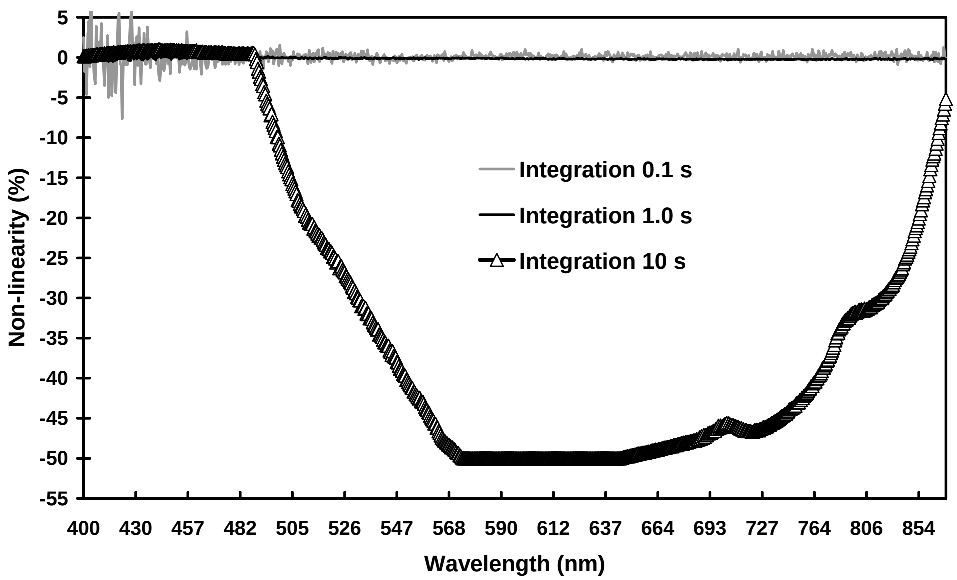


Figure 4.2.5-1 Results of a linearity measurement on a PDA spectroradiometer using the double aperture technique for three different signal integration times.

<u>Indirect methods</u>

Other methods for assessing linearity depend on the use of reference artifacts such as lamps and filters, and the measurement of another quantity such as distance or transmittance; they are therefore classed as indirect methods. These methods are more straightforward to implement than flux superposition techniques and therefore much more widely used.

In the case of a spectroradiometer, the simplest technique by which to assess system linearity is to use a source whose output can be varied in a known or predictable fashion. One way to achieve this is to use a lamp that obeys the inverse square law:

$$E_e = I_e \frac{\cos\theta}{d^2}, \tag{4.2.5-3}$$

where E_e is irradiance, I_e is radiant intensity, θ is the angle of illumination to normal of irradiated area, and d is the distance of the point source from irradiated area (Fig. 4.2.5-2).

The procedure is as follows: a suitable source is mounted on an optical bench perpendicular to and at a measured distance d_1 from the measurement plane of the spectroradiometer, and the spectral irradiance $E(\lambda)_1$ is measured. The source is then moved further away from the measurement plane to a new position d_2, and the new irradiance $E(\lambda)_2$ measured. If the source obeys the inverse square law and the system is linear, then

$$\frac{E(\lambda)_1}{E(\lambda)_2} = \frac{d_2^2}{d_1^2}. \tag{4.2.5-4}$$

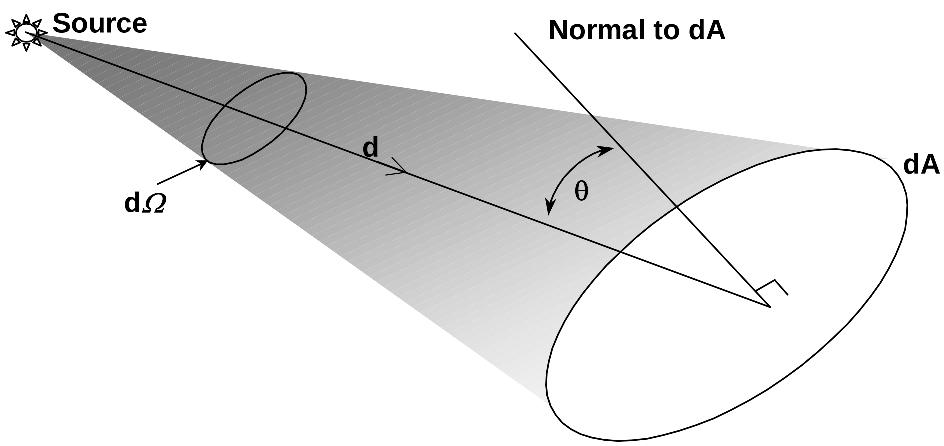


Figure 4.2.5-2 Inverse square law for a point source.

Thus, any system nonlinearity between $E(\lambda)_2$ and $E(\lambda)_1$ is given by

$$\text{Nonlinearity (\%)} = 100 \times \left\{ \left[\frac{E(\lambda)_1 \times d_1^2}{E(\lambda)_2 \times d_2^2} \right] - 1 \right\}. \tag{4.2.5-5}$$

In order for this method to work, the following is necessary:

a) All distances must be sufficiently large, or the source sufficiently small, that the source is effectively a point source (the inverse square law only applies for a point source). In practice this means that either a long optical bench is needed or a small, usually low-power, lamp must be used.

b) Distances must be known with sufficient accuracy. A consistent measurement error of 1 mm when using this method for a lamp placed at 1 m and 0.5 m, for example, will result in an apparent nonlinearity of 0.2% for a system that is actually linear. This means that a good-quality optical bench, with calibrated bench scale, must be used, and that it must be mounted perpendicular to the measurement plane of the spectroradiometer.
c) The source must emit radiation uniformly within the cone defined by the entrance aperture of the spectroradiometer at the closest distance used. If this condition is not fulfilled, then the measured irradiance will not obey the inverse square law.

Although the inverse square law method is relatively straightforward, it is limited in terms of the range of incident power levels that can be studied—to meet a typical requirement to assess linearity over three or four orders of magnitude requires a very long optical bench and the use of several different lamps. Thus, it is not widely used with conventional spectroradiometers. However, it can be a useful technique for evaluating the effect of using different integration times with array spectroradiometers, particularly at high power levels where nonlinearities are most likely to occur. The procedure is to make measurements of the time-integrated spectral irradiance at the same two distances for a short, intermediate, and long integration time, and determine the linearity for each condition. Care needs to be taken that the output signal is not significantly affected by noise at any of the integration times used. Any change in the measured linearity with integration time is likely to indicate saturation of the detector array, and can be used to set an upper limit to the time-integrated signal that can be measured reliably. A nonlinearity that does not vary with integration time is likely to be due to more complex effects such as filling of trap sites in the detector material (as described in Ref. 51); but with modern arrays these effects are relatively rare.

Another method by which to vary the output from a source in a known or predictable manner is by use of a calibrated, adjustable luminance gauge. In its simplest form, a luminance gauge generally consists of a lamp placed inside a small integrating sphere, but with variable devices the lamp is often placed outside the sphere on a micrometer-controlled stage and illuminates the interior of the sphere through an adjustable aperture (see Fig. 4.2.5-3). By changing the distance between the lamp and the entrance port and by adjusting the size of this port, it is possible to change the luminance at the exit port by a factor of about 100 without significantly changing the spectral distribution of the radiation.

In order to use such a device for checks on the linearity of a spectroradiometer, it is first necessary for it to be traceably calibrated for output luminance at all the aperture and lamp-distance settings that are to be used. Once this has been done, measurements can be made using the spectroradiometer at these same aperture and distance settings and the results ratioed to those for the maximum luminance setting. For a linear system, the ratio will be constant as a function of wavelength and will be equal to the ratio in the luminances at these settings. As in the case of an investigation using the inverse square law, care needs to be taken that noise on the output at low signal levels is not misinterpreted as nonlinearity. In addition, since different areas of the sphere are illuminated as the lamp distance or aperture size are varied, a slight spectral shift may occur in the output; for particularly demanding applications it may be necessary to investigate this effect using a conventional scanning monochromator/detector system prior to use of the gauge for linearity evaluations.

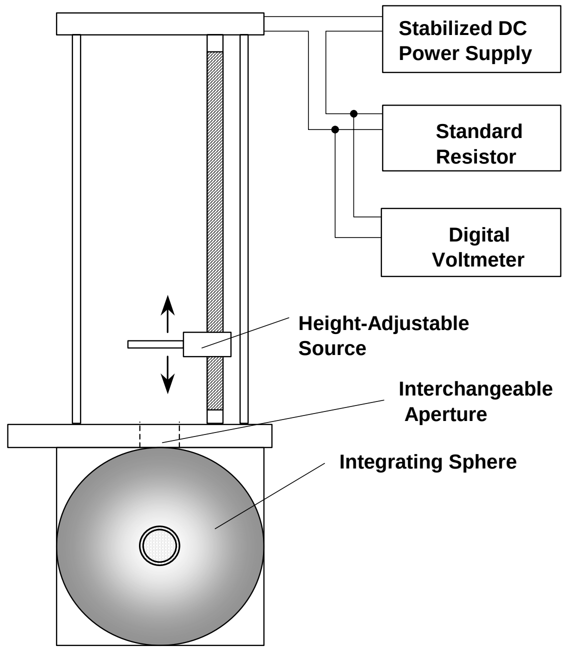


Figure 4.2.5-3 Schematic of variable luminance gauge.

Although the use of a variable luminance gauge of this type for investigating system linearity is relatively straightforward, it is difficult to generate sufficiently high luminance levels to check performance at the limits of the working range of most spectroradiometers. In addition, these devices, particularly when calibrated, tend to be quite expensive. Thus, the use of this approach is not widespread. However, a luminance gauge is an excellent transfer standard for calibrating the responsivity of an array spectroradiometer, as discussed in Sec. 4.2, and if already used for this purpose, the additional cost involved in calibrations at a range of settings to allow its use for linearity checks is probably well justified.

An example of this type of evaluation performed on a 1024-element PDA spectroradiometer is given in Fig. 4.2.5-4. In this case, the results from three combinations of aperture and distance setting are shown, covering approximately 30:1 in terms of luminance level (setting A is that for the highest luminance). The results have been expressed as

$$\text{Nonlinearity} = 100\times\left[\frac{L_A(\lambda)/L_X(\lambda)}{L_A/L_X}-1\right], \tag{4.2.5-6}$$

where $L_X(\lambda)$ is the measured spectral radiance at setting X using the spectroradiometer, and L_X is the calibrated luminance at setting X.

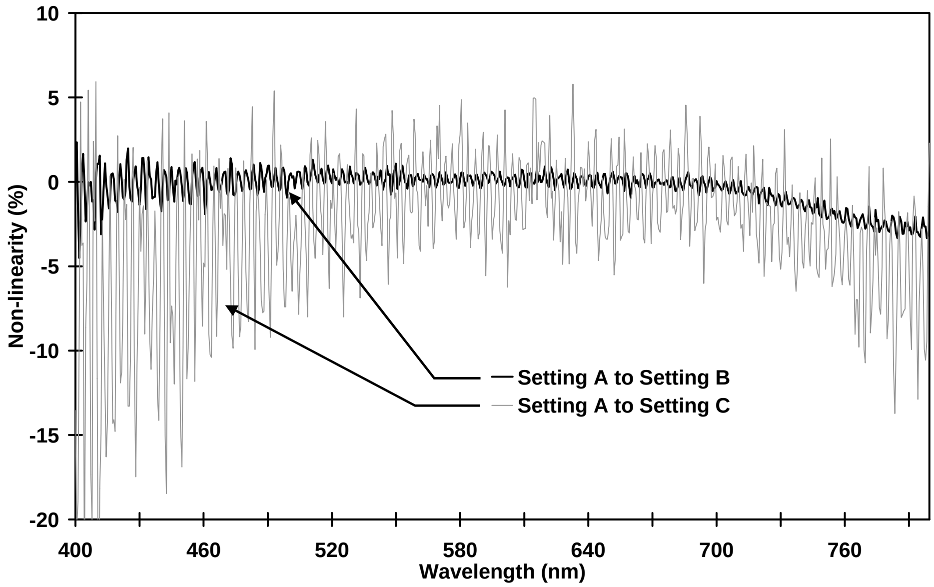


Figure 4.2.5-4 Example of the results of a linearity check on a PDA spectroradiometer using a variable luminance gauge.

The following conclusions can be drawn from this example:

1. Saturation of the system is apparent at long wavelengths (it should be noted that signal levels are highest in this region).
2. The magnitude of the saturation effect is similar when going from luminance A to either luminance B or C; thus, the signal corresponding to that measured at long wavelengths at setting B can be taken as a "safe" upper limit, below which the system is likely to behave linearly.
3. The measured spectral radiance values are dominated by noise at the lowest luminance setting (setting C), particularly in the blue spectral region, and this may account for the apparent nonlinearity in the blue. Further investigations (e.g., using longer integration times) would be necessary to confirm whether this is a "real" effect.

Methods using sources are well suited to investigations of spectroradiometer linearity, but, as already indicated, they are generally limited in terms of the maximum level and range of levels that can be studied and the reference artifacts are relatively expensive. A more widely used approach, which is also applicable to spectrophotometers, is based on calibrated neutral density filters (see Refs. 47, 48, and 60 for examples). Neutral density filters are widely available as either 2-in. squares or cuvette size; the former are usually most useful for spectroradiometers and the latter for spectrophotometers. The procedure is:

1. Make a measurement without the filter using a stable source, **S**(λ), allowing for dark signal.
2. Insert the (traceably calibrated) filter in a suitable location in the beam, in the correct orientation and position, and repeat the measurement $F(\lambda)$, again allowing for dark signal.
3. Calculate the measured filter transmittance $T_M(\lambda)$ from the ratio $F(\lambda) / S(\lambda)$.
4. Compare the results with the calibrated transmittance, $T_C(\lambda)$.
5. The nonlinearity between reading F and reading S at any specified wavelength is given by

$$\text{Nonlinearity} = 100 \times \left(\frac{T_M}{T_C} - 1 \right). \tag{4.2.5-7}$$

6. Repeat using other filters (or filters in combination) to cover the full range of levels of interest.

Normally, the calibration of the filter(s) will be performed at a limited number of wavelengths. The linearity can only be assessed at these same wavelengths—neutral filters are never completely neutral, and erroneous results will be obtained if the calibrated transmittance at one wavelength is assumed to apply at all other wavelengths as well.

A number of precautions are necessary when using this method:

- The transmittance will depend on the pathlength through the filter, which in turn is affected by its orientation and the cone angle of the incident radiation. Generally, filters are calibrated using quasi-collimated radiation and with the filter perpendicular to the beam.
- For a nonuniform filter, the transmittance will also vary with the size and position on the surface of the incident radiation.
- Inter-reflections between filters (if they are being used in combination) or between the filter and the other elements of the system (lenses, apertures, polychromator, etc.) may affect measured transmittance.
- Surface contamination of the filter will also affect its transmittance and may also result in increased degradation of the filter with time (i.e., large changes in filter transmittance with time).
- As with all other methods, noise and/or poor measurement repeatability can be misinterpreted as being indicative of nonlinear behavior. For this method, repeated measurements should be made so that the effect of nonreproducibility of the filter position, for example, can be correctly allowed for.

Clearly, it is important to locate the filter at a suitable position in the system so as to reproduce the calibration conditions as precisely as possible. With a spectroradiometer this can be difficult to achieve—most systems are designed to collect radiation over a fairly large cone angle, and illumination of the filter under quasi-collimated conditions may not be possible. In this case, it may be necessary to specify the required illumination conditions to the calibration laboratory so as to more closely match the spectroradiometer input optics. With a transmittance spectrophotometer, on the other hand, the situation is much more straightforward and it should generally be possible to place the filter(s) in the normal sample position. With reflectance spectrophotometers it may be possible to place neutral filters at a suitable position in the illumination optics, but usually it is preferable to use calibrated neutral (gray) tiles or Spectralon samples instead of filters—these can then be placed in the normal reflectance sample position.

An example of such a linearity investigation using neutral density filters for a 512-element PDA spectrophotometer is shown in Fig. 4.2.5-5 and summarized in Table 4.2.5-1. Two filters were used in this case, one with a transmittance of approximately 50% and the other approximately 15%. The results show that this spectrophotometer is linear to better than the measurement repeatability for both filters.

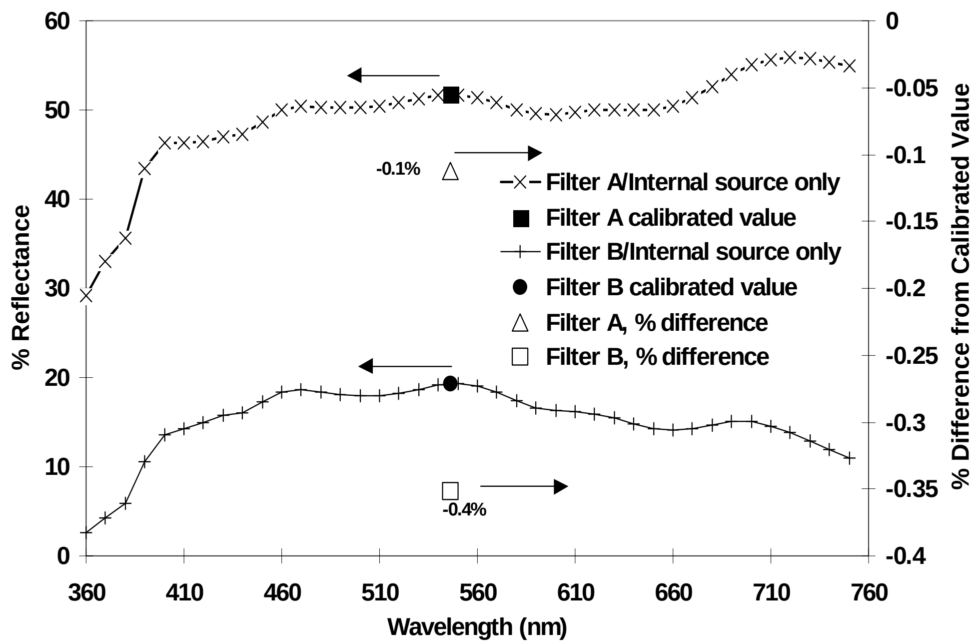


Figure 4.2.5-5 Example of the results of a linearity investigation using neutral density filters for a 512-element PDA spectrophotometer.

Table 4.2.5-1 Results of a linearity investigation for a 512-element PDA spectrometer.

Filter	Calibrated transmittance value at 550 nm (%)	Measured transmittance value at 550 nm (%)	% difference	System repeatability (%)
A	51.7	52.6	-0.1	~ 0.5
B	19.3	18.9	-0.4	~ 0.5

Other indirect methods which are sometimes used to assess the linearity of a system include the following:

1. **Beer's law—**The transmittance, *T*, of a solution is related to the concentration of the solute as follows:

$$\log (1/T) \propto \text{concentrate of solute.} \tag{4.2.5-8}$$

The usual form of application of this law to test for detector linearity involves cuvettes of varying path lengths or solutions of varying concentration. It is most commonly used by chemists to check the performance of spectrophotometers.

2. **Use of polarizers—**A pair of polarizers provides a filter with variable transmittance:

$$T = T_0 \cos^2 \alpha, \tag{4.2.5-9}$$

where α is the angle between the two planes of polarization and T_0 is the transmittance for $\alpha = 0°$.

If the angle α can be measured or set accurately, the predicted transmittance can be used to test for detector linearity. This method is prone to a number of sources of error and is therefore not widely used.

3. **Use of rotating sector discs and choppers—**Attenuators based on rotating sector discs and choppers are spectrally nonselective, insensitive to polarization, and relatively easy to construct. Provided that the detector with which they are used obeys Talbot's law (i.e., that the mean signal produced by periodically fluctuating radiation is the same as that obtained if the incident flux is distributed uniformly throughout the period), then the measured transmittance of the sector disc will be linearly proportional to the angular opening. In order to check the linearity of a system over a range of levels, therefore, it is necessary to have a series of sector disCs of varying sizes of angular opening and to be able to measure accurately the area of these openings. Alternatively, several similar rotating discs may be used in combination and their individual transmittances combined. Thus, this method is very similar in basic concept to the use of calibrated neutral density filters. However, complications can arise due to possible beating effects between the clock frequency and the rate of signal fluctuation, and this method is therefore not recommended for array systems.

4.2.5.3 Linearity and Dynamic Range

The term dynamic range can be used in many different contexts when discussing array spectrometers. For the purposes of this guide, we take dynamic range to be the range of output signal levels over which the system will operate reliably. It is usually limited at the lower end by noise and dark current, and at the upper end by system nonlinearity and saturation effects. Thus, in order to specify the dynamic range of a system, it is necessary to investigate its dark current characteristics (see Sec. 3.1.7) and its linearity.

Manufacturers generally state the dynamic range of their systems in terms of the number of bits; e.g., a 16-bit dynamic range. This relates to the performance of the A/D converter alone.

4.2.5.4 Recommendations

1. Linearity should be assessed for a range of integration times and at varying incident power levels in order to determine safe operating limits. Any change in the measured linearity with integration time is likely to indicate saturation of the readout registers, and can be used to set an upper limit to the time-integrated signal that can be measured reliably. A nonlinearity that does not vary with integration time is likely to be due to more complex effects such as filling of trap sites in the detector material; but with modern arrays these effects are relatively rare.
2. The linearity assessment should ideally cover the full range of wavelengths over which the system is to be used; linearity may vary with wavelength and performance may be poor at the limits of the responsivity range of the array elements.
3. Flux superposition methods are the most fundamental and accurate means by which to assess linearity and are recommended for state-of-the-art applications.
4. Indirect methods are more prone to error but are simpler and cheaper to perform. They are recommended for the majority of users of this guide.
5. The choice of method will depend on the details of the application, but in most cases a source-based approach or the use of neutral density filters will be recommended for a spectroradiometer, neutral density filters or varying solute concentrations for a transmittance spectrophotometer and neutral reflectance standards (tiles or Spectralon plaques) for a reflectance spectrophotometer.
6. Whichever indirect method is chosen, it is important that the reference artifact(s) are traceably calibrated under conditions similar to those in which they will be used on the spectrometer.

4.2.6 Dark Reading and External Stray Light

4.2.6.1 Introduction

With all spectrometers, some signal is recorded even when the instrument is in darkness, due to thermal generation, etc. (see Sec. 1.2.3) For an array device, this “dark reading” will be further increased by the offset level and fixed pattern noise.

It is important to distinguish between the signal recorded when the system is in darkness (the true dark reading) and that recorded when the irradiating source is switched on but radiation is prevented from directly reaching the polychromator by use of a shutter. The latter is sometimes—confusingly—also referred to as the *dark reading*, but it actually also includes the external stray light (see Sec. 4.2.3), so we will therefore call this the *background reading*. The external stray light is then equal to the difference between the background and dark readings. If correct measurement results are to be obtained, it is essential to subtract both the dark reading and the external stray light from the signal recorded when the system is irradiated.

4.2.6.2 Calibration methods and recommendations

The software for most array-based spectrometers is written so that a dark reading must be taken before measurements can be made. In the case of a spectroradiometer, the dark reading is taken by placing a completely opaque cover over the entrance to the polychromator, if this is accessible, or in front of the input optics if these are fixed to the polychromator. For a spectrophotometer, a piece of matte black card or similar nontransmitting or reflecting sample is placed in the normal sample position. In the majority of cases, only one dark reading is required by the software, and this is then automatically subtracted from all subsequent signal measurements, although some systems may require a dark reading to be taken before each measurement. If only one dark reading is made, it is recommended that occasional measurements are made under the same conditions as used for this dark reading; the recorded signal should be zero, and any nonzero signal would indicate a shift in the dark signal. In addition, for most systems a new dark reading will be required if the signal integration time is changed, and the software will generally ensure this is taken.

However, knowledge of the dark reading alone is not sufficient; as mentioned above, when making measurements it is also necessary to allow for any external stray light. Most spectrophotometers are designed with good screening between the various system components and have well-blackened walls, etc., so the levels of external stray light are negligible. In spectroradiometers, on the other hand, there is significant potential for light from the source to be bounced off floors, walls, ceiling, or system components, so as to reach the input slit of the polychromator by indirect routes; the levels of external stray light can therefore be high. Furthermore, the amount of external stray radiation present will vary depending on the geometry of the source and the precise positions of any reflecting surfaces; therefore, it needs to be regularly reassessed. There are three approaches that can be used:

1. The level of external stray light can be measured by placing a baffle close to the source, which is just large enough to prevent direct radiation form the source from entering the system. Screens can then be positioned so as to reduce the external stray light to negligible levels.
2. If the effective elimination of external stray light is not practical, a stray light reading can be made and subtracted from all subsequent (dark-reading-corrected) measurements.
3. A measurement of the background reading can be made instead of a dark reading. In practice this means placing the opaque screen immediately in front of the source rather than the spectrometer.

The first approach is recommended whenever possible. When this is not feasible, the latter approach has the advantage that the dark reading and external stray light are both automatically subtracted from the signal readings (assuming that this is how the software has been configured). The disadvantage is that the user then has no information on the levels of external stray light that are present and is therefore less able to judge the impact that small changes in source geometry or position will have on the measurement results; i.e., measurement uncertainties will be higher.

4.3 Multichannel Raman and Emission Spectrometers

The previous section discussed general issues in calibrating instruments for spectrometry. With Raman spectrometers and other instruments for emission spectroscopy (e.g. atomic emission spectroscopy and fluorescence spectroscopy) there are some specific issues, which are discussed below.

Multichannel Raman spectrometers primarily use gratings for dispersion. They use silicon CCDs and operate in the visible and near-IR (though silicon detector arrays are also used in UV resonance Raman spectrometers). With the improved performance of present detectors, the availability of a variety of laser sources and the introduction of holographic laser line rejection filters and optic–optic probes, rugged compact dispersive spectrometers now rival (and, in many cases, surpass) the performance of FT-Raman instruments. Dispersive systems with 2D CCD detectors (multichannel Raman) tend to be most popular where the CCD sensitivity has most advantage, i.e., in low-light applications. Such applications often demand high accuracy and efficiency in the calibration process. Although the effects involved are essentially similar to those discussed in Sec. 4.2, there are specific issues and a specific body of literature related to Raman spectrometers. As with all spectrometry instruments, wavelength and intensity calibration are the prime concerns. Stray light is not directly an issue since notch filters are used to reject light at the excitation laser wavelength, though Rayleigh scattering and sample fluorescence tend to be important in the visible and UV.

4.3.1 Wavelength Calibration

Raman

Wavelength calibration for multichannel Raman spectrometers has been discussed in detail by Carter et al.,[61,62] Tseng et al.,[30] and Fountain et al.[63] These authors also provide references to earlier literature to the calibration of Raman instruments in general. Vickers[64] has given an overview, with examples and recent data. The main issues separate into two categories: (1) issues related to instrumental factors, such as source/grating changes, collection optics, thermal distortion, etc., and (2) issues related to wavelength fitting (interpolation) and choice of wavelength standards.

Reference 61 discusses the instrumental factors in detail. It is shown (also in Ref. 30) that small misalignments in the collection optics can cause significant shifts in the Raman spectrum and in the symmetry of calibration lines. Errors can be as large as 40 cm^{-1}. In use it is normal for samples to be changed frequently, and this entails movement of the Raman source; however, misalignments can be minimized if the collection lens is kept fixed. The authors discuss the use of a sample stage that allows coordinated movement of both the laser beam and the sample. Use of optics will help to minimize the misalignments caused by source or sample changes.[30] Temperature drifts can also cause frequency shifts. It is usually necessary to perform wavelength calibration after grating changes, but not necessarily after each rotation (to a different wavelength band).

In Raman spectroscopy the output spectrum needs to be interpreted in terms of the Raman shifts relative to the excitation wavelength. Wavelength calibration can be performed either by (a) using absolute wavelength standards (e.g., atomic emission lines), in which case the laser wavelength also needs accurate absolute measurement (so that shifts can be calculated), or (b) using Raman shift standards that produce lines with known shifts relative to the laser line.

In many applications the Raman shift needs to be measured to an accuracy <1 cm^{-1} (corresponding to 0.02 nm at 400 nm, 0.04 nm at 600 nm, or 0.1 nm at 1000 nm). Raman shift standards can have wide (5–10 cm^{-1}) line widths, leading to calibration uncertainties, but have the advantage that the laser wavelength does not need to be determined (especially useful if a variety of sources or a tunable laser are in use).

Although the narrower lines achievable with emission sources tend to give higher wavelength accuracies, the situation can arise where line widths are comparable with the pixel dimension (since the spectral waveband covered by a CCD line can be large). Errors in determining the center wavelengths can result. In some instruments the center wavelength is simply determined by positioning a cursor by eye. This can lead to errors when the spectral lines are sparsely sampled, as discussed in Ref. 61. References 30 and 63 discuss a technique to interpolate the data by adding zeros to the Fourier transform (which is apodized to force the line shape to a Gaussian); the peak is then found by fitting a polynomial. The center-of- gravity method is also useful for cases where there are few samples over the line shape (see also Sec. 4.2.2). However, any of these methods will not correct for errors due to inter- or intrapixel uniformity, that is, in the assumption that all pixels have the same response, which is flat across the pixel. Reference 63 discusses a technique where several pairs of line and Raman spectra are obtained with small changes in grating position in between. This gives a range of sampling positions over the pixel and tends to average out the sampling errors. The authors note a significant improvement in the precision of wavelength measurement with this "blurring" technique. Note that the grating changes (wavelength shifts) do not need to be measured since the spectra are not superimposed (it is the separations between the wavelength pairs that are pooled). Recent work by Vickers[64] indicates that even with the technique discussed above, the wavelength accuracy is limited to a few tenths of 1 cm^{-1}. This may be due to the effects of nonuniformities in the response of the CCD pixels.

A variety of wavelength shift and atomic line standards have been discussed in Ref. 62 and some updated wavelength values are given. Neon lines (e.g., from a commercial "pencil lamp") are commonly used for line standards (see, for example, Ref. 65), and laser plasma lines can also be convenient since they can be obtained by detuning the laser and scattering off a solid sample (such as imidazole). The most common laser source is Ar-ion, and Ar^+ lines are tabulated in Ref. 62.

Note that accurate measurement of the laser wavelength is often more conveniently performed using a shift standard (without the need for removal of the laser notch filter). Knowing the absolute value of the line wavelengths produced by the shift standard (from the emission line calibration), as well as knowledge of the shifts, allows determination of the laser wavelength. Another (less accurate) method is to inject some of the laser light directly into the spectrometer (bypassing the notch filter).[66]

An ASTM standard[67] gives further information on Raman shift standards, and McCreery et al.[68] have listed information on the Web.

The choice of standard will depend on the wavelength range of interest; as many lines as possible (at least six) of suitable intensity (i.e., neither saturated nor in the noise level) should be available over this range. The peaks should be fairly evenly spread over the wavelength range, with wavelengths that are reproducible and accurately known. Also, the optical configuration should be as similar as possible to that used with samples. The relatively few lines and lack of accurate shift values for shift standards has to be balanced against the need for laser line calibration and the sparsely sampled nature of atomic lines; and in some laboratories, both types of standards are used in conjunction.

The algorithm used to interpolate between the calibration wavelengths and an unknown wavelength also needs to be considered. A simple algorithm is usually required since many applications require frequent changes in experimental configuration and, hence, frequent calibrations. Many commercial instruments use a polynomial fit to the pixel positions to derive the *wavenumber* (in cm^{-1}). Carter et al.[62] have studied the errors involved and concluded that higher-order polynomials are more susceptible to errors, and that the lowest order should be used that gives an acceptable fit between pixel number and wavenumber. They found that the method gave acceptable accuracy for routine work.

In fact, the relation between wavelength and pixel number is usually nearly linear over a limited wavelength range, although data processing can be inconvenient with some commercial instrument software. Carter et al. found that a linear interpolation in wavelength gave acceptable accuracy over a moderately wide range. A linear interpolation in wavelength was also used by Fountain et al.[63] Tseng et al.[30] used a polynomial fit to wavelength to achieve good accuracy over a wider wavelength range, but this approach is likely to be too time consuming for routine measurements unless specifically provided for in the instrument software.

Fluorescence

A standard method for wavelength calibration of fluorescence spectrometers is to a glycogen suspension or colloidal silica sample illuminated with a low-pressure mercury arc.[69] A fluorescence standard based on a single crystal of dysprosium-activated yttrium aluminum garnet has recently been proposed.[70] When excited by UV radiation (from 250 to 500 nm) the crystal emits in four bands in the range 450 to 800 nm.

4.3.2 Intensity Calibration

Raman

Raman spectroscopy is a single-beam technique and does not have the inherent response calibration of dual-beam absorbance spectroscopy. Hence, Raman spectra are often not calibrated for wavelength variations in spectrometer efficiency or detector response, nor for pixel response nonuniformity. This leads to distortion of line profiles and intensity values. Calibration for the above effects is performed using broadband sources of known spectral shape, but it is important that the Raman sampling geometry is preserved. This has led to the recent investigation of fluorescent chemical[71] and glass standards[72, 73] that are placed in the sample position and excited by the Raman laser. Quinine is a common broadband fluorescent standard, but it requires UV excitation and is not applicable to visible or NIR spectroscopy. Luminescent glass standards are under development. The present accuracy of intensity calibration is limited to 10–20%, but future developments may improve on this.[63]

The disadvantage of luminescence standards is that a different fluorophore is needed for each laser wavelength. Also, they are inconvenient for some online processing applications. Use of white-light tungsten halogen bulbs is a simple and attractive alternative,[64] but it is almost essential that optic–optics are used to relay the light to the spectrometer without change in the spectrometer collection geometry.[74] At least one manufacturer[75] offers a tungsten-halogen calibration accessory (see also Ref. 66). Note that an absolute calibration is only needed for quantitative work, usually only the shape of the lamp spectrum needs to be known and reproducible. It has been suggested[76] that the shape of the spectrum for a tungsten halogen lamp can be calculated from Planck's law for a blackbody radiator. However, this assumes that the temperature of the tungsten filament is known. Depending on the calibration accuracy involved, it will often be more practical to use a lamp with a traceable calibration (see the discussions in Secs. 4.2 and 6.1.1 on calibration effects for continuum light sources).

Fluorescence

Similar problems in intensity calibration arise in fluorescence spectroscopy. After examining other methods, Gardecki and Maroncelli[77] proposed the use of secondary emission standards based on six common fluorophors dissolved in neat solvents. These give smooth emission profiles which together span the region 300–800 nm.

4.4 Imaging Spectrometers

Imaging spectrometers are required to be calibrated in both spatial (imaging) and spectral performance. The main parameters are

- Focus and aberration of the spatial image (i.e., the MTF), which may depend on wavelength and may be different in the two directions (along-scan and across-scan).
- Response nonuniformity and pixel offsets (fixed pattern noise and thermal dark signal)—as in any imager
- Spectral resolution
- Stray light
- Noise, dynamic range
- Wavelength accuracy and its stability over time—as in any spectrometer
- Straightness and parallelism of near-monochromatic spectral line images. Departures from straightness are called *smile*. Smile is the change in apparent wavelength with across-scan position in the image (or field position).
- Straightness and parallelism of line spectra from white point sources at various positions in the image (in the field). Departures from straightness are called *frown*. Frown is the change in apparent spatial (or field) position with wavelength.

Testing for spectral resolution, stray light, noise, etc., are essentially as described previously for imagers and spectrometers. Pixel responsivity and offsets can be calibrated using flat-field targets at single or multiple wavelengths (pixel response nonuniformity will usually be wavelength dependent). The illumination source needs to be stable over the scan time if temporal changes in brightness are not to be confused with nonuniformities in the along-scan direction. Frequent nonuniformity calibration is particularly important for infrared instruments.

The remaining parameters—MTF, wavelength accuracy, smile, and frown—are discussed below. The basic principles of imaging spectrometer calibration have been described in two recent papers. The calibration of a high-precision instrument (in this case a spaceborne remote sensing instrument) has been described by Cutter et al.,[78] while Martinsen et al.[79] describe a low-resolution instrument for food inspection.

Note that in an imaging spectrometer each image from the CCD (or other detector array) will be a "slice" of the X-Y-λ data cube (see Sec. 2.2.6) and will have spatial (across-scan) information in one direction and wavelength (dispersion) in the other. These directions will be either the rows or columns of the detector (depending on the orientation of the scan direction and slit relative to the array. For example, the spectrum will be dispersed either in the row or the column direction).

4.4.1 MTF Testing

MTF testing can be performed either using a scanning slit or spatial targets (such as speckle patterns or bar targets). The basic principles are described in Sec. 1.2.14. The along-scan MTF will tend to be limited by the pixel dimension. The across-scan MTF will tend to be limited by the product of the slit width and the scan speed. Variability in the scan speed will produce an along-scan MTF that varies over the image. It will also produce distortions that can be detected with a bar or grid target.

4.4.2 Wavelength Accuracy

Wavelength accuracy can be measured on single images using a spectral line source or external monochromator as in nonimaging spectrometers (Sec. 4.2.2), the source being arranged to fill the entrance slit. Because of the smile effect, the spectral "lines" will be bowed (to an extent depending on the design and alignment of the spectrometer), and so measurements should always be made at the same location on the spatial axis (column or row number).

4.4.3 Smile

In the case where the scanning motion is fixed and uniform (as in a remote sensing instrument), smile (the variation in apparent wavelength with position) can be measured on single image slices. As for wavelength calibration, the line source is arranged to fill the entrance slit (see the recommendations in Sec. 4.2.2). Normally the measurement is a check that the instrument is within calibration and that correction for smile is not needed. In cases where the slit is mechanically scanned (e.g. by a stepper motor), then movement of the spectral scale with slit position should be checked by measuring the position of spectral peaks on several image slices (in Ref. 79 this was done by examining 10 equally spaced slices). It is not necessary to fill the slit. For example, a pencil-type lamp can be positioned at right angles to the slit and a scan made along its length.

4.4.4 Frown

If a moving slit illumination source is available, then this can be used to measure frown (the variation in position with wavelength). In the case where the scanning motion is fixed and uniform (as in a remote sensing instrument), frown can be measured on single image slices. The external slit source is positioned perpendicular to the entrance slit of the spectrometer. The illumination should be white, i.e., broadband, so that each image from the detector array is a thin line extending along the dispersion direction. Any tilt or bowing of this line indicates frown. The slit illuminator is then moved in a direction perpendicular to its length (so its image moves along the entrance slit) so as to investigate changes in frown with across-track position. If a moveable slit is not available or a rapid measurement is needed, then the slit source can be replaced by a set of parallel slits (Ref. 79 used a white card with black lines). As with smile, several image slices should be checked with mechanically scanned systems. Normally the measurement is a check that the instrument is within calibration and that correction for frown is not needed.

References

[1] M. R. Shortis, W. L. Snow, B. A. Childers and W. K. Goad, "The influence of storage media on the accuracy and repeatability of photogrammetric measurements using CCD cameras," *Proc. SPIE*, vol. 2067, pp. 80-92 (1993).

[2] A. Cumani, A. Guidicci, P. Grattoni and G. Pettiti, "Characterization of commercial solid-state TV cameras for non-contact metrology," *Proc. SPIE*, vol. 2067, pp. 115-122 (1993).

[3] A. M. Amin and G. Petrie, "Geometric calibration of thermal video frame scanners," *Proc. SPIE*, vol. 2067, pp. 44-54 (1993).

[4] G. Buyuksalih, "Geometric calibration models of infrared cameras and scanners for photogrammetric and remote sensing applications," *Opt. Eng.*, vol. 42(7), pp. 1923-1934 (2003).

[5] X. Wang and R. Ning, "Accurate and efficient image intensifier distortion correction algorithm for volume tomographic angiography," *Opt. Eng.*, vol. 37(3), pp. 977-983 (1998).

[6] H. A. Beyer, "Determination of radiometric and geometric characteristics of framegrabbers," *Proc. SPIE*, vol. 2067, pp. 93-103 (1993).

[7] G. Luo, O. Chutape and H. Fang, "Experimental study on nonuniformity of line jitter in CCD images," *Appl. Opt.*, vol. 40(26), pp. 4716-4720 (2001).

[8] A. F. Habib and M. Morgan, "Automatic calibration of low-cost digital cameras," *Opt. Eng.*, vol.42(4), pp. 948-955, (2003).

[9] A. Asundi and Z. Wensen, "Unified calibration technique and its applications in optical triangular profilometry," *Opt. Eng*., vol. 38(16), pp. 3556-3561 (1999).

[10] K. Barnard and B. Funt, "Camera calibration for color research," *Proc. SPIE*, vol. 3644, pp. 576-585 (1999).

[11] A. J. Blanksby and M. J. Loinaz, "Performance analysis of a color CMOS photogate image sensor," *IEEE Trans. Elec. Dev*., vol. 47(1), pp. 55-64 (2000).

[12] G. Putnam, S. Kelly, S. Wang, W. Davis, E. Nelson and D. Carpenter, "Photography with an 11-megapixel 35-mm format CCD," *Proc. SPIE,* vol. 5017, pp. 371-383 (2003).

[13] G. C. Holst, *Testing and Evaluation of Infrared Imaging Systems*, JCD Publishing Co., Maitland, USA, 1993.

[14] D. L. Perry and E. L. Dereniak, "Linear theory of nonuniformity correction in infrared staring sensors," *Opt. Eng*., vol. 32(8), pp. 1854-1859 (1993).

[15] J. M. Mooney, F. D. Shepherd, W. S. Ewing, J. E. Murguia, J. Silverman, "Responsivity nonuniformity limited performance of infrared staring cameras," *Opt. Eng*., vol. 28(11), pp. 1151-1161 (1989).

[16] A. F. Milton, F. R. Barone, M. R. Kruer, "Influence of nonuniformity on infrared focal plane array performance," *Opt. Eng*., vol. 24(5), pp. 855-862 (1985).

[17] D. A. Scribner, M. R. Kruer, C. J. Gridley, K. Sarkady, "Physical limitations to nonuniformity correction in IR focal plane arrays," *Proc. SPIE*, vol. 865, pp. 185-203 (1987).

[18] D. A. Scribner, M. R. Kruer, C. J. Gridley, K. Sarkady, "Measurement, characterization, and modeling of noise in staring infrared focal plane arrays," *Proc. SPIE*, vol. 782, pp. 147-161 (1987).

[19] D. A. Scribner, K. Sarkady, J. T. Caulfield, M. R. Kruer, G. Katz, C. J. Gridley, C. Herman, "Nonuniformity correction for staring IR focal plane arrays using scene-based techniques," *Proc. SPIE*, vol. 1308, pp. 224-233 (1990).

[20] C. T. Elliott, N. T. Gordon, R. S. Hall and G. Grimes, "Reverse breakdown in long wavelength lateral collection $Cd_xHg_{1-x}Te$ diodes," *J. Vac. Sci. Technol.*, vol. A8(2), pp. 1251-1253 (1990).

[21] N. Bluzer and A. S. Jensen, "Current readout of infrared detectors," *Opt. Eng*., vol. 26(3), pp. 241-248 (1987).

[22] "The spectroradiometric calibration of light sources," CIE Publication 63, published by the CIE (1984).

[23] J. R. Moore, "Sources of error in spectroradiometry," *Light. Res. Technol.*, vol. 12(4), pp. 213-220 (1980).

[24] "Spectroradiometry of pulsed optical radiation sources," CIE Publication 105, published by the CIE (1993).

[25] H. K. Roscoe, D J. Fish and R. L. Jones, "Interpolation errors in UV-visible spectroscopy for stratospheric sensing: implications for sensitivity, spectral resolution and spectral range," *Appl. Opt*., vol. 35(3), pp. 427-432 (1996).

[26] J. Cho, P. J. Gemperline and D. Walker, "Wavelength calibration method for a CCD detector and multichannel fibre-optic probes," *Appl. Spectrosc*., vol. 49(12), pp. 1841 - 1845 (1995).

[27] M. Lindrum and B. Nickel, "Wavelength calibration of optical multichannel detectors in combination with single and double grating-monochromators," *Appl. Spectrosc*., vol. 43(8), pp. 1427-1431 (1989).

[28] M. Diem, F. Adar and R. Grayzel, "A simple algorithm to convert diode array spectral data to linear wavelength or wavenumber scales," *Comp. Enhanced Spectrosc*., vol. 3, pp. 29-33 (1986).

[29] J. J. Sullivan and B. D. Quimby, "Characterisation of a computerised photodiode array spectrometer for gas chromatography—atomic emission spectrometry," *Anal. Chem.*, vol. 62(10), pp. 1034-1043 (1990).

[30] C. H. Tseng, J. F. Ford, C. K. Mann and T. J. Vickers, "Wavelength calibration of a multichannel spectrometer," *Appl. Spectrosc*., vol. 47(11), pp. 1808-1813 (1993).

[31] S. Scopatz, G. Neel, E. Romesburg and M. Zivitz, "Calibrating a detector array based spectroradiometer with sub-pixel precision," *Proc. SPIE*, vol. 1055, pp. 306-313 (1989).

[32] P. E. Berlo and G. A. Locascio, "Ultraviolet-visible photodiode array spectrophotometer wavelength calibration method. A practical computer algorithm," *Analyst*, vol. 116, pp. 313-316 (1991).

[33] Standard practice for describing and measuring performance of ultraviolet, visible and near infrared spectrophotometers, ASTM Designation : E 275 – 83.

[34] "Operational qualification and performance verification of UV-visible spectrophotometers," HP Technical note, available on Hewlett-Packard CD-ROM *Explore the World of UV-Visible Spectroscopy* ('98 Edition).

[35] C. Burgess and K. D. Mielenz, eds., "The role of photodiode array spectrometry in analytical chemistry," in *Advances in Standards and Methodology in Spectrophotometry*, Elsevier Science Publishers B.V., 1987.

[36] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch and E. Rovira, "Wavelength calibration transfer between diode array UV-Visible spectrophotometers," *Appl. Spectrosc.*, vol. 49(5), pp. 593-597 (1995).

[37] J. C. Travis, G. W. Kramer, M. V. Smith and S. J. Choquette, "Challenges in providing standard reference materials for chemical and pharmaceutical process analysis," *Anal. Chem. Acta*, vol. 380(2-3), pp. 115-126 (1999). See also J. C. Travis and D. L. Duwer, "Holmium oxide solution as a prototype intrinsic standard in molecular absorption spectrophotometry, *Proc. SPIE*, vol. 4826, pp. 103-112 (2003); and S. J. Choquette, J. C. Travis, L. E. O'Neal, C. Zhu, and D. L. Duwer, "A rare-earth-oxide glass for the wavelength calibration of near-infrared dispersive and Fourier-transform spectrometers," *Spectroscopy*, vol. 16(4), pp. 14-19 (2001).

[38] J. F. Verrill, "Advances in spectrophotometric transfer standards at the National Physical Laboratory," in *Spectrophotometry, Luminescence and Colour; Science and Compliance*, edited by C. Burgess and D. G. Jones, Elsevier Science Publishers B.V., pp. 49-63 (1995).

[39] V. R. Weidner, R. Mavrodineanu, K. D. Mielenz, R. A. Velapoldi, K. L. Eckerle and B. Adams, "Holmium oxide solution wavelength standard from 240 to 640 nm - SRM 2034," NBS Special Publication 260-102 (1986).

[40] H. H. Schlemmer and M. Machler, "Diode-array spectrometer: an optimised design," *J. Phys. E: Sci Instrum.*, vol. 18(11), pp. 914-919 (1985).

[41] A. Springsteen, "Standards for the measurement of diffuse reflectance - an overview of available materials and measurement laboratories," *Anal. Chem. Acta*, vol. 380, pp. 379-390 (1999).

[42] J. F. Verrill, P. C. Knee and J. C. O'Halloran, "A study of the uniformity and thermochromism of surface colour standards," CIE 119-1995-23rd Session, New Delhi.

[43] F. J. J. Clarke, "High accuracy spectrophotometry at the National Physical Laboratory," NBS Special Publication 378, Accuracy in Spectrophotometry and Luminescence Measurements Conference Proceedings (1972).

[44] H. J Kostkowski, "Reliable spectroradiometry," Spectroradiometry Consulting, P.O. Box 2747, La Plata, MD 20646-2747, USA (1997).

[45] M. Battuello, T. Ricolfi and L. Wang, "Realisation of the ITS-90 above 962°C with a photodiode-array radiation thermometer," *Metrologia*, vol. 32(5), pp. 371-378 (1995/96).

[46] "Measuring the stray light performance of UV-visible spectrophotometers," HP Technical note available on Hewlett Packard CD-ROM *Explore the World of UV-Visible Spectroscopy* ('98 Edition).

[47] Y. Talmi, "Spectrophotometry and spectrofluorimetry with the self-scanned photodiode array," *Appl. Spectrosc*., vol. 36(1), pp. 1-18 (1982).

[48] P. J. Aiello and C. G. Enke, "Wide dynamic range array detector for absorbance and rotation spectrometry," Amer. Chem. Soc. Symposium Series on Multichannel Image Detectors, pp. 57-73 (1983).

[49] "Standard test method for estimating stray radiant power of spectrophotometers by the opaque filter method," ASTM Designation: E 387-84 (Reapproved 1995).

[50] K. D. Mielenz, V. R. Weidner and R. W. Burke, "Heterochromic stray light in UV absorption spectrometry: a new test method," *Appl. Opt*., vol. 21(18), pp. 3354-3356 (1982).

[51] E. F. Zalewski and J. Geist, "Silicon detector nonlinearity and related effects," *Appl. Opt.*, vol. 22(8), pp. 1232-1236 (1983).

[52] C. L. Sanders, "Accurate measurements of and corrections for nonlinearities in radiometers," *J. Res. Natl. Bur. Stand. Sec. A*, vol. 76, pp. 437-453 (1972).

[53] W. Budde, *Physical Detectors of Optical Radiation*, Academic Press, (1983).

[54] K.D. Mielenz and K.L. Eckerle, "Spectrophotometer linearity testing using the double aperture method," *Appl. Opt*., vol. 11(10), pp. 2294-2303 (1970).

[55] W. Budde, "Mulitdecade linearity measurements on silicon photodiodes," *Appl. Opt*., vol. 18(10), pp. 1555-1558 (1979).

[56] C.L. Sanders, "A photocell linearity tester," *Appl. Opt*., vol. 1, pp. 207-210 (1962).

[57] J.C. Zwinkels and D.S. Gignac, "Automated high precision variable aperture for spectrophotometer linearity testing," *Appl. Opt*., vol. 30(13), pp. 1678-1686 (1991).

[58] A. Reule, "Testing spectrophotometer linearity," *Appl. Opt*., vol. 7(6), pp. 1023-1028 (1968).

[59] A. T. Hattenburg and J. B. Shumaker, "Linearity study of a diode array radiometer," *Appl. Opt*., vol. 23 (19), pp. 3257-3258 (1984).

[60] Y. Talmi and R. W. Simpson, "Self-scanned photodiode array: a multichannel spectrometric detector," *Appl. Opt*., vol. 19(9), pp. 1401-1413 (1980).

[61] D. A. Carter and J. E. Pemberton, "Frequency/wavelength calibration of multipurpose multichannel Raman spectrometers. Part I: instrumental factors affecting precision," *Appl. Spectrosc*., vol. 49(11), pp. 1550-1560 (1995).

[62] D. A. Carter, W. R. Thompson, C. E. Taylor and J. E. Pemberton, "Frequency/wavelength calibration of multipurpose multichannel Raman spectrometers. Part II: calibration fit considerations and calibration standards," *Appl. Spectrosc*., vol. 49(11), pp. 1561-1576 (1995).

[63] A. W. Fountain III, T. J. Vickers and C. K. Mann, "Factors that affect the accuracy of Raman shift measuremnets on multichannel spectrometers," *Appl. Spectrosc*., vol. 52(3), pp. 462-468 (1998).

[64] http://chemweb.chem.fsu.edu/editors/tvickers/web/vickers.htm.

[65] D. P. Strommen and K. Nakamoto, *Laboratory Raman Spectroscopy*, Wiley-Interscience, New York (1984).

[66] J. M. Tedesco and K. L. Davis, "Calibration of dispersive Raman process analyzers," *Proc. SPIE*, vol. 3537, pp. 200-212 (1998).

[67] "Guide for Raman Shift Standards for Spectrometer Calibration," in *Annual Book of ASTM Standards* (ASTM, W. Conshohocken, PA), E1840, vol. 2 (1996).

[68] http://www.chemistry.ohio-state.edu.

[69] "Standard test method for spectral bandwidth and wavelength accuracy of fluorescence spectrometers," ASTM standard E388-72 (reapproved 1993), (ASTM, Conshohocken, PA).

[70] Photon Technology International, http://www.pti-nj.com.

[71] K. Iwata, H. Hamaguchi and M. Tasumi, "Sensitivity calibration of multichannel Raman spectrometers using the least-squares fitted fluorescence spectrum of quinine," *Appl. Spectrosc.*, vol. 42(1), pp. 12-14 (1988).

[72] K. G. Ray and R. L. McCreery, "Simplified calibration of instrument response function for Raman spectrometers based on luminescent intensity standards," *Appl. Spectrosc.*, vol. 51(1), pp. 108-116 (1997).

[73] K. J. Frost and R. L. McCreery, "Calibration of for Raman spectrometer instrument response function with luminescence standards," *Appl. Spectrosc.*, vol. 52(12), pp. 1614-1618 (1998).

[74] M. Fryling, C. J. Frank and R. L. McCreery, "Intensity calibration and sensitivity comparisons for CCD/Raman spectrometers," *Appl. Spectrosc.*, vol. 47(12), pp. 1965-1974 (1993).

[75] Kaiser Optical Systems, Inc., http://www.kosi.com.

[76] M. G. Shim and B. C. Wilson, "Development of an in vivo Raman spectroscopic system for diagnostic applications," *J. Raman Spectrosc.*, vol. 28, pp. 131-142 (1997).

[77] J. A. Gardecki and M. Maroncelli, "Set of secondary emission standards for calibration of the spectral responsivity in emission spectroscopy," *Appl. Spectrosc.*, vol. 52(9), pp. 1179-1189 (1998).

[78] M. A. Cutter, D. R. Lobb, T. L. Williams and R. E. Renton, "Integration and testing of the compact high-resolution imaging spectrometer (CHRIS)," *Proc. SPIE*, vol. 3753, pp 180-191 (1999).

[79] P. Martinsen, P. Schaare and M. Andrews, "A versatile near infrared imaging spectrometer," *J. Near Infra. Spectrosc.*, vol. 7, pp. 17-25 (1999).

CHAPTER 5

CALIBRATION EQUIPMENT

In this chapter we give a summary of calibration equipment, together with approximate prices and possible suppliers (for those artifacts that are not widely available). Note that these lists are not necessarily exhaustive and that other suppliers may also be able to provide similar artifacts. In addition, the National Physical Laboratory can supply most of the artifacts described, with calibration where appropriate, and can provide further guidance on the suitability of particular artifacts for specific applications. Additional information can be found in manufacturers' catalogues.

5.1 Sources

Reference 1 gives an extensive discussion of light sources for use in spectroscopy, and Ref. 2 gives a general introduction to a large variety of sources for the UV, visible, and infrared.

5.1.1 Continuum Sources

The lowest calibration uncertainties will usually be achieved using specially designed tungsten lamps, but despite this these are not always the most appropriate choice. The following points must be considered when selecting a standard source:

> *Type of source*. The standard source should generally be as similar as possible to the test sources to be measured.
>
> *Stability and repeatability*. It is important that the calibration values assigned to a reference standard should be maintained (within reasonable limits) throughout its working life. Some types of sources may be subject to sudden, unpredictable changes in output (e.g., most discharge lamps if not carefully selected), while others may require special handling in order to preserve their calibration values.
>
> *Reproducibility*. This refers to how well the alignment of the source can be repeated on each occasion of use and the effect any uncertainty in alignment has on the calibration. Other factors such as temperature may also affect reproducibility. If a source is to be a useful calibration standard, its conditions of use must be sufficiently reproducible that they do not unduly affect the calibration values assigned to it. In this context, the requirements for good reproducibility will be more stringent for a standard for absolute

spectral responsivity calibration than for one to be used only to calibrate the relative spectral responsivity.

Robustness and ease of use. It is often the case that a reference source is calibrated at one location (e.g., a national standards laboratory or other approved calibration laboratory) and subsequently used elsewhere (e.g., the quality control department of a lamp manufacturer). It is therefore important that the calibration values should not change as a result of transport between the two locations; i.e., the source should be robust. Even with lamps specially designed for use as standards, however, hand carriage is always recommended. Normal postage or air freight should be avoided if at all possible. Sources that are difficult to use (e.g., those that require very high starting voltages) typically show poor reproducibility. Thus, although ease of use is not an absolute requirement, it is usually a definite advantage.

It will be clear from the foregoing that many factors need to be considered when selecting a calibration source. By careful selection, identifying appropriate alignment and operating procedures, and by using groups of several sources rather than relying on just two or three of each type, it is often possible to use as calibration standards sources of the same type as those to be measured.

However, in situations where the test sources are not suitable for use as standards, where a wide range of different types of test source are to be measured or where the calibration system is well characterized, enabling accurate assessment of systematic errors, it is often necessary or preferable to use a source specifically designed for use as a standard.

5.1.1.1 CIE Standard Illuminants

The CIE (the International Commission on Illumination) has recommended two illuminants for general purposes[3] and these are sometimes used in the calibration of cameras and other optical radiation instrumentation. The color temperatures and spectral power distributions (but not the intensities) are defined as

Illuminant A	color temperature 2856 K
Illuminant D65	color temperature 6500 K

Illuminant A is a blackbody radiator and is intended to represent typical domestic tungsten-filament lighting. Illuminant D65 is intended to represent average daylight.

The CIE draws an important distinction between a standard illuminant and a standard source: an illuminant is defined in terms of its spectral distribution and may not be possible to realize in practice, whereas a standard source is defined in terms of a physical artifact whose spectral output approximates, but is not identical to, the spectral values defined for the corresponding illuminant. In the case of Illuminant A, a source has been defined (CIE Source A)—this is a tungsten filament lamp operating at a color temperature of 2856 K. No source has been defined for Illuminant D65, and although D65 simulators are available from several manufacturers, these give only a poor approximation to the defined spectral power distribution (especially in the UV region) and they should therefore be used with caution.

5.1.1.2 Selection and Operation of Tungsten Lamps

Tungsten filament lamps are probably the most widely used continuum calibration sources in the visible and near-infrared regions of the spectrum. They fulfill all of the conditions outlined above (being stable, reproducible, robust and easy to use), and in addition they have a continuous, smoothly varying spectral power distribution, which makes them extremely versatile. Various different types of filament construction or optical configuration have been developed to enable tungsten lamps to be used as standards for both of the input optics arrangements described earlier. An example of one such design is shown in Fig. 5.1.1-1.

The spectral power distribution of the radiation from a tungsten lamp is relatively close to that of a Planckian radiator, at least in the visible part of the spectrum. It is, therefore, convenient to characterize it by its color temperature, and for many less-demanding applications it is acceptable to use a lamp set to a specific color temperature as a relative spectral power standard (the spectral distribution is calculated from the color temperature using Planck's law). For applications requiring a lower uncertainty of measurement, direct calibration of the relative or absolute spectral power distribution of the lamp will be necessary.

Vacuum lamps are frequently used as standards at color temperatures up to about 2400 K and have the advantage that there can be no noise on the output due to turbulence in the gas flow. At higher temperatures the rate of evaporation of tungsten from the filament becomes unacceptably high and gas-filled lamps are used. These fall into two main groups. In the 2000 K to ~2900 K range, an inert gas filling (typically argon with 10% to 15% nitrogen) at approximately atmospheric pressure is used to inhibit tungsten evaporation and so extend the life of the lamp.

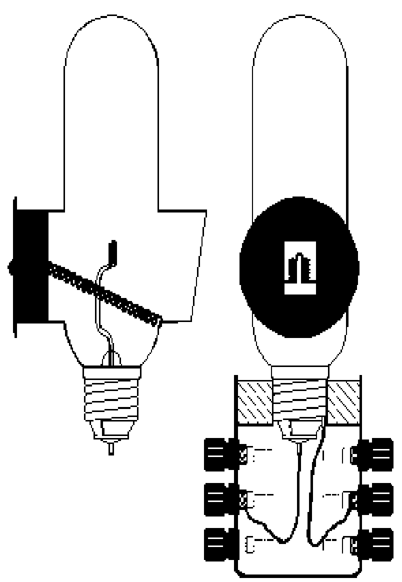


Figure 5.1.1-1 NPL/Polaron tungsten transfer standard.

Above ~2900 K, tungsten halogen lamps are normally used. In these, the gas filling is at between 7 and 10 times atmospheric pressure, resulting in a considerable further reduction in the rate of tungsten evaporation. The gas is pure argon or krypton with a small quantity of a halogen added, which supports a transport cycle that keeps the bulb wall free from deposited tungsten. This transport cycle is essential with the relatively small envelopes required to contain the high pressure within these lamps. Without it, excessive blackening would occur, resulting in a shorter working life. It must be remembered that in order to maintain the halogen cycle the temperature of the bulb wall (and hence the operating temperature of the lamp) must be maintained above a certain minimum level. Operation below this level will lead to rapid blackening of the envelope and possible "tail erosion" of the filament (erosion of the colder parts of the filament assembly by the highly active halogen).

The required color temperature, and hence the type of lamp to be used, is an important point to consider when specifying a standard source, since it determines the proportion of the total power emitted into each region of the spectrum. For example, at 2000 K the ratio between power at 700 nm and at 300 nm is $\sim 1.7\times 10^4 : 1$, while at 3200 K this same ratio is $\sim 80 : 1$. The low proportion of power emitted at short wavelengths even at 3200 K (the highest practical operating temperature for a calibration standard) means that tungsten lamps are difficult to use for measurements at wavelengths much below 300 nm and that there can be significant systematic errors if they are used in this spectral region.

Lamps with glass envelopes transmit radiation over the entire visible spectrum and at wavelengths up to about 2000 nm in the infrared. Below 350 nm, glass absorbs to a certain extent and standard sources intended for use below 350 nm should have either a silica window or an envelope made entirely from silica.

When a *new* tungsten or tungsten halogen lamp is lit up, its characteristics change quite rapidly at first but, after an initial period of operation, tend to become more settled and, unless the lamp is faulty, the rate of change becomes relatively slow and steady. The object of aging a lamp for use as a standard is to take it into this more stable state of operation. Aging behavior varies from lamp to lamp, depending on the design and on the operating temperature, but as a general rule, initial aging for 5% to 10% of the nominal life at the operating current or voltage is recommended.

Once calibrated, care should be taken not to subject lamps to stress by switching them on or off at full current. At switch-on, the current should be gradually increased and after use gradually reduced. After switch-on, lamps need a warm-up period to allow them to reach a state of stable thermal equilibrium. This period typically lies between 5 and 15 minutes, depending partly on the thickness of the filament.

The electrical properties of tungsten and tungsten halogen lamps can be extremely stable, provided that the lamps are not subjected to mechanical shock or excess current or voltage. Either ac or dc operation is satisfactory, so long as the required accuracy and degree of regulation can be achieved. Lamps should always be operated on the same form of current that was used for their calibration. As the measurement techniques involved are easier and less prone to error, dc operation is usually preferred.

Lamps may be operated at either constant current or constant voltage. The degree of stabilization required for current control is twice that needed for voltage control (at a color temperature of about 2856 K, a change of 1% in current or 2% in voltage corresponds to a change of ~20 K in color temperature, ~8% in luminous intensity, and about ±1% change in the slope of the spectral power distribution across the visible), but with modern power supplies adequate precision can be achieved in either mode. Indeed, current control is preferable in many cases, since it is often difficult to exclude the potential drop across socket contacts from a measurement of the lamp voltage. Whichever of the two quantities is controlled, the other is usually monitored as a check on the behavior of the lamp.

For tungsten lamps the polarity may be reversed on each occasion that the lamp is lit or a fixed polarity may be maintained. Where a lamp is of symmetrical construction, polarity reversal has been found in many cases to reduce the drift rate of the lamp, and so extend the length of time for which it can be used before recalibration is required. Asymmetric lamps, such as ribbon lamps and tungsten halogen lamps, must always be operated at fixed polarity (usually positive). If it is decided to use polarity reversal for a symmetrical lamp, this regime must be followed consistently. Running the lamp on one polarity for an extended period (several hours) and then changing the polarity will result in a rapid change in output.

Gas-filled lamps with large envelopes tend to exhibit short-term (typically 0.2 to 2 Hz) convection-induced fluctuations of output, or noise. Fluctuations in intensity of 0.5% are not uncommon, and in some cases they may be as great as 3% or 4%. Where such fluctuations are present it may be necessary to use an average value collected from several consecutive scans of the same source, and this feature is commonly available with most array systems.

Tungsten halogen lamps operate at internal gas pressures of 7 to 10 atmospheres, and can therefore explode. It is desirable that they should be operated inside a lamp house or that screens are used to protect the operator from possible explosion. Operation in a lamp house may increase the aging rate (the lamp house will age as well as the lamp) and will also affect the calibration values by increasing the filament temperature and by causing radiation from the surface of the lamp house behind the source to be reflected in the forward direction. Such lamps should therefore always be used in the lamp house in which they were calibrated.

Lamps made for general use are normally not suitable as transfer standards. Hooks providing intermediate support for a lamp filament are a potential source of instability since contact pressure and hence thermal loss can vary during the life of the lamp. Thus, the filaments of lamps designed for use as standards either have no intermediate supports at all, or have all such supports welded to the filament. In lamps made for general lighting service, such hooks are never welded and these lamps are therefore best avoided.

Lamps used as calibration standards for spectral irradiance and spectral radiant intensity are usually operated cap down (as shown in Fig. 5.1.1-1) and calibrated in a specified horizontal direction. It must be possible to set up the lamp on each occasion of operation in exactly the same position relative to the system input optics. This can be achieved by the use of a telescope to align the plane of the filament assembly, by the alignment of sighting marks on the front or back of the lamp or its housing, or by the reflection of a laser beam from a suitable plane surface such as the front window (where appropriate). Although it is not essential for a source of irradiance to obey the inverse square law if it is always used at the distance at which it was calibrated, it must be constructed so that the calibration distance can be measured precisely from a reference point on the lamp, the lamp enclosure, or the lamp mount. It is often useful, however, if a lamp used for this purpose does obey the inverse square law, at least over a limited range of distances, so that it can be used to provide a range of irradiance levels.

Lamps calibrated as standards of spectral radiant intensity are generally assumed to obey the inverse square law, and it must therefore be possible to define the effective light center. The area of the filament should also be small compared with the distance between the lamp and detector. As a general rule, a minimum distance of 20 times the largest dimension of the filament should be used. However, even for these lamps it is always advisable to check for compliance with inverse square law behavior in critical applications.

Another requirement for a source used as a standard of irradiance is that its field should be uniform, preferably to better than ~0.25% over the irradiated area or angle (typically at least 6^{o}). Coiled filaments, especially single coils arranged in a regular pattern, can show rapid changes of intensity with angle of view as the front turns mask the rear turns to a greater or lesser extent. Uniformity is improved if the lamp envelope is diffusing, but if clear, it should be of good optical quality. Lamps are often grit blasted to achieve more uniform irradiance, but the surface is then vulnerable to contamination and consequent discoloration. Furthermore, where a lamp has a diffusing envelope, it is almost impossible to define the position of the light center; i.e., the inverse square law is unlikely to be obeyed.

Ribbon filament lamps can be used as spectral radiance standards when higher power levels are required, e.g., if the effective wavelength range of the array system extends down to 350 nm or even 300 nm. These are normally operated cap down, with the ribbon vertical, and have a plane window of glass or silica to permit good optical imaging. The calibration applies to radiation about an axis, normally horizontal, from a specific area of the ribbon. The calibrated area must be readily identifiable, and for this reason a pointer is often provided adjacent to the ribbon, or a small notch may be cut into the ribbon itself. Because the ribbon can move relative to the lamp base during warm-up, the alignment should at least be checked once the lamp has reached its final operating temperature.

5.1.1.3 Selection and Use of Other Types of Continuum Sources

Where standards for lower power levels are required there are typically two basic types. The most fundamental is a plane surface of known luminance or spectral radiance factor, such as a barium sulphate or magnesium oxide plaque or a calibrated white opal, which is irradiated using a standard of known spectral irradiance (see Fig. 5.1.1-2). The diffuser acts as a secondary source and is usually irradiated normally and viewed at 45^{o}.

In the more general case, where a source of radiant intensity I_e irradiates a diffusing surface at an angle θ_A to the normal, which is then viewed at angle θ_B to the normal, the radiance is given by

$$L = \frac{I\sigma(\theta_A, \theta_B)\cos\theta_A}{\pi d^2}, \tag{5.1.1-1}$$

where $\sigma(\theta_A, \theta_B)$ is the radiance factor of the diffuser under these conditions of irradiance and view.

A less direct but widely employed standard is a luminance gauge. This often consists of a small integrating sphere coated internally with a white diffusing material of high reflectance, such as barium sulphate paint, with an illuminating source located either inside or outside the sphere. An example of this is described in Sec. 4.2.5.

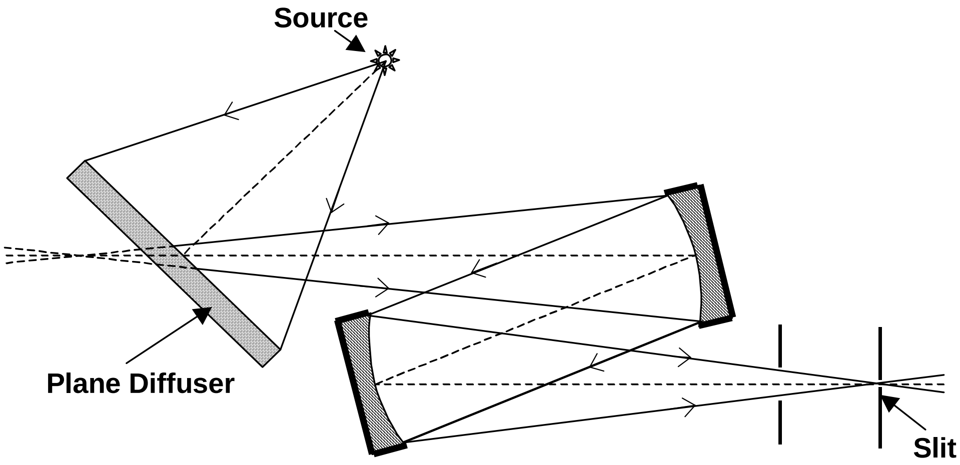


Figure 5.1.1-2 Use of a plane diffuser and a lamp to provide a standard of spectral radiance.

Whatever type of radiance source is chosen, the reference direction and alignment method, the location and size of the calibrated area, and the solid angle subtended by the optical system of the detector must all be specified. When using the integrating-sphere-type source as a standard of spectral radiance, it is also necessary to ensure it has been calibrated against another standard of known spectral radiance. It is not sufficient to calculate the spectral distribution from the color temperature, since this can lead to very large errors, particularly in the blue spectral region (the reflectance of most integrating sphere coatings falls rapidly in the blue).

Tungsten and tungsten halogen lamps radiate very little power in the UV region of the spectrum even when operated at high color temperatures. Where spectral standards are required to be below ~300 nm, a different type of lamp is required. Deuterium discharge lamps are the usual choice, covering a wavelength range from ~ 370 nm down to 200 nm, and indeed with special windows of magnesium fluoride into the vacuum ultraviolet. The form of the spectral power curve complements that of a tungsten lamp, falling off at longer wavelengths (Fig. 5.1.1-3). For wavelengths above 380 nm, spectral emission lines intrude on the continuous spectrum, so deuterium lamps are not suitable as standards beyond this point. However, there is still a useful overlap of the measurable spectrum of tungsten and deuterium lamps between 300 nm and 380 nm.

The relative spectral emission characteristic of a deuterium discharge is a line-free continuum in the UV region. It is quite reproducible from lamp to lamp (within ~ ±10%), even between lamps from different manufacturers, provided similar-quality synthetic fused silica window material is used. Furthermore, it is not very sensitive to discharge current (<5% change between 200 and 500 mA). Absolute emission levels are much more variable between different lamps. For any individual lamp, the absolute level is approximately proportional to the discharge current, and this determines the required stability of the constant current power supply. A stability of 0.1% is readily achieved with economically priced but dedicated power supplies.

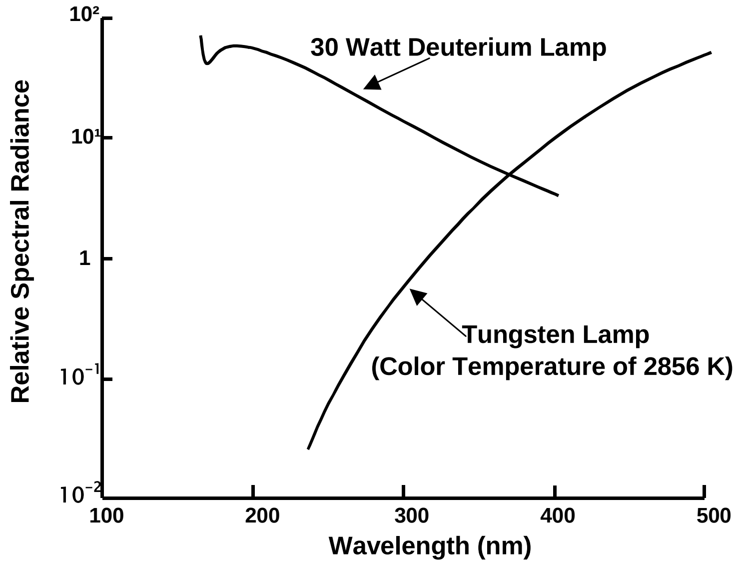


Figure 5.1.1-3 Spectral distribution of a tungsten lamp and a deuterium lamp.

The radiance of the emission from a circular aperture deuterium lamp is a maximum at the center of the emitting aperture and decreases in a radially symmetric manner towards the edge. For use as a radiance standard it is therefore essential to be able to clearly define a specified, ideally circular, emitting area, as the integrated radiance will vary with field size. Typically, the radiance decreases by a factor of 2 between 0.3-mm and 0.9-mm-diameter field sizes. Even if the lamp is to be used only as a relative spectral radiance standard, the field size must be specified, as there is also a variation of relative spectral distribution with field size. (The effect is smaller in this case, usually not exceeding 2% between field sizes of 0.3 mm and 0.9 mm over the range 200 to 350 nm.). Used as radiance standard with a field size of 0.6 mm or less, the directional variation of radiance is usually small (<1% over $\pm$ 3^{o}). The variation of spectral irradiance with angle of view can be much greater, however; 1% per degree is quite common. This arises from several sources, including contributions to the observed emission from the discharge glow around the cathode and asymmetry of the main discharge around the aperture. In order to obtain the best repeatability when using any design of deuterium lamp as an irradiance standard, it is recommended that a lamp with a good quality, squarely mounted window should be selected. Using a helium-neon laser or similar arrangement it is then possible to align the window perpendicular to the optical axis to much better than 1^{o} on each occasion of use, enabling reproducible results to be obtained.

Although the deuterium lamp is not a perfect point source, it approaches this ideal more closely than many other sources. Deviations from the inverse square law are usually less than 2% over distances from 500 mm down to 150 mm, a typical working range for deuterium irradiance standards. It should be noted that it can be difficult to accurately determine the emission point by visual inspection of the lamp; calculation based on the assumption that the lamp obeys the inverse square law may often be advantageous.

The emission level in a deuterium lamp depends, among other factors, on the temperature and pressure in the discharge. As it may take from 10 to 20 minutes or more for internal thermal equilibrium to be established after ignition, a warm-up period of about 30 minutes is normally recommended.

Most noise and other random fluctuations are caused by changes in the position of the discharge, either on the cathode or, less frequently, at the anode. Short-term noise is generally low on most modern deuterium lamps, and peak-to-peak fluctuations are normally 0.1% or less in a 1-Hz bandwidth. However, reproducibility of level from ignition to ignition is one of the most frequently encountered problems in selecting deuterium lamps for use as standards. A series of tests must be carried out in which the lamp is repeatedly ignited and its emission monitored over a period of 60–90 minutes. A suitable cooling down period between ignitions of at least 30 minutes is needed. A minimum of at least 20 ignitions is desirable. In addition to establishing reproducibility of level after the 30 minutes warm-up, drift, noise, and other random fluctuation effects can be checked. It is also advisable to check for large drifts in absolute spectral power over a period of about 100 hours (lamps with drifts >5%–10% over this period are not suitable as standards). These measurements should normally be carried out after an initial aging of 100 hours; and once all the tests have been completed, the lamp will have been run for a total of about 250 hours.

Deuterium lamps produce potentially hazardous levels of UV radiation. Care should be taken to protect the eyes and skin from excessive exposure.

5.1.2 Line Sources

As mentioned in Sec. 4.2.2, a variety of emission line sources are available, chiefly for wavelength calibration of spectrometers (see, for example, Chapter 10 of Ref. 1).

5.1.2.1 Pencil Discharge Lamps

One of the most common lamps for wavelength calibration is a pencil-sized low-pressure cold cathode discharge source (see, for example, Ref. 4). These can usually be held with simple laboratory clamps and operated in any orientation. Common fillings are Xe, Ar, Ne, Kr, and Hg. Low-pressure Hg(Ar) lamps are often used because of their stable intensity level and long life. Some manufacturers also offer a Hg(Ne) lamp. When this lamp is operated at normal laboratory ambient temperature, the output is very similar to that of a Hg(Ar) lamp (giving characteristic Hg lines). However, when forced air cooled the lamp shows additional Ne lines out to 2651 nm in the NIR. Zinc and cadmium lamps are also available (and are often used for calibration in the UV). A battery-powered Hg(Ar) source can be used for low-cost or portable equipment.

Pencil lamps usually have a short initial warm-up period but need ~30 minutes for stabilization. AC power supplies are not usually recommended for array-based instruments because of the possibility of beating effects between the fluctuating lamp output and the clocking signals for the array. Note, however, that prolonged use in a single-polarity dc mode will shorten the lifetime of the lamp due to electrophoresis effects.

5.1.2.2 Hollow-Cathode Discharge Lamps

Hollow cathode discharge lamps are more intense, have a larger source area than pencil-type lamps and are available with cathode materials of more than 60 different elements or alloys (e.g., Cd, Cs, Hg-Cd, K, Na, Ne, Tl, Zn, and Hg). The elements in the cathode are sputtered into the discharge and emit line spectra. They operate in AC mode but the average output (over several seconds) is usually stable. As the power to the lamp increases, so does the width of the spectral lines. Due to the differing energies of the electrons in the various zones of the discharge, the source region tends to be inhomogeneous. Hollow cathode lamps are used extensively in atomic absorption spectroscopy (AAS); further details can be found in Ref. 5.

5.1.2.3 Electrodeless Discharge Lamps (EDLs)

Electrodeless discharge lamps can be used if particular calibration wavelengths cannot be achieved with either of the above types of discharge lamp. They give high radiant intensities and narrow line widths. Typically the lamp is made of fused quartz and contains the gas of interest or a metal halide with a rare gas filler. Excitation is via a microwave generator that induces an inductively coupled discharge. In general, these lamps are better able to produce stable radiation of sharp spectral lines than arc lamps and, unlike the inductively coupled plasma (ICP) source which operates at atmospheric pressure, a low-pressure plasma is generated in the EDL (see, for example, Ref. 5).

5.1.2.4 Deuterium Lamps

Deuterium lamps have reasonably strong emission lines at 486 nm and 656 nm, and these are often used for wavelength checks on spectrometers.

5.1.3 Other Sources

There are a large number of other sources that are used in connection with electro-optical systems. A brief discussion is given below. Further details can be found in Ref. 2 and manufacturers' literature.

- Arc lamps, such as Xe arcs, provide an intense continuum radiation (with superimposed spectral lines) and are useful for spectroscopy (particularly as an excitation source) and as an illuminator for optic–optic systems. The intensity and position of the arc tend to be unstable because of convection currents inside the lamp and arc "wander" on the electrodes. Stabilizing controllers are available from some manufacturers but tend to be relatively expensive. In general, arc sources need to be operated vertically or else convection currents cause uneven heating of the quartz envelope (which shortens the lifetime).
- LEDs provide moderately broadband illumination and have the advantage of compactness and low cost. They operate at low voltages and can be pulsed. A stable light source can be achieved by controlling the supply current, but since the output is

temperature sensitive it is usually desirable to use optical feedback. Illumination levels are low and so LEDs are not commonly used for spectrometric work.

- Phosphor sources (e.g. tritium betalights[6]) are useful for stable low-intensity broadband illumination.
- Blackbodies are often used as spectral radiance standards at wavelengths greater than about 2.4 µm and they can also be used as a near-uniform flat-field source. A large range of temperatures are available. They are rarely used for calibrations in the visible (except in cases when the very lowest uncertainty is required).
- Other continuum sources for the IR are divided into two types: (a) metal oxide, silicon carbide (e.g. Globar), or ceramic rigid sources, and (b) metal filament (coilform) sources. Sources of instability are air turbulences in front of the source (reduced by using an insulating enclosure), voltage fluctuations and material changes (e.g. oxidation). Further details are given, for example, in Ref. 7.

5.1.4 Suppliers of Light Sources

Source type	Use	Effective wavelength range (nm)	Suggested suppliers of uncalibrated artifact	Calibration type
Mercury, neon, argon discharge lamps	Wavelength calibration	250 – 775	Most good optical equipment suppliers	None required
Specially-designed tungsten lamp transfer standards (coiled or ribbon filament)	Spectral responsivity calibration, stray light checks[1]	350 – 2500	Polaron, Gigahertz-Optik, Gamma Scientific or other specialist lamp manufacturers	Spectral irradiance / radiance
Tungsten halogen lamps and other normal production lamps	Spectral responsivity calibration, stray light checks[1]	300 – 2500	Most lamp manufacturers or good optical equipment suppliers	Spectral irradiance / radiance
Deuterium discharge lamp	Spectral responsivity calibration, stray light checks[1]	200 – 400	Cathodeon, Hamamatsu, Heraeus	Spectral irradiance / radiance
Variable luminance gauge	Linearity, spectral responsivity calibration	350 – 2500	NPL, Labsphere, Glen Spectra, LOT Oriel or other good optical equipment suppliers	Spectral radiance

[1]Used with stray light filters described below

5.2 Filters

5.2.1 Neutral Density Filters

Neutral density filters are used for testing linearity and for controlling the level of illumination (e.g., when performing flat-fielding or responsivity tests). There are two basic types: reflective metal film and absorptive. Metal film filters are neutral over a large wavelength range (typically showing ~10% variation over the 400- to 2400-nm range), but since they reflect the unwanted radiation, great care is needed to make sure that the reflected light does not have a path onto the detector—especially when stacking several filters together. However, there is the advantage that the unwanted radiation does not heat up the filter. Either glass or fused quartz substrates can be used. Some manufacturers will provide a transmittance versus wavelength calibration curve.

Absorptive filters are usually glass and show a greater variation of transmittance with wavelength; however, stacking of filters is easier than with metal film filters and the cost is usually lower. When using dense filters and intense illumination (e.g., from a Xe arc) care has to be taken that the filter is not damaged by heat caused by the absorbed radiation.

A number of precautions are necessary when using neutral density filters:

- The transmittance will depend on the path length through the filter, which in turn is affected by its orientation and the cone angle of the incident radiation. Generally, filters are calibrated using quasi-collimated radiation and with the filter perpendicular to the beam.
- For a nonuniform filter, the transmittance will also vary with the size and position on the surface of the incident radiation.
- Inter-reflections between filters (if they are being used in combination) or between the filter and the other elements of the system (lenses, apertures, monochromator, etc.) may affect measured transmittance (especially with metal film filters, as discussed above).
- Surface contamination of the filter will also affect its transmittance and may also result in increased degradation of the filter with time (i.e., large changes in filter transmittance with time).

5.2.2 Other Filters

Interference, edge, and notch filters are used to isolate wavelength regions and to measure stray light (as discussed in Sec. 4.2). The precautions needed are generally as for neutral density filters.

5.2.3 Suppliers of Filters

Filter type	Use	Effective wavelength range (nm)	Suggested suppliers for uncalibrated artifact	Type
Holmium oxide glass	Wavelength calibration	270 – 700	Unicam, Optiglass, NPL, NIST	Wavelength peaks
Holmium oxide solution	Wavelength calibration	230 – 700	Most good spectrophotometry suppliers	Wavelength peaks
Crystalline	Wavelength calibration	250 – 1800	McCrone Research Associates[2]	Wavelength peaks
Edge	Stray light check	Varies with filter	Most good optical equipment suppliers	None required
Notch	Stray light check	Varies with filter – blocks one narrow wavelength range	Most good optical equipment suppliers	Transmittance
Liquid	Absorbance / stray light check / linearity	Varies with solution	Most good spectrophotometry/ chemicals suppliers	Transmittance
Neutral density	Linearity/ dynamic range	350 – 800	Most good optical equipment suppliers	Transmittance

[2] Available from NPL

5.3 Ceramic Tiles

Tile type	Use	Effective wavelength range (nm)	Suggested suppliers for uncalibrated artifact	Type and *approximate cost of calibration*
Holmium oxide-doped	Wavelength calibration	350 – 700	Ceram, Avian Technologies	Reflectance
Reflectance	Linearity, responsivity	300 – 3000	Ceram, Avian Technologies	Reflectance

5.4 Grids and Resolution Charts

Type	**Use**	**Effective wavelength range (nm)**	**Suggested suppliers for uncalibrated artifact**	**Type and *approximate cost of calibration***
Chrome on glass or quartz grid targets	Geometrical Calibration	visible	Pyser-SGI	Some targets are supplied with UKAS or NPL certificates for an additional cost
Resolution Targets	Geometrical Calibration	visible	Most good optical equipment suppliers	-

5.5 Cells for Absorption Spectroscopy

In absorption spectroscopy liquid samples are commonly placed in cells or cuvettes, or transfer pipes can continuously flow liquid through a sample chamber. These cells are usually of glass or quartz. A variety of sizes, window materials light path, etc., are available and some manufacturers will provide a transmission calibration and polarization check.

In optically thin cells it may be necessary to avoid optical interference effects arising from multiple reflections; and in optically thick cells, differences in the refractive index between the sample and an empty cell may cause changes in the optical path (e.g., focus shifts).

It is important to keep cells clean and free from scratches (Ref. 8 provides some simple guidelines).

5.6 Contact Details

Avian Technologies	www.aviantechnologies.com
Hamamatsu	www.hamamatsu.com
Optiglass	www.optiglass.com
Cathodeon	www.cathodeon.co.uk
Heraeus	www.heraeus-noblelight.com
Polaron	www.polaron-group.co.uk/ components/spec_lamps.htm
Ceram Research	www.ceram.co.uk
Labsphere	www.labsphere.com
Pyser-SGI	www.pyser-sgi.com
Gamma Scientific	www.gamma-sci.com
L.O.T.-Oriel Ltd	www.lotoriel.co.uk
Unicam	www.unicaminstruments.com
Gigahertz-Optik	www.gigahertz-optik.de
National Physical Laboratory	www.npl.co.uk/optical_radiation
Glen Spectra Ltd	www.isa-gs.co.uk
NIST	www.nist.gov

References

[1] H. J Kostkowski, *Reliable Spectroradiometry*, Spectroradiometry Consulting, P.O. Box 2747, La Plata, MD 20646-2747, USA (1997).

[2] *The Infrared and Electro-Optical Systems Handbook*, *Vol. 1: Sources of Radiation,* G. J. Zissis, ed., Infrared Information Analysis Center, Ann Arbor/SPIE Optical Engineering Press, Bellingham, WA (1993).

[3] "CIE standard illuminants for colorimetry," CIE Standard S 005/E-1998, republished as ISO 10526-1999/CIE S 005-1998.

[4] Ultra Violet Products Ltd, http://www.uvp.com.

[5] B. Welz and M. Sperling, *Atomic Absorption Spectroscopy, Third Edition*, Wiley-VCH, Weinheim (1999).

[6] Saunders-Roe Ltd., http://www.betalight.com.

[7] J. Coates, "Vibrational spectroscopy: instrumentation for infrared and Raman spectroscopy," *Appl. Spectrosc. Rev*., vol. 33(4), pp. 267-425 (1998).

[8] Hellma (England) Ltd., http://www.hellma.co.uk/.

CHAPTER 6

COMPARISON OF INTERNATIONAL STANDARDS

Many countries around the world have established scales for the measurement of optical radiation. These include luminous intensity, luminous flux, illuminance, luminance, spectral radiance and irradiance, spectral total flux, spectral responsivity, spectral reflectance and transmittance. Many countries also operate measurement accreditation services. Laboratories with accredited calibration services have been assessed by the accreditation body to ensure that the measurements are both traceable to national standards and carried out under a formal quality system. A calibration certificate from an accredited laboratory provides the user with confidence in the results contained within it, through the assurance that the laboratory is regularly inspected, its measurement results are subject to periodic audits and its measurement standards are fully traceable to national standards.

For the purposes of international trade it is important that (1) calibration certificates and reports from accredited laboratories are accepted internationally without the need for further assessment, recalibration, or testing; and (2) the measurement scales established at laboratories around the world are in agreement with each other to within known limits.

The first point has been addressed through the establishment of multilateral Mutual Recognition Agreements between accreditation bodies. Most of the major accreditation bodies are signatories to these agreements, thereby ensuring the formal recognition of the technical equivalence of their accreditation services and the calibration certificates issued by accredited laboratories.

In order to address the second point, in October 1999 the directors of the national metrology institutes (NMIs) of 38 Member States of the Metre Convention and representatives of two international organizations signed a Mutual Recognition Arrangement (MRA) for national measurement standards and for calibration and measurement certificates issued by NMIs. The objective was to establish the degree of equivalence of different national measurement systems and to make information on the relationship between national standards widely available. Since then, a number of other institutes have signed. The MRA of the Comité International des Poids et Mesures (CIPM) has now been signed by the representatives of 59 institutes—from 44 Member States, 13 Associates of the Conférence Générale des Poids et Mesures (CGPM), and 2 international organizations—and covers a further 77 institutes designated by the signatory bodies.

The MRA is a response to a growing need for an open, transparent and comprehensive scheme to give users reliable quantitative information on the comparability of national metrology services and to provide the technical basis for wider agreements negotiated for international trade, commerce, and regulatory affairs. It defines "Key Comparisons," which are organized by the Consultative Committees of the CIPM, which are intended to provide data on the degree of equivalence between national measurement systems. The outcome of these comparisons is entered onto the key comparison database (KCDB) that is maintained by the Bureau International des Poids et Mesures (BIPM) in Paris and is accessible over the internet. In this way information can be made available on the degree of equivalence between national standards laboratories to anyone requiring it. In practice this will only be important to those making measurements close to the state-of-the-art, as differences between national standards realized from the definitions of SI are normally small.

Further information on the MRA is available on the BIPM website at http://kcdb.bipm.org, organized into three main parts:

1. the MRA itself
2. the Joint Committee of the Regional Metrology Organizations and the BIPM (JCRB)—its role, composition, documents and membership
3. the BIPM key comparison database (KCDB)—which includes the results of key and supplementary comparisons and the calibration and measurement capabilities (CMCs) of the national metrology institutes signatories to the MRA and the other designated institutes.

Work to complete all the Key Comparisons will take some years, but good progress is already being made. Photometric Key Comparisons have now been completed; the results are available on the KCDB and are summarized in Fig. 6-1. Measurements are also underway for Key Comparisons of spectral irradiance, spectral radiance, regular spectral transmittance, spectral diffuse reflectance, and spectral responsivity; the current status of these comparisons is given on the KCDB and results will be posted when these become available.

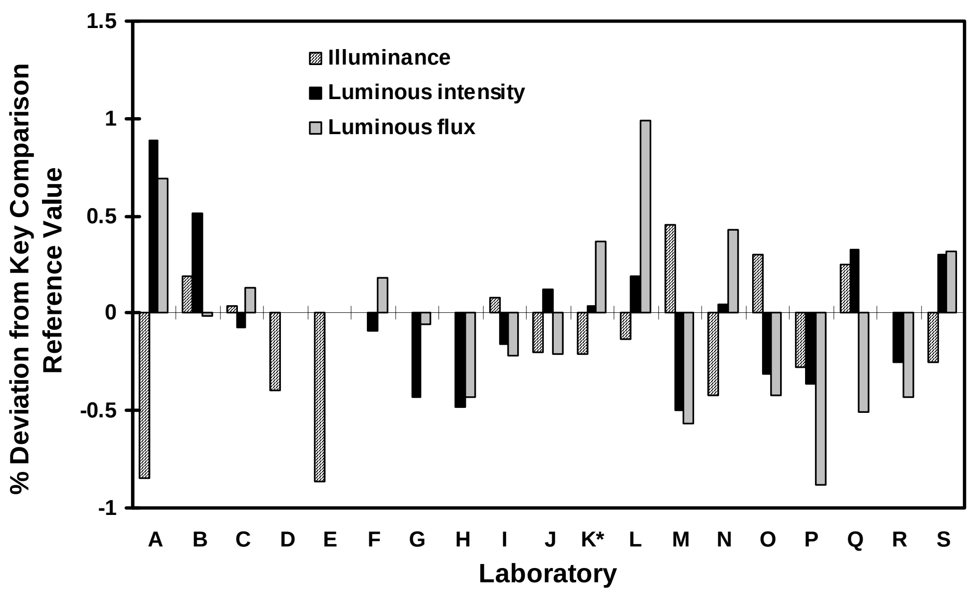


Figure 6-1 Results of the recent photometric key comparisons.

APPENDIX A

REFERENCE GUIDE

A.1 Calibration Services and Contacts

The calibration of artifacts discussed in Chapters 4 and 5 should be traceable to national standards. As described in Chapter 6, the best guarantee of traceability is to use a calibration laboratory that has been accredited by a body such as UKAS, and this approach is strongly recommended for all the artifacts mentioned in this guide.

A.1.1 Accreditation

Further information on accredited laboratories is available from the following:

UKAS Directory of Accredited Laboratories (contact United Kingdom Accreditation Service, 21-47 High Street Feltham, Middlesex, TW13 4UN. Tel. 020 8917 8400). Also available at **www.ukas.org.**

The U.S. National Institute of Standards and technology (NIST) operates a National Voluntary Laboratory Accreditation Program (NVLAP), the NIST website (**www.nist.gov**) gives details of the laboratories involved.

The European co-operation for Accreditation (EA) (**http://www.european-accreditation.org/**) gives links to directories of accredited laboratories in each European country.

The Asia Pacific Laboratory Accreditation Cooperation (APLAC) groups accreditation bodies in the Asia Pacific region responsible for accrediting calibration, testing and inspection facilities. Its website is **http://www.aplac.org/** .

The International Laboratory Accreditation Cooperation (ILAC) website (**http://www.ilac.org/**) maintains links to directories of all accredited laboratories in member countries.

A.1.2 Calibration

Information on calibration and related topics can be obtained from the following:

http://www.npl.co.uk —the site for the U.K. National Physical Laboratory. This gives details of optical radiation measurement and advisory services, the Optical Radiation Measurement (ORM) club and newsletter, programs, events, and contacts, as well as links to the sites of other national calibration laboratories.

http://members.eunet.at/cie —the site for CIE, the international commission on illumination. This gives news on optical calibration issues and events, details on technical committees, and links to other sites.

http://www.nist.gov —the site of the U.S. National Institute for Science and Technology. **http://www.nist.gov/srd/online.htm** gives a collection of online data, including emission line wavelengths.

http://www.astm.org —the site for the American Society for Testing and Materials. This has a searchable database of test standards and these can be ordered online by credit card.

http://www.iso.ch —the home page for the International Organization for Standardization (ISO).

A.2 Bibliography—Including General Articles, Books, and Useful Websites

A.2.1 General

A useful website for all aspects of engineering is the Edinburgh Virtual Library. One hundred twenty-two records on optics technology can be accessed via the Electrical, Electronic and computer Engineering topic, followed by selecting Light and Optical Technology. The address is **http://www.eevl.ac.uk.**

Research papers on detectors and optical instrumentation can be found in the following journals:

Proceedings of SPIE (list of volume titles and abstracts available from http://spiedl.org)
Optical Engineering
Applied Optics

Research papers on spectroscopy can be found in

Applied Spectroscopy (abstracts available from **http://www.s-a-s.org/**)
Analytical Chemistry
Spectroscopy

A.2.2 Arrays

The following websites provide information on CCD camera manufacturers and CCD manufacturers. Most detector manufacturers have their own websites (often containing useful tutorials and technical notes, as well as product information).

http://www.pacificsites.com/~brooke/ACCD.shtml
http://canopus.saao.ac.za/~wpk/ccdworld.html
http://www.aps.anl.gov/xfd/Detectors/pages/ccd.html

Books, general papers

A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices* (Kluwer, Dordrecht, 1995).

J. M. Harnly and R. E. Fields, "Solid-state array detectors for analytical spectrometry," *Appl. Spectrosc.*, vol. 51(9), pp. 334A-351A, (1997).

Selected Papers on CCD and CMOS Imagers, M. G. Kang, Ed., (SPIE Press, Bellingham, WA, 2003).

J. Janesick, *Scientific Charge-Coupled Devices* (SPIE Press, Bellingham, WA, 2001).

A. Rogalski, *Infrared Photon Detectors*, (SPIE Optical Engineering Press, Bellingham, WA, 1995).

A.2.3 Instrumentation

The website **http://color.psych.ucsb.edu/psychtoolbox/measure.html** provides information on a variety of manufacturers of light measurement equipment.

The website **http://www.s-a-s.org/** is the site of the U.S. Society for Applied Spectroscopy.

The website **http://spectroscopymag.com** gives useful articles and links to several spectroscopy resources

ASTM standards:
There are several recent standards that are relevant to the calibration of spectrometers:

- E1655-97el, Standard practices for infrared multivariate, quantitative analysis
- E388-72(1998), Standard test method for spectral bandwidth and wavelength accuracy of fluorescence spectrometers
- E1840-96, Standard guide for Raman shift standards for spectrometer calibration

Useful books and papers

Charge-Transfer Devices in Spectroscopy, J. V. Sweedler, K. L. Ratzlaff and M. B. Denton, Eds., (VCH Publishers Inc., NY, 1994).

C. Burgess and T. Frost, eds., *Standards and Best Practice in Absorption Spectrometry* (Blackwell Science, Oxford, UK, 1999).

H. J. Kostkowski, *Reliable Spectroradiometry*, Spectroradiometry Consulting, P.O. Box 2747, La Plata, MD 20646-2747, USA (1997).

T. L. Williams, *The Optical Transfer Function of Imaging Systems* (Institute of Physics Publishing, Bristol, 1999).

G. C. Holst, *Testing and Evaluation of Infrared Imaging Systems* (JCD Publishing Co., Maitland, USA, 1993).

G. C. Holst, *CCD Arrays, Cameras and Displays* (JCD Publishing, Maitland, 1996).

The Infrared and Electro-Optical Systems Handbook, in particular, Vol. 1: Sources of Radiation, G. J. Zissis, ed., and Vol. 3: Electro-Optical Components, W. D. Rogatto, Ed, (Infrared Information Analysis Center, Ann Arbor/SPIE Optical Engineering Press, Bellingham, USA, 1993).

C. Poynton, *A Technical Introduction to Digital Video* (John Wiley & Sons, 1996).

P. Colarusso, L. H. Kidder, I. W. Levin, J. C. Fraser, J. F. Arens and E. N. Lewis, "Infrared spectroscopic imaging: from planetary to cellular systems," *Appl. Spectrosc*., vol. 52(3), pp. 106A-120A (1998).

J. R. Ferraro, K. Martin and R. J. Jarnutowski, "Five year summary of commercial instrumentation for absorption spectroscopy (1992-1996)," *Spectroscopy*, vol. 12(8), pp. 18-47 (1997).

J. Tellinghuisen, "Statistical error calibration in UV-visible spectrophotometry," *Appl. Spectrosc.*, vol. 54(3), pp. 431-437 (2000).

A.2.4 Calibration

http://www.npl.co.uk is the site for the U.K. National Physical Laboratory. This gives details of optical radiation measurement and advisory services, the Optical Radiation Measurement (ORM) club and newsletter, programs, events and contacts as well as links to the sites of other national calibration laboratories.

http://members.eunet.at/cie is the site for CIE, the international commission on illumination. This gives news on optical calibration issues and events, details on technical committees and links to other sites. The U.K. arm of the CIE, the CIE-UK can be found at **http://www.cibse.org/cieuk/** .

http://www.nist.gov is the site of the U.S. National Institute for Science and Technology. **http://www.nist.gov/srd/online.htm** gives a collection of online data, including emission line wavelengths.

http://www.astm.org is the site for the American Society for Testing and Materials. This has a searchable database of test standards and these can be ordered on-line by credit card.

http://www.iso.ch is the home page for the International Organization for Standardization (ISO).

A.3 List of Abbreviations

AAS	atomic absorption spectroscopy
AC	alternating current
ADC	analog to digital converter
ADU	ADC unit
AES	atomic emission spectroscopy
AOTF	acousto-optical tunable filter
APS	active pixel sensor
AR	anti-reflection (coating)
ASIC	application specific integrated circuit
BDI	buffered direct injection
BIPM	Bureau International des Poids et Mesures
BSI	British Standards Institute
CAD	cascode amplifier per detector
CARS	coherent anti-Stokes Raman spectroscopy
CCD	charge-coupled device
CCPR	Consultative Committee for Photometry and Radiometry
CDS	correlated double sampling
CFA	color filter array
CGPM	Conférence Générale des Poids et Mesures
CID	charge injection device
CIE	Commission Internationale de l'Éclairage or International Commission on Illumination
CIPM	Comité International des Poids et Mesures
CTE	charge transfer efficiency of a CCD
CTF	contrast transfer function
CTI	charge transfer inefficiency of a CCD
CTIA	capacitative transimpedance amplifier
CVF	charge-to-voltage conversion factor
DC	direct current
DI	direct injection
DFT	discrete Fourier transform
DNL	differential nonlinearity (of an ADC)
DSNU	dark signal nonuniformity
E	exposure
EBI	equivalent background illumination (for an intensified array)
EDL	electrodeless discharge lamp
EM	electron multiplying (CCD)
ESF	edge spread function
EPROM	electroically programmable read-only memory
ε	charge transfer inefficiency for one stage of a CCD
FLIM	fluorescence lifetime imaging microscopy
FT	Fourier transform
FWHM	full width half maximum
GD	glow discharge
GLP	good laboratory practice

GMI	gate modulation input
GR	genetic regression
GRF	geometric response function
HCTE	horizontal charge transfer efficiency of a CCD
HCTI	horizontal charge transfer inefficiency of a CCD
HPLC	high-performance liquid chromatography
IC	integrated circuit
ICCD	intensified CCD
ICP	inductively coupled plasma
IEC	International Electrotechnical Commission
I_{DD}	power supply current
INL	integral nonlinearity (of an ADC)
λ	wavelength
LCTF	liquid crystal tunable filter
LE	linearity error
LED	light-emitting diode
LPDA	linear photodiode array
LSF	line spread function
LUT	look-up table
LWIR	long-wave infrared
MAPS	monolithic active pixel sensor
MCP	microchannel plate
MWIR	mid wave infrared
MRA	mutual recognition arrangement
MTF	modulation transfer function
NIR	near-infrared
NIST	National Institute for Science and Technology
NMI	National Metrology Institute
NPL	National Physical Laboratory
OECF	opto-electronic conversion function
OES	optical emission spectroscopy
OTF	optical transfer function
p	pixel pitch
PCR	principle components regression
PDA	photodiode array
PDS	piecewise direct standardization
PLS	partial least squares
PRNU	photoresponse nonuniformity
PSF	point spread function
QE	quantum efficiency
R	responsivity
RF	radio frequency
SERS	surface enhanced Raman spectroscopy
SFD	source follower per detector
SNR	signal-to-noise ratio
SPRNU	spectral photoresponse nonuniformity
SWIR	short-wave infrared

TDI	time delay and integration
T_{int}	integration / exposure time
UKAS	United Kingdom Accreditation Service
UV	ultraviolet
VCTE	vertical charge transfer efficiency of a CCD
VCTI	vertical charge transfer inefficiency of a CCD
V_{OFFSET}	offset voltage
V_{SAT}	saturation output voltage
V_{SS}	substrate bias voltage
V_{Video}	video voltage

APPENDIX B

GLOSSARY OF TERMS

Absorbance
Measure of the ability to absorb radiation: negative log (base 10) of transmittance [−log 1/ T]. It is the product of absorption coefficient and path length, or, in chemical analysis, the product of extinction coefficient, path length and concentration.

Active pixel sensor
An imaging detector array, usually fabricated using CMOS technology, having photogate or photodiode detector elements directly addressed by MOS transistors. Each pixel also has a gain transistor, hence the name.

Active thickness
The thickness of the silicon in which electrons are photo-generated and collected as signal. In back-illuminated devices this is the actual thickness of the device silicon; but in front-illuminated devices it is generally the thickness of a lightly doped epitaxial layer deposited on much thicker material (or substrate) that is highly doped (p+) to cause almost immediate recombination of any charge photo-generated within it.

Advanced inverted-mode operation (AIMO) device
An improved inverted-mode device structure developed by EEV to achieve peak signal levels higher than available with the basic IMO device. *See also* **Inverted-mode operation device** and **Multi-phase pinned**.

Amplifier (on-chip)
A single- or multiple-transistor structure fabricated with and forming an integral part of the array detector. It is designed to convert the accumulated charge into a voltage signal usable by external circuitry. *See also* **Output node**.

Anti-blooming drain (of a CCD)
The drain structure used to remove excess overload-generated charge from the pixels of an image section, usually located in the column isolation region. *See also* **Blooming, Fixed anti-blooming, Gated anti-blooming**.

Aperture response (of a pixel)
If a small (~1-μm diameter) spot of light is scanned across a pixel, the resulting plot of signal versus distance is known as the aperture response. The scans are usually arranged to be either in the row or the

column direction and may or may not be deconvolved with the spot profile. The aperture response will vary with the wavelength of the spot illumination.

Background reading
The signal recorded when the irradiating source is switched on, but radiation is prevented from reaching the input to the system by use of a shutter or similar obstruction

Back-illuminated (back-thinned)
A CCD fabricated on one surface of silicon material that is subsequently processed for illumination from the reverse side, thereby avoiding transmission loss in the electrode layer (particularly significant at short wavelengths or with low-energy x-rays). This requires the silicon to be reduced to a thin layer, which is usually achieved with chemical etching, together with surface passivation and an optional anti-reflection coating.

Bandwidth (of a spectrometer)
Also termed *instrument bandwidth* or *spectral width*, bandwidth is the width (usually the full width at half height, or FWHM) of the line spread function. This function is the convolution of the slit function and the pixel aperture response. Bandwidth is measured in wavelength units (usually nm).

Beer's Law
Relationship between the amount of light absorbed by an analyte (absorbance, A) and its concentration (c), pathlength (b), and absorption coefficient (a), written as $A = a\ b\ c$.

Bias
The electrical voltage supplies necessary for the operation of a semiconductor device.

Binning
The addition of signal charge from more than one pixel into a given location. Binning may be performed in either horizontal (row) or vertical (column) directions by performing the summation onto the output node, into a summing well or into the read-out register.

Binning capacity (of a CCD)
The storage capacity of the read-out register and/or summing well relative to that of the pixel.

Blackbody
Ideal thermal detector that absorbs completely all incident radiation, whatever the wavelength, direction of incidence or polarization. It emits radiation in accordance with Planck's radiation law.

Bloomed
An optical surface provided with an anti-reflection coating.

Blooming
When the illumination level is sufficient to generate more charge than can be stored in a pixel, the spread of excess charge to adjacent pixels is known as blooming.

Channel
A region of semiconductor defined by implantation for charge transport or conduction. In an image sensor

this might be immediately below the surface of the electrode structure, hence, surface channel; or deeper within the material, i.e., buried channel.

Charge-coupled device (CCD)
A semiconductor device in which generated electronic charge is accumulated and transferred by the application of electrical potentials to insulated electrodes or gates. In an imaging CCD the charge is generated by received photons, either optical or x-ray.

Charge dumping
The operation of removing unwanted charge from a CCD, performed either using a specific dump-gate and drain structure or using gated anti-blooming.

Charge packet
An isolated quantity of charge (electrons), typically consisting of the integrated photo-generated charge from a single pixel of the image, which is stored and transported as signal in a CCD.
.

Charge Transfer Efficiency (CTE)
A measure of the ability of the CCD to transfer correctly a charge packet from the point of generation to the device output. It is defined as the fraction of the charge initially stored in a CCD element that is transferred to an adjacent element by a single clock cycle. The charge not transferred remains in the original element, possibly in trapping states, and may subsequently transfer to the next element (and beyond) with later clock cycles. The value for CTE is not constant but varies with signal size, temperature and clock frequency. Note that in some definitions the value is per electrode-to-electrode transfer, rather than element-to-element.

Charge transfer inefficiency (CTI)
Mathematically, (1 − CTE).

Chromaticity
The location of a color within the chromaticity diagram. This diagram represents the unit plane in a tristimulus space (i.e., ignoring brightness).

Clamp-and-sample
A practical circuit technique for implementing correlated double sampling.

Clamping
The action of setting a signal voltage to a predetermined reference level prior to the subsequent excursion of useful signal as a voltage difference.

Clock cycle
The basic increment in the set of clock pulses necessary to transfer charge from one element to the next.

Clock electrode
The part of the basic CCD structure, usually fabricated from polysilicon that, by means of the clock pulses applied to it and its neighbors, effectively controls the storage and movement of signal charge in the silicon beneath it.

Clock phase
A term used to identify (1) the different clock pulses, in the set of similar but time-shifted pulses, used to operate a section of a device; (2) the terminal of the device to which these pulses are applied, and (3) the electrodes connected to this terminal. A clock phase is usually designated by the Greek symbol ϕ, followed by a number that generally indicates sequence (both in time for the pulses and spatially for the electrodes), and preceded by a letter to indicate the device section to which it applies.

Clock pick-up
Attenuated clock pulse waveforms present in the output of the device and therefore superimposed on any signal, but not considered to be noise because of the constant repetitive nature. It is a result of coupling through parasitic capacitance in the CCD or its external connections.

Clock pulse
The pulsed voltage waveform (of which several are generally required) applied to an array detector for operation.

Clocking
A term used to describe operation of an array detector by applying clock pulses to achieve collection, transfer and detection of signal charge.

Color space
A system for ordering colors that represents the similarity relationships between them. There are a variety of different color spaces, but all are three dimensional.

Color temperature
A term used to describe the color appearance of a light source. It is the temperature of a black body or Planckian radiator having the same (or nearly the same) chromaticity as the source.

Column
A line of CCD elements in the direction perpendicular to the read-out register and generally transferring charge to one particular element in the read-out register.

Column isolation
The region between two adjacent CCD columns necessary to separate the signal charge in the pixels of each.

Continuum source
A source of optical radiation that emits at all wavelengths within a given spectral region and shows no discontinuities in output.

Contrast transfer function (CTF)
Similar to the modulation transfer function (MTF) but used to describe performance with a square-wave input.

Correlated double sampling (CDS)
A technique for reducing the noise associated with the charge detection process by subtracting a first output sample taken just after reset from a second taken with charge present.

Cuvette
Transparent receptacle in which sample solutions are introduced into the light path of spectrometers. Usually, two sides are equal, i.e., 1 × 1 cm, while the third dimension is elongated, possibly as long as 15 cm. For UV work, the material is quartz. Visible work permits the use of glass or plastic cuvettes.

Dark reference pixels
Pixels of a CCD active area that are made insensitive to illumination (e.g. with an aluminum shield) specifically for the purpose of dark signal compensation. *See also* **Over-scanning (virtual pixels)**.

Dark signal
The output signal of an image sensor with zero illumination. This is typically thermally generated electrons within the semiconductor material that are accumulated in each element of the device and transferred to the output during read-out.

Dark signal nonuniformity (DSNU)
The spatial variation of the dark signal within an image sensor.
Deep-depletion device
A device with the active thickness increased from normal and with a corresponding increase in the depth of depletion (i.e., to minimize the depth of the remaining field-free region).

Defects (cosmetic)
Pixels of an image sensor that show performance not representative of the average for the device. Defects present with no illumination (white spots or columns) are pixels with greater than a defined limit above the average thermal charge generation; while dark defects (black spots or columns) in an illuminated image are pixels with photo-response less than a defined limit below the local average for the device. Test specifications and data sheets define particular limits or thresholds for the discrimination and counting of defects.

Depletion region
The depth of silicon below the surface that is normally devoid of carriers and in which there is an electric field normal to the surface to direct photo-generated charge to the nearest pixel.

Dispersion (of a spectrometer)
The scaling factor between measured wavelength (usually in nanometers) and focal plane distance (usually in millimeters or pixels).

Dither clocking
A method of dark signal reduction where the electrode used for charge storage in a CCD element is periodically switched between the different phases.
Drain
An n-type region in the array device structure to which a positive bias is applied to extract electrons.

Dump gate/drain
A device structure, normally alongside the read-out register, for the removal of unwanted signal charge. Application to the dump gate of an electrical potential above a threshold enables the fast transfer of charge to a drain.

Edge filter
A filter that absorbs all radiation below a specified wavelength and has close to 100% transmittance above this wavelength. Sometimes also called a cut-on filter.

Electronic shutter
Some array detectors have the facility to drain charge out of the photo-site for part of the integration time, thus giving a variable exposure time. Alternative names are “charge reset” and “gated antiblooming.”

Field-free region
The depth of active silicon below the depletion region of an array detector in which electrons (as minority carriers) can diffuse laterally away from the point of photo-generation before being attracted by the depletion region of a pixel and then collected as signal.

Fill factor
The ratio of the light sensitive area to the total pixel area.

Fixed anti-blooming
An anti-blooming drain designed to remove any excess charge above a pre-set maximum signal level. *See also* **Gated anti-blooming**.

Fixed pattern noise
Pixel-to-pixel variations in the dark signal of an array detector due to electronic pick-up or crosstalk effects.

Frame transfer CCD
A CCD designed with the active area separately clocked in two sections: one for image acquisition and one for temporary storage and subsequent transfer to a readout register. In some CCDs a shield (*see* **Store shield**) is incorporated to make the storage region insensitive to illumination. Integration of a subsequent frame may take place simultaneously with readout.

Frame transfer time
The time period in which a complete CCD image can be clocked from the imaging region to the storage area.

Front face illuminated
The conventional mode of imaging CCD operation where the incident radiation is transmitted to the charge generating silicon via polysilicon clock electrodes. *See also* **Back-illuminated**.

Full frame CCD
A mode of CCD operation where all active elements are used for imaging.

Full well capacity
An alternative term for storage capacity.

Gate
A general term for the control electrode in any MOS structure, e.g., transistor, CCD, etc.

Gate protection
Voltage-limiting components integrated on chip in the gate connections of the device to reduce its susceptibility to damage from electrostatic discharge (ESD).

Gated anti-blooming
An anti-blooming drain designed with a gate electrode to control (via the applied voltage) the maximum signal level at which anti-blooming occurs; it can also be used for charge dumping purposes. *See also* **Fixed anti-blooming** and **Dump gate/drain**.

Grating
A reflective surface covered with evenly spaced microscopic grooves, the purpose of which is to separate the individual wavelengths from "white" light. The distance between grooves and the angle of the faces are determined by the wavelengths to be separated.

Guard ring/drain
A drain structure sometimes used in the peripheral regions of a CCD or diode array to collect photo-generated charge that would otherwise diffuse into the active regions of the device as spurious signal.

Image smear (frame shift smear)
The generation of charge in other than the correct pixels during the time that charge signals are being transferred through an illuminated section of the device, e.g. for readout. The effect can be minimized by reducing the ratio of transfer to integration times and eliminated by using a shutter to block the illumination during transfer.

Integration
The accumulation of photo-generated charge within the pixels of an image sensor.

Integration time
The time or proportion of the operating cycle of the array detector in which charge is accumulated.

Interlaced Scan
A method of reading out an imaging sensor so that even rows of the image are displayed in one TV field and the odd rows in the next TV field.

Inverted
A mode of operation where the electrode voltage of a CCD is held sufficiently negative of the substrate bias so that holes are attracted to the silicon surface, thereby locally inverting the n-type buried channel to p-type and also suppressing the surface component of dark signal.

Inverted-mode operation (IMO) device
A conventional CCD structure fabricated with additional implants to allow integration with all clock phases at zero and the whole surface inverted, thereby achieving very low levels of dark signal. **See also Inverted, Pinned, Multi phase Pinned**.

Kell Factor
The ratio of the perceived vertical resolution (in resolved lines) on a TV display to the number of raster lines. This factor takes into account the fact that a test target will not be precisely aligned with the raster

pattern and assumes a random phase relation. The Kell factor is usually taken as 0.7.

Lag
Signal readout at the end of an integration time, which is influenced by optical absorption at earlier times (previous integrations).

Lenticular array
See **Microlens array**.

Limit of detection
Lowest signal that can be seen above noise level of an instrument.

Linearity
A device is linear if its output varies in direct proportion to the input optical radiation.

Linearity error
The departure from linear behavior, usually measured as a percentage either of the signal or the saturation signal.

Line read-out section
An alternative name for read-out register.

Microlens array
A method to increase the effective pixel area (fill factor) is to manufacture a small lens (microlens or lenslet) on each pixel, thus forming a microlens array (also known as a lenticular array).

Modulation transfer function (MTF)
A term traditionally used to define the resolution performance of any optical component. It is the ratio of the depth of output signal modulation to that present in a sinusoidal bar-pattern input of a given spatial frequency.

Multi-phase pinned (MPP)
Alternative description of IMO CCDs, used particularly in the U.S.

Notch filter
A filter with zero transmittance at one specified wavelength and high transmittance at other wavelengths.

Nyquist frequency
The highest frequency (either spatial or temporal) that can be reliably sampled without aliasing effects. The sampling theorem states that the Nyquist frequency = 1/(2 × sampling interval).

Open electrode structure
In the normal operation of a front-face-illuminated CCD the incident light is transmitted to the silicon via the electrode structure. To minimize the inherent absorption of these layers, particularly at short wavelengths, CCDs may be fabricated with patterned electrodes where a part of each pixel is left unobscured and therefore “open” to direct illumination.

Output drain(s)
Drain connection of the amplifier transistors of a CCD.

Output gate(s)
Final electrodes in the read-out register, generally held at fixed bias to screen the output node from the pulsed electrodes, and thereby minimize clock pick-up.

Output node
The on-chip circuit point of a CCD or diode array whose capacitance to substrate is used to convert the accumulated and transferred charge to a voltage that is subsequently buffered by an amplifier to give an output signal. The node may be formed by the fabrication of a diode structure within the silicon.

Output register
An alternative name for readout register.
Output responsivity (charge to voltage conversion factor)
A measure of the transfer characteristic of the CCD output node and on-chip amplifier of a CCD. This is normally expressed in units of μV per electron of detected charge.

Overscanning (virtual pixels)
The action of clocking a read-out register by more cycles than the number of register elements. This is sometimes performed to provide blank or zero charge read-outs to use for output offset compensation in the electronic processing chain.

Parallel transfer
Transfer of the rows of charges toward the read-out register.

Peak signal
The maximum signal that can be stored, transferred, and read out with specified performance. The value quoted may be determined by either saturation or by some parameter reaching a specified limit, e.g., nonlinearity.

Phosphor (scintillator)
A chemical compound with the ability to photoluminesce when irradiated by optical or x-ray photons. The emitted wavelength is usually visible light, and, if appropriately selected, such compounds may be used to effectively wavelength shift incident radiation to a region of enhanced CCD response.

Photo-response nonuniformity (PRNU)
The spatial variation of the photon-induced signal generating process within an image sensor.

Pinned
Another term for “inverted” as the accumulated holes “pin” the surface potential of a CCD, photodiode or implanted gates to that of the substrate regardless of the applied voltage to the CCD electrode(s) or gates.

Pixel
Short for “picture element”: an element in the detector array in which photo-generated charge is collected as signal.

Potential well
A term used to describe the charge storage region (of an element) derived from the shape of the potential distribution in the underlying silicon.

Progressive scan
An array sensor that transfers out the entire image during one exposure cycle is called a progressive scan (or noninterlaced) imager.

Quantum efficiency (QE)
A measure of the sensitivity of an image sensor to input illumination. It is defined as the proportion of the incident photons that generate signal charge, and is normally expressed as a percentage. It is wavelength dependent. *See also* **Responsivity**.

Readout rate
The clock frequency of the read-out register, divided by the horizontal binning factor if appropriate.

Readout register
A CCD analogue shift register fabricated adjacent to the column structure of the device, used for serial read-out of each row of charge signals. Also sometimes referred to as the serial register.

Resolution (of a spectrometer)
Also termed “pixel pitch” or “spectral sampling interval,” resolution is the spacing (in wavelength units) between samples of the spectrum. Note that the effective spectral pixel width will be less than the pixel pitch if there are dead spaces between pixels, and will be larger than the pixel pitch when there is crosstalk between pixels. Although the term is common in spectroradiometry, use of “spectral sampling distance” is to be preferred.

Responsivity
An absolute measure of the sensitivity of an image sensor to input illumination. This is normally a function of the wavelength of the incident radiation and is typically expressed in units of mA/W at a given wavelength

Row
A line of CCD pixels in the direction parallel to the readout register.

Run-in or Run-out pixels
Some CCDs are provided with a greater number of register elements than pixels in the imaging row to provide stability of readout clocking and improved integrity of the signal pixels or for system offset compensation (e.g. line-by-line clamping).

Saturation
The absolute maximum signal level possible in the device, usually determined by factors such as the onset of clipping, (e.g., with antiblooming), charge spreading (e.g., in non-antibloomed devices), or gross nonuniformity.

Scanning
The process where the wavelength range of the system is viewed in order, usually from lowest to highest

wavelength. This usually occurs when the grating is rotated about its axis.

Scientific sensor
A high-performance CCD or other detector array that is not usually compatible with standard TV formats and is typically operated at reduced temperatures to achieve a wide dynamic range.

Sensitivity
See **Limit of Detection**.

Serial register
Another name for readout register

Serial transfer
Transfer along the readout register to the output.

Shift register
A temporary storage or delay circuit in which the output is released a fixed number of clock pulses after the input

Signal
The output of the detector due to its response to light emerging from the sample holder or reference cell.

Signal-to-noise Ratio
The numerical ratio of the total signal at 100% T to the noise of the instrument.

Slit function
The relative spectral transmittance of a polychromator (monochromator) for a given spectral adjustment and a given setting of its slit width(s)

Slit width
Size of opening of slit through which light enters a spectrometer. Often fixed or automatically programmed.

Spectral range (of a spectrometer)
The full wavelength range over which a spectrometer can measure optical radiation. For a fixed grating (or prism) this will be the number of pixels in a line multiplied by the spectral sampling interval, but will be greater than this if the grating is moveable.

Spectral response
The variation of responsivity with optical wavelength.

Spectroscopic sensor
A detector array sensor designed for use in applications where the input illumination is spectrally dispersed in one dimension.

Stitching
A photo-lithographic technique used to define large-area CCDs where a projected mask image of limited

area is stepped and repeated.

Storage capacity
A measure of the peak signal capability of the device, generally expressed as the number of electrons per pixel.

Store shield
An opaque layer (typically aluminium) fabricated on a frame transfer CCD to make the storage region insensitive to illumination. It may also be extended to cover the output amplifier and elements of the image section to give dark reference pixels (e.g., edge columns and/or rows at top and bottom of the image section).

Stray light
Any radiation reaching the detector by other than the desired optical path. It can be generated outside the instrument (in a spectrometer this is termed homochromic stray light) or inside the instrument (in a spectrometer this is termed heterochromic stray light)

Substrate
The underlying body of a semi-conductor material, usually held at fixed applied bias (Vss), also the term used to describe the highly doped base material on which an epitaxial layer is deposited.

Summing well (electrode)
A separately connected electrode at the end of the readout register before the output gate, which may be used to accumulate several charge packets prior to output and thus facilitate horizontal binning.

Time delay and integration (TDI)
A method of CCD operation where the image illumination is scanned across the CCD synchronously with the clocking of the active area electrodes. Each pixel in any column of the CCD therefore acquires a temporally integrated intensity from a point of the image. These may be binned into the register or separately read out.

Transition region/elements
One or more columns/rows that form the boundary between the dark reference pixels and the useable imaging pixels. The signal level in these elements is undefined and therefore should not be used.

Transmittance
Ratio of the radiant power transmitted by a sample to the radiant power transmitted by a blank in an equivalent cell or by some other means of compensation for solvent absorption, reflection losses, etc.

Trap (Trapping state)
Another name for the generation-recombination center in a semiconductor that is generally the cause of dark signal and charge transfer inefficiency in a device. The term trap is also used to describe a type of image defect resulting from a localized imperfection in the buried channel of a CCD. This operates by capturing a fixed quantity of charge from a charge packet (or the whole charge packet if less than this quantity) as it is clocked past the defect site, and then slowly releasing the charge to be collected by the following potential wells as they are clocked past the same site.

Tristimulus
Tristimulus colorimetry is a set of techniques for matching the color of a stimulus with the amounts of three specified primary colors.

APPENDIX C

QUICK GUIDE TO RECOMMENDED CALIBRATION PROCEDURES

C.1 Spectrometers

C.1.1 Wavelength Calibration

1. A number of calibration features (i.e., emission lines, or transmittance or reflectance peaks) spanning the entire wavelength range of interest should be used.
2. A cubic spline fitting algorithm generally gives an acceptable fit to the wavelength scale, but if possible several algorithms should be tested and the best selected on the basis of the results obtained.
3. Determination of the position of the peak wavelength by use of the FWHM or center- of-gravity methods should give acceptable results in most cases.
4. Where possible, the entrance slit width used for the wavelength calibration should be sufficient that each calibration feature covers several elements in the array. The bandwidth used during routine measurements should be matched to that used for the wavelength calibration or vice versa.
5. Where possible, the array output should be checked for nonuniformities. It may be necessary to exclude regions near “bad” pixels from the wavelength calibration.
6. Wavelength calibration should be performed using emission lines from low-pressure discharge lamps wherever possible. Where this is not feasible (as is often the case with spectrophotometers) glass filters or solutions providing relatively narrow absorption features should be used. For reflectance spectrophotometers, ceramic tiles or sintered materials, which have been doped with rare-earth metal oxides, may be used.
7. The emission lines from a low-pressure discharge lamp always occur at well-documented wavelengths; calibration of such lamps is therefore not necessary.
8. When using a low-pressure discharge lamp for wavelength calibration, it is essential that the spectrometer optics, including the polychromator, are overfilled and uniformly illuminated. Ideally the lamp is placed at the normal test lamp position; direct illumination of the slit should only be used if absolutely essential, and if this approach is taken, it is crucial to ensure that the polychromator optics are overfilled.

9. Filters, solutions or reflectance standards must be traceably calibrated and care taken to maintain cleanliness. The calibration bandwidth should be the same as that for the user's system.

C.1.2 Responsivity

1. Aside from an initial or diagnostic check on an instrument, spectral responsivity calibration is generally only necessary for spectroradiometer systems or other applications where it is necessary to measure the relative amounts of radiation present at each individual waveband across the spectrum or the absolute amount of radiation at one or more wavebands.
2. Suitable input optics are essential. For irradiance measurements or radiance measurements on nonuniform or polarized sources, a diffuser arrangement should be used. For radiance measurements on diffuse sources, direct imaging of the source onto the entrance slit is acceptable.
3. A group of suitably calibrated sources should be selected to perform the calibration; these standard sources should be stable, reproducible, rugged and user friendly, and the calibration should be traceable to national standards.
4. Wherever possible, the selected calibration sources should be similar in terms of size, shape and spectral distribution to the types of source to be measured. Where this is not possible or appropriate, specially designed and calibrated tungsten lamps (for the visible and near-infrared) or deuterium lamps (for the UV) are usually a good choice.
5. The short-term reproducibility (repeatability) of the system should be assessed by measuring the spectrum of a stable lamp (not necessarily calibrated) at the start, during and at the end of a batch of measurements to give an indication of system drift.
6. The spectral responsivity of the system should be calibrated on each occasion of use until a sufficiently long history of long-term reproducibility exists. After this, the system need only be calibrated often enough to stay within the required error limits.

C.1.3 Stray Light

1. A series of edge or notch filters should be used spanning the wavelength range of interest.
2. Calibration of the transmittance profile of edge filters is not necessary. Notch filters should ideally be calibrated, although for many purposes it is sufficient to assume zero transmittance at the notch wavelength (this may give a slight over-estimate of the level of stray light).
3. When using calibrated notch filters, the calibration bandwidth and the measurement bandwidth should be the same.
4. The filters should be placed in a convenient position between the source and the entrance slit. This may not always be easy, but is almost always possible.
5. The source used for stray light checks will affect the results obtained. For spectrophotometry the source will usually be automatically selected by the instrument software. For a spectroradiometer a tungsten or deuterium lamp is usually the best choice,

the final selection being determined by the wavelength range and the types of test source to be measured.

6. Stray light errors are almost impossible to correct. If an instrument shows significant levels during the above tests, the optimal approach is to restrict its use to those regions and/or sources for which acceptably low stray light is seen. The use of a stray light filter may be possible in some cases. If neither approach is feasible, an allowance for potential levels of stray light should be made in the uncertainty budget for the instrument.

C.1.4 Linearity

1. Linearity should be assessed for a range of integration times and at varying incident power levels in order to determine safe operating limits.
2. The linearity assessment should ideally cover the full range of wavelengths over which the system is to be used; linearity may vary with wavelength and performance may be poor at the limits of the responsivity range of the array elements.
3. Flux superposition methods are the most fundamental and accurate means by which to assess linearity, and are recommended for state-of-the-art applications. Indirect methods are more prone to error but are simpler and cheaper to perform. They are recommended for the majority of users of this guide.
4. The choice of method will depend on the details of the application, but a source-based approach or the use of neutral density filters will usually be recommended for a spectroradiometer, neutral density filters, or varying solute concentrations for a transmittance spectrophotometer and neutral reflectance standards (tiles or Spectralon plaques) for a reflectance spectrophotometer.
5. It is important that the reference artifact(s) are traceably calibrated under conditions similar to those in which they will be used on the spectrometer.

C.1.5 Dark Reading and External Stray Light

1. In a spectroradiometer the dark reading is taken by placing a completely opaque cover over the entrance to the polychromator, if this is accessible, or in front of the input optics if these are fixed to the polychromator. For a spectrophotometer, a piece of matte black card or similar nontransmitting or reflecting sample is placed in the normal sample position.
2. Usually only one dark reading is required by the software, and this is then automatically subtracted from all subsequent signal measurements. If only one dark reading is made occasional measurements should be made under the same conditions—the recorded signal should be zero and any nonzero signal indicates a shift in the dark signal. For most systems a new dark reading will be required if the signal integration time is changed.
3. It is also necessary to allow for any external stray light. Most spectrophotometers are designed so that the levels of external stray light are negligible. However, in spectroradiometers, there is significant potential for light from the source to reach the input by indirect routes. The amount of external stray radiation will vary depending on

the geometry of the source and the positions of any reflecting surfaces, therefore it needs to be regularly reassessed. There are three approaches that can be used:

- Placing a baffle close to the source, just large enough to prevent direct radiation form the source from entering the system. Screens can then be positioned so as to reduce the external stray light to negligible levels. This approach is recommended whenever possible.
- If the effective elimination of external stray light is not practical, a stray light reading can be made and subtracted from all subsequent (dark-reading-corrected) measurements.
- Measurement of the background reading instead of a dark reading. In practice, this means placing the opaque screen immediately in front of the source rather than the spectrometer.

C.2 Imagers

1. Flat fielding (for correction of response nonuniformity) should be performed with illumination conditions (wavelength and cone angle) as close as possible to those in normal use and at an illumination level bright enough to give good signal to noise ratio but not high enough to give saturation. Flat fielding at several illumination levels may be needed to detect traps.
2. Dark images should, if possible, have the same exposure time as the images to be corrected. A check should be made on dark charge stability (flickering pixels).
3. For linearity measurements the recommendations discussed in Sec. 8.1.4 for spectrometers are also applicable to imagers (see also Sec. 1.2.11).
4. Other measurements can be made using those methods from Sec. 1.2 that are most applicable. Note that a typical imager will have its optical performance influenced by the camera lens and any band-limiting filters.
5. Warm-up effects can influence both the camera electronics (e.g. gain) and the geometric stability. For accurate work this may require investigation by the user and the derivation of a suitable start-up procedure.
6. Checks should be made for spurious image artifacts (c.f. Sec. 3.1.18) and electronics effects (e.g., pick-up or bandwidth effects).

INDEX

About the Authors

Gordon Hopkinson has been a member of the technical staff of Sira Ltd. since 1983. During that time he has been involved in many projects using CCDs, active pixel sensors, and IR arrays for space applications and also for industry and defense. He first started working on solid state detector arrays while at the University of Durham, U.K., in the late 1970s, and used some of the first arrays manufactured in Europe (by Plessey Ltd.) for astronomical spectroscopy. Afterwards he moved to Leicester University to help establish the use CCDs for X-ray astronomy.

In recent years, as well as working on custom CCD developments, instrument design, testing and sensor procurement, he has been especially interested in the effects of radiation damage on sensor systems and has been an active participant in the IEEE Nuclear and Space Radiation Effects Conference and the Radiation Effects on Components and Systems (RADECS) conference—as a contributor, reviewer, session chair, awards chair and short course instructor. He has authored more than 30 papers on CCDs and similar detectors.

Teresa Goodman has worked in the Optical Radiation Measurement group at NPL (the U.K.'s national measurement standards laboratory) for about 23 years, with a range of responsibilities that have covered research, development, maintenance, and dissemination of measurement scales and standards for photometry, radiometry, spectroradiometry, spectrophotometry, color, and appearance. Her major research projects have included the first U.K. realization of the candela according to its radiometric definition, and the establishment of the U.K. scale of spectral total flux, together with the development of a calibration facility and transfer standards for its dissemination. She is currently leading work related to the development of scales for mesopic photometry, dosimetry for UV and blue light phototherapy, and human factors. Other responsibilities include "knowledge transfer" activities such as editing the NPL Optical Radiation Measurement Newsletter, *ORM News*, chairing the NPL ORM Club, laboratory assessments, and preparation of papers, reports, and best practice guides.

Teresa is also heavily involved in international and national standardizing activities and is director of Division 2 of the International Commission on Illumination (CIE). This Division has responsibility for producing recommendations and best practice guidance for all aspects of the measurement of light, color, and optical radiation.

Stuart Prince joined NPL in 1994 and works in the Optical Radiation Measurement Group. He has worked on a number of projects, including evaluation of photometric testing facilities, development of facilities for spectroradiometric measurements on deuterium lamps, research into new source technologies for photometry and spectroradiometry, and establishment of a portable (array-based) spectroradiometer facility. Stuart also took part in the comparison of the UV spectral scale with PTB at BESSY, Berlin. He has investigated methods to increase take up of traceability in the optical radiation metrology area and provided support for the formulation of the 2004-2007 Optical Radiation Metrology Programme. Most recently, he has contributed to a best practice guide for traceable measurements on pulsed sources and has set up a facility to perform spectral measurements on pulsed sources, using a CCD array spectroradiometer.

Before joining NPL, he worked in the scientific sector for 11 years (THORN EMI Electronics and BP Research), and during this time obtained a B.Sc. in Applied Chemistry.