A Users' Guide to IR Detectors

Glenn D. Boreman

Professor of Optics & Electrical Engineering

School of Optics/CREOL
University of Central Florida
Orlando, FL 32816-2700
USA

BOREMAN@CREOL.UCF.EDU

Abstract

This paper will guide the first-time user toward proper selection and use of infrared detectors for applications in industrial inspection, process control, and laser measurements.

1. Why Use Infrared?

There are a variety of sources of infrared radiation, including lasers, light-emitting diodes, and solid bodies that are heated to incandescence. For temperatures of engineering interest, say from somewhat below room temperature to a few thousand degrees Kelvin, we find that the majority of the radiation emitted by a body will be in the infrared, that is, at longer wavelengths than what can be seen by the human eye. The infrared region thus begins at a wavelength of 0.77 micrometers and extends toward the millimeter-wave band. For most applications, we typically speak of infrared radiation as being between 1 and 10 micrometers in wavelength.

2. Thermal Detectors and Photon Detectors

Detectors are transducers: they produce a measurable electrical output in response to radiation intercepted by the sensing element. The three overall quantities of concern for all detectors are responsivity, response speed, and sensitivity. Responsivity is a measure of output per unit input. Because there are various input and output quantities, there are a variety of ways to specify responsivity: for example, in volts per watt or in amps per photon per second. Responsivity allows prediction of the magnitude of the sensor's response, given a radiometric calculation of flux on the sensor. Thus, responsivity determines the voltage levels involved at the interface between the detector and the preamplifier that follows. The response speed of a detector is pertinent because any signal of interest will vary with time. How fast can the signal flux vary and still have the detector follow the variation? The Fourier transform of the time-domain impulse response is the transfer function, the relative response of the sensor as a function of temporal frequency. Sensitivity is a separate quantity from responsivity. While responsivity is a measure of the output level for a given level of input flux, sensitivity specifies the signal-to-noise ratio (SNR) that the user can expect for a given input flux level. The SNR is a crucial parameter in the determination of image detectability, that is, whether a given feature in the image can be reliably discerned above the noise.

Laser Metrology for Precision Measurement and Inspection in Industry, Armando Albertazzi, Jr., Editor, Proceedings of SPIE Vol. 4420 (2001) © 2001 SPIE · 0277-786X/01/$15.00

The two primary classes of detectors are thermal detectors and photon detectors. Both kinds of detectors respond to absorbed photons, but their mechanism of response differs, leading to differences in response speed and responsivity as a function of wavelength. Thermal detectors absorb the energy of the photon as heat. This heat causes a temperature rise in the sensing element. The sensing element has some temperature- dependent electrical property, such as resistance. The change in this electrical property, as a function of input flux level, is measured by an external circuit. Photon detectors use the energy of the photon not as heat, but to increase the energy of a charge carrier, so that the carrier makes an electronic transition across a forbidden energy gap. This is typically a transition of an electron from the valence band to the conduction band in a semiconductor material. The excitation of these carriers into a higher energy state affects the sensor's electrical properties. The change of electrical properties as a function of input flux level is also measured by external circuitry.

3. Response Speed and Bandwidth

We first compare the properties of thermal and photon sensors in terms of response speed. Thermal detectors are typically slow because a finite time is required for the sensing element to rise in temperature after the absorption of energy. Typical time constants for thermal detectors are in the range of tenths of seconds to milliseconds. A trade-off exists between response speed and response magnitude for thermal detectors. Because of their relatively long thermal time constants, waiting a longer time generally produces a larger response from a thermal detector. Photon detectors are fast, because an electronic transition is virtually instantaneous upon photon absorption. Typical time constants for photon detectors are in the microsecond range and shorter. Response speed for photon detectors is often determined by the resistance-capacitance (*RC*) product of the readout circuit that interfaces to the photon detector.

A related quantity is the frequency bandwidth Δf for the sensor, usually taken to be

$$\Delta f = 1/2\tau \quad . \tag{1}$$

where τ is the response time of the sensor. A fast time response implies a broad bandwidth. Conversely, it requires more measurement time to make a measurement with a narrow bandwidth than one with a wide bandwidth.

4. Spectral Responsivity

We can also compare the two main classes of sensors in terms of spectral responsivity $\mathcal{R}(\lambda)$. A thermal detector has a flat response, when plotted in energy-based units such as volts per watt. As seen in Fig. 1, the ideal thermal sensor has a responsivity that is independent of wavelength because a watt of radiation at $\lambda = 1$ μm will cause the same temperature rise in the sensing element as a watt at $\lambda = 10$ μm. There is no long-wavelength cutoff behavior in thermal sensors because no energy gap is inherent in the mechanism. Practically, the wavelength range for a thermal sensor is limited by the spectral absorption of the sensor material and the spectral transmission range of the window material that typically covers the sensor.

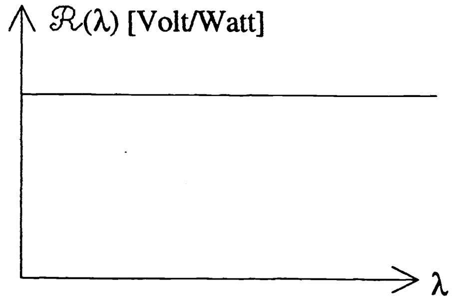

Fig. 1. Spectral responsivity for a thermal detector.

A photon detector has an inherent nonuniformity of response as a function of wavelength. For a photon to be absorbed by the sensor and impart its energy to an electron, its energy must be sufficient to lift the electron across the energy gap. A photon has an energy $\mathcal{E}$ [Joules/photon] is inverse with wavelength:

$$\mathcal{E} = hc/\lambda \ . \quad (2)$$

As seen in Fig. 2, given a photon sensor with an energy gap $\mathcal{E}_{gap}$, photons with wavelength longer than the long-wavelength cutoff $\lambda_{cut} = hc/\mathcal{E}_{gap}$ are not absorbed and not detected. The wavelength where the photon has just enough energy to bridge $\mathcal{E}_{gap}$ corresponds to the long-wavelength cutoff $\lambda_{cut} = hc/\mathcal{E}_{gap}$, and is the longest wavelength that will be detected by a sensor of a given material. We can solve for the long-wavelength cutoff of the photon detector in terms of $\mathcal{E}_{gap}$:

$$\mathcal{E}_{gap}[\mathrm{eV}] = \frac{hc}{\lambda_c} = \frac{6.6\times10^{-34}\,\mathrm{J\,s}\;\;3\times10^{8}\;\;\mathrm{m/s}}{\lambda_c\;\mu\mathrm{m}\times10^{-6}\;\mathrm{m}/\mu\mathrm{m}}\times\frac{1\;\mathrm{eV}}{1.6\times10^{-19}\,\mathrm{J}} = \frac{1.24}{\lambda_c\;[\mu\mathrm{m}]} \ . \quad (3)$$

If Si, with $\mathcal{E}_{gap}$ = 1.12 eV, is used as a sensor, photons with wavelengths shorter than 1.1 μm will be detected. If InSb, which has a smaller $\mathcal{E}_{gap}$ = 0.22 eV, is used, the sensor will detect photons with wavelengths shorter than 5.6 μm.

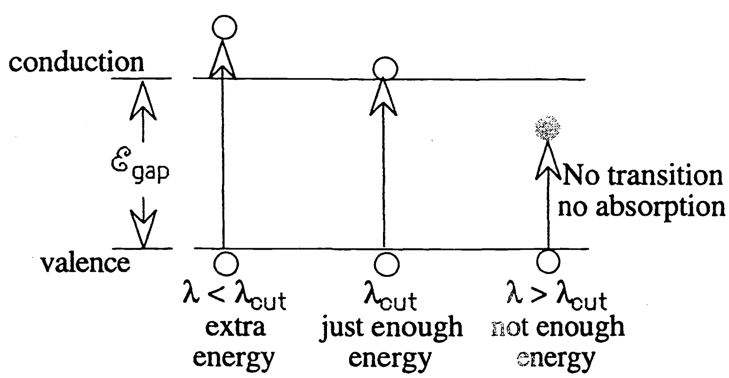

Fig. 2. Origin of cutoff wavelength for a photon detector.

An interesting issue arises when considering the spectral responsivity curves for photon detectors, as seen in Fig. 3. The responsivity will surely go to zero for wavelengths greater than λ_{cut}, but also the curves can be plotted with respect to either energy-derived or photon-derived units. Energy-derived units (such as Volt/Watt) are based on the number of Joules, while an analogous set of units can be derived on the basis of a number of photons/sec. Conversion between the two sets of units is easily accomplished using the relationship of Eq. (2) for the amount of energy carried per photon, but it should be realized that the conversion between the two sets of units, depends on λ. A longer-wavelength photon carries less energy than a short-wavelength photon. At an infrared wavelength of 10 μm, the photon energy is approximately 2×10^{-20} J/photon, while at a visible wavelength of 0.5 μm, the photons are a factor of 20 more energetic, having approximately 4×10^{-19} J/photon. This units conversion can also be thought of in terms of how many photons/s it takes to make 1 W of power. At 10 μm, there are 5×10^{19} photons/s in 1 W. For the more energetic photons at a shorter wavelength of 0.5 μm, fewer photons/s (2.5×10^{18}) are required to make 1 W. When wavelength-dependent responsivity is plotted, use of either sets of units will yield curves of different shape.

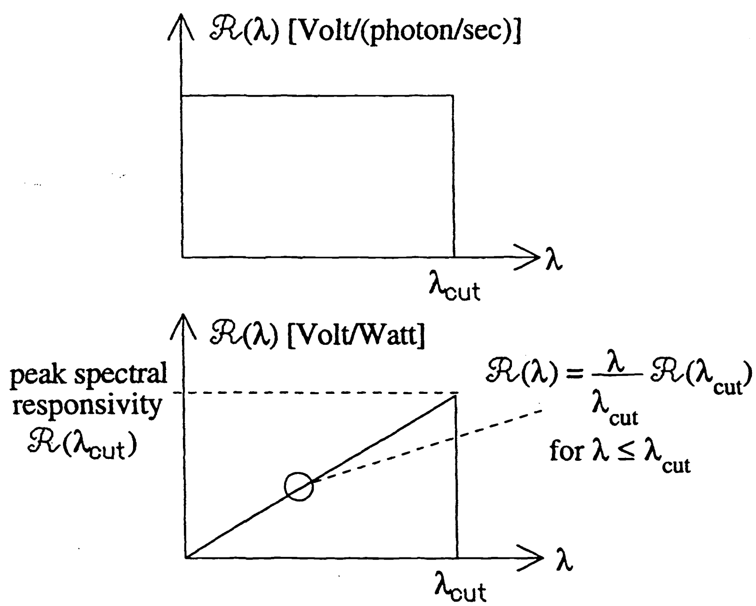

Fig. 3. Spectral responsivity for an ideal photon detector, plotted in energy and photon units.

An ideal photon detector has a flat $\mathcal{R}(\lambda)$ out to λ_{cut} when plotted in photon-derived units (e.g., Volt/(photon/sec). Each photon produces the same amount of response, as long as the photon has enough energy to bridge the gap. If $\mathcal{R}(\lambda)$ is plotted as a function of energy-derived units, a linear increase in responsivity is seen up to cut. Photon detectors are most naturally described in photon-derived units but for historical reasons are often plotted with respect to energy-derived units. Because the proportionality between these two sets of units, $\mathcal{E} = hc/\lambda$, depends on wavelength, these two sets of units will produce responsivity curves of different shape. There is no difference in the performance of the sensor, the different curve shape is just caused by the units. When described in energy-derived units, an ideal photon detector has a spectral responsivity of the form

$$\mathcal{R}(\lambda) = \frac{\lambda}{\lambda_c}\mathcal{R}(\lambda_c) \text{ when } \lambda \le \lambda_c; \;\; \mathcal{R}(\lambda) = 0, \text{ when } \lambda > \lambda_c \quad . \tag{4}$$

The significance of the spectral responsivity from a design viewpoint is that it is used to calculate the output from a detector in response to flux from a spectrally distributed source

$$\begin{gathered} \text{Output} = \int_0^{\infty} \phi_\lambda(\lambda)\, \mathcal{R}(\lambda)\, \mathrm{d}\lambda \\ [\mathrm{V}] = [\mathrm{W}/\mu\mathrm{m}][\mathrm{V}/\mathrm{W}][\mu\mathrm{m}] \end{gathered} \quad . \tag{5}$$

In the overlap integral of Eq. (5), there is a contribution to detector output only at those wavelengths where both flux and responsivity are nonzero. The spectral flux falling on the detector can be calculated from the Planck equation, if the source is a blackbody. For the important special case of a laser source, the spectral flux can be approximated by a delta function

$$\phi_\lambda(\lambda) = \phi_{\text{laser}}\,\delta(\lambda - \lambda_{\text{laser}}) \quad . \tag{6}$$

and the detector output is then given by

$$\text{Output} = \phi_{\text{laser}}\,\mathcal{R}(\lambda_{\text{laser}}) \quad . \tag{7}$$

5. Examples of Thermal Detectors

5.1. Bolometers

The bolometric sensor is a common type of thermal detector. A bolometer is a resistor with a temperature-dependent resistance. Photons are absorbed on the sensor surface, and the energy of the absorbed photons causes a temperature rise in the sensor. A change in resistance is sensed using a voltage-divider circuit, as seen in Fig. 4. A bias current is required to sense the change in resistance. The load resistor is often an element identical to the sensor but shielded from radiation. This allows the circuit to be insensitive to changes in ambient temperature.

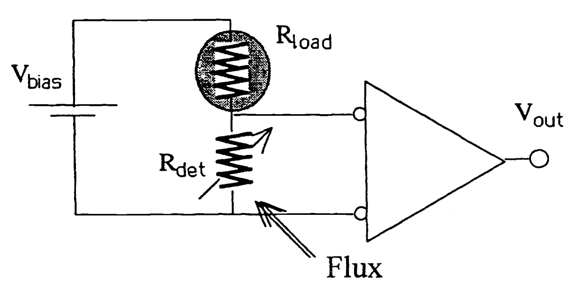

Figure 4. Bolometric readout circuit.

An important characteristic bolometers is α, the fractional change in resistance per degree of temperature change

$$\alpha \equiv \frac{1}{R}\frac{\partial R}{\partial T} \quad . \tag{8}$$

Typically α is positive, around 0.5%/°C for metals, because resistance of a metal rises with increased temperature. For semiconductors, the resistance decreases with increasing temperature, giving a negative value for α around – 5%/°C for semiconductors. Typical bolometer resistances are 10 k to 1 M Ohm. The primary advantage of bolometers is their wide spectral response, from visible to the long-wave IR. They are useful in applications where cryogenic cooling is not feasible (from a weight, power, or cost viewpoint) but where IR detection is required at wavelengths that would require photon detectors to be cooled. Time constants are typically in the range of 1 to 100 ms, with minimum detectable powers in the range of 10^{-8} to 10^{-10} W. Better sensitivities are obtained for the longer time constants. Bolometers have been recently demonstrated in large focal-plane-array configurations, allowing development of uncooled IR imaging sensors.

5.2. Pyroelectrics

Another useful thermal detector is the pyroelectric. These sensors are fabricated from materials (such as triglycerine sulfate) that have a permanent dipole moment even in the absence of an applied electric field. The magnitude of this dipole moment is temperature dependent. The sensing mechanism is based on the fact that when photons are absorbed, the temperature of the element is changed, and there is motion of bound charge on the surface of the material corresponding to the change in dipole moment. If this material is placed between the plates of a capacitor, as seen in Fig. 5, the motion of bound charge induces a current flow in the circuit connected to the plates in response to dT/dt. A pyroelectric detector responds only to a change in temperature, thus cooling is generally not required.

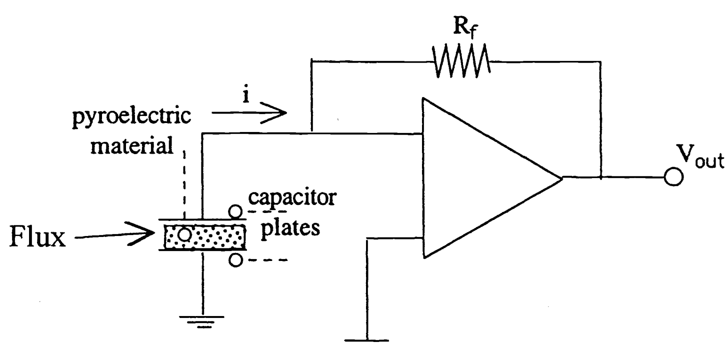

Fig. 5. Pyroelectric sensor configuration.

The value chosen for the feedback resistance R_f affects both the responsivity and the response time. A large R_f (e.g., 10^9 Ohms) gives slow response but large responsivity. Conversely, a small R_f (e.g., 10^3 Ohms) gives faster response but at the expense of a smaller responsivity. Better noise performance is achieved by pyroelectrics than by most other uncooled thermal detectors, with minimum detectable powers in the range of 10^{-9} to 10^{-10} W. The time response is quite fast for a thermal detector, with time constants of 100 ns or shorter possible, although with relatively small responsivity. For many laser-measurement applications, the disadvantage of small responsivity is outweighed by the fast time response and good sensitivity, along with the broad spectral response typical of a thermal detector. Pyroelectric sensors are widely used in laboratory instrumentation, and have also been fabricated as large uncooled focal plane arrays.

6. Examples of Photon Detectors

6.1. Photoconductives

The mechanism for a photoconductive (PC) detector is the generation of electron-hole pairs by photon absorption in a semiconductor. A PC detector is constructed from homogeneous material, without a p-n junction, so that the device simply changes conductivity in response to applied photons, and cannot generate an open-circuit voltage. A PC detector must have a bias voltage applied to render the change in conductivity measurable, in a similar configuration to the voltage-divider circuit seen in Fig. 4. The spectral response of a PC detector is determined by the energy gap of the semiconductor. Some of the more common materials used are shown below, along with their cutoff wavelengths. Photoconductors typically have minimum detectable powers in the range of 10^{-10} to 10^{-12} Watts, with time constants in the μs to ns range.

Table 1. Typical cutoff wavelengths for photon detectors.

Material	λ_{cut} [μm]
CdS	0.52
CdSe	0.69
GaAs	0.80
Si	1.1
GaAlAs	1.3
Ge	1.9
PbS	3.0
PbSe	4.5
InSb	5.6
HgCdTe	12 (depending on mixing ratios)

6.2. Photovoltaics

The mechanism of a photovoltaic (PV) detector is that an absorbed photon generates a hole-electron pair at a p-n junction in a semiconductor. The range of cutoff wavelengths is that seen in Table 1, consistent with Eq. (3) for any give material. As seen in Fig. 6, a built-in electric field exists in the vicinity of the junction. This causes an immediate separation of the hole and the electron once they are generated, allowing the photovoltaic detector to develop a voltage across an open circuit. While a PV detector can be operated in a mode where it is biased, a bias voltage is not required for the sensor to operate. This is the main operational difference between PC and PV detectors.

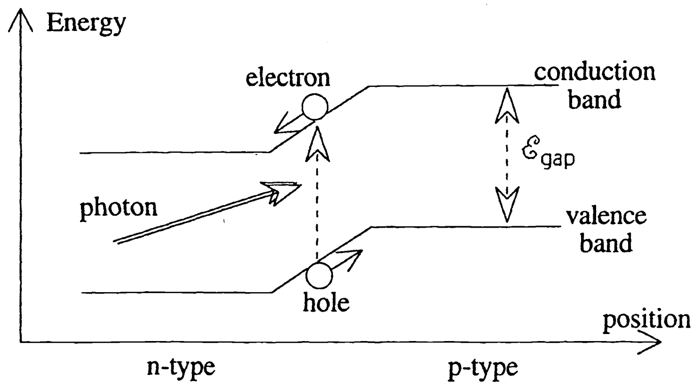

Figure 6. Photovoltaic detector.

The p-n junction in the photovoltaic detector is a diode, so the current-voltage (i-v) characteristic of the PV detector is that of a diode, as seen in Fig. 7. For the situation where the detector is in the dark, we have

$$i_{\text{dark}} = i_0\left(e^{qV/\beta kT} - 1\right) \tag{9}$$

where β is an empirical factor of the specific diode, q is the electronic charge, and k is Boltzmann's constant. With photons incident on the detector, the total diode current becomes

$$i = i_0\left(e^{qV/\beta kT} - 1\right) - \eta\phi q \tag{10}$$

where the negative sign comes from the direction of the photogenerated current compared to the convention for positive current flow in a diode. The factor η is the quantum efficiency (electrons per photon) and ϕ is the photon flux (photons/sec) incident on the detector.

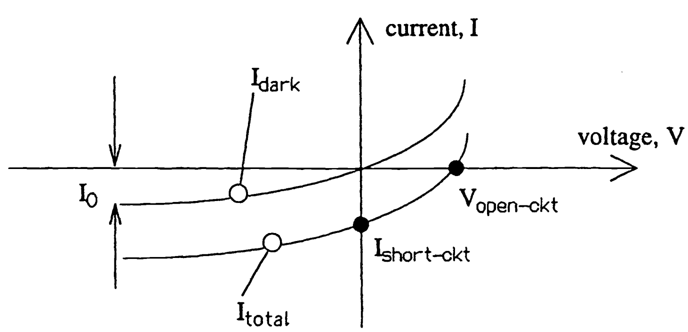

Fig. 7. Typical i-v plot for a photovoltaic detector.

There are three different operating modes that are commonly employed for the PV detector, the open-circuit mode, the short-circuit mode and the reverse-bias mode. Circuit interfacing to PV detectors is typically done using operational amplifiers. In the open-circuit mode of the PV (Fig. 8), the total current through the device is held at zero, and photon irradiation produces an open-circuit voltage. The open-circuit voltage is proportional to the natural log of the photon flux providing a large (but nonlinear) dynamic range:

$$v_{\text{open-circuit}} = \frac{\beta kT}{q} Ln\left[\frac{\eta\phi_q q}{i_0} + 1\right] \quad . \tag{11}$$

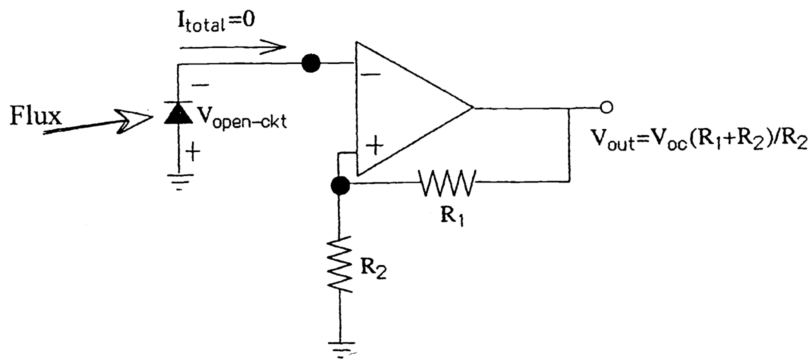

Fig. 8. Open-circuit mode for a PV detector.

In the short-circuit mode of operation, the current flow through the diode is measured when the two terminals of the diode are kept at the same voltage, as seen in Fig. 9. A short-circuit current is seen to be linearly proportional to the photon flux

$$i_{\text{short-circuit}} = -\eta\phi_q q \quad . \tag{12}$$

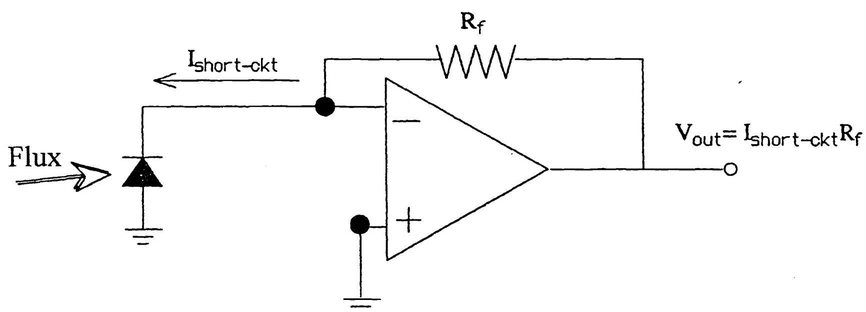

Fig. 9. Short-circuit operation of a photovoltaic detector.

Photovoltaic detectors can also be operated in a reverse-bias mode, as seen in Fig. 10. In this case the current is also linearly proportional to photon flux,

$$i_{\text{short-circuit}} = i_0 - \eta\phi_q q \quad . \tag{13}$$

The advantage of this configuration is that the reverse bias decreases the capacitance of the p-n junction by increasing the separation of the space-charge layer. A decreased capacitance lowers the RC time constant, leading to a faster response. The performance of photovoltaics is similar to that for photoconductives, with minimum detectable powers in the range of 10^{-10} to 10^{-12} W. Time constants are typically in the range of μs to ns, with reverse biased diodes providing faster response of the PVs at about 100 ps.

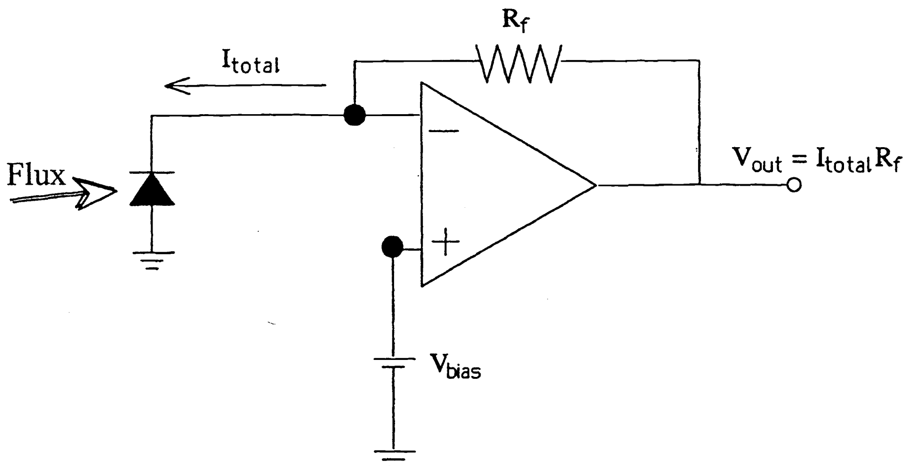

Fig. 10. Reverse-bias operation of a photovoltaic detector.

7. Sources of Noise

There are many sources of noise that must be considered in the design of an IR-sensing system. These noise sources include external interference effects, amplifier noise, and the noise arising in the sensor itself. In general, interference noise can be controlled by proper electrical and optical shielding. Amplifier noise is often times a limiting factor in a particular experimental setup, especially if a simple amplifier is used. However, by careful amplifier design it is generally possible to reduce this contribution to the point where it is small compared to the internally-generated noise in the sensor.

Internally-generated noises typically get worse with increased detector area, increased detector bandwidth, and with increasing temperature of the sensor. The sensor area and the data bandwidth are parameters that are determined by the job that the system has to perform. As seen in Fig. 11, knowing the sensor area and the focal length of the lens used will give the system field of view (the angular coverage) according to

$$FOV = \frac{\sqrt{A_{\mathrm{det}}}}{f} \quad . \tag{14}$$

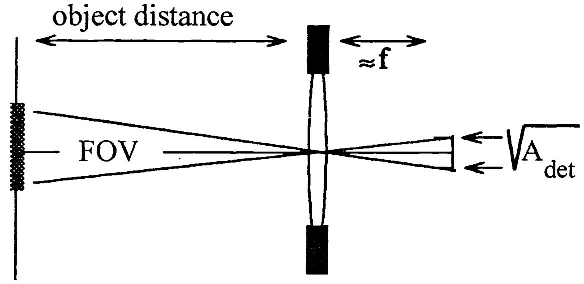

Fig. 11. Determining the FOV.

The area of the sensor should be no larger than that necessary to cover the region of object space you want to collect information about. The focal length of the lens can sometimes be decreased so as to give the same coverage using a smaller detector. Similarly, the temporal bandwidth of the measurement should be no larger than necessary to pass the information desired. Extra bandwidth (a faster-than-necessary time response) will just bring in more noise.

Also, the sensor is sometimes cooled to decrease the noise. This is typically required for photon detectors with cutoff wavelengths longer than 1 μm. The charge carriers have a thermal energy that, at room temperature, will allow them to cross over the energy gap of the semiconductor. Cooling to cryogenic temperatures (180 Kelvin for 3-5 μm detectors and 77 Kelvin for 8-12 μm detectors) reduces this so-called dark current noise.

Oftentimes, the sensor noise will get worse at low frequencies. For this reason, it is common to modulate the source to be detected at a frequency of a few kHz to avoid this high-noise region near dc.

8. Specification of Signal-to-Noise Performance

We want a means to specify the noise performance of a detector, so that we can predict the signal-to-noise ratio (SNR) which will be obtained when a given amount of power falls on the detector. The noise-equivalent power (NEP) is the amount of flux [W] that would produce an output equal to the standard deviation of the noise. Thus, NEP is the input flux required to produce SNR = 1. Noise-equivalent power is thus interpreted as the "minimum detectable power," although SNR = 1 is only a reference level, and the actual input flux required for acceptable sensor operation will depend on the SNR requirements for the specific application. The definition of NEP is

$$NEP = \frac{\phi_{det}}{SNR} , \tag{15}$$

which shows that NEP is the required flux on the detector for SNR =1. Sensitivity of the sensor is better for smaller NEP, in that the SNR produced is higher for a given sensor flux level.

We now consider a numerical-calculation example with NEP. The SNR is typically defined as a peak signal to root-mean-square (rms) noise, so that the NEP is proportional to the square root of the sensor bandwidth Δf. Suppose that a sensor has been measured to have an NEP of 7.5×10^{-7} W using a bandwidth of 500 kHz. Using this NEP we can predict the SNR when 1 μW of signal power falls on the detector as SNR = 1 μW/0.75 μW = 6.7. If the bandwidth can be changed from 500 kHz to 20 Hz, the NEP will change by a factor of the square root of the ratio of the bandwidths, yielding a new NEP of 4.7 nW. The system sensitivity was improved by reduction of the noise bandwidth.

The normalized detectivity, D*, is a figure of merit often used by manufacturers to specify the performance of a detector. D* is inversely proportional to the NEP, so that bigger D* corresponds to better sensitivity. D* is also proportional to the square root of the detector area and the square root of the measurement bandwidth, and has units of cm $Hz^{1/2}$ $Watt^{-1}$.

$$D^* = \frac{\sqrt{A}\sqrt{\Delta f}}{NEP} \tag{16}$$

Normalization with respect to detector area and measurement bandwidth cancels out the dependence of NEP on these quantities. We have seen in the example above that NEP is proportional to the rms noise and hence (for white noise) to the square root of the measurement bandwith. This dependence will cancel in the calculation of D*, yielding a figure of merit independent of the bandwidth used to make the measurement. Similarly, D* is independent of the area of the sensor used.

The normalization used in Eq. (16) yields a figure of merit independent of both the measurement bandwidth and the detector area. Thus, D* is most useful in the comparison of the merits of different detector materials and fabrication processes, without consideration of a particular bandwidth or area inherent in a given application. Manufacturers of detectors often specify the performance of their generic detector products using D*. Given the D*, an end user can calculate the SNR using

$$SNR = \frac{\phi_{\text{det}}}{NEP} = \frac{\phi_{\text{det}}\, D^*}{\sqrt{A_{\text{det}}}\sqrt{\Delta f}} \quad . \tag{17}$$

The D* is a consideration when choosing an appropriate detector technology for a particular application. However, the NEP remains the fundamental quantity for the calculation of SNR. Calculation of SNR requires a specification of the detector area and the bandwidth needed for a particular application. Although their effect on D* has been normalized out according to Eq. (5.33), both the detector area and the bandwidth affect the final SNR achieved by the sensor system.

We now consider a numerical-calculation examples using D*, remembering that when the detector area and bandwidth change, NEP changes but D* does not. Supppose a detector is specified to have a D* of 3.3×10^9 cm $Hz^{1/2}$ W^{-1}. If the sensor has dimensions 1 mm × 1 mm, and is used in a 1-MHz bandwidth its NEP will be 3×10^{-8} W. If this same detector material is fabricated into a 50 μm × 50 μm detector, and is used in a bandwidth of 500 kHz, we find the NEP for this situation (given that the D* is independent of area and bandwidth) to be 1×10^{-9} W.

References

Boreman, G, *Basic Electro-Optics for Electrical Engineers*, SPIE Press, 1998.
Dereniak, E., and G. Boreman, *Infrared Detectors and Systems*, Wiley, New York, 1996.
Keyes, R., *Optical and Infrared Detectors*, Springer-Verlag, New York, 1980.
Boyd, R., *Radiometry and the Detection of Optical Radiation*, Wiley, New York, 1983.
Wolfe, W., and G. Zissis, *The Infrared Handbook*, SPIE Press, 1978.