By Dean Scribner, Penny Warren, Jon Schuler, Michael Satyshur, and Melvin Kruer

INFRARED COLOR VISION:

AN APPROACH TO SENSOR FUSION

Sensor fusion is the science of combining information from multiple sensor responses. One approach, visualization, fuses various images and adds false color to define features that might otherwise be hidden and to aid the viewer in deciphering a scene.

The proliferation of powerful, inexpensive computers and new sensor technologies responsive to the non-visible regions of the electro-magnetic spectrum has made many new imaging applications possible. Research groups working in image processing are showing an increased interest in combining the information from various sensor combinations, such as IR and visible- or multiple-IR bands.

Although there are several approaches to sensor fusion, the one discussed here is color visualization, the combining of multiple bands of information into color images for a viewer. Color visualization is a new method of presenting IR and visible imagery. The supporting technology for color visualization is currently under development. The IR phenomenology involved with combining various bands is still under study and the algorithms used for color visualization are under development.

The multiband approach discussed here fuses only two or three bands of color. The challenge is to generate an intuitively

1047-6938/98/8/0027/06 $0015.00 © OSA

meaningful color image, for a specific viewer and application. A simple approach where each band of the system is assigned to a visible color is an interesting starting point, but there are many physical and system factors that complicate sensor fusion. It is expected that this multiband approach will greatly enhance object detection and scene comprehension. Besides obvious advantages for military applications, the multiband approach should prove useful for other applications such as medical imaging, remote sensing, and night driving.

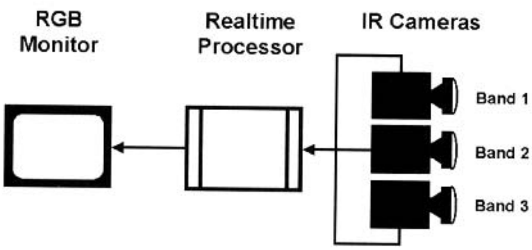

Figure 1. Simplified IR color vision system using three separate cameras and a real time processor that remaps the raw IR imagery into an RGB image for display on a conventional monitor.

IR color vision

One limitation of single-band IR imaging systems is that the contrast between objects and backgrounds is minimal under certain conditions. This difficulty arises because of fluctuations in target and background temperature differences. The radiant intensities of targets can vary from appearing to have strong positive contrast to strong negative contrast and, at certain times of the day, pass through a point of zero contrast. These variations can cause severe detection problems. By using two or three IR bands to create a multidimensional color space, the occurrence of zero contrast becomes less probable.

The goal of multiband IR visualization is to display all of the information from two or more bands in a single composite image, thereby overcoming the contrast limitations that exist with single-band imagery. By using two or more IR bands to create a color space, an associated composite color image can be mapped into the visible color bands for display (see Fig. 1). In this color space, the likelihood that an object's color is different from that of the background is very high.

Since a VGA monitor can display eight bits each of red, green, and blue (providing a color palate of approximately 16 million colors), a simplistic three-band color fusion method is to scale and normalize the pixel intensities for the three IR bands, and to map them to red, green, and blue values. While this gives reasonably good results in some cases, there are a number of special cases where additional considerations must be taken into account to achieve more meaningful and intuitive imagery. One approach is to mimic similar processes that exist in biological systems, which is especially appealing because of the large body of literature that exists on the topics of color science[1] and human color vision.[2]

Three processing functions of particular interest are color-and-shape discrimination,[3] color contrast enhancement, and color constancy.[4] Researchers in pattern recognition often use color-and-shape discrimination because color can help resolve contrast ambiguities that severely limit performance in applications such as automatic target recognition. We will not address automatic target recognition in this discussion because of its complexity, and instead will focus on color contrast enhancement and color constancy.

Supporting technologies

Recent sensor technology and physical phenomenology drive the color processing algorithms that are used for multiband color visualization, and advances are creating many new possibilities for multiband imaging systems. As shown in Figure 1, two or three cameras can be combined to form color imagery. Specific camera technologies that are now becoming available in production quantities at reasonable costs are: long-wavelength IR (LWIR) cameras using uncooled detector arrays; medium-wavelength IR (MWIR) cameras using InSb detector arrays; short-wavelength IR (SWIR) cameras using InGaAs detector arrays; and low-light-level cameras operating in the near-IR (NIR) and visible bands using silicon CCD arrays.[5]

One of the major problems encountered when attempting to assemble a multiband sensor using a camera configuration like that shown in Figure 2, is that single-band images must be registered to achieve a good

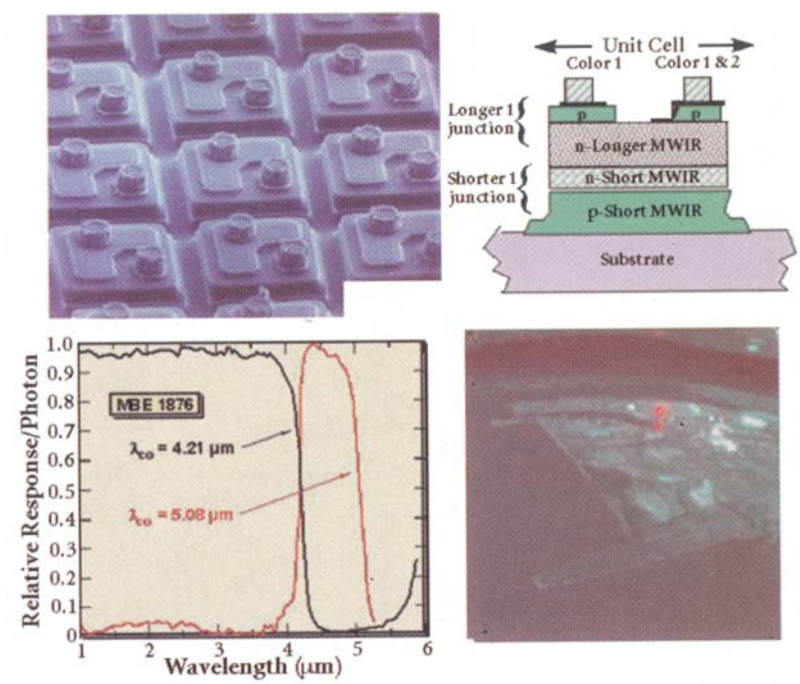

Figure 2. Clockwise starting from top left: a photo-micrograph of several detectors in a dualband IR focal plane array; a schematic of the detector structure of an individual pixel; a dualband color image using color fusion; and the spectral response for each band.

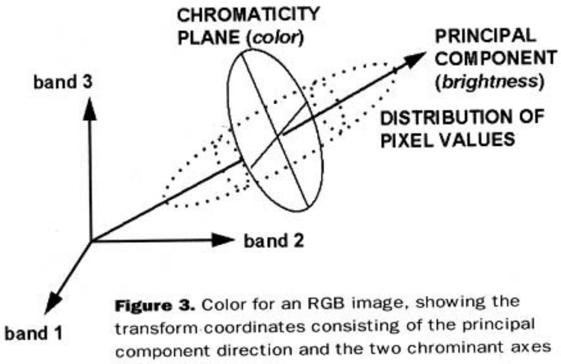

Figure 3. Color for an RGB image, showing the transform coordinates consisting of the principal component direction and the two chrominant axes in the chromaticity plane.

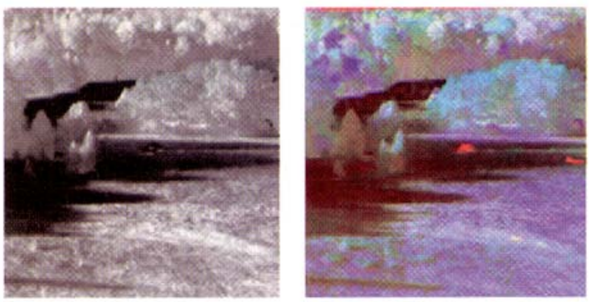

Figure 4. A three-band color fusion image of an MWIR band (4.5–5.0 mm) assigned to blue, and two LWIR bands (8–9 mm and 9–10 mm) assigned to green and red. The data has been digitally processed to maximize the color contrast. The vehicle, virtually hidden in the single-band image, is an obvious magenta-and-red in the composite color image.

color fused image. Even if the cameras have matching fields-of-view, it is usually not enough to simply boresight the cameras. Instead, a complete pixel registration is required. Differences in rotation, lens magnification, chromatic aberration, and distortion can all lead to pixel misregistration. Generally two non-registered images can be registered using computer-processing techniques to stretch the images and interpolate the pixel values. This computationally intensive operation can only be achieved in realtime using a customized processor.

A more elegant approach to multiband imaging is to perform the sensor fusion at the focal plane array level. This eliminates the need for pixel registration, provided that an appropriate dualband optic exists or can be fabricated. A number of efforts are under way to develop dualband IR focal plane arrays. Currently, arrays have been fabricated that image in the LWIR/MWIR or MWIR/MWIR mode. In both cases, the technology relies on the use of HgCdTe detector arrays in a stacked diode structure (see Fig. 3). Some important features of dualband IR focal plane arrays are collocated detectors and simultaneous integration of the signals from each band. Combining this with full *staring* efficiency (as opposed to *scanning*) provides very high signal-to-noise ratios. The two detectors in each unit cell are electrically hybridized to a silicon readout integrated circuit using an indium bump-bond technique. The development of this technology is very challenging because a number of requirements need to be met to realize the full potential of dualband arrays. Specifically, these arrays need to meet the same specifications as single-band arrays, but without band-to-band crosstalk. Inherent in this design is the difficulty of operating two parallel detector circuits inside each unit cell.

Multi-band IR visualization: Phenomenology considerations

There are several IR phenomenology factors that affect color processing for visualization. First, because IR emissions are based on the Planck distribution function, the flux ratios between pixel values in any two bands as a function of temperature are inherently nonlinear. A simple way to linearize such data is to represent it as apparent temperature, a technique used in IR radiometry. Second, the IR photon flux radiated from an object is also dependent on the surface emissivity. Just as surfaces in the visible band have spectral reflectance characteristics, described as $0 < \rho(\lambda) < 1.0$, surfaces are characterized in the IR wavelengths as having a spectral emissivity $0 < \varepsilon(\lambda) < 1.0$. It is known from Kirchhoff's law for radiation (and because of the conservation of energy law), that at a given temperature the spectral emissivity of a point on the surface of a thermal radiator in a given direction is equal to the spectral absorbance for the incident radiation coming from that direction. Therefore, at any given spectral wavelength, $\rho(\lambda) = \varepsilon(\lambda) - 1$. Note that an ideal blackbody emitter has by definition an emissivity of unity.

The photon flux incident on an IR sensor can have both solar-reflected and self-emitted components. Obviously the solar-reflected component would be only significant during the daytime. However, it should be noted that in multi-band images, any of the bands composed of both solar-reflected and self-emitted photons can cause some fundamental problems in terms of analyzing colors. A simple visual example makes this clear.

A gray iron rod appears gray when it is cold and viewed under typical daytime illumination. If one end of the rod is heated to several hundred degrees Celsius and viewed under dark conditions, then it glows amber. In the latter case, the photons are all self-emitted. In the former, they are all reflected. A direct analogy can be drawn to IR color visualiza-

Figure 5. An airborne image taken with a dualband MWIR/MWIR sensor of an industrial plant on a river. The exhaust flume is very saturated in the red band and the storage tank in the near-field is bright cyan as a result of strong solar reflections.

Figure 6. A nighttime dualband image taken with a low-light level camera (responsive in both the visible- and NIR-bands) and an uncooled LWIR camera. In the single-band low-light level image, there is some blooming and a number of dark areas. In the LWIR image, these same areas have better contrast. The composite color image shows that the imagery from both bands can be combined with essentially no loss of information in a clear and intuitive color image.

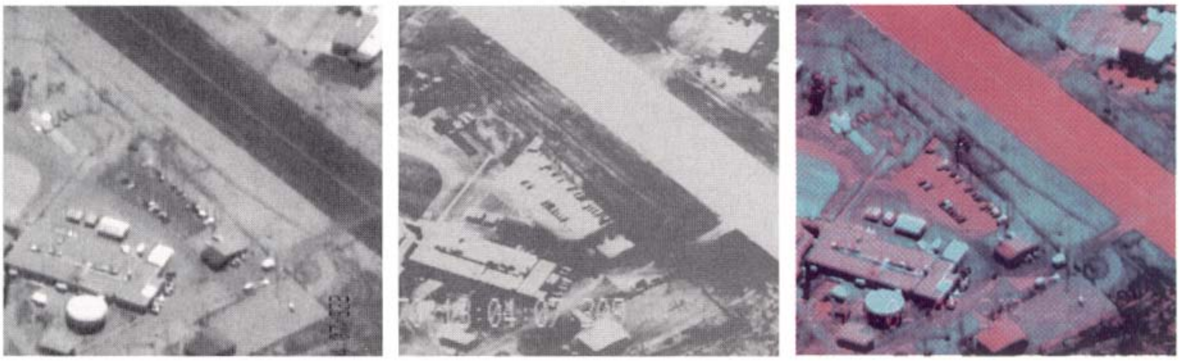

Figure 7. A daytime image taken from an aircraft flying over an industrial site in a desert setting using a TV camera and an LWIR FLIR. All the information from both bands is seen clearly in the composite color image. Color aids in clearly delineating buildings, parking areas, vegetation, and barren soil.

tion, except the incident flux is generally dominated by self-emitted photons in the LWIR. However in the MWIR, the flux distribution is more evenly distributed. For example, during the day, the photon flux in the 4–5 μm band is typically 25% solar-reflected and 75% self-emitted. In the 3–4 μm band the situation is reversed: 75% of the photons are solar-reflected and 25% are self-emitted.

Basics of IR color processing algorithms

In color processing for multi-band IR imagery,[6] it is important to first define the color space representation. As a starting point, consider the color space for representing visible color imagery, where the 3-D color space is based on the red, green, and blue spectral bands to which the photoreceptors of the human retina are sensitive. This allows data to be analyzed using traditional vector space methods, and results to be displayed on an RGB monitor. If each band has a dynamic range of n bits, then a 3-D vector space has a dynamic of $3n$ bits, or about 16 million colors.

This is very impressive when compared to a single-band, monochrome image where the human eye only discerns about 100 shades of gray at any instant in time. Clearly, combining bands into color images provides a method for powerfully discriminating objects from backgrounds using detectors with limited dynamic range. For IR color vision, the pixel values from three IR bands can be similarly represented in a 3-D vector space (see Fig. 4, page 29).

One problem that arises when representing IR data and mapping it to RGB space, is that the pixel values from band-to-band are often highly correlated for those IR wavelengths that are predominantly self-emitting. An example of two correlated bands is the MWIR and LWIR at nighttime. Simple scaling or normalizing of the data in the RGB-space will not remove the correlation. Note that other IR bands of importance are the SWIR and near-IR, which are predominately used in daytime imaging systems that view solar-reflected photons. Typically, these bands are anti-correlated when compared to corresponding MWIR and LWIR. To decorrelate the MWIR and LWIR imagery, and thus increase the color contrast, a simple Hotelling transform[7] could be performed; however, this would cause serious problems in terms of achieving color constancy. A modified approach is to find the principal component direction based on the scene statistics, and to use this as a primary axis for creating a transform space, namely a brightness direction and a chromaticity plane (see Fig. 3). The principal component direction is taken to be the brightness direction, and the chromaticity plane is orthogonal to it. The data are then decorrelated in the transform space by normalizing the pixel values.

In standard color representation, the chromaticity plane is either described as polar coordinates (hues being from 0–360°, and saturation being a positive value in the radial direction), or in rectangular coordinates (chrominant axes 1 and 2, sometimes described as the red-green and the yellow-blue directions). In either case, a transform space, with one axis being a principle component direction, is very useful because it allows certain types of color processing to be performed in a separable manner.

A simple representation of the three-band IR case and accompanying color contrast enhancement described above is fairly straightforward because the number of bands matches that of the human eye. Displaying two bands of IR to a human observer, who has three types of photoreceptors, is an inherently poor fit. Because the human color processing system works with color opponents (namely red-green and yellow-blue), it is possible to use a variant of this to display a dualband IR image.

In the work shown in these figures, a red-green color opponency was chosen. In actual fact, the longer spectral wavelength IR band was input to a red display channel, and the shorter wavelength band was input to both the green and the blue display channels, creating cyan. In essence, this is the same as red-green color opponency. In this manner, the relative intensities between the two bands at each pixel can be represented as a chromatic continuum starting as red, going through gray, and ending as cyan. At the same time, a brightness value can be assigned in an orthogonal direction. Therefore, each pixel has a chrominant value (red-cyan) and a brightness value (black-white). There are a number of statistical processing operations that can be applied to the image data to map it into a color space (chrominant-brightness).

It is not enough to only enhance color contrast. To have an intuitively meaningful image, the image color must be consistent regardless of time-of-day, illumination, and temperature. In essence, color constancy involves computing the surface properties of the scene constituents, regardless of the spectral distribution of the illuminant. For visible wavelengths, this means computing the surface reflectivity. Essentially this is an inverse problem, where the radiation reflected from a surface in each band is known, but the illuminant and the surface reflectivity are not. A complete solution to this problem would be to find the intensity of the illuminant, $I(\lambda)$, and the reflectivity, $\rho(\lambda)$, associated with each pixel.

For the case of only self-emitted IR radiation (for example, the LWIR band at nighttime), the radiant intensity is known and the surface temperature and the surface emissivity are not. In a manner similar to that used for visible color processing, the objective of IR color constancy is to compute the spectral emissivity value for each pixel. Thus, an approach to IR color constancy is to make a reasonable estimation of the surface emissivity in each band. If all of the energy in each band is from only self-emitted radiation, then the color constancy computation can proceed in a manner directly analogous to that of the case of only solar-reflected radiation. The major phenomenological assumption inherent in this case is that objects and backgrounds are at thermal equilibrium, that is, a static ambient condition. In general, this is typical of nighttime backgrounds that are not being dynamically heated by solar flux. Although the static, ambient condition is not always true, it is a very important case, as it is typically the condition in which objects are most difficult to discriminate from backgrounds.

As previously discussed, IR color constancy computations can be complicated by the presence of both solar-reflected and self-emitted radiation. Aside from the two distinct cases mentioned above (only-solar-reflected and only-self-emitted radiation), there is the combined case where both are present in one or more bands. To varying degrees, this effect will be greatest in the MWIR bands during the daytime. When this occurs, it greatly complicates computing color constancy, as projections onto the chromaticity plane are no longer valid.

One approach to addressing this complication is to make a simple estimate of the fractional amount of radiation received from reflected-solar and self-emitted IR radiation. The approach is a two-step process. First, an estimate of the spectral composition of the radiation received at the sensor is made, then the estimate of the composition is used to represent the data as a combination of two distinct components. In a pure mathematical sense this approach is internally inconsistent and cannot be used to solve an inverse problem in a rigorous sense. However, by having a good *a priori* knowledge of the background emissivity (or reflectivity) statistics, it is possible to make fairly good estimates of the illuminant and radiant intensities. This is particularly true if the bands are well separated because the spectra of the solar illuminant and an ambient blackbody emitter are very different—for example, the band combination of SWIR, MWIR, and LWIR. The pixel values can then be represented as a system of equations written in terms of the two separate components.

A basic assumption is that the detected IR radiation can be represented as a linear combination of a solar spectral distribution and a terrestrial, ambient blackbody spectral distribution. For example, the radiation can be represented as the weighted sum of a 6,000 K blackbody spectra (the Sun) and an ambient temperature blackbody (*i.e.*, 290 K). These two distributions can be normalized and thought of as a set of basis functions, although they are not necessarily orthogonal.

The challenge is to generate an intuitively meaningful color image, for a specific viewer and application.

Examples of multiband IR images

Using the processing techniques described above, a three-band LWIR image of a vehicle in a wooded clearing is shown in Figure 4. Note that in a low-light-level single-band image, the vehicle has very low contrast against the neighboring background, whereas in the color image, the vehicle has good color contrast and is easily discernible. Figures 5 (page 29), 6, and 7 show other applications of applying the technique discussed above.

Performance predictions and band selection

One important question that arises regarding the concept of color vision is how much performance improvement it provides compared to single-band imaging systems. One method we have used to make a quantitative comparison of spatio-spectral discrimination (composite color processing) versus single-band discrimination (spatial-only processing), is to derive multi-dimensional matched filters and apply them to detect targets embedded in the background imagery. Single-band filters (2-D spatial) can be derived using an estimated power spectrum[8] for the expected background, and an energy spectrum to represent the target shape. A color filter (3-D spatio-spectral) can be derived by considering the target and the background as a 3-D space in which the third dimension at each pixel is the spectral value of each pixel. This is directly analogous to the human ability to discriminate color and shape.

A rigorous analysis of target detection is usually based on receiver-operator curves (ROC), in which detections and false alarms are compiled based on varying levels of threshold settings. Basically, a plot is constructed of the number of false alarms versus the number of missed detections. Figure 8 gives an example of several ROC curves comparing spatio-spectral processing and spatial-only processing, using matched filters applied to a sample image. The various points along each ROC curve correspond to different threshold settings. Data are shown for three individual bands—LWIR, MWIR, and SWIR labeled R, G, and B, respectively. The spatio-spectral result from the color processing is labeled C. A number of multi-band IR images have been processed using this technique. Typically, the color is found to be 3–10 times better than single-band IR performance.

A critical issue in the successful use of color fusion systems is selecting the appropriate spectral bands. Surface reflectivities and emissivities as a function of spectral wavelength are a fundamental factor in defining the color of an object. The reflectivity of dark green camouflage paint, green cloth camouflage netting, conifers, and grass have been compared. In the visible, the spectral reflectances of all the materials are quite similar, as expected, because they are all essentially green. Beyond the visible, there are major differences in overall reflectivities. Although there are a number of small, distinct spectral features associated with each material, large differences in target contrast are obtained in integrating the absolute sensor response over wide spectral bands.

Although the image improvement may be apparent to a casual viewer, psycho-visual tests are necessary to assess the performance enhancement for color fused imagery over monochrome imagery. The impact of color fusion on factors important to military aviation is being studied at the Naval Postgraduate School.

Summary

The goal of IR color vision is to display multiple bands in one image in a way that improves the viewer's scene comprehension and emphasizes relevant information.

Improved sensor technologies, such as dualband IR focal plane arrays, are being used to create novel image visualization across the electromagnetic spectrum. Biological investigations have inspired many of the techniques used for color processing. Color contrast enhancement and color constancy are goals of the algorithms used to present the imagery. The self-emissive nature of objects in the IR presents an additional level of consideration for the color processing algorithms. Performance prediction studies show that multiple bands improve target detection and reduce false alarm rates. Psycho-visual tests indicate that color fused imagery improves factors important to military applications. The choice of bands for a particular application, with respect to scene phenomenology, is central to the development of IR color vision.

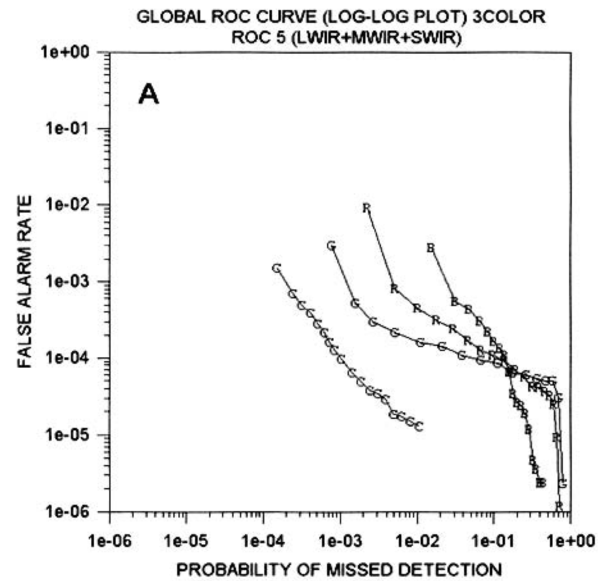

Figure 8. Receiver-operator-curves from a matched filter operation. Data is shown for three individual spectral bands—LWIR, MWIR, and SWIR, labeled R, G, and B, respectively. The various points along each ROC curve correspond to different threshold settings. The spatio-spectral result from the composite color processing of all bands simultaneously is labeled C.

These new focal plane array and computer technologies, recent studies of physical phenomenology in various bands, new algorithms, and performance prediction studies provide encouraging indications that, in the future, color visualization of multiple bands will be a beneficial technique.

References

1. G.W. Wyzecki and W.S. Stiles, *Color Science* (John Wiley & Sons, New York, N.Y., 1967).
2. R.M. Boynton, *Human Color Vision* (Optical Society of America, Washington, D.C., 1992).
3. S. Zeki, "The visual image in mind and brain," Sci. Am. **267** (3), 68–76 (1992).
4. G. Healey *et al., Physics-Based Vision: Principles and Practice, Color* (Jones and Bartlett, Boston, Mass., 1992).
5. For an update on recent developments in IR cameras, see SPIE **3379** (SPIE Press, Bellingham, Wash., April 1998).
6. D. Scribner *et al.*, "Infrared Color Vision: Separating Objects from Backgrounds," Proc. SPIE **3379** (SPIE Press, Bellingham, Wash., 1998) pp. 14–17.
7. R.C. Gonzalez and P. Wintz, *Digital Image Processing*, 2nd ed. (Addison-Wesley, Reading, Mass., 1987).
8. A. Papoulis, Chap. 12, *Signal Analysis* (McGraw-Hill, New York, N.Y., 1977).

Dean Scribner, Penny Warren, Jon Schuler, Michael Satyshur, and Melvin Kruer are research physicists at the Naval Research Lab., Washington, D.C.