
Solid-State Color Night Vision: Fusion of Low-Light Visible and Thermal Infrared Imagery

Allen M. Waxman, Mario Aguilar, David A. Fay, David B. Ireland, Joseph P. Racamato, Jr., William D. Ross, James E. Carrick, Alan N. Gove, Michael C. Seibert, Eugene D. Savoye, Robert K. Reich, Barry E. Burke, William H. McGonagle, and David M. Craig

■ We describe an apparatus and methodology to support real-time color imaging for night operations. Registered imagery obtained in the visible through near-infrared band is combined with thermal infrared imagery by using principles of biological opponent-color vision. Visible imagery is obtained with a Gen III image intensifier tube fiber-optically coupled to a conventional charge-coupled device (CCD), and thermal infrared imagery is obtained by using an uncooled thermal imaging array. The two fields of view are matched and imaged through a dichroic beam splitter to produce realistic color renderings of a variety of night scenes. We also demonstrate grayscale and color fusion of intensified-CCD/FLIR imagery. Progress in the development of a low-light-sensitive visible CCD imager with high resolution and wide intrascene dynamic range, operating at thirty frames per second, is described. Example low-light CCD imagery obtained under controlled illumination conditions, from full moon down to overcast starlight, processed by our adaptive dynamic-range algorithm, is shown. The combination of a low-light visible CCD imager and a thermal infrared microbolometer array in a single dual-band imager, with a portable image-processing computer implementing our neural-net algorithms, and color liquid-crystal display, yields a compact integrated version of our system as a solid-state color night-vision device. The systems described here can be applied to a large variety of military operations and civilian needs.

CURRENT NIGHT OPERATIONS are enabled through imaging in the visible–near-infrared band, as provided by Gen III image intensifier tubes in night-vision goggles, and in the thermal infrared (IR) bands, supported by a variety of forward-looking infrared (FLIR) imaging devices (both scanners and IR focal-plane arrays) displayed on monitors, the cockpit heads-up display, or combiner optics [1, 2]. These dual sensing modalities are complementary, in that the intensifier tubes amplify reflected moon light and starlight (primarily yellow through near-infrared light), whereas the FLIR senses

thermally emitted light (in the mid-wave or long-wave infrared) from objects in the scene. Each sensing modality has its own limitations, which at times can be disorienting [3], while alternating between these modalities can be difficult, confusing, and distracting [4]. However, there is much utility in fusing this complementary imagery in real time into a single image product. This article describes a methodology to provide such fused imagery in color and in real time.

Prior to our work [5, 6], existing methods for visible/infrared image fusion were based on taking local measures of image contrast, choosing between the vis-

ible and infrared image on a pixel-by-pixel basis, and attempting to maximize contrast [7, 8]. The result is a grayscale fused-image product that combines features (and noise) present in each of the separate image bands. Texas Instruments Corporation (now Raytheon Systems) has developed a similar system for the grayscale fusion of intensified visible and FLIR imagery (the methods of which are proprietary). This system has been tested by the U.S. Army Night Vision and Electronic Sensors Directorate (NVESD) under its Advanced Helicopter Pilotage Program [9].

Recognizing that color vision evolved in animals for survival purposes, we describe in the following section a methodology, based on biological opponent-color vision, to fuse registered visible and infrared imagery in real time in order to create a vivid color night-vision capability, as shown in the section entitled "Dual-Band Visible/Infrared Imagers and Fusion Results." Utilizing full (24-bit digital) color allows for simultaneous presentation of multiple fused-image products. The user's visual system can then exploit this coloring to aid perceptual pop-out of extended navigation cues and compact targets [10, 11]. The ability to generate a rich color percept from dual-band imagery was first demonstrated experimentally in the visible (red and white imagery) domain by E.H. Land [12, 13], and motivated his famous retinex theory of color vision [14], which itself lacked any notion of opponent color.

In the latter part of the article we summarize our work on the development of low-light-sensitive CCD cameras, which are sensitive from the ultraviolet through near infrared, and which operate at thirty frames per second in controlled illumination conditions from full moon to overcast starlight. These solid-state imagers possess extremely high quantum efficiency and low read-out noise, which together yield an extreme low-light sensitivity and support a large intrascene dynamic range. Their utility is increased by our retina-like computations that enhance visual contrast and adaptively compress dynamic range in real time. These CCDs for night vision emerge from technology originally developed at Lincoln Laboratory [15, 16] for high-frame-rate applications (i.e., adaptive optics and missile seekers). They represent the beginning of the technology curve for

solid-state visible night vision, and they are complemented by emerging solid-state uncooled thermal infrared imagers [17], as well as a variety of cryogenically cooled infrared focal-plane arrays.

We conclude with a discussion on the importance of conducting human perception and performance testing on natural dynamic scenes in order to assess the true utility of visible/infrared fusion and color night vision for enhanced situational awareness and tactical efficiency.

Visible/Infrared Fusion Architecture

The basis of our computational approach for image fusion derives from biological models of color vision and visible/infrared fusion. In the case of color vision in monkeys and man, retinal cone sensitivities are broad and overlapping, but the images are quickly contrast enhanced within bands by spatial opponent processing via cone-horizontal-bipolar cell interactions creating both ON and OFF center-surround response channels [18]. These signals are then color-contrast enhanced between bands via interactions among bipolar, sustained amacrine, and single-opponent color ganglion cells [19, 20], all within the retina. Further color processing in the form of double-opponent color cells is found in the primary visual cortex of primates (and the retinas of some fish). Opponent processing interactions form the basis of such percepts as color opponency, color constancy, and color contrast, though the exact mechanisms are not fully understood. (See section 4 of Reference 21, and Reference 22, for development of double-opponent color processing applied to multispectral infrared target enhancement.)

Fusion of visible and thermal infrared imagery has been observed in several classes of neurons in the optic tectum (evolutionary progenitor of the superior colliculus) of rattlesnakes (pit vipers) and pythons (booid snakes), as described by E.A. Newman and P.H. Hartline [23, 24]. These neurons display interactions in which one sensing modality (e.g., infrared) can enhance or depress the response to the other sensing modality (e.g., visible) in a strongly nonlinear fashion. These tectum cell responses relate to (and perhaps control) the attentional focus of the snake, as observed by its striking behavior. This discovery pre-

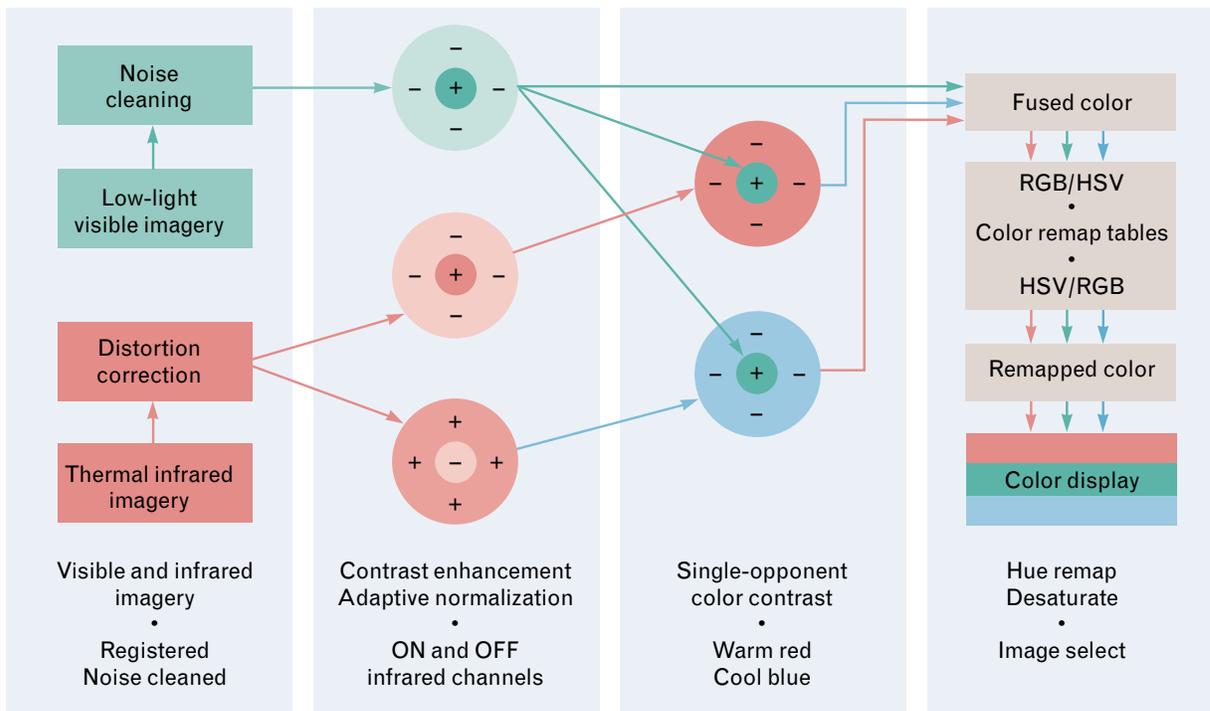


FIGURE 1. Neurocomputational architecture for the fusion of low-light visible and thermal infrared imagery, based on principles of opponent processing within and between bands, as motivated by the retina.

dates the observation of bimodal visual/auditory fusion cells observed in the superior colliculus [25]. Moreover, these visible/infrared fusion cells are suggestive of ON and OFF channels feeding single-opponent color-contrast cells, a strategy that forms the basis of our computational model.

There are also physical motivations for our approach to fusing visible and infrared imagery, revealed by comparing and contrasting the different needs of a vision system that processes reflected visible light (in order to deduce reflectivity ρ) versus one that processes emitted thermal infrared light (in order to deduce emissivity ϵ). Simple physical arguments show that spectral reflectivity and emissivity are linearly related, $\rho(\lambda) = 1 - \epsilon(\lambda)$, which also suggests the utility of ON and OFF response channels. Thus it is not surprising that FLIR imagery often looks more natural when viewed with reverse polarity (black hot as opposed to white hot, suggestive of OFF-channel processing [18]). This simple relation strongly suggests that processing anatomies designed to determine reflectivity may also be well suited for determining emissivity; therefore, computational models of these

anatomies will also be well suited for determining both reflectivity and emissivity.

Figure 1 illustrates the multiple stages of processing in our visible/infrared fusion architecture. These stages mimic both the structure and function of the layers in the retina (from the rod and cone photodetectors through the single-opponent color ganglion cells), which begin the parvocellular stream of form and color processing. The computational model that underlies all the opponent processing stages utilized here is the feed-forward center-surround shunting neural network of S. Grossberg [26, 27]. This model is used to enhance spatial contrast within the separate visible and infrared bands, to create both positive (ON-IR) and negative (OFF-IR) polarity infrared contrast images, and to create two types of single-opponent color-contrast images. These opponent-color images already represent fusion of visible and infrared imagery in the form of grayscale image products. However, the two opponent-color images together with the enhanced visible image form a triple that can be presented as a fused color image product.

The neurodynamics of the center-surround recep-

tive fields is described at pixel ij by the equations,

$$\begin{aligned} \frac{dE_{ij}}{dt} = & -AE_{ij} + (1 - E_{ij})[CI^C]_{ij} \\ & - (1 + E_{ij})[G_S * I^S]_{ij}, \end{aligned} \quad (1)$$

which, in equilibrium, yields

$$E_{ij} = \frac{[CI^C - G_S * I^S]_{ij}}{A + [CI^C + G_S * I^S]_{ij}}, \quad (2)$$

where E is the opponent processed enhanced image, I^C is the input image that excites the single pixel center of the receptive field (a single pixel center is used to preserve resolution of the processed images), and I^S is the input image that inhibits the Gaussian surround G_S of the receptive field. Equation 1 describes the temporal dynamics of a charging neural membrane (cf. capacitor) that leaks charge at rate A and has excitatory and inhibitory input ion currents determined by Ohm's law. The shunting coefficients ($1 \pm E$) act as potential differences across the membrane, and the input image signals modulate the ion-selective membrane conductances. Equation 2 describes the equilibrium of Equation 1 that is rapidly established at each pixel (i.e., at frame rate), and defines a type of nonlinear image processing with parameters A , C , and size of the Gaussian surround. The shunting coefficients of Equation 1 clearly imply that the dynamic range of the enhanced image E is bounded, $-1 < E < 1$, regardless of the dynamic range of the input imagery. When the imagery that feeds the center and Gaussian surround is taken from the same input image (visible or infrared), the numerator of Equation 2 is the familiar difference-of-Gaussians filtering that, for $C > 1$, acts to boost high spatial frequencies superimposed on the background. The denominator of Equation 2 acts to adaptively normalize this contrast-enhanced imagery on the basis of the local mean. In fact, Equation 2 displays a smooth transition between linear filtering (when A exceeds the local mean brightness, such as in dark regions) and ratio processing (when A can be neglected as in bright regions of the imagery). These properties are particularly useful for processing the wide-dynamic-range visible imagery obtained with low-light CCDs, as de-

scribed in the latter part of the article. Equation 2 is used to process separately the input visible and infrared imagery. These enhanced visible and ON-IR images are reminiscent of the lightness images postulated in Land's retinex theory [14] (also see Grossberg on discounting the illuminant [26]).

A modified version of Equation 1, with an inhibitory center and excitatory surround, is also used to create an enhanced OFF-IR image (i.e., a reverse-polarity enhanced infrared image). After reducing noise in the imagery (both real-time median filtering and non-real-time boundary-contour and feature-contour system processing [26, 21] have been explored), and correcting for distortion to ensure image registration, we form two grayscale fused single-opponent color-contrast images by using Equation 2 with the enhanced visible feeding the excitatory center and the enhanced infrared (ON-IR and OFF-IR, respectively) feeding the inhibitory surround. In analogy to the primate opponent-color cells [20], we label these two single-opponent images $+Vis - IR$ and $+Vis + IR$. In all cases, we retain only positive responses for these various contrast images. Additional application of Equation 2 to these two single-opponent images serves to sharpen their appearance, restoring their resolution to the higher of the two images (usually visible) used to form them. These images then represent a simple form of double opponent-color contrast between visible and ON/OFF-IR.

Our two opponent-color contrast images are analogous to the infrared-depressed-visual and infrared-enhanced-visual cells, respectively, of the rattlesnake [23, 24]; they even display similar nonlinear behavior. In fact, because the infrared image has lower resolution than the visible image (in the snake and in man-made uncooled infrared imagers), a single infrared pixel may sometimes be treated as a small surround for its corresponding visible pixel. In this context, our opponent-color contrast images can also be interpreted as coordinate rotations in the color space of visible versus infrared, along with local adaptive scalings of the new color axes. Such color-space transformations were fundamental to Land's analyses of his dual-band red and white colorful imagery [12-14].

To achieve a natural color presentation of these op-

ponent images (each is an eight-bit grayscale image), we assign the following color channels (eight bits each) to our digital imagery: (1) enhanced *Vis* to green, (2) $+Vis - IR$ to blue, and (3) $+Vis + IR$ to red. These channels are consistent with our natural associations of warm red and cool blue. Finally, these three channels are interpreted as RGB (red, green, blue) inputs to a color remapping stage in which, following conversion to HSV (hue, saturation, value) color space, hues are remapped to alternative (i.e., more natural) hues, colors are desaturated, and the images are then reconverted to RGB signals to drive a color display. The result is a *fused color* presentation of visible/infrared imagery.

Dual-Band Visible/Infrared Imagers and Fusion Results

We have developed several dual-band imaging systems to collect registered visible and long-wave infrared (LWIR) imagery in the field at night, as shown in Figure 2 [6]. In our first-generation system, shown in Figure 2(a), the visible imagery is obtained by using a Gen III image intensifier tube optically coupled to a conventional CCD (supporting a resolution of 640×480 pixels), while the thermal infrared imagery is obtained by using an uncooled ferroelectric detector array developed by Texas Instruments Corp. (supporting a resolution of approximately 320×240 pixels). The two fields of view (about 30° wide) are matched and imaged through a dichroic beam splitter. In our second-generation system, shown in Figure 2(b), we utilize a Lincoln Laboratory low-light CCD to acquire visible imagery at a resolution of 640×480 pixels, in conjunction with the uncooled LWIR camera.

An alternative LWIR imager we currently use in our third-generation system is the silicon microbolometer array originally developed by Honeywell Corp. [17]. For long-standoff distance (narrow field of view) imagery, we plan to use a cryogenically cooled infrared imager. In the field we record synchronized dual-band time-stamped imagery on two Hi-8 videotape recorders for later processing back in our lab. We also perform real-time computations on Matrox Genesis boards using new TMS320C80 multi-DSP chips from Texas Instruments. For compact portable systems, a head-mounted display could

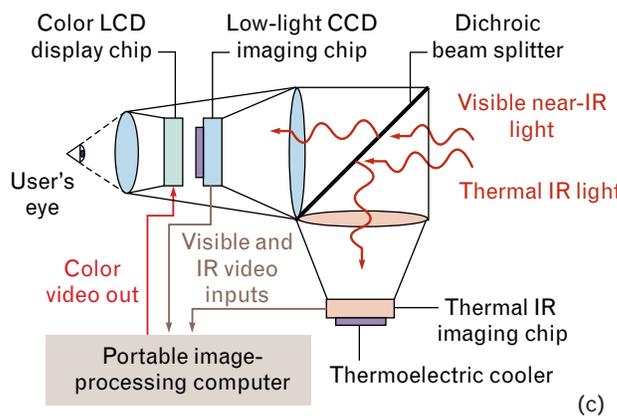
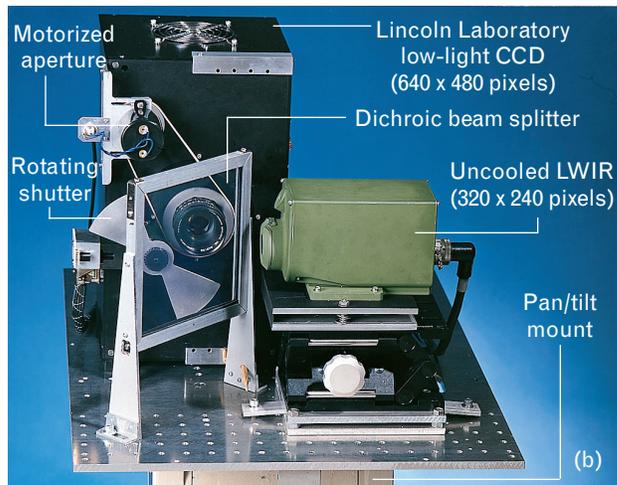
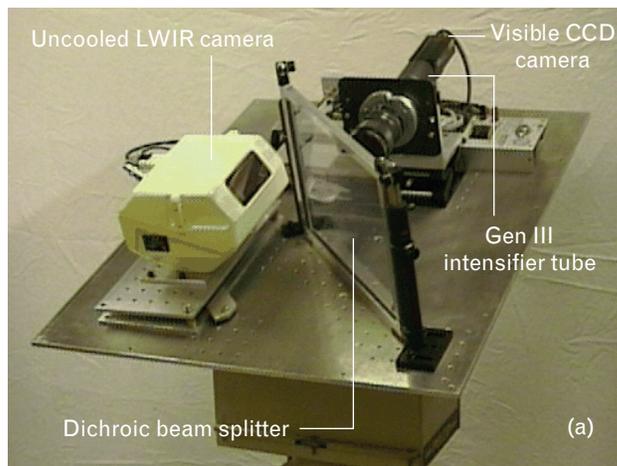


FIGURE 2. Dual-band visible/long-wave-infrared (LWIR) imagers. (a) Sensor pod consisting of a Gen III intensified CCD, an uncooled LWIR imager, and a dichroic beam splitter. (b) Sensor pod consisting of a Lincoln Laboratory low-light CCD, an uncooled LWIR camera, and a dichroic beam splitter. (c) Design of a monocular solid-state color night-vision scope.

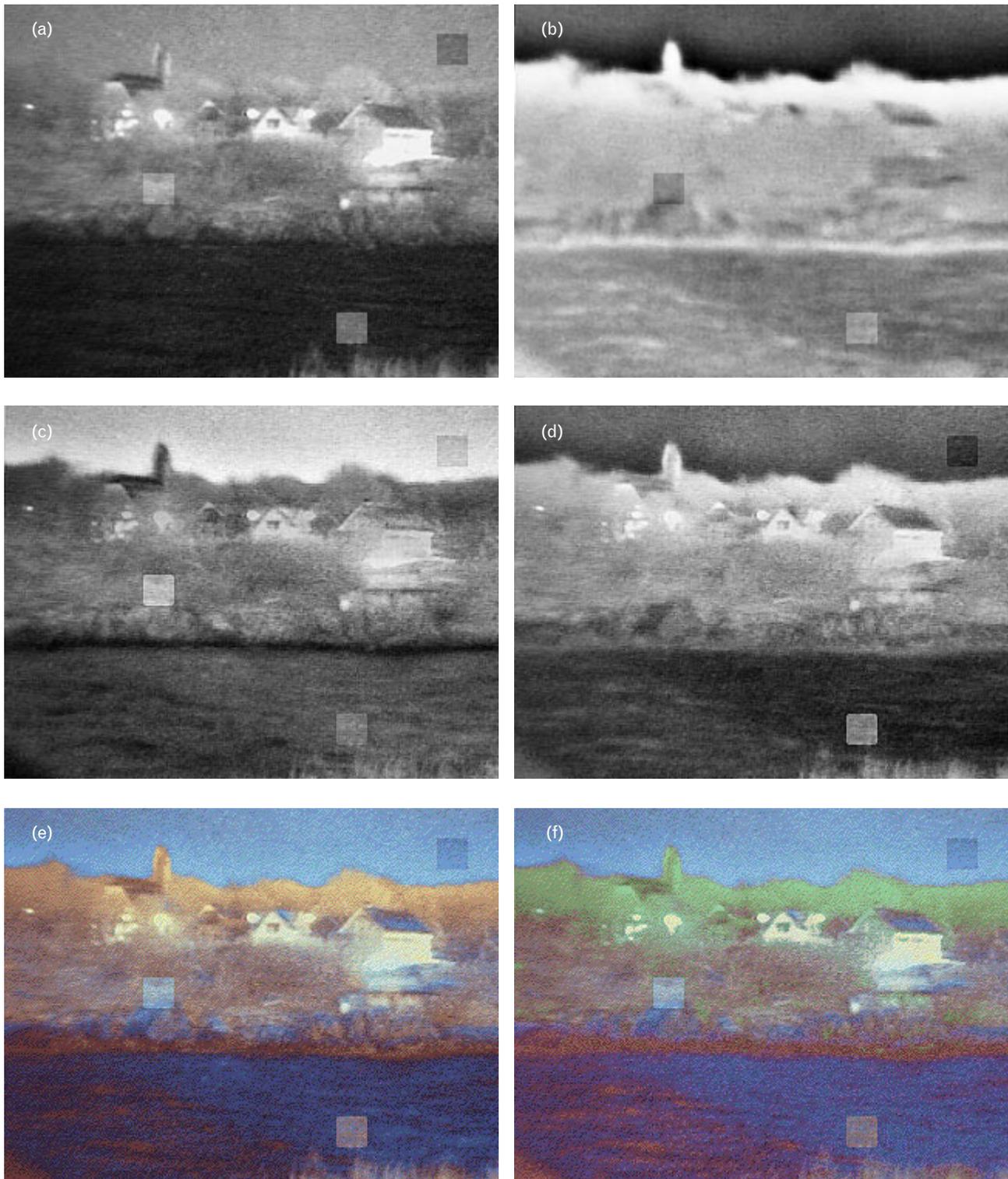


FIGURE 3. Dual-band imagery of the town of Gloucester, Massachusetts, at dusk, with three embedded low-contrast square targets. (a) Enhanced visible image taken with a Gen III intensified CCD; (b) enhanced thermal IR image taken with an uncooled IR camera; (c) gray fused opponent-color (blue channel) image; (d) gray fused opponent-color (red channel) image; (e) color fused image; (f) remapped color fused image. Note that the color fused images support the perceptual pop-out of all three embedded targets from the background.

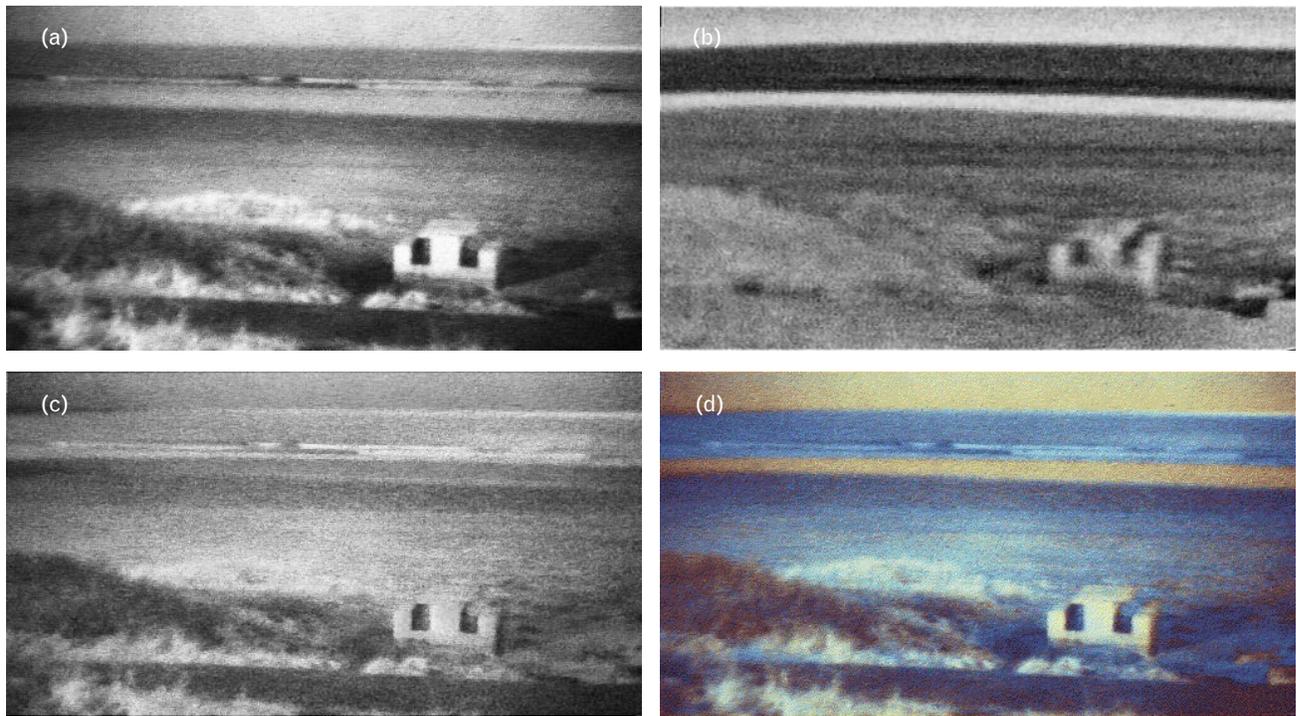


FIGURE 4. Nahant beach on the Atlantic Ocean in overcast near-full-moon illumination conditions. Dual-band visible and thermal IR imagery are combined to create grayscale and color fused images of the night scene. (a) Intensified visible image, (b) thermal IR (FLIR) image, (c) gray fused image, and (d) color fused image.

utilize a solid-state high-resolution color LCD display on a chip, such as the displays being manufactured by Kopin Corp., along with a low-power fusion processor utilizing custom application-specific integrated circuits (ASIC). For vehicle-based applications in which the user is behind a windscreen (which does not transmit thermal infrared light), the dual-band sensor is placed in an external turret or pod with an appropriately transmissive window, while the real-time fusion results are displayed on color helmet-mounted displays or on a monitor.

We are planning to shrink our dual-band sensor to a size of several inches, which would be suitable for use as a hand-held or helmet-mounted color night-vision device (or mounted as a gunsight) for the soldier on the ground. Conceptually, a compact dual-band color night-vision scope could be laid out according to Figure 2(c), in which much of the camera electronics is remotely located away from the low-light CCD imager and microbolometer array [5].

A dual-band visible/LWIR scene of Gloucester, Massachusetts, is shown in each of the scenes in Fig-

ure 3, which includes three embedded low contrast (15% or less) square targets that modulate brightness but do not alter texture in the original visible and infrared images. This imagery was taken under dusk illumination conditions (no moon) with our first-generation system, shown in Figure 2(a), in January 1995. Note the complementary information present in the visible and infrared imagery, where the horizon and water line is obvious in the infrared but not in the visible image, while the ground detail is revealed in the visible but not the infrared. The enhanced visible, enhanced thermal infrared, both opponent-color contrast (i.e., fused gray), fused color, and remapped fused color images are shown in Figure 3. In the fused color images in Figures 3(e) and 3(f), the horizon is clearly rendered, as are the houses and shrubs on the ground, the water line on the rocks, and ripples on the water surface. The enhanced contrast afforded by the color now supports the perceptual pop-out of all three embedded targets, one of which (in the water) is weakened in the gray fused image (c) and one (on the land) is lost in the gray fused image (d). Note that the

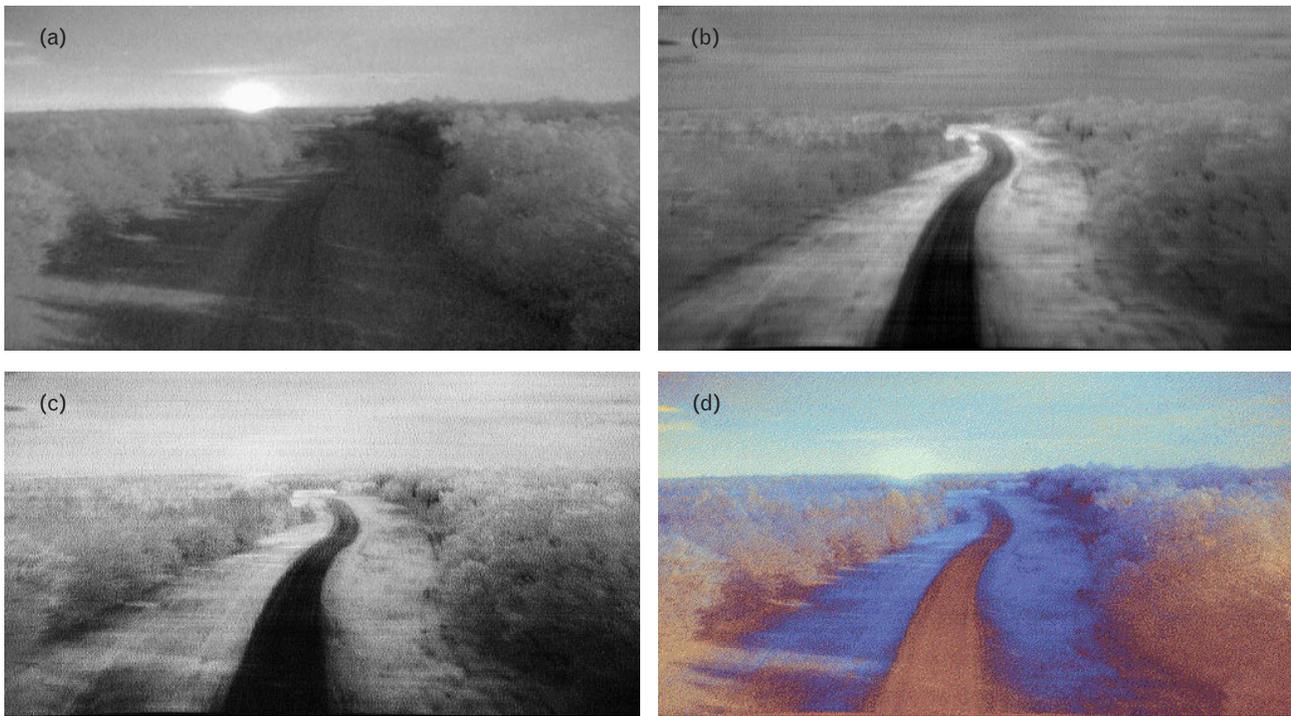


FIGURE 5. Road-scene imagery collected during a helicopter flight provided by the U.S. Army Night Vision and Electronic Sensors Directorate (NVESD). Dual-band visible and FLIR imagery are combined to create grayscale and color fused images of the night scene. (a) Intensified visible image, (b) thermal IR (FLIR) image, (c) gray fused image, and (d) color fused image.

fused color imagery inherits the higher resolution of the visible image. In the remapped color fused image (f), the trees and shrubs corresponding to brown in the fused image (e) have been remapped to a greenish hue with low saturation, and the blue water is brightened. In practice, the class of color remap selected by the user (in real time) will depend on the kind of mission undertaken.

Figure 4 illustrates a scene taken at Nahant beach on the Atlantic Ocean, on an overcast night with near-full moon in January 1995. We illustrate (a) the enhanced visible, (b) the enhanced thermal infrared, (c) the gray fused (blue channel) opponent color, and (d) the unremapped color fused imagery. In the color fused image, notice how the water and surf easily segment from the sand, and how the horizon is clear over the water. A concrete picnic table and asphalt bicycle path are also in the foreground. Real-time processing of this scene is quite dramatic, and the incoming waves are clearly apparent. Notice that the gray fused imagery displays an enhanced surf but a weak horizon. Clearly, even the low resolution and low sensitiv-

ity of the uncooled infrared imager seem adequate in modulating the visible imagery into a color fused result. It will be of great interest to assess the utility of such a night-vision system for search-and-rescue operations at sea.

Figures 5 and 6 present fusion results on data provided by the U.S. Army NVESD, Advanced Helicopter Pilotage Program. Here an intensified CCD provides low-light visible imagery, and a cryogenically cooled first-generation FLIR provides high-quality thermal infrared imagery. In many respects, the FLIR imagery is more useful than the visible imagery. By inspecting the original visible (a) and original infrared (b) images, however, we can clearly see how the sensors complement each other. The gray fused result (c) is shown next to the color fused result (d). In Figure 5 we see that the color fused result (d) displays a clearer horizon, clearer tree shadows across the road, and a better sense of depth down the road than does the gray fused result (c). In Figure 6, both fused results show a strong horizon, but the color fused result (d) reveals more detail near the top of the tower and the

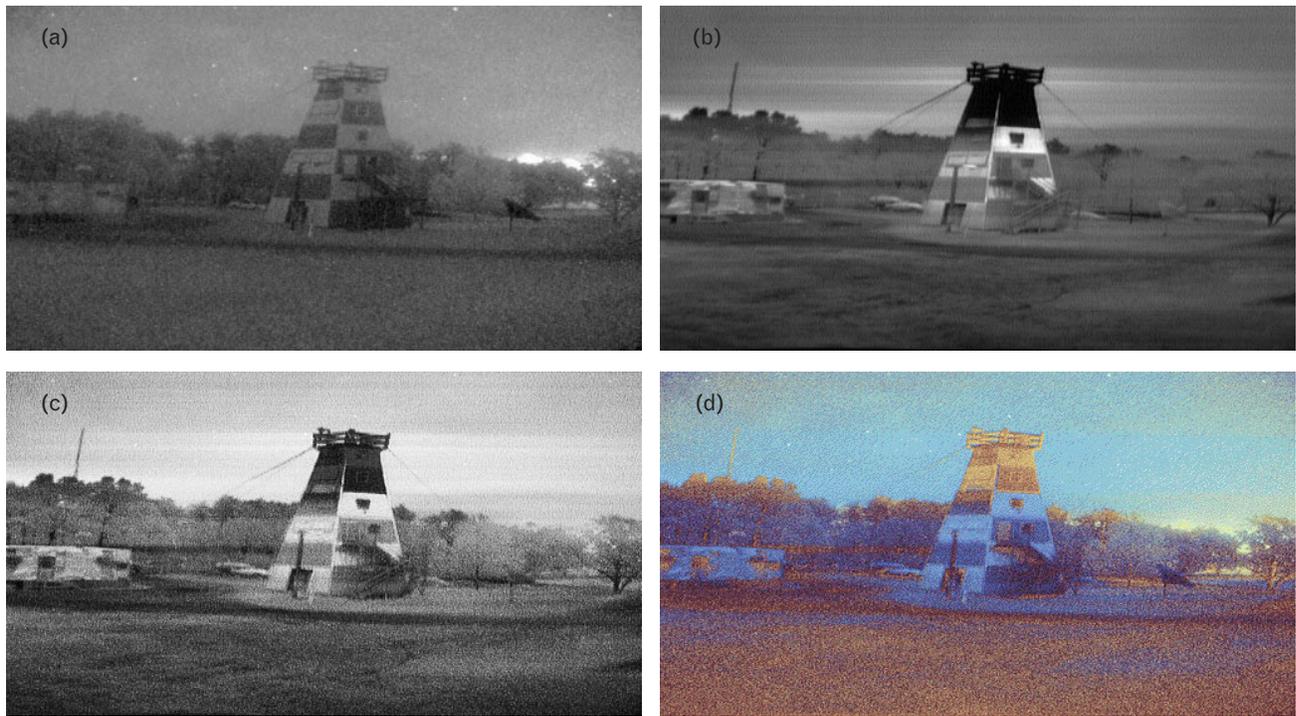


FIGURE 6. Tower-scene imagery provided by the U.S. Army NVESD. Dual-band visible and FLIR imagery is combined to create grayscale and color fused images of the night scene. (a) Intensified visible image, (b) thermal IR (FLIR) image, (c) gray fused image, and (d) color fused image.

communication dish on the ground, whereas the gray fused result (c) reveals more detail on the trailer.

Color Remapping of Fused Imagery

Figures 5 and 6 show the results of color fusion as produced by the direct output of the opponent-color processing described earlier in the section on the visible/infrared fusion architecture. Alternatively, these fused channels can provide input to a final color remapping stage, as shown in the architecture diagram in Figure 1. Color remappings are essentially transformations in the HSV color space, designed to render the fused imagery in more natural coloring. We have developed separate color remappings for different classes of scenes (not for each individual scene), such as forested imagery like Figures 5 and 6, and for flight over water. We expect that different color remappings will be required for desert, ice, and urban class scenes. Figure 7 shows fused and color-remapped examples from the Army helicopter pilotage program images shown in Figures 5 and 6. We have demonstrated real-time fusion with color

remapping on videotaped imagery from an Army helicopter during night flight over forest and water, processing intensified-CCD and FLIR imagery with 640×480 -pixel resolution at thirty frames per second with two TMS320C80 processors.

Figure 8 shows another interesting example involving the penetration of a smokescreen. This imagery, from an unclassified Canadian defense study, was taken during the day with a conventional CCD visible camera and a first-generation FLIR thermal imager. Clearly, the visible image provides the scenic context, whereas the infrared image highlights hot targets. The gray fused imagery puts this complementary information together nicely, but the color fused and remapped result clearly separates the hot vehicles (tow truck and helicopter) from the men running through the smoke and the background.

Perceptual Testing

We have provided fused imagery for several kinds of human performance testing. A.M. Waxman et al. [28] studied human reaction time to detect artificial



FIGURE 7. Color fused remappings of the road-scene imagery in Figure 5 and the tower-scene imagery in Figure 6. Color remapping transforms the red/blue thermal colors of fused imagery to more natural and familiar hues.

targets (i.e., contrast modulations) embedded in a real scene, as shown in Figure 3. Reaction times were compared for gray and color fused, as well as original sensor imagery. A. Toet et al. [29] provided visible and infrared imagery of a person walking among shrubs and sand, taken during early morning hours in which both visible and thermal contrast in the scene were low. Our gray and color fused results were compared to those of Toet and J. Walraven [30], as well as

the individual sensor imagery, for the task of detecting and localizing the person designated as the target in the scene. Twenty-seven frames were used to test the subject population. P.M. Steele and P. Perconti [31] conducted tests on military helicopter pilots by using Army helicopter imagery (both stills and video sequences) to assess accuracy and reaction time in detecting objects and the horizon, as well as determining image quality. Steele and Perconti compared our

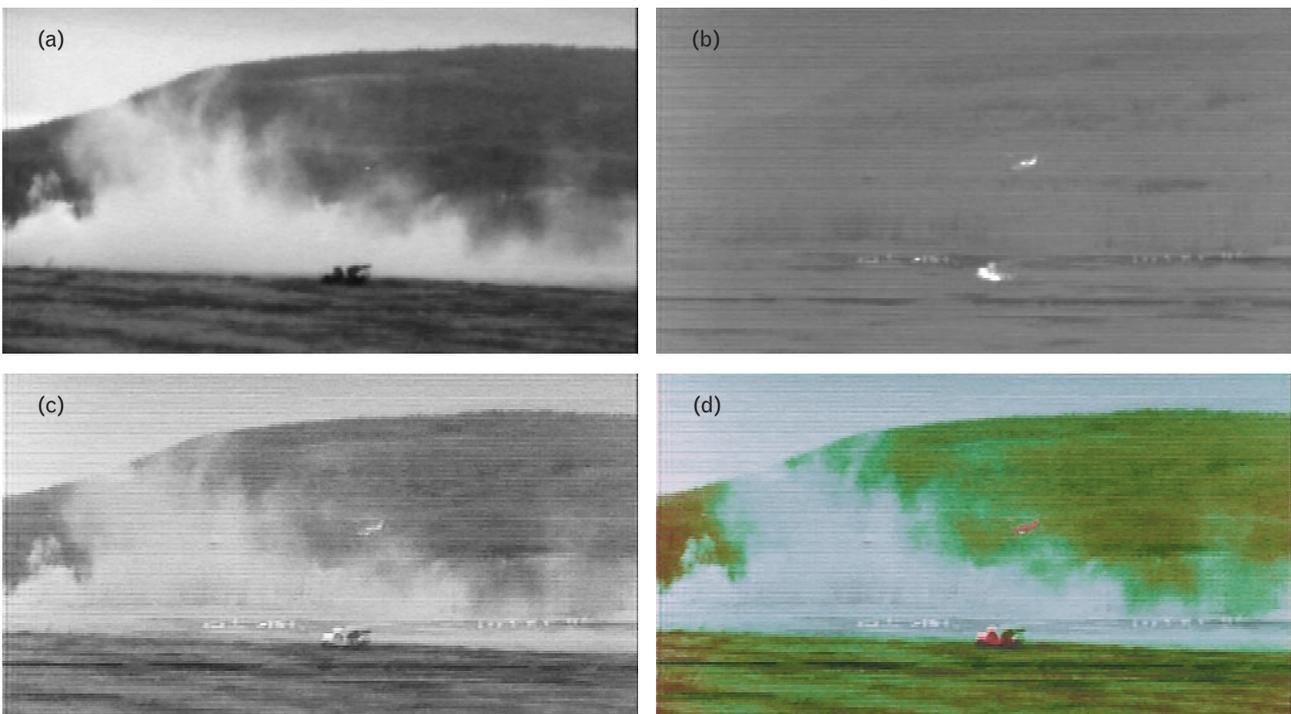


FIGURE 8. Smokescreen penetration and target pop-out is achieved through the color fusion of visible CCD and FLIR imagery in this daytime scene (imagery provided through the Canadian Defense Research Establishment, Valcartier, Québec, as part of a NATO study). (a) Intensified visible image, (b) thermal IR (FLIR) image, (c) gray fused image, and (d) color fused image.

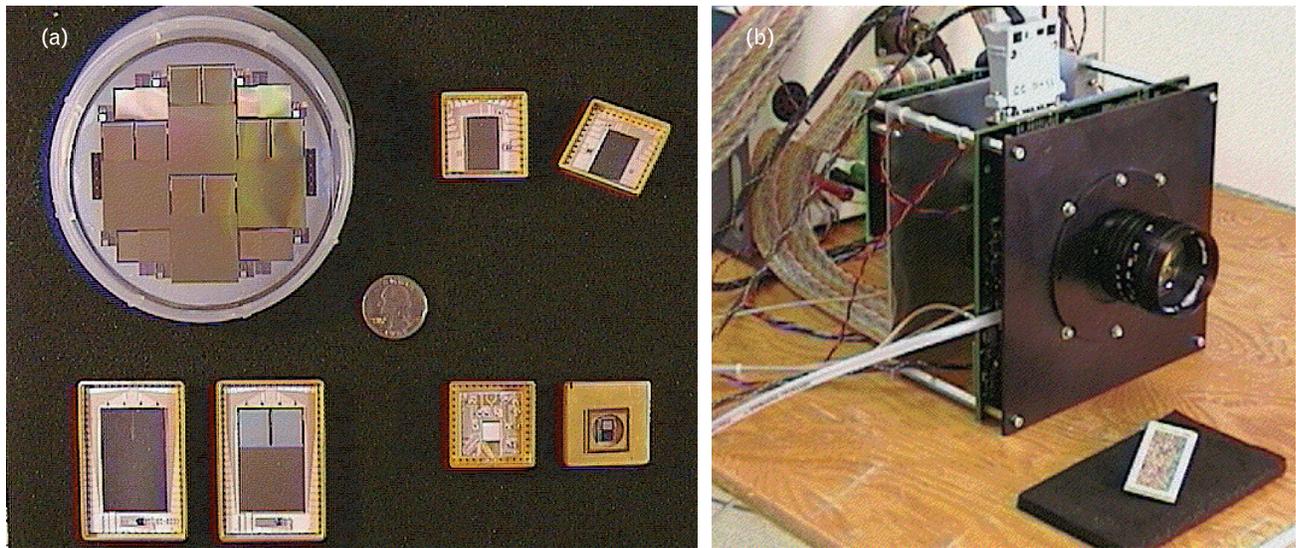


FIGURE 9. Low-light CCD imagers. (a) Thinned wafer and packaged multiported CCDs with formats 1K x 1K, 512 x 512, and 128 x 128 pixels; (b) prototype low-light camera using a four-port 128 x 128 back-illuminated CCD, and an analog circuit multichip module (shown in the foreground).

gray and color fused imagery to proprietary gray fused results from Texas Instruments and a trivial color assignment scheme from the Naval Research Laboratory, as well as the original sensor imagery. In all of these tests, our color fused imagery showed clear improvements in human performance over the original sensor imagery as well as every alternative fusion method it was compared to.

Prototype 128 × 128 Low-Light CCD Imager

Solid-state, thinned, back-illuminated, multiported frame-transfer CCD imagers offer enormous benefits over electro-optic intensifier tubes, including excellent quantum efficiency (>90%), broad spectral sensitivity (0.3–1.1 μ), high spatial resolution, sensitivity in overcast starlight, enormous dynamic range, anti-blooming capability, and near-ideal modulation transfer-function characteristics. Such CCDs with integrated electronic shutters have been fabricated and tested at Lincoln Laboratory [15, 16]. Our near-term target CCD imager has 640 × 480 pixels and sixteen parallel read-out ports; it also supports twelve-bit digital imagery at less than $5e^-$ read-out-noise level, operates at thirty frames per second with integrated electronic shuttering and blooming drains, and requires only thermoelectric cooling (as does the noncryogenic uncooled thermal LWIR imager).

Nearly all of these capabilities have already been developed and demonstrated in different devices. We are currently integrating them into a single imager for night-vision applications.

Figure 9(a) illustrates a variety of low-light CCD imagers (a coin in the center of the image provides size comparison) including (upper left) a wafer patterned with four large 1K × 1K-pixel imagers and four smaller 512 × 512-pixel imagers, thinned to ten microns for back illumination; (lower left) two 1K × 1K imaging chips inside open packages with one mounted for front illumination and the other for back illumination; (upper right) two 512 × 512 imaging chips mounted in open packages; and (lower right) a mounted and sealed 128 × 128-pixel four-port imager and an empty package showing the thermoelectric cooler upon which the imager is mounted. Figure 9(b) shows our first laboratory prototype low-light CCD camera built around a back-illuminated four-port 128 × 128 pixel imager. This camera operates in the dark at thirty frames per second or less (and was actually designed to operate in excess of five hundred frames per second with adequate lighting). In front of the camera is a multichip module containing all the analog circuitry for the four read-out ports; the relatively small size of this module illustrates the potential to build far more compact cameras. Further

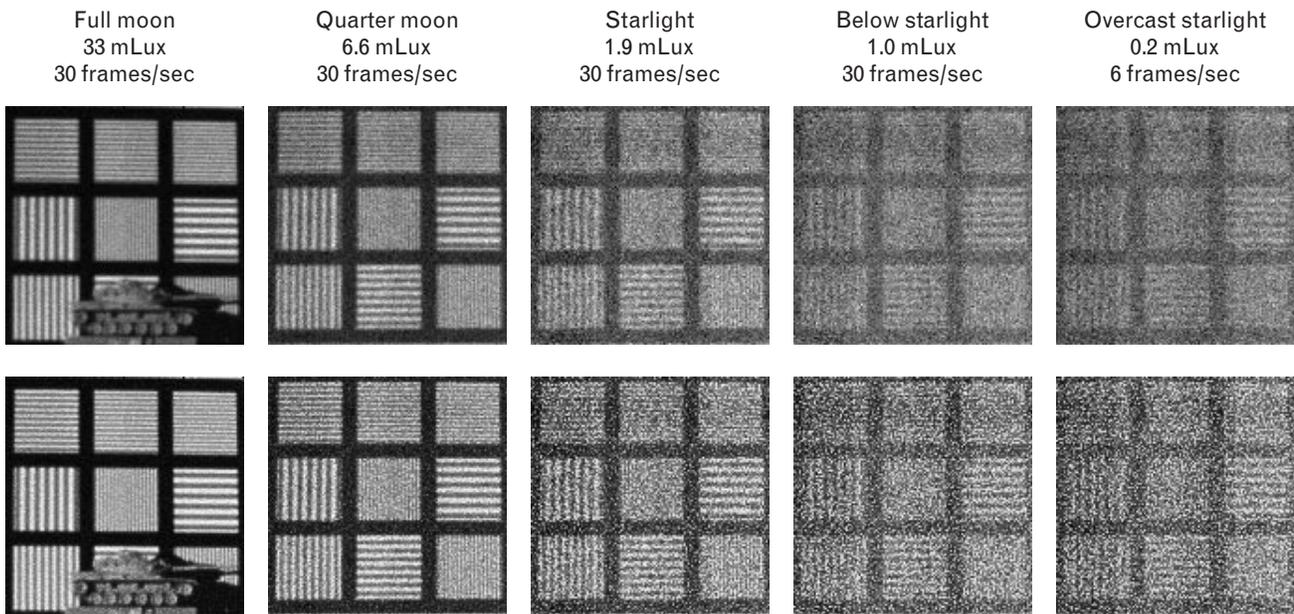


FIGURE 10. Low-light CCD imagery taken at video frame rates under controlled illumination conditions as indicated. The top row shows raw twelve-bit imagery scaled so that minimum and maximum map to zero and 255. The bottom row shows the corresponding eight-bit imagery obtained from center-surround shunt neural processing of the original twelve-bit imagery.

size reduction can be realized through the use of ASICs for the read-out and timing circuitry. This camera operates at thirty frames per second with a measured read-out noise of about $5e^-$.

Figure 10 illustrates imagery obtained in the laboratory with the camera shown in Figure 9(b), under controlled lighting conditions from full moon down to overcast starlight (as measured at the scene with a photometer calibrated for a Gen III intensifier tube,

using a calibrated light source with a blue-cut filter). The scene consists of a 50% contrast resolution chart, and a toy tank in the full-moon example. All images, except for overcast starlight, were taken at thirty frames per second; for overcast starlight the frame rate was reduced to six frames per second. We can obtain better quality imagery at starlight or below by reducing the frame rate below thirty frames per second, thereby integrating photons directly on the imager

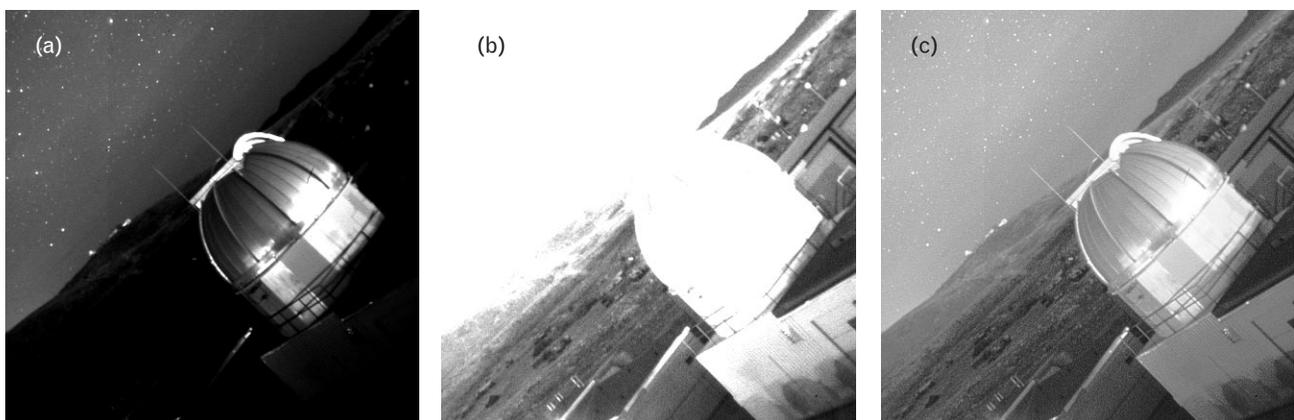


FIGURE 11. Low-light CCD imagery taken at White Sands, New Mexico, under starlight conditions, originally $1K \times 1K$ pixels. (a) The high end of the twelve-bit dynamic range; (b) the low end of the twelve-bit dynamic range; (c) the entire eight-bit dynamic range after center-surround shunt neural processing of the original twelve-bit imagery captures all scenic details.

without the penalty of accumulating additional read-out noise. Across the top row of Figure 10 we show the original twelve-bit imagery scaled such that the minimum pixel value is set to zero and the maximum pixel value is set to 255 on an eight-bit grayscale display. This scaling is possible only because of the simplicity of the scene and uniformity of lighting. Across the bottom row of Figure 10 we show the corresponding images after processing the twelve-bit data with the center-surround shunt processing of Equation 2. In all cases we can see that contrast has been enhanced and the dynamic range has been adaptively compressed to only eight bits. All images were processed exactly the same, without individual adjustments.

Figure 11 shows an example of a 640×480 pixel low-light CCD image. The original image, taken at White Sands, New Mexico, in 1994 under starlight conditions, is approximately $1K \times 1K$ pixels, digitized to twelve bits (4096 gray levels). This high-resolution imagery was taken at a relatively low frame rate (five frames per second), in order to maintain low read-out noise over the imager's four read-out ports. Figures 11(a) and 11(b) are the same image shown at opposite ends of the twelve-bit dynamic range. At the high end of the dynamic range, Figure 11(a) shows the stars in the sky and the horizon, but nothing is visible on the ground. At the low end of the dynamic range, Figure 11(b) shows the presence of vehicles on the ground, but the sky and dome are saturated white. This enormous dynamic range is a tremendous asset for night

imaging, since the moon and cultural lighting can dominate the high end, while objects and shadows on the ground may be apparent only at the low end of the dynamic range (and would ordinarily be lost due to the automatic gain control of an intensifier tube). The center-surround shunting neural networks of Equation 2 can exploit the contrast inherent in the wide-dynamic-range CCD imagery while adaptively normalizing the local data to a dynamic range well suited to only 256 gray levels (i.e., an eight-bit display range). And the computations can be carried out in real time, even at high data rates. Figure 11(c) shows the result of this neural processing, where we can easily see the stars in the sky, the buildings on the horizon, the vehicles on the ground, and the telescope dome without any saturation at either end of the dynamic range.

640 × 480-Pixel Low-Light CCD Imager

The low-light CCD technology described in the previous section has been recently scaled up to produce a 640×480 -pixel, eight-port imager with twelve-bit dynamic range, able to operate at thirty frames per second below starlight illumination conditions. These imagers also contain blooming drains at each pixel, to prevent charge spreading among neighboring pixels in the presence of a brightness overload, and pixel binning to reduce read-out noise when the signal-to-noise ratio supports only lower-resolution imagery. Figure 12 illustrates this new camera, the packaged

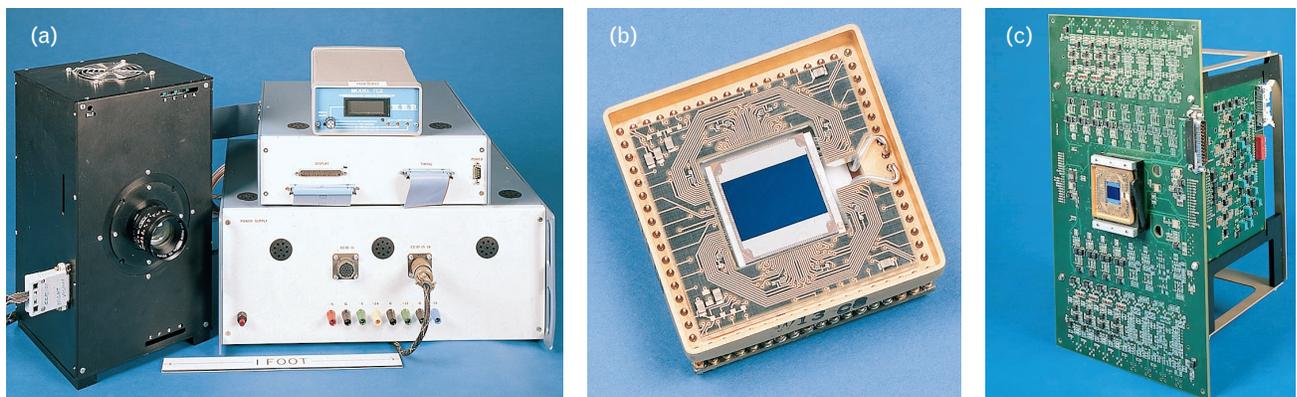


FIGURE 12. (a) The new CCD camera with associated electronics, including thermoelectric-cooler controller (top), digitizer/multiplexer (middle), and power supply (bottom). (b) Packaged CCD imager mounted on a two-stage internal thermoelectric cooler. (c) CCD imager inserted into camera chassis with read-out electronics. This camera is utilized in the dual-band imaging pod shown in Figure 2(b).

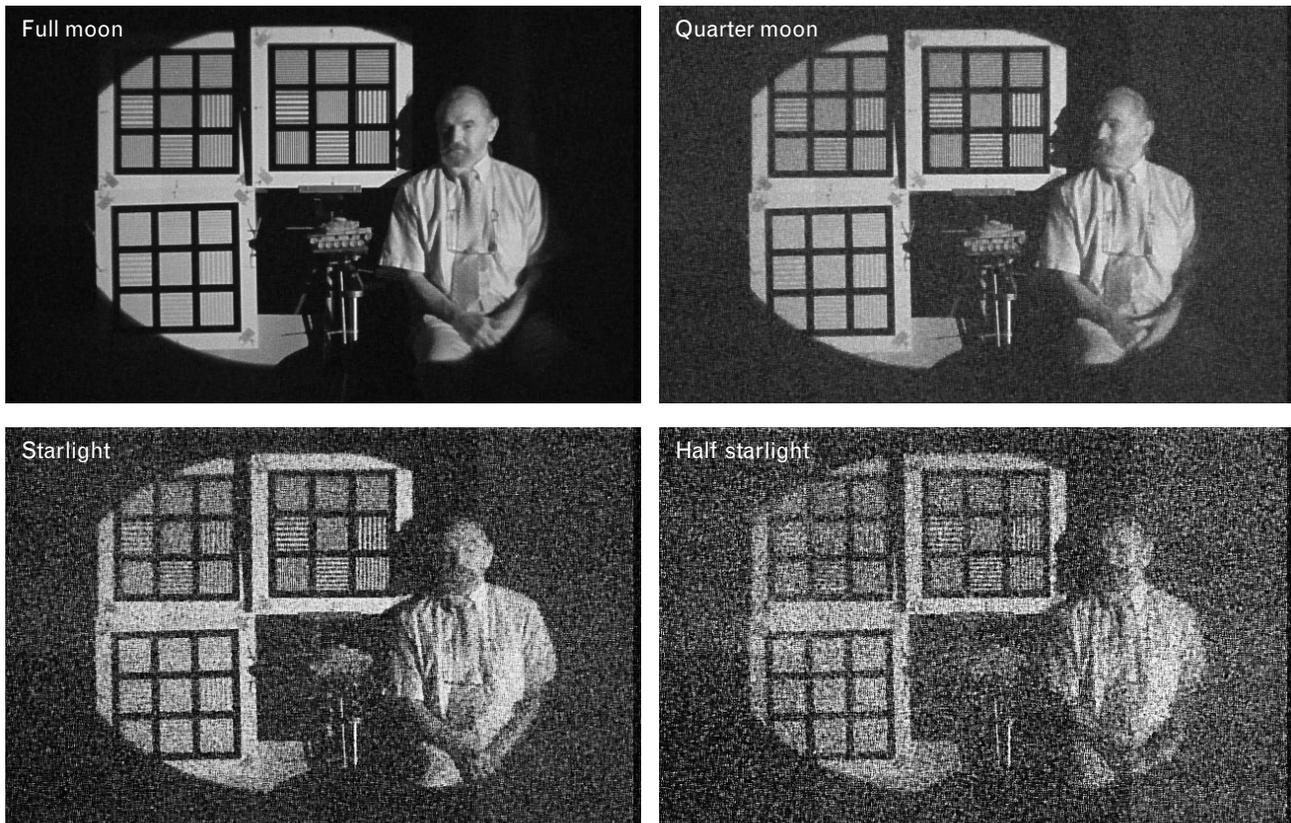


FIGURE 13. Low-light CCD imagery taken with the Lincoln Laboratory CCD camera shown in Figure 12, at a resolution of 640×480 pixels. Under laboratory-controlled scene illumination ranging from full moon (33.3 mLux) in the upper left to half starlight (1.0 mLux) in the lower right, imagery was captured at thirty frames per second with an $f/1.4$ lens, and adaptively processed to reduce noise, enhance contrast, and compress dynamic range to eight bits. In the background of each image are three resolution charts of 100% contrast (upper right), 50% contrast (upper left) and 20% contrast (lower left), and a toy tank (lower right).

640×480 -pixel imager, and the read-out electronics. Figure 13 shows examples of this imagery taken under controlled illumination conditions from full moon down to half starlight. The imagery shown in the figure incorporates adaptive processing for noise reduction, contrast enhancement, and dynamic-range compression.

This low-light CCD visible camera has been integrated with an uncooled LWIR camera, shown in Figure 2(b), and a multi-C80 color fusion processor for demonstration as a driver's night-vision enhancement system (see Reference 32 for our earlier work on electronic imaging aids for night driving). Figure 14 shows an example of color fused imagery derived from the CCD/LWIR imager pod of Figure 2(b). This imagery was collected in March 1998 at the Lincoln Laboratory Antenna Test Range under approximately quarter-moon illumination conditions. Figure

14(a) shows the adaptively processed low-light visible CCD imagery, Figure 14(b) shows the processed uncooled IR imagery, Figure 14(c) shows the color fused imagery before remapping, and Figure 14(d) shows the fused imagery following a remapping of color designed to render the trees green.

Conclusions

We have described a novel approach to achieve color night-vision capabilities through fusion of complementary low-light visible and thermal infrared imagery. Our approach to image fusion is based on biologically motivated neurocomputational models of visual contrast enhancement, opponent-color contrast, and multisensor fusion [33]. Example imagery illustrates the potential of the approach to exploit wide-dynamic-range visible imagery obtained with new low-light CCD cameras, and to create a natural

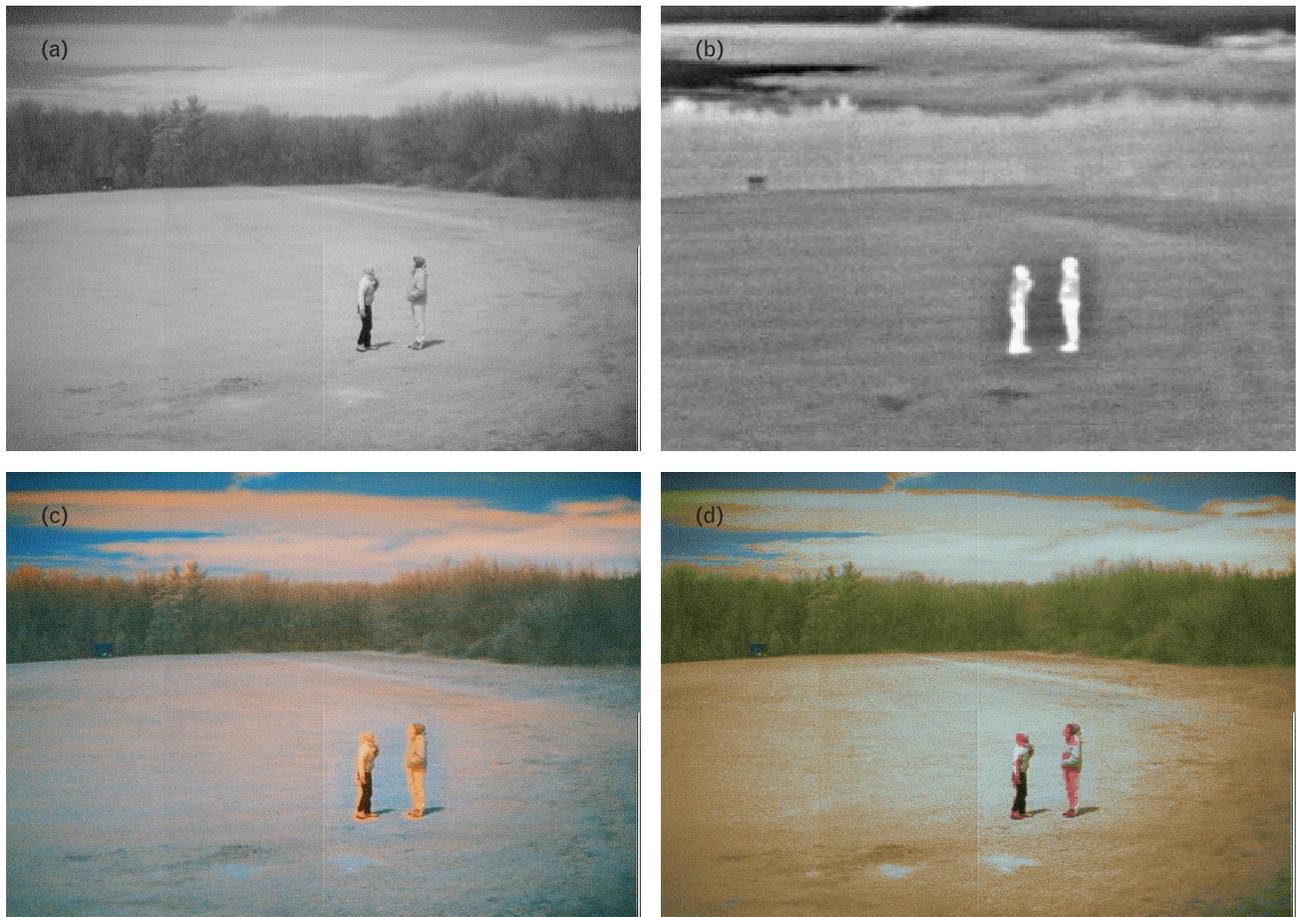


FIGURE 14. Color fused imagery derived from the CCD/LWIR imager pod shown in Figure 2(b), collected under approximately quarter-moon conditions. (a) Adaptively processed low-light visible CCD imagery, (b) processed uncooled IR imagery, (c) color fused imagery before remapping, and (d) fused imagery following a remapping of color to render the trees green.

color scene at night that supports the perceptual pop-out of extended navigation cues and compact targets.

We have conducted psychophysical testing on static imagery to assess the utility of color versus gray visible/infrared fusion in terms of human reaction time, accuracy, and false-alarm rate for detection of embedded low-contrast targets and extended navigation cues (relevant to enhancing situational awareness and tactical efficiency) [34]. Related tests have been carried out on dynamic image sequences of natural visible and infrared night scenes, before and after real-time fusion is carried out. Our most recent dual-band fusion system, constructed around the Lincoln Laboratory low-light CCD camera shown in Figure 12 and a Lockheed-Martin uncooled microbolometer camera, incorporates both image fusion and moving target detection/cueing. It was demonstrated at Fort

Campbell, Kentucky, under starlight illumination conditions (1.5 mLux) in field and water operations with Army Special Forces 5th Group.

We anticipate that solid-state, visible/infrared fusion, color night-vision systems will offer many advantages over existing monochrome night-vision systems in use today. They will play increasingly important roles in both military operations and civilian applications in the air, on the ground, and at sea.

Acknowledgments

This work has been supported by the Defense Advanced Research Projects Agency, the Office of Special Technology, the Air Force Office of Scientific Research, and the Office of Naval Research.

REFERENCES

1. C.G. Bull, "Helmet Mounted Display with Multiple Image Sources," *SPIE* **1695**, 1992, pp. 38–46.
2. A.A. Cameron, "The Development of the Combiner Eyepiece Night Vision Goggle," *SPIE* **1290**, 1990, pp. 16–29.
3. J.S. Crowley, C.E. Rash, and R.L. Stephens, "Visual Illusions and Other Effects with Night Vision Devices," *SPIE* **1695**, 1992, pp. 166–180.
4. J. Rabin and R. Wiley, "Switching from Forward-Looking Infrared to Night-Vision Goggles: Transitory Effects on Visual Resolution," *Aviation, Space, and Environmental Medicine* **65**, Apr. 1994, pp. 327–329.
5. A.M. Waxman, D.A. Fay, A.N. Gove, M.C. Seibert, and J.P. Racamato, "Method and Apparatus for Generating a Synthetic Image by the Fusion of Signals Representative of Different Views of the Same Scene," U.S. Patent No. 5,555,324, 10 Sept 1996; rights assigned to MIT.
6. A.M. Waxman, D.A. Fay, A.N. Gove, M.C. Seibert, J.P. Racamato, J.E. Carrick, and E.D. Savoye, "Color Night Vision: Fusion of Intensified Visible and Thermal IR Imagery," *SPIE* **2463**, 1995, pp. 58–68.
7. A. Toet, L.J. van Ruyven, and J.M. Valetton, "Merging Thermal and Visual Images by a Contrast Pyramid," *Opt. Eng.* **28** (7), 1989, pp. 789–792.
8. A. Toet, "Multiscale Contrast Enhancement with Applications to Image Fusion," *Opt. Eng.* **31** (5), 1992, pp. 1026–1031.
9. D. Ryan and R. Tinkler, "Night Pilotage Assessment of Image Fusion," *SPIE* **2465**, 1995, pp. 50–67.
10. J.M. Wolfe, K.R. Cave, and S.L. Franzel, "Guided Search: An Alternative to the Feature Integration Model of Visual Search," *J. Experimental Psychology: Human Perception and Performance* **15** (3), 1989, pp. 419–433.
11. S. Grossberg, E. Mingolla, and W.D. Ross, "A Neural Theory of Attentive Visual Search: Interactions of Boundary, Surface, Spatial, and Object Representations," *Psychol. Rev.* **101** (3), 1994, pp. 470–489.
12. E.H. Land, "Color Vision and the Natural Image. Part I," *Proc. Natl. Acad. Sci.* **45** (1), 1959, pp. 115–129.
13. E.H. Land, "Experiments in Color Vision," *Sci. Am.* **200**, May 1959, pp. 84–99.
14. E.H. Land, "Recent Advances in Retinex Theory and Some Implications for Cortical Computations: Color Vision and the Natural Image," *Proc. Natl. Acad. Sci. USA* **80**, Aug. 1983, pp. 5163–5169.
15. C.M. Huang, B.E. Burke, B.B. Kosicki, R.W. Mountain, P.J. Daniels, D.C. Harrison, G.A. Lincoln, N. Usiak, M.A. Kaplan, and A.R. Forte, "A New Process for Thinned, Back-Illuminated CCD Imager Devices," *Proc. 1989 Int. Symp. on VLSI Technology, Systems and Applications, Taipei, Taiwan, 17–19 May 1989*, pp. 98–101.
16. R.K. Reich, R.W. Mountain, W.H. McGonagle, J.C.-M. Huang, J.C. Twichell, B.B. Kosicki, and E.D. Savoye, "Integrated Electronic Shutter for Back-Illuminated Charge-Coupled Devices," *IEEE Trans. Electron Devices* **40** (7), 1993, 1231–1237.
17. R.E. Flannery and J.E. Miller, "Status of Uncooled Infrared Imagers," *SPIE* **1689**, 1992, pp. 379–395.
18. P.H. Schiller, "The ON and OFF Channels of the Visual System," *Trends in Neuroscience* **15** (3), 1992, pp. 86–92.
19. P.H. Schiller and N.K. Logothetis, "The Color-Opponent and Broad-Band Channels of the Primate Visual System," *Trends in Neuroscience* **13** (10), 1990, pp. 392–398.
20. P. Gouras, "Color Vision," chap. 31 in *Principles of Neural Science*, 3rd ed., E.R. Kandel, J.H. Schwartz and T.M. Jessell, eds. (Elsevier, New York, 1991), pp. 467–480.
21. A.M. Waxman, M.C. Seibert, A.N. Gove, D.A. Fay, A.M. Bernardon, C. Lazott, W.R. Steele, and R.K. Cunningham, "Neural Processing of Targets in Visible, Multispectral IR and SAR Imagery," *Neural Networks* **8** (7/8), 1995, pp. 1029–1051 (special issue on automatic target recognition; S. Grossberg, H. Hawkins, and A.M. Waxman, eds).
22. A.N. Gove, R.K. Cunningham, and A.M. Waxman, "Opponent-Color Visual Processing Applied to Multispectral Infrared Imagery," *Proc. 1996 Meeting of the IRIS Specialty Group on Passive Sensors 2, Monterey, Calif., 12–14 Mar. 1996*, pp. 247–262.
23. E.A. Newman and P.H. Hartline, "Integration of Visual and Infrared Information in Bimodal Neurons of the Rattlesnake Optic Tectum," *Science* **213** (4508), 1981, pp. 789–791.
24. E.A. Newman and P.H. Hartline, "The Infrared 'Vision' of Snakes," *Sci. Am.* **246** (Mar.), 1982, pp. 116–127.
25. A.J. King, "The Integration of Visual and Auditory Spatial Information in the Brain," in *Higher Order Sensory Processing*, D.M. Guthrie, ed. (Manchester University Press, Manchester, U.K., 1990), pp. 75–113.
26. S. Grossberg, *Neural Networks and Natural Intelligence*, chaps. 1–4 (MIT Press, Cambridge, Mass., 1988), pp. 1–211.
27. S.A. Ellias and S. Grossberg, "Pattern Formation, Contrast Control, and Oscillations in the Short-Term Memory of Shunting On-Center Off-Surround Networks," *Biol. Cybernetics* **20** (2), 1975, pp. 69–98.
28. A.M. Waxman, A.N. Gove, M.C. Seibert, D.A. Fay, J.E. Carrick, J.P. Racamato, E.D. Savoye, B.E. Burke, R.K. Reich, W.H. McGonagle, and D.M. Craig, "Progress on Color Night Vision: Visible/IR Fusion, Perception and Search, and Low-Light CCD Imaging," *SPIE* **2736**, pp. 96–107.
29. A. Toet, J.K. Ijspeert, A.M. Waxman, and M. Aguilar, "Fusion of Visible and Thermal Imagery Improves Situational Awareness," *SPIE* **3088**, 1997, pp. 177–180.
30. A. Toet and J. Walraven, "New False Color Mapping for Image Fusion," *Opt. Eng.* **35** (3), 1996, pp. 650–658.
31. P.M. Steele and P. Perconti, "Part Task Investigation of Multispectral Image Fusion Using Grayscale and Synthetic Color Night-Vision Sensor Imagery for Helicopter Pilotage," *SPIE* **3062**, 1997, pp. 88–100.
32. A.M. Waxman, J.E. Carrick, D.A. Fay, J.P. Racamato, M. Aguilar, and E.D. Savoye, "Electronic Imaging Aids for Night Driving: Low-Light CCD, Thermal IR, and Color Fused Visible/IR," *SPIE* **2902**, 1996, pp. 62–73.
33. A.M. Waxman, A.N. Gove, D.A. Fay, J.P. Racamato, J.E. Carrick, M.C. Seibert, and E.D. Savoye, "Color Night Vision: Opponent Processing in the Fusion of Visible and IR Imagery," *Neural Networks* **10** (1), 1997, pp. 1–6.
34. M. Aguilar, D.A. Fay, W.D. Ross, A.M. Waxman, D.B. Ireland, and J.P. Racamato, "Real-Time Fusion of Low-Light CCD and Uncooled IR Imagery for Color Night Vision," *SPIE* **3364**, 1998, pp. 124–135.



ALLEN M. WAXMAN is a senior staff member in the Machine Intelligence Technology group, where his research focuses on neural networks, multisensor fusion, pattern recognition, and night vision. He also holds a joint appointment as an adjunct associate professor in the Department of Cognitive and Neural Systems at Boston University. He received a B.S. degree in physics from the City College of New York, and a Ph.D. degree in astrophysics from the University of Chicago. Prior to joining Lincoln Laboratory in 1989, he performed research at MIT, the University of Maryland, the Weizmann Institute of Science (Israel), the Royal Institute of Technology (Sweden), and Boston University. In 1992 he was corecipient (with Michael Seibert) of the Outstanding Research Award from the International Neural Network Society for work on 3D object learning and recognition. In 1996 he received the Best Paper Award from the IRIS Passive Sensors Group for work on image fusion and color night vision. He holds three patents and has authored over eighty publications.



MARIO AGUILAR is a staff member in the Machine Intelligence Technology group. His research interests are in data fusion for night vision and data mining. Before joining Lincoln Laboratory in 1986, he developed decision support systems for the stock market. He received a B.S. degree in computer science from Jacksonville State University, and a Ph.D. degree in cognitive and neural systems from Boston University.



DAVID A. FAY is a staff member in the Machine Intelligence Technology group. His recent research has been on real-time processing of imagery for night vision. He received a B.S. degree in computer engineering and an M.A. degree in cognitive and neural systems, both from Boston University. He has been at Lincoln Laboratory since 1989.



DAVID B. IRELAND is an assistant staff member in the Machine Intelligence Technology group. His research work is in the field of color night vision. He has been at Lincoln Laboratory since 1982.



JOSEPH P. RACAMATO, JR. is a staff specialist in the Machine Intelligence Technology group, where his research is in multisensor night-vision imaging and data-collection systems. He has also worked on several projects for other groups at Lincoln Laboratory, including the acoustic detection and signal processing project and the MX parallel computer architecture project. He has been at Lincoln Laboratory since 1981.



WILLIAM D. ROSS is a staff member in the Machine Intelligence Technology group. His recent research has been on image processing, pattern recognition, and multisensor fusion algorithms for both real-time color night vision and interactive 3D site visualization. He received a B.S. degree in electrical engineering from Cornell University and a Ph.D. degree in cognitive and neural systems from Boston University. His postdoctoral research at Boston University and at the University of North Carolina focused on developing and applying neurocomputational models of biological vision. He has been at Lincoln Laboratory since January 1998.



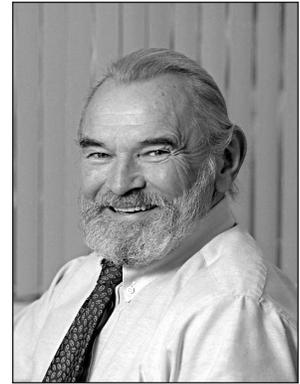
JAMES E. CARRICK is a real-time software development engineer at The MathWorks, Inc., in Natick, Massachusetts, and a former assistant staff member in the Machine Intelligence Technology group. His current work involves the generation of production-quality code from high-level block diagrams of control systems. He received a B.S. degree in electrical engineering from the University of Wisconsin at Madison.



ALAN N. GOVE is a former staff member in the Machine Intelligence Technology group. He received a B.S. degree in computer science and a B.A. degree in mathematics from Brown University, an M.S. degree in computer science from the University of Texas at Austin, and a Ph.D. degree in cognitive and neural systems from Boston University. He is currently employed as a software manager at Techtrix International Corporation in South Africa, where he is working on camera-based forest-fire detection.



MICHAEL C. SEIBERT is a staff member at the Lincoln Laboratory KMR Field Site, Kwajalein, and a former staff member in the Machine Intelligence Technology group. His research interests are in vision and neural networks, and in 1992 he was the corecipient (with Allen Waxman) of the Outstanding Research Award from the International Neural Network Society. He received a B.S. degree and an M.S. degree in computer and systems engineering from the Rensselaer Polytechnic Institute, and a Ph.D. degree in computer engineering from Boston University.



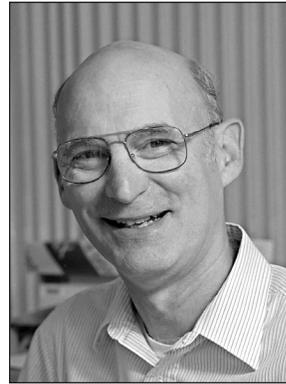
EUGENE D. SAVOYE is the former leader of the Microelectronics group. He received a Ph.D. degree from the University of Minnesota, and joined RCA's David Sarnoff Research Laboratories in 1966. In 1970 he became Manager of Advanced Technology at RCA in Lancaster, Pennsylvania, where he initiated RCA's early engineering work on charge coupled devices (CCDs). In 1983, as Director of CCD and Silicon Target Technology, he initiated and directed RCA's corporate program to develop high-performance CCD imagers for television, which resulted in the world's first all-solid-state studio-quality TV cameras. For this work he received the 1985 Emmy Award for Outstanding Technical Achievement. He joined Lincoln Laboratory in 1987, and in 1990 he became leader of the Microelectronics group, with responsibility for the development of advanced silicon imaging devices, including CCD imagers aimed at demanding applications in astronomy, surveillance, and advanced night-vision systems. Dr. Savoye retired from Lincoln Laboratory in October 1997, and is currently a consultant in the area of sensors for electronic imaging.



ROBERT K. REICH is the assistant leader of the Microelectronics group. His area of research is in the design of high-frame-rate and low-noise optical detector arrays. He received a B.S. degree from Illinois Institute of Technology, and M.S. and Ph.D. degrees from Colorado State University, all in electrical engineering. He has been at Lincoln Laboratory since 1987, and he is a senior member of the IEEE.



BARRY E. BURKE is a senior staff member in the Microelectronics group. His research interests are in the area of CCD imagers. He received a B.S. degree in physics from the University of Notre Dame, and a Ph.D. degree in applied physics from Stanford University. He has been at Lincoln Laboratory since 1969, and he is a senior member of the IEEE.



WILLIAM H. MCGONAGLE is an associate staff member in the Microelectronics group. His research interests are in analog and digital circuit design relating to CCD devices. He received a B.S.E.E. degree from Northeastern University, and he has been at Lincoln Laboratory since 1960.



DAVID M. CRAIG is an assistant staff member in the Submicrometer Technology group. He joined Lincoln Laboratory in 1980 after graduating from Quincy College, in Quincy, Massachusetts. His research efforts have been in the development of software and instrumentation for a 193-nm integrated-circuit lithography system. He is currently working on CCD camera development for the Microelectronics group, concentrating primarily on imaging software and electronics hardware and firmware.