A Low Cost Gun Launched Seeker Concept Design for Naval Fire Support

Donald E. Maurer, Eric W. Rogala, Isaac N. Bankman, Bradley G. Boone, Kathryn K. Vogel, Christopher Parris

The Low Cost Gun Launched Seeker Program is a component of the Navy's effort to develop effective weapons for surface fire support missions. The objective is to improve the performance of precision-guided munitions against hardened targets by developing a low-cost terminal seeker, including detection and target selection algorithms, to autonomously acquire a target and provide commands to guide the projectile. This technology will have widespread application to the time-critical engagement of both ground- and sea-based threats by providing a smart-weapons capability under the control of the field commander. This article provides an overview of the required system capabilities, discusses the design requirements imposed on such a system by these capabilities, and summarizes APL's effort to develop a system that meets mission objectives.

INTRODUCTION

This article describes a proposed target detection system, consisting of an infrared (IR) seeker with target detection, classification, and selection algorithms, designed to meet the mission objectives of the Low Cost Gun Launched Seeker (LCGLS) Program. This system is intended to improve the performance of projectiles like the Extended Range Guided Munitions (ERGM) against hardened targets by developing a low-cost, uncooled IR staring focal plane array (FPA) terminal seeker. The seeker will autonomously acquire a target and provide commands to guide the projectile. A ground-based forward observer, air controller, or airborne asset locates and classifies a target and issues a "call for fire." Target location and target templates are loaded into the LCGLS and the projectile is launched.

The seeker must survive the initial 15,000-g launch acceleration. It navigates to the target using the Global Positioning System-Aided Inertial Navigation System. At approximately 5 km from the threat, the seeker enters search mode. It acquires the target at approximately 2–3 km, begins tracking, and provides commands to guide the projectile. Biased proportional navigation begins to create a lofted trajectory and steep approach at a terminal velocity of approximately 150 m/s. At the optimum point, an inexpensive standoff fuze initiates the warhead.

The seeker has been proposed to track fixed, relocatable, and mobile land and sea targets with sufficient accuracy to achieve a small circular error probability (CEP). If the seeker is activated below the cloud level,

nominally at 300 to 900 m based on cloud-free line-of-sight statistics, it has about 2–6 s to find the target and, if necessary, complete a single divert. If the cloud ceiling is higher, the seeker could be activated earlier and more time would be available to find the target. Thus, some of the major constraints include the ability to survive a high-*g* launch, limited space and weight for the seeker, environmental factors affecting seeker performance, and terminal engagement geometry, which determines the seeker field of view (FOV) and the time available for target detection. These constraints limit the feasible optical designs and the complexity and computational requirements of the detection algorithms. An additional programmatic constraint is cost (preferably less than $10,000 per round for the seeker).

The technology will have widespread application to the time-critical engagement of both ground- and sea-based threats by providing a smart-weapons capability under the control of the field commander. Many of these scenarios share the same compressed time line in the terminal engagement process, which requires substantial real-time image processing functionality supporting highly responsive terminal guidance, target separation from clutter, and aimpoint selection. The ability to accurately deliver munitions in support of precision attack using tactical targeting sources will give the commander greater flexibility and autonomy, thus reducing the time required to respond in a dynamic threat situation.
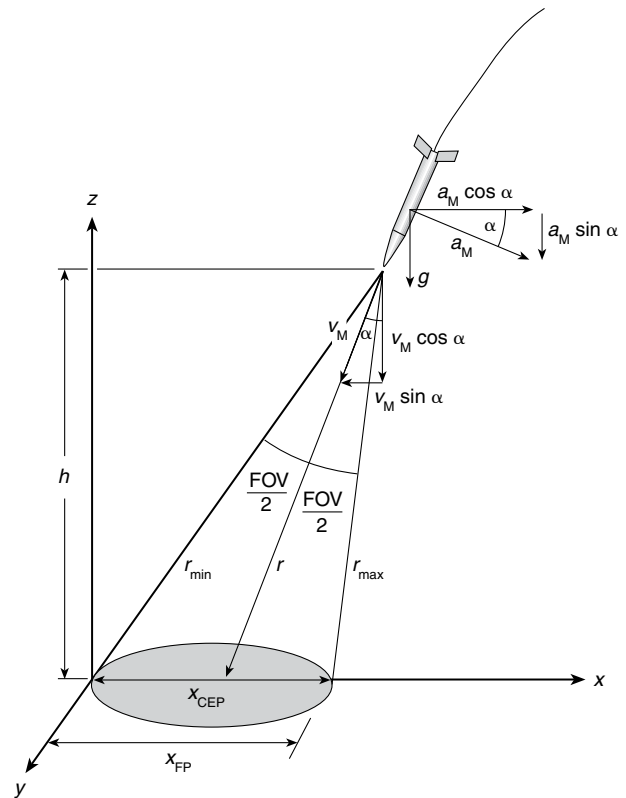
Although originally designed as an enhancement of ERGM, the LCGLS technology is applicable to a wide range of weapons programs where a sensor or seeker system is needed. Examples include the Rapid Ordnance Defense System, Joint Direct Attack Munitions, Forward Air Support-Marine-Gun Launched Loitering Munition, as well as Affordable Weapon, Precision Guided Mortar, and other smart submunitions concepts. Unmanned aerial vehicles, precision targeting, battle damage assessment, and identification friend or foe (IFF) are also potential application areas.

The focus of this article is on the algorithms being developed for the system. However, to provide context, the desired functional capabilities of the LCGLS are summarized as well as the project's requirements. Earlier work focused on conceptual designs and preliminary analyses to determine if they met requirements. In view of the promising outcome of these efforts, we began to develop a set of algorithms that will autonomously detect and track the desired targets. Ground targets pose especially challenging automatic target recognition problems and continue to be the focus of our work. A preliminary study of the algorithm execution time is discussed here, followed by a demonstration of the algorithm's typical performance for a variety of ground target images, including buildings, bunkers, and airfields.

FUNCTIONAL CAPABILITIES

In the conceptual design we considered a broad mission encompassing increasingly challenging target scenarios. Subsequently we focused on the two most critical functional capabilities: target segmentation and terminal aimpoint selection, i.e., the acquisition of hard targets against complex natural backgrounds. Additional capability against moving targets, although not necessarily more challenging, will likely involve in-flight targeting updates, which require a data link. More challenging scenarios include selecting high-priority targets in a clutter background, especially near the forward edge of the battle; controlling salvo distributions for maximizing lethality over large areas; and avoiding countermeasures and jamming.

The FOV needed to support these capabilities is determined by the terminal engagement geometry illustrated in Fig. 1. The diameter of the footprint that must be contained in the FOV at initial acquisition by the seeker consists of a combination of 10 m for the ERGM baseline terminal guidance accuracy for hand-over and 10 m for targeting (root sum squared). Ordinarily this diameter corresponds to a 50% target containment requirement. However, we use a 90% containment probability corresponding to 26 m in order to
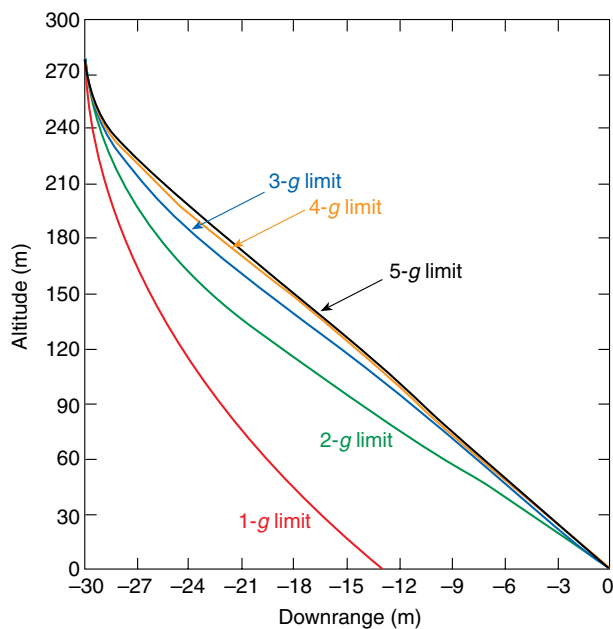

Figure 1. Terminal engagement scenario (a_M = acceleration perpendicular to the projectile boresight, v_M = terminal velocity, x_{FP} = diameter of the FOV footprint, x_{CEP} = circular error probability, and α = seeker angle).

establish a conservative FOV. For simplicity, assume the seeker is pointed straight down ($\alpha = 0$). Then it can be shown that the FOV must be greater than 10° for acquisition ranges as low as 300 m. Of course, if target acquisition and track occurs at a shorter range, the FOV must be correspondingly larger. Therefore, we nominally assume a minimal 14° FOV. This should allow target containment in the 90% CEP down to a range of approximately 275 m.

A major consideration is the airframe on which the seeker must operate. ERGM, like most guided projectiles, has a relatively low divert capability. The airframe has been designed to remove small heading errors throughout flight as it is guided to a predetermined point. The ERGM divert capability is a function of vehicle speed but is typically about ±2 g. A study was conducted early in the program to determine whether ERGM or a similar platform could provide enough divert capability during the terminal phase to remove the IR-measured guidance errors given a few seconds of homing time. The results of the study were used to assess terminal guidance policy options (terminal velocity, acquisition range, approach angle, and hit-to-kill versus shoot-to-kill).

A two-degree-of-freedom (2-DOF) simulation, containing a representation of the ERGM kinematics, was developed to parametrically study the terminal homing phase. This simulation was used to conduct a trade study of intercept geometry and seeker look angle as a function of projectile divert capability, projectile velocity, and position and attitude errors at acquisition. A sample output in Fig. 2 shows miss distance as a function of

the g capability of the airframe. In this case the initial position error (delivery plus targeting) was 30 m, initial heading error was zero, and the projectile velocity was 150 m/s. The 2-DOF results showed that a ±2 g airframe is marginal for a hit-to-kill scenario if the target is acquired at approximately 300 m (such as if the cloud ceiling is low). Better performance is achieved for earlier acquisition and hence longer homing times; thus, the target must be detected and classified as soon after seeker activation as possible in order to optimize the limited divert capability. This, in turn, limits the available processing time.

SEEKER DESIGN

Initial assumptions made to minimize cost, complexity, and form-factor include body mounting the seeker and optics; using lightweight optical designs; z-plane chip packaging; and uncooled FPA technology to reduce seeker form-factor and obviate cryogenic cooling. The use of body-mounted optics (compared with gimbaled optics) simplifies the design, reduces cost and weight, and increases survivability during launch. Z-plane chip packaging is a state-of-the-art technique that reduces FPA and post-processor form-factors. It consists of stacked thinned-silicon integrated circuits coupled to the IR FPA. Northrop Grumman has fabricated such a chip under the DARPA-sponsored Flexible Manufacturing Program.[1] An uncooled bolometric FPA does not require cryo-cooling, which is costly and takes up space (although the uncooled FPA may require a thermoelectric cooler for temperature stabilization). Nonuniformity compensation may be required, however, to improve target selection and clutter rejection.

The Detector

Although the experience of processing measured IR image data may ultimately help determine the exact spectral band, long-wave IR (LWIR) is preferred for a number of reasons. For instance, LWIR attenuation is generally less than mid-wave IR (MWIR) at short ranges (about 2.5 km), although, for ranges less than 1 km, the sensor should not be significantly sensitivity-limited in either band under most atmospheric conditions. Although most energy is in the LWIR band for expected targets near equilibrium, differential clutter contrast is generally greater in the MWIR band. Furthermore, even though MWIR has better diffraction-limited spatial resolution, a well-designed seeker modulation transfer function should not be dominated by the optics, but by detector size. Finally, and most importantly, uncooled array technology is currently available only for LWIR.

Target-background contrast, ΔT, in the LWIR at the FPA yields an output signal-to-noise ratio (SNR) given by

Figure 2. A 2-DOF simulation showing miss distance as a function of g capability of the airframe.

$$\text{SNR} \cong \frac{\tau_0 \tau_{\text{atm}} \Delta T}{\text{NETD}},$$

where τ_0 is the mean optics transmissivity over the prescribed spectral band, τ_{atm} is the mean atmospheric transmissivity over the same band, and NETD is the noise equivalent temperature difference.

The mean ΔT, based on the contrast temperature range for a typical military vehicle (T-72 tank) against a grassy background, over various azimuths, is about 1.0 K for measured data and 1.5 K for modeled data (personal communication, S. E. Motz, U.S. Army Missile Command, 24 Apr 1998). Significant factors for determining the actual scene contrast (and hence the local scene contrast SNR), however, are diurnal variations in ΔT, which can sometimes be zero, principally as a result of the combined effect of differential solar loading and the constituent materials, their orientation, and their associated emissivity. This effect, called thermal crossover, is particularly noticeable in "pillbox"-type building structures.

The extinction coefficients, β, for LWIR versus MWIR for various atmospheric obscurants (including gases, haze, fog, rain, snow, and dust) are not significantly different overall. At an altitude of 900 m (the cloud-free line of sight about 90% of the time, worldwide) with $\beta = 0.7$ km^{-1}, corresponding to weather that is "poor" or better 80% of the time, and NETD = 0.1 K, a typical uncooled bolometer FPA[2] has a SNR variation between 8 and 12. The larger the entrance aperture of the seeker, the better the SNR. The entrance aperture size, however, will be limited by the airframe to approximately 10 cm.

Now, to ensure optimal resolution for identification of a vehicle-sized target at the greatest expected acquisition range, we assume that 8 pixels are required for target identification,[3] and that the minimum target dimension is 5.18 m ($\approx$10% CEP). Then the number of pixels required in the FOV ranges between 85 for a vertical descent to 176 for a relatively extreme deviation from vertical. Therefore, a 256×256 array should be sufficient.

Optics

The most stringent design requirements include the ability to survive a high-g shock at launch; minimal size, weight, and volume; and low cost. The favored design, a $f/1.54$ system with a FOV of 22°, a clear aperture of 20 mm, and 79.9% ensquared energy on the center 50-μm pixel, has the largest FOV and the most compact size of the candidate designs studied. The length of the system is 49.8 mm and it can fit in a cylinder of volume 35.2 cm^3.

Requiring both LWIR operation and shock survivability severely limits the choice of viable lens materials. First, the material must be adequately transparent throughout the entire LWIR spectral band. Of the few materials that are chemically stable enough to withstand the intended storage and use environment, only those that can survive the mechanical stress of launch shock are acceptable. A preliminary structural analysis using various materials, including polyethylene, zinc sulfide, zinc selenide, and germanium, determined that all these should survive. It is important to note, however, that the lenses were not designed to image during the shock, and that, in the analyses, *the mechanical stress was not applied as an impulse*. Neither did the analyses account for material or surface characteristics; these could substantially lower the allowable stress.

All the materials being investigated would require a thin-film antireflective coating on both surfaces to increase transmission to an acceptable level and to provide for a specific wavelength passband optimized to the required LWIR camera response. For the front optical element, this coating must survive the stresses of being exposed to a variety of atmospheric conditions at relatively high speeds. Also, if there is excessive aerodynamic heating of the front optic, the impact on both the optical and mechanical integrity of the lens has to be taken into account. For instance, as germanium is heated to greater than 50°C, the transmission through the bulk material drops, and the LWIR emissivity increases. When germanium reaches 250°C, it is essentially opaque in the LWIR and thus no longer a viable lens material. We believe, however, that if optics are not exposed until the terminal phase of trajectory, they will not be subjected to atmospheric flow at subsonic speeds for more than a few seconds, and hence these considerations were not a major concern.

DETECTION ALGORITHMS

The analyses summarized in the preceding sections provide a positive initial assessment of the feasibility of fielding a gun-launched imaging IR system. The next critical step is to demonstrate algorithms that accurately detect desired targets with a high probability of success. Limited processing time is, perhaps, the most significant constraint. The projectile approaches impact at a high rate of speed; the time between seeker activation and impact is only a few seconds. In order to have sufficient divert capability, the target must be selected as soon as possible. Moreover, the frame rate is on the order of 60 Hz to avoid scale change during tracking; therefore, if the frames are processed much slower than real time, it will be difficult to correlate an object found in one image with the same object several frames later. Since there is no uplink, the algorithms must be capable of autonomous operation. Such a system cannot rely on detailed targeting information to guide its search; thus, standard techniques requiring terrain mapping or detailed target libraries are not feasible.

The fundamental philosophy guiding our algorithm design is to use different processing modalities operating in parallel to provide detection capability under a greater range of image conditions than would otherwise be possible. Figure 3 shows the baseline detection architecture implementing this approach. The proposed system includes two complementary parallel processes consisting of region-based and edge-based detection and segmentation. By fusing the outputs of these detection algorithms, we hope to enhance robustness and attain greater discriminatory power. The fused output will reduce false alarms and provide regions of interest (ROIs) for further analysis. A set of features for analyzing these ROIs will evolve as the algorithms develop. Ideally, the only ROIs passed to the classifier would be those containing target objects. Since the seeker has a restricted FOV—it is boresighted and the projectile has limited maneuverability—it cannot search an extended footprint. If it does not detect its intended target, the seeker must default to any man-made object it can find. Our first objective, therefore, is to demonstrate the capability of this system to differentiate between man-made objects and natural features. If this is successful, more refined classification algorithms and techniques for detecting moving targets will be developed.
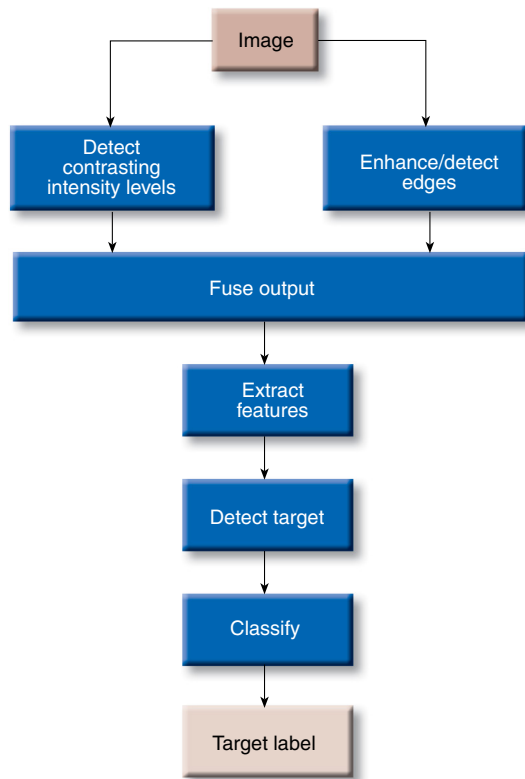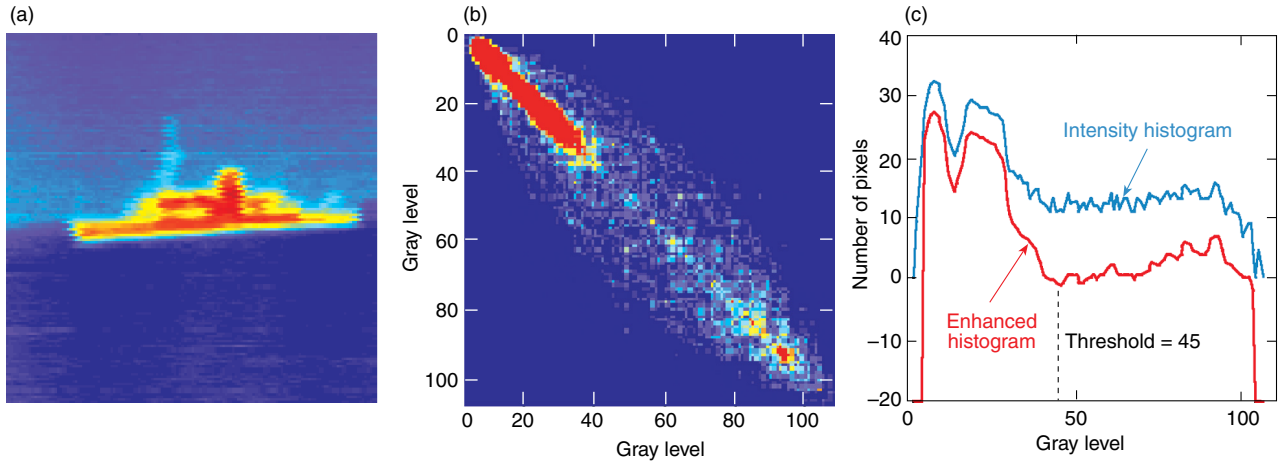
Region-Based Processing

Histogram Enhancement

The region-based detector uses intensity to segment the image. Ideally the histogram of an image consists of two modes, one corresponding to the background and the other to the target. Thus a threshold, chosen to be the gray level between the modes at which the histogram is minimal, can be used to partition the images into target and background classes. However, the histogram may not exhibit this simple form; therefore histogram enhancement techniques have been developed to accentuate the desirable features and thus extend the range of images to which this technique can be applied. Rounds and Sutty[4] describe a histogram enhancement technique based on co-occurrence matrices. In general, a co-occurrence matrix measures the frequency with which pairs of nearby pixels have similar intensity. An example is the matrix whose $(m, n)^{th}$ entry is the number of pixel pairs (i, j) and $(i-1, j+1)$ for which the first pixel has gray level m and the second has gray level n. Similar directional co-occurrence matrices are obtained for pairs consisting of a pixel and any one of its neighbors. Since pixel pairs within a target region tend to have similar intensity, these regions contribute to larger entries near the main diagonal, and the co-occurrence matrix has a more prominent modal structure along the main diagonal than the ordinary histogram. This is illustrated in Figs. 4a and 4b.

The algorithm then calculates an *enhanced* histogram by averaging the co-occurrence matrix entries within a narrow band containing the main diagonal. To find the threshold between target and background, the two modes are first detected and then the threshold can be chosen to be the intensity value between the two peaks at which the histogram assumes its minimal value.

The simplest peak detector compares the histogram value $H(u)$ at a test cell u with its values at adjacent cells and declares a peak if $H(u-1) < H(u)$ and $H(u+1) < H(u)$. If the histogram is rough, however, this method can produce many false alarms. Therefore, normally, a window is centered at the test cell u, and $H(u)$ is compared with the average value of the histogram over this window. The length of the window can be adapted to the data at hand. For peak detection in the LCGLS application, this technique is extended by centering a window of length $2N+1$ at the test cell and computing two histogram averages m and M (assume $m \leq M$), one for each half-window on either side of the test cell. The peak detection threshold is $T = RM + (1-R)m$, where R is a non-negative parameter that determines how much larger $H(u)$ must be than the adjacent histogram values in order to be classified as a peak. This reduces to simple window averaging if $R = 0.5$. When the test cell is located at the center of

Figure 3. Baseline target detection architecture.

Figure 4. Calculating the enhanced histogram: (a) the original image, (b) matrix obtained by averaging the directional co-occurrence matrices, and (c) the enhanced histogram compared with the standard intensity histogram.

a relatively symmetric peak, $M \approx m$, and the response is, again, essentially equivalent to window averaging. A further smoothing process removes smaller peaks that may otherwise cause false alarms. The improvement provided by this technique is evident when comparing the enhanced histogram (red curve) with the standard intensity histogram (blue curve) shown in Fig. 4c.
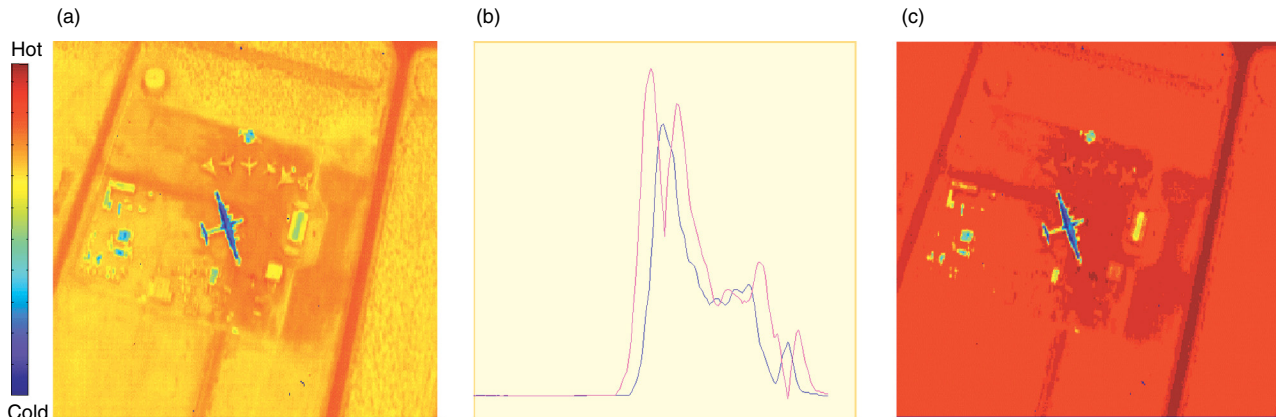
Quantization

Analyses of histograms obtained from ground target imagery collected at the Naval Air Warfare Center show that, in general, useful information is contained in all significant peaks. Therefore, the technique described above has been generalized to a multithreshold approach by defining a set of intensity levels $T_0, T_1, \ldots, T_n$ corresponding to the location of peaks exceeding the detector threshold. The raw image is quantized by replacing the value of each pixel whose intensity lies in the half-open interval $[T_i, T_{i+1})$ by i. The quantized image retains the salient target features, but there are only about a dozen intensity levels. This represents a significant data reduction and, hence, faster processing. Figure 5 is a typical example. Figure 5a shows a raw image (color-enhanced) of aircraft of various sizes, buildings, and roads in the desert. Figure 5b is a plot of the enhanced histogram (blue curve) showing several significant peaks. The magenta curve is the threshold, determined as described above. There are at least four peaks exceeding the threshold (the scale of the figure is too gross to show details of the structure on the far left of the histogram). The quantized image is shown in Fig. 5c.

Filtering

Although quantization reduces the dynamic range, objects may still be composed of several intensity layers. An attribute signature comprising the geometric and intensity characteristics of these layers provides a powerful feature that can be used to filter the quantized

Figure 5. Quantization: (a) raw image and (b) enhanced histogram (blue curve is the co-histogram obtained from the raw image data; magenta curve is the threshold). Each peak exceeding the threshold corresponds to a gray level in the quantized image (c).

image to remove nontarget regions such as vegetation and terrain. The attribute signature, for example, may specify acceptable ranges in size, shape, and intensity; and it could also include important information linking these features. A compact data structure called a component tree[5] has been developed to implement filtering based on attribute signatures (see the boxed insert).

Filtering an image reduces to a decision process on the associated component tree that determines which nodes are removed. This process provides a degree of robustness since the test criteria can be fuzzy or expressed in linguistic rather than quantitative terms. The only restriction is imposed by computational limitations. If the processing is applied to a raw image with 256 gray levels, for example, the component tree is quite complex. If, however, the image is first quantized with the multithreshold algorithm, the number of gray levels is reduced to a dozen or so, greatly simplifying the component tree; consequently, more complex filter criteria can be implemented without degrading the filtering speed. Research is continuing to determine optimal features for defining filter criteria and how best to perform the filtering and segmentation step to produce as robust an algorithm as possible within the LCGLS constraints.

Edge-Based Processing

To provide greater segmentation capability, an algorithm that extracts edge information has been implemented to work in parallel with, and complement, the region-processing algorithm. This edge-based technique

---

FILTERING COMPONENT TREES

A connected component of an image is a connected region, all of whose pixels are greater than or equal to a given intensity level. Two components of different levels are linked if they are connected to each other and there is no component of an intermediate level connected to both. The component tree derived from an image consists of nodes corresponding to each connected component, and two nodes are connected by an edge if the corresponding components are linked. The leaves correspond to regional maxima, i.e., connected regions, all of whose pixels have the same intensity and such that all surrounding pixels have lower intensity. The root node corresponds to the entire image. These definitions are illustrated in the figure. The height of the component tree depends on the number of gray levels, and the branching structure depends on the definition of connectivity.

Standard definitions include four- or eight-connectivity, although the concept has been generalized in the image processing literature.[6] Two pixels are four-connected if and only if they have a common edge, and are eight-connected if they have a common edge or vertex. Since four-connectivity is the more stringent condition, the four-connected component tree of a given image is usually more complex than the eight-connected component tree of the same image. Each node has an associated gray level of its corresponding component, the location of a representative pixel, and a set of attributes characterizing the component. The representative pixel locates the component in the image, and the gray level value is a simple feature describing the component. In general, the node description may consist of a vector $\boldsymbol{f} = (f_1, \ldots, f_n)$ of attributes, or features, that capture salient information about the associated component. Besides area, typical features such as perimeter, compactness, aspect ratio, and convexity may be useful.

The component tree supports a class of nonlinear filters, called connected filters, that preserve only those components satisfying a given criterion based on the component's attributes. The criterion may, for example, require these attributes to lie within specified limits. A connected filter never introduces new components to the image, and if an image component is preserved by one application of the filter it is preserved by any number of repeated applications. In mathematical morphology, a filter satisfying these properties is called a thinning.

If, in addition, the filter is increasing (i.e., whenever a component C satisfies the criterion, so does any superset), the filter is an opening. Connected component filters either preserve a component without alteration or remove it entirely, so there is no distortion as occurs with openings and closings using structure elements[7]—a desirable property for segmentation.

Component tree filtering is a decision process that classifies nodes as active or inactive. A node is active if and only if the filter preserves its associated component. A connected filter that ignores the links is called flat and can be implemented by classifying each node as active or inactive. If the filter criterion is increasing, every node above an active node is also active; in this case, therefore, the filter can be implemented by starting at the root and moving up the tree, branch by branch. Once an inactive node has been found, it is unnecessary to check higher nodes. This increasing property facilitates rapid implementation of this type of filter. An example of a flat filter criterion is the following: *The area of the node is greater than 12.* A second type of connected filter, called a signature filter, applies to an entire branch. The attribute signature of a branch is a sequence of node attributes. If the attribute signature satisfies a specified criterion, the branch is declared active. For example, the following is a signature filter criterion: *In order to be active, a branch must contain a node whose corresponding connected component has area between 10 and 40 pixels and eccentricity between 0.4 and 0.6.*

Segmenting an image consists of subtracting the filtered image from the original. Thus, the filter criterion must be chosen so as to inactivate nodes associated with target image components. One way to implement this process is to first apply a signature filter to activate all branches containing no target-like components. For instance, a branch would be activated if it contains no nodes whose connected component is convex and has an area within expected target bounds. If the objective is to locate the moving tank shown in the figure, any branch whose gray-level structure does not indicate a hot engine would also be activated. This may, of course, require some general *a priori* knowledge, such as target type and an estimate of the range to go.
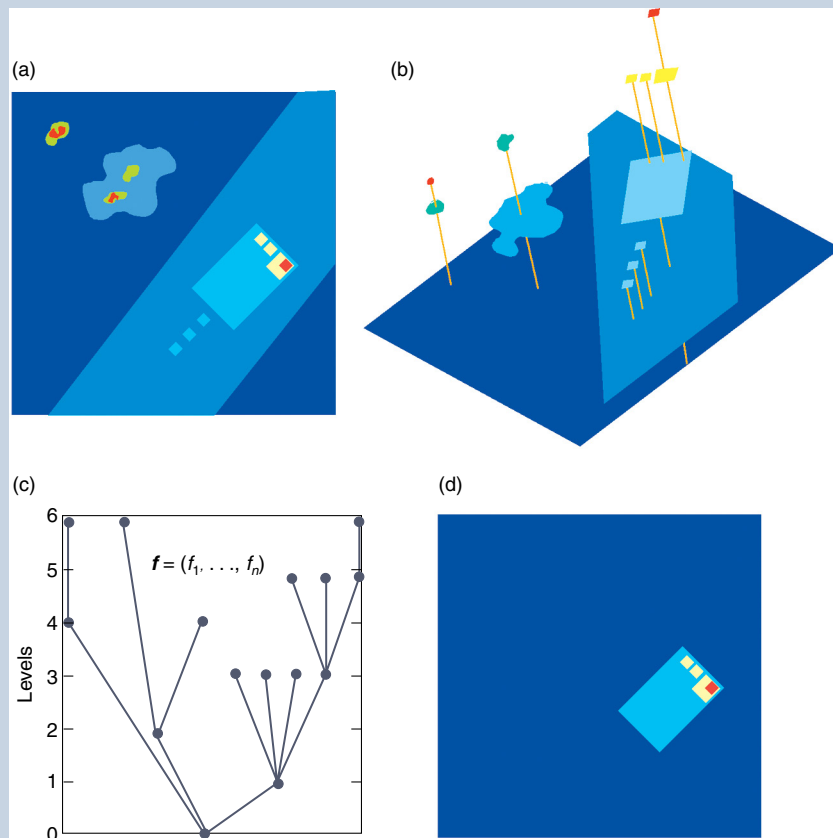
Next, a flat filter can be applied to determine which nodes in the remaining branches to activate. In the figure,

---

can detect a variety of potential targets in diverse backgrounds. The algorithm requires no user intervention or input during its execution and does not depend on a library of predefined targets. It is designed to identify man-made objects, structures, and vehicles in a fully automated manner using a single frame of passive IR imagery.

To achieve this objective, the algorithm is composed of nine steps or subalgorithms, where the output of one serves as input to the next. The process starts with a noise-reduction filter (Step 1) and then Step 2 enhances the edges for the detection process (Step 3). The adaptive edge segmentation procedure (Step 4) binarizes the edge detector result, while also performing a size discrimination process (Step 5) where objects that are too small or too large to be the target are

filtered out. In Step 6, the remaining objects are thinned to produce a contour of 1 pixel width. A chain code algorithm (Step 7) then strings contour pixels together for analysis. In Step 8, the shape discrimination algorithm identifies objects by desired shape attributes. Finally, Step 9 applies a target likelihood index to rank the potential targets according to their shape attributes. A more detailed discussion of the edge-based processing steps follows.

The noise reduction algorithm (Step 1) is a simple median filter that reduces random noise and eliminates dead pixels. The advantage of the median filter over many other noise removal filters is that it preserves the edge contrast while eliminating noise considerably. The edge enhancement algorithm (Step 2) utilizes gradient and variance information. At each pixel the enhanced

Part (a) is a schematic quantized image of a tank with a hot engine. The image in (b) shows the connected components. Each color corresponds to an intensity level from cold (deep blue) to hot (red). The leaves in the component tree (c) correspond to the regional maxima; each leaf determines a branch terminating at the root node (the image plane shown in dark blue). Part (d) shows the result of filtering the image as discussed in the text.

for example, the tank is on a road that is clearly not part of the tank. All nodes up to and including the one corresponding to the road can be activated by filtering with the criterion that activates nodes whose area is greater than the expected area of the tank. In the figure, the gun barrel is composed of several regional maxima, which, although part of the tank, would be inactivated by this filtering scheme.

value is computed using the local intensity gradient and local variance in horizontal and vertical directions, producing two images that highlight edges in two orthogonal directions.

Step 3, edge detection, processes the edge-enhanced image and yields a binary image where only significant edge pixels are retained. This step first removes pixels where the intensity variations are negligible, such as those associated with background noise. Then the edge detection algorithm identifies local maximum pixels on the edges of each image by comparing each pixel to its immediate neighborhood and saving only those pixels whose values are larger than their neighbors. The final result is a binary image, B, obtained as the logical OR of the horizontal and vertical direction results, where edge pixels are set to the value 1.

While the first three steps identify most edge pixels, they do not guarantee a closed contour around each object. Therefore, Step 4 is an adaptive edge-filling algorithm that uses both the binary image of edge detection and the original noise-filtered image from the first step. The edge discontinuities in the binary edge image are filled by inspecting each pixel with a value of 0 and converting it to a 1 if it fulfills desired neighborhood criteria. The final result is an edge-thickened version of B in which the contours are connected. A labeling algorithm assigns a unique value to all pixels that form the contour of an object. This step provides a means to distinguish distinct objects that will subsequently be analyzed separately to determine if they are likely to be targets. Step 5, size discrimination, is implemented using the number of pixels in each contour. At this step we assume some *a priori* knowledge of the target type and that an estimate of the range to go is available (e.g., from GPS or passive ranging). Small objects that have fewer pixels than a first threshold and very large objects that have more pixels than a second threshold are removed. In this manner many background objects are discarded before feature extraction.

The remaining objects are thinned in Step 6 using a standard thinning algorithm that retains the medial axis of the contours by setting most pixels to zero depending on neighborhood criteria. The final result is a set of contours, 1 pixel in width. To analyze the shape of the resulting contours, each one is represented with a one-dimensional array using the chain code algorithm (Step 7). This encodes the transition between adjacent pixels of the contour with a unique code for each of the eight possible directions.

The shape discrimination algorithm (Step 8) removes any remaining nontarget objects using the chain code information. The principal discriminator in our initial study was straight edges. Man-made objects such as tanks, bunkers, bridges, and planes are characterized by straight edges, whereas most natural objects tend to have a more arbitrary appearance. Furthermore, man-made objects tend to have adjacent straight edges. The chain codes were analyzed to find sequences where the directional value remained constant over a predetermined number of pixels. This analysis eliminated background objects that may have contained a few pixels in a line. Moreover, since we cannot expect perfectly straight edges, we allow for some variation from a straight line. Typically, a sequence of pixels having the same direction, except for one or two values, is assumed to be a straight edge. The algorithm also identifies linear sequences that adjoin one another to form straight edges. This information is used to calculate the target likelihood index (Step 9). The index is higher when the candidate contour contains more adjoining or connected straight edges. Potential targets are ranked with respect to this index.

ALGORITHM TIMING STUDIES

The ability to implement these algorithms is as important as developing them. To accomplish this goal, two different—yet complementary—technologies are being investigated: microprocessors and field-programmable gate arrays (FPGA). These are fundamentally different devices that require different programming paradigms to get good results. The microprocessor has a fixed architecture (or configuration) that performs a set of operations as determined by the manufacturer (e.g., Intel, Motorola) during the design process. Modern designs use super-scalar architectures and instruction pipelines to increase performance by taking advantage of fine-grain parallelism in a short sequence of operations. The FPGA has an architecture that is determined at programming time by the end user. This process allows pipeline lengths and computation units to be determined by the algorithms being implemented.

To determine which technology is truly best for a particular algorithm, the algorithm must be implemented and its execution times compared. This is not always practical, however. A rule of thumb is that operations that do not lend themselves to pipeline operations or that are not data-independent tend to work better in a microprocessor. The reason is the disparate clock speeds between FPGA and microprocessors. FPGA manufacturers are now boasting clock speeds as high as 300 MHz. However, in practice the best speeds achieved are about half the maximum of manufacturer claims. The FPGA functions well with algorithms that can be decomposed into many small independent computations because of the configurable architecture.

The two main challenges in implementing the algorithms so far have been avoiding repeated memory scans and adapting some nonlinear operations to vector-style processing. Because of the size of the

images being processed, the raw image data and temporary memory areas do not stay in the L1 or L2 cache of a microprocessor. Consequently, repeated data scans can require relatively long times because of the time required to access main memory. Also, pieces of the algorithms with strong data interdependence cannot be made to take advantage of special properties of an FPGA or modern microprocessors.

The microprocessor currently being used is a Motorola PowerPC 7400 with the AltiVec™ processing unit (a special vector processor added by Motorola to accelerate digital signal processor [DSP] and image processing applications). The PowerPC 7400 is on a Power Macintosh G4 AGP motherboard with a 1-Mb half-speed L2 cache and runs at 400 MHz. Early benchmarks showed that the PowerPC 7400 had more promise of being adequate than other microprocessors and DSPs. In fact, it was 4 to 5 times faster without using the AltiVec unit than the DSP initially being used. The AltiVec unit has increased the processing speeds of several algorithms by a factor of 2 or more—even before optimization.

As can be seen in Fig. 6, the algorithms that have been benchmarked execute in very short periods of time, with the exception of the multiblob. The most important part of reducing execution time has been smart implementation. Algorithms need to be structured so that repeated memory accesses are not necessary and so that computations can be easily pipelined. To date, however, the timing studies are incomplete and a great deal of work still needs to be done. Although the findings presented here are preliminary, they do represent the result of optimal implementations on the PowerPC 7400.
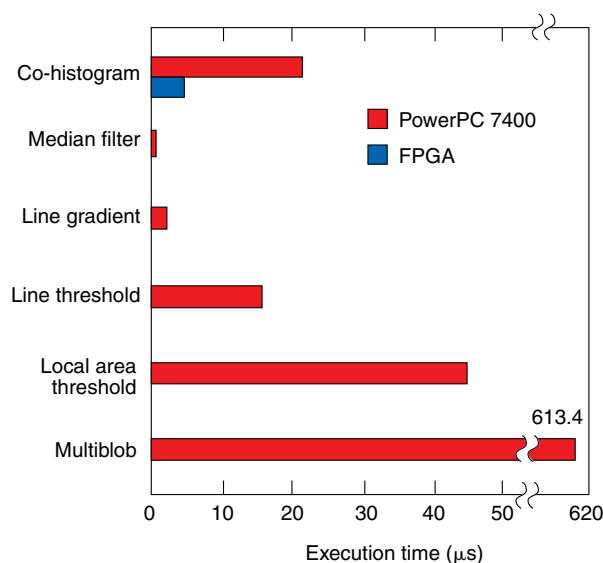

Figure 6. Preliminary algorithm execution times.
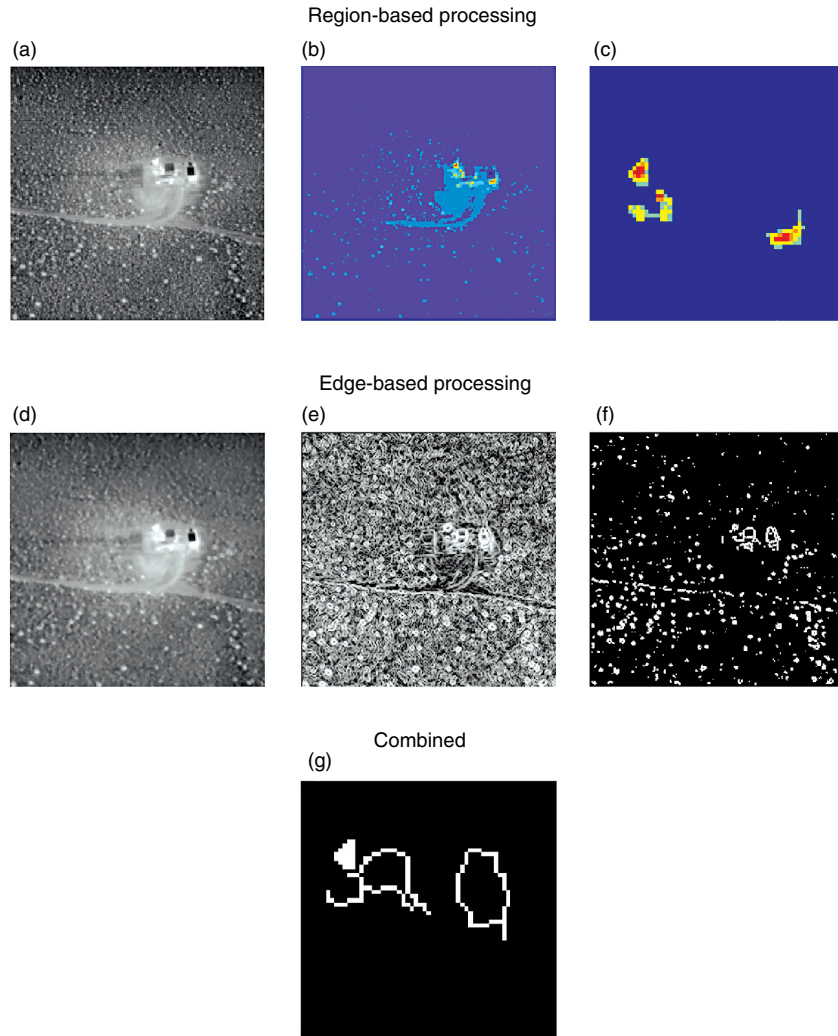
RADAR SITE SCENARIO

An initial demonstration of the detection algorithms was carried out using a sequence of 120 frames simulating action against an air defense unit, as shown in Fig. 7a. The radar antenna was the primary target. After acquisition, a correlation tracker was initiated to see if the antenna could be tracked until the end of the sequence. The seeker began functioning at a nominal range-to-target of 600 m. Figure 7 shows the stages in applying region-based and edge-based processing. The multithreshold applied to the enhanced image histogram produces the quantized image shown in Fig. 7b, having seven gray levels.

The following criterion was used to filter the associated component tree: a node is active if and only if its area is less than 20 pixels, greater than 1,000 pixels, or its gray level is less than 4. Components remaining after this filtering step were considered target-like. This criterion was sufficient to eliminate all natural background features in the image. The remaining objects include the radar antenna (ROI 1) and some very hot reflections from buildings (ROIs 2 and 3). The buildings themselves were not detected by the region-based processing because they were connected to a large component that also contained background and was eliminated by the filter's size criterion.

Figures 7d through 7f show steps in the edge-based processing. Many of the edges and contours that can be seen in Fig. 7e are eliminated when the background threshold is applied. Figure 7g shows the final result of combining the region- and edge-based techniques. Additional filtering of the segmented ROI in Fig. 7c using criteria involving common geometric image-processing features, such as convexity and eccentricity, eliminated the hot reflections from the buildings.

Table 1 compares the values of several of these features for each of the three ROIs (1, 2, and 3, from left to right). In all cases the values for the antenna (object 1) are markedly different than those for the remaining two objects. Thus, if the radar antenna is the intended target in this image, a criterion based on its intensity and geometric structure would be able to discriminate between it and the background. For the edge-based processing, thinning and size and shape discrimination are able to remove all objects except the two dark buildings. The clutter background has been eliminated, and we are left with the desired targets. This was accomplished with no direct user input and no dependence on target structure apart from some mild assumptions concerning target size and intensity, and that the target is composed of straight edges and corners.

The importance of the fusion step can be seen in Fig. 7g. The edge-processing algorithm detected the major buildings, but not the radar antenna. By combining the complementary information from both region and

Region-based processing

(a)   (b)   (c)

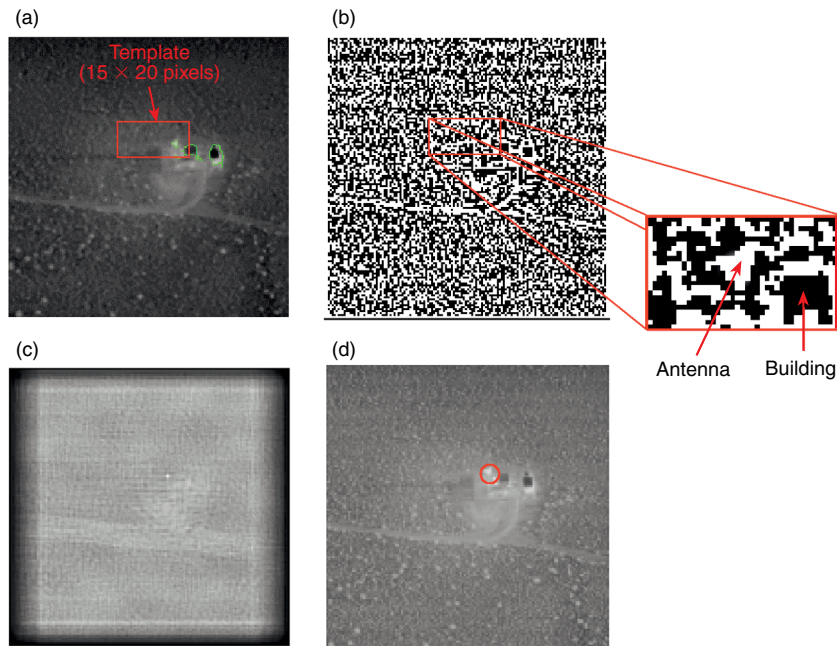Edge-based processing

(d)   (e)   (f)

Combined

(g)

Figure 7. Detection processing steps: (a) original image, (b) quantized image, (c) segmented image (enlarged), (d) median filtered image, (e) edge-enhanced image, (f) edge-segmented image, (g) fused object set.

Table 1. Feature characteristics for ROIs in the segmented image.

Feature	Object 1	2	3
Area (pixels)	26	49	40
Perimeter (pixels)	15	29	25
Compactness	8.65	20.02	15.63
Aspect ratio	0.69	1.78	1.82
Convexity	1.93	0.76	0.86

edge processing, a more complete set of target objects is obtained. Thus, by fusing the outputs of these complementary processors, the discrimination power of more sophisticated image processing techniques can be attained by simple high-throughput processors working in parallel.

The radar antenna was detected in a single frame, and two additional frames were used to verify the detection. Once this validation was complete, tracking began. As shown in Figs. 8a and 8b, the image frame in which the target has been detected is binarized, and a small region centered at the estimated target centroid is used as a template. Each subsequent image frame is binarized and correlated, with the template constructed from the previous image in order to update the target location (Fig. 8c is a typical correlation surface). Then the template is updated to account for changes in the image due to the seeker's motion. In this way the antenna was tracked until the last frame in the sequence (Fig. 8d).

Although the algorithms have not been tested on an extensive set of images, this example is typical of the performance obtained on the real imagery we have examined so far. The ground clutter, while not extreme,

Figure 8. Correlation and tracking: (a) template of a region of the binarized raw image centered on the estimated target centroid, (b) subsequent binarized and correlated image frame, (c) typical correlation surface, (d) updated template and image resulting from the seeker's motion.

is typical of a desert environment and consists primarily of creosote bushes that can have intense IR signatures.

CONCLUSIONS

This article provides an overview of the desired system capabilities of the LCGLS and summarizes work done at the Laboratory to address design constraints and assess system feasibility. The preliminary design includes

- Body-mounted optics and lens materials that can survive the high *g* force at launch
- An uncooled bolometric FPA and lightweight optics to reduce space and weight requirements
- Using the LWIR spectral band to minimize atmospheric degradation

In a demonstration using a sequence of real imagery to simulate closing on a radar site, the target detection system was able to acquire and track the designated target. Moreover, algorithms that have been benchmarked are generally fast.

While these results are promising, more in-depth analyses are required. The algorithms need further development to optimize detection, segmentation, and classification for speed and computational efficiency. A set of features that are robust against a wide variety of target and image conditions and backgrounds must still be defined. Moreover, an automatic procedure must be developed for combining the region- and edge-based algorithms, or fusing their outputs. Finally, the performance of the system needs to be validated in end-to-end flight simulation studies. If successful, such a technology could be inserted into a wide range of weapons.

REFERENCES

[1]Carson, J. C., Shostak, D. R., and Ludwig, D. E., "Smart Sensors Integrate Detectors and Processors," *Laser Focus World*, 113–115 (Jul 1996).
[2]Herring, R. J., and Howard, P. E., "Design and Performance of the ULTRA 320 3 240 Uncooled Focal Plane Array and Sensor," *SPIE* **2746**, 2–12 (1996).
[3]Holst, G. C., "Target Discrimination," Chap. 2, in *Electro-Optical Imaging System Performance*, JCD Publishing, Winter Park, FL, pp. 412–440 (1995).
[4]Rounds, E. M., and Sutty, G., "Segmentation Based on Second-Order Statistics," *SPIE* **205**, 126–132 (1979).
[5]Jones, R., "Connected Filtering and Segmentation Using Component Trees," *Comput. Vis. Image Underst.* **75**(3), 215–228 (1999).
[6]Goutsias, J., et al. (eds.), "Computational Imaging and Vision," Vol. 18 of *Mathematical Morphology and Its Applications to Image and Signal Processing*, Kluwar, The Netherlands (2000).
[7]Weeks, A. R., "Fundamentals of Electronic Image Processing," *SPIE/IEEE Series on Imaging Science & Engineering*, pp. 294–386 (1996).

THE AUTHORS

DONALD E. MAURER is a member of the Senior Professional Staff in APL's Air Defense Systems Department. He holds a B.A. from the University of Colorado and a Ph.D. from the California Institute of Technology. Before joining APL in 1982, he held positions at the Institute for Defense Analyses and the Center for Naval Analyses. At APL, Dr. Maurer has been involved in modeling, simulation, and algorithm development in support of various programs. He was the Principal Investigator for an IR&D program in distributed data fusion, where he developed a metric to measure the similarity between ocean surveillance track pictures. As technical lead for the High-Speed Anti-radiation Missile Receiver Sensitivity Improvement Program, he developed acquisition and target detection algorithms. Currently, he is involved in developing and evaluating radar-to-IR hand over algorithms in support of the Navy Theater-Wide program. Dr. Maurer is a reviewer for the *Mathematical Reviews* and is a member of the American Mathematical Society, the New York Academy of Sciences, and the Society of the Sigma Xi. His e-mail address is donald.maurer@jhuapl.edu.

ERIC W. ROGALA, a member of APL's Senior Professional Staff, is with the Electro-Optical Systems Group in ADSD. He received B.S. and M.S. degrees in optical sciences from the University of Rochester in 1990 and 1992, respectively. In 1999 he received a Ph.D. in optical engineering from the University of Arizona's Optical Sciences Center for his theoretical and experimental work on a novel phase-shifting interferometer capable of mapping the profile and complex index of refraction of a surface. In March 2000 Dr. Rogala joined ADSD's Electro-Optical Systems Group and has been involved in designing and analyzing optical systems as well as developing pattern recognition software. His current interests include automated target recognition in passive IR imagery, automated analysis of solar flare images, boresight error analysis in sapphire dome for SM-2 Block IVA, modeling of 3-D laser radar images, and phase-shifting interferometry of biological cells. His e-mail address is eric.rogala@jhuapl.edu.

ISAAC N. BANKMAN, a member of the APL Principal Professional Staff, is Supervisor of the Imaging and Laser Systems Section in ADSD's Electro-Optical Systems Group. He received a B.S. in electrical engineering from Bosphorus University, Turkey, in 1977, an M.S. in electronics from the University of Wales, Britain, in 1979, and a Ph.D. in biomedical engineering from Technion University, Israel, in 1985. He joined APL in 1990 and worked on signal processing algorithms for transient detection, image processing algorithms for object recognition, and image registration algorithms. In the past 6 years he has developed analytical models of ladar and radar signatures of ballistic missile parts and related discrimination algorithms. His recent work also includes ladar sensing and laser communication in marine applications, image analysis for missile guidance, and radar and passive IR fire control systems for ship self-defense. His e-mail address is isaac.bankman@jhuapl.edu.

BRADLEY G. BOONE is a Principal Professional Staff physicist in the RF Engineering Group of the APL Space Department. After earning his Ph.D. in physics from the University of Virginia in 1977, he joined APL and became involved in a variety of advanced missile guidance projects and was Principal Investigator on numerous IR&D projects in optical signal processing, superconducting electronics, and pattern recognition. After serving as a section supervisor (1983–1996), he was group supervisor for the Electro-Optical Systems Group from 1997 to 2000. He has published over 30 technical papers and one book, and holds four U.S. patents. He has taught extensively for the G.W.C. Whiting School of Engineering and was visiting professor in the Electrical and Computer Engineering Department in 1990–1991. He is currently working on optical communications for deep space applications. His e-mail address is brad.boone@jhuapl.edu.

KATHRYN K. VOGEL is a Senior Professional Staff engineer in the Electro-Optical Systems Group of the Air Defense Systems Department. She received a B.S. degree from Drexel University in 1983 and an M.S. degree from The Johns Hopkins University in 1988, both in electrical engineering. Since joining APL in 1983, Ms. Vogel has worked on the test, evaluation, and simulation of infrared (IR) and inertial systems primarily for missile guidance. Her experience includes coordinating IR seeker test activities, analysis of laboratory and flight data, seeker performance evaluation, inertial component testing, and modeling and simulation of IR seeker dynamics and inertial systems. Currently, she is working on the test and evaluation of the Standard Missile-3 Kinetic Warhead. Her e-mail address is kathryn.vogel@jhuapl.edu.

CHRISTOPHER PARRIS received a Bachelor's degree in electrical engineering from Tennessee Technological University in 1995. After earning his Master's degree in 1997, he was employed by the Naval Surface Warfare Center in Dahlgren, Virginia, working in the Advanced Signal Processors Group in the Area Systems Branch. His main interest is high-speed digital signal processors for radar and image processing. His e-mail address is parriscp@nswc.navy.mil.