
Real-Time 3D Ladar Imaging

Peter Cho, Hyrum Anderson, Robert Hatch, and Prem Ramaswami

■ A prototype image processing system has recently been developed which generates, displays, and analyzes three-dimensional (3D) ladar data in real time. It is based upon a suite of novel algorithms that transform raw ladar data into cleaned 3D images. These algorithms perform noise reduction, ground-plane identification, detector response deconvolution, and illumination pattern renormalization. The system also discriminates static from dynamic objects in a scene. In order to achieve real-time throughput, we have parallelized these algorithms on a Linux cluster. We demonstrate that multiprocessor software plus Blade hardware results in a compact, real-time imagery adjunct to an operating ladar. Finally, we discuss several directions for future work, including automatic recognition of moving people, real-time reconnaissance from mobile platforms, and fusion of ladar plus video imagery. Such enhancements of our prototype imaging system can lead to multiple military and civilian applications of national importance.

RECENT ADVANCES in ladar imaging technology have opened many new possibilities for intelligence gathering and information visualization.* Over the past several years, Lincoln Laboratory researchers have developed arrays of single-photon-counting avalanche photodiodes (APD) and integrated them onto various surveillance platforms [1, 2]. The resulting three-dimensional (3D) imaging systems have demonstrated significant potential for both localized and synoptic surveillance. For example, ladars have yielded detailed views of vehicles hidden under dense foliage [3]. High-resolution height maps of entire cities have also been generated via ladar mapping in under two hours. These impressive results have firmly established the military value of ladar sensing in general and APD technology in particular.

One critical performance metric of any surveillance system is its operating speed. In future conflicts, detailed 3D images of cities, jungles, and deserts will most likely prove invaluable to the warfighter, particularly if they are delivered in a timely fashion. Lin-

coln Laboratory has launched an effort to develop the necessary hardware and software to meet this need for speed. In this article we report upon a prototype system that generates, displays, and analyzes 3D imagery in real time.

We first discuss the real-time system's experimental setup. The data stream comes from a ladar previously developed under the DARPA Jigsaw program, which is currently being utilized for indoor laboratory tests. We next describe how raw angle-angle-range input is transformed into cleaned point cloud output by a sequence of image processing algorithms. These include Cartesian integration, ground-plane detection, APD response deconvolution, illumination pattern renormalization, and static voxel identification. We subsequently quantify the time needed to run these algorithms on a single processor. Real-time throughput is shown to be achievable by parallelizing the work load across multiple machines. The software and hardware that implement this parallelization on the prototype are then discussed. Finally, we mention several interesting extensions of this work that can be pursued in the near future, and we close by listing long-range applications of real-time 3D ladar imaging to problems of national importance.

*The contraction of "laser radar" into the term *ladar* has become part of the optical engineering vernacular. It is synonymous with *lidar*, which also frequently appears in the technical literature.



FIGURE 1. Jigsaw ladar, data acquisition computer, Global Positioning System (GPS), inertial navigation unit (INU), and operator.

Experimental Setup

The prototype system is based upon the ladar pictured in Figure 1. This sensor was originally designed for foliage penetration experiments as part of the DARPA Jigsaw program, and its operating parameters are listed in Table 1. The ladar was rapidly developed in 2003 and flown on board a helicopter to search for stationary targets under trees. Following the conclusion of the outdoor field tests, the sensor was brought back to Lincoln Laboratory, where it is now being employed for surface reconnaissance experiments.

Figure 2 illustrates the scene in view of the Jigsaw ladar within Lincoln Laboratory's optical systems test range. The ceilings and walls of this fifty-meter-long hallway are colored black to minimize laser light reflections. Its height and width are also reasonably well matched to the ladar's conical field of view. The test range provides a controlled setting where large quantities of 3D data can be safely and efficiently collected.

Many background objects of different shapes and sizes residing in the test range make for good stationary targets to image with the ladar. As detecting, identifying, and tracking mobile targets represent key objectives for a real-time system, human subjects have also been introduced into experimental scenes. For example, the three people posing in Figure 3 prior to donning safety goggles were imaged by Jigsaw as they moved around the test range and played with various objects. Note that one subject wore a dark trench coat while another sported a large sombrero on his head.

Table 1. Jigsaw Ladar Operating Parameters

Wavelength	532 nm
Pulse repetition frequency	16 kHz
Avalanche photodiode array size	32 × 32 pixels
Field of view	11°
Range resolution	30 cm

Jigsaw data of the human subjects were collected in the test range in 2005. Figure 4 exhibits one representative example of raw Jigsaw imagery. The 3D image is color coded according to range. Cool-colored volume elements (voxels) are located closer to the ladar, while warm-colored voxels are positioned farther down range, as shown in Figure 4(a). The point cloud's cross-sectional pattern is generated by two counter-rotating Risley prisms that spread laser beam light throughout the hallway [3]. Since the prisms' angular velocities slightly differ, the entire cloud precesses over a 12.25 second period.



FIGURE 2. Ladar's view of the optical systems test range.



FIGURE 3. Human subjects standing next to various test-range objects.

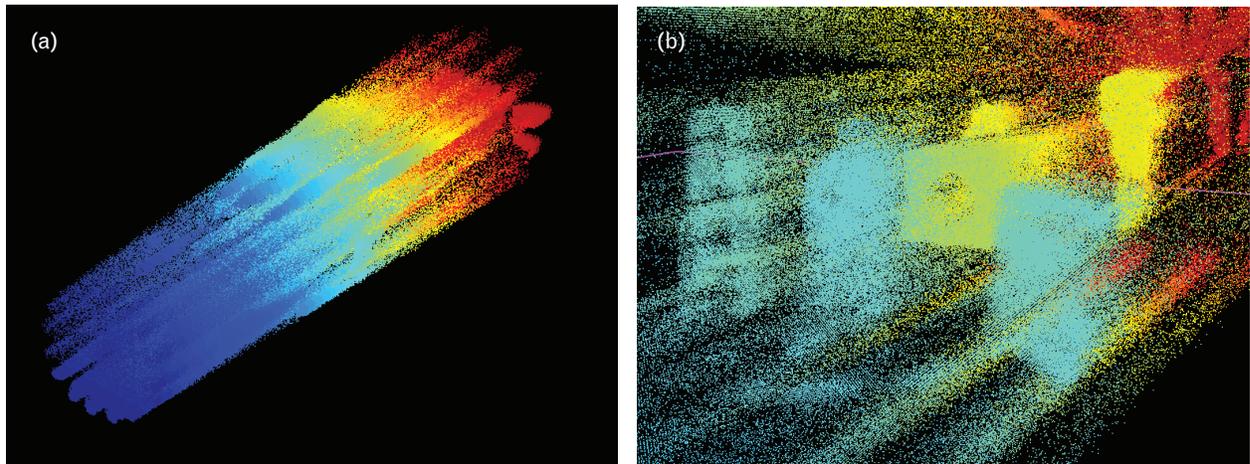


FIGURE 4. (a) Raw Jigsaw data colored according to range relative to the ladar; (b) view from inside the raw point cloud of genuine signal and background noise.

Most of the points appearing in the spinning point cloud originate from APD dark counts [1]. However, after zooming into the cloud's center, we start to see genuine signal, as shown in Figure 4(b). Significant processing obviously needs to be performed on the raw data to convert them into useful imagery.

Coincidence processing algorithms developed in the past have laid the foundation for decreasing noise and increasing interpretability of ladar data. But previous approaches to 3D imagery generation did not extract all useful information content, nor did they strive for peak efficiency. Refining and extending these earlier efforts represented our first technical challenge. We present our new set of image processing algorithms in the following section.

Image Processing Algorithms

The algorithm flow that transforms raw ladar input into cleaned 3D output follows a series of basic steps, as the diagram in Figure 5 indicates. A machine first converts angle-angle-range data for a single laser pulse into Cartesian locations. It then bins the XYZ points into $(2.5 \text{ cm})^3$ voxels inside a nominal $8 \text{ m} \times 25 \text{ m} \times 5 \text{ m}$ volume of interest. Converting from continuous to discrete coordinates allows 3D information to be stored in a one-dimensional hash table with voxel locations mapped to unique integer keys. Utilization of this sparse data structure is well suited to single photon counting, for the vast majority of the volume's voxels remain empty over time. After 4000 laser pulses are

transmitted and received during a $\frac{1}{4}$ -second interval, a single 3D image is formed by aggregating all non-empty voxels in the hash table.

The machine next starts to clean the integrated image by pruning voxels containing low numbers of counts, which most likely correspond to noise. We have empirically determined that setting a lower global bound of four voxel counts optimally enhances signal to noise. The computer also excises voxels located near the angular periphery of the sensor's conical field of view. Like any periodic oscillator, the Risley pattern's direction vector dwells longer at its turning points than at its center, causing dark counts to accumulate near the pattern's periphery. We crop away 1.25° from the conical output to suppress these noisy regions without sacrificing much useful ladar signal. Genuine targets then become observable in the partially cleaned image, as exhibited in Figure 6.

It is useful to anchor the Jigsaw point cloud to a fixed reference frame. The machine consequently searches for the ground within the target scene. In the benign setting of the optical system test range, the ground surface is simply represented by a Z-plane. As illustrated in Figure 7, its height is readily extracted from the lowest peak in the Z-distribution obtained after integrating over X- and Y-coordinates in the partially cleaned point cloud. Since the ground's location is time invariant, it should theoretically be unnecessary to compute the cloud's height distribution more than once. But drifts in Jigsaw's Global Positioning

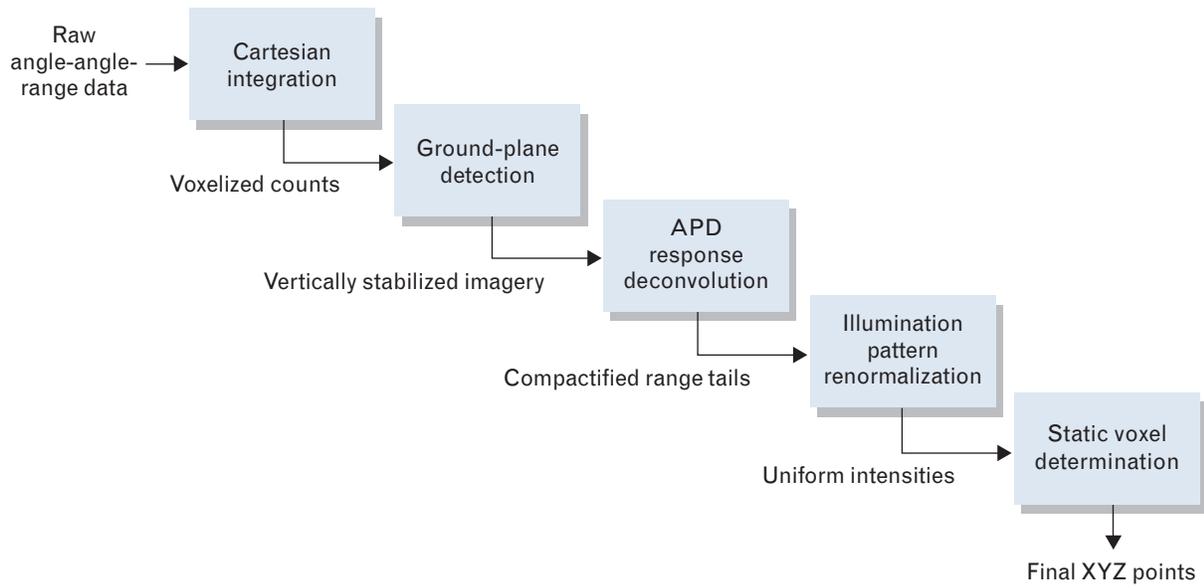


FIGURE 5. Image processing algorithm flow that transforms raw Jigsaw input into cleaned three-dimensional (3D) output. Intermediate results from each stage in the flow are presented in Figures 6 through 12.

System (GPS) and inertial navigation unit (INU) cause point positions to wander over time. To counter the dominant Z-direction drift, we must recalculate the ground's location at every image time, or else ac-

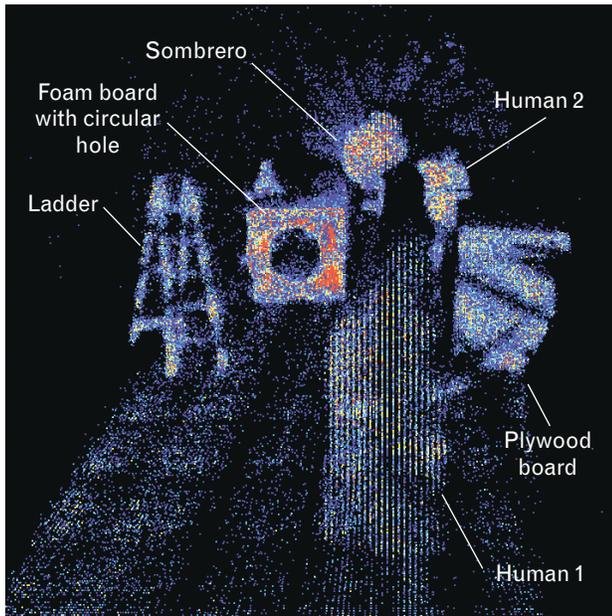


FIGURE 6. Voxelized counts output from the first $\frac{1}{4}$ -second Cartesian integration step in the algorithm flow of Figure 5. Following preliminary cleaning of the 3D image, human subjects and test-range objects are much more apparent than in the raw input displayed in Figure 4.

cept target positions relative to a fixed sensor in place of absolute geolocations. In either case, the 3D image becomes vertically stabilized, and it can be displayed relative to a ground-plane grid, as shown in Figure 8.

The machine next corrects for imperfect timing responses of the APD detectors. If the responses were ideal delta functions, all reconstructed scattering surfaces would be infinitesimally thin. But actual APD response functions exhibit fast rises followed by slow

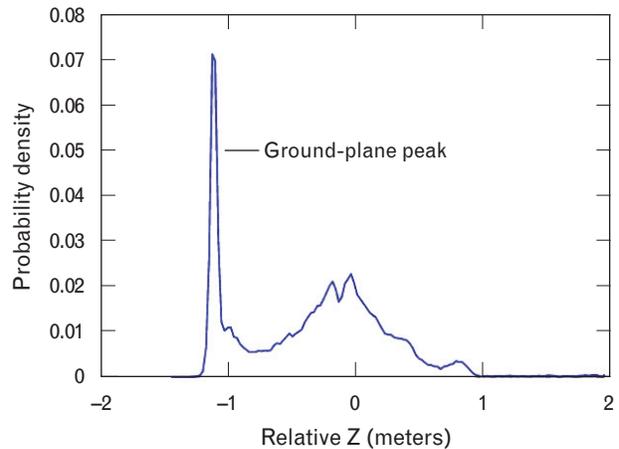


FIGURE 7. Height distribution derived from the partially cleaned point cloud in the ground-plane detection step of Figure 5. The strong peak marks the location of the ground plane.

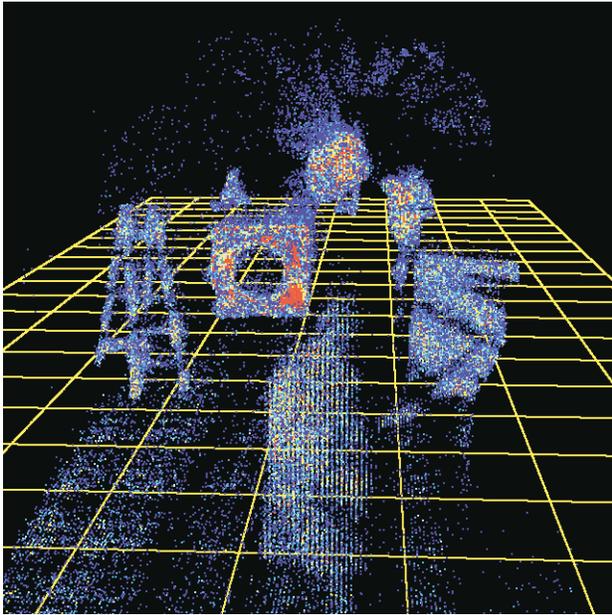


FIGURE 8. Vertically stabilized point cloud appearance after subtraction of below-ground noise. Each cell within the ground-plane grid is 1 m² in size.

decays. This behavior induces comet-like range tails in Jigsaw imagery. Such tails are clearly visible in the front and overhead views of the test range presented in Figure 9.

The long range tails might seem to pose a serious problem for automated imagery generation and display, but the machine can actually exploit them. Range tails clearly mark interesting objects in the scene. On the other hand, isolated noise points arising from APD dark counts do not exhibit such tails. This observation helps motivate our deconvolution algorithm (presented in the sidebar entitled “Detector Response Deconvolution”), which the machine employs to rapidly remove range tails and enhance signal to noise. After comparing deconvolved results in Figure 10 with their uncorrected progenitors in Figure 9, we see that the deconvolution algorithm yields a striking improvement in 3D image quality.

The Jigsaw ladar’s illumination pattern poses another challenge for imagery generation. Its slow pre-

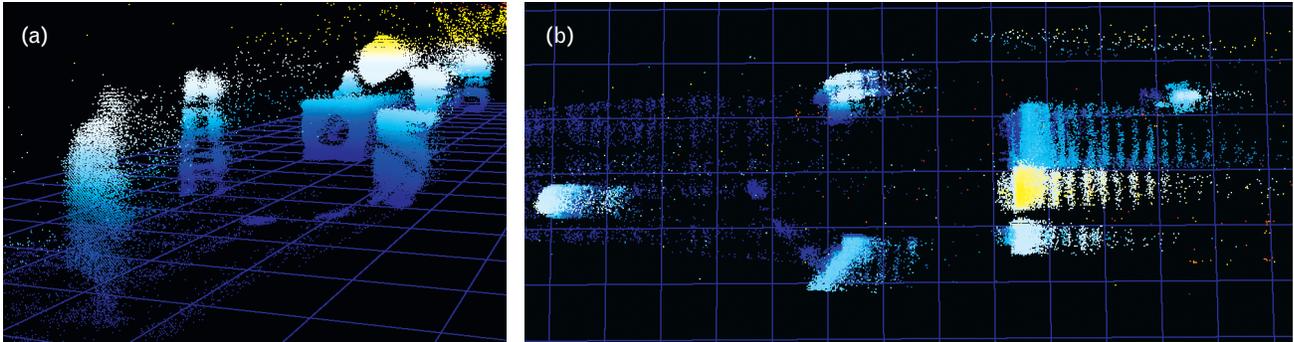


FIGURE 9. Range tails resulting from imperfect avalanche photodiode (APD) response. (a) Side and (b) overhead views of the same 3D scene are colored according to height.

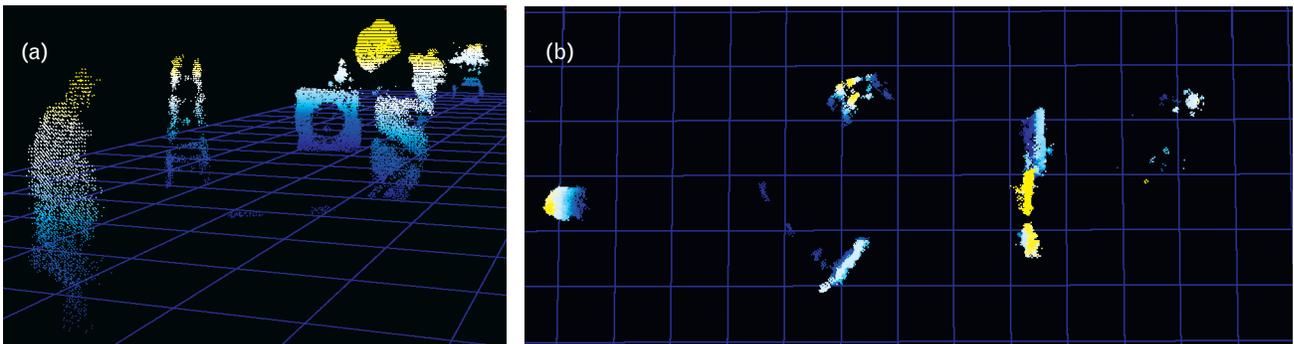


FIGURE 10. Range tails are suppressed and signal to noise is amplified following the APD response deconvolution step in Figure 5. (a) Side and (b) overhead views, similar to those in Figure 9, are presented here for direct comparison.

DETECTOR RESPONSE DECONVOLUTION

AVALANCHE photodiodes measure times of flight for single photons to travel from a laser transmitter, scatter off some downrange target, and return to the receiver. If their response functions were ideal, all reconstructed 3D surfaces for opaque targets would be precisely one voxel deep, as shown in Figure A(1). But actual avalanche photodiode (APD) responses are never perfect. Solid-state physics effects in these devices such as spatially dependent electric fields, random locations for doping impurities, and nonuniform electron mobility smear the times when returning photons are detected. Photons can never be sensed sooner than the minimal travel time set by the ratio of a target's fixed range to the constant speed of light. But variable responsiveness does in-

duce detector delays that give rise to range tails in output ladar imagery, as seen in Figure A(2).

Standard deconvolution methods are not well suited for removing these artifacts. They rely upon knowledge of APD response functions, which is difficult to precisely measure, and they do not distinguish between genuine signal and background noise. Furthermore, conventional deconvolution is computationally expensive. We have consequently devised a new technique that takes full advantage of sparse hash-table storage of APD count information. It rapidly identifies range tails and compacts their contents into scattering voxels' true spatial positions.

The machine first sorts hash-table entries by range and locates the voxel closest to the end of the

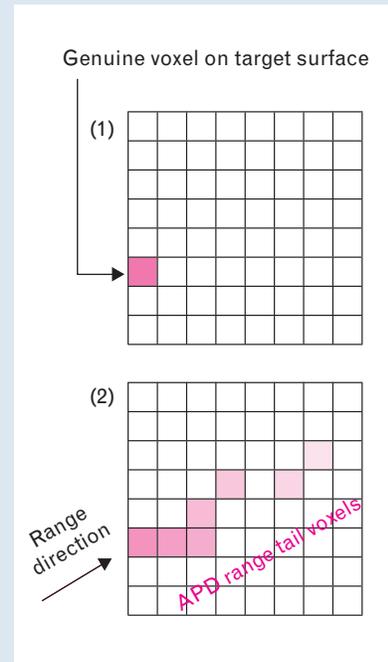


FIGURE A. Ideal versus real APD detector response. (1) If avalanche photodiodes were perfect devices, all reconstructed surfaces would be one voxel deep. (2) Actual APD response induces range smearing.

cession allows the sensor to cover the entire conical field of view. But the precession also introduces time dependence into output imagery, which has nothing to do with genuine motion in the scene. Moreover, the sixteen-petalled rosette pattern produced during $\frac{1}{4}$ second by Jigsaw's Risley prisms is complex, as shown in Figure 11. Local intensities vary over the rosette by three orders of magnitude. Illumination flux needs to be taken into account before a post-deconvolution thresholding operation is performed to reduce surviving noise. But exactly determining the instantaneous illumination pattern is computationally expensive. As a compromise between accuracy and speed, our machine utilizes a time-averaged bull's-eye pattern,

exhibited alongside the rosette in Figure 12, to remove illumination variation.

Representative examples of final processed images taken from a one-minute data sequence are displayed in Figure 13. Each image is color coded according to range relative to the fixed ladar. Compared to raw Jigsaw input, the noise content of the cleaned output is dramatically diminished. As a result, various test-range targets such as the human subjects, the sombrero, and the ladder are readily recognizable. The locations of these objects relative to the sensor and each other are easily measured via the ground-plane grid.

In Figures 13(a) and 13(b) we see obvious angular gaps within the broad plywood board and ladder

range window. It then employs a 3D generalization of the midpoint line algorithm developed in computer graphics for efficiently rendering segments [1] to search uprange for partner voxels containing counts, as shown in Figure B(1).^{*} If more than three such partners are found less than one meter upstream, the machine concludes the voxel belongs to the range tail of some genuine target. To increase signal to noise, the machine multiplies the voxel's counts by ten and transfers them to the voxel found furthest upstream, as shown in Figure B(2).

^{*} To maximize deconvolution algorithm speed, we neglect the small cross-range dependence of the uprange direction vector as it varies over the volume of interest. This approximation is quite good for 11° field-of-view Jigsaw imagery.

After iterating this compactification and amplification procedure over every nonempty voxel in the volume of interest, all range tails are flattened into thin scattering surfaces, as shown in Figure B(3). As the results in Figures 9 and 10 demonstrate, this deconvolution technique greatly improves output 3D image quality.

Reference

1. J.D. Foley, A. van Dam, S.K. Feiner and J.F. Hughes, *Computer Graphics: Principles and Practice* (Addison-Wesley, Reading, Mass., 1987).

FIGURE B. Deconvolution technique. (1) For each non-empty voxel, search uprange for other voxels containing counts. (2) If the number of uprange voxels exceeds a certain threshold, amplify the original's counts and transfer upstream. (3) Iterate procedure to compact all range tails and increase signal to noise.

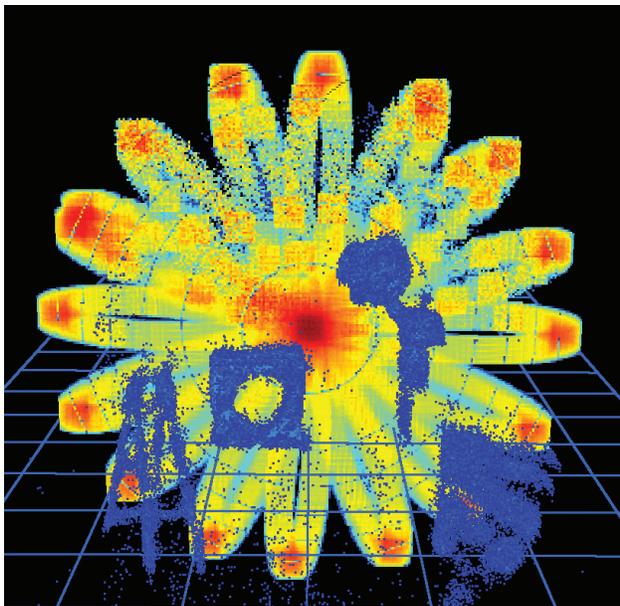
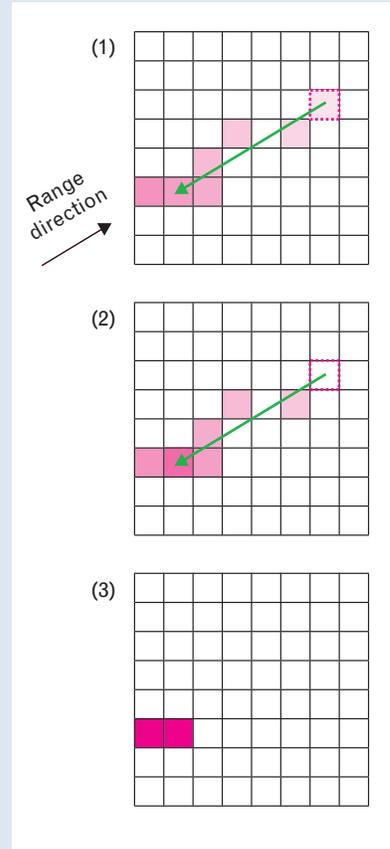


FIGURE 11. Jigsaw ladar's rosette illumination pattern superimposed onto the 3D scene. The false coloring is based on a logarithmic intensity scale.

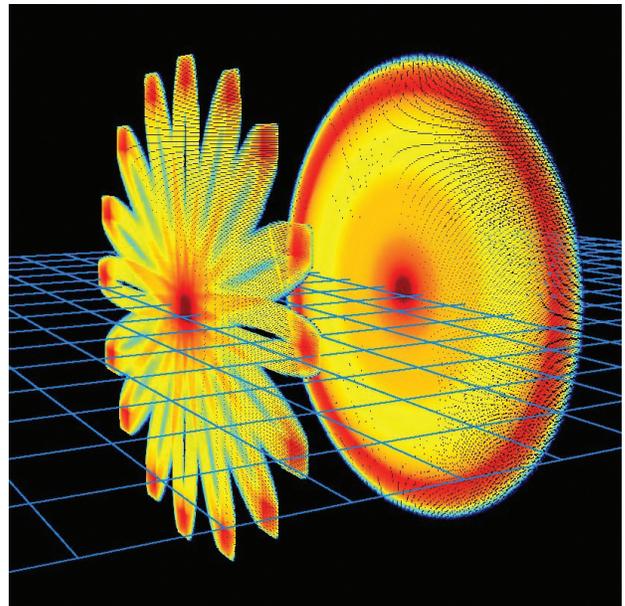


FIGURE 12. Time-averaged bull's-eye approximation used in place of the rosette illumination pattern within the renormalization step in Figure 5.

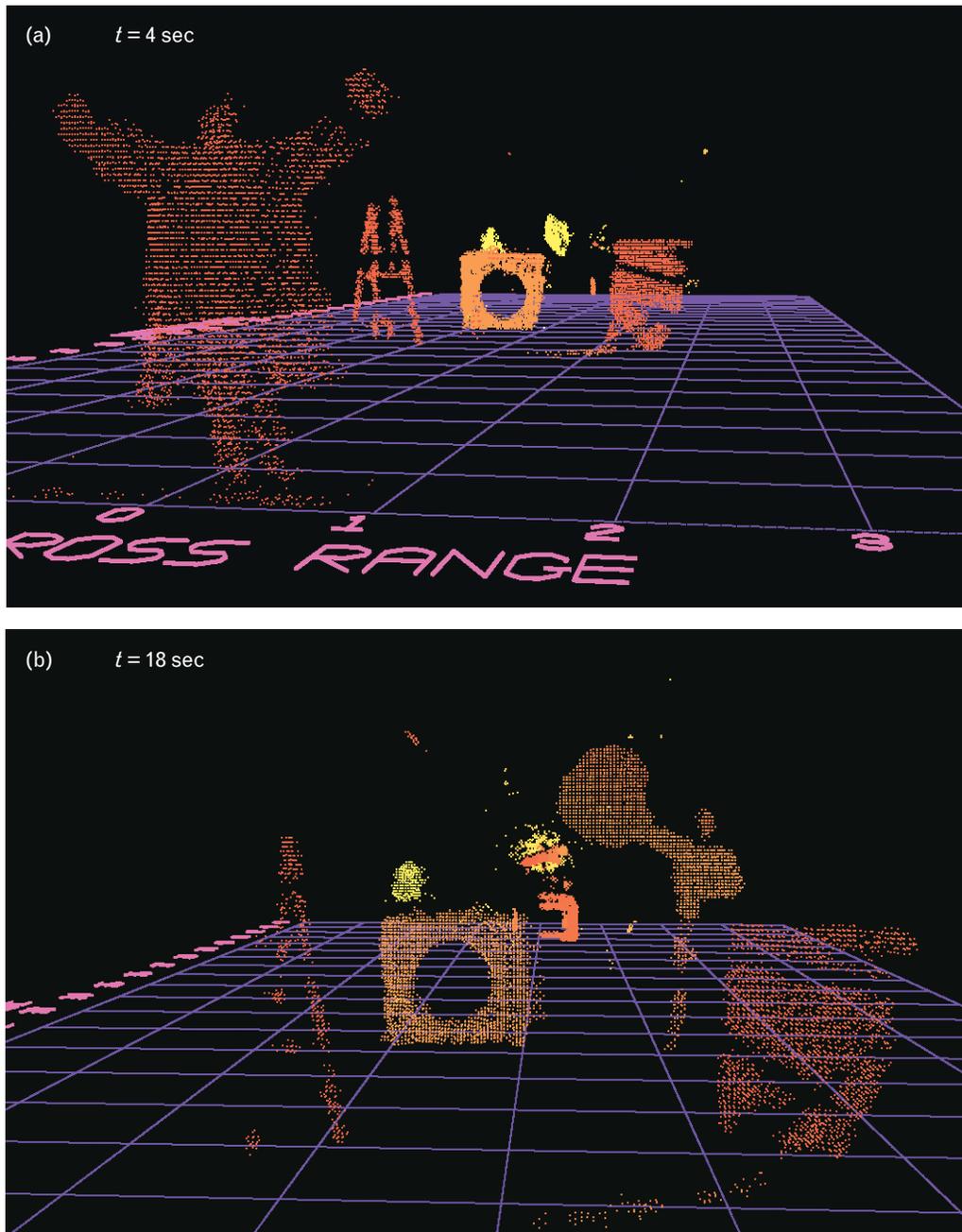
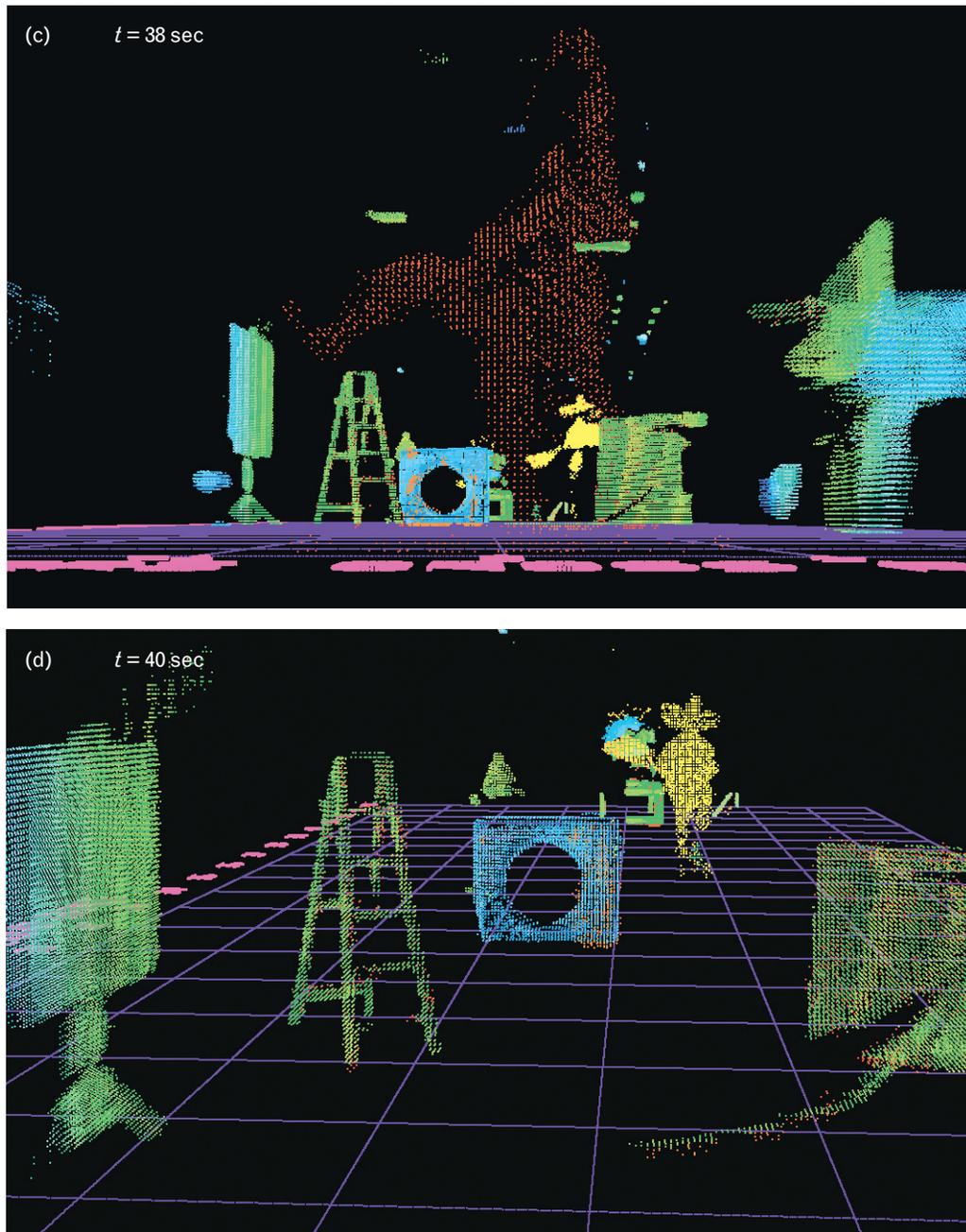


FIGURE 13. Representative examples of final processed $\frac{1}{4}$ -second imagery output from the algorithm flow of Figure 5. XYZ points are colored according to periodic range. (a) Trench-coat subject stretching out his arms near the start of the 25 m range window. (b) Second human subject brandishing his sombrero on the right. The ladder on the left is momentarily shadowed by the trench coat subject. (c, following page) Trench coat subject leaning back and kicking up. By this time in the sequence, targets automatically identified as static are marked by cool colors (green, blue), while mobile objects are indicated by warm colors (red, orange, yellow). (d, following page) Sombrero subject walking down range toward the rear wall of the test range.

target arising from illumination pattern nulls. When viewing the entire 3D sequence as a four-frame-per-second movie, we mentally perform temporal averaging and fill in the missing instantaneous regions. But

it would clearly be preferable if the system integration time could be extended beyond $\frac{1}{4}$ second in order to fill illumination gaps and accumulate fine detail.

Imaging dynamic scenes always introduces a ten-



sion between maximizing light collection and minimizing motion blur. To resolve this exposure time conflict, our algorithms integrate ladar data over two separate intervals. As we have described, the machine accumulates every 4000 pulses of laser returns into $\frac{1}{4}$ -second images. It also performs a separate 25-second integration to search for voxels that repeatedly produce strong signals. If the number of long-term counts in a fixed voxel exceeds an empirically determined lower limit of thirty, the machine assumes its position in the

volume of interest lies upon the surface of some stationary object. The voxel is then permanently lit up whether or not it is detected in subsequent $\frac{1}{4}$ -second intervals. All such static voxels within the entire 3D volume of interest are displayed in Figure 14, while one particular zoomed view of a stationary target is presented in Figure 15.

Static targets show up quite clearly in long-exposure 3D imagery. On the other hand, targets that move on time scales shorter than 25 seconds do not

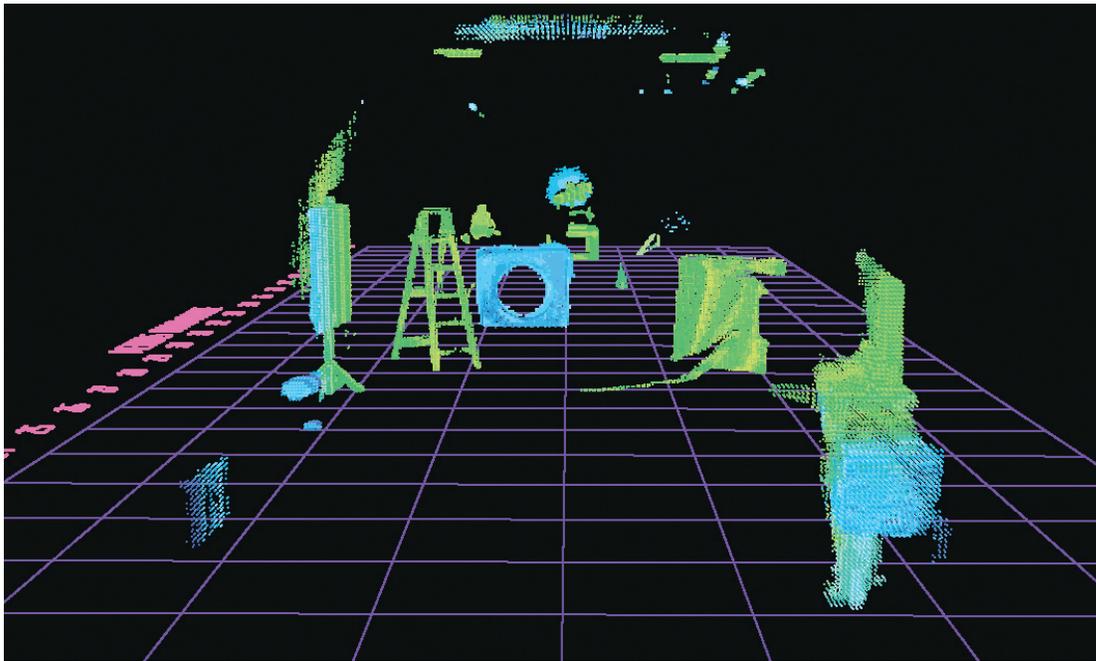


FIGURE 14. Voxels identified as static within the $8 \times 25 \times 5$ m volume of interest by the end of the one-minute data sequence. Given the complex temporal dependence of the ladar's illumination pattern, these classification results require a nontrivial degree of machine intelligence.

generally add enough counts to fixed voxels for them to be tagged as stationary. Hence this dual time-scale technique distinguishes moving targets from stationary ones. As Figures 13(c) and 13(d) illustrate, the machine uses cool colors to indicate range to stationary objects. After less than one minute, the computer effectively finds all static voxels and subsequently associates warm colors with only mobile targets.

The viewing directions in Figures 13 and 14 basi-

cally coincide with that of the sensor. But because ladar images are truly three-dimensional, they can be mathematically rotated and observed from any perspective, as seen in Figure 16. In particular, we can choose to view the entire scene from a vantage point looking back toward the illuminator. It is important to note that such visualization capabilities of active 3D ladar sensing are very difficult to replicate by passive 2D imaging.

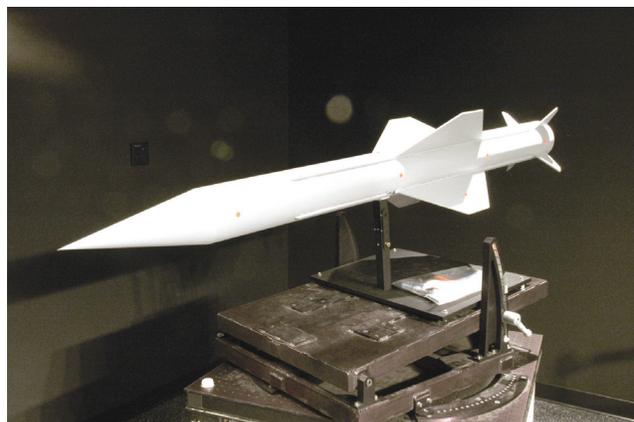
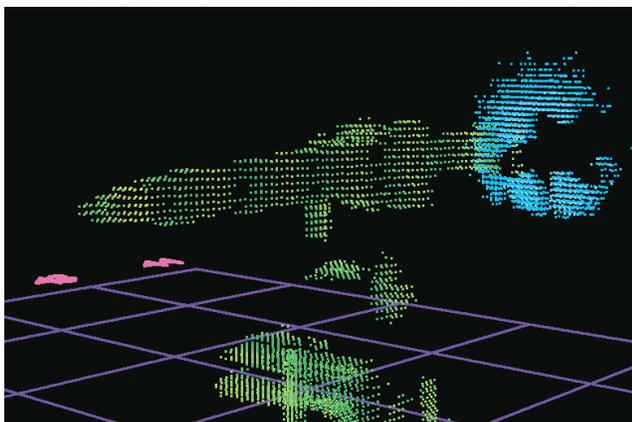


FIGURE 15. (a) Close-up of a missile model located at the back of the test range in Figure 14. A shadow cast by the model onto the near wall is visible in the ladar output. (b) Conventional photograph of the missile model shown for comparison.

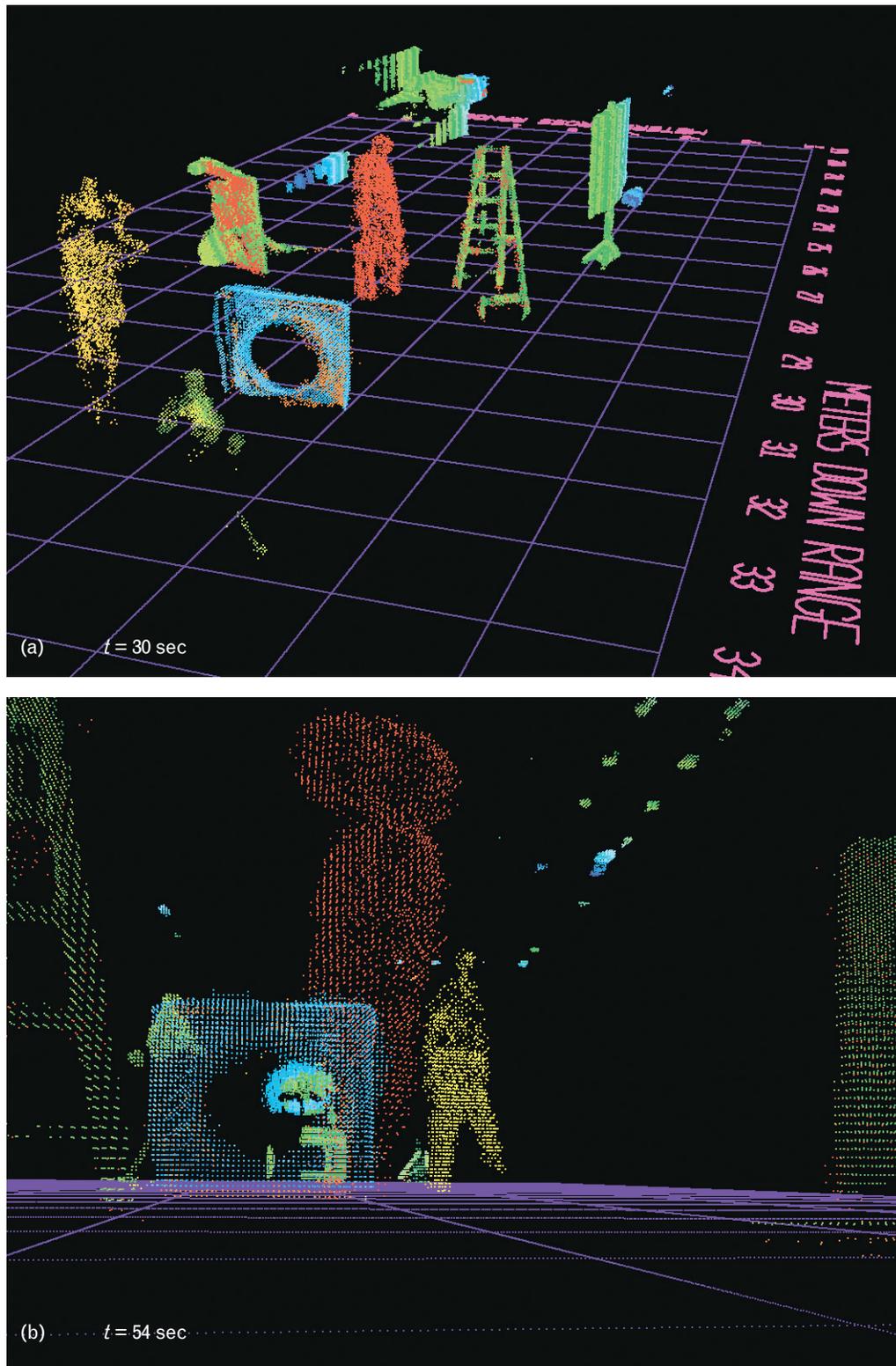


FIGURE 16. More examples of 3D images taken from additional one-minute data sets. (a) View looking back toward the ladar. (b) View looking upward from the floor. Recessed ceiling lights in the test range, which are clearly visible in the upper right part of this image, indicate the size of small details that can be resolved by the ladar system.

Table 2. Single Processor Timing Results Measured on a 3 GHz Pentium Machine

Algorithm task	Processing time/real time
Raw data input	0.1
Cartesian integration	3.0
Ground detection	0.2
Response deconvolution	3.2
Static voxel determination	1.1
XYZ file output	0.7
Total	8.2

Software and Hardware Parallelization

The processing that transforms raw Jigsaw data into clean 3D imagery is computationally intensive. During the course of developing our algorithms, we made every reasonable effort to optimize their computer code implementation.* Table 2 lists the times required by our streamlined code to perform various algorithmic steps. As indicated in the table's bottom line, a single 3 GHz computer takes 820 seconds to process 100 seconds of raw data. To achieve real-time throughput on commercial off-the-shelf hardware, the processing must be parallelized across multiple machines.

* We wrote all our programs in C++, compiled them with GCC [4], and ran them under Redhat Linux 9.0.

Fortunately, the algorithm flow naturally parallelizes in time. This basic observation underlies the multiprocessor architecture diagrammed in Figure 17. The first 4000 pulses coming out of Jigsaw are routed to the first compute node. While that machine is generating the first image, the next 4000 pulses are routed to the second compute node. After their generation, the images from each compute node are forwarded to a gather node where proper temporal ordering is enforced. Final output is subsequently directed to a data viewer machine for 3D display.

To speed up throughput by an approximate factor of ten, we need approximately ten compute nodes. Since each compute node takes nearly two seconds to process ¼ second of ladar data, the parallelized architecture of Figure 17 inevitably introduces some latency delay. But real-time ladar imagery generation can be achieved provided no machine ever acts as a bottleneck.

While this approach to parallel processing is theoretically simple, its practical implementation is not. We first must consider bandwidth requirements for communication links between nodes in the network. These are most stringent at the head of the processing chain where Jigsaw data must be relayed at the rate

$$16,000 \frac{\text{pulses}}{\text{sec}} \times 1024 \frac{\text{APD pixels}}{\text{pulse}} \times 2 \frac{\text{bytes}}{\text{APD pixel}} = 33 \frac{\text{Mbytes}}{\text{sec}}$$

Moreover, no packets transmitted between computers

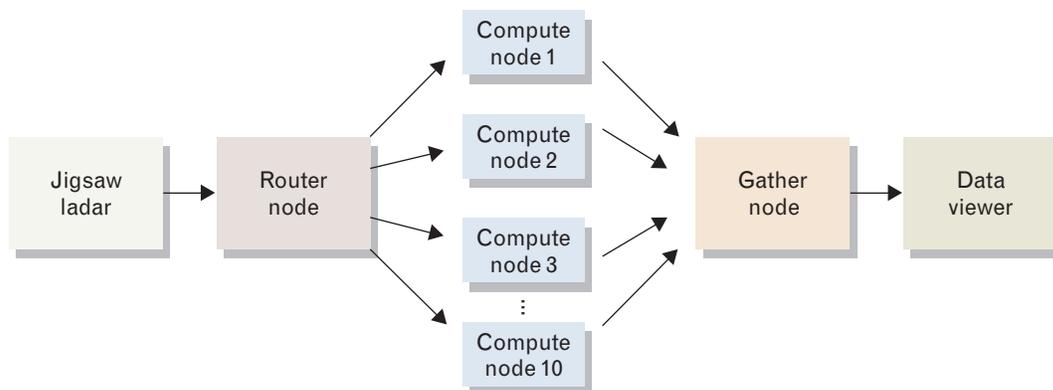


FIGURE 17. Parallel processing architecture implemented to achieve real-time throughput. Raw data from the Jigsaw ladar are routed to multiple compute nodes. Individual processed images are subsequently gathered together and temporally ordered. Final output is sent to a data viewer for 3D display.

should ever be lost due to communication hang-ups. We therefore let fail-safe TCP/IP handle all internode communication over gigabit Ethernet links. This backbone choice can sustain 90 Mbytes/sec transfer rates, which comfortably exceeds our system's maximum throughput requirement. Its generality also worked well for prototype development and enables future system growth.

The multiple compute nodes in Figure 17 represent the parallelized system's core. Figure 18 provides a closer look inside one such node. On the left, we see a dedicated socket reader thread that constantly monitors the input gigabit Ethernet connection for incoming packets. As soon as any new data are detected, the reader thread performs a copy from the socket into a first-in first-out (FIFO) buffer. Temporarily storing data in the FIFO buffer buys time for the slow image generation process (0.2 Mbytes in approximately 2 sec) to catch up with the much faster input burst (8.25 Mbytes in 0.25 sec). Once an image is formed, it is passed to a socket writer thread for decomposition into packets prior to transmission across the output gigabit Ethernet connection. The writer thread also isolates the main processing thread from network traffic. This separation prevents the main thread from stalling when downstream sockets become busy or unavailable.

To guarantee real-time throughput, we have instrumented our parallelized codes with checks for timing slowdowns. For example, a partially generated image is jettisoned if new data from the router node enter the FIFO buffer. Such data dumping rarely occurs, and its visual impact is minor. But compensating for

occasional large timing fluctuations is mandatory to prevent image processing from gradually slipping behind sensor collection.

Since all $\frac{1}{4}$ -second images are constructed from independent sets of 4000 pulses, parallelizing their generation is relatively straightforward. Each compute node transmits its short time-scale results to the gather node pictured in Figure 17. However, incorporating static voxel determination into the multiprocessor architecture is nontrivial, for information extracted by the separate compute nodes must be shared. So each compute node also transmits information about voxels containing two or more counts prior to initial global thresholding to an aggregate node not pictured in the figure.

The aggregate node performs long time-scale integration and identifies static voxels. Its output is sent to the gather node where stationary and dynamic voxels are fused. As computation of the former is slower than that of the latter, the gather node allows asynchronous combining of compute and aggregate node results. Static information for one image may be merged with dynamic information from a later image. Since static voxels do not change over time, such asynchronous merging is acceptable.

To test the parallelized multiprocessor architecture, we implemented a simplified version on Lincoln Laboratory's LLGrid cluster [5]. A simple simulator was written to take the place of the Jigsaw ladar. Displaying 3D imagery on the tail-end data viewer was also initially neglected. LLGrid access then enabled rapid development of communication software that links together all other nodes in the network. It also

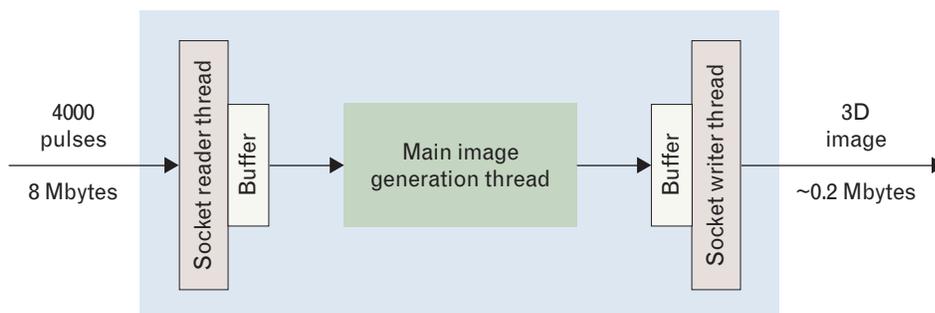


FIGURE 18. Data flowing through and threads going on within an individual compute node. The main image generation thread produces an output $\frac{1}{4}$ -second image that is more than an order of magnitude smaller in size than the 4000 pulses of raw input data.



FIGURE 19. IBM Blade server cluster installed at bottom of the Jigsaw computer rack.

demonstrated the feasibility of real-time 3D imagery generation. (For more information on the LLGrid, see the article entitled “Interactive Grid Computing at Lincoln Laboratory,” by Nadya Travinin Bliss et al., in this issue.)

Our testing experience with LLGrid helped bring into focus hardware issues related to building a dedicated cluster. In particular, the stand-alone network must be powerful enough to perform all necessary image processing in real time. Yet it must also be small enough to fit inside a mobile vehicle for future outdoor reconnaissance experiments. These considerations motivated us to purchase an IBM Blade server cluster containing thirteen dual processors inside a 12" × 17.5" × 28" enclosure. As can be seen in Figure 19, its compact size permitted mounting the cluster as a Jigsaw adjunct inside the rack already holding the sensor's data acquisition computer and GPS/INU.

After the parallelized software and cluster hardware were set up, we compared the rates at which raw data go into and 3D images come out of the Blade. Timing results are plotted in Figure 20(a) where we see a 3D image emerging every $\frac{1}{4}$ second on average. The Blade

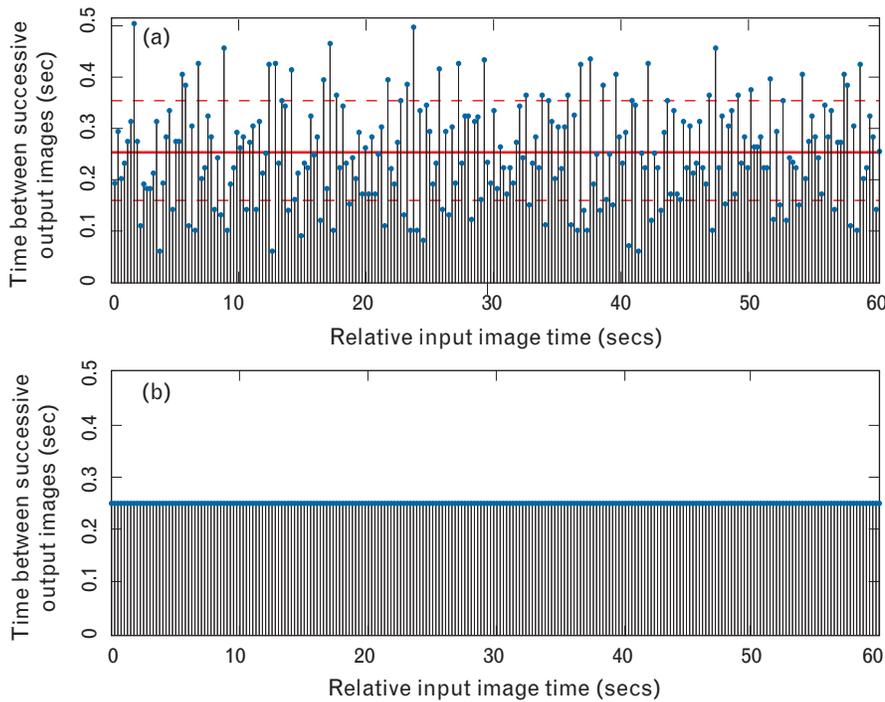


FIGURE 20. Blade server cluster timing results. (a) On average, a 3D image is generated every 0.25 ± 0.1 second. (b) After inclusion of a final first-in first-out (FIFO) buffer, images are generated every 0.25 ± 0 second.

adjunct consequently keeps up with the Jigsaw ladar, but its output rate is uneven. Subsecond timing fluctuations stemming from variable scene complexity in imagery display are quite displeasing to the eye. Some mechanism is therefore needed to eliminate such jerkiness.

We chose to add a final FIFO buffer where every image is parked for up to one second before being displayed. This design choice increases the total system latency delay to approximately three seconds. But as Figure 20(b) proves, all timing jitter is absorbed by the last buffer. Final images thus appear on the 3D data viewer precisely four times per second.

The real-time ladar imaging system has been run continuously in the optical systems test range for intervals approaching an hour in duration. It has also been successfully demonstrated in several live tests. Hence the combination of imagery generation algorithms, multiprocessor communication software, and cluster computing hardware presented in this article has yielded an operational real-time 3D imaging capability.

Future Work

The development of the prototype system opens many new directions for future research. We briefly discuss here extensions of our work that can now be pursued. If carried through to their full potential, these machine intelligence, mobile reconnaissance, and sensor fusion capabilities can lead to important military and civilian applications.

A first obvious direction for future work is automated exploitation of four-dimensional (4D) geometry. The algorithms that have been developed to distinguish stationary and mobile objects do so only at the level of individual voxels. It would be more useful if a machine could identify point cloud subsets by using collective shape and speed information. Template matching has successfully been performed on static Jigsaw imagery in the past [6–8]. But template-based recognition will have to be generalized for time-dependent data containing nonrigid bodies. People moving around complicated scenes represent automation targets of special importance. We expect humans can be identified in 4D imagery via their general volume, shape, and speed characteristics, as Figure 21 suggests.

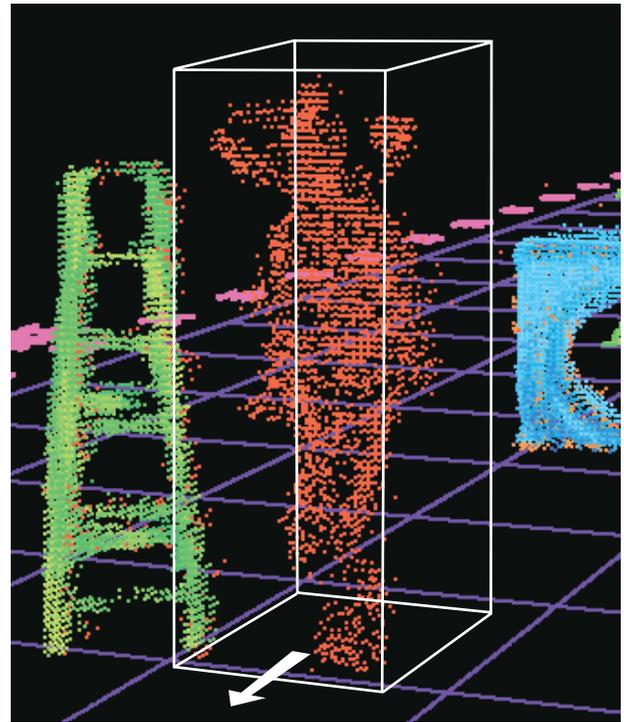


FIGURE 21. Static and dynamic targets should automatically be recognized on the basis of their shapes and speeds. As the white bounding box and motion arrow suggest, humans might be identified via isolated point clouds occupying $1\text{ m} \times 1\text{ m} \times 2\text{ m}$ volumes, moving at approximately one meter per second.

However, much more work needs to be done before such ambitious machine intelligence goals will be realized.

Further algorithm development is also needed to enable the prototype real-time system to operate on a mobile platform. To counter GPS drifts, the prototype currently relates all target locations to the ladar's position, which has so far remained fixed inside the optical systems test range. This stationary sensor constraint must be relaxed in the future. Disentangling *a priori* unknown target motion from imperfectly measured sensor movement at subvoxel accuracy poses a major challenge. We suspect rapid recognition of static objects in the ladar's field of view will play a central role in solving this 4D mobile reconnaissance problem.

Ladar and video data fusion represents another important research direction. These two imaging modalities exhibit complementary strengths. Active ladars are good at generating high-resolution 3D geometry maps, while passive cameras are better at measuring

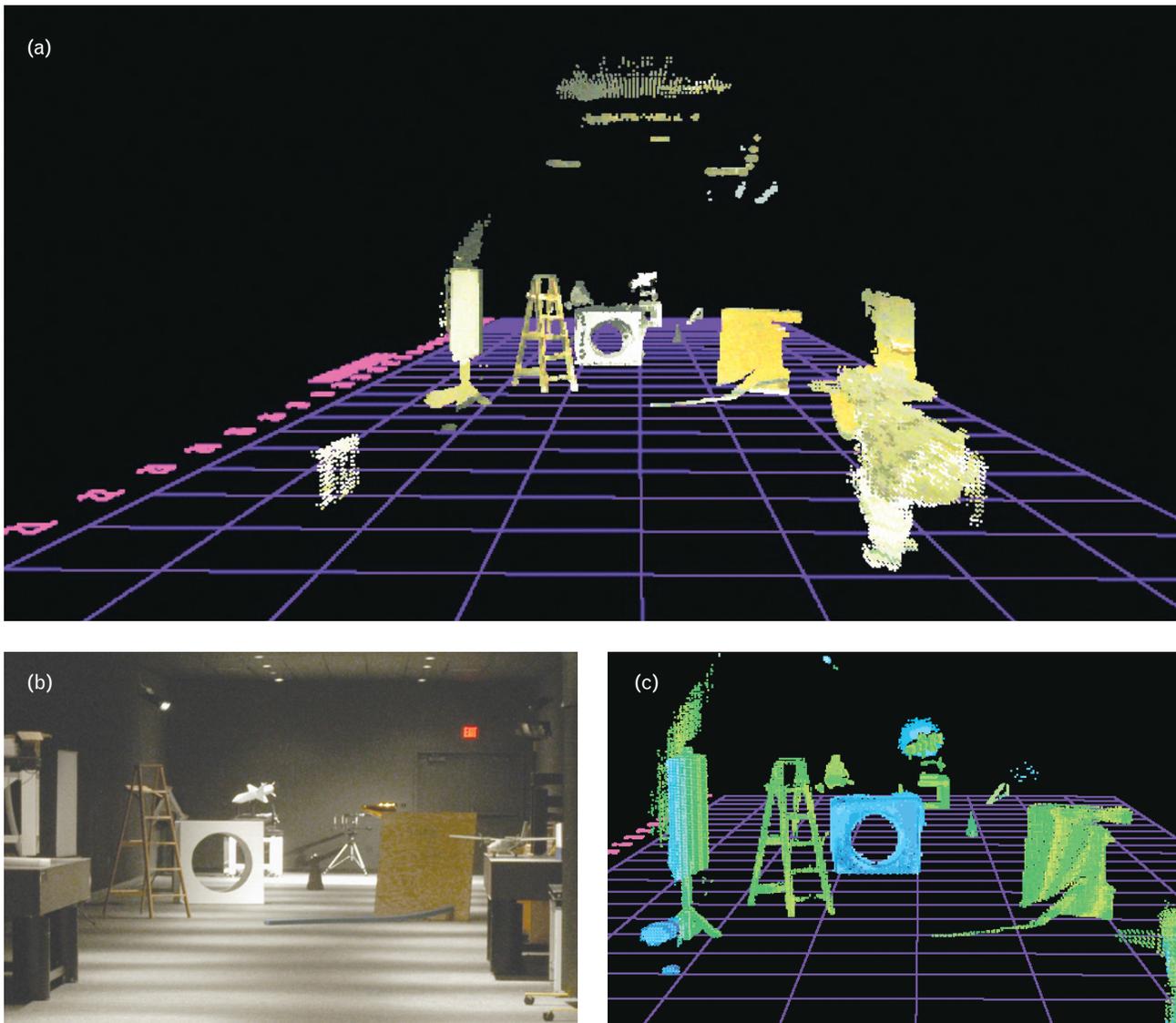


FIGURE 22. (a) Fused intensity and geometry image obtained after draping (b) a portion of the photograph in Figure 2 onto (c) the static voxel point cloud of Figure 14. Original colors have been brightened for display purposes.

2D intensity fields. Several investigators have focused in the past on multisensor imagery fusion in general and draping optical imagery onto ladar point clouds in particular [9–11]. Building upon these earlier efforts, we have developed programs that combine independent XYZ and RGB data sets. Figure 22 displays one example of fused ladar and video results. Transforming this off-line-generated proof of concept into a functioning real-time ladar/video system will require much more hardware and software development.

In closing, we list several potential applications of real-time 3D ladar imaging of national importance:

- Homeland defense (e.g., monitoring of sensitive site perimeters, ports of entry, and border crossings)
- Persistent urban surveillance (e.g., detecting, identifying, and tracking moving vehicles)
- Obscurant penetration (e.g., imaging tanks under trees)
- Robot navigation (e.g., planning routes through obstacle courses)
- Autonomous satellite operation (e.g., refueling, repairing, and upgrading spacecraft in orbit).

Future extensions of the work reported in this article can contribute to all these long-range applications by helping people and machines better understand complicated environments, maintain situational awareness, and monitor time-critical events.

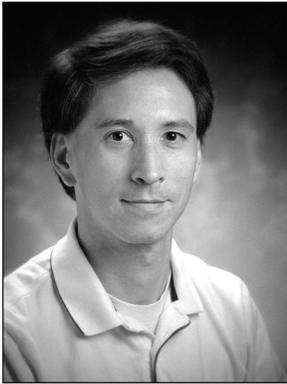
Acknowledgments

Many colleagues have generously shared their ideas, expertise, software, and data with us. We especially wish to thank the following Lincoln Laboratory members and groups for their help: Joseph Baldassini and the Space Control Systems group; Vadim Grinshpun, Vincenzo Sestito, and the Space Situational Awareness group; Bill Beavers, Mark Czerwinski, and the Aerospace Sensor Technology group; Robert Garnier, Richard Heinrichs, David Ireland, Leaf Jiang, Nirmal Keshava, Richard Marino, Michael O'Brien, Patrick Quinn, Gregory Rowe, David Schue, John Shelton, Luke Skelly, Ted Square, and the Active Optical Systems group; Robert Bond, Tim Currie, Hahn Kim, Andrew McCabe, Michael Moore, Albert Reuther, and the LLGrid team; Chet Beals, Dana Bille, Keith Ente, Donald Gaudette, Ned Rothstein, Scott Walton, and the Publications department.

- Target Detection and Recognition Using 3D Laser Radar Imagery," *Linc. Lab. J.* 15 (1), 2005, pp 61–78.
8. B.C. Matei, Y. Tan, H.S. Sawhney, and R. Kumar, "Rapid and Scalable 3D Object Recognition Using Lidar Data," *SPIE* 6234, 2006, in press.
 9. D.A. Fay, J.G. Verly, M.I. Braun, C.E. Frost, J.P. Racamoto, and A.M. Waxman, "Fusion of Multisensor Passive and Active 3D Imagery," *SPIE* 4363, 2001, pp. 219–230.
 10. A.M. Waxman, D.A. Fay, B.J. Rhodes, T.S. McKenna, R.T. Ivey, N.A. Bomberger, V.K. Bykoski and G.A. Carpenter, "Information Fusion for Image Analysis: Geospatial Foundation for Higher-Level Fusion," *Proc. 5th Int. Conf. on Information Fusion, Annapolis, Md., July 2002*.
 11. W.-Y. Zhao, D. Nistér, and S.C. Hsu, "Alignment of Continuous Video onto 3D Point Clouds," *Proc. 2004 IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition, 2, Washington, 27 June–2 July 2004*, pp. 964–971.

REFERENCES

1. B.F. Aull, A.H. Loomis, D.J. Young, R.M. Heinrichs, B.J. Felton, P.J. Daniels and D.J. Landers, "Geiger-Mode Avalanche Photodiodes for Three-Dimensional Imaging," *Linc. Lab. J.* 12 (2), 2000, pp. 383–396.
2. R.M. Marino, T. Stephens, R.E. Hatch, J.L. McLaughlin, J.G. Mooney, M.E. O'Brien, G.S. Rowe, J.S. Adams, L. Skelly, R.C. Knowlton, S.E. Forman and W.R. Davis, "A Compact 3D Imaging Laser Radar System Using Geiger-Mode APD Arrays: System and Measurements," *SPIE* 5086, 2003, pp. 1–15.
3. R.M. Marino and W.R. Davis, "Jigsaw: A Foliage-Penetrating 3D Imaging Laser Radar System," *Linc. Lab. J.* 15 (1), 2005, pp. 23–35.
4. See the GNU Compiler Collection home page at <http://gcc.gnu.org>.
5. A.I. Reuther, T. Currie, J. Kepner, H.G. Kim, A. McCabe, M.P. Moore and N. Travinin, "On-Demand Grid Computing with gridMatlab and pMatlab," *Proc. High Performance Computing Modernization Program Users Group Conf. 2004, Williamsburg, Va., 8 June 2004*.
6. E. Sobel, J. Douglas, and G. Ettinger, "Automated Identification and Classification of Land Vehicles in 3D LADAR Data," *SPIE* 5426, 2004, pp. 92–101.
7. A.N. Vasile and R.M. Marino, "Pose-Independent Automatic



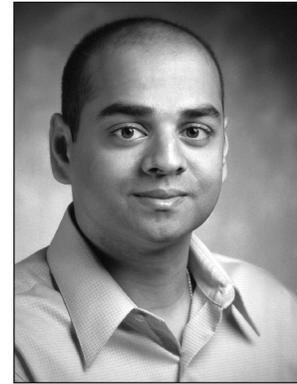
PETER CHO is a staff member in the Aerospace Sensor Technology group. He received a B.S. degree in physics from the California Institute of Technology (Caltech) in 1987 and a Ph.D. in theoretical particle physics from Harvard University in 1992. As a DuBridge Research Fellow in physics at Caltech from 1992–1996, he worked on diverse problems in high-energy physics. He returned to Harvard as a theoretical physics postdoctoral fellow in 1996. Since he joined Lincoln Laboratory in 1998, his research has spanned a broad range of subjects in machine intelligence, computer vision, and multisensor imagery fusion.



HYRUM ANDERSON is an associate staff member of the Aerospace Sensor Technology group where he researches image formation, fusion, and automated analysis of ladar, radar, and video imagery. Prior to joining the Laboratory in 2003, he studied microwave remote sensing at Brigham Young University, where he received B.S. and M.S. degrees in electrical engineering.



ROBERT HATCH is an associate staff member in the Active Optical Systems group. He is a 1998 summa cum laude graduate of Northeastern University, where he received a B.S. degree in engineering technology. His research interest is in the development of real-time data acquisition systems for focal plane arrays. He has been at the Laboratory since 1986.



PREM RAMASWAMI is an assistant staff member in the Active Optical Systems group. He received a B.S. degree in electrical and computer engineering and biomedical engineering from Carnegie Mellon University. His current area of research is in signal processing and system optimizations for airborne 3D laser radar applications. He is currently interested in pursuing graduate studies in the field of biomedical engineering with an emphasis in embedded system technology, medical devices, and rehabilitation engineering.