RSTA Research of the Colorado State, University of Massachusetts and Alliant Techsystems Team *

J. Ross Beveridge
Colorado State University
ross@cs.colostate.edu

Allen Hanson
University of Massachusetts
hanson@cs.umass.edu

Durga Panda
Alliant Techsystems Inc.
panda@rtc.atk.com

Abstract

The complementary nature of LADAR, FLIR and color data for ATR is being exploited by new algorithms in a three stage recognition system. The stages are initial detection, target class and pose hypothesis generation, and precise model to multisensor coregistration matching. Coregistration globally aligns 3D target models with range, IR and color imagery while simultaneously refining registration parameters between sensors. This model directed approach is expected to improve ATR performance for occluded targets, targets seen at unusual angles, and targets in cluttered settings.

Color is used for initial target detection under daylight conditions and camouflage learned from training generalizes across vehicles and distinguishes targets from natural terrain. Target class and pose hypothesis generation will draw upon existing LADAR boundary matching work extended to tolerate more occlusion, clutter and viewpoint variation. New model to multisensor coregistration algorithms appear robust in early tests and are the basis for future coregistration matching. A new interactive 3D visualization environment allows inspection of multisensor data, coregistration, and monitoring of recognition.

1 Introduction

Colorado State University, the University of Massachusetts and Alliant Techsystems are one of the teams working under BAA-9301 on the UGV Program's Reconnaissance, Surveillance and Target Acquisition (RSTA) effort. The project is led by Colorado State University and this paper introduces the project, describes its rationale and goals, and summarizes project activities and results to date.

Each team member brings particular expertise to the project and is working on particular subprojects. Alliant Techsystem's has state-of-the-art LADAR and FLIR based ATR capabilities which are being extended in this project to perform target class and pose hypothesis generation. The University of Massachusetts is using new machine learning techniques to perform initial target detection. Colorado State University, along with coordinating the project, is developing new geometrically precise techniques for matching 3D object models to multisensor data.

The rationale for this project lies in the complementary nature of range, IR and color data, and the conviction that exploiting this complementarity will enable automated systems to perform reliably under adverse operational conditions. The project goal is to develop 3D model-based techniques for integrating range, IR and color data and demonstrate that these 3D model-based techniques improve ATR performance under adverse conditions. The project rationale and goals are expanded upon in Sections 2 and 3.

Adverse conditions include occluded targets, targets viewed against structured or cluttered backgrounds, targets viewed at uncharacteristic angles, and objects viewed under unusual or unfavorable thermal conditions. To test the utility of combined LADAR, FLIR and color recognition, all three forms of data had to be collected under a RSTA like scenario. In conjunction with Martin Marietta's RSTA team, we collected over 400 range, IR and color images at Fort Carson, Colorado in November 1993.

To support algorithm development and evaluation a recognition testbed is being assembled. The testbed will integrate modules performing the three basic subtasks associated with our approach to recognition: 1) detection, 2) pose and class hypothesis generation, and 3) multisensor matching consistent with global geometric constraints implied by known sensor and object geometry. An overview of the recognition testbed is presented in Section 4. The first integrated tests of the recognition testbed are expected in Fall 1994. Testing on major components has already begun and progress on these efforts are summarized in Sections 5.

2 Project Rationale

Existing object recognition systems neglect the complementary nature of LADAR, FLIR and color CCD imagery. While LADAR provides direct but often noisy

*This work was sponsored by the Advanced Research Projects Agency (ARPA) under grant DAAH04-93-G-422, monitored by the U. S. Army Research Office.

3-D surface information, FLIR and color CCD data provides surface appearance and surface boundary information distorted by perspective projection. To fully exploit this complementarity, the data from these three sensors must be precisely registered with respect each other and with respect to the hypothesized object in the scene. The model-based approach of this project will use known object model and sensor geometry to constrain registration between sensors and between sensors and the object. This combined sensor and object registration is here called *coregistration*.

Traditionally, sensor registration has been treated separately from recognition. First sensor data is registered using only very low-level correlation information, and then a model is matched to this data. Typically matching uses image space templates and variations in viewpoint are d by multiple templates. These approaches fail to exploit the rich 3D geometric constraints derivable from object model and sensor geometry. Under adverse conditions, templates can proliferate and low-level sensor registration errors can cause recognition to fail.

The recognition algorithms being developed for this project will simultaneously coregister 3D models to LADAR, FLIR and Color CCD as part of the matching/recognition process. Coregistration computes the best-fit pose of a 3D object model relative to multiple sensors and includes free variables to allow for refinement of sensor-to-sensor registration as well as the 6 pose parameters defining the position and orientation of the object relative to the sensor suite.

Relative to current approaches to ATR, the benefits of our coregistration approach include freedom from multiple 2D templates of 3D objects and correspondingly fewer restrictions upon possible viewing angles. Sensor fusion driven by model-based geometric constraints has the added benefit that it provides a precise mapping between sensor readings and model features which is refined based upon constraints derived from the object model.

Using coregistration, reliable identification should be possible under conditions for which a single sensor is inadequate. Moreover, the precise mappings between sensor readings and model features can be expected to permit better discrimination than is possible when multisensor data is first analyzed independently and then fused using algebraic techniques. One reason for this is that independent techniques demand sensor data already be accurately registered, and will err if this registration is not correct.

Coregistration matching introduces a computational demand not present for simpler 2D template approaches. It is therefore important to intelligently couple coregistration matching with a hypothesis generation capability which focuses attention upon specific regions of interest where objects of interest are likely to appear. For this reason, developing sensor hypothesis generation capabilities using IR, color, and range is an important part of this project.

For hypothesis generation, as well as final object discrimination, color is a useful adjunct to FLIR. Color has been neglected in ATR, and a thorough integration of color with IR and range data is a significant part of our effort. Tests have begun of new machine learning techniques to distinguish military camouflage from natural terrain in data collected last fall at Fort Carson in Colorado.

3 Project Goal

Our goal is to develop a multisensor coregistration ATR system which solves recognition problems essentially impossible to solve using only a single sensor. Appearance and 2D edge constraints from FLIR and color sensors, 3D edge and surface constraints from LADAR, and surface appearance and geometric constraints derived from BRL-CAD models are all brought together through model feature and sensor data coregistration. This coregistration globally aligns model features to sensor features as part of a matching process aimed at finding the most globally consistent correspondence between object model and sensor features. Our overall approach, highlighting the coupling through coregistration, is illustrated in Figure 1.

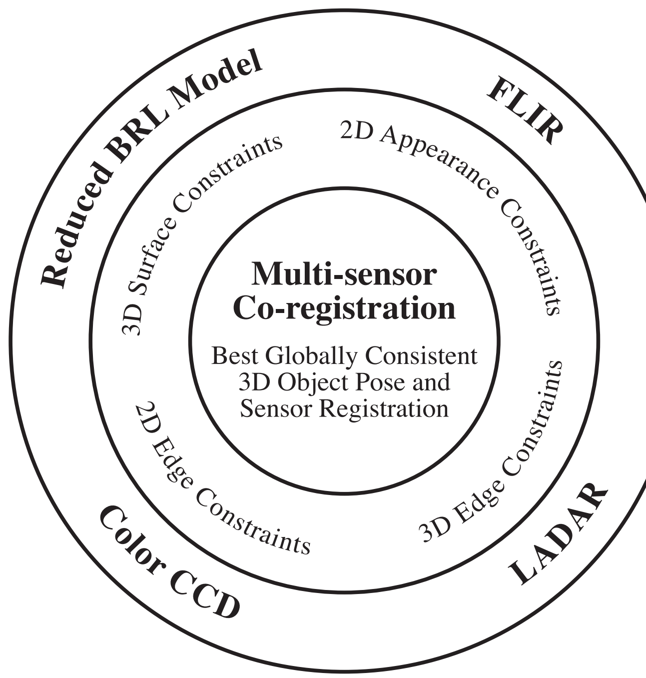

Figure 1: Multisensor and 3D object model constraints combined through coregistration matching.

The new techniques and algorithms required for ATR have much broader applicability, and a longer range goal is to advance the underlying basis of multisensor object recognition. Specifically, the mathematical basis for coregistration matching is applicable to any multisensor recognition task involving precisely modeled objects and calibrated sensors.

A single global constraint equation expressed as a sum-of-squared errors between corresponding 3D object model and sensor features is at the heart of coregistration. The free variables signify the 3D pose of the object relative to one sensor and the registration mapping between that particular sensor and the other sensors.

A second major subgoal is match evaluation metrics which exploit coregistration constraints. The most obvious constraint is the quality of the global least-squares fit itself. Equally important is omission: are significant portions of the model missing from the data. More sophisticated dependencies include compensation for occlusion.

While occlusion is essentially impossible to detect directly in FLIR or color data, it can be detected in LADAR by testing for significant discontinuities in range. Consequently, match evaluation can differentiate between missing versus occluded FLIR or color features based upon whether or not LADAR finds evidence of occlusion. From the standpoint of match quality, the two cases are quite different. Occluded features are to be expected while missing features are a cause for concern and reduced match confidence.

4 The Recognition Testbed

To integrate the software components being produced by different subprojects, a Recognition Testbed is being constructed. A summary diagram of this system is shown in Figure 2. The inputs to the run-time system are sensor images from FLIR, LADAR and color sensors. Additional inputs come from off-line components which provide vehicle model information, terrain model information and decision-trees used in FLIR and color based hypothesis generation. The 3D vehicle models are reduced from their full BRL-CAD detail to simpler 3D representations appropriate for matching. When available, terrain models will also be registered to imagery and used to provide range estimates to points in IR and color imagery.

One of our challenges is to adapt current state-of-the-art FLIR and LADAR techniques to fit our new approach. Existing software components, some originally designed to perform ATR tasks on their own, are now being put into service as components of our larger coregistration matching system. We are utilizing current system components in two basic ways. First, with modification, current FLIR and LADAR techniques promise to provide a strong hypothesis generation front-end for our coregistration matching system. Second, the statistical error models and match evaluation scores developed for separate FLIR and LADAR systems can provide a principled multisensor match evaluation measure.

5 Activities and Subprojects

The first major activity carried out in this project was the LADAR, FLIR and color data collection mounted last Fall jointly with Martin Marietta at Fort Carson. Prior to this collection, no suitable data existed for testing our new LADAR, FLIR and color based approach. The result of this data collection effort is over 400 LADAR, FLIR and color images of four different military vehicles. Vehicles are both out in the open and terrain occluded. This activity is described further in Section 5.1 and [Beveridge *et al.*, 1994].

Another major activity is to explore the use of color as an alternative to FLIR for daytime target detection. Results on Fort Carson data show that multivariate decision tree learning techniques can discriminate camouflage from outwardly similar terrain. The current work uses non-parametric combination of the basic red, green and blue values of the color signal and realtime detection can be supported by encoding the decision tree as a lookup table. This subproject is described further in Section 5.3 and elsewhere in these proceedings [Buluswar *et al.*, 1994].

Past work of our team has shown that bounding contours in LADAR can be used to perform recognition. Adapting and extending this work to perform target class and pose hypothesis generation is now a major project activity. Hypothesis generation places different demands on boundary contour matching. It makes the problem simpler in that it is no longer critical to avoid suggesting false hypotheses. However, the goal of extending performance to problems involving significant occlusion, odd viewpoints, and cluttered backgrounds, demands that the contour matching cover a much broader range of possible contours and fractions of contours. This subproject is described further in Section 5.4.

New 3D object model to multisensor coregistration algorithms are a keystone of this project. These coregistration algorithms will be used for matching models to multisensor data in a way which lets global model and sensor geometry constrain the relationship between matched features. Measures of match quality will thus account for global geometric consistency. Our past work on model matching [Beveridge, 1993] has emphasized matching subject to global geometric consistency as enforced by a single least-squares constraint between all model and corresponding image features. In early versions of this work [Beveridge *et al.*, 1991; Beveridge *et al.*, 1990], essentially 2D models were fit to 2D image data subject to a single best-fit 2D similarity transformation. Later versions utilized the 3D sensor pose work of Kumar [Kumar, 1989; Kumar, 1992; Kumar and Hanson, 1994] to perform fitting of 3D object models to corresponding features in a 2D image [Beveridge and Riseman, 1992; Beveridge and Riseman, 1994]. Currently the work of Kumar is being extended to handle both registration between multiple sensors as well as 3D pose between the sensors and object model. The resulting least-squares fitting procedure is what we have chosen to call 'coregistration'. Coregistration is summarized in Section 5.2 and reported in more detail elsewhere in these proceedings [Schwickerath and Beveridge, 1994].

To work with range, IR and color data, particularly since the geometric relationship between data and model is a primary concern, it is has been crucial to develop tools which permit these relationships to be visualized. Consequently, over the past six months a new interactive 3D visualization environment called 'RangeView' has been developed. LADAR range data and 3D object models are embedded in the same 3D space which a user can view from any vantage point. Either color or IR data

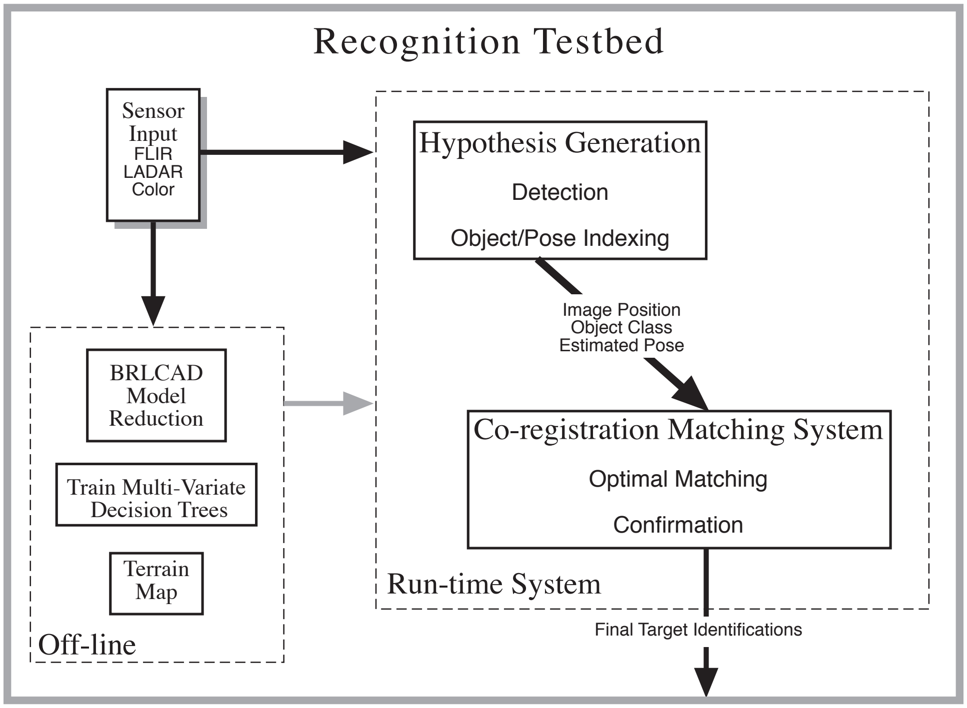

Figure 2: Overview of Recognition Testbed. The three sections, hypothesis generation, coregistration matching, and off-line components reflect the breakdown of our project into major focus areas.

may be overlaid on top of the range data. Interactive tools permit hand registration by selected control points. More interestingly, selection of corresponding points on object models and in multiple images permits the coregistration algorithm to be used to generate both the best placement of the object and the associated best registration between images. This tool is further described in Section 5.5 and elsewhere in these proceedings [Goss *et al.*, 1994].

5.1 Fort Carson Data Collection

A data collection effort was mounted by Martin Marietta, Colorado State University, and Alliant Techsystems in order to meet immediate needs of the the RSTA effort. The collection took place in the first week in November 1993 at Fort Carson. The Fort Carson Colorado Army National Guard Depot made several vehicles available and provided drivers who placed the vehicles on the National Guard test range.

The data collection effort was highly constrained in terms of time, resources, vehicles and terrain. This limited the amount of data and ground truth information which could be collected. These limitations not withstanding, over 400 LADAR, FLIR and color images were collected. There is a 50 page report [Beveridge *et al.*, 1994] describing each image, vehicle array, and ancillary information such as time of day and weather conditions. It also includes a copy of the original data collection plan and examples of the data. Members of the BAA-9301 RSTA community can acquire the data from the Martin Marietta FTP server.

The Fort Carson data meets all the basic minimum requirements to permit algorithm testing under conditions set forth in the RSTA scenario specified in the original BAA-9301. Specifically, it includes FLIR, LADAR and color imagery for military vehicles positioned in natural terrain. The Alliant Techsystems LADAR generates 24 by 120 pixels with a 3 by 5 degree field of view. To simulate the nominal 1 foot per pixel range called for in the planned RSTA LADAR, vehicles were placed about 400 feet from the sensors at Fort Carson. Modestly wide angle lens were used with the FLIR and color cameras so that 'pixels on target' values for these sensors would also be comparable to those expected in the 0.5 to 1.0 kilometer range with the RSTA sensor suite.

The Fort Carson data contains 39 distinct LADAR, FLIR and color sensor triples obtained from 10 different vehicles arrays. The different vehicles arrays include vehicles in full view, partially occluded and at odd angles to the sensors. The data collected is now being widely used by the most of the RSTA BAA-9301 co-contractors.

5.2 Model to Multisensor Coregistration

Appearance of 3D object models varies with viewpoint, and pixels from multiple sensors typically are not in one-to-one correspondence. Knowledge of sensor parameters and relative sensor positions provides moderately accurate estimates of the pixel to pixel registration. However, small variations in relative sensor position can lead to significant misregistration between pixels. This is of concern when matching objects which are small in terms of absolute image size. To get around registration problems and 3D variations in appearance, ATR systems commonly assume sensor registration is exactly known or determined using low-level correlation. Variation in 3D appearance is typically accounted for by sampling expected viewpoints to produce a set of templates represented in image space.

Rather than assume perfect registration obtained prior to matching object models, and rather than build a suite of viewpoint specific templates, this project is developing new methods for simultaneously computing the best-fit pose of a 3D object model relative to multiple sensors. This fitting process includes free variables to allow for refinement of sensor-to-sensor registration as well as the 6 pose parameters defining the position and orientation of the object relative to the sensor. This coregistration fitting process is based upon extending earlier single sensor pose work performed by Kumar [Kumar, 1992; Kumar and Hanson, 1994]. Kumar proposed a set of equations for computing the "best" pose, given a matching between model edges and image edges of a single image. Here this formulation is broadened to include corresponding point and edge features in FLIR and color images as well as range images.

Partial constraints between sensors which limit movement of one relative to others make coregistration an interesting problem. A family of coregistration algorithms can be defined with each individual algorithm arising out of different assumptions about relative sensor movement. At one extreme, if sensors are assumed to be perfectly registered, then coregistration devolves into sensor to object pose computation. At the other extreme, if sensors move freely and independently, then there is no coupling and the result is an independent sensor pose problem for each sensor. The interesting problems lie between these two extremes.

The first coregistration algorithm developed under this project assumes that a LADAR and CCD (or FLIR) sensor are constrained to the same relative orientation, but can translate one relative to the other in a common image plane. The setup is illustrated in Figure 3. Figure 3a illustrates the 3D geometry of an object, a FLIR or CCD sensor and a LADAR sensor. The sensors, together, are free to rotate and translate relative to the object. The sensors are constrained so as to permit only translation in a common image plane. These 3D constraints permits translation of FLIR or color images relative to LADAR images as illustrated in Figure 3b.

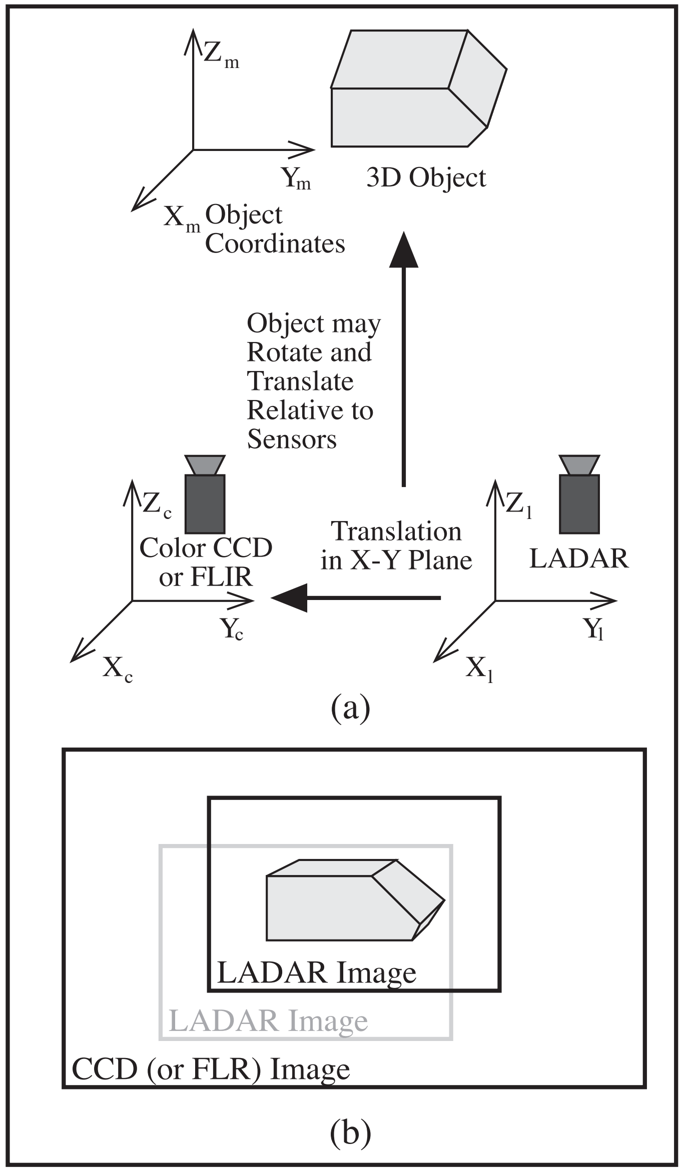

Figure 3: Coregistration with sensor-to-sensor planar translation. a) 3D scene geometry and transformation constraints, b) resulting allowed pixel translation between images.

The image translation case approximates what can be expected in the RSTA sensor suite. Sensors are bolted to the same platform and their relative image registration mappings are known to within several pixels. However, very small rotations due to vibration will induce small motions (perhaps 1 to 5 pixel) in one image relative to the other. Given that objects at a distance are being viewed, these small sensor rotations may be approximated as small image translations.

The image translation coregistration work is described in detail elsewhere in these proceedings [Schwickerath and Beveridge, 1994]. Experiments have been conducted using synthetic data to test the robustness of the algorithm. It has also been used to perform coregistration on LADAR and CCD data collected at Fort Carson. Three batteries of experiments have been conducted: A) sensitivity to noise in initial coregistration estimate, B) sensitivity to noisy image data, C) noisy data and noisy initial coregistration estimates. Four different model feature configurations were used to simulate the relative model and sensor geometry associated with the planned RSTA sensor suite viewing vehicles at half a kilometer.

The coregistration algorithm almost always converges to the correct 8 coregistration parameters when initial estimates are within about 45 degrees of the true orientation and error in translation are not more than 50% off from the true values. Unusual geometric configurations can cause the algorithm to converge to other local minima on the coregistration error surface. However, this is not common, and the convergence behavior is more than adequate to support the needs of coregistration matching under the RSTA scenario. Initial pose estimates generated using LADAR contours should be accurate to within 10 to 20 degrees, if not better. Sensor registration will be known to within a quite small uncertainty.

Adding Gaussian noise to image measurements with $\sigma = 1$ (pixels) leads to recovered object pose estimates with roughly 0.01 radians of rotational error and about 5 meters of translational error for objects at 500 meters. For $\sigma = 5$, these errors rise to about 0.1 radians and 30 meters. Adding noise to the initial coregistration estimates as well as the image data increases the average error on the recovered coregistration parameters by about an order of magnitude. However, the averaging is somewhat misleading because 100 trials with independent samplings of the noise process are used and, while most trials converge to the correct parameters, a few converge to incorrect locally optimal parameter sets far off from the true answer.

The next coregistration variant to be developed will assume a fixed distance between sensors and unconstrained relative orientation. This algorithm has the potential of performing coregistration between sensors operating on two separate vehicles. It is reasonable to assume two GPS equipped vehicles at 250 to 1000 meter separation will have essentially known relative translation. However, relative orientation may be off by as much as 5 degrees and need correction.

The reason for constructing these multisensor coregistration algorithms is to support coregistration matching. Development of new search procedures for finding optimal correspondences between object model and sensor features using coregistration to determine the globally best geometric configuration is just beginning. Two approaches are being explored. One repeatedly matches features to their nearest neighbors in object/sensor space and has been used successfully to match LADAR data [Bevington *et al.*, 1992; Bevington, 1992]. The other is to extend our local search matching work previously developed for single sensor problems [Beveridge, 1993; Beveridge and Riseman, 1994]. These local search algorithms explore perturbations in correspondence space. They require more computation than the nearest neighbor techniques. However, they also find optimal solutions for a wider range of initial conditions.

5.3 Color Target Detection

There are three basic reasons why it is important to study the use of color as a feature for performing target detection. First, new machine learning techniques have been developed [Brodley and Utgoff, 1994] which learn non-parametric combinations of features and they have been successfully applied to difficult pixel classification tasks [Draper *et al.*, 1994] under varied lighting conditions. Second, although most RSTA activities are conducted at night, for those conducted during daylight, it makes no sense to disregard visible light. Moreover, one can argue color is most useful precisely when FLIR is least useful. As illustrated by the conditions for the UGV Demo B, hot summer sun heats terrain features and gives off significant solar reflection for the 3 to 5 micron RSTA FLIR. Under such conditions, color is probably more distinctive and more stable. Finally, granting that RSTA should be conducted at night, the underlying mechanisms being developed for color detection will transfer to multiband IR sensors as these sensors become available.

Color detection begins by classifying each pixel in the image as belonging to one of two classes – target or non-target. This determination is based upon non-parametric combinations of red, green and blue values learned by a multivariate decision tree learning procedure described below and in more detail elsewhere in these proceedings [Buluswar *et al.*, 1994]. Following pixel classification, morphological techniques are used to discard single or several pixel target responses and leave target regions. Preliminary tests have shown generalization across vehicles and across lighting conditions, specifically direct sun versus overcast.

To perform the pixel classification, each pixel is described by a linear combination (weighted sum) of features and the weights associated with each feature (in this case, the red, green and blue color values). This linear combination of the features and weights is known as a Linear Threshold Unit (LTU), and serves as a multivariate decision criterion. If the LTU for a given pixel is positive, it is labeled as a positive candidate, otherwise it is considered a non-target. The weights for each feature are learned using the Recursive Least Squares (RLS) algorithm over a given set of training pixels. The training pixels are a hand labeled subset of the total image pixels. Graphical user interface tools make selection of training pixels by area a relatively simple procedure. In the color target detection currently being studied there are only two class labels: target and non-target.

Once the weights are learned, the LTU creates a hyperplane that separates the set of instances. If a set of instances is linearly separable, a single LTU test is sufficient for separation. Typically, however, a large set of instances in complex scenes is seldom linearly separable. In order to classify such instances a decision tree must be built, which recursively creates hyperplanes of different orientations that separate the non-homogeneous partitions resulting from a previous attempt at separation. A homogeneous partition resulting from a separation is, by definition, non-separable, and is labeled as a class. A decision tree therefore consists of nodes that are either decisions or classes.

Based on promising results from the approach described, we believe that color, even in complex outdoor scenes, can be used as an informative feature. Early results are presented in Figure 4. A multi-variate decision tree was trained to recognize the M60 coloring in an image similar to that shown in the upper image. This tree was then used to classify *all* pixels in both images shown. A simple morphology operator was used to remove stay target classified pixels and the result was that only pixel truely on target were labeled as target pixels.

The color detection generates four regions of interest as indicated by the rectangular boxes. To better pre-

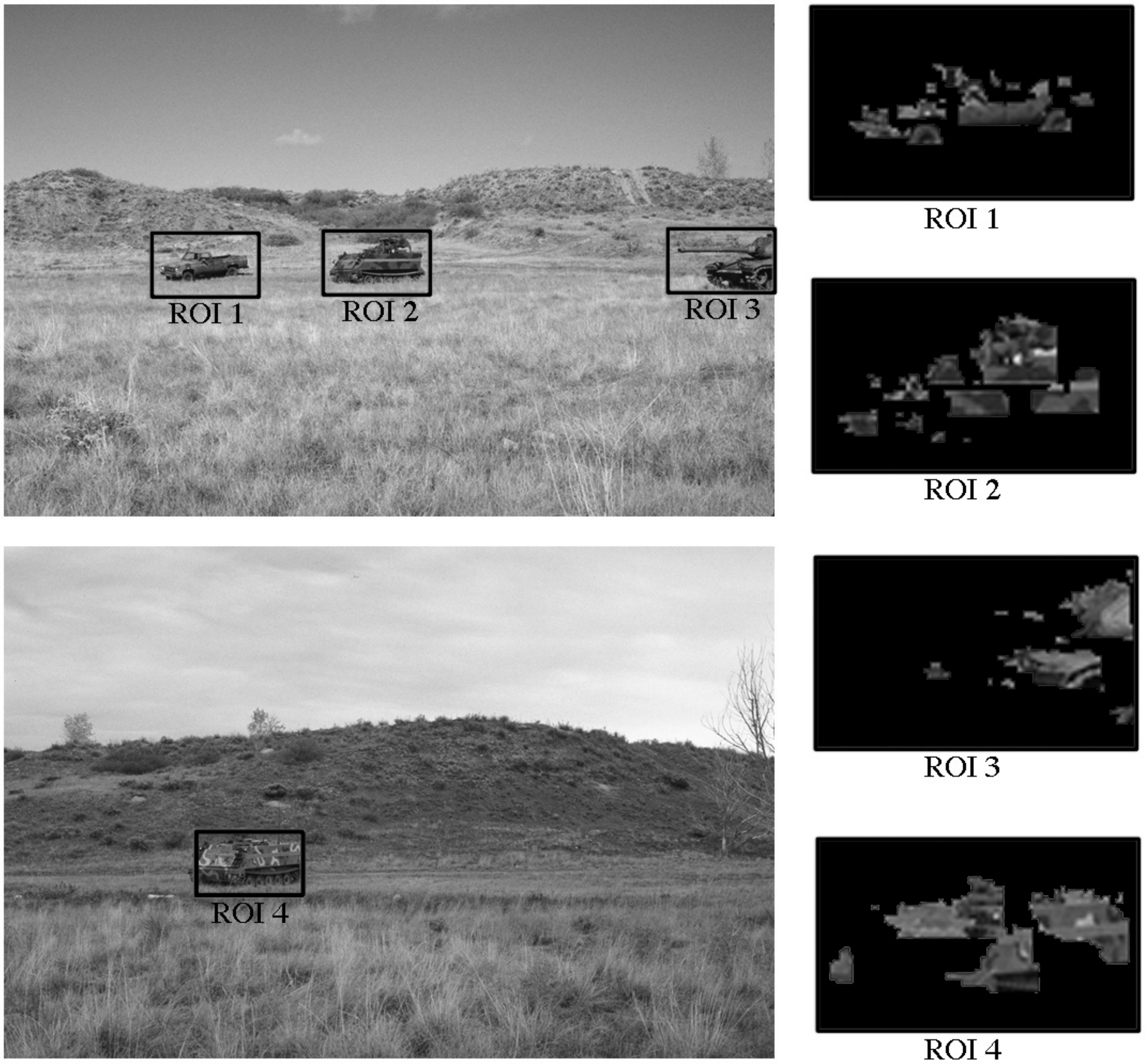

Figure 4: Example of color ROI detection.

cisely which pixels on the targets are being identified, these ROIs are expanded on the right and only those pixels selected by the detection procedure are shown. The absence of false positives in this early test is encouraging. It is all the more so considering the lower image was taken on a different day, under cloudy rather than sunny conditions, and shows a different vehicle with somewhat lighter camouflage.

The performance of any feature-based classifier will depend, to a large extent, on the set of features used, the methods used to extract those features, and the classifier used. This project is studying established techniques for the above, as well as experiment with new ideas, and in the process, do a comparative study of the performance of several well-known classifiers in the context of ATR. Neural networks, instance-based classifiers, univariate decision trees, minimum-distance classifiers, genetic algorithms, multiple-class classifiers, etc., will be used with features ranging from low-level pixel data (R-G-B, H-S-V, etc.), to higher level spatial and regional characteristics, such as texture, fractal dimension, context, etc. In addition, we shall attempt to integrate image data with that from other sensors such as FLIR in order to reinforce the accuracy of the regions extracted.

5.4 LADAR Boundary Contours for Hypothesis Generation

Alliant Techsystem's LADAR Recognition System (LARS) has demonstrated state-of-the-art target identification performance of hundreds of frames of both real and synthetic imagery. The LARS suite, summarized in Figure 5, uses a non-segmenting model-based approach, which efficiently exploits both the 2-D (boundary matching) and 3-D (surface matching) shape information contained in LADAR signatures. Templates are derived from BRL models of the expected target set. No training imagery is required. Since LARS does not perform segmentation, it avoids information loss and provides robust performance in low SNR scenarios, an important consideration for low LADAR visibility conditions. Under the LARAA program, LARS consistently attained target identification performance in the mid-to-upper 90%

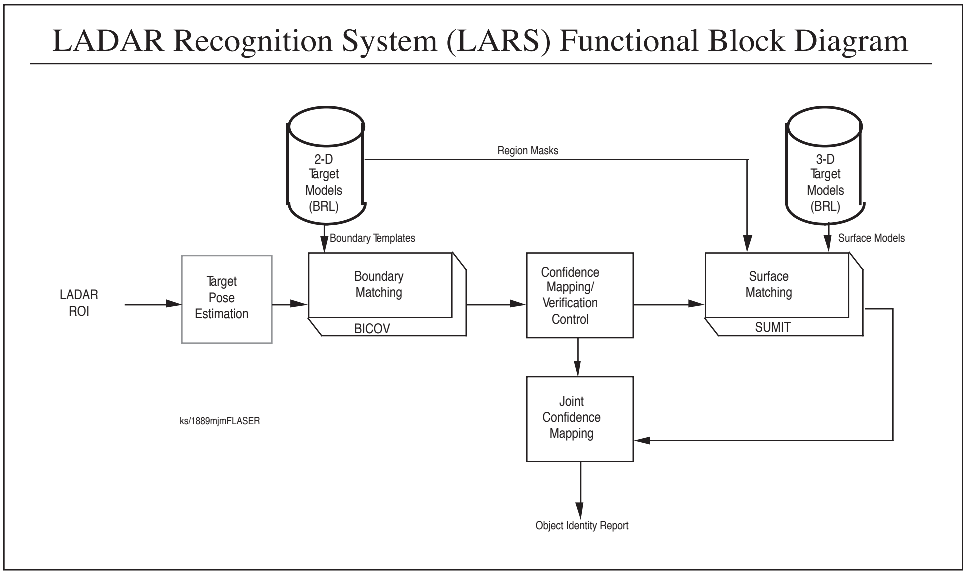

Figure 5: Block diagram of Alliant Techsystems LADAR Recognition System.

range.

LARS operates on individual absolute range images corresponding to pre-cued ROIs. The output of LARS is an ordered list of the most likely target hypotheses at a specific pose, paired with a likelihood ratio confidence. A LADAR-only recognition result could be obtained by selecting the highest confidence hypothesis.

As shown in Figure 5, LARS processes the 2D and 3D signature information separately. First, the 2D boundary matching process, based on a pixel probing approach called BICOV (Boundary Interval Coincidence Verification), exploits the target shape information present in the range discontinuity along the target boundary. This generates a preliminary target class/pose hypothesis list. A subset of the best matches are used to drive the 3D surface matcher (known as SUMMIT), which exploits the topography of the targets surface. The separation of matching stages is done mainly in the interest of computational efficiency. A-priori knowledge of target class and aspect (as provided by the boundary matcher) greatly constrains the 3D surface matcher search space and simplifies the SUMMIT algorithm complexity as well.

A certainty accrual mechanism is used to combine the BICOV and SUMMIT match scores. This level of fusion does not take account of co-constraint between the boundary and surface matching steps. However, such co-constraints will be exploited by subsequent coregistration matching. It is the goal of this subproject to test and extend the boundary matching capabilities of BICOV as a means of efficiently generating the initial object and class pose hypotheses. This is a different use of the LADAR matching from that for which it was initially designed and optimized. Adjusting and modifying the BICOV process to better meet the needs of hypothesis generation is the principle objective of this subproject.

5.5 RangeView: A 3D Model to Multisensor Data Visualization System

To is difficult to fully understand multisensor data, particularly when one of the sensors is a ranging device, without the aid of some form of visualization system. The norm currently is to display range data as though it were an image with values equalling depth, or as a 3D smooth surface viewed from a fixed vantage point. Neither of these is adequate for fully realizing the 3D character of the data, and even less so for understanding the 3D relationships between this data and a 3D model or between this 3D data and other sensors such as IR and color cameras.

To support our work on multisensor recognition in general and coregistration in particular, it is essential to be able to visualize 3D models and data in proper relation to each other. It likewise important when developing recognition algorithms to be able to monitor the progress of the recognizer while it is finding an object, and to verify the correctness of the results. Monitoring allows a user to interactively inspect in 3D both intermediate states as well as final recognized models. In particular, it permits one to assess the geometric plausibility of the object pose and sensor registration produced by the recognizer. A prototype system, called RangeView, has been developed for realtime 3D display and interaction with both multisensor data and 3D object models.

Rangeview combines range imagery, color imagery, thermal (infrared) imagery, and CAD models of objects to be recognized. Range imagery is used to create a partial three-dimensional representation of a scene. Optical imagery is mapped onto this partial 3D representation. In its role as the graphical users interface to a coregistration matching system, coregistered object models will be displayed in spatial relation to the scene, and the registered scene and object may be visually inspected from any viewpoint. The system also allows the user to manually register range data with optical and CAD model data for use in evaluating the fidelity between models and data as well as the results of automatic recognition. Representations of all three types of sensor data can be displayed simultaneously with model data in the same 3D coordinate space.

Interactive 3D inspection is supported using a high speed graphics workstation. The operator is provided with a set of controls which allow 3D data to be viewed from any angle and at any desired level of magnification. Animation is also provided for multiple frame image sequences. Multiple views from the same or different source imagery can be displayed simultaneously.

We foresee eventual use of this technology in a fielded system for operator verification of Automatic Target Recognition (ATR) results. Display of a target model registered to a partial 3D scene could allow an operator to verify a target with great confidence. In time-critical applications, multiple simultaneous views could be scanned rapidly by the operator.

References

[Beveridge and Riseman, 1992] J. Ross Beveridge and Edward M. Riseman. Hybrid Weak-Perspective and Full-Perspective Matching. In *Proceedings: IEEE 1992 Computer Society Conference on Computer Vision and Pattern Recognition*, pages 432 – 438. IEEE Computer Society, June 1992.

[Beveridge and Riseman, 1994] J. Ross Beveridge and Edward M. Riseman. Optimal Geometric Model Matching Under Full 3D Perspective. In *Second CAD-Based Vision Workshop*, pages 54 – 63. IEEE Computer Society Press, February 1994. (Submitted to CVGIP-IU).

[Beveridge *et al.*, 1990] J. Ross Beveridge, Rich Weiss, and Edward M. Riseman. Combinatorial Optimization Applied to Variable Scale 2D Model Matching. In *Proceedings of the IEEE International Conference on Pattern Recognition 1990, Atlantic City*, pages 18 – 23. IEEE, June 1990.

[Beveridge *et al.*, 1991] J. Ross Beveridge, Rich Weiss, and Edward M. Riseman. Optimization of 2-Dimensional Model Matching. In Hatem Nasr, editor, *Selected Papers on Automatic Object Recognition (originally appeared in DARPA Image Understanding Workshop, 1989)*, SPIE Milestone Series. SPIE, Bellingham, WA, 1991.

[Beveridge *et al.*, 1994] J. Ross Beveridge, Durga P. Panda, and Theodore Yachik. November 1993 Fort Carson RSTA Data Collection Final Report. Technical Report CSS-94-118, Colorado State University, Fort Collins, CO, January 1994.

[Beveridge, 1993] J. Ross Beveridge. *Local Search Algorithms for Geometric Obejct Recognition: Optimal Correspondence and Pose.* PhD thesis, University of Massachuesetts at Amherst, May 1993.

[Bevington *et al.*, 1992] James Bevington, Randy Johnston, Joel Lee, and Richard Peters. A Modular Target Recognition Algorithm for LADAR. In *Proc of the 2nd Automatic Target Recognizer Systems and Technology Conference*, pages 91 – 104, Fort Belvoir, VA, March 1992.

[Bevington, 1992] James Bevington. Laser Radar ATR Algorithms, Phase III Final Report. Technical report, Alliant Techsystems, May 1992. Prepared under CC-NVEO Contract No. DAAB07-87-C-F109.

[Brodley and Utgoff, 1994] C. E. Brodley and P. E. Utgoff. Goal-directed Classification Using Linear Machine Decision Trees. *Machine Learning*, page (to appear), 1994.

[Buluswar *et al.*, 1994] Shashi Buluswar, Bruce A. Draper, Allen Hanson, and Edward Riseman. Nonparametric Classification of Pixels Under Varying Outdoor Illumination. In *Proceedings: Image Understanding Workshop*, page (to appear), Los Altos, CA, November 1994. ARPA, Morgan Kaufmann.

[Draper *et al.*, 1994] B. Draper, C. E. Brodley, and P. Utgoff. Goal-directed Classification Using Linear Machine Decision Trees. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(9):(to appear), September 1994.

[Goss *et al.*, 1994] Michael E. Goss, J. Ross Beveridge, Mark Stevens, and Aaron Fuegi. Visualization and Verification of Automatic Target Recognition Results Using Combined Range and Optical Imagery. In *Proceedings: Image Understanding Workshop*, page (to appear), Los Altos, CA, November 1994. ARPA, Morgan Kaufmann.

[Kumar and Hanson, 1994] Rakesh Kumar and Allen R. Hanson. Robust Methods for Estimating Pose and a Sensitivity Analysis. *CVGIP:Image Understanding*, 11:(to appear in November), 1994.

[Kumar, 1989] Rakesh Kumar. Determination of Camera Location and Orientation. In *Proceedings: Image Understanding Workshop*, pages 870 – 881, Los Altos, CA, June 1989. DARPA, Morgan Kaufmann Publishers, Inc.

[Kumar, 1992] Rakesh Kumar. *Model Dependent Inference of 3D Information From a Sequence of 2D Images.* PhD thesis, University of Massachusetts, Amherst, February 1992.

[Schwickerath and Beveridge, 1994] Anthony N. A. Schwickerath and J. Ross Beveridge. Model to Multisensor Coregistration with Eight Degrees of Freedom. In *Proceedings: Image Understanding Workshop*, page (to appear), Los Altos, CA, November 1994. ARPA, Morgan Kaufmann.