Computer Science
Technical Report

Colorado State University

Progress on Target and Terrain Recognition Research at Colorado State University.*

J. Ross Beveridge, Bruce A. Draper and Kris Siejko

December 16, 1995

Technical Report CS-96-106

Computer Science Department
Colorado State University
Fort Collins, CO 80523-1873

Phone: (970) 491-5792 Fax: (970) 491-2466
WWW: http://www.cs.colostate.edu

*This work was sponsored by the Advanced Research Projects Agency (ARPA) under grants DAAH04-93-G-422 and DAAH04-95-1-0447, monitored by the U. S. Army Research Office.

Progress on Target and Terrain Recognition Research at Colorado State University. *†

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

Bruce A. Draper
University of Massachusetts
bdraper@cs.umass.edu

Kris Siejko
Alliant Techsystems Inc.
siejko@rtc.atk.com

Abstract

A target recognition capability is described which performs: color detection, target type and pose hypothesis generation, and target verification by 3D alignment of target models to range and optical imagery. The term 'coregistration' is introduced to describe target, range and optical sensor alignment. The following key verification components are described and demonstrated: target-model feature extraction, model-driven edge detection, range, optical and target coregistration, and coregistration space matching. As a precursor to future incorporation of terrain data, the ability to match terrain features to imagery from the UGV Demo C test site is demonstrated.

1 Introduction

Our goal is the development of new Automatic Target Recognition (ATR) algorithms which are more robust with respect to scene clutter, target occlusion and variations in viewing angle. The heart of our approach is to fuse range and optical imagery, color or IR, using global geometric constraints. Constraints derive from known sensor, target and scene geometry. This may be thought of as model-based sensor fusion, and contrasts with more traditional approaches which attempt to fuse data based upon low-level queues only [EG92].

The roots of our approach lie in past alignment based object recognition research [Low85, HU90, BR95] which has demonstrated the value of algorithms which precisely vary 3D object to sensor alignment as part recognition. While this paradigm is dominant in many domains, it is surprisingly absent from work on ATR. Instead, ATR is dominated by systems which employ fixed sets of image space templates or probe sets: each derived from a slightly different viewpoint. We believe that integrating 3D alignment into the recognition is superior to using large sets of image-space templates, and we are therefore developing new algorithms to accomplish this in the context of ATR.

Geometrically precise alignment techniques are computationally expensive. To limit its use, up-stream processing focuses attention so alignment is applied sparingly as a final means of resolving conflicting hypotheses. Consequently, there are two other significant efforts associated with this project. The first is a target detection effort being led by the University of Massachusetts. The second is a hypothesis generation effort being led by Alliant Techsystems.

Recently we have also begun looking at ways to add terrain model constraints into the recognition process. A first step has been taken by adapting existing matching capabilities to a restricted but quite important problem. To date, precise determination of the pointing angle of the Unmanned Ground Vehicle (UGV) relative to the terrain map has been determined by hand. At the end of this report there is a brief overview of some recent experiments suggesting a practical way of automating this process.

2 Alignment and ATR

While adapting the alignment paradigm to ATR might at first seem to be a simple transfer from on application domain to another, it is not. By their nature, ATR problems are more difficult than those typically solved using alignment algorithms. In ATR, image resolution is typically low. Targets viewed in color imagery are textured, in FLIR appearance is highly variable, and in range imagery geometric form is often complex. CAD models of targets are typically available, but often contain excessive detail. Terrain features introduce clutter and targets are often partially occluded. These factors, plus the fundamental ambiguities associated with perspective mapping of small objects into optical imagery, make direct application of current alignment algorithms infeasible.

To counter some of these difficulties, optical imagery may be supplemented with range imagery. Direct 3D

*This work was sponsored by the Advanced Research Projects Agency (ARPA) under grants DAAH04-93-G-422 and DAAH04-95-1-0447, monitored by the U. S. Army Research Office.

†Appears also in the Proceedings of the 1996 ARPA Image Understanding Workshop.

data resolves many ambiguities inherent in optical imagery. However, the introduction of a second sensor complicates the problem by introducing the need to fuse data from heterogeneous sensors. Fusion might be accomplished by processing range and optical imagery separately and combining evidence after the fact. However, doing so throws away the possibility of coupling evidence through the known 3D sensor and target geometry.

To exploit constraints derived from scene and target geometry, we are developing new algorithms which geometrically align 3D target models with both range and optical image data. This alignment is performed so as to maintain global geometric constraints associated with known sensor and scene geometry. Some geometric constraints are precisely calibrated while others are not completely specified. For example, the 3D position and orientation of the target relative to the sensors obviously varies. Also, as long as separately mounted range and optical sensors are used, exact pixel registration between images can be expected to vary. Thus, estimates of 3D object pose as well as image registration should be refined as part of the alignment process. As a shorthand, we have coined the term **coregistration** to describe this process of simultaneously refining these estimates based upon corresponding target and sensor features.

3 Components

To develop a complete end-to-end Target Recognition capability, recognition is divided into the three stages outlined below:

Color Target Detection Given color imagery, or imagery from any multi-band sensor, determine regions of interest (ROIs) where targets might be present. While initiating ATR with a detection phase is standard, two things are novel about the approach taken here. First, new machine learning technology for building multi-variate decision trees is being adapted to the problem of target detection. Second, color imagery is being utilized. Results to date often show camouflaged vehicles can be distinguished from natural backgrounds even when each is in a gross sense the same color: green camouflage against green grass and brush. Color detection is relatively mature and has been integrated and demonstrated running on the Unmanned Ground Vehicle. The color detection effort is led by the University of Massachusetts. The general approach to target detection as laid out in [BDHR94] is based upon more general work on the use of learned multivariate decision trees in computer vision [DBU94].

Hypothesis Generation Given regions of interest generated by the color detection process, or any other detection algorithm, this second stage hypothesizes what type or types of vehicles may be present and at what positions and orientations relative to the sensors. To provide this capability, a LADAR boundary template probing algorithm is being utilized which is itself an accomplished ATR algorithm [BJLP92]. This algorithm does the best it can to reduce the possibilities and then its top hypotheses are passed onto the final coregistration verification stage. Adding this third stage takes the pressure of making the final decision regarding target type off the boundary probing algorithm and means the algorithm operates under different performance constraints: it needs to generate a hypothesis which is approximately correct, but need not rank the top alternatives perfectly. Alliant Techsystems is leading the work on this algorithm.

Coregistration Target Verification This is the most intricate, computationally demanding, and novel aspect of our system. It takes as input the target ID and pose hypotheses generated by the previous step, and refines each based upon coregistration of the target model, optical imagery, and range imagery. The output is an exact match between sensor features and 3D target features and a quality of match measure based upon the associated 3D alignment of features. First instantiations of all the subcomponents needed to perform coregistration verification have been developed and integrated into a single testbed. Each of these subcomponents represents a separate research project is described further below.

To accomplish target verification through coregistration, a set of component technologies have been developed. These are summarized below:

Target Model Feature Prediction CAD models of targets do not explicitly represent the types of information required to do matching in optical and range imagery. An on-line algorithm for generating sampled-surfaces for matching to range data and 3D silhouette features suitable for matching to optical imagery has been developed.

Model-Driven Image Feature Extraction In optical imagery, bottom-up feature extraction is problematic at best. Consequently, a model-driven edge detection and line extraction process has been implemented which seeks locally optimal silhouette features in the optical imagery.

Range, Optical and Target Coregistration Extending past work on 3D pose determination [KH94], a new least-squares algorithm has been developed which determines the 3D pose of the target relative to a range and an optical sensor and simultaneously adjusts the registration mapping between the sensor image planes [SB94]. A recent extension of this work to perform median filtering has dramatically improved the quality of the results [Ant96b, Ant96a].

Coregistration Space Matching A match quality measure has been formalized to evaluate alternative coregistration estimates based upon fidelity of the target model to the sensor data. A local search operator in the space of coregistration mappings is then used to find better matches between target models and sensor features. This work is further described below, in [BSS96] and elsewhere in these proceedings [Mar96a].

Each of the components outlined above is a focus of research and the progress in each area is summarized in the following sections.

4 Progress on Key Components

4.1 Color Detection

The essential elements of this work along with results on data collected by ourselves and Martin Marietta at Fort Carson [BPY94a] were reported in the previous Image Understanding Workshop Proceedings [BDHR94]. Since this initial description of the color detection work, the following has been accomplished:

1. An improved way of coalescing individual pixel detections into ROIs has been implemented.

2. The algorithm has been successfully integrated with the other Reconnaissance, Surveillance and Target Acquisition software running on the UGV and has been demonstrated to run reliably in the field.

3. The algorithm is being formally evaluated in an independent effort being led by Ted Yachik of LGA.

4.1.1 Operating Scenario

It is best to start by reviewing the basic operating scenario for the color detection system. First, it is assumed that training imagery is obtained prior to a fielded mission, and based upon this training data the system learns to discriminate between color values produced by camouflaged vehicles and values produced by background terrain. The result of this training is a color lookup table (LUT) indicating, for each possible RGB color pixel value, whether it is more likely to be produced by a target or background.

In fielded operation, the system performs real-time color lookup on all pixels coming in and classifies them as target or background. Then, an ROI extraction process sums responses over fixed sized windows in the image and extracts ROIs: one ROI for each local maximum in this summed response image which is over a minimum threshold. When integrated with the RSTA package on the UGV, the results of the color detection were combined with those of a traditional FLIR detection algorithm.

IR and color detection complement each other, since false positives do not tend to correlate. For IR, false positives are typically produced by objects such as rocks, which heat up in the sun, or reflect solar energy back into the IR sensor [1]. For color, typically cool objects such a shrubs and trees tend to generate false positives.

Min	Max	Median	Mean	S.D.
3	41	13	15	9.5

Table 1: Detection Statistics on 51 Demo C Test Images. No true target was missed in this test.

Perhaps the most important factor in evaluating the usefulness of color detection concerns the degree to which training generalizes to variations in field conditions. The current system, using a single LUT, has been demonstrated to generalize across times of day, lighting conditions, weather and vehicles. Results demonstrating this on color imagery obtained from 35mm film have been previously reported [BDHR94, BHP95], and more recent results obtained using the color CCD sensor on the UGV are summarized below.

The current system does not generalize across sensors, but instead is trained to the specific response of a particular sensor. This is primarily a matter of experience, since generalization between sensors requires effort be devoted to the problem of cross-sensor color calibration. In principle such calibration can be done, but there are a variety of subtle issues involved which make this its own topic for future research.

4.1.2 Experience Running on SSV-B

One way to illustrate our confidence in the potential of the color detection system is to simply recount our first experiences with the system running as part of the UGV RSTA package. After a significant software integration, the system was finally integrated and debugged by mid June, 1995. On Tuesday, June 13, 18 training images were collected using SSV-B. On Wednesday morning, June 14, an hour was taken to select 14 image chips: 3 indicating typical background colors and 11 showing vehicles. A color LUT was built, loaded on the vehicle, and the system was tested from 1 to 5 PM on 51 new images. These 51 images included targets not in the training data: both with brown and green camouflage and viewed from vantage points different from those in the training data.

The key result was that over the 4 hour period, under both cloudy and sunny conditions, viewing 4 different targets from 2 different vantage points, the system *never missed a target*. This first field result was positive beyond our expectations. While perfect performance such as this is not a realistic expectation in general, it is suggestive of the strength of the system. Tight timing constraints associated with scheduling of SSV-B leading up to Demo C prevented further field testing.

Because the system was tuned to work with FLIR,

[1] The RSTA FLIR operates in the 3 to 5 micron band and is sensitive to reflected thermal energy.

(a) (b)

Figure 1: Color Detection Example in UGV Data from Demo C Test Site. a) Image with ROI boxes overlaid, b) Summed detection values from which ROIs are derived.

a high false positive rate was considered acceptable as a way of reducing the change of missed targets. Table 1 provides statistics summarizing the detection performance over the 51 test images. The columns present the minimum number of detections on a single image, maximum number of detections on a single image, the median number across the image set, the average number, and the standard deviation.

To illustrate how these detection ROIs appear, the ROIs found for a typical image from the June tests at the Demo C site are shown in Figure 1a [2]. The summed response producing these ROIs are shown in Figure 1b. Because each ROI is relatively small, even for those images with high numbers of detections, the color detection algorithm is focusing attention on a very small percentage of the total image.

4.2 Hypothesis Generation

The LADAR boundary probing system developed by Alliant Techsystems has been modified to run as a stand alone system on a Sparc workstation. James Steinborn at Colorado State, and Kris Siejko at Alliant Techsystems, have been working to make the system operational in this stand-alone mode running under Solaris. They have also been making enhancements, including a new visualization tool developed by Jim Steinborn which enables us to better understand the geometric structure of the probe templates.

Preliminary tests of system have been performed on LADAR images from the Fort Carson data set [BPY94b]. Templates for the M113 and M60 generated at 52 meters were selected for all the tests. Templates generated at closer and farther ranges were tried as well, but the results changed little. The probing algorithm does its own internal template scaling based upon the LADAR data itself. There were a total of 72 templates: 36 of each vehicle sampled at 10 degree aspect increments.

The results for this test are summarized in Table 2 [3]. The hypothesis generation algorithm produces a set of ranked hypotheses. The top five are indicated as pairs: (vehicle)/(aspect angle). The best hypothesis, in terms of the true type and aspect of the vehicle, are indicated in boldface in Table 2.

The correct hypotheses are not the highest ranked hypotheses based upon the boundary probe result alone. However, a hypothesis for the correct target within 10 degrees of the true aspect angle does appear in the top five for all 4 images. While the performance of the boundary probing could no doubt be improved through more careful tuning to the Fort Carson data, this experiment demonstrates the system is capable of focusing attention upon a small set of aspect and target hypotheses.

4.3 Model Feature Prediction

Highly detailed Constructive Solid Geometry (CSG) models of target vehicles are available in BRL-CAD format [U. 91]. These detailed models are a tremendous asset. However, much work is required to transform these CSG models into features appropriate for matching to sensor data. Over the past year, Mark Stevens has developed a semi-automated system for transforming the CSG to a polygonal representation [SBG95, Ste95]. He has also developed a fully automated system for extracting edge and surface information from these polygonal models. This later system is summarized here and more fully described elsewhere in these proceedings [Mar96a].

For optical imagery, 3D face boundaries likely to generate observable edges in imagery are extracted from the 3D target models. This is done on-line given an estimate of the target pose and lighting. Target pose is produced

[2] Many of the figures in this paper look much better in color and can be accessed through the CSU homepage: `http:://www.cs.colostate.edu/~vision/`

[3] The LADAR and optical imagery for image 4 appears in Figure 3.

No.	File	Annotation	Hyp 1	Hyp 2	Hyp 3	Hyp 4	Hyp 5
1	nov21115L1	Clear Angle On	M113/10	**M113/40**	M113/120	M113/30	M113/50
2	nov40755L1	Nose Down Profile	M113/120	M113/110	M113/60	M113/50	**M113/70**
3	nov31055L1	Clear Profile	M113/120	M113/60	M113/50	M113/70	**M113/80**
4	nov31170L1	Head On Nose Down	M113/30	M113/20	M113/40	M60/0	**M113/10**

Table 2: Hypothesis Generation Results for 4 Fort Carson Images. The target in all cases is an M113. The top five hypotheses are shown in ranked as (vehicle)/(aspect angle). The hypothesis indicated in boldface is within 10 degrees of the true aspect angle.

by the hypothesis generation phase described above and is refined as part of matching as discussed below in Section 4.6. Lighting information for outdoor scenes may be derived from co-lateral information regarding time-of-day, time-of-year, place and weather.

For range imagery, sampled surfaces are extracted from the 3D model using a process which simulates the operation of the actual range sensor. The target model is transformed into the range sensor's coordinate system using the initial estimate of the target's pose and rays are cast into the scene and intersected with the 3D faces of the target model. Sampling geometry is selected to reflect the characteristics of the actual range device.

4.3.1 Silhouettes and Internal Structure

To match 3D rather than 2D image space, features must be 3D and not flattened 2D templates. Therefore, the model feature prediction system determines those 3D features within the target model responsible for generating the target silhouette for a given pose estimate. These 3D features naturally accommodate modest changes in viewing angle. If the expected pose estimate changes significantly, new features may be generated. The first version of the feature prediction system extracts only features associated with the silhouette. More recent work, described in more detail elsewhere in these proceedings [Mar96a], extends this method to add significant internal structure as well.

The silhouette prediction algorithm begins by assigning a unique color to each face in the target model. This color acts as an index into a hash table of 3D faces. The model is then rendered from the hypothesized viewing angle using orthographic projection and a hardware Z-buffer. A target model with 250 faces can be rendered in 1.2 seconds on a Sparc 10 with a ZX hardware accelerator.

Pixels adjacent to the unique background color indicate faces contributing to the target silhouette. These faces are in turn checked to determine which face boundaries (edges) contribute to the silhouette. These edges are then clipped to retain only those segments lying on the silhouette. Since rendering is orthographic, parametric values for clipping measured in the image space may be applied directly to the corresponding 3D edges. The final result is a list of 3D edges representing the silhouette of the target model for a given viewing angle.

4.4 Model-Driven Image Feature Extraction

Bottom-up line extraction, such as performed by the Burns algorithm [BHR86], is unreliable in imagery such as that shown in Figure 2b. To overcome the difficulty inherent in this imagery, a more model-driven approach is required. To accomplish this, we combine two ideas from the literature: model-driven edge detection [FL87, FL88] and directionally tuned gradient filters [Can86]. The quality of a straight line segment denoting an extended edge is defined to be a function of the gradient magnitude under that edge. A gradient mask tuned to the specific expected orientation of the segment is used. The placement of the segment is perturbed until a locally optimal placement is found.

4.4.1 Placing Silhouette Edges

Initially, 3D silhouette edges are projected into the color image based upon the known intrinsic sensor parameters and the estimated pose of the target. Figure 2b shows the projection of a 3D silhouette onto the image plane using the sensor calibration parameters [BHP94]. The pose estimate comes from the hypothesis generation stage.

For each projected silhouette feature, a search is initiated in the image for the locally best corresponding line segment. A gradient mask tuned to the the particular expected orientation of each silhouette edge is created by rotating the first derivative of a Gaussian. There are many precedents for tuned edge masks including Canny [Can86] and Torres [TP86] and their use for bottom-up edge detection [Shu94, FA91]. An example of such a filter, displayed as an image, is shown in Figure 2a.

The placement of the silhouette edge is locally perturbed so as to maximize a function of the underlying tuned gradient response. To do this with subpixel accuracy, a commonly used graphics anti-aliasing technique known as Pineda Arithmetic [Pin88] is used to weight the contribution of individual pixels. A weighting value for each pixel is created (see Figure 2d) and the response for the silhouette line is the weighted sum of responses at each pixel. Additional details on this work appear elsewhere in these proceedings [Mar96b].

a. Mask **b.** Silhouette Line **c.** Gradient Response **d.** Weight

Figure 2: Gradient Mask and Response

4.5 Range, Optical and Target Coregistration

Anthony Schwickerath has developed a new least-median-squares algorithm for determining the best coregistration estimate based upon a set of corresponding sensor and target features. The motivation for this work and full mathematical development has been presented previously [SB94, BHP95]. Recently, median filtering has been included in the algorithm. In addition, a match error for ranking alternative correspondence mappings [Ant96b] has been developed.

The complete system has now been demonstrated on optical and range imagery collected at Fort Carson using high quality range and optical target model features generated using the system described above. Results of these improvements are reported in more detail elsewhere in these proceedings [Ant96a]. Here, allow us to summarize briefly what coregistration does, and show one result using the new median filtering capability.

4.5.1 Review and Median Filtering Example

Coregistration addresses a fundamental problem arising when optical and range imagery is collected from separate co-located sensors. Off-line calibration can largely determine the sensor-to-sensor registration, but some small variations may be expected during field operations. Thus, multi-sensor pose determination must include degrees of freedom to express movement of the target relative to the sensors, as well as degrees of freedom to permit fine adjustments to the image-to-image registration between the sensors.

Our specific formulation of this problem introduces a coplanarity constraint which limits the freedom of movement of the range sensor relative to the optical sensor. Thus, the range reference coordinate system may translate in the common x-y image plane of the two sensors, but otherwise the two sensors move together. Thus, six degrees of freedom express the position and orientation of the target relative to the sensor suite and two degrees of freedom permit translation of the optical image plane with respect to the range image plane.

This choice of parameterization may at first seem odd. Given two sensors on a common platform, it is their relative pointing angles, not their relative spacing, which is most likely to vary. However, the pixel-to-pixel movement between the two image planes is virtually indistinguishable in the two cases when rotations are small. The advantage of this translation formulation is that it does *not* introduce a second rotation term into the coregistration formulation, which would in turn add unnecessary nonlinearities.

Figure 3 [4] shows a result of the median filtering coregistration algorithm on a pair of range and optical images containing an M113. This is the same data for which results of the hypothesis generation algorithm were reported in Table 2: image NOV31170L1 [5]. The top image shows silhouette model features overlaid on the optical image. The body of the figure shows the coregistered range features from nine different vantage points in order to provide alternative views of the 3D structure. The filled rectangles are features included in the least-median-squares match: light representing image points and dark representing model points. The outlined rectangles are features excluded from the match. The coregistration estimate shown in Figure 3 has correctly skewed the target aspect slightly to the left, even though the initial estimate was exactly head-on.

4.5.2 Correspondence Space Matching

The space of possibly corresponding target and sensor features is inherently combinatoric and the absolute number of possible correspondences is highly dependent upon the accuracy of the initial coregistration estimate. To determine candidate features, range and optical model feature are projected into the imagery and all image features within some local area are marked as potential matches. The size of these areas grows with uncertainty in the coregistration estimate. For median filtering to be effective, the initial estimate must be accurate enough to ensure that over 50% of all paired features are part of the true match.

In principle, search in this combinatorial space of correspondence mappings could be conducted using local search in a manner analogous to that set for in Beveridge's dissertation [Bev93]. However, as a practical matter, the fine grained sampling of range points causes the combinatorics to explode. While such a local search

[4] This figure is being improved so greytones work out properly

[5] The unusual angle of the vehicle makes this an interesting and challenging problem.

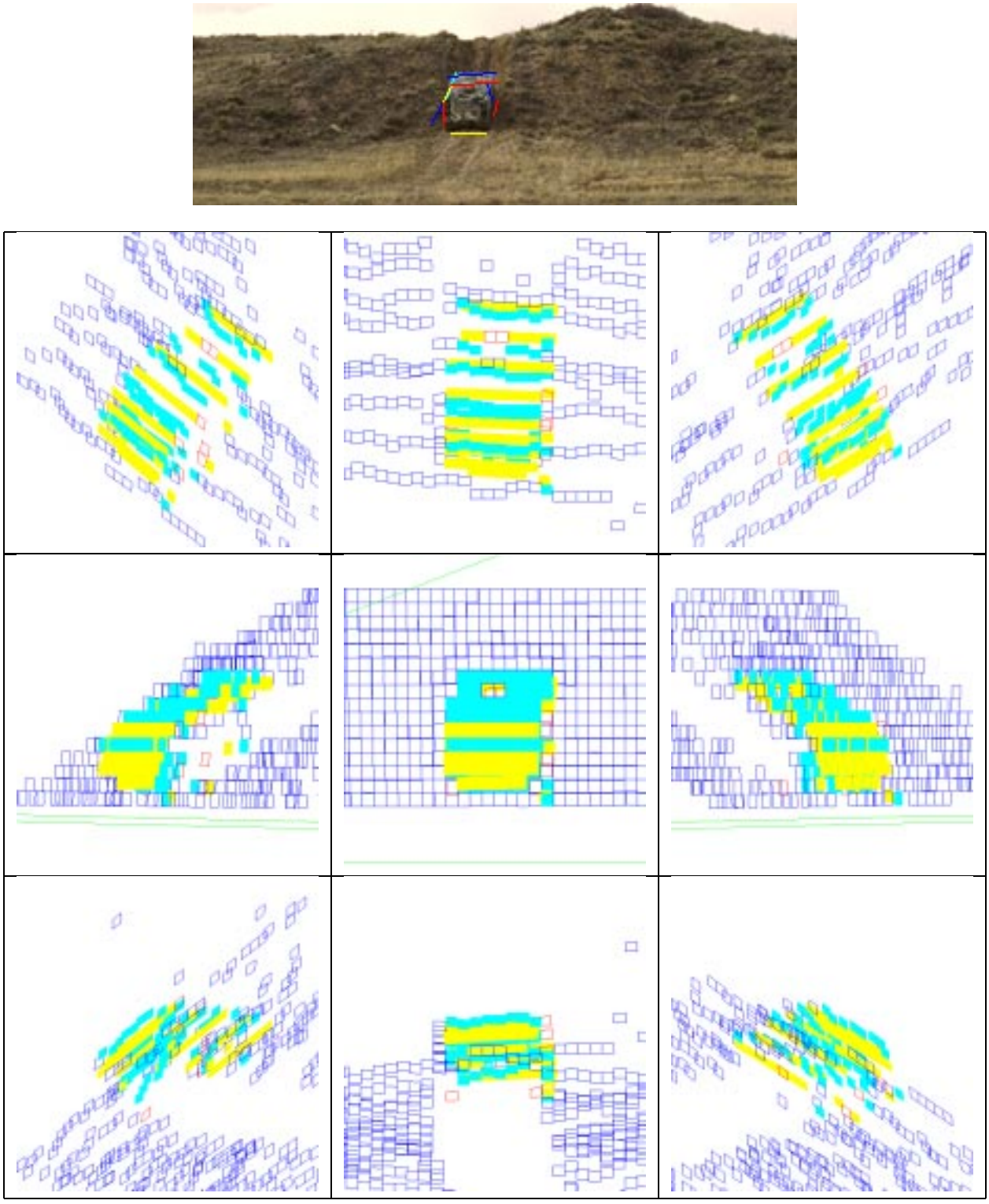
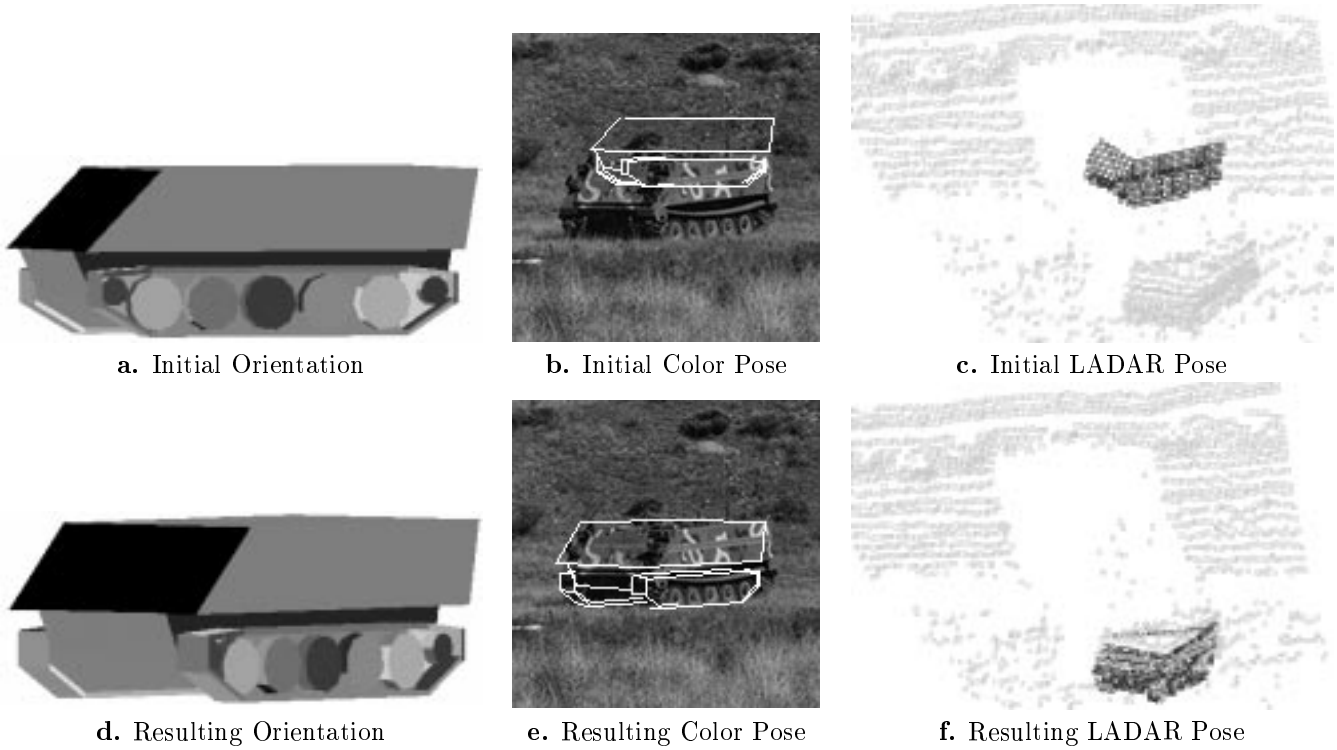


Figure 3: Coregistration median filtering results for m113 in a pair LADAR and Color Images. The top image shows silhouette features in color image using the coregistration estimate. The bottom nine views show 3D registration of target and range features displayed from different viewpoints.

procedure has been implemented, it's usefulness is currently limited, and will only become practical if some form of initial grouping is applied to the range features. This is trivial in the case of the target models, where bounded faces are already represented. For the range data, we are experimenting with a scan-line range segmentation technique [X. 94].



a. Initial Orientation **b.** Initial Color Pose **c.** Initial LADAR Pose

d. Resulting Orientation **e.** Resulting Color Pose **f.** Resulting LADAR Pose

Figure 4: Coregistration-space Local Search Results for Shot 20 Array 5 (M113 APC)

4.6 Coregistration Space Matching

As an alternative to search in the space of correspondence mappings between target model and sensor features, this section introduces a local search algorithm which operates in the space of coregistration estimates. While increasing the number of potentially matching features increases the size of the correspondence space exponentially, the dimensionality of the coregistration space is fixed: $\Re^8$ for the case treated here. This work is being pursued as an extension of Mark Stevens's work on model prediction and model-driven feature extraction. A more detailed account of this work appears elsewhere in these proceedings [Mar96a]. What is presented here is an overview with an example.

Early experiments with search in coregistration-space suggests an ability to correct quite large errors in the initial coregistration parameters. While this approach is still very early in its development and has been applied to only four pairs of range and optical imagery, this initial experience is quite encouraging.

If any failing has been observed so far, it is a weakness in generating the final highly precise match. In other word, search leads to a much better but not perfect coregistration estimate. This suggests the two approaches complement each other. the first getting to a near correct estimate followed by median filtering producing a highly accurate final result. These two algorithm are implemented within the same testbed and combined testing will begin shortly.

The search algorithm locally minimizes an error function which measures the relationship between the model and data features. This measurement takes into account both range and optical features, but treats the two cases somewhat differently. For the optical features, the error is a function of the gradient measurement used in the model-driven edge detection described in Section 4.4. For range, the error measure is a function of the Euclidean distance from points on the target model sampled surface to their nearest neighbor in the range image data.

The local search itself samples each of the 8 dimensions of the coregistration-space about the current estimate. Clearly, the step-size used in this sampling is important. The general strategy implemented moves from coarse to fine sampling as the algorithm converges upon a locally optimal estimate. The initial scaling of the sampling interval is determined automatically based upon moment analysis applied to the initial target-model and sensor data sets. The search continually takes the best along each dimension until it reaches an optima. When no further progress is possible along any dimension, the resulting 8 values are returned as the locally optimal coregistration estimate.

Figure 4 shows a sample result of the coregistration-space search applied to an image containing the m113. The left column shows the initial coregistration estimate, both target pose and sensor registration, provided by the hypothesis generation algorithm. The right column shows the result of the local search algorithm. Figures 4a and 4d show how the process has corrected the orienta-

tion of the target.

Figures 4b and e show both the silhouette and internal features. Observe how closely the features in Figures 4e correspond to the color image. Also note that change in appearance reflects the 3D rotation of the target model: the change between Figures 4b and 4 e in a could not be produced by a 2D affine transformation of a 2D template.

Figures 4c and 4f show the movement of the target model sampled points relative to the actual range data. It is this range data which is providing much of the constraint used to correct the orientation of the target. The dark grey rectangles are model points, the lighter grey rectangles are range image points. Because the search aligns the two, it is difficult to distinguish model from data in Figure 4f, while they are easily distinguished in the initial configuration: Figures 4c.

5 Horizon Line Orientation Correction

There are both immediate and longer range benefits to be realized if terrain-based constraints are integrated into the ATR process. One of the most immediate needs is strikingly evident in the RSTA function of the UGV program. The current SSVs developed by Martin Marietta use GPS to determine position and inertial navigation to determine orientation. All but the most expensive inertial systems only measure true orientation to within, speaking loosely, 1 degree. For some purposes this is a modest error, but when attempting to register imagery from a 4 degree field of view sensor with stored terrain maps, such an error introduces a 128 pixel error in a 512x512 image.

Consequently, a normal part of current SSV operation is something called HOC (Horizon Line Orientation Correction) or LOC (Landmark Orientation Correction). This is a by-hand procedure in which known points on the horizon or surveyed points on the terrain map are fed to the RSTA system in order that it can recover, to within several pixels, the true relationship between the stored terrain map and the live video imagery. Clearly, this is done once the vehicle is stopped at an observation point, and must be repeated each time the vehicle is moved.

While the need for such precise registration may at first not be obvious, it is a pre-condition for the use of most FLIR target detection algorithms. This is because these algorithms require an initial range-to-target estimate for every pixel in the FLIR image. In the absence of an active ranging sensor, these estimates can be derived from terrain maps, but only if precise registration is established.

Recently, Christopher Graves and Christopher Lesher have begun work on automating the terrain-to-imagery registration task. A proof-of-concept horizon line matching experiment has been conducted using imagery collected by Lockheed-Martin at the UGV Demo C test site. The geometric matching system originally developed as part of Beveridge's thesis work is being used [Bev93] in this test and results are presented elsewhere in these proceedings [J. 96]. The results prove the feasibility of using local search matching as a tool to automate the orientation correction process in domains where horizon structure is distinctive.

References

[Ant96a] Anthony N. A. Schwickerath and J. Ross Beveridge. Coregistering 3D Models, Range, and Optical Imagery Using Least-Median Squares Fitting. In *Proceedings: Image Understanding Workshop*, page (to appear), Los Altos, CA, February 1996. ARPA, Morgan Kaufman.

[Ant96b] Anthony N. A. Schwickerath and J. Ross Beveridge. Coregistration of Range and Optical Images Using Coplanarity and Orientation Constraints. In *1996 Conference on Computer Vision and Patter Recognition*, page (submitted), San Francisco, CA, June 1996.

[BDHR94] Shashi Buluswar, Bruce A. Draper, Allen Hanson, and Edward Riseman. Non-parametric Classification of Pixels Under Varying Outdoor Illumination. In *Proceedings: Image Understanding Workshop*, pages 1619–1626, Los Altos, CA, November 1994. ARPA, Morgan Kaufmann.

[Bev93] J. Ross Beveridge. *Local Search Algorithms for Geometric Obejct Recognition: Optimal Correspondence and Pose.* PhD thesis, University of Massachuesetts at Amherst, May 1993.

[BHP94] J. Ross Beveridge, Allen Hanson, and Durga Panda. Integrated color ccd, flir & ladar based object modeling and recognition. Technical report, Colorado State University and Alliant Techsystems and University of Massachusetts, April 1994.

[BHP95] J. Ross Beveridge, Allen Hanson, and Durga Panda. Model-based fusion of flir, color and ladar. In Paul S. Schenker and Gerard T. McKee, editors, *Proceedings: Sensor Fusion and Networked Robotics VIII*, Proc. SPIE 2589, pages 2 – 11, October 1995.

[BHR86] J. B. Burns, A. R. Hanson, and E. M. Riseman. Extracting straight lines. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI–8(4):425 – 456, July 1986.

[BJLP92] 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, mar 1992.

[BPY94a] J. Ross Beveridge, Durga P. Panda, and Theodore Yachik. November 1993 Fort Carson RSTA data collection final report. Technical Report CS-94-118, Computer Science Dept., Colorado State University, January 1994.

[BPY94b] 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.

[BR95] J. Ross Beveridge and Edward M. Riseman. Optimal Geometric Model Matching Under Full 3D Perspective. *Computer Vision and Image Understanding*, 61(3):351 – 364, 1995. (short version in IEEE Second CAD-Based Vision Workshop).

[BSS96] J. Ross Beveridge, Mark R. Stevens, and N. A. Schwickerath. Toward target verification through 3-d model-based sensor fusion. *IEEE Transactions on Image Processing*, page (Submitted), 1996.

[Can86] John Canny. A computational approach to edge detection. *PAMI*, 8(6):679–698, November 1986.

[DBU94] 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.

[EG92] R. O. Eason and R. C. Gonzalez. Least-Squares Fusion of Multisensory Data. In Mongi A. Abidi and Rafael C. Gonzalez, editors, *Data Fusion in Robotics and Machine Intelligence*, chapter 9. Academic Press, 1992.

[FA91] William T. Freeman and Edward H. Adelson. The Design and Use of Steerable Filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-13(9):891–906, September 1991.

[FL87] Pascal Fua and Yvan G. Leclerc. Finding Object Boundaries Using Guided Gradient Ascent. In *Proceedings: Image Understanding Workshop – 1987*, pages 888–891. DARPA, Morgan Kaufmann, February 1987.

[FL88] Pascal Fua and Yvan G. Leclerc. Model Driven Edge Detection. In *Proceedings: Image Understanding Workshop – 1988*, pages 1016–1020. DARPA, Morgan Kaufmann, April 1988.

[HU90] Daniel P. Huttenlocher and Shimon Ullman. Recognizing Solid Objects by Alignment with an Image. *International Journal of Computer Vision*, 5(2):195 – 212, November 1990.

[J. 96] J. Ross Beveridge and Christopher Graves and Christopher E. Lesher. Local Serach as a Tool for Horizon Line Matching. In *Proceedings: Image Understanding Workshop*, page (to appear), Los Altos, CA, February 1996. ARPA, Morgan Kaufman.

[KH94] Rakesh Kumar and Allen R. Hanson. Robust methods for estimating pose and a sensitivity analysis. *CVGIP:Image Understanding*, 11, 1994.

[Low85] David G. Lowe. *Perceptual Organization and Visual Recognition.* Kluwer Academic Publishers, 1985.

[Mar96a] Mark R. Stevens and J. Ross Beveridge. Interleaving 3D Model Feature Prediction and Matching to Support Multi-Sensor Object Recognition. In *Proceedings: Image Understanding Workshop*, page (to appear), Los Altos, CA, February 1996. ARPA, Morgan Kaufman.

[Mar96b] Mark R. Stevens and J. Ross Beveridge. Optical Linear Feature Detection Based on Model Pose. In *Proceedings: Image Understanding Workshop*, page (to appear), Los Altos, CA, February 1996. ARPA, Morgan Kaufman.

[Pin88] Juan Pineda. A Parallel Algorithm for Polygon Rasterization. In *Proceedings of Siggraph '88*, pages 17–20, 1988.

[SB94] Anthony N. A. Schwickerath and J. Ross Beveridge. Model to Multisensor Coregistration with Eight Degrees of Freedom. In *Proceedings: Image Understanding Workshop*, pages 481 – 490, Los Altos, CA, November 1994. ARPA, Morgan Kaufmann.

[SBG95] Mark R. Stevens, J. Ross Beveridge, and Michael E. Goss. Reduction of BRL/CAD Models and Their Use in Automatic Target Recognition Algorithms. In *BRL/CAD Symposium 95*, 1995.

[Shu94] Alexander Shustorovich. Scale Specific and Robust Edge/Line Encoding with Linear Combinations of Gabor Wavelets. *Pattern Recognition*, 27(5):713–725, 1994.

[Ste95] Mark R. Stevens. Obtaining 3D Silhouettes and Sampled Surfaces from Solid Models for use in Computer Vision. Master's thesis, Colorado State University, September 1995.

[TP86] Vincent Torre and Tomaso A. Poggio. On Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(2):147–164, March 1986.

[U. 91] U. S. Army Ballistic Research Laboratory. *BRL-CAD User's Manual*, release 4.0 edition, December 1991.

[X. 94] X. Y. Jiang and H. Bunke. Fast Segmentation of Range Images into Planar Regions by Scan Line Grouping. *Maching Vision and Applications*, 7(2):115 – 122, 1994.