MODEL-BASED INTERPRETATION OF RANGE IMAGERY

Darwin T. Kuan
Robert J. Drazovich

Advanced Information & Decision Systems
201 San Antonio Circle, Suite 286
Mountain View, CA 94040

ABSTRACT

This paper describes a model-based approach to interpreting laser range imagery. It discusses the object modeling, model-driven prediction, and feature-to-model matching aspects of the problem. The model objects are represented by a viewpoint-independent volumetric model based on generalized cylinders. Predictions of 3-D image features and their relations are generated from object models on multiple levels. These predictions give guidance for goal-directed shape extraction from low level image features. Interpretation proceeds by comparing the extracted image features with object models in a coarse to fine hierarchy. Since the 3-D information is available from the range image, the actual measurements are used for feature-to-model matching. A limited prototype system has been developed, preliminary results on prediction and interpretation are shown, and future research directions are discussed.

I INTRODUCTION

Range images offer significant advantages over passive reflectance images because they preserve the 3-D information of the scene viewed from the sensor. While 3-D information can be obtained from 2-D images only with extensive inference (due to the ambiguities introduced by the 2-D projection of the 3-D scene and the additional effects of illumination, surface reflectivity, and geometric shape), they can be easily calculated from 3-D range images. Therefore, range data is becoming an increasingly important source of information for a variety of applications including automatic 3-D target classification, autonomous vehicles, robot vision, and automatic inspection. In this paper we discuss 3-D object recognition for vehicle objects in air-to-ground laser range imagery.

Much of the past work on range data analysis has emphasized a data-driven approach [5], [7]. The ACRONYM system [2] is a powerful model-based vision system capable of doing symbolic reasoning among 3-D models and 2-D images. In this research, we have attempted to extend the principles of the ACRONYM approach to the analysis of 3-D imagery.

Research sponsored by the Air Force Office of Scientific Research (AFSC), under Contract F49620-82-C-0071. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation herein.

This paper describes a model-based approach to interpreting laser range imagery. The full system includes 3-D image feature extraction, geometric object modeling, model-driven prediction, and image feature-to-model matching. The overall structure of the system is shown in Figure 1. Further discussion of this system is provided in [4]. The extraction of low level 3-D image features from the range data was previously reported in [3]. This paper emphasizes the object modeling, model-driven prediction, and image feature-to-model matching aspects of the system.

In the geometric modeling system (see Figure 1), object models are represented by a single, viewpoint-independent representation based on generalized cylinders. Model priority index information and attachment relations are explicitly specified in the models to facilitate 3-D image feature prediction and extraction. The prediction process predicts physical edge types (occluding, convex, or concave edges), surface properties (planar or curved), cylinder contour, cylinder obscuration (visible, occluded), and invariant shape properties (parallel, collinear, connectivity). These knowledge-based predictions are on multiple levels and are very powerful for directing the feature extraction algorithms' search for particular features in a limited region. The object classification task proceeds from coarse to fine by first comparing gross object features (e.g., object length, height, extreme points, etc.) and then finer component features (e.g., cylinder volume, position, orientation, etc.) extracted from laser imagery with a model using a set of rules that produces a likelihood value to indicate the goodness of match. Since the 3-D information are available from the range image, the actual measurements (e.g., length, width, volume) are used for matching.

II MODEL REPRESENTATION

The system uses a viewpoint-independent volumetric representation for 3-D object modeling. The volume primitives we use are generalized cylinders [1]. A generalized cylinder is defined by a space curve, called the axis, and planar cross-section functions defined on the axis. A complex object is represented in terms of a set of individual cylinders and their spatial relations. The 3-D object model has a hierarchical representa-

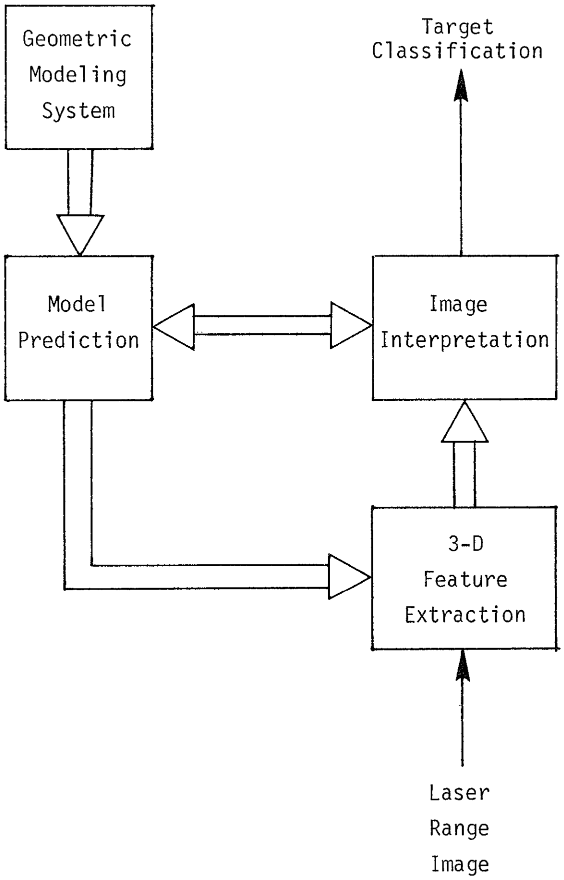
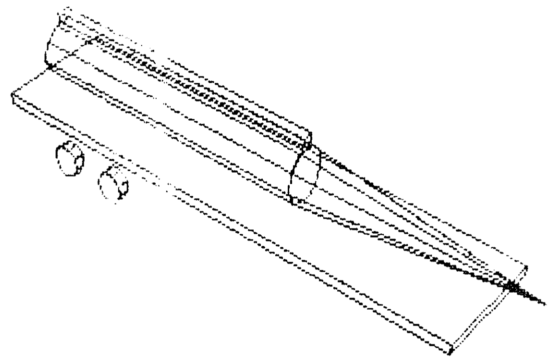
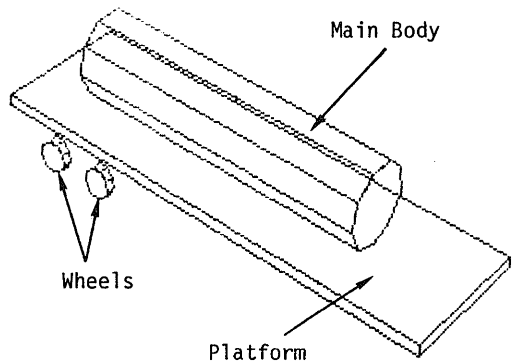


Figure 1: Major Components of a 3-D Object Classification System



Component Level Model of Missile Launcher



Component Level Model of Missile Launcher Decoy

Figure 2: 3-D Models of a Missile Launcher and a Decoy

tion with coarse to fine details that enables successful refinement of analysis and also provides a prediction generation mechanism at multiple levels. These levels might include the scene, object, component and sub-component levels. Figure 2 shows the models of a missile launcher and a missile launcher decoy at the component detail level.

The components (cylinders) in the same model detail level may vary in importance for recognizing the object. For example, the gun barrel of a tank is unique in vehicle models and provides sufficient evidence to distinguish a tank from a truck or other vehicles. Therefore, to recognize a tank, we may first look for the gun barrel in the image. This kind of knowledge for goal-directed feature extraction is explicitly represented in our object models by using a model priority index. Another aspect of the model priority index is determined by the geometric properties of each cylinder. For example, elongated cylinders and large cylinders show distinct cylinder properties that are easy to distinguish from other cylinders. These distinguished pieces can be used for fast model access and selection. The model priority index is viewpoint-independent and provides a mechanism for efficient model access and selection. However, this index does not give us any information about the visibility of the cylinder or the ease of cylinder extraction.

Shapes that are occluded are generally more difficult to extract from images. This knowledge can be used to direct the feature extraction algorithm to look for reliable and easy to obtain features first. This kind of planning information is encoded in an obscuration priority index that indicates the occlusion and visibility relations among cylinders. The obscuration priority index is similar to the priority algorithm [6] used for hidden surface elimination. The idea is to arrange all cylinders in the scene in priority order based on their depth and obscuration relations. Cylinders closer to the viewpoint and not obscured by other cylinders will have higher priority indices. Cylinders that are totally obscured are explicitly indicated, and no effort will be spent trying to find them in the range image. This obscuration priority index is purely geometrical and determined from the object orientation and viewpoint. The combination of model priority index and obscuration priority index gives a new priority order that not only indicates the importance of a particular cylinder for object recognition, but also compares the ease of cylinder extraction in the range image. Due to the performance limitation on occluded cylinder extraction, we currently use the obscuration priority index as the base index for cylinder extraction and matching. If two cylinders have the same obscuration priorities (they don't obscure each other), then the model priority indices are compared to determine the order of feature extraction sequence. In general, the priority index is valid for an interval of the viewing angle and thus can be used based on rough object orientation estimates.

Physical edges such as occluding, convex and concave edges can be distinguished in the range image. The prediction of the physical edge types of a cylinder contour strongly constrain the possible interpretations of each edge segment. In order to use information, our models should have the capability to predict the types of physical edges. To do this, the attachment relations between cylinders in the object are explicitly specified. The specification of these attachment relations facilitates the prediction of physical edge type and will be discussed in more detail in the edge level prediction section below.

III MODEL

Prediction is the process of making estimates about the image appearance of objects using model knowledge and given some information about the objects' relative position, orientation, and shape in the image. Predictions first give guidance to image feature extraction processes for goal-directed shape extraction; then they provide mechanisms for feature-to-model matching and interpretation. The best features for prediction are those invariant features that will always be observable in the image independent of the object's orientation and sensor position. Examples of these invariant features in range imagery are physical edge type (i.e., occluding, convex, concave), surface type (planar, curved), cylinder contour edge type (occluding or concave), collinear and parallel relations and connectivity.

The prediction process proceeds hierarchically in a top-down fashion. Global object features and spatial relations among object components viewed from the sensor are first predicted. These include the object side-view characteristics, and the obscuration priority index. From the obscuration and model priority indices, the system chooses a cylinder with the highest priority. The occluding contour of this cylinder is identified and the physical edge types of the cylinder contour boundaries are predicted. At the cylinder contour level, there are only two types of edges, occluding and concave. The convex edge only occurs as the internal edge of a cylinder, hence it is irrelevant in the cylinder-level feature prediction and extraction. Once a cylinder is extracted, its volume properties and relative position and orientation to the object coordinate system are used for interpretation. If this information is not sufficient for target classification, the cylinder contour of the next highest priority cylinder is predicted. The same procedure continues until the object recognition task is achieved or all the object components have been examined. Details of these predictions developed for different levels are discussed in the following paragraphs.

A. <u>Object-Level Prediction</u>

The object level predictions provide global object features and spatial relations among object components viewed from the sensor. Examples are the dimensions and side-view characteristics of the object in the image, the spatial relationships between object components and the occluding relations among object components. The structural relationships (object-level prediction) can be used for global structure matching after we extract and analyze the individual cylinders of an object or it can be used similar to a junction dictionary [8] to guide the search of cylinder features. For objects with self-occlusion, the second approach may be more appropriate because structural knowledge is actively used for both occluded cylinder feature extraction and interpretation.

B. <u>Cylinder-Level Prediction</u>

Cylinder-level predictions provide goal-directed guidance for cylinder extraction from low level image features. This is the most important prediction level in the system because cylinders are the basic symbolic primitives used to perform image feature-to-model matching.

A generalized cylinder is hierarchically characterized by first defining its axis and then the cross-sections along the axis. For totally visible cylinders (cylinders with a high obscuration priority index), cylinder level prediction is accomplished by using the hierarchy in defining generalized cylinders. The properties of the two major cylinder boundaries along the cylinder axis are first predicted. These predictions include parallel relations (relative angles in general), physical edge types (concave, occluding), length, distance between two segments, and the extent of overlap. These predictions are sufficient to guide the coarse extraction of cylinders. After extracting the two major cylinder boundaries along the axis, other boundaries on the cylinder contour can be predicted in limited regions relative to the two major boundaries. Heuristic rules utilizing convexity, neighboring relations, the goal distance, and relative position information are used to find the complete cylinder contour from incomplete edge segments. A complete example will be given in the processing example section.

Due to the internal structure of the vehicle objects, most cylinders are partially obscured. The prediction of occluded cylinder contours in the image can be generated by polygon clipping algorithms [9] according to the obscuration priority index. Again, the same hierarchical cylinder extraction process can be used. Additional relations on the two major boundaries such as missing segments due to occlusion and collinear relations between disjoint segments can also be identified. In addition, the relative structural relations of the extracted cylinders (with higher obscuration index than the current cylinder) constrain the search region. The common boundaries between the current cylinder and the previously extracted cylinders can be used as landmarks for registration. We have not implemented this partially occluded cylinder extraction algorithm and it requires further research efforts.

C. Surface-Level Prediction

Surface-level predictions describe the cylinder surface appearances and their spatial relations in the image. The convex edges are useful at this level for grouping edges into the boundaries of a surface patch. These surface primitives can be grouped together to form a cylinder according to surface-level predictions from the model. Due to the low resolution nature of air-to-ground laser imagery, surface properties are not easy to extract and to use (sometimes a single surface patch only has a few points). However, for industrial applications where high resolution range images are available, surface-level predictions impose strong constraints on cylinder extraction.

D. Edge Level Prediction

Edge-level predictions assign physical edge types to each edge segment and thus strongly constrain the possible interpretation of each edge segment for cylinder contour extraction.

Prediction of the physical edge type of a cylinder contour is made possible by explicitly specifying the attachment relations between cylinders in the model. For example, if cylinder A is supported by cylinder B, the two touching faces of the cylinders are explicitly labeled in the object model. The occluding edge type is predicted for those cylinder contour segments that do not belong to a labeled face. The concave edge type is predicted for those segments that belong to a labeled face, and are inside the other surfaces with the same label. The convex edges correspond to the internal edges of cylinders, and thus are not useful for cylinder extraction. The physical edge type limits the search space for cylinder grouping, but more importantly, it can be used to verify the correctness of the extracted cylinder.

IV INTERPRETATION

Interpretation proceeds by comparing the image features on multiple levels to the object models according to a set of if-then rules. Each rule for comparison produces a "goodness" measure of the system's confidence in how well the two features match. If a single object model has a much larger likelihood than others, the target in the range image is classified as an instance of that object. Besides the classification of the object, object position and orientation information are also available and can be used for higher level scene interpretation.

The set of rules for interpreting image features in terms of models can be divided into two classes according to the level of detail they compare. The first class of rules looks for general features and global characteristics; i.e., the object-level features such as the object length, width, and height. Since the 3-D surface data can be obtained from the range image through a coordinate transformation, the actual length measurements are available and can be compared directly with the true model parameters. A typical rule for object length matching is shown in the first rule of Figure 3. The rule assigns a negative likelihood to models that exceed the tolerance interval and prunes these objects from further consideration. For those object models within the tolerance interval, the rule returns a likelihood value as the goodness measure. Another set of general features is the extreme positions of the image object. For example, the lowest components of a truck are its wheels. The second rule in Figure 3 is one example. Other extreme positions such as the heights of the front and rear points of the side-view projection image can also be used for comparison. Note that these rules are domain-independent and can be derived from the object models.

1. Object Length

<u>if</u> |LENGTH(FEATURE) - LENGTH(MODEL)| > DELTA

<u>then</u> -.5

<u>else</u> $.5\left(1 - \frac{|\text{LENGTH(FEATURE)} - \text{LENGTH(MODEL)}|}{\text{LENGTH(MODEL)}}\right)$

2. Minimum Height Point Location

<u>if</u> #POINTS(MIN.REAR.HEIGHT(MODEL)) > 0

<u>and</u> #POINTS(MIN.REAR.HEIGHT(FEATURE)) > 0

<u>then</u> .5

<u>else</u> .-5

3. Relative Cylinder Orientation

<u>if</u> |RELATIVE_ORIENTATION(FEATURE) - RELATIVE_ORIENTATION(MODEL)| > THETA

<u>then</u> -0.5

<u>else</u> 0.3

Figure 3: Example Interpretation Rules

The second class of rules compares finer object details at the cylinder level. The system first tries to extract a single cylinder from the range image by some heuristic rules and predictions of invariants (e.g., antiparallels, angles). Once a cylinder is extracted, its 3-D properties (length, width, length, and volume) and relative position and orientation in the object coordinate system are compared with the model. These cylinder-level features not only provide finer detail for feature-to-model matching, but also put strong constraints on the internal structure of the object. These constraints are often sufficient to make a unique interpretation of the image object.

V PROCESSING EXAMPLE

To assess the feasibility and capability of rule-based interpretation for classifying vehicle targets from extracted 3-D features, a sample set of rules was developed and tested on extracted 3-D image feature information obtained by using the techniques described in [3]. A synthetic laser range image of a missile launcher decoy is shown in Figure 4. This image has size 64x64 and is provided by The Analytic Science Corporation. The sensor viewing angle is known, but nothing is assumed about the decoy's orientation. The approach used for 3-D feature extraction is to first transform the range image (in a sensor-centered coordinate system) to the surface data (in a world coordinate system) from knowledge of the sensor position. The object is then separated from the background by an object-ground segmentation algorithm. Once the object segment is extracted from the image, the ground projection and side-view projection images of the object segment are generated. These projection images are useful for extracting gross object features and major object structures. The object orientation can be estimated from the orientation of the ground projection image, since vehicle targets are usually elongated. The side-view projection image can be used to locate major object structure positions such as wheel and missile positions of a missile launcher. After extracting those global features, a 3-D edge detection algorithm is used to extract physical edge segments for fine feature-to-model matching. Our 3-D edge feature extraction algorithm directly calculates the physical angle of the object surface from surface data. Convex and concave edges can be distinguished according to the value of the physical edge angle. This physical edge angle image is not only useful for physical edge detection, but also provides relative surface orientation information for extracting planar and curved surfaces. The extracted occluding edge segments (in solid line) and concave edge segments (in dashed line) are shown in Figure 5.

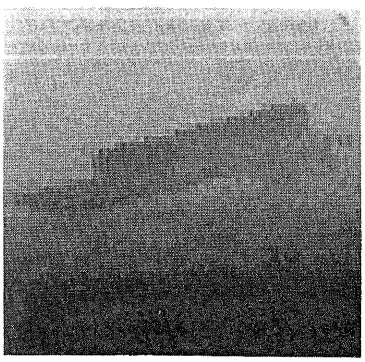


Figure 4: Synthetic Range Imagery of a Missile Launcher Decoy

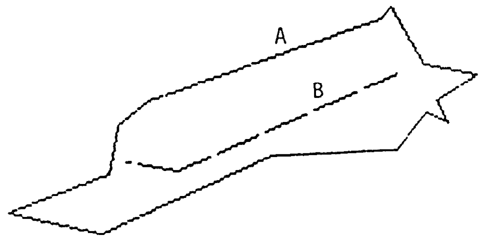


Figure 5: Line Segments of Occluding and Concave Edges of Figure 4

The rules are then applied to compare this extracted feature information with two vehicle models. These were a missile launcher and a decoy. They are chosen because their similarities in size and structure make the selection and correct classification a non-trivial task. The results are shown in Figure 6. The first three rules check the object's length measurements and the system prefers the decoy slightly. No classification can be made at this stage. The second set of rules checks the general features of each model and tries to resolve the object's front and rear ambiguity since we are not sure which end of the image feature is the vehicle front. Reasonable classification is achieved at this level. These seemingly simple rules are able to classify similar objects as a result of rich 3-D information provided by the range data.

RULE	MISSILE LAUNCHER		DECOY	
OBJECT LENGTH	.443		.489	
OBJECT WIDTH	.418		.499	
OBJECT HEIGHT	.400		.489	
SUBTOTAL	1.261		1.477	
	FRONT	REAR	FRONT	REAR
MINIMUM HEIGHT LOCATION	.5	-.5	.5	-.5
OBJECT FRONT EXTREME HEIGHT	-.5	-.5	.48	.48
OBJECT REAR EXTREME HEIGHT	-.5	-.5	.46	.46
TOTAL	.711	-.239	2.917	1.917

Figure 6: Likelihood Weights Associating Rules and Object Models

If the gross object features and major structures do not provide sufficient evidence for classification, the system tries to extract finer details at the cylinder level based on model predictions. The cylinder extraction algorithm finds the major cylinder boundaries along the cylinder axis by using the prediction that these two segments will appear parallel in the image and that one of them is a concave edge. This prediction constrains the possible edge segments for the major cylinder boundaries. The cylinder extraction algorithm successfully finds that two edge segments (with labels A and B in Figure 5) satisfy the cylinder prediction and that they have a significant amount of overlap between them. Using these two edge segments as two sides, the cylinder extraction algorithm tries to find the complete cylinder contour. It first determines the direction of edge segment B (a concave edge can have two edge directions) from the direction of edge segment A (occluding edge) and decides that B should be used as an occluded edge for the cylinder we consider. Then it labels one end point of segment A as the starting point and one corresponding end point of segment B as the goal point (in the correct case, the goal point is closer to the end point of segment B). The algorithm starts from the starting point and searches for nearby edge segments with the correct direction. If one nearby edge segment has the correct direction and is of sufficient length, heuristic rules are used to decide whether to include the segment in the cylinder contour. Examples of these heuristics are: 1) the convexity rule for checking the convexity of the cylinder contour angle and 2) the goal distance rule for avoiding the new starting point being too far away from the goal point (compared to the distance between the previous starting and goal points. The algorithm proceeds until the new starting point is the goal point. Then the algorithm defines the other end point of segment A as the starting point and the other end point of segment B as the goal point and starts over again. This cylinder extraction algorithm successfully finds a complete cylinder contour from the edge segments. The length, width, and height of this cylinder are then extracted by the same techniques used for gross object feature extraction on the object level. These features strongly restrict the possible object models for matching. Future research efforts are required to extend this hierarchical cylinder extraction algorithm to deal with partially occluded cylinders as discussed in the cylinder-level prediction section.

VI CONCLUSIONS

A limited prototype system for range image interpretation has been developed. Important issues on 3-D image feature extraction, model representation, prediction, and interpretation are discussed. Directions for future research on occluded cylinder extraction and prediction are proposed. Because the approach is domain-independent and somewhat independent of the types of range sensor used, it can be applied directly to other applications such as robot vision and autonomous vehicles.

REFERENCES

[1] Agin, G.J. and T.O. Binford, "Computer Description of Curved Objects." IEEE Trans. Computer, C-25, (1976) 439-449.

[2] Brooks, R.A., "Symbolic Reasoning Among 3-D Models and 2-D Images." Artificial Intelligence, Vol. 17, (1981) 285-348.

[3] Kuan, D.T., "Three-dimensional Feature Extraction," IEEE Computer Vision and Pattern Recognition Conference, Arlington, VA, June 1983.

[4] Kuan, D.T. and R.J. Drazovich, "Intelligent Interpretation of 3-D Imagery," AI&DS Tech. Report 1027-1, Mountain View, CA, February 1983.

[5] Nevatia, R. and T.O. Binford, "Description and Recognition of Curved Objects," Artificial Intelligence, Vol. 8, pp. 77-98, February 1977.

[6] Newman, W.M. and R.F. Sproull, Principles of Interactive Computer Graphics. New York: McGraw-Hill, 1973.

[7] Oshima, M. and Y. Shirai, "A Scene Description Method Using Three-Dimensional Information," Pattern Recognition, Vol. 11, pp. 9-17, 1979.

[8] Sugihara, K., "Range Data Analysis Guided by a Junction Dictionary," Artificial Intelligence, Vol. 12, pp. 41-69, 1979.

[9] Weiler, K., and P. Atherton, "Hidden Surface Removal Using Polygon Area Sorting," Computer Graphics, Vol-11, pp. 214, Summer 1977.