IU at the University of Utah: Building 3-D Models from Sensed Data

William B. Thompson and **Thomas C. Henderson**
Department of Computer Science
University of Utah
Salt Lake City, UT 84112
`http://www.cs.utah.edu/projects/robot/`

Abstract

Current work on image understanding at the University of Utah is focused principally on using computer vision and related sensing techniques to aid in problems associated with the construction of geometric models from sensed data. Projects are now underway in high-precision, constraint-based model construction; sensor modeling for virtual environments; interpretation of sonar range data; and calibrated image generation.

1 Overview

The current emphasis of image understanding research at the University of Utah is on improved modeling techniques and sensing strategies for recovering geometry information about physical objects and environments. We are concentrating largely on man-made structures. This allows the use of powerful domain-specific knowledge to be used in the modeling process. Man-made objects are designed for a purpose and exhibit properties that reflect both that purpose and conventional design and construction practices. Exploiting this information can yield models that are more useful and more accurate than would otherwise be possible.

Specific activities include:

- *High-precision, constraint-based model construction.*
  - Model construction.
    Domain-specific information can be used to increase the geometric accuracy of models reconstructed from sensed data.
  - Quantifying modeling accuracy.
    Meaningful measures of modeling accuracy require domain-specific information about the relevance and intent of geometric features.
  - An application in reverse engineering of spare parts.
    Geometric modeling reconstruction from sensed data can be used to generate CAD models from existing parts, providing an important tool for DOD maintenance and repair activities.
- *Sensor Modeling for Virtual Environments.*

  Virtual environments for training and design require the simulation of virtual sensors that behave like their physical counterparts.
- *Sonar.*

  Inferring spatial structure from sonar arrays involving wide field of view sensors requires sensor models quite different from those normally associated with range images.
- *Calibrated Image Generation*

  Images of objects with known shapes in known positions, taken with calibrated cameras, will be made available in order to provide test data for image understanding systems performing classification, pose estimation, and stereo surface reconstruction operations.

2 High-Precision, Constraint-Based Model Construction from Sensed Position Data

The creation of geometric scene models is an essential component in computer vision systems for tasks ranging from reconnaissance to robot navigation. To date, only limited attention has been paid to the accuracy of the shape information recovered from vision and range sensors. This has the effect of limiting the usefulness of vision-based model construction for applications in which geometric precision is critical to performance.

Our emphasis is on the creation of accurate and useful geometric models from sensed data about the position of 3–D surface points, though the methods we describe have application to other sensing modalities as well. Geometric representations and inference methods used by most image understanding systems are designed for generality. In contrast, we approach the model generation problem with representations and data fitting methods specific to the domain of problems

This work was supported by the Advanced Research Projects agency under Army Research Office grant number DAAH04-93-G-0420 and by National Science Foundation grant IRI-9196146, with partial funding from the Defense Advanced Research Projects Agency.

for which the model is intended. This leads directly to significant improvements in modeling accuracy. For systems which are not fully automated, this approach also facilitates the creation of systems that are much more natural and easy to use by end-users.

2.1 Using geometric constraints to increase modeling precision

Modeling accuracy depends on effective use of properties that distinguish the geometry of interest from effects due to sensing errors. Standard signal processing approaches can deal with this problem only for simple and statistically well behaved signals and sensor noise. The sensors used to acquire data with which to construct a geometric model are seldom so well behaved. Compensating for this is the fact that the "signal" (i.e., the actual shape for which a model is desired) has a great deal of domain-specific structure. Each application domain will have shapes that are common and shapes that are unlikely or impossible. This is particularly true when the objects and environments being modeled are man-made.

Domain-specific geometric structure can be exploited most easily in the modeling process if the domain-specific modeling primitives are utilized. When used to model man-made structures, such representations will be far more natural to practitioners in the domain than will more generic primitives. In addition, most application areas involving geometric representations have a rich, pre-existing set of domain-specific tools which only operate with the appropriate representations. There is another advantage to the use of domain-specific representations that may be less obvious but is the key to obtaining better precision. When fitting models to noisy sensor data, the best noise immunity is usually obtained by using modeling primitives with the fewest degrees-of-freedom required to describe the shapes of interest [Thompson *et al.*, 1996a]. The geometric primitives commonly used in image understanding systems are either intrinsically unable to represent many shapes accurately or are so general as to be able to represent essentially any shape equally well. In the latter case, the representation gives no help in pulling out the underlying shape, since in and of themselves they provide no information to distinguish signal from noise.

When modeling man-made objects or environments, the use of domain-specific geometric primitives gives additional advantages (Figure 1). Man-made objects and environments are designed by people to serve some purpose. For a given domain, the design process is almost always characterized by a set of widely used common practices. While not all geometry within the domain necessarily satisfies this set of *pragmatic* constraints, most of it usually does. To take a simple example, it would be foolish to model an indoor building environment without using the knowledge that most large surfaces are likely to be flat and organized in a rectilinear manner. The geometric primitives used in design systems often reflect these pragmatics. As a result, the introduction of pragmatic constraints into systems which recover models from sensed data is facilitated by using the same primitives.

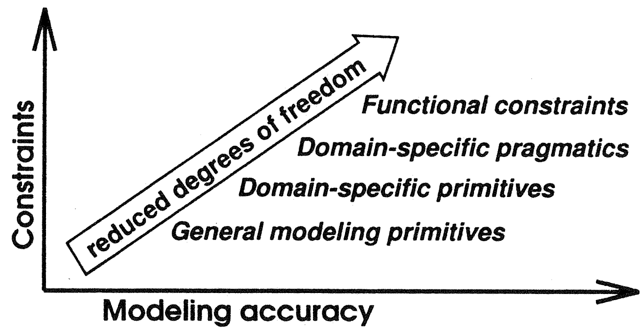

Figure 1: Constraints reduce degrees-of-freedom and increase modeling accuracy.

Domain-specific modeling primitives can lead to an even more powerful type of constraint for many applications. The shape of man-made objects is designed to fulfill a specific *intent*. In fact, one of the most important themes in the development of modern CAD systems is the use of representations which combine geometry and intent [Cunningham and Dixon, 1988, Shah, 1991]. In manufactured parts, a hole in which a bearing assembly is to be installed needs to be treated very differently from a portion of metal removed in order to lighten the part. In an indoor building environment, hallways and offices have somewhat similar geometric properties but very different function. Automated recognition of design intent given only the finished product is well beyond the state-of-the-art. Human users in an application domain, however, can often make good guesses as to intent. Interactive model generation systems can be given this information, which will often provide additional constraints on the underlying shapes being modeled.

2.2 Determining the similarity of geometric models

In order to evaluate the effectiveness of methods which create geometric models of objects from sensed data, it is necessary to be able to quantify the quality of the reconstructed model. This requires a measure of how closely the reconstructed model approximates the "true" geometry. Image understanding methods for surface reconstruction typically use functional approximation, minimizing an L_2 or L_∞ norm by making relatively simple assumptions about the nature of the possible surface shapes and corrupting sensor noise. Shape comparisons using functional norms do not easily extend to complex objects made up of multiple surface patches. Not the least of the problems is that for two models to be compared, they must first be represented in a common coordinate system. Thus, the registration between two models is a central part of the comparison process. Complex shapes, particularly those involving significant concavities, introduce discontinuities into the measurement space, further complicating the computations.

Comparing geometric models only in terms of the closeness of corresponding surface points ignores two issues critical to evaluating methods for reconstructing models from sensed

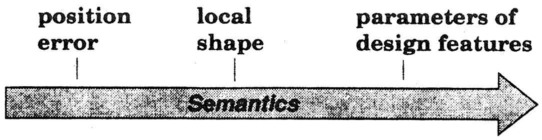

Figure 2: Meaningful measures of accuracy depend on semantic information about the task.

data:

- What are meaningful ways to compare similarity?
- What is the "true" model being reconstructed?

When generating models of man-made objects, it is seldom meaningful to evaluate precision independent of an understanding of the function of the objects and environments being modeled. At a minimum, geometric accuracy will almost always be more important over some portions of objects than over others. Other shape properties such as surface roughness are not well captured by norm-like measure of geometric tolerances. Whether or not such *local shape* properties should be considered in quantifying the accuracy of a reconstructed model can only be determined in the context of the purpose for which the modeling is being done. Since man-made objects are often naturally described in terms of hierarchical structures closely tied to design intent, some measures of modeling quality are only possible by considering these more complex descriptions. In general, the more meaningful measures of modeling precision require information about the *semantics* of the model and the problem domain, not just the relevant geometry [Thompson *et al.*, 1996b] (Figure 2).

2.3 Applications in reverse engineering for support of DOD maintenance and repair activities

We have demonstrated the usefulness of domain-specific geometric modeling primitives in an application relevant to the reverse engineering of mechanical parts [Owen *et al.*, 1994, Thompson *et al.*, under review]. (For related methods, see [Sobh *et al.*, 1994a, Sobh *et al.*, 1994b, Sobh and Owen, 1995, Sobh *et al.*, 1995]). The Department of Defense has a formidable problem maintaining a large number of hardware systems. Spare parts inventories can be exhausted well before de-commissioning of the relevant pieces of equipment. Additional spares are often difficult or impossible to obtain from the original suppliers of the equipment. Complicating the problem, a substantial portion of the contracts under which DOD hardware has been acquired have failed to require documentation sufficient for another supplier to replicate needed parts.

One solution to this problem is to create new parts based on an analysis of existing parts. *Reverse engineering* techniques can be used to create CAD models of a part based on sensed data acquired using three-dimensional position digitization techniques. Part-to-CAD reverse engineering allows

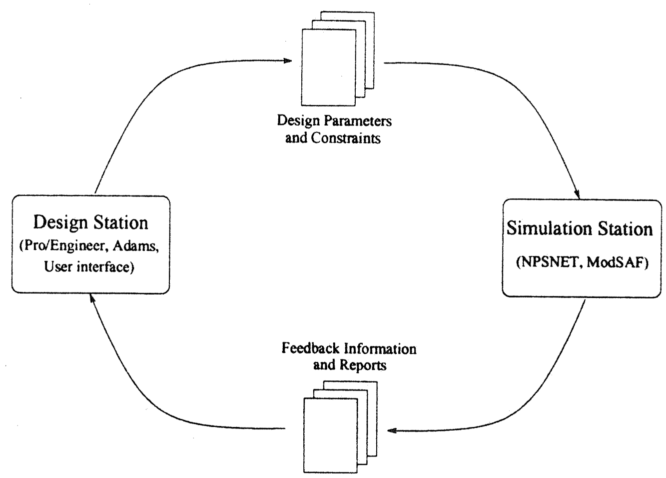

Figure 3: Simulation-based design cycle.

up-to-date NC fabrication plus easier modification of the design than would otherwise be possible. While commercial systems exist to assist in this process, they are hard to use and often fail to produce a CAD model of sufficient quality.

Our approach to reverse engineering uses manufacturing design features [Drake and Sela, 1989] as geometric primitives. As a result, we can generate models that can be used in the extensive collection of feature-based tools that exist – something not possible with any of the commercial reverse engineering systems. In addition, we are able to utilize pragmatic and functional constraints to significantly improve modeling accuracy over what would otherwise be possible.

3 Sensor Modeling for Virtual Environments

Sensor modeling has been extensively studied for autonomous and semi-autonomous systems which interact with the physical world. Sensor modeling is equally important in virtual reality systems, though it has received far less attention in that context. In a virtual reality system, a user interacts with a synthetic environment via an immersive human-computer interface. This interface requires real sensors to determine actions of the users and virtual sensors to correctly present to the user the effects of interacting with the simulated environment.

The uses of virtual environments for training purposes are well known. Increasingly important is the use of virtual environments in *simulation-based design*, in which virtual reality tools are used for modeling, testing and analyzing new systems before attempting to build them. This is particularly useful if creating the virtual world is cheaper or less hazardous that implementing a prototype. Simulation-based design uses standard design processes to "fabricate" a simulation. Users interact with the simulation in the same way that they would with a physical prototype, providing input on necessary refinements to the design (Figure 3).

Virtual environments for either training or simulation-based design must behave in a manner that accurately mimics the physical world being simulated. Figure 4 shows the

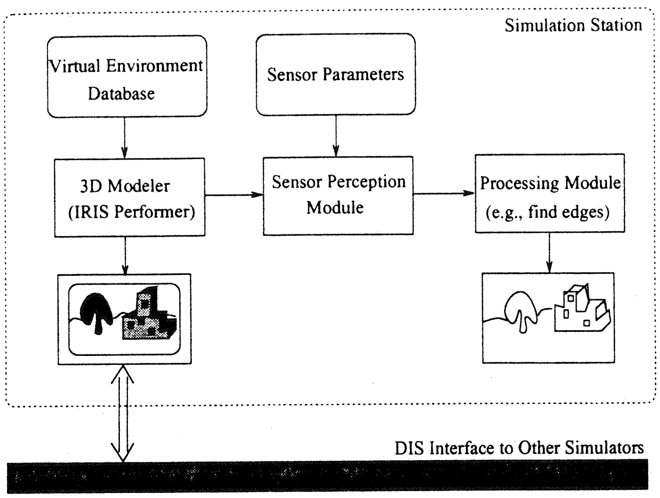

Figure 4: Sensing modules for visual presentation.

processing modules involved in generating appropriate visual displays. More than simple graphical rendering is involved. The effectiveness of the simulation will often depend on accurately recreating:

- *Camera specifications.* Geometry, resolution, distortion.
- *Communications channel effects.* Bandwidth and update rate, noise.
- *Environmental effects.* Time of day, lighting, weather.

[Dekhil *et al.*, 1996] discusses a system currently under construction for easily specifying these properties and then using them to quickly generate the appropriate simulation modules.

4 Sonar

Sonar sensors are frequently used to help guide mobile robots in indoor environments. Arrays of these sensors are often treated as if they were a range imaging device, albeit with a low angular resolution compared to the resolution for range. In fact, there are important differences between the imaging model needed to analyze sonar arrays and that appropriate for true range images. Individual sonar sensors have a much wider field of view than is normally associated with a "pixel" in a range or visual image. For example, the commonly used Polaroid sensor returns the distance to the closest reflecting point within a 22.5° wide cone.

Figure 5 shows a simple example of how this affects the interpretation of the readings from elements in a sonar array. The figure shows a view from above of range measurements to a flat wall using a wide-beam and a narrow beam sensor. In either case, there is a family of wall positions and orientations consistent with a single measurement. Note, however, that the family is different depending on which sensor is used. The computation needed to determine the actual wall position and orientation, given two wide field of view sensors of the sort shown on the left in Figure 5 is described analytically in

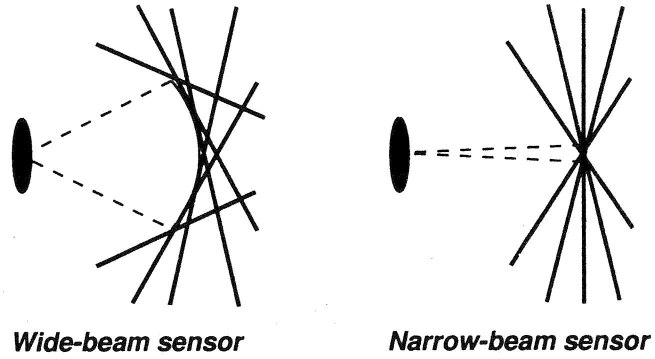

Figure 5: Sensing modules for visual presentation.

[Henderson *et al.*, 1996b, Henderson *et al.*, 1996a]. One important consequence of the analysis is an understanding of the spatial relationship between two wide field of view sensors needed to resolve positional ambiguities associated with flat surfaces. The requirements for sensor positioning turn out to be very different from those associated with narrow field of view sensors.

5 Calibrated Image Generation

Though often discussed, calibrated data with which to quantitatively evaluate the performance of general purpose computer vision algorithms is still not widely available. The creation of imagery for this purpose requires more than just camera calibration. Since the basic purpose of vision systems is to create a description of the scene under view, evaluation necessarily requires information about the "ground truth" nature of the scene.

We are creating a data set consisting of imagery and sufficient collateral information to support the evaluation of computer vision methods for model-based and exemplar-based object recognition and pose estimation and for depth reconstruction from binocular stereo [Owen *et al.*, 1996]. Standard visual calibration methods are used to generate camera models [Faugeras, 1993]. Unique to this effort, most of the objects will be designed and manufactured in our own facility [Thompson and Owen, 1994]. This means that we not only know the true geometry, we can manipulate it in any way that we want in order to test aspects of vision algorithms. In addition, we can make accurate replicas of each object so that physical objects can be distributed to sites interested in evaluating active vision systems that can not be tested on pre-acquired imagery. The parts making up this collection of objects are shown in Figure 6. Models for each object are generated directly from the CAD system used to design them, and are available in a variety of formats.

Object pose is determined with a coordinate measuring machine (CMM), which is a precision contact position sensing device. Objects are placed in the field of view of the cameras in an arbitrary orientation. Known locations on each visible object are measured using the CMM. Together with the CAD model of the object, this is sufficient to specify orientation. Fi-

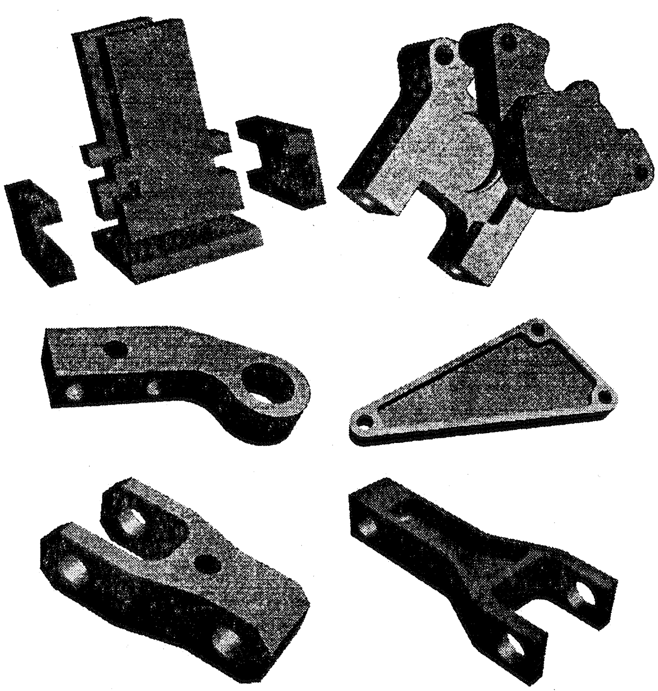

Figure 6: Objects used to create the calibrated imagery data set.

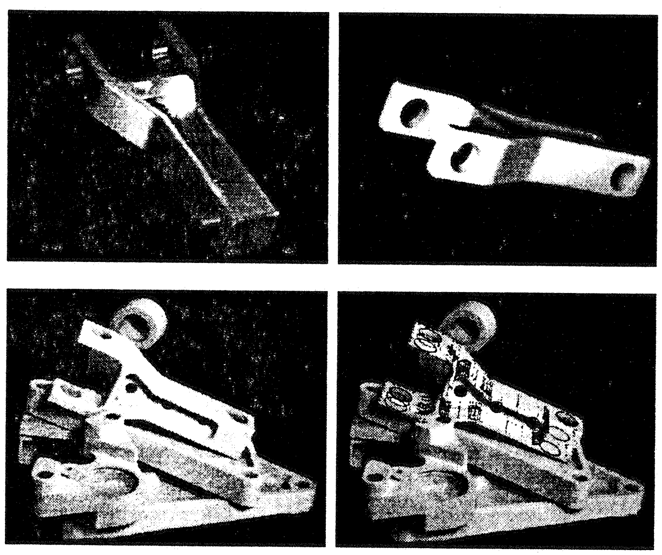

Figure 7: Calibrated imagery.

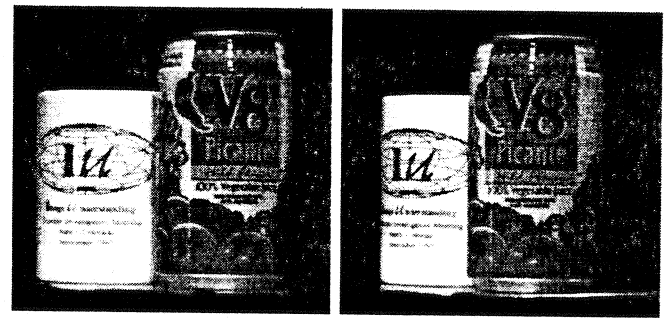

Figure 8: Stereo image pair.

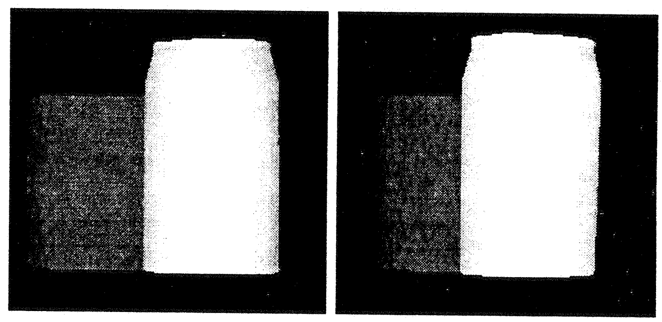

Figure 9: Intensity-coded range images corresponding to Figure 8

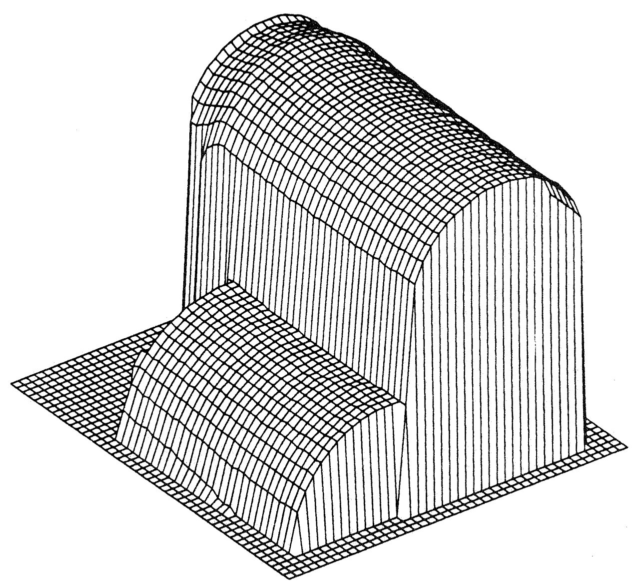

Figure 10: Depth map associated with right image in Figure 8

nally, the orientation is represented in the camera coordinate systems by measuring the position of the optical camera calibration target and solving for the appropriate transformation.

Figure 7 shows examples of imagery in the data set. The upper two frames are images of a single part in two different orientations. The parts are made of machined aluminum. On the left, the view is of a shiny metal part. The upper right image shows the part after it has been painted with a talc like powder. The data set includes both painted and unpainted versions of each view. The lower left image in Figure 7 shows a "jumble" of parts. To verify our calibration and pose measuring procedures, in the lower left frame the model of one part has been back projected onto the image to show that it lines up accurately with the image of the same object.

To provide more imagery with which to evaluate stereo algorithms, the data set includes stereo pairs and "true" depth maps for objects in addition to those shown in Figure 6. Figures 8–10 show one example. The correct range images for these stereo pairs are generated by using a laser scanner to measure the 3–D position of a dense sampling of surface points on the objects in the field of view. These 3–D points are then back projected through the camera models for each camera and a hidden surface algorithm is used to generate synthetic range images.

References

[Cunningham and Dixon, 1988] J. J. Cunningham and J. R. Dixon. Designing with features: The origin of features. In V. A. Tipins and E. M. Patton, editors, *Computers in Engineering*. ASME, 1988.

[Dekhil *et al.*, 1996] M. Dekhil, J. Marti, T. C. Henderson, and S. Jacobsen. Sensing in the virtual world – work in progress. In *Proc. ARPA Image Understanding Workshop*, February 1996.

[Drake and Sela, 1989] S. Drake and S. Sela. A foundation for features. *Mechanical Engineering*, 111(1), January 1989.

[Faugeras, 1993] O. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, Cambridge, MA, 1993.

[Henderson *et al.*, 1996a] T. C. Henderson, B. Brüderlin, M. Dekhil, L. Schenkat, and L. Veigel. Sonar sensing strategies. In *Proc. IEEE International Conference on Robotics and Automation*, April 1996.

[Henderson *et al.*, 1996b] T. C. Henderson, M. Dekhil, L. Schenkat, and L. Veigel. Flat surface reconstruction using minimal sonar readings. In *Proc. ARPA Image Understanding Workshop*, February 1996.

[Owen *et al.*, 1994] J. C. Owen, P. P. J. Sloan, and W. B. Thompson. Interactive feature-based reverse engineering of mechanical parts. In *Proc. ARPA Image Understanding Workshop*, pages 1115–1124, November 1994.

[Owen *et al.*, 1996] J. C. Owen, H. J. de St. Germain, S. Stark, and W. B. Thompson. Calibrated imagery for quantitative evaluation of IU pose-estimation and stereo algorithms. In *Proc. ARPA Image Understanding Workshop*, February 1996.

[Shah, 1991] J. J. Shah. Assessment of features technology. *Computer-Aided Design*, June 1991.

[Sobh and Owen, 1995] T. M. Sobh and J. C. Owen. A sensing strategy for the reverse engineering of machined parts. *Journal of Intelligent and Robotic Systems*, August 1995.

[Sobh *et al.*, 1994a] T. M. Sobh, J. C. Owen, and M. Dekhil. A dynamic recursive approach for autonomous inspection and reverse engineering. *Journal of Robotics and Autonomous Systems*, October 1994.

[Sobh *et al.*, 1994b] T. M. Sobh, J. C. Owen, M. Dekhil, C. Jaynes, and T. C. Henderson. Industrial inspection and reverse engineering. In *Proc. IEEE 2nd CAD-Based Vision Workshop*, February 1994.

[Sobh *et al.*, 1995] T. M. Sobh, J. C. Owen, M. Dekhil, and T. C. Henderson. Industrial inspection and reverse engineering. *Computer Vision, Graphics, and Image Processing – Image Understanding*, May 1995.

[Thompson and Owen, 1994] W. B. Thompson and J. C. Owen. "Hard-copy" benchmark suite for image understanding in manufacturing. In *Proc. ARPA Image Understanding Workshop*, pages 221–227, November 1994.

[Thompson *et al.*, 1996a] W. B. Thompson, H. J. de St. Germain, T. C. Henderson, and J. C. Owen. Constructing high-precision geometric models from sensed position data. In *Proc. ARPA Image Understanding Workshop*, February 1996.

[Thompson *et al.*, 1996b] W. B. Thompson, R. F. Riesenfeld, and J. C. Owen. Determining the similarity of geometric models. In *Proc. ARPA Image Understanding Workshop*, February 1996.

[Thompson *et al.*, under review] W. B. Thompson, J. C. Owen, and H. J. de St. Germain. Feature-based reverse engineering of mechanical parts. *Computer-Aided Design*, (under review).