

# Report on Range Image Understanding Workshop, East Lansing, Michigan, March 21–23, 1988

## Ramesh Jain

Artificial Intelligence Laboratory, Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, Michigan, USA

## Anil K. Jain

Computer Science Department, Michigan State University, East Lansing, Michigan, USA

|  |    |  |    |
|--|----|--|----|
| <b>Contents</b>  |    | 4.1 Matching .....   | 53 |
|  |    | 4.2 Modeling .....   | 53 |
| <b>1 Introduction</b>                                    | 46 | <b>5 Sensor Integration</b>  | 54 |
| 1.1 Format of the Workshop .....                         | 46 | <b>6 Range Sensing for Navigation</b>  | 54 |
| <b>2 Issues in Sensing and Sensors</b>                   | 47 | 6.1 System Parameters, Navigational Tasks,<br>and Representation.....                  | 54 |
| 2.1 General Background .....                             | 47 | 6.2 Case 1: An Underwater Surveyor .....   | 55 |
| 2.2 Popular 3-D Range Sensors.....                       | 47 | 6.2.1 Sensors .....  | 56 |
| 2.2.1 Stereo .....                                       | 47 | 6.2.2 Representation .....   | 56 |
| 2.2.2 Structured light.....                              | 47 | 6.2.3 Sensor fusion, object (landmark) rec-<br>ognition, finding navigable space ..... | 56 |
| 2.2.3 Time-of-flight .....                               | 47 | 6.2.4 Self-localization .....  | 56 |
| 2.3 Other 3-D Sensing Techniques .....                   | 48 | 6.2.5 Obstacle detection.....  | 56 |
| 2.4 Needs of Five Major Application Areas ..             | 48 | 6.2.6 Map creation.....  | 56 |
| 2.5 Example: ERIM Range Sensor Specs .....               | 48 | 6.3 Case 2: Surveying an Urban Environment.  | 56 |
| 2.6 Status of Moire Technology .....                     | 48 | 6.3.1 Sensors .....  | 56 |
| 2.7 Commonly Cited Problems in Range Sens-<br>ing.....   | 48 | 6.3.2 Representation .....   | 57 |
| 2.8 Future Efforts .....                                 | 49 | 6.3.3 Early processing .....   | 57 |
| <b>3 Early Processing</b>                                | 49 | 6.3.4 Sensor fusion .....  | 57 |
| 3.1 Issues in Early Processing of Range Im-<br>ages..... | 49 | 6.3.5 Object (landmark) recognition .....  | 57 |
| 3.2 Definition of “Early” Processing.....                | 50 | 6.3.6 Self-localization .....  | 57 |
| 3.3 Surface Geometry .....                               | 50 | 6.3.7 Servoing, obstacle detection, and<br>navigable space .....                       | 57 |
| 3.4 Early Processing Algorithms .....                    | 51 | 6.3.8 Map creation.....  | 58 |
| 3.5 Summary.....   | 52 | <b>7 Applications Group Report</b>   | 58 |
| <b>4 Object Recognition</b>                              | 52 | <b>8 Appendix</b>  | 59 |
|  |    | 8.1 Overview Speakers .....  | 59 |
|  |    | 8.2 List of Participants.....  | 59 |
|  |    | 8.3 Workshop Groups and Group Chairs.....  | 60 |

*Address reprint requests to:* Ramesh Jain, Artificial Intelligence Lab, EE and CS Depts., The University of Michigan, Ann Arbor, MI 48109-2122.

This workshop was sponsored by the National Science Foundation.

## 1 Introduction

Most computer vision research has concentrated on using digitized gray-scale intensity images as sensor data. It has proven to be extraordinarily difficult to program computers to understand and describe these images in a *general purpose* way. One important problem is that digitized intensity images are rectangular arrays of numbers which indicate the *brightness* at individual points on a regularly spaced rectangular grid and contain no *explicit* information that is directly usable in *depth perception*. Yet human beings are able to correctly infer depth relationships quickly and easily among intensity image regions whereas automatic inference of such depth relationships has proven to be remarkably complex. In fact, many famous visual illusions, such as Kanizsa's triangle, vividly demonstrate that humans impose 3-D surface structure on images to interpret them. Computer vision researchers recognized the importance of surfaces in the understanding of images. The popularity of *shape from . . .* approaches in the last decade is the result of this recognition.

In recent years digitized range data has become available from both active and passive sensors, and the quality of this data has been steadily improving. Range data is usually produced in the form of a rectangular array of numbers, referred to as a *depth map* or *range image*, where the numbers quantify the distances from the sensor plane to the surfaces within the field of view along the rays emanating from a regularly spaced rectangular grid. Not only are depth relationships between depth map regions explicit, but the three-dimensional *shape* of depth map regions approximates the three-dimensional shape of the corresponding object *surfaces* in the field of view. Therefore, the process of recognizing objects by their shape *should* be less difficult in depth maps than in intensity images due to the explicitness of the information. For example, since correct depth map information depends only on geometry and is independent of illumination and reflectivity, intensity image problems with shadows and surface markings do not occur. Nonetheless, it seems that existing vision techniques have influenced many investigators and this has led to restricted approaches to processing range data. The range image understanding problem is a well-posed problem as contrasted with the ill-posed intensity image understanding problem.

Many tasks for machine vision systems will be significantly simplified by range images. The last few years have seen growing interest in the appli-

cation of the range images in industrial applications. It is expected that the vision systems based on range images will be very useful in autonomous navigation systems also.

Considering the growing interest in range image understanding, the NSF funded our proposal to organize a workshop on this topic. The aim of the workshop was to bring together active researchers in a range image understanding to discuss emerging themes in this field that may play a key role in automation. In addition to the researchers in range image understanding, we invited a few researchers interested in designing range sensors also. The participants were drawn from academic institutions, research centers, and industry. The aim of the workshop was to identify major research areas that should be addressed by researchers for making range image understanding systems useful in industrial and other applications.

### 1.1 Format of the Workshop

The workshop lasted two and a half days. The emphasis in this workshop was on group discussions. The workshop began with seven overview talks, one on general range image understanding by Ruzena Bajcsy, and one on each of the major areas in range image understanding and its applications that are active:

- Early processing (Paul Besl)
- Object recognition (Robert Bolles)
- Sensor integration (Jake Aggarwal)
- Navigation (Takeo Kanade)
- Inspection (Joe Mundy)
- Range sensors (Avi Kak)

These talks were expected to set the tone for the discussions at the workshop.

The second day of the workshop was devoted to group meetings. The research issues were divided into three classes:

- Short term research and development work that may result in applications in the next three to five years.
- Long term research issues that may result in applications in five to ten years.
- High risk areas.

The purpose of the group discussions was to make recommendations for the research in the specific areas.

The recommendations were presented by group leaders on the third day in the general meeting.

Consensus recommendations were formed during that discussion. These recommendations are presented in the following sections.

The list of participants of the workshops, different groups, overview speakers, and group leaders is given in Section 8, the Appendix.

## 2 Issues in Sensing and Sensors

Nelson Corby, GE  
 Kevin Harding, ITI  
 Avi Kak, Purdue  
 Robert Sampson, ERIM  
 George Stockman (Chair), MSU

### 2.1 General Background

Sensors connect intelligent processes, whether in man or machine, to the external 3-D world. Engineering approaches have separated sensor design from the design of the overall intelligent system. In some well-specified applications such design has led to successful systems. Techniques are getting better and cheaper, and there seems to be good potential for many future industrial applications.

For some very difficult problems, such as autonomous navigation or object detection in significant clutter, higher level processes have not yet been proved successful. Therefore, it is not clear exactly what kind of data is necessary and what top-down information must be employed. Uncertainty in the overall system design is thus problematical for specification of the sensing module. A related problem is that of *multisensor fusion*: data from different sensors must be integrated in such a way as to maintain a continuous/cohesive representation of the real world. Fusion may be done at a low level, for instance at the  $(x,y,z)$  *point level*, or at a higher *structural/symbolic level*. A final notion to mention is that of *active versus passive sensing*. A common definition of active sensing would apply to any sensor that provided and controlled its own illumination—such as triangulating with structured light or timing a laser pulse. At a higher level we may define active sensing to mean that the sensing is *goal-directed*; that is, the intelligent consumer of the data is gathering that data for hypothesis testing.

Having made the given caveats, the report proceeds, concentrating on the sensor as a separate system. Top-down control of the sensor via the interface to higher level modules is sometimes addressed, but only to a modest degree.

### 2.2 Popular 3-D Range Sensors

According to Kak, the three most popular 3-D sensing techniques are:

1. Stereo.
2. Structured light.
3. Time-of-flight (sensors which measure phase change, rather than time, as a function of distance traveled are lumped into this category).

The most important characteristics of these techniques are as follows.

**2.2.1 Stereo.** The chief advantage is that it is passive. The main disadvantage is that it apparently cannot produce dense range measurements over uniform regions due to lack of features or presence of too many features for correspondence. The cost of two intensity sensors (cameras) is low, but hardware needed to produce the correspondences add significantly to the cost of a system. There has been a great deal of R&D on stereo sensing, and the technique is of common use in making maps from aerial imagery. Progress in robotics has been disappointingly slow, but researchers seem to agree that there is more potential to tap. In certain industrial environments, it is easy to use active illumination to aid in solving the correspondence problem.

**2.2.2 Structured light.** The main advantages are simplicity and the resulting low costs, making this the most common technique in industry. Various forms of structured light may be projected onto a scene, such as a ray, a sheet, a grid, or even cylinders, etc. Measurements may be dense or sparse depending upon the light pattern and on whether or not it is scanned over the scene. The active illumination can be a disadvantage in natural or hostile environments and even in industrial environments where specular reflections can be a problem. There is also the problem of lack of data due to one object shadowing another from the projected light.

**2.2.3 Time-of-flight.** The main advantage is that a dense and (almost) complete range image is output and that little or no image processing is required to get range. There is also the growing capability of obtaining several bits of registered reflectance data, thus allowing simple sensing of most of the so-called *intrinsic image*. The chief disadvantage is cost. Most sensors developed so far have been *one of a kind* and have six-figure price tags. Lower prices are expected as successful applications increase the market, but chances for low prices are dim because of the costly optical and mechanical parts used in current designs. More details are described in the sections below.

### 2.3 Other 3-D Sensing Techniques

There are other 3-D sensing techniques which are likely to find increasing application. These include:

1. Depth-from-focus.
2. Moire fringe pattern interpretation.
3. The use of acoustic or radar signals to measure distances via timing of returned signals.

### 2.4 Needs of Five Major Application Areas

Different application problems place different requirements on the sensor system. Requirements can be defined in terms of many parameters such as working volume, materials in the scene, object placement, precision of measurements, and measurement rate. A summary of application areas and the spatial resolution and data rates required was done by Nelson Corby and Joe Mundy of General Electric and is given in Table 1.

### 2.5 Example: ERIM Range Sensor Specs

There are several companies that have manufactured time-of-flight range sensors. Most of these have been designed with a specific application in mind. However, the technology is likely to be applied in other cases which were not considered during design. For the purpose of informing readers of current capabilities, the following data are given for the most recent sensor designed at ERIM.

- 16 in.  $\times$  16 in.  $\times$  16 in. envelope
- 12 bit range values of 0.004 in. (0.1 mm) resolution
- Data rate of 0.25M rangels per second
- Reflectance data as well as range

### 2.6 Status of Moire Technology

Moire technology shows good promise for inspection of smooth surfaces. For purposes of informa-

tion, some of the characteristics of a Moire sensor are outlined.

- Advances
  - uses white light image
  - low cost
  - gives *snapshot* data; that is, no scanning needed
- Resolution
  - 12 bits of depth obtainable, for example, 0.0004 in. steps over 2 in.
  - x,y resolution depends on camera used
- Data rate
  - 100K range pts/sec can be obtained
- Limitations
  - maximum surface slope about 30–60 deg
  - surface steps must be below 0.06 in. ambiguity interval
  - shadows created due to different projection and viewing angles
  - limited by dynamic range of reflectance

### 2.7 Commonly Cited Problems in Range Sensing

A diversity of problems and limitations can be cited. These can be related to the components of the equipment, the physical phenomena used, the overall system, or even the environment of the application. A few problems are worthy of special note. First is the problem of dynamic range. For solder joint inspection the dynamic range in surface reflection can approach 100,000 to 1, thus placing severe constraints on the detector design. A second important problem is that of resolution and work volume requirements—sensors typically can achieve a good resolution over a limited work volume. For instance, if a sensor can deliver only a 12 bit value for range, this allows a resolution of 0.004 in. over 16 in. If a dimension of 32 in. is required, then the

**Table 1.** Sensing requirements for various applications

| Application area        |   | Spatial resolution    | Data rate                       |
|-------------------------|---|-----------------------|---------------------------------|
| Integrity and placement | Detect missing or misaligned components in product assemblies; emphasis is on assembly and configuration control. | 0.01 to 0.1 in.       | 100K to 1M range pts/sec        |
| Metrology               | Provides precision measurements of machine parts and electronic patterns.   | 1 to 100 $\mu$ in.    | 10 to 100K range pts/sec        |
| Surface inspection      | Exhaustively scan part and assembly surfaces to detect surface flaws and component defects.                       | 0.1 to 10 mils        | 100K to 100M range pts/sec      |
| Modeling                | Extract a 3-D solid model from multiple range views; model supports design and engineering simulation.            | 0.1 to 10 mils        | 10K to 1M range pts/sec         |
| Navigation              | Provide scene data for mobile robot or autonomous vehicle.  | Perhaps 1 in. to 1 ft | Perhaps 10K to 1M range pts/sec |

sensor must be moved (or reset in some other way). If such scanning must be done to enlarge the work envelope, then the sensor design is complicated and both data rates and overall system precision may be decreased. A related problem is that of providing variable work envelopes in order to allow smarter sensing of larger scenes without excess data or scanning times.

- Component limitations
  - dynamic range of light intensity can be 100,000 to 1
  - drift due to changes of environment, such as temperature
  - stuck bits or loss of sync often observed
- Design limits
  - multiple reflections
  - shadows
  - ambiguity interval
  - motion of sensor or objects being sensed
  - resolution versus duration of sensing
  - spatial resolution and mixed pixels
  - missing data or artifact
  - power consumption
- System limits
  - sensor-to-application match
  - cost of sensor
  - representation gap between sensed data and application process
  - fusion of range data with other sensor data
  - need for variable sampling in same scan

### 2.8 Future Efforts

Previous sections identified some state-of-the-art technology and some of the problems yet existing in that technology. This section identifies a few lines of attack where new work or continued development effort is needed.

- Breakthroughs needed
  - sensors are needed which give *snapshot* range images, that is, without scanning. These would be immune to motion problems and the high cost components used in current sensors.
  - sensors should work in a closed loop with the higher level processes of image and scene understanding.
- Engineering for breakthroughs
  - continued progress on the development of sensors that deliver registered range and reflectance is needed.
  - sensors are needed which directly deliver higher level primitives such as edges or surface patterns.

- sensors are needed which provide variable speed, resolution, and work envelope.

- Miscellaneous

- more experience is needed in the use of range sensors in applications
- documentation, in the form of texts or handbooks, is needed to educate users in the capabilities of 3-D range sensors and how to fit them to specific application problems.

## 3 Early Processing

The early range image processing (EP) discussion group consisted of:

Gerald Medioni, USC (Chair)

Paul Besl, General Motors

Patrick Flynn, MSU

Richard Hoffman, Northrop

Gerhard Roth, NRC-Canada and McGill

Ishwar Sethi, Wayne State

### 3.1 Issues in Early Processing of Range Images

In a general machine perception system structure, a range imaging sensor is recognized as a single contributor of sensory information within a suite of other possible sensors. Early processing of range image data may take place in isolation from other sensor data, but it is also possible and sometimes desirable to process imaging sensor data from multiple sensors at the pixel level, usually under the constraint that all simultaneously processed pixels come from registered images. Since processing data from multiple sensors falls in the domain of sensor integration algorithms and strategies, the EP group discussed early range image processing in isolation from other types of sensor information.

One “fact” that was painfully apparent in many of the discussions and presentations is that no single range imaging sensor, no single early processing strategy, no single object recognition scheme or navigation technique or inspection algorithm, and no single sensor integration approach will be adequate in the near term for the wide variety of practical applications that would benefit from range sensing technology. As a result, range imaging system design tends to be application driven and will probably remain so until a significant number of successful range imaging systems are being used. In other words, the application requirements will generally specify the type of range imaging sensor that will be used. For example, specifications for range accuracy, depth of field, field of view, image acquisition time, standoff distance, lateral sampling in-

tervals, optical properties of scene materials, and other application specifications can severely constrain the choice of range sensing technology. The range imaging sensor and the image features needed by higher level application-specific processes then place fundamental constraints on the choice of early processing algorithms. The EP group attempted to address many of the common issues involved in potential application.

### 3.2 Definition of “Early” Processing

A definition of early range image processing was needed since the interface between early processing and sensing can be muddled with emergence of smarter sensors. In addition, the distinction between early range image processing and later processing can also be difficult to make. An early range image processing algorithm is any geometric signal processing algorithm that organizes (abstracts) discrete range sensor data into a more convenient form for application-specific (task-oriented) reasoning. Early processing algorithms do not directly address range estimation nor do they infer the existence of objects in the image data. The goal of an early processing algorithm is to accept input from a range imaging sensor and to extract geometric primitives or features relevant to higher level application-dependent processes. Ideally, early processing completely partitions, or segments, the range data into geometric primitives so that almost all image data points are grouped with some geometric representation.

### 3.3 Surface Geometry

It is common in early processing to assume that most visible surfaces of interest are piecewise smooth and more specifically,  $C^2$  almost everywhere (first and second derivatives exist except on sets of zero area: points or curves). This assumption arises from the fact that matter is cohesive. Although visual texture due to surface reflectance variations are common in the everyday world, significant physical surface texture, such as the texture present in trees, mountain ranges, and clouds, is not as common in many potential range imaging application environments. When physical surface texture is important, fractal or Markov random field (MRF) models have been explored and will require further investigation with respect to application requirements.

Assuming an underlying piecewise smooth surface geometry exists that approximates the range image data well, how is it recovered, estimated, or analyzed? Many researchers agree that knowledge of the differential properties of a range image is use-

ful for many purposes. For example, edge detection requires estimation of at least the first partial derivatives, and estimation of second partial derivatives and mixed partials is commonly done. The majority of early processing algorithms seem to lie somewhere between (a) first partial derivative estimation at the simple end of the EP spectrum (e.g. a Sobel operator) and (b) computation of a complete principal (Darboux) frame field for the entire range image at the complex end of the EP spectrum. Knowledge of the principal frame field of a range image implies that one knows the surface normal and principal directions at each point in the image and possesses a line of curvature surface parameterization that approximates the data well. If this information is known, then any other surface information, such as surface curvatures, can be easily computed. The goals of many early processing approaches are summarized as the (a) detection of  $C^0$  (jump),  $C^1$  (crease), and  $C^2$  (curvature) discontinuities and their junctions and (b) the approximation or characterization of the smooth  $C^2$  surface patches whose boundaries are the aforementioned discontinuities. There are many methods to create approximate surface patches without using differential properties of the pixel data. These methods rely mostly on surface fitting.

Differential properties are most often estimated either explicitly or implicitly via local window surface fits (a.k.a. local facet model). Such methods yield adequate results quickly, but the depth, or range, direction is a preferred direction in space. For isotropic computation, surface normals can also be estimated by diagonalizing the covariance matrix of points in a neighborhood and selecting the eigenvector associated with the minimum eigenvalue. Curvature can then be computed from surface normal estimates. The consensus seemed to be that surface normals and curvature can be computed fairly reliably at the majority of points in a range image by either method, but that there are no existing reliable low level methods that work everywhere, especially near crease edges or curvature edges.

In order to estimate differential properties of range image data, it is common to use nonadaptive Gaussian or binomial weighted presmoothing operators [low pass (FIR) filters] to filter out unwanted noise. Linear filters unavoidably round sharp edges, so it is also common to use nonlinear filters that tend to (a) preserve edges and (b) smooth nonedge image regions. The K-nearest-neighbor smoothing approach with  $K = 5$  is a good inexpensive adaptive alternative to constant coefficient  $3 \times 3$  smoothing operators. Other methods exist in the

literature, but none have proven to be the unique best solution for all applications.

### 3.4 Early Processing Algorithms

The goal of any preliminary filtering operator in range image processing is to preserve geometry while attenuating noise. Whereas no combination of smoothing and derivative estimation have proven themselves superior to other techniques, there is a serious lack of an optimality criterion. It was suggested in the early processing group discussions that surface curvature types from synthetic range images might allow a testing criterion. For example, if the ideal surface curvature types are known at every pixel in a range image, any scheme for estimating partial derivatives might be evaluated using the percentage of misclassified pixels. A “perfect” algorithm would have a zero misclassification rate for a series of preselected test images with different types of surface geometry and different levels of noise. In general, standardized segmentation error metrics are needed to help advance the state of the art. No quantitative metrics are measured on standard test images in most of today’s research environments.

The output from an early processing algorithm should be representations of geometric primitives: points, curves, surfaces, and/or volumes. If symmetry is present in the data, an early processing algorithm should recognize the symmetry and encode it in the extracted geometric primitive. For example, lines (dihedral crease edges), circles, planes, spheres, cylinders, cones, and undeformed superquadrics can be recognized directly by low level processes and used to limit the possible degrees of freedom of the objects. Feature points, such as polyhedral vertices, cone tips, and isolated umbilic points, also provide powerful constraints for higher level processes when available. Other types of more general shape primitives might also be extracted, such as B-spline space curves and tensor-product surfaces. Other shape description representations, such as extended Gaussian images (EGIs) and 4-vector EGIs, are also useful for describing elliptic and hyperbolic surfaces. Geometric representations are critical concerns for higher level processes and the selection of particular representations will dictate requirements to early processing algorithms.

Multi-resolution methods were discussed and it was decided by the group that a multi-resolution paradigm is probably not relevant to early range image processing unless it provides computational advantages. That is, the best, finest resolution results should be returned by early range image processing algorithms. If multi-resolution concepts are

useful for later matching representations, then later processes can coarsen the models for the given algorithm or application.

As far as general purpose system integration issues are concerned, a range imaging sensor should return a model of its processes and a model of possible noise sources to an early processing algorithm when queried. For example, range image noise is typically non-normal and is definitely nonstationary and might usually consist of random measurement noise, quantization noise, and systematic sensor noise (e.g., a small amount of ripple) as well as “outlier” noise, which causes bad range readings due to depth discontinuities and finite laser beam width, steep relative surface slopes, specular reflection, absorption, multiple reflections, etc. Similarly, the sensor/noise model and a model of the early processing algorithms and their possible errors should be passed on to higher level processes. A general purpose early processing algorithm will need to adapt itself to situations that depend on which sensor is used and which application is seeking geometric features from the image data. Although such issues may not impact today’s planned practical systems, we must begin thinking along these lines to move toward general purpose systems.

Range images can be represented in (raster)  $r(i,j)$  form (range as a function of two integer indices) or in (scattered) 3-D  $(x,y,z)$  form. Several comments indicated that all early processing should be performed in the  $r(i,j)$  image form since the uncertainty in range  $r$  is much larger than the uncertainty in the actual 3-D ray corresponding to the  $(i,j)$  indices. There may be a global uncertainty about where the range imaging sensor is pointing, but the point-to-point uncertainty is minimal. Hence, all probabilistic second moment information can be well approximated by a  $\sigma_r(i,j)$  range variance image and a  $2 \times 2 \sum_{ij}$  covariance matrix for the overall sensor direction. That is, one number per point plus four numbers per image. In contrast, the  $(x,y,z)$  form of a range image would require a  $3 \times 3$  covariance matrix (nine numbers) for each point. As Kalman filtering algorithms become more common for integrating estimates of point positions from multiple views acquired at different times, these memory and efficiency concerns may be critical. As a general guideline, range image data should be processed in the sensor coordinate frame. Geometric primitives extracted from the image will need to be transformed to the relevant global coordinates with their related uncertainty information.

Is multiple view integration an early process or a later process? It appears to depend on the applica-

tion. For instance, in autonomous land vehicle (ALV) range images, each scan line (and possibly each point) may be coming from a different viewpoint during image acquisition. In such applications, image data must be corrected before other early processing algorithms can begin. Hence, such viewpoint corrections are a key feature of early processing stages of the system. In contrast, an automated geometric modeling system may perform integration of extracted geometric primitives from different views after all the early processing is done. In general, if multiple view integration is required for geometric primitive extraction, then it is early processing; otherwise, it is not. High level multiple view integration algorithms should tie into sensor integration algorithms in a natural unified manner.

A never ending source of difficulty (and often confusion) is the set of early processing algorithm parameters (window sizes, thresholds, tolerances, etc.). It is often difficult to verify the results of others without knowing the algorithm parameters they used. As a general rule of thumb, the number of parameters should be minimized. For each parameter used, the range of values tested and the effects should be clearly documented by those describing the algorithm for others. Besides edge detection thresholds, a classic case is fitting unconstrained quadrics to range data from cylinders. Sometimes a long, but enclosed, ellipsoid is generated; other times, a long nonclosed hyperboloid is generated. A threshold is needed to decide when the surface is a cylinder and when it is something else. In addition to parameter set listings, the EP group agreed that all journals in the field should insist on explicit concise pseudo-code listings or flow charts as algorithm documentation in papers describing algorithms or the results of algorithms.

### 3.5 Summary

In order to make progress as an engineering and scientific discipline, the computer vision community, and the range imaging understanding interest group in particular, should begin to share software and image data in a more open manner. We need to avoid duplication of effort and to enable quantitative comparisons of different approaches. It was suggested that journals require that new algorithms be run on at least one standardized image and compared to at least one other well-accepted algorithm. As examples, the Alvey consortium of universities and industry in England standardized on C and Sun workstations making it possible for a large group to share results and source code. In the numerical analysis community, people can send CSNet messages to netlib at Argonne National Labs and re-

ceive numerical software back in FORTRAN. The workshop group agreed that an image database should be set up to allow easy access to range imagery given the scarcity of good sensors. For software, it is suggested that standard algorithm libraries be begun in (ANSI) C and (Common) LISP languages. Candidates for standard software libraries include the following: adaptive/nonadaptive smoothing and derivative estimation functions; jump/crease/curvature discontinuity edge detection/tracking/linking; region growing based on various uniformity predicates; regression code for curve and surface fitting of various types of geometric entities (e.g., B-splines); standardized rotation matrix manipulation codes (e.g., quaternions, etc.); device independent 3-D display algorithms for points, curves, surfaces, and volumes (silicon graphics interfaces); device independent range image generation algorithms from polyhedra, superquadrics, Bezier patches, or non-uniform rational B-splines (NURBS) for different types of scanner configurations; and geometric operations such as free-form surface-surface intersections and Boolean operations.

Early processing of time-varying range imagery is an area that requires more research. A complicating feature of most range imaging sensors is long image acquisition times (greater than 0.25 sec). This is a general feature of most range imaging sensors that will not be alleviated in the near term except at great expense in sensor technology. Dynamic algorithms must be able to handle this artifact of range imaging sensors.

## 4 Object Recognition

The object recognition discussion group consisted of:

Thomas O. Binford (Chair)  
 Robert Bolles  
 Francis Quek  
 Richard Weiss

Recognition and interpretation integrate vision system functions. Recognition must account for results from preceding analysis, that is, sensing, early analysis, and any multisensor integration. Figure 1 shows a convenient hierarchy in model based vision. It shows dimension of representation, not control. On the left side is modeling, on the right is observation, while matching connects models with observations. Application scenarios provide a background for requirements which depend on the task:

1. The ALV scenario involves cross country navi-



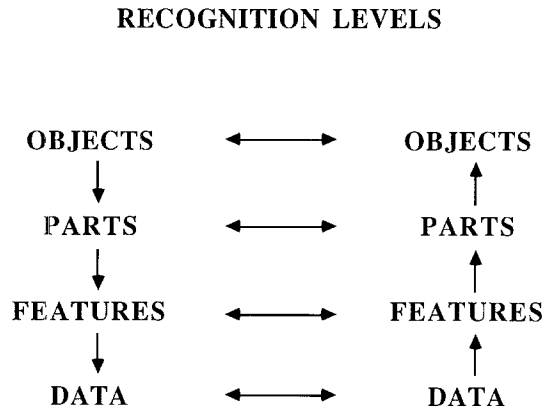


Figure 1. Model based vision.

gation outdoors. Objects vary greatly within object classes, terrain, vegetation, and structures. This is an extreme of the generic object class requirement.

2. Industrial robot vision tasks require handling one or a few objects at any time. However, large companies have many parts in production, for example, as many as 150,000. Programming many objects provides some motivation for generic object class capabilities.
3. Industrial inspection requires finding a wide variety of flaws, some of which occur infrequently, in the face of cosmetic marks which do not affect function. The number of different inspection tasks may be large. Inspection motivates generic object class treatment.
4. Space applications may not require generic models. The environment on board will be heavily documented. Planetary exploration does encounter terrain which makes generic methods valuable.
5. Home and service robots work in complex environments with complex objects. Generic models appear valuable.
6. The attraction of robots in warehousing is flexibility of routes and stacks. Automatic retrieval of objects from storage may be very complex, verifying correct placement in bins. Navigation is of intermediate difficulty, that is, there may be obstacles which come and go.

These scenarios motivate criteria for system design.

1. Number of object models: from one to as many as 1,000,000.
2. Variability of models: from classes of identical objects from a "cookie cutter," to object classes with variability such as humans, who vary

greatly one from another, and each of whom varies greatly over time, changing clothes, growing old, etc.

3. Complexity of background: simple or uniform on one hand to "any."
4. Computational complexity: one model per object, that is, order  $n$  complexity for  $n$  objects, at the brute force extreme, to log  $n$  complexity with indexing.

Much range image analysis and over vision is aimed toward recognition of few, identical objects. These application scenarios put value on recognition algorithms for many objects, object classes, and complex backgrounds, with low complexity.

#### 4.1 Matching

We summarize briefly a selected set of issues from group discussion. Among three components of recognition, modeling, observation, and matching, there was consensus that matching is a weak link, has highest priority, and is most neglected. Issues relate to performance in time and performance in quality of recognition or interpretation.

We discuss recognition within this interpretation paradigm: hypothesis generation, hypothesis management, hypothesis verification, and refinement. Performance in time motivates study of the following issues:

1. Structural indexing for hypothesis generation is a part of efficient recognition.
2. Control of the recognition process involves a group of issues: execution strategy, such as resource allocation; structuring recognition for parallel and distributed computation; data structures for matching; and algorithm synthesis based on detailed model based complexity.
3. Currently, the key to successful system building is using special case simplifications. A step toward this is study of general methods to use domain-specific information.

Quality of performance further motivates other important issues. One key is refinement of pose and parameters. A second is accurate implementation of geometric and probabilistic constraints in solution.

#### 4.2 Modeling

The power and generality of recognition depends on accurate modeling overall. First is the modeling of objects and scene. A new issue is generic modeling of object class. Another issue is building models and knowledge base. Second is modeling of observ-

ables, that is, generating models for sensors and operators.

### 5 Sensor Integration

Tom Henderson (Chair)  
 Ramesh Jain  
 Y.T. Chien  
 Jake Aggarwal  
 Ruzena Bajcsy

Multisensor integration is the combination of evidence or sensed data from several sensors. Most aspects of the sensor integration problem can be characterized in terms of Figure 2, where

- $s_i$  is the  $i$ th sensor,
- $X_i$  is the  $i$ th sensor data,
- $C_i$  is a measure of the uncertainty of  $X_i$ ,
- $f_i$  is the  $i$ th parameter produced,
- $t$  is the time parameter,
- $R$  is the composite model or representation,
- $K$  is the knowledge used in the problem.

Sensors deliver:

1. position and properties, and
2. absence of objects.

Integration techniques and the level of integration are functions of:

- the physical relation (or coupling) of the sensors, and
- the physical properties which characterize the object.

Levels of integration must be considered:

- signal/point
- neighborhood
- large region or surface level
- object level
- scene level

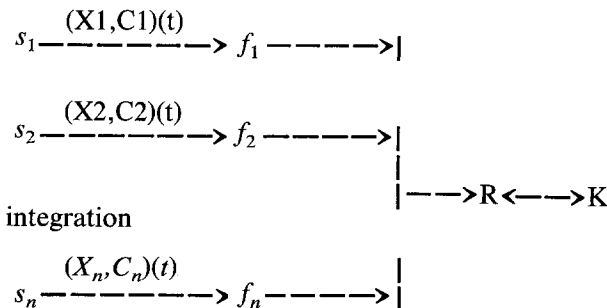


Figure 2. Sensor integration.

Major issues:

1. The representation of the composite model: This representation should be such that information acquired from disparate sensors can be integrated and assimilated. This representation should capture details and should represent symbolic information, also. One area of study is the use of explicit 3-D connectivity data structures as opposed to image data structures and the determination of appropriate applications for such techniques.
2. The combination of measurements is an estimation technique and well understood; however, inference techniques which operate on derived features need study. There are many approaches to uncertainty management. It is not clear which one performs better under given real world conditions. We require a good scientific study to determine the scope and limitations of uncertainty management techniques.
3. It is important to determine how the nature of the properties of problems influences the techniques selected or the parameters. This requires the study of requirements models and the related physical sensor models. It is essential to get at the cost of obtaining information with sensors and the development of an information based complexity model for sensing systems. This is particularly urgent for multisensor systems.
4. It is necessary to define the time-varying requirements of multisensor systems. This includes defining strategies for the sequence of application of sensors which complement each other, as well as the sequence of motions required for individual sensors.
5. In many cases, it is possible to obtain similar information using two different types of sensors. It is not clear that we understand the scope and limitations of sensors well enough to determine which sensor should be used in a given situation (this includes masking off parts of the sensed data).

### 6 Range Sensing for Navigation

C.M. Brown (Chair)  
 T. Kanade  
 W.E.L. Grimson  
 U.K. Sharma

#### 6.1 System Parameters, Navigational Tasks, and Representation

The concept of navigation is ill-specified and potentially extremely broad, encompassing tasks like

Table 2.

|   |                                 |                            |
|---|---------------------------------|----------------------------|
| Aligned and coupled                         | Identity (on element placement) | Pointwise integration      |
| Coupled, not aligned                        | Fixed rigid motion transform    | Signal ratio integration   |
| Not coupled, not aligned but known relation | Rigid motion transform          | Feature level integration  |
| Unknown transform                           | Unknown                         | Symbolic level integration |

finding space in a robotic environment, following an underwater course while maintaining a known pose and accounting for known and unknown currents, driving at speed over urban or open terrain with or without accurate or inaccurate maps. Thus, *the role of, the requirements for, and the necessary future research and development on range sensors for navigation all depend on the navigational task at issue.*

In order to keep this truism from being our last word on the subject, we decided to try to identify important and representative navigational tasks and system parameters that could affect choices, and to consider two case studies that illustrated various issues and answers. We also tried to address the topics of other groups and make explicit what capabilities in those areas we would like to see. In particular, we outlined what we should like to see in the areas of sensors, early processing, object recognition, and sensor fusion.

Autonomous navigation is not usually called for unless there are considerable practical problems involved, and so often the constraints that make it necessary also make it harder. The system or mission assumptions that seemed most crucial were the following.

1. Autonomy time—how long the navigator runs open loop. We considered times from hours (land vehicle reconnaissance mission) to weeks (undersea monitoring).
2. Weight and power restrictions. These can vary from loose (for a land vehicle) to stringent (for a Mars rover).
3. Task specifications.

The navigational tasks we considered in some detail were the following.

1. Finding navigable space.
2. Recognizing landmarks.
3. Localizing self in relation to an environment.
4. Servoing motion and detecting obstacles.

The sensor technologies we considered were the following.

1. Laser ranging.
2. Passive stereo or structured light triangulation.
3. Active camera tracking and focusing, including low level reflexes.
4. Sonar.

An important question is the representations to be used for navigation. Perhaps one of the more surprising outcomes of this group's discussions was the agreement that a small set of representations would probably suffice. There are interesting issues involving the algorithms that manipulate them, but the basic data structures we thought necessary are the following.

1. Occupancy maps or elevation maps—discrete, iconic representations of local topography, possibly *annotated* with symbolic information or labels.
2. Planar-faced and planar-polyhedral representations of geometry, perhaps labeled. More complex shapes can be approximated by these representations.
3. Materials should be represented explicitly when known or discovered. This information can be represented as labels on geometric elevation representations.
4. Topological connectivity (“(has been navigated” or “can be navigated” connectivity) between areas represented as occupancy or elevation maps.
5. Coordinate systems are basic.

## 6.2 Case 1: An Underwater Surveyor

A small underwater robot is to construct an elevation map of an area, annotating it with results of other sensors (where a chemical level exceeds a threshold, where a certain material is found, etc.). The robot has a range sensor that returns ranges  $R_{ij}$  for a local area (perhaps the output is a one-dimensional scan line). The robot has an orientation and orientation-derivative detector, and a velocity detector. This is a minimal scenario in several cases. The weight and power restrictions mean that “smart” sensing technology can be traded for processing, which is weight- and space-efficient. There

are no real-time constraints, so processing can be complex. The sensor technology itself can be relatively primitive. We noted that the underwater robot technology has grown quite sophisticated and that more cooperation between land and sea robotics workers would be a good idea.

The following paragraphs relate this case to the requirements for the subtasks mentioned above and for the subjects addressed by other working groups.

**6.2.1 Sensors.** The basic system is a sonar ranger. If close-up work, such as oil rig inspection, is called for, then there exist underwater laser systems that could be used for the requisite high resolution sensing. This subscenario calls for different mobility requirements and generally more sophistication. Ambiguity intervals would become an issue, for instance. There may be millimeter-wave radar for underwater use. Materials analysis from the sensor is a desideratum. High resolution and speed are not issues, since in this case we are not time limited.

**6.2.2 Representation.** The desired output is an annotated elevation map.

Spurious output detection is necessary. Registration of the results is a primary problem, but this is covered under self-localization below. Averaging or combining multiple readings from one location may be called for. Nothing in the way of segmentation, surface fitting, dealing with occlusion, etc. is needed.

**6.2.3 Sensor fusion, object (landmark) recognition, finding navigable space.** Not needed.

**6.2.4 Self-localization.** This issue is quite important, since unknown currents may affect the robot's location, and the output is a quantitative map of the area. There are engineering solutions (a buoy with a transponder that uses an outside reference source, like a navigation satellite). Failing that, the issue would come down to matching currently available sensing data against the map derived so far, which could be rather difficult.

**6.2.5 Obstacle detection.** Not needed in the simple scenario, and relatively easy to do using another echolocation sensor, except insofar as avoidance maneuvers could make self-localization more an issue.

**6.2.6 Map creation.** This is not difficult if the self-localization issue can be solved and if the sensors return the requisite data. The basic technique is temporal fusion of data, done in software.

### 6.3 Case 2: Surveying an Urban Environment

Here we assumed a large autonomous vehicle carrying several depth and imaging sensors, whose job it was to investigate a small deserted area of urban terrain and record differences between the current situation and the situation expected from an on-board map. The task could include checking if buildings still existed, if their identifications on the map were correct, if roads are still open, etc.

**6.3.1 Sensors.** The assumptions here invite more sophistication to be built into the sensors, in order to relieve the higher level processing. The following desiderata all raise considerable research and technological issues. This case may call for the following capabilities.

1. Stability for the sensor platform is necessary, to provide something like a vestibulo-ocular reflex (VOR) to keep the sensor fixated on a spot or to isolate it from vehicle motions.
2. A laser ranger should provide five times the speed, five times the distance, and five times the depth resolution of current sensors. It should have a vertical angular extent of 60 deg.
3. Effective techniques based on traditional image sensing might be possible. The technology to provide structured (or "unstructured") light for triangulation or stereo ranging over a wide area does not seem as daunting as laser-ranging technology. Multispectral scanners (or "pokers") for material identification are useful but a different issue. Recent work in active vision, such as "depth from parallax" or kinetic depth determination, (related to depth from flow), is aimed at duplicating the (very effective) processing performed by humans. Fixation and object tracking reflexes can simplify object recognition and also depth and egomotion calculations. Stereo is still a hot topic and could yield at any time. Our group felt that the potential power of techniques based on high-resolution optical images should not be overlooked.
4. Auto-registration of input through time would be an interesting option. This capability would compensate for vehicle motion as it happened. The goal would be to have the sensor emit true  $(x,y,z)$  information (point locations in a global fixed coordinate system) instead of  $R(i,j,t)$  information (range from the sensor in direction  $(i,j)$  at time  $t$ ). At present, the hardware does not exist to do automatic motion compensation. Further, it is important not to throw away covariance information (error ellipsoids are oriented toward the ranger, wherever it is). However, the hardware

situation may be changing, and taking account of the effect of errors is not a particularly difficult problem). The desired effect is the “stable world” we perceive when we move our eyes.

5. Many times it is desirable to work on  $R(i,j,t)$  data, rather than on  $(x,y,z)$  data, because the latter involves more parameters and therefore more computational time. Also, most of the time low level processing is done in the local coordinate system, and only important features are converted into the global coordinates. Usually, such features contain only 5% of the image points. Therefore, we save about 95% of the coordinate translation time by not converting (or not demanding the sensor to convert) the entire image into the global coordinate system.

If we expect the sensor to give  $(x,y,z)$  information in the global coordinate system, the sensor can perform “translation” based on inertial navigation system (INS) data to get the global value at each pixel. Therefore, the image acquisition time in local  $(x,y,z)$  will be more than that in global  $(x,y,z)$ , and much more than  $R(i,j,t)$ . Most of the time,  $R(i,j,t)$  suffices (e.g., obstacle detection, surface normal computation) but, sometimes  $(x,y,z)$  global or local is needed (e.g. for motion analysis, etc.).

In summary, there are three formats:

- (a)  $R(i,j,t)$  data
- (b)  $(x,y,z)$  local data
- (c)  $(x,y,z)$  global data

Depending upon the task and the penalty paid in time to the sensor for (2) and (3), one can determine which format is best.

**6.3.2 Representation.** Annotated, planar polyhedral representations for objects and annotated elevation maps for topography, linked topologically to indicate navigability (perhaps by simple “stay on the road” servoing), seem to be all that is required.

**6.3.3 Early processing.** Segmentation into regions of spatial or material homogeneity is the primary desideratum here. Simple planar regions with dihedrals labelled as to convex, concave, or obscuring type would be adequate. Reflexes, such as obstacle avoidance, can operate before segmentation. Multiresolution methods may well be useful.

**6.3.4 Sensor fusion.** The usefulness of combining depth and intensity data for more reliable segmentation has been demonstrated in several contexts. The primary reason for sensor fusion is to do more reliable segmentation. Insufficient information from a single sensor, such as only TV cam-

era or only range sensor, may lead to incomplete or incorrect object recognition. For example, range data obtained from a laser scanner, and reflectance data obtained from a TV camera provide complementary information. Range data provides important clues on the geometry of an observed scene. However, it does not provide any information about the physical properties of the scene objects such as color or reflectance. On the other hand, it is extremely difficult to extract geometrical information from TV data. Therefore, both types of data need to be analyzed. Doing this correctly involves understanding the physics of the problem, and thus how one sensor’s output is related to another’s. However, the existence of multiple sensors also raises the possibility of switching modes as called for by the task, so flexibility in sensor choice is a desideratum. Fusion can be done with help from the structural level, not just at the iconic level. That is, semantic labels or segments can be useful to constrain matches between sensor outputs. Last, temporal fusion is important in this scenario, as in almost all navigational tasks. Using a pre-existing map can help, but sophisticated techniques like Kalman filtering can be swamped by the registration problems that are inherent in real-world multiframe data. Thus, robust temporal combination techniques are another important research area.

**6.3.5 Object (landmark) recognition.** Landmark recognition (or self localization) is helpful so that the vehicle (or robot) can use triangulation to one or more landmarks to correct position errors caused by drift in the vehicle’s inertial navigation system. Recognition of specific objects from a small catalog is not considered to be a problem these days. Much more relevant in any real-world situation, including this scenario, is the ability to recognize a “generic object.” That is, to recognize to which class (e.g., “gas station,” “vehicle”) a particular instance belongs. Rule-based techniques may be applicable here, but construed broadly, this is almost *The Vision Problem*.

**6.3.6 Self-localization.** This is *The Matching Problem*, in a particular context, with particular representations. The knowledge or context provided by previous movements and known landmarks may make this problem slightly easier.

**6.3.7 Servoing, obstacle detection, and navigable space.** Servoing to guide motion with respect to road width constraints and to sense obstacles ahead (at slow speeds) can be provided by a skirt sonar sensor. Stationary obstacles ahead can show

up with minimal processing as “tombstones,” whose material properties may be important to know. This problem is currently under study in a number of places, and most autonomous robots have reasonable obstacle avoidance and servoing capabilities today. As motion speeds get faster, there are increasing demands on the sensor speeds and distances.

There seems to be no agreement today on how best to attack the navigable space problem. Whether space should be represented explicitly or implicitly (absence of known masses) is a basic question. Our assumptions here are consistent with either representation, but also make our problem easier, since the map would presumably yield a potential navigable space to be verified.

**6.3.8 Map creation.** The main goal in this case is no more difficult than any other computer vision undertaking, given the capabilities outlined above. That is, it is a big problem presenting many difficult intellectual and technical problems. Depending on the semantic content of the desired map, its component of geometric representation, and the disparity between the reality and the expectations, the task will be more or less reasonable given today’s knowledge. The contribution of range information to the job is to provide information about the three-dimensional world, largely in service of navigation itself. However, an unannotated, iconic elevation map could perhaps be produced by ranging alone. Further, the fusion of depth and image data can yield more reliable segmentation and thus improve navigation, recognition, and mapping.

## 7 Applications Group Report

Anil Jain, Michigan State University  
 Martin Levine, McGill University  
 C.W. “Ron” Swonger, ERIM (Chair)

The Applications Group of the Range Image Understanding Workshop focused upon examination of issues affecting the current and near-future routine “production” use of range image based machine vision for substantially unconstrained applications (i.e., applications where “image understanding” is a nontrivial undertaking). This focus inherently led to dismissal from serious consideration of passive stereo techniques and some other methods such as shape-from-shading methods. The group therefore intentionally considered uses beyond the research, development, test, and evaluation (RDT&E) phase of the classical high

technology system life cycle. Uses in industry, aerospace/defense and civil government were considered, but the intent of the group was to offer an integrative or “big picture” assessment of application issues across these communities.

First, (Table 3) a number of application areas were characterized as being “present” or “future” with respect to routine production use *with many image data*. Counterexamples were later solicited from all workshop attendees, but none were volunteered.

Next, (Table 4) several broad generalizations were offered concerning the application needs and benefits of range image data for image understanding problems. One observation in Table 4 was an assessment of the approximate performance level of range image *sensors*, stated in terms of five cost/performance parameters, required in order for a significant increase to occur in the industrial/commercial/government use of range image understanding systems. These performance/cost levels should again be taken as the rough “knees of the curves” of the number of sensors which might be sold nationally versus the indicated parameter values. No sensor is currently known to be available which is close to meeting all of the cost/performance values indicated.

The value of range image data in achieving integrated iconic (i.e., pixel-domain) and symbolic processing was also noted in Table 4. It should also be noted that active “laser radar” or LIDAR sensor designs are currently evolving which offer signifi-

**Table 3.** Production applications of range imagery

|   |
|---|
| Present   |
| Assembly control and verification                                 |
| Weld seam tracking  |
| Sheet metal part/assembly metrology                               |
| Electronics inspection—leads, SMDs, etc.                          |
| Future  |
| Contoured object inspection/verification—<br>e.g., turbine blades |
| Unconstrained material handling                                   |
| Wheel alignment   |
| General geometric flaw inspection                                 |
| Automated geometric modeling of prototype parts                   |
| Electronics inspection—solder, wirebonds, etc.                    |
| Food processing—quality control, sorting, grading                 |
| Vehicle guidance  |
| Automatic target recognition (ATR)                                |

**Note:** No attempt was made to characterize or compare these sensors other than the indicated categorization. No endorsement is implied and no completeness or accuracy is guaranteed for this listing. It is quite unclear whether the list provided will grow or shrink over the next few years given the technological issues, compelling needs, and economic difficulties within the machine vision supplier and user communities worldwide.

**Table 4.** Range image understanding application observations

---

The density of range data required depends upon:  
 The richness and lack of constraint of scenes to be observed.  
 The required variety and selectivity/accuracy of output decisions or actions.

Improvements most needed for practical application of range image sensors are:

|                                     |   |
|-------------------------------------|---|
| Cost                                | <\$25 K                                   |
| Frame rate                          | <1 second                                 |
| Tolerance of reflectance variations | >10 <sup>4</sup> :1                       |
| Size (scanning head)                | <¼ ft <sup>3</sup>                        |
| Range noise                         | Standard deviation <0.1% of work envelope |

Range images offer significantly greater opportunity for achieving the long sought and critical integration of iconic and symbolic processing for 3-D scene interpretation.

---

cant potential for multimodal sensing and associated information fusion at the image (pixel) level. Specifically, image sensors can be built to provide fully registered images of scenes which directly encode range, panchromatic reflectance, multispectral reflectance (material properties), surface roughness, radial velocity, vibration, polarization, temperature, and/or other properties for each pixel in the scene.

Thirdly (see Table 5), the group undertook to prioritize the few most significant obstacles currently limiting the practical “production use” of range image understanding systems. Cost, algorithmic concepts, data availability, and processing hardware issues each enter into that list.

Finally, a brief list was compiled (Table 6), from the knowledge of the group members and immediately available references, of organizations known to currently build range image sensors (including point, contour or profile, and array or whole image sensors).

**Table 5.** Prioritized obstacles to application of range imagery

---

1. Sensor cost for required resolutions and data rates.
2. Difficulty of interpreting substantially unconstrained scenes:
  - (a) Segmentation.
  - (b) Object recognition.
3. Inadequate availability to researchers of multimodal registered image data.
4. Inadequate availability of special purpose hardware for real-time geometric feature computation (e.g. Gaussian curvature).

---

**Table 6.** Known builders of range image sensors (current, industry applicable)

---

| Organization                        | Point | Profile | Array          |
|-------------------------------------|-------|---------|----------------|
| Boulder Electro-Optics <sup>a</sup> |       |         | ●              |
| Chesapeake Laser                    |       |         | ●              |
| Cyberoptics                         | ●     | ●       |                |
| Diffrauto                           | ●     | ●       |                |
| Digital Signal                      |       |         | ●              |
| ERIM                                |       | ●       | ●              |
| G.E.                                |       |         | ●              |
| Hymark (Canada)                     |       | ●       | ● <sup>b</sup> |
| Industrial Technology Institute     |       |         | ●              |
| Keyence (Japan)                     | ●     |         |                |
| NRC                                 |       |         | ●              |
| Odetics                             |       |         | ●              |
| Oldelft (Netherlands)               |       | ●       |                |
| Perceptron                          | ●     | ●       | ●              |
| Photonic Automation                 |       |         | ●              |
| RVSI                                | ●     | ●       |                |
| Servo-Robot (Canada)                |       | ●       | ● <sup>b</sup> |
| Selcom                              | ●     |         |                |
| Siemens (W. Germany)                |       |         | ●              |
| Synthetic Vision Systems            |       | ●       |                |
| Technical Arts                      |       | ●       | ● <sup>b</sup> |

---

<sup>a</sup> Acquired by Melles Griot. Machine no longer available.

<sup>b</sup> Nodding mirror option.

**Caveats:**

Undoubtedly incomplete.  
 Accuracy not guaranteed.

## 8 Appendix

### 8.1 Overview Speakers

- Avi Kak (Purdue Univ.): Sensing
- Paul Besl (General Motors Research): Early Processing
- Robert Bolles (Stanford Research Institute): Object Recognition
- Jake Aggarwal (Univ. Texas): Sensor Integration
- Takeo Kanade (Carnegie Mellon Univ.): Navigation
- Joe Mundy (General Electric): Application
- Ruzena Bajcsy (Univ. Pennsylvania): 3-D Vision

### 8.2 List of Participants

This list of participants gives affiliation of people and groups (identified by letter codes: SI—sensor integration, EP—early processing, OR—object recognition, RS—range sensing, N—navigation, APP—applications) in which they participated.

1. Jake Aggarwal (Univ. Texas) (SI)
2. Ruzena Bajcsy (Univ. Penn) (SI)
3. Paul Besl (General Motors Research) (EP)
4. Tom Binford (Stanford Research Institute) (OR)

5. Robert Bolles (Stanford Research Institute) (OR)
6. Chris Brown (Rochester) (N)
7. Y.T. Chien (National Science Foundation) (SI)
8. Nelson Corby (General Electric) (RS)
9. Patrick Flynn (Michigan State Univ.) (EP)
10. Eric Grimson (Massachusetts Inst. Tech.) (N)
11. Kevin Harding (Industrial Technology Inst.) (RS)
12. Tom Henderson (Utah) (SI)
13. Richard Hoffman (Northrop) (EP)
14. Anil Jain (Michigan State Univ.) (APP)
15. Ramesh Jain (Univ. Michigan) (SI)
16. Avi Kak (Purdue) (RS)
17. Takeo Kanade (Carnegie-Mellon Univ.) (N)
18. Martin Levine (McGill Univ.) (APP)
19. Gerardo Medioni (Univ. S. California) (EP)
20. Joe Mundy (General Electric) (APP)
21. Francis Quek (Univ. Michigan) (OR)
22. Gerhard Roth (Nat. Research Council—Canada) (EP)
23. Robert Simpson (Env. Research Inst. of Michigan) (RS)
24. Iswar Sethi (Wayne State) (EP)
25. Uma Kant Sharma (FMC Corp.) (N)
26. George Stockman (Michigan State Univ.) (RS)
27. Ron Swonger (Env. Research Inst. of Michigan) (APP)
28. Richard Weiss (Univ. Mass) (OR)

### 8.3 Workshop Groups and Group Chairs

#### Range Sensing

George Stockman (Chair), MSU  
 Nelson Corby, GE  
 Kevin Harding, ITI

Avi Kak, Purdue  
 Robert Sampson, ERIM

#### Early Processing

Gerard Medioni, USC (Chair)  
 Paul Besl, General Motors (Reporter)  
 Patrick Flynn, MSU  
 Richard Hoffman, Northrop  
 Gerhard Roth, NRC-Canada and McGill  
 Ishwar Sethi, Wayne State

#### Object Recognition

Thomas O. Binford (Chair)  
 Robert Bolles  
 Francis Quek  
 Richard Weiss

#### Sensor Integration

Tom Henderson (Chair)  
 Ramesh Jain  
 Y.T. Chien  
 Jake Aggarwal  
 Ruzena Bajcsy

#### Navigation

C.M. Brown (Chair)  
 T. Kanade  
 W.E.L. Grimson  
 U.K. Sharma

#### Applications

C.W. "Ron" Swonger, ERIM (Chair)  
 Anil Jain, Michigan State University  
 Martin Levine, McGill University