AIAA-87-1674
Computational Themes in Applications of Visual Perception

R. Jain, B. G. Schunck and T. Weymouth,
University of Michigan,
Ann Arbor, MI

Second AIAA/NASA/USAF Symposium on Automation, Robotics and Advanced Computing for the National Space Program

March 9-11, 1987/Arlington, VA

Computational Themes in Applications of Visual Perception

Ramesh Jain
Brian G. Schunck
Terry Weymouth

Computer Vision Research Laboratory
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109-2122

abstract

The paper summarizes the current research in the Computer Vision Research Laboratory at the University of Michigan. The laboratory concentrates on developing generic vision algorithms for industrial applications. Generic vision algorithms can be applied to a wide variety of inspection problems. The paper includes a discussion of the current state of the machine vision industry and provides recommendations for improving the transfer of vision technology from research to practice.

1 Introduction

The University of Michigan recently formed a laboratory, called the Computer Vision Research Laboratory, within the Department of Electrical Engineering and Computer Science for research in computer vision. The Computer Vision Research Laboratory will evolve through interaction among the researchers of various related disciplines and through inductrial and government contacts.

The research program in the Computer Vision Laboratory concentrates on the development of generic vision algorithms that can be applied to a variety of situations. Generic perception algorithms are simple, reliable modules that hide the details of perception in a packaged hardware or software component that can be used by applications specialists who are not themselves expert vision researchers. Generic vision algorithms are not developed for specific applications, but are designed to solve a vision task that is part of many different inspection problems that occur in different applications in different industries. In developing generic vision algorithms, researchers leverage their efforts by providing solutions to classes of vision tasks while removing the details of the application from the research efforts. This insulates the vision researchers from the details of the application and allows them to concentrate on understanding the vision problems.

Another aspect of research on generic vision algorithms is that the algorithms are classified according to properties that are meaningful in the context of the emerging knowledge of vision fundamentals, as opposed to classification according to applications areas. There are no automotive vision algorithms or aerospace vision algorithms; vision algorithms are not specific to any one industry. Vision algorithms are defined by the properties of the vision processing task; that is, in terms that depend on the vision processing itself, rather than on some intended use for the algorithm.

The nature of the problems of applying vision to specific applications has not been widely perceived. The current macroeconomic system of machine vision is inappropriate. The solution is to place machine vision technology in the hands of the end users who are familiar with the requirements of their environment. Vision technology should be like any other instrumentation: plug it in and use it to make measurements. It is not practical for the developers of machine vision applications to perform the development specific to the application and to develop the "instrumentation" required for the development at the same time. To change the current, untenable system, vision researchers must produce generic vision algorithms that can be used by vision systems developers in a variety of applications.

Researchers are currently addressing problems related to vision engineering but without much, if any, concern for designing machine vision systems. The focus of the Computer Vision Research Laboratory on vision algorithms that are generic, rather than applications specific, meets needs not addressed by other research institutions. The Computer Vision Research Laboratory provides a focus for the synergistic interaction of researchers aided by interactions with industry. The goal of research in the laboratory is to address fundamental problems blocking advances in the design of machine vision systems for various applications. The research addresses basic issues in computational vision research and the design of systems that can perform vision tasks in real time for a given application. New engineering techniques for designing systems are being developed that can perform tasks in an unstructured environment. The research may also advance the understanding of biological vision systems.

2 Research Problems

Visual perception is the process of interpreting measurements such as light intensity and range that relate to the projection of surfaces in a scene onto the image plane. The central problem in vision is the reconstruction of surface structure and properties from the projection onto the image plane. This paper outlines some of the research problems currently being explored in the Computer Vision Research Laboratory.

- Dynamic vision
- Range images
- Knowledge-based perception systems
- Computer architectures for vision algorithms
- Sensor integration
- Bridging the gap between research results and applications

3 Dynamic Vision

Motion provides valuable clues to surface structure. Algorithms for interpreting sequences of images can provide measurements that are very useful for applications. Image flow research has many applications including passive sensing systems for autonomous mobility, machine inspection of surface structure, passive sensors for aircraft and satellite docking systems, and image compression.

3.1 Background

The last few years have seen increasing interest in dynamic scene analysis. The input to a dynamic scene analysis system is a sequence of images. A major problem in a computer vision system is to recover the information about objects in a scene from images. This problem cannot be solved without some assumptions about the world. A sequence of frames provides the additional dimension of time for recovering the information about a 3-dimensional world that is lost in the projection process. Multiple views of a moving object acquired using a stationary camera may allow recovery of the structure of the object [45]. A mobile camera may be used to acquire information about the structure of the stationary objects in a scene using optical flow [11], axial motion stereo [25,32], and other methods [21].

Many researchers in the psychology of vision support the recovery of information from image sequences, rather than from a single image [11,31]. Gibson [11] argued in support of active information pick-up by the observer in an environment. Neisser [31] proposed a model in which the perceptual processes continually interact with the incoming information to verify expectations formed on the basis of available information. In computer vision systems, the power of exploiting motion, even with such noise sensitive approaches as difference pictures and accumulated difference pictures, has been demonstrated on complex, real world scenes [20]. Many researchers are addressing the problem of recovering information in dynamic scenes; but due to the legacy of static scenes, most researchers are approaching the recovery problem using just two or three frames of a sequence. This restricts the results to quasi-dynamic scene analysis, rather than true dynamic scene analysis. The information recovery process requires constraints about the scene. The analysis based on small numbers of frames rests on assumptions that ignore the most important information in dynamic scenes.

The viewpoint of the Computer Vision Research Laboratory is that image understanding is a dynamic process. Dynamic vision algorithms cope with the error-filled visual world by exploiting redundant information in the image sequence. Current research efforts are developing a qualitative approach to vision that uses only relative information available in a sequence to infer relationships between objects in a scene.

3.2 Segmentation

In many dynamic scene analysis systems, the goal is to recognize moving objects and to find their motion characteristics. If the scene is acquired using a stationary camera, then segmentation generally refers to the separation of moving components from stationary components in the scene and identification of individual moving objects based on velocity or some other characteristic. For the case of a moving camera, the segmentation task may be the same as above or may involve further segmentation of the stationary components of the scene by exploiting the motion of the camera. Most research efforts for the segmentation of dynamic scenes have been concerned with the extraction of the images of the moving objects observed by a stationary camera. It has long been argued by researchers in perception [11,45] that motion cues aid the segmentation process. Computer vision techniques for segmenting dynamic scenes perform well in comparison to those for the segmentation of stationary scenes. The segmentation of moving camera scenes into their stationary and nonstationary components has received attention only recently [22]. The major problem in the segmentation of moving observer scenes is that every surface in the scene shows motion in the image. For separation of moving object images, the motion component assigned to various surfaces in the images due to the motion of the camera must be removed. The fact that the image motion of a surface depends on its distance from the camera and the surface structure complicates the situation.

Jain developed techniques for the segmentation of dynamic scenes [19,20,22]. These techniques have been used in many applications. Sandia Laboratories is developing systems for tracking objects for Army applications using techniques based on differencing using a likelihood ratio. Recently research led to a new approach for segmentation using accumulative difference pictures that may be implemented easily on special hardware [23].

3.3 Image Flow

Image flow is the velocity field in the image plane that arises due to the projection of moving patterns in the scene onto the image plane. The motion of patterns in the image plane may be due to the motion of the observer, the motion of objects in the scene, or both. The motion may also be apparent motion where a change in the image between frames gives the illusion of motion [36].

The intent of this research is to discover new models for image flow that will yield new algorithms for image flow estimation and analysis. Constraint equations that better model the interactions between changes in object position and the illumination and surface reflectance characteristics will naturally result in better algorithms and better understanding of existing algorithms.

Prior image flow research developed an image flow equation for smooth patterns of image irradiance and smooth velocity fields [27]. The equation was extended first to image irradiance patterns with discontinuities and then to velocity fields with discontinuities [55,56,60]. Future research in image flow constraint equations will aim to increase our understanding of image flow characteristics. This will lead to new algorithms for image flow estimation that incorporate the new continuity equations. Restrictions on the situations in which existing continuity equations can be used will be discovered and these insights will improve the performance of existing image flow estimation algorithms by pin pointing the situations where the algorithms cannot be used.

3.4 Motion Stereo

Depth determination is a continuing problem in computer vision. Research in this area is motivated by the need for target tracking, autonomous vehicles control, visual prostheses for the blind, realistic flight trainers, and models of the human visual system. There is a plethora of depth determination techniques. Many different stereo systems for depth determination have been developed just in the last few years.

Stereo information can also be obtained using a single moving camera. Jain and O'Brien [25,32] map images into a complex log space where the movement of the objects in two dimensions due to the camera motion becomes a translation along one axis in the new space. Given this phenomenon, the motion correspondence problem is greatly reduced, since only a small strip of the new space needs to be searched [18]. This constraint is similar to the epipolar constraint in stereo. In addition, the transform is scale and rotation invariant. It also has an analog in the human visual system: the mapping of the retinal space into the striate cortex is very closely approximated by the CLM. Figure 1 shows an image and its mapping.

Figure 1: An image and its complex log mapping are shown in this figure. The mapping is similar to the retino-striate mapping in primate visual systems. This mapping is very useful in segmentation and depth recovery in dynamic vision.

Optical flow has been studied with the aim of recovering information about the environment and the motion of the observer. The *egomotion complex logarithmic mapping* (ECLM) exploits some characteristics of optical flow without computing it. The mapping combines scale, projection, and rotation invariances of complex log mapping with the characteristics of optical flow. This mapping is useful in segmentation of image sequences to recover images of moving objects. The distance to stationary objects can also be computed from the egomotion complex logarithmic mapping. Using this mapping, motion stereo and segmentation can be achieved in one step. The feasibility of the approach has been demonstated. The next step is to study its application in dynamic scenes. This research should yield important results for the navigation of mobile robots.

3.5 Motion Trajectories

Iterative algorithms for determining trajectories of points in an extended frame sequence using *path coherence* are being developed. The emphasis in this approach is to exploit motion characteristics for establishing correspondence without assuming rigidity of objects. A greedy exchange algorithm has been developed to implement this idea [41]. The results of applying this algorithm to a sequence from the *Superman* movie are shown in Fig. 2.

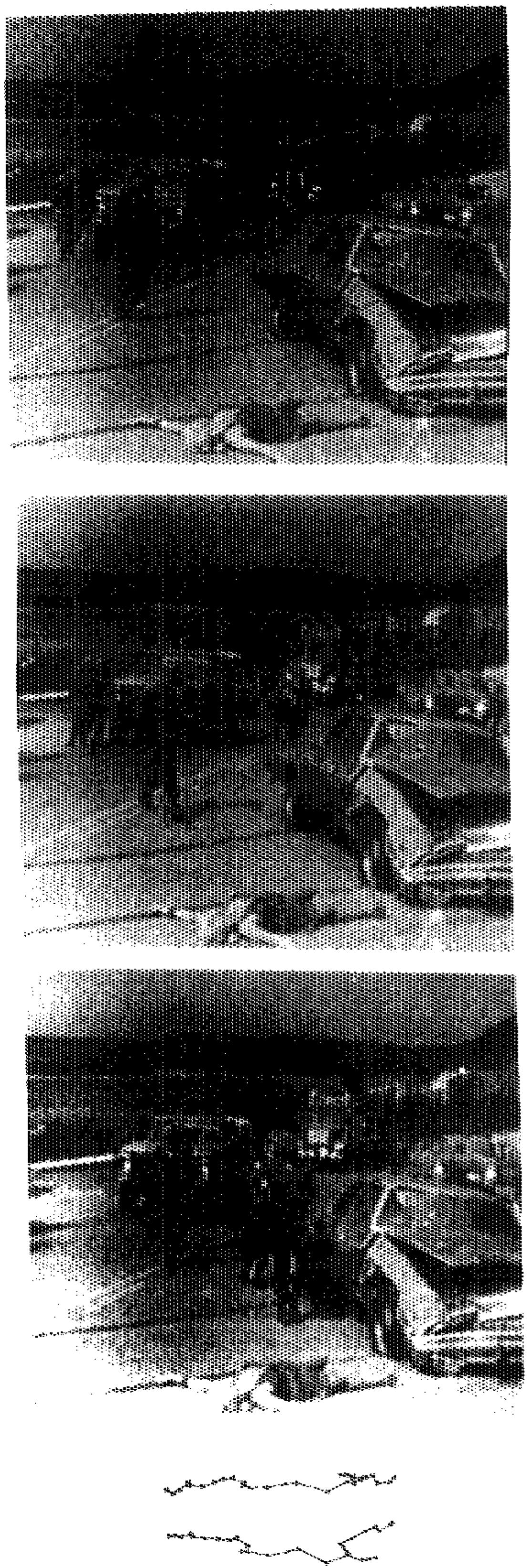

Figure 2: Three frames from the *Superman* sequence are shown in this figure. Points on the head and belt of the three soldiers running towards the camera were tracked using a greedy exchange algorithm. The trajectories are also shown here.

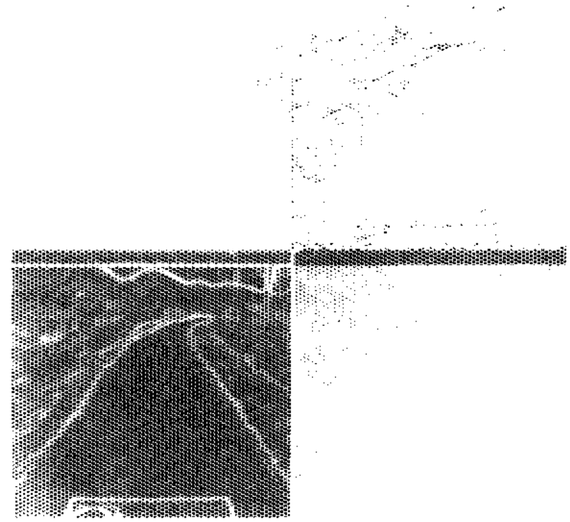

Figure 3: This figure shows results of our road edge detection algorithm. For the road shown in the top left corner, the road edges are shown in the right bottom corner.

Another approach to finding trajectories that will determine motion events is being studied. In this approach, motion parameters are recovered using constraints imposed by the equations for the motion of points [15]. The algorithm uses successive refinement to determine the trajectories of points. A very attractive feature of this approach is that it determines discontinuities and tries to use smoothness only for the known smooth path in establishing correspondence.

The determination and role of events in dynamic scenes is being studied. A motion event may occur due to a change in the motion parameters or due to occlusion. A major thrust of other research is to develop techniques that will recover qualitative information about depth and motion of objects using constraint propagation.

3.6 Navigation

Techniques for the guidance of autonomous mobile robots or vehicles are being developed. This requires techniques to recover information about the environment using vision and then to use knowledge based techniques to control the navigation. The observation that the location of road vanishing points does not change significantly from frame to frame is used to develop an algorithm for finding road boundaries [29]. The algorithm uses the hypothesize and test paradigm. Results of this approach are shown in Fig. 3.

4 Range Images

The long term goal of this project is to develop techniques that will be useful in object recognition and navigation using range information.

The last few years have seen increasing attention to the analysis of range images. Range images may be obtained with passive methods, such as binocular stereo, or with range sensors. Range images contain explicit information about surfaces. This explicit information facilitates recognition and location tasks in many applications. The goal in range image understanding is to find robust symbolic surface descriptions that are independent of viewpoint. Techniques to characterize surfaces in range images are being developed [5,4,24].

Surfaces are segmented using local features, such as Gaussian and mean curvatures and related differential geometric measures. The signs of the curvatures at each point in the image are used to assign one of eight basic surface types to each point in the image. The next step is to develop techniques to identify surfaces by grouping points using the surface type, spatial proximity, and other criteria. All grouping processes must conform to the sensed information since the ultimate truth is in the sensed data. This stimulus bound approach uses symbolic surface descriptions in the segmentation of images. Figure 4 shows a result of this segmentation approach. The symbolic surface descriptions will play important role in many applications.

4.1 Recognition Methods

Object recognition is a major motivation in most image-understanding systems. Despite strenuous efforts, only limited success has been achieved.

The object recognition task can be classified based on difficulty in several ways. One way is based on the degree of uncertainty allowed in the object's position and orientation. This can range from no uncertainty, to uncertainty only in its 2-D position, through uncertainty in both its 2-D position and rotation, up to uncertainty in its 3-D position and orientation.

The task can be further classified based on the complexity of interactions allowed between objects. In the simplest case, each object to be recognized must be completely visible and surrounded by background. In a more complex case, objects are allowed to touch but not overlap. In the most general case, objects are allowed to partially occlude one another.

Most work on the occluded-objects problem has concentrated on the case where there is only one object type in the scene. Algorithms developed for this problem can be extended to cases of multiple objects types, simply by running the algorithm multiple times, once for each possible object type. This is not ideal, as it does not consider or take advantage of the similarities and differences among possible objects. Also the recognition time increases linearly with the number of possible objects, after a fixed time for preprocessing. This can become a problem as the number of possible object types increases.

A new method for object recognition called the *feature indexed hypotheses method* is being developed. This method breaks the recognition process into two phases: hypotheses generation and hypothesis verification. By using features that occur multiple times in the possible object set, the number of features in the search can be greatly reduced generating a small number of false hypotheses that are easily rejected. By carefully selecting the features, the recognition time growth rate can be reduced to the square root of the number of possible objects. This method also has the advantage that unique features which are difficult or impossible to find if the possible object set contains many similar objects are not required. The method's feasibility is demonstrated by the results of a prototype two-dimensional occluded-parts recognition system.

5 Knowledge-Based Perception Systems

Most complex tasks require specialized knowledge, image understanding is no exception. The last few years have seen an increasing application of knowledge in computer vision systems. The application of knowledge at different levels in a vision system is being studied.

5.1 Knowledge-Based Algorithms

An image interpretation system is a program that derives a scene description from a time-varying image. The input signal is a rectangular grid of signal samples taken at regular time intervals $I(x, y, t)$. The sample can be taken in several frequency ranges (for example, red, green, and blue) and combined into a time-varying vector field $C(x, y, t)$. There is wide agreement that this characterizes the environmental data used by such a system [9,17]. On the other hand, it is much more difficult to formulate a description of the output of an interpretation system. In this project, the output of an interpretation system is a network representing the scene being viewed, an approach common to other interpretation systems. The nodes of such a network are groups of objects within the scene: individual objects, object parts, or references to clusters of image events. Arcs in the network are labeled with relations between the objects. Since primitive object parts are usually depicted as geometric solids or collections of joined surface patches [10], these representations are included in the representation used in this work.

However, there is no need to limit ourselves to only that type of representation. In fact, the study of which types of primitives are useful and how they relate to the processing of the image is part of our current research. The issues of which relations to use, how to vary those relations with time, and how to incorporate the variations in object descriptions that occur over time are also being studied.

The image sequence is not the only source of data used by an interpretation system. A large database of relational and descriptive information, including information about processes and procedures, is also necessary for image interpretation [13,30]. These data are a symbolic representation of knowledge about the world, especially as that knowledge pertains to the interpretation problem; thus, the database is usually referred to as a knowledge base. For example, if it is known that the camera is upright, level, and outdoors, then it can reasonably be expected that the bright area at the top of the image is the sky. Such simple rules and other more complex inferences form the greater part of knowledge for interpretation. The problem of selecting and representing the appropriate knowledge is a difficult one. Understanding of image interpretation must be achieved by building systems of processes that work in restricted domains. In building such systems, the rules and processes which can eventually be incorporated into a knowledge base are learned. The problem then is to organize the information about the object and the procedures for using that information so that they can work in situations where the answers are less than certain.

The issues are basically these:

- Organization must be imposed on the information used for image interpretation, because that information is richly detailed and complex, and because large amounts of information are required.
- Varied types of information (for example, relations, procedures, and structural descriptions) need to be made easily accessible at many levels of abstraction.
- There must be a way to control the selection of which information is considered, which programs are activated, and what information is passed among programs.
- Interpretation processes need to deal with errors from image data and introduced by processes producing partial interpretation.
- All components of the interpretation system should be able to deal with the changing nature of dynamic data and the reorganization of interpretations that occur when a process creates new hypotheses.

6 Vision Computer Architecture

The real-time application of image understanding algorithms requires computer hardware that allows the algorithms to be executed at high speed. The NCUBE machine is being used to study hypercube architectures for the implementation of vision algorithms.

This research is concerned with developing techniques to perform computer vision computations on loosely and tightly coupled parallel processors. The size of data sets and the speed at which they can be created in current problems swamp even the fastest computers. The processing power required to keep up with the input is greater than 100 MIPS. This level of processing power is possible only with some degree of parallelism. However, a rule of thumb is that parallel processors can rarely sustain an efficiency level of greater than 20% [38]; thus the machine needs to have a peak power of about 500 MIPS. Normally the number of operations far exceeds this and the size of the input data set is often much greater than 512 times 512 bytes, either because multiple views are required, or because the sensor has a higher resolution (4096 times 4096, for example), or more than one modality is in use.

In the past, vision researchers have been forced to use SIMD meshes (such as the Goodyear MPP or pyramids) to achieve this rate of computation. Hypercubes offer a better alternative than SIMD meshes for several reasons.

- They can efficiently simulate these SIMD machines
- Their interconnections are better for operations such as image rotation, that are inefficient on meshes or pyramids
- The hypercube is better suited to higher level symbolic tasks since it is an MIMD machine, the nodes have far more memory, the interconnections are much better for the more random information exchange, and the node processors have a better instruction set

As an example of efficiency in vision tasks, a hypercube with p processors can rotate an n times n image 90° in $\mathcal{O}(n^2 \log p/p)$ time, versus $\mathcal{O}(n^2/\sqrt{p})$ for a mesh or pyramid.

Hypercube architectures for MIMD organizations are better than architectures with SIMD organizations for operations at the intermediate levels of vision processing where the sequence of operations for different regions in an image are usually different because the sequence is governed by properties of the region. For this type of processing, a loosely coupled multiprocessor system is ideal. If dynamic scenes are involved, then massive parallelism is essential [1].

7 Sensor Integration

Representations and techniques for combining vision, range, tactile, and other sensory information are being studied. Representations that will allow easy inference using multi-sensors are not known. Representations that will facilitate combining multi-sensor information for planning and recognition tasks using partial information obtained from each source are key to the success of sensor fusion. Techniques to combine uncertain and imprecise information are being developed using nonmonotonic and probabilistic approaches. Methods that try to combine beliefs and disbeliefs using uncertainty calculus for solving problems in distributed problem solving systems are being studied.

Computer vision and knowledge-based systems have to deal with uncertain and imprecise data in almost all phases of reasoning. Many approaches, such as Bayesian methods, fuzzy logic, Dempster-Schaefer theory, and qualitative approaches, have been proposed for dealing with uncertainty. All of these approaches have their problems and advantages. Unfortunately, it is not clear which method is best for which types of applications. With the aim to understand the strenghts and weaknesses of each approach, a language called ULOG is being developed that will implement uncertain variables in a PROLOG type environment. The ULOG language allows a user to change the algorithm used

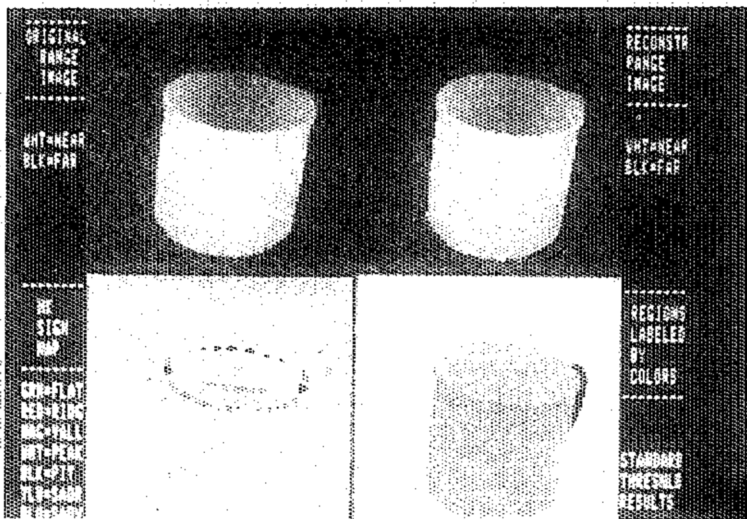

Figure 4: A range image and its segmentation using Besl's and Jain's algorithm are shown in the top-left and bottom-right corners respectively. Using the polynomial descriptors of the segmented surfaces, the original image is reconstructed. This is shown at top-right.

Figure 5: This figrre shows an image showing the classification of solder joints by our algorithm.

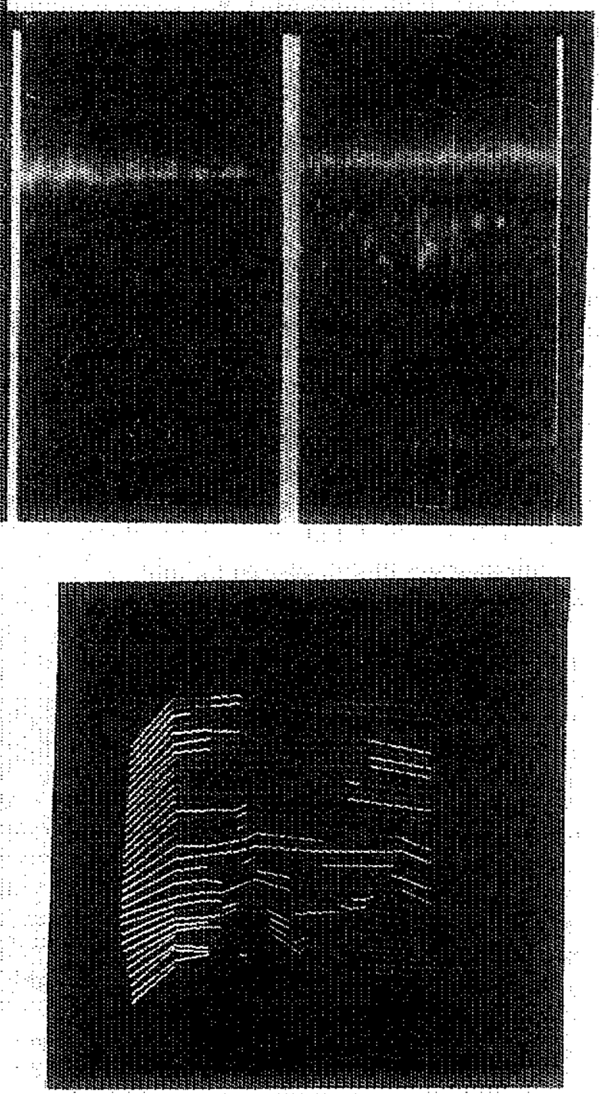

Figure 6: This figure shows an image showing two SEM images of a section of a wafer and the 3-D surface structure as reconstructed using our sem stereo algorithm.

for uncertainty management to experiment with different approaches in order to find the best approach for a given problem.

8 Applications

Several research projects are closely applied with specific applications. Normally, the research projects are focussed on the development of generic algorithms that can be applied to a wide variety of applications. All of the research work on generic algorithm can lead to applications. However, occasionally some application is investigated in the Computer Vision Research Laboratory because the application provides a focus for learning about some class of vision algorithms. Research projects that are closely associated with applications are discussed in the following sections.

8.1 Solder Joint Inspection

This project is for improving the reliability of solder joints and to reduce the cost of manufacturing by reducing the fault rate. Techniques are being developed for inspection of solder joints to identify not only good/bad joints but also to identify the nature of defects. By identifying the nature of defects and using trend analysis, the system may be able to indentify the process parameters that may result in defective joints [7]. An algorithm based on surface characteristics classifies the joints. The performance of the current version of the algorithm is 97-98% correct results. A knowledge-based approach to the task is being implemented to improve the performance of the classifier still further and to identify process parameters [3]. Some results from this project are shown in Figure 5. Recent extensions to this research project include inspection algorithms for circuit boards with surface mounted components.

8.2 Semiconductor Wafer Inspection

This project is a part of the center of excellence in semiconductor manufacturing. The project is concerned with the in-process inspection of semiconductor wafers to identify defects. The objective is to identify defects to control subsequent processes; if possible, the system will take corrective action and adjust the defective process parameters at the stage that is be responsible for the defect. The corrective actions are automatically performed by an expert system that controls the processes. The inspection is done with scanning electron microscopy (SEM) and optical microscopy. Current activities include the study of SEM images to implement stereo approaches for those images to perform analysis of those images to inspect the nature of edges and other features on a wafer. A model-based stereo algorithm has been implemented to reconstruct the 3-D structure of a wafer section. Figure 6 shows 3-d surface structure recovered by our algorithm. Knowledge-based controls are being developed for the interface between the vision system and the expert system in the automated semiconductor manufacturing project.

References

[1] Agrawal, D. P. and R. Jain, "A pipelined pseudoparallel architecture for dynamic scene analysis," *IEEE Trans. Computers* 31 (1982), 952–962.

[2] Barrow, H. G. and J. M. Tenenbaum, "Recovering intrinsic scene characteristics from images," *Computer Vision Systems*. ed. A. R. Hanson and E. M. Riseman. New York: Academic Press, 1978.

[3] Bartlett, S. L., C. L. Cole, and R. Jain, "Expert system for visual solder joint inspection," in *Proc. IEEE Conf. Art. Int. Applications*, March 1987.

[4] Besl, P. G. and R. Jain, "Segmentation through Symbolic Surface Descriptions," Accepted for publication in *IEEE Trans. Pattern Analysis and Machine Intelligence*.

[5] Besl, P.G. and R. Jain, "Invariant Surface Characteristics for 3-D object recognition in depth maps," *Comp. Vision, Graph. and Image Proc.* 33 (1986), 33–80.

[6] Besl, P.G. and R. Jain, "Three-dimensional object recognition," *Computing Surveys* 17 (1985), 75–145.

[7] Besl, P. G., E. Delp, and R. Jain, "Automatic visual solder joint inspection," *IEEE J. Robotics and Automation* 1 (1985), 42–56.

[8] Binford, T., "Survey of Model Based Image Analysis Systems," *International Journal of Robotics Research* 1 (1982), 18–64.

[9] Brady, M., "Computational Approaches to Image Understanding," *Computing Surveys* 14 (1982), 3–71.

[10] Brooks, R., "Symbolic Reasoning Among 3-D Models and 2-D Images," Technical Report STAN-CS-81-861, Department of Computer Science, Stanford University, Stanford, California, June 1981.

[11] Gibson, J. J., *The ecological approach to visual perception*. Boston: Houghton Mifflen, 1979.

[12] Grosky, W. I., and R. Jain, "A pyramid-based approach to segmentation applied to region matching," Accepted for publication in *IEEE Trans. Pattern Analysis and Machine Intelligence*.

[13] Hanson, A., et al. "A Methodology for the Development of General Knowledge-Based Vision Systems," Technical Report, COINS Department, University of Massachusetts, September 1985.

[14] Haynes, S. M. and R. Jain, "Detection of Moving Edges," *Comp. Vision, Graph. and Image Proc.* 21 (1983), 345–367.

[15] Haynes, S. M. and R. Jain, "Event detection and correspondence," *Optical Engineering* 25 (1986), 387–393.

[16] Horn, B. K. P. and B. G. Schunck, "Determining optical flow," *Artificial Intelligence* 17 (1981), 185–203.

[17] Horn, B. K. P., *Robot Vision.* Cambridge: M. I. T. Press, 1986.

[18] Jain, R., Bartlett, S., and N. O'Brien, "Motion Stereo Using Egomotion Complex Logarithmic Mapping," Accepted for publication in *IEEE Trans. Pattern Analysis and Machine Intelligence.*

[19] Jain, R., D. Militzer, and H.-H. Nagel, (1977) "Separating non-stationary from stationary scene components in a sequence of real world TV-images," in *Proc. Int. Joint Conf. Artificial Intelligence*, pp. 612–618.

[20] Jain, R. and H.-H. Nagel, (1979) "On the analysis of accumulative difference pictures from image sequences of real world scenes," *IEEE Trans. Patt. Anal. and Mach. Intel.* 1 (1979), 206–214.

[21] Jain, R., "Direct computation of the focus of expansion," *IEEE Trans. Patt. Anal. and Mach. Intel.* 5 (1983), 58–64.

[22] Jain, R. "Segmentation of frame sequences obtained by a moving observer," *IEEE Trans. Patt. Anal. and Mach. Intel.* 6 (1984), 624–629.

[23] Jain, R. "Difference and Accumulative Difference Pictures in Dynamic Scene Analysis," *Image and Vision Computing* 2 (1984), 99–108.

[24] Jain, R., T. Sripradisvarakul, and N. O'Brien, "Symbolic Surface Descriptors for 3-D Object Recognition," in *Proc. SPIE*, January 1987.

[25] Jain, R. and N. O'Brien, "Ego-Motion Complex Logarithmic Mapping," in *Proc. SPIE*, November 1984.

[26] Jerian, C. P. and R. Jain, "Determining motion parameters for scenes with translation and rotation," in *Proc. Workshop on Motion*, Toronto, 1983.

[27] Kayaalp, A. E. and R. Jain, "A knowledge-based automatic on line wafer inspection system," in *Proc. Vision 85*, pp. 5.117-5.130, 1985.

[28] Knoll, T. F. and R. Jain, "Recognizing partially visible objects using feature indexed hypothesis," *IEEE J. Robotics and Automation* 2 (1986), 3–13.

[29] Liou, S. P. and R. Jain, "Detecting Road Edges Using Hypothesized Vanishing Points," Accepted for publication in *Comp. Vision, Graph. and Image Processing.*

[30] Macworth, A., "Vision Research Strategy: Black Magic, Metaphors, Mechanisms, Miniworlds, and Maps," *Computer Vision Systems.* ed. A. Hanson and E. Riseman. New York: Academic Press, 1978.

[31] Neisser, U., *Cognition and Reality.* San Francisco: W. H. Freeman, 1976.

[32] O'Brien, N. G. and R. Jain, "Axial motion stereo," in *Proc. Workshop on Computer Vision*, Annapolis, Maryland, 1984.

[33] Rao, A. R., R. Jain, and T. E. Weymouth, "Knowledge-based vision systems," Submitted to *Computing Surveys.*

[34] Rao, A. R. and R. Jain, "Knowledge representation and control in computer vision systems," Accepted for publication in *IEEE Expert.*

[35] Schunck, B. G. "Motion segmentation and estimation," Doctoral Thesis, Department of Electrical Engineering and Computer Science, M. I. T., 1983.

[36] Schunck, B. G. "The motion constraint equation for optical flow," in *Proc. Int. J. Conf. Pattern Recognition*, pp. 20–22, Montreal, Canada, 1984.

[37] Schunck, B. G. "Motion segmentation and estimation by constraint line clustering," in *Proc. Workshop on Computer Vision*, pp. 58–62, Annapolis, Maryland, 1984.

[38] Schunck, B. G. "Surface-based smoothing of optical flow fields," in *Proc. Conference on Intelligent Systems and Machines*, pp. 107–111, Oakland University, Rochester, Michigan, 1984.

[39] Schunck, B. G. and B. K. P. Horn, "Constraints on optical flow computation," in *Proc. Conf. Pattern Recognition and Image Processing*, pp. 205–210, 1981.

[40] Schunck, B. G. "Image Flow: Fundamentals and Future Research," in *Proc. Conf. Computer Vision and Pattern Recognition*, San Francisco, 1985.

[41] Sethi, I.K. and R. Jain, "Finding trajectories of feature points in a monocular image sequence," *IEEE Trans. Patt. Anal. and Mach. Intel.* 9 (1987), 56–73.

[42] Sethi, I.K., "A fast algorithm for recognizing nearest neighbors," *IEEE Trans. Systems, Man, and Cybernetics* 11 (1981), 245–248.

[43] Sethi, I.K. and R. Jain, "Determining three dimensional structure of rotating objects using a sequence of monocular views," Technical Report, Wayne State University, 1984.

[44] Shah, M. A. and R. Jain, "Detecting Time Varying Corners," *Computer Vision, Graphics, and Image Processing* 28 (1984), 345–355.

[45] Ullman, S. *The Interpretation of Visual Motion.* Cambridge: M. I. T. Press, 1979.

[46] Weymouth, T. E., J. S. Griffith, A. R. Hanson, E. M. Riseman, "Rule Based Strategies for Image Interpretation," in *Proc. AAAI*, pp. 429–432, August 1983.

[47] Weymouth, T. E., "Using Object Descriptions in a Schema Network for Machine Vision," Ph. D. Dissertation, University of Massachusetts, April 1986.