

(12)

DTIC FILE COPY

# COMPUTER VISION RESEARCH AND ITS APPLICATIONS TO AUTOMATED CARTOGRAPHY

Final Technical Report for SRI Project 5355

Covering the period 11 June 1984 to 31 May 1986

March 1987

Contract Amount: \$3,654,877  
Effective Date: 10 December 1982  
Expiration Date: 30 September 1985  
Extension Date: 31 May 1986

By: Martin A. Fischler, Program Director, Perception  
Project Leader, 415/859-5106

Artificial Intelligence Center  
Computer and Information Science Division

Prepared for:

Defense Advanced Research Projects Agency  
1400 Wilson Boulevard  
Arlington, Virginia 22209

Attention: Lt.Col. Robert L. Simpson, Jr., Program Manager  
Information Processing Techniques Office

SRI Project 5355  
Contract No. MDA903-83-C-0027  
ARPA Order No. 3862 and AMD 8  
Program Code No. 3D30, Program Element 61101E

DTIC ELECTE  
APR 2 1987  
S A D

This document has been approved for public release and sale; its distribution is unlimited.

87 4 1 333

AD-A178 815  
SRI International



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A178815

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188 Exp. Date: Jun 30, 1986	
1a. REPORT SECURITY CLASSIFICATION unclassified		1b. RESTRICTIVE MARKINGS n/a			
2a. SECURITY CLASSIFICATION AUTHORITY n/a		3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited			
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE n/a					
4. PERFORMING ORGANIZATION REPORT NUMBER(S)  5355-FR		5. MONITORING ORGANIZATION REPORT NUMBER(S)			
6a. NAME OF PERFORMING ORGANIZATION  SRI International		6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION  DCASMA		
6c. ADDRESS (City, State, and ZIP Code)  333 Ravenswood Avenue Menlo Park, California 94025		7b. ADDRESS (City, State, and ZIP Code)  1250 Bayhill Drive San Bruno, California 94066			
8a. NAME OF FUNDING / SPONSORING ORGANIZATION  DARPA		8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER  Contract MDA903-83-C-0027		
8c. ADDRESS (City, State, and ZIP Code)  1400 Wilson Boulevard Arlington, Virginia 22209		10. SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO. 61101E	PROJECT NO. 5355	TASK NO. 3D30	WORK UNIT ACCESSION NO. 3862 & AMD8
11. TITLE (Include Security Classification)  Computer Vision Research and Its Applications to Automated Cartography					
12. PERSONAL AUTHOR(S) Martin A. Fischler					
13a. TYPE OF REPORT Final		13b. TIME COVERED FROM 6/11/84 TO 5/31/86	14. DATE OF REPORT (Year, Month, Day) 1987 March		15. PAGE COUNT 19
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)  image understanding, computer vision, automated cartography, feature extraction, stereo compilation, linear delineation		
FIELD	GROUP	SUB-GROUP			
19. ABSTRACT (Continue on reverse if necessary and identify by block number)  The SRI Image Understanding program is a broad effort spanning the entire range of machine vision research. Three major concerns are: (1) to develop a computational description of the physics and mathematics of the vision process; (2) to develop a knowledge-based framework for interpreting sensed (imaged) data; and (3) to develop a machine-based environment for effective experimentation, demonstration, and evaluation of our theoretical results, as well as providing a vehicle for technology transfer. This final report summarizes progress in these and related areas.					
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION unclassified		
22a. NAME OF RESPONSIBLE INDIVIDUAL Martin A. Fischler			22b. TELEPHONE (Include Area Code) 415/859-5106		22c. OFFICE SYMBOL EK262

## Preface

The nominal expiration date for this contract was 30 September 1985. A no-cost extension to May 1986 was requested for administrative purposes, to publish final results, and to ensure our ability to satisfy some open commitments concerned with technology transfer, but technical work on the contract was largely terminated in September 1985. The technical report issued in September 1985 [Fischler 85a] contains a relatively complete description of our most recent efforts in this program; this report provides a final update and program summary.



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A1	

# 1 Introduction

The goal of this research program was to obtain new solutions to fundamental problems in computer vision, particularly to such problems as stereo compilation, feature extraction, and general scene modeling that are relevant to the development of an automated capability for interpreting aerial imagery and the production of cartographic products.

To achieve this goal, we engaged in investigations of such basic issues as image matching, scene partitioning, shape representation, and physical modeling. However, it is obvious that high-level high-performance vision requires the use of both intelligence and stored knowledge (to provide an integrative framework), as well as an understanding of the physics and mathematics of the imaging process (to provide the basic information needed for reasoned interpretation of the sensed data). Thus, a significant portion of our work was devoted to developing new approaches to the problem of "knowledge-based vision." Finally, vision research cannot proceed without a means for effective implementation, demonstration, and experimental verification of theoretical concepts; we have developed an environment incorporating some of the newest and most effective computing tools currently available.

The research results described in this final report are partitioned into three topic areas: (1) three-dimensional scene modeling and stereo reconstruction; (2) feature extraction: scene partitioning, semantic labeling, and the representation of natural scenes; and (3) computing environments and technology transfer.

## 1.1 Three-Dimensional Scene Modeling and Stereo Reconstruction

Our goal in this research area was to develop automated methods for producing a 3-D scene model from several images recorded from different viewpoints. The standard approach to this problem is to use stereo compilation — a technique that involves finding pairs of corresponding scene points in two images (which depict the scene from different spatial locations) and using triangulation to determine scene depth [Barnard 82]. Various factors associated with viewing conditions and scene content can cause the matching process to fail; these factors include occlusion, projective or imaging distortion, featureless areas, and repeated or periodic scene structures. In this section we discuss some of the ways we devised for dealing with these problems: more effective methods for image matching, interpolation for filling in "holes" caused by matching failure, and some exciting and radically new methods for 3-D modeling which avoid the need for local matching.

### 1.1.1 Baseline Stereo System

We have implemented a complete state-of-the-art stereo system (STEREOSYS) to produce dense three-dimensional data from stereo pairs of intensity images using automated area-based stereo compilation. This system operates in several passes over the data, during which it iteratively builds, checks, and refines its model of the three-dimensional world, as originally represented by the pair of images.

Our research strategy had been to implement a baseline system that performs conventional stereo compilation, then to replace pieces of the system with improved modules as we developed them. We evaluated new developments by testing the "updated" baseline system [Hannah 85a] against a "challenge data base" [Hannah 85b] of image areas where conventional stereo techniques encounter difficulty.

Our system includes routines to perform the following operations automatically:

- Construct resolution hierarchies for stereo images.
- Select "interesting" points for sparse matching.
- Search 2-D regions for corresponding points (sparse matching).
- If necessary for uncalibrated imagery, compute relative camera parameters from sparse matches.
- Compute epipolar lines.
- Locate epipolar matches, using disparity estimates from sparse matches when available.
- Evaluate matched points for believability.
- Interpolate between matched points.
- Display images and results in left-right stereo, red-green stereo, or as a monocular disparity field.
- Compute range data and x-y-z coordinates for matched point pairs.
- Display terrain data in perspective with hidden lines removed.

(For a more complete description of the components of STEREOSYS, see [Hannah 85c] or [Hannah 86].)

Precise quantitative evaluation of the accuracy of STEREOSYS was difficult because we were not able to obtain stereo data sets with known ground truth

with which to compare our results. We did, however, have the results of an interactive stereo compilation algorithm, the Digital Interactive Mapping Program (DIMP), produced and operated by the U.S. Army Engineer Topographic Laboratories (ETL) [Norvelle 81]. Detailed comparisons of results on this and other data sets are presented in [Hannah 85a]. Overall, we have found that the results of STEREO SYS agree quite well with DIMP results and with human perceptions. In addition, STEREO SYS remedies some of the obvious problems we had seen with existing systems, such as DIMP's tendency to extrapolate itself off track — and of course, STEREO SYS is fully automatic, while comparable systems, such as DIMP, require interactive operation.

Overall, we have found that STEREO SYS performs very well on the low-resolution aerial imagery for which it was designed. It has also been applied to narrow-angle ground-based imagery with reasonable results: STEREO SYS has difficulties when processing areas that violate its premises about the continuity of the world, but experiments linking it to an edge-based matcher [Baker 82] appear to solve the most severe versions of these problems.

STEREO SYS has been used extensively at SRI and is suitable for transport to other VAX systems (however, this is research code with the corresponding limitations).

### 1.1.2 New Methods for Stereo Compilation

The conventional approach to recovering scene geometry from a stereo pair of images is based on the matching of distinctive scene features, as well as on the satisfaction of constraints imposed by the viewing geometry (e.g., the epipolar constraint). Typically, three steps are required: determination of the relative orientation of the two images, computation of a sparse depth map, and derivation of a dense depth map for the given scene.

In the first step, points corresponding to unmistakable scene features are identified in each of the images. The relative orientation of the two images is then calculated from these points. This is, in part, an unconstrained matching task. Corresponding image features must be found. Without a priori knowledge, such a matching procedure knows neither the approximate location (in the second image) of a feature found in the first image, nor the appearance of that feature. However, it is often the case that appearance will vary little between images and that the images were taken from similar positions relative to the scene.

Recovery of the relative orientation of the images reduces the need for two-dimensional matching to a one-dimensional matching problem; a scene feature identified in the first image is found by a one-dimensional search along a (epipolar) line

in the second image. Identification of this feature in the second image makes it possible to calculate the feature's disparity, and hence its relative scene depth.

Identification of corresponding points in the two images is typically based on correlation techniques. Area-based correlation processes may be applied directly to the raw image irradiances or to images that have been preprocessed in some manner. Edges (e.g., as identified by the zero crossings of the Laplacian of their image irradiances) have also been used to obtain correspondences.

The outcome of this second step is a sparse map of the scene's relative depth at those points that were identified in both images of the stereo pair.

A sparse depth map does not define the scene topography. The third and final step in recovering the topography of the scene is "filling in" this sparse map to obtain a dense depth map of the scene. Typically, a surface interpolation or approximation method is used as a means for calculating the dense depth map from its sparse counterpart [Smith 84b]. A surface approximation model may be formulated to provide desirable image properties (such as the lack of additional zero crossings — in the Laplacian of the image irradiances — that are artifacts of the surface approximation model), but often the surface model is based on a priori requirements for the fitted surface, such as smoothness.

The problems encountered in the first two steps — recovery of the relative orientation of the images and computation of the sparse depth map — are dominated by the problems of image matching. False matches that arise from repetitive scene structures, such as windows of a building, or from image features that are not distinctive (at least, on the basis of local evidence) occur more frequently in the unconstrained matching environment than in the constrained environment. In recovering the relative orientation of the images, we can use redundant information in an effort to reduce the influence of false matches; this is more difficult in the case when the sparse depth map is computed. Furthermore, we have little choice as to which features we may use for sparse depth mapping; if we choose not to use a feature, we cannot recover the relative depth at that scene point (without invoking semantic or contextual knowledge).

The selection of suitable features for determining image correspondence is difficult in itself. Correlation techniques embed assumptions that are often violated by the best image features. Area-based correlation techniques usually reflect the premise that image patches are of a scene structure that is positioned at one distinct depth, whereas edges that arise at an object's boundaries are surrounded by surfaces at different scene depths. Edge-based techniques are based on the assumption that an edge found in one image is not "moved" by the change in viewing position of the second image, whereas zero crossings found at boundaries of objects whose surface gradients are tangential to the line of sight contradict this assump-

tion. These would seem minor problems, were it not for the accuracy required of the matching process. Typically, the spatial resolution of disparity measurements must be an order of magnitude better than the image's spatial resolution. Stereo matching appears to require features with properties that are often incompatible with what is practical in realistic situations.

The third step, derivation of a dense depth map from a sparse one, is still far short of having an adequate solution. Most approaches employ "blind" interpolation, since no effective methods are currently in use for extracting depth from the irradiance data in the individual images of the stereo pair (although some of the new work described in the next section might alter this situation).

In summary, we see that the most demanding steps in the stereo compilation process are the final two: computation of a sparse depth map, and derivation of its dense counterpart. We have developed a new approach to stereo compilation which involves combining these steps to recover a dense relative-depth map of the scene directly from the image data [Smith 85 & 86]. We use image irradiance profiles as input to an integration procedure that returns the corresponding dense relative-depth profile. This procedure does not match image points (at least, not in the conventional sense), nor does it "fill in" data to obtain the dense depth map. It avoids the need to make the restrictive assumptions usually required for stereo image matching, and it directly uses the image irradiance data in recovering the dense depth map.

Other innovative approaches to stereo compilation that we have developed include:

(a) A technique [Quam 84] that merges matching and interpolation in the context of a coarse-to-fine hierarchical control structure; one of the images is geometrically warped to improve the performance of a cross-correlation-based matching technique. A surface interpolation algorithm [Smith 84b] is used to fill holes whenever the matching operator fails.

(b) A stochastic optimization approach [Barnard 86] that provides a dense array of disparities without the need for interpolation. It uses a simulated annealing algorithm to find a 3-D model which best satisfies the goals of matching points with similar intensities while ensuring that the resulting surfaces are as smooth as possible.

### **1.1.3 New Methods for 3-D Modeling Using Methods Which Do Not Depend On Stereo Correspondence**

We have noted the fact that it will not always be possible to find corresponding scene points in the two images of a conventional stereo pair, and yet — to recover a

dense scene model — we need to determine the depth at every scene point. Because interpolation will not always provide an acceptable answer when matching fails, we have developed a number of new techniques for recovering scene depth which do not require establishing stereo correspondence.

A significant body of work exists in the area of extracting depth from the shading and texture visible in a single image (e.g., [Smith 83a & 83b] and [Pentland 84a]), these different techniques make a variety of distinct assumptions about the nature of the scene, the illumination, and the imaging geometry. In Strat and Fischler [85], we show that the distinct assumptions that are used by each of the different schemes must be equivalent to providing a second (virtual) image of the original scene, and that all of these different approaches can be translated into a conventional stereo formalism. In particular, we show that it is frequently possible to structure the problem as that of recovering depth from a stereo pair consisting of a conventional perspective image (i.e., the original image) and an orthographic image (the virtual image); we provide a new algorithm needed to accomplish this type of stereo reconstruction task.

In Pentland [85a] we show how focal gradients (image “blur”), which result from the limited depth of field inherent in most optical systems, can be used to recover scene depth. The advantages of this technique are that it is fast and computationally simple, makes no special assumptions about the scene, and avoids the stereo matching problem. Mathematical analysis and experiments indicate that the accuracy achievable by this technique is comparable to what can be expected from the use of stereo disparity or motion parallax to determine scene depth.

For most purposes concerned with the analysis of imaged data, determination of an array of depths (e.g., as obtained by conventional stereo methods) is only the first step in the construction of a scene description. The conventional approach next compiles largely continuous surfaces from the discrete depth information, and then attempts to partition these surfaces into coherent 3-D objects. Aside from some still unsolved theoretical problems, this process is computationally expensive and time consuming. In Bolles and Baker [85], we describe a new method for using camera motion through a scene to obtain a 3-D model in which higher-level scene attributes are directly accessible.

This technique is based on considering a dense sequence of images as forming a solid block of data. Slices through this solid at appropriately chosen angles intermix time and spatial data in such a way as to simplify the partitioning problem: These slices have more explicit structure than the conventional images from which they were obtained. We believe that this work is a very important development; it offers a completely new and direct method for accessing information about scene objects without requiring a completely bottom-up analysis process.

## **1.2 Feature Extraction: Scene Partitioning, Semantic Labeling, and the Representation of Natural Scenes**

Creating a scene description from a photographic image requires the ability to perform two basic operations: (a) partitioning the image into independent or coherent pieces, and (b) assigning names or semantic labels to these pieces.

### **1.2.1 Scene Partitioning**

The partitioning operation, necessary to reduce the computational complexity of the subsequent scene analysis steps, has proven to be extremely difficult — the performance of automated systems is still far inferior to that of humans. In part, this disparity in performance is due to the fact that humans appear to use contextual knowledge and past experience in such tasks, while most available computational techniques employ only the local intensity patterns visible in the image, i.e., they perform “syntactic partitioning.” For practical as well as theoretical reasons, we have carried out an investigation to determine the competence limits of a purely syntactic approach to partitioning and, simultaneously, to construct an operational system that approaches these limits. This investigation has resulted in a very high performance system described in a paper by Laws [85].

Barnard [84b] describes one of a number of investigations that attempt to provide a theoretical basis for the partitioning process. In this paper, Barnard explores the idea that partitioning decisions result in alternative descriptions of a scene, and that the preferred partitioning is the one that provides the “simplest” description. In a paper by Fischler and Bolles [83], partitioning is viewed as an explanation of how the image is related to the scene from which it was derived; it is shown that completeness and stability of explanation, as well as simplicity, are necessary partitioning criteria, since these attributes are necessary for an explanation to be believable.

### **1.2.2 Feature Delineation and Semantic Labeling**

In Fua and Hanson [85, 86a & 86b], we describe an approach to the problem of converting a syntactically partitioned image (e.g., one provided by Laws' segmentation system) into a semantic description. This work has produced a system that can extract cultural objects from aerial imagery; it uses geometric reasoning to identify semantically significant arrangements of straight line segments in the borders of the supplied partition. Emphasis is placed on using generic models that characterize significant kinds of geometric relationships and shapes, thereby avoiding the well-known drawbacks inherent in the use of specific object templates. An important

feature of this system is the generation of an explanation for any detected discrepancy between the hypothesized object models and the initial partition. In principle, this technique should ultimately permit intelligent compensation for anomalies due to imaging or environmental effects that would be recognized by a well-briefed human analyst; for example, on the basis of illumination effects consistent with the known sun position, the system can identify two contrasting regions of a peaked roof as belonging to a single house.

### 1.2.3 The Representation and Recognition of Natural Forms

Our research in this area addressed two related problems: (1) representing natural shapes such as mountains, vegetation, and clouds, and (2) computing such descriptions from image data. A key step toward solving these problems is to obtain a model of natural surface shapes.

A model of natural surfaces is extremely important because we face problems that seem impossible to address with standard descriptive computer-vision techniques. How, for instance, can we describe and recognize the shape of leaves on a tree? Or grass? Or clouds? When we attempt to describe such common natural shapes using standard representations, the result is a model of impractical complexity.

Furthermore, how can we extract 3-D information from the image of a textured surface when we have no models that describe natural surfaces and how they evidence themselves in the image? The lack of such a 3-D model has restricted image texture descriptions to being ad hoc statistical measures of the image intensity surface.

Fractal functions, a novel class of naturally arising functions, are a good choice for modeling natural surfaces because many basic physical processes (e.g., erosion and aggregation) produce a fractal surface shape, and because fractals are widely used as a graphics tool for generating natural-looking shapes. Additionally, in a survey of natural imagery, we found that a fractal model of imaged 3-D surfaces furnishes an accurate description of both textured and shaded image regions, thus providing justification for the use of this physics-derived model.

Progress relevant to computing 3-D information from imaged data has been achieved by use of the fractal model. A test has been derived to determine whether the fractal model is valid for a particular set of image data, an empirical method for computing surface roughness from image data has been developed, and substantial progress has been made in the areas of shape-from-texture and texture segmentation. Characterization of image texture by means of a fractal surface model has

also shed considerable light on the physical basis for several of the texture partitioning techniques currently in use, and made it possible to describe image texture in a manner that is stable over transformations of scale and linear transforms of intensity. The computation of a 3-D fractal-based representation from actual image data has been demonstrated. This work has shown the potential of a fractal-based representation for efficiently computing good 3-D representations for a variety of natural shapes, including such seemingly difficult cases as mountains, vegetation, and clouds.

Finally, the fractal model of surface shape has been used in a technique for 3-D shape estimation that treats shading and texture in a unified manner. Previously, shape-from-shading and texture methods have had the serious drawback that they are applicable only to smooth surfaces, while real surfaces are often rough and crumpled. We have extended one class of such methods to more realistic surfaces by using the fractal surface model, constructing a method for estimating 3-D shape that treats shading and texture in a unified manner [Pentland 84a]

We have constructed a representational system that combines the fractal functions described above, for use in describing 3-D texture, and superquadric functions (defined below) for describing shape in a concise and natural manner.

The idea behind this representational system is to provide a vocabulary of shapes and transformations that will allow us to model an object world as a relatively simple composition of component "parts," much as people seem to do. The most primitive notion in this representation may be thought of as analogous to a "lump of clay," a modeling primitive that may be deformed and shaped, but that is intended to correspond roughly to our naive perceptual notion of "a part." For this basic modeling element we use a parameterized family of shapes known as *superquadrics*. This family of functions includes cubes, cylinders, spheres, diamonds, and pyramidal shapes as well as the round-edged shapes intermediate between these standard shapes. Superquadrics are, therefore, a superset of the modeling primitives currently in common use.

These basic "lumps of clay" are used as prototypes that are then deformed by stretching, bending, twisting or tapering, and then combined using Boolean operations to form new, complex prototypes that may, recursively, again be subjected to deformation and Boolean combination [Pentland 86b, 86c & 86d].

We have also made significant progress toward the reliable recovery of these modeling primitives from image data. We have developed theoretical results that show how such descriptive primitives may be recovered in an overconstrained, and therefore reliable, manner [Pentland 86e].

This research has contributed to the development of a computational theory of vision applicable to natural surface shapes, compact representations of shape useful

for describing natural surfaces, and real-time regeneration and display of natural scenes.

## 2 Computing Environments and Technology Transfer

Machine vision is largely an experimental science; progress in this science depends on having available massive amounts of computing power, and on methods for manipulating and displaying images, their transformations, and depictions of the corresponding scene content. Technology transfer must often be in the form of machine code that can run in a compatible computing environment.

As part of our previously described work, we have built a research environment based on the VAX-11/780 computer and have made this environment available to appropriate university and government institutions. This environment includes standard utilities (e.g., low-level image operators) and advanced scene-modeling and recognition techniques (e.g., the Hannah Baseline Stereo System, the Laws Segmentation System, and the Generalized Hough Transform).

More recently, we have constructed a powerful computing environment based on the Symbolics 3600 series of LISP machines and an SRI product called Image-Calc. Using this environment as a substrate, we have developed a number of scene analysis and rendering systems. For example, a system called Terrain-Calc [Quam 85] can be used to synthesize realistic sequences of perspective stereo views of real-world terrain from stored geometric and photometric models. This system has a sophisticated graphical interface that allows the user to specify an arbitrary flight path over a modeled piece of terrain. A sequence of views (single images or stereo pairs, as desired), spaced at equal distances along the flight path, can be generated at about 1 frame/minute; up to 60 frames can be displayed at 16 frames/second. This system is revolutionary in its flexibility, computational efficiency, and the quality of the renderings it produces — given that it does not employ any special-purpose hardware.

The computational demands of practical machine-vision applications frequently exceed the capacity of conventional serial computer architectures. For this reason, attempts have been made to reduce computation time by decomposing serial algorithms into segments that can be executed simultaneously on parallel hardware architectures. Because many classes of algorithms do not readily decompose, one seeks some other basis for parallelism. We have investigated the use of techniques that exhibit natural parallelism. For example, in Fischler and Firschein [87] we show that “guessing” the answer to a problem and then checking its validity is a

useful approach, and that a number of important vision algorithms can be viewed as having this structure; a parallel architecture capable of executing such algorithms is described.

### **3 Acknowledgments**

The following researchers have contributed to the work described in this report: H. Baker, S.T. Barnard, R.C. Bolles, O. Firschein, M.A. Fischler, P. Fua, M.J. Hannah, A.J. Hanson, K. Laws, D.H. Marimont, A.P. Pentland, L.H. Quam, G.B. Smith, T. Strat, and H.C. Wolf.

## Bibliography

- [Baker 82] H.H. Baker, *Depth from Edge and Intensity Based Stereo*, Ph.D. Thesis, Report STAN-CS-82-930, Stanford University Computer Science Department (September 1982). Also published as Report AIM-347 Artificial Intelligence Laboratory, Stanford University, Stanford, California, (1982).
- [Barnard 82] S.T. Barnard and M.A. Fischler, "Computational Stereo," *ACM Computing Surveys*, Vol. 14, No. 4, pp. 553-572 (December 1982). Also published in *Proc. DARPA Image Understanding Workshop*, Washington, D.C., pp. 157-167 (23 April 1981), and as Technical Note 236, Artificial Intelligence Center, SRI International, Menlo Park, California (March 1981).
- [Barnard 83] S.T. Barnard and A.P. Pentland, "Three-Dimensional Shape from Line Drawings," *Proc. 8th Int. Int. Conf. on Artificial Intelligence*, Karlsruhe, West Germany, pp. 1062-1064 (8-12 August 1983). Also published in *Proc. DARPA Image Understanding Workshop*, Arlington, Virginia, pp. 282-284 (23 June 1983), and as Appendix A of [Fischler 83c].
- [Barnard 84a] S.T. Barnard, "Choosing a Basis for Perceptual Space," *Proc. IEEE Workshop on Computer Vision: Representation and Control*, Annapolis, Maryland, pp. 225-230 (30 April-2 May 1984). Also published in *Computer Vision, Graphics, and Image Processing*, Vol. 29, No. 1, pp. 87-99 (January 1985), as Technical Note 315, Artificial Intelligence Center, SRI International, Menlo Park, California (January 1984), and as Appendix A of [Fischler 84a].
- [Barnard 84b] S.T. Barnard, *An Inductive Approach to Figural Perception*, Technical Note 325, Artificial Intelligence Center, SRI International, Menlo Park, California (September 1984). Also published as Appendix F of [Fischler 85a].
- [Barnard 86] S.T. Barnard, "A Stochastic Approach to Stereo Vision," *Proc. 5th National Conf. on Artificial Intelligence*, Philadelphia, Pennsylvania, pp. 676-680 (11-15 August 1986). Also published as Technical Note 373, Artificial Intelligence Center, SRI International, Menlo Park, California (April 1986).
- [Bolles 85] R.C. Bolles and H.H. Baker, "Epipolar-Plane Image Analysis: A Technique for Analysing Motion Sequences," *Proc. DARPA Image Understanding Workshop*, Miami Beach, Florida, pp. 137-148 (9-10 December 1985). Also published in *Proc. 3rd Int. Symp. on Robotics Research*, Paris, France, pp. 41-48 (October 1985), in *Proc. IEEE 3rd Workshop on Computer Vision: Representation and Control*, Bellaire, Michigan, pp. 168-178 (13-16 October 1985), and as Appendix E of [Fischler 85a].
- [Bolles 87] R.C. Bolles, H.H. Baker, and D.H. Marimont, "Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion," *International Journal of Computer Vision* (March 1987, in press).
- [Fischler 83a] M.A. Fischler, "The SRI Image Understanding Research Program," *Proc. DARPA Image Understanding Workshop*, Arlington, Virginia, pp. 24-27 (23 June 1983).

- [Fischler 83b] M.A. Fischler and H.C. Wolf, "Linear Delineation," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Washington, D.C., pp. 351-356, (19-23 June 1983).
- [Fischler 83c] M.A. Fischler, *Computer Vision Research and its Applications to Automated Cartography*, First Semiannual Technical Report, SRI Project 5355, Artificial Intelligence Center, SRI International, Menlo Park, California (July 1983).  
Includes [Barnard 83], [Smith 83b], [Pentland 84b], and [Fischler 83e].
- [Fischler 83d] M.A. Fischler and H.C. Wolf, "Machine Perception of Linear Structure," *Proc. 8th Int. Jnt. Conf. on Artificial Intelligence*, Karlsruhe, West Germany, pp. 1010-1013 (8-12 August 1983). Also published as *A General Approach to Machine Perception of Linear Structure in Imaged Data*, Technical Note 276, Artificial Intelligence Center, SRI International, Menlo Park, California (February 1983).
- [Fischler 83e] M.A. Fischler and R.C. Bolles, "Perceptual Organisation and the Curve Partitioning Problem," *Proc. 8th Int. Jnt. Conf. on Artificial Intelligence*, Karlsruhe, West Germany, pp. 1014-1018 (8-12 August 1983). Also published as Appendix D of [Fischler 83c]. Slightly expanded version of "Perceptual Organisation and Curve Partitioning," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Washington, D.C., pp. 38-46 (19-23 June 1983), also published in *Proc. DARPA Image Understanding Workshop*, Arlington, Virginia, pp. 224-232 (23 June 1983).
- [Fischler 84a] M.A. Fischler, *Computer Vision Research and its Applications to Automated Cartography*, Second and Third Semiannual Technical Reports, SRI Project 5355, Artificial Intelligence Center, SRI International, Menlo Park, California (September 1984).  
Includes [Barnard 84a], [Strat 84a], [Strat 84b], [Hannah 84], [Smith 84b], [Quam 84], [Pentland 86e], [Laws 85], and [Smith 84a].
- [Fischler 84b] M.A. Fischler, "The SRI Image Understanding Research Program," *Proc. DARPA Image Understanding Workshop*, New Orleans, Louisiana, pp. 8-13 (3-4 October 1984).
- [Fischler 85a] M.A. Fischler, *Computer Vision Research and its Application to Automated Cartography*, Fourth and Fifth Semiannual Technical Reports, SRI Project 5355, Artificial Intelligence Center, SRI International, Menlo Park, California (September 1985).  
Includes [Hannah 85a], [Smith 85b], [Strat 86], [Pentland 87], [Bolles 85], [Barnard 84b], [Fua 85a], [Pentland 86a], [Quam 85], and [Fischler 87].
- [Fischler 85b] M.A. Fischler, "The SRI Image Understanding Research Program," *Proc. DARPA Image Understanding Workshop*, Miami Beach, Florida, pp. 40-44 (9-10 December 1985).
- [Fischler 87] M.A. Fischler and O. Firschein, "Parallel Guessing: A Strategy for High-Speed Computation," *Pattern Recognition* (1987, in press). Also published as Technical Note 338, Artificial Intelligence Center, SRI International, Menlo Park, California (September 1984), and as Appendix J of [Fischler 85a].

- [Fua 85a] P. Fua and A.J. Hanson, "Object Labeling Using Generic Knowledge," in [Fischler 85a], Appendix G, pp. 1-10 (September 1985).
- [Fua 85b] P. Fua and A.J. Hanson, "Locating Cultural Regions in Aerial Imagery Using Geometric Cues," *Proc. DARPA Image Understanding Workshop*, Miami Beach, Florida, pp. 271-278 (9-10 December 1985).
- [Fua 86] P. Fua and A.J. Hanson, *Using Generic Geometric Knowledge to Delineate Cultural Objects in Aerial Imagery*, Technical Note 378, Artificial Intelligence Center, SRI International, Menlo Park, California (March 1986).
- [Hannah 84] M.J. Hannah, *Description of SRI's Baseline Stereo System*, Technical Note 342, Artificial Intelligence Center, SRI International, Menlo Park, California (October 1984). Also published as Appendix D of [Fischler 84a].
- [Hannah 85a] M.J. Hannah, *Evaluation of STEREOSYS vs. Other Stereo Systems*, Technical Note 365, Artificial Intelligence Center, SRI International, Menlo Park, California (October 1985). Also published as Appendix A of [Fischler 85a].
- [Hannah 85b] M.J. Hannah, *The Stereo Challenge Data Base*, Technical Note 366, Artificial Intelligence Center, SRI International, Menlo Park, California (October 1985).
- [Hannah 85c] M.J. Hannah, *SRI's Baseline Stereo System*, *Proc. DARPA Image Understanding Workshop*, Miami Beach, Florida, pp. 149-155 (9-10 December 1985).
- [Hannah 86] M.J. Hannah, *Users's Guide for the STEREOSYS Stereo System*, Technical Note, Artificial Intelligence Center, SRI International, Menlo Park, California (1986, in preparation).
- [Laws 85] K.I. Laws, "Goal-Directed Textured-Image Segmentation," *Proc. SPIE Conf. on Applications of Artificial Intelligence II*, Vol. 548, Arlington, Virginia, (9-11 April 1985). Summarizes Technical Note 334, Artificial Intelligence Center, SRI International, Menlo Park, California (September 1984), which has also been published as Appendix H of [Fischler 84a].
- [Marimont 86] D.H. Marimont, "Projective Duality and the Analysis of Image Sequences," *Proc. Workshop on Motion: Representation and Analysis*, Kiawah Island, South Carolina, pp. 7-14 (7-9 May 1986).
- [Norvelle 81] F.R. Norvelle, *Interactive Digital Correlation Techniques for Automatic Compilation of Elevation Data*, Report ETL-0272, U.S. Army Engineer Topographic Laboratories (October 1981).
- [Pentland 84a] A.P. Pentland, "Local Shading Analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 2, pp. 170-187 (March 1984), reprinted in [Pentland 85a], pp. 40-77 (1985). Also published as Technical Note 272, Artificial Intelligence Center, SRI International, Menlo Park, California (November 1982).
- [Pentland 84b] A.P. Pentland, "Fractal-Based Description of Natural Scenes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 6, pp. 661-674 (November 1984), reprinted in R. Chellappa and A.A. Sawchuk (eds.), *Digital Image Processing and Analysis: Volume 2: Digital Image Analysis*, pp. 368-381 (1985), and in [Pentland 85a], pp. 227-252 (1985). Also published in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Washington, D.C., pp. 201-209 (19-23 June 1983), in *Proc. DARPA Image Understanding Workshop*, Arlington, Virginia, pp. 184-192 (23 June 1983), and as Appendix C of [Fischler 83c].

Slightly expanded version of "Fractal-Based Description," in W. Richards (ed.), *Natural Computation*, MIT Press, also published in *Proc. 8th Int. Jnt. Conf. on Artificial Intelligence*, Karlsruhe, West Germany, pp. 973-981 (8-12 August 1983).

- [Pentland 85a] A.P. Pentland (ed.), *From Pixels to Predicates: Recent Advances in Computational and Robotic Vision*, Ablex, Norwood, New Jersey (1985).  
Includes [Pentland 84a, 84b, and 86f].
- [Pentland 85b] A.P. Pentland, "On Describing Complex Surface Shapes," *Image and Vision Computing*, Vol. 3, No. 4, pp. 153-162 (November 1985).
- [Pentland 86a] A.P. Pentland, "Perceptual Organization and the Representation of Natural Form," *Artificial Intelligence*, Vol. 28, No. 3, pp. 293-331 (May 1986). Also published as Technical Note 357, Artificial Intelligence Center, SRI International, Menlo Park, California (July 1985, revised July 1986), and as Appendix H of [Fischler 85a], and summarized in *Optics News* (October 1985).
- [Pentland 86b] A.P. Pentland, "On Perceiving 3-D Shape and Texture," *Computational Models in Human Vision*, Rochester, New York (19-21 June 1986).
- [Pentland 86c] A.P. Pentland, "Part Models," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Miami Beach, Florida, pp. 243-249 (22-26 June 1986).
- [Pentland 86d] A.P. Pentland, "Parts: Structured Descriptions of Shape," *Proc. 5th National Conf. on Artificial Intelligence*, Philadelphia, Pennsylvania, pp. 695-701 (11-15 August 1986).
- [Pentland 86e] A.P. Pentland, "Shading into Texture," *Artificial Intelligence*, Vol. 29, No. 2, pp. 147-170 (August 1986). Also published in *Proc. 4th National Conf. on Artificial Intelligence*, Austin, Texas, pp. 269-273 (6-10 August 1984), in *Proc. DARPA Image Understanding Workshop*, New Orleans, Louisiana, pp. 179-183 (3-4 October 1984), as Appendix G of [Fischler 84a], and in [Pentland 85a], pp. 253-267 (1985).
- [Pentland 86f] A.P. Pentland, *Recognition by Parts*, Technical Note 406, Artificial Intelligence Center, SRI International, Menlo Park, California (December 1986).
- [Pentland 87] A.P. Pentland, "A New Sense for Depth of Field," *IEEE Trans. on Pattern Analysis and Machine Intelligence* (1987, to appear). Also published in *Proc. 9th Int. Jnt. Conf. on Artificial Intelligence*, Los Angeles, California, pp. 988-994 (18-23 August 1985), and as Appendix D of [Fischler 85a].
- [Quam 84] L.H. Quam, "Hierarchical Warp Stereo," *Proc. DARPA Image Understanding Workshop*, New Orleans, Louisiana, pp. 149-155 (2-3 October 1984). Also published as Appendix F of [Fischler 84a].
- [Quam 85] L.H. Quam, "The Terrain-Calc System," *Proc. DARPA Image Understanding Workshop*, Miami Beach, Florida, pp. 327-330 (9-10 December 1985). Also published as Appendix I of [Fischler 85a].
- [Smith 83a] G.B. Smith, "Shape from Shading: An Assessment," *Proc. NASA Symp. on Mathematical Pattern Recognition and Image Analysis*, Houston, Texas, pp. 543-576 (1-3 June 1983). Also published as Technical Note 287, Artificial Intelligence Center, SRI International, Menlo Park, California (May 1983).

- [Smith 83b] G.B. Smith, "The Relationship Between Image Irradiance and Surface Orientation," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Washington, D.C., pp. 14-19 (19-23 June 1983). Also published in *Proc. DARPA Image Understanding Workshop*, Arlington, Virginia, pp. 243-248 (23 June 1983), as Technical Note 306, Artificial Intelligence Center, SRI International, Menlo Park, California (September 1983), and as Appendix B of [Fischler 83c].
- [Smith 84a] G.B. Smith and H.C. Wolf, "Image-to-Image Correspondence: Linear-Structure Matching," *Proc. 2nd Annual NASA Symp. on Mathematical Pattern Recognition and Image Analysis*, Houston, Texas, pp. 467-487 (6-8 July 1984). Technical Note 331, Artificial Intelligence Center, SRI International, Menlo Park, California (July 1984), and as Appendix I of [Fischler 84a].
- [Smith 84b] G.B. Smith, "A Fast Surface Interpolation Technique," *Proc. DARPA Image Understanding Workshop*, New Orleans, Louisiana, pp. 211-215 (2-3 October 1984). Also published as Technical Note 333, Artificial Intelligence Center, SRI International, Menlo Park, California (August 1984), and as Appendix E of [Fischler 84a].
- [Smith 85a] G.B. Smith, "Recovery of Surface Shape from Multiple Images," *Proc. 3rd Annual NASA Symp. on Mathematical Pattern Recognition and Image Analysis*, College Station, Texas, pp. 217-238 (10-11 June 1985).
- [Smith 85b] G.B. Smith, "Stereo Reconstruction of Scene Depth," *Proc. Computer Vision and Pattern Recognition*, San Francisco, California, pp. 271-276 (19-23 June 1985). Also published as Appendix B of [Fischler 85a].
- [Smith 86] G.B. Smith, "Stereo Integral Equation," *Proc. 5th National Conf. on Artificial Intelligence*, Philadelphia, Pennsylvania, pp. 689-694 (11-15 August 1986). Also published as Technical Note 379, Artificial Intelligence Center, SRI International, Menlo Park, California (September 1986).
- [Strat 84a] T.M. Strat, "Spatial Reasoning from Line Drawings of Polyhedra," *Proc. IEEE Workshop on Computer Vision: Representation and Control*, Annapolis, Maryland, pp. 219-224 (30 April-2 May 1984). Also published in *Proc. DARPA Image Understanding Workshop*, New Orleans, Louisiana, pp. 230-235 (2-3 October 1984), and as Appendix B of [Fischler 84a].
- [Strat 84b] T.M. Strat, "Recovering the Camera Parameters from a Transformation Matrix," *Proc. DARPA Image Understanding Workshop*, New Orleans, Louisiana pp. 264-271 (2-3 October 1984). Also published as Appendix C of [Fischler 84a].
- [Strat 86] T.M. Strat and M.A. Fischler, "One-Eyed Stereo: A General Approach to Modeling 3-D Scene Geometry," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp. 730-741 (November 1986). Also published in *Proc. 9th Int. Jnt. Conf. on Artificial Intelligence*, Los Angeles, California, pp. 937-943 (18-23 August 1985), in *Proc. DARPA Image Understanding Workshop*, Miami Beach, Florida, pp. 363-372 (9-10 December 1985), as Appendix C of [Fischler 85a], and as *One-Eyed Stereo: A Unified Strategy to Recover Shape from a Single Image*, Technical Note 367, Artificial Intelligence Center, SRI International, Menlo Park, California (5 November 1985).