

AD-A040 196

NORTH CAROLINA STATE UNIV RALEIGH SIGNAL PROCESSING LAB F/6 9/2
COMPUTER VISION USING ENCODED STEREO IMAGES.(U)
JAN 77 J N ENGLAND

UNCLASSIFIED

SPL-18

ARO-13989.1-EL

DAA629-76-8-0133

NL

| OF |
AD
A040196



END

DATE
FILMED

6-77

020-13989.1-EL(9)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER	2. JOINT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) COMPUTER VISION USING ENCODED STEREO IMAGES.		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL MEMO
6. AUTHOR(s) J.N./England		7. PERFORMING ORG. REPORT NUMBER SPL-18
8. AUTHORING ORGANIZATION NAME AND ADDRESS Electrical Engineering Dept. NC State Univ. Raleigh, NC 27607		9. CONTRACT OR GRANT NUMBER(s) DAAG 29-76-G-0133
10. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		11. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
12. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. REPORT DATE 1777 Jan 77
14. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		15. NUMBER OF PAGES 7
15. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) NA		16. SECURITY CLASS. (of this report) Unclassified
16. SUPPLEMENTARY NOTES The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.		17. DECLASSIFICATION/DOWNGRADING SCHEDULE NA
17. KEY WORDS (Continue on reverse side if necessary and identify by block number) SCENE ANALYSIS, COMPUTER VISION, STEREOSCOPIC IMAGES		18. TECHNICAL MEMO 9
18. ABSTRACT (Continue on reverse side if necessary and identify by block number) The methods of depth determination used in scene analysis are discussed. Previous schemes incorporating a single view of the scene are reviewed, including methods requiring a special illumination source. A review of the work using two (stereoscopic) images is presented. Finally, a method for extracting objects from a pair of run length coded images is developed. The procedure relies on feature extraction and correlation techniques developed specifically for operation on objects in run length coded images.		19. JUN 6 1977 RECEIVED

DD FORM 1473

EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD NO.

DDC FILE COPY

AD A 040196

408 851/58

SIGNAL PROCESSING LABORATORY

REPORT #18

Computer Vision Using Encoded Stereo Images

by

J. N. England

January 1977

NORTH CAROLINA STATE UNIVERSITY
ELECTRICAL ENGINEERING DEPARTMENT

ACCESSION for	
NTIS	White Section ☒
DOC	Buff Section ☐
UNANNOUNCED	☐
JUSTIFICATION.....	
BY.....	
DISTRIBUTION/AVAILABILITY CODES	
Dist.	AVAIL. and/or SPECIAL
<i>R</i>	

This work was supported by Grant DAAG 29-76-G-0133, U.S. Army
Research Office

ABSTRACT

The methods of depth determination used in scene analysis are discussed. Previous schemes incorporating a single view of the scene are reviewed. These include methods requiring a special illumination source. A review of the work using two (stereoscopic) images is presented. Finally, a method for extracting objects from a pair of run length coded images is developed. The procedure relies on feature extraction and correlation techniques developed specifically for operation on objects in run length coded images.

I. Introduction

The use of three dimensional (as opposed to the more common two dimensional) information in scene analysis has been shown to be quite beneficial. Perhaps the clearest indication of the value of obtaining depth information is in the work recently reported (1) at the Third International Joint Conference on Pattern Recognition (3IJCPR). Although the approach to determine depth is different than that taken here it is striking to note the ease with which separation and recognition of objects within a scene can be accomplished when using both range and intensity data.

II. Depth Determination

The analysis by computer of a scene from an image or images of that scene is an important problem that has been approached by many researchers with many different methods. We shall restrict our discussion here to methods that have depth determination as a key prelude to the separation and classification of objects within the scene.

A. Single views

Only a few attempts have been made to determine the explicit three dimensional shapes of objects from a single view without using any special form of illumination. Horn's use of shading (2,3) to derive shape information is certainly the most notable. This cue to depth is useful, however, only for smooth, uniformly colored, smoothly curved objects.

In order to extract depth information from a scene several researchers have tried special illumination of the scene in question. One of the more innovative of these methods is that

of Will and Pennington (4). They illuminated an object with a gridded light source and proceeded to extract differently oriented surfaces through filtering in the Fourier transform domain. Shirai and Tsuji (5) sequentially illuminated the scene from different directions and thereby obtained information about orientation of surfaces. This method, although it relies on a single image viewpoint, is akin to the information obtained from multiple viewpoints.

The use of various rangefinding techniques (again with a single image viewpoint) has met with some success. Agin and Binford (6) and Shirai and Suwa (7) used what is, in effect, a cutting plane of light to illuminate the scene and computed range by using an image obtained from a viewpoint located at some angle with respect to the illuminating plane. This triangulation method is similar to stereoscopic (two-image) range determination without the problems of correlation between the two images since points are uniquely identified by the illumination scheme.

A true rangefinder using a single beam of light to measure range and reflection from the scene has been used at Stanford Research Institute (1,8). The phase shift of the reflected beam is used to determine range to the illuminated point.

B. Multiple views

Whenever two or more views of a scene are available we may obtain depth information by noting the disparities between points within the views. If we know the geometry of the situation involved, a straight-forward solution can be obtained providing the corresponding locations of points within the available images are known. It is finding the corresponding locations that provide the

real challenge. Certainly one way to identify the same point in different views is to use the excellent image processing and pattern recognition ability of a human to point out to the computer system the appropriate locations. This, however, is practical only in limited situations or for debugging an automatic system as in (9).

In order to simplify the task a number of researchers have started with idealized line drawings of the scene in question (10,11, 12) and have chiefly used the vertices of the planar polygons so indicated as reference points. Unfortunately this simply shifts part of the problem away since the extraction of crude line drawings, much less idealized ones, from real images is no mean task in itself.

In the area of true stereoscopy with real images most work has gone into determining suitable feature points (as the vertices, above) for the correlation process. One exception which should be noted, however, is the work of Marr and Poggio (13). They have used a cooperative highly-interconnected parallel processing network to extract depth information even from random dot stereograms where no monocularly visible form for registration is available. The use of a network of this type is, as they point out, vastly more complicated than the modes of computation normally available in non-biological systems.

Otherwise, the problem has been approached on a feature point extraction and cross-image matching basis. This process is very similar to the registration problem in satellite and spacecraft imagery and work in stereo computer vision has drawn on this field.

Quam's work in registration (14,15) laid the foundation for the work at the Stanford Artificial Intelligence Laboratory by several researchers. Hannah (16) took the cross-correlation techniques and added the ideas of sampling across the image, the adaptive threshold scheme of Barnea and Silverman (17), and the notion of local coherency to develop an effective system. Local coherency simplifies the search for a match by using the idea that once a match B for target A is found, the match for target which is near A will be found near B. This was extended somewhat by Thompson (18).

Pingle and Thomas (19) have developed a feature extractor to identify targets (specifically corners) which have a high probability of being matched.

III. Object identification from stereo views.

We now present a procedure for extracting three-dimensional objects from stereo views of scenes. We shall restrict the objects to consist predominantly (but not totally) of planar polygonal surfaces.

If we use a run-length coded data structure (20) for the two images we may apply the feature extractor of (20) starting in the upper left corner of one of the images. Due to the nature of this feature extractor, we will assume for the time being that we have found a polygon vertex and will place its X, Y coordinates in a vertex list. We now apply the feature extractor to the other image again starting in the upper left corner of the image. If we have arranged our camera geometry correctly we expect to find the matching vertex in the second image at approximately the same Y coordinate as the target. Using the correlation scheme of (20) on the two run-length coded regions, we store the candidate X, Y coordinates and correlation score in a second vertex list. When all features within $\pm \Delta Y$

of the target Y have been scored we choose as a match that feature which the highest score. Returning again to the first image we follow a region edge to the right through the data structure as in (20). At the next end of the edge, we once again apply the feature extraction and correlation check. If our second target Y value is near to the first, we may only have to add a few features to the candidate list for correlation checking.

As a check, we should now try an edge following between the two chosen points in the second image. If no edge connects these two we must rechoose the next lower scoring features. If the only edge found includes a low scoring candidate we may assume some problem (occluded vertex, etc.) exists and we must backtrack and try edge following in a different direction, deleting a target if no success is found in this manner.

The above procedure should be repeated around a region to obtain a list of vertices and their X, Y, and Z values. If these vertices are roughly coplanar, we have identified a planar polygon and should move on to an adjacent region.

Adopting the notion of a jump discontinuity as in (1) we may proceed to separate objects in the scene on that basis. At this time, we should also check for coplanar adjacent polygons and merge them as necessary. This type of artificial separation may exist due to shading irregularities on the surface.

IV. Implementation

The above procedure has been developed as a first cut at the problem of object depth determination and extraction in run-length coded images. Implementation of these ideas during Spring 1977 will illuminate difficulties in the procedure and suggest pertinent revisions.

Publication of the results of the actual implementation and revisions is hoped for.

REFERENCES

- [1] D. Nitzan and R.O. Duda, "Low-level processing of registered intensity and range data", Proc. Third Int'l. Conf. on Pattern Recognition, IEEE 76CH1140-3C (Nov. 1976) 598-601.
- [2] B.K.P. Horn, "Shape from shading", MIT, MAC TR-59 (Nov. 1970).
- [3] B.K.P. Horn, "Image intensity understanding", MIT AI Lab AIM-335 (Aug. 1975).
- [4] P.M. Will and K.S. Pennington, "Grid Coding: A preprocessing technique for robot and machine vision", Artificial Intelligence 2 (1971) 285-318.
- [5] Y. Shirai and S. Tsuji, "Extract of the line drawings of 3-dimensional objects by sequential illumination from several directions", 2IJCAI (1971) 71-79.
- [6] G.J. Agin and T.O. Binford, "Computer description of curved objects", 3IJCAI, (1973) 629-638.
- [7] Y. Shirai and M. Suwa, "Recognition of polyhedrons with a range finder", 2IJCAI (1971) 80-87.
- [8] D. Nitzan, A.E. Brain and R.O. Duda, "Measurement and use of reflectance and range data for machine perception", SRI Tech. Note 128 (March 1976).
- [9] Y. Yakimovsky and R. Cunningham, "A system for extracting 3-dimensional measurements from a stereo pair of TV cameras", NASA-CR-147149 (May, 1976).
- [10] S. Ganapathy, "Reconstruction of scenes containing polyhedra from a stereo pair of views", Stanford AI Lab AIM-272 (Dec. 75).
- [11] G. Lafue, "Computer recognition of three-dimensional objects from orthographic views", Carnegie-Mellon Univ. Inst. of Physical Planning Res. Rep. 56 (Sept. 1975).
- [12] S.A. Underwood and C.L. Coates, "Visual learning from multiple views", Univ. of Texas at Austin Elec. Res. Ctr. Tech. Rept. 158, (May 1974).
- [13] D. Marr and T. Poggio, "Cooperative computation of stereo disparity", MIT AI Lab Memo 364, (June 1976).
- [14] L.H. Quam, "Computer comparison of pictures", Stanford AI Lab AIM-144 (May 1971).
- [15] L.H. Quam and M.J. Hannah, "Stanford automatic photogrammetry research", Stanford AI Lab AIM-254 (Dec. 1974).
- [16] M.H. Hannah, "Computer matching of areas in stereo images", Stanford AI Lab AIM-239 (1974).

- [17] D.I. Barnea and H.F. Silverman, "A class of algorithms for fast digital image registration", IEEE Trans. Comp. C-21, No. 2 (Feb. 1972) 179-186.
- [18] C. Thompson, "Depth perception in stereo computer vision", Stanford AI Lab, AIM-268 (Oct. 1975).
- [19] K.K. Pingle and A.J. Thomas, "A fast, feature-driven stereo depth program", Stanford AI Lab AIM-248 (May 1975).
- [20] J.N. England, "Run length coding for image analysis", NC State Univ. E.E. Dept, Sig. Proc. Lab Report 17 (Jan. 1977).