NASA Contractor Report 165804

Computational Theory of Line Drawing Interpretation

Andrew P. Witkin, Principal Investigator
Artificial Intelligence Center
Computer Science and Technology Division

SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

CONTRACT NAS1-16282
DECEMBER 1981

NASA
National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

NASA Contractor Report 165804

Computational Theory of Line Drawing Interpretation

Andrew P. Witkin, Principal Investigator
Artificial Intelligence Center
Computer Science and Technology Division

SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

CONTRACT NAS1-16282
DECEMBER 1981

NASA
National Aeronautics and
Space Administration

Langley Research Center
Hampton Virginia 23665

N82-17897 #

CONTENTS

ILLUSTRATIONS

I INTRODUCTION AND OVERVIEW

This is SRI International's final report to the National Aeronautics and Space Administration for Contract No. NAS1-16282, Computational Theory of Line Drawing Interpretation. A central problem for visual perception in man and machine is the recovery of the properties and three-dimensional structure of visible surfaces. Our research has focused on the recovery of these intrinsic scene characteristics by emphasizing the role of geometric cues of the sort conveyed by line drawings, rather than relying on analytic photometry and detailed lighting models. Our interest in the three-dimensional interpretation of line drawings stems from our belief that human perception of grey-level imagery relies heavily on geometric cues of just the sort line drawings capture.

Three key components of a computational theory for line drawing interpretation are:

(1) Line Classification--ascertaining the type of physical boundary each line represents.

(2) Line Interpretation--determining the three-dimensional space curves that correspond to surface contours and to the surface normals along extremal boundaries.

(3) Surface Interpolation--reconstructing smooth surfaces consistent with the boundary conditions established by line interpretation.

We have made significant progress in each of these areas, at the level of implementation as well as that of computational theory.

A model for three-dimensional line interpretation and surface orientation has been refined and implemented [1]. The model recovers the three-dimensional conformation of image boundaries by optimizing a smoothness metric, then takes the reconstructed space curve as a boundary condition for surface interpolation. This technique was applied to various boundary curves and simple test surfaces for which its results were in reasonable accord with human perception.

A theory for the recovery of surface shape from surface-marking geometry, developed by the author of this report at M.I.T [2], was refined and extended by him while working at SRI International [3].

A new intensity-based approach to the classification of edges was developed and implemented. Using basic properties of scenes and images, signatures were deduced for each of several edge types, expressed in terms of correlational properties of the image intensities in the vicinity of the edge. A computer program was developed that evaluates image edges as compared with these prototype signatures. It was shown to discriminate extremal boundaries effectively from cast shadow boundaries in cases where the traditional junction cues were absent from the image. We believe that this technique may also be directly applicable to the detection and correction of some common flaws in satellite imagery.

In addition, a major survey and synthesis of work in computational vision was prepared, in part with NASA support, and has appeared in IEEE Proceedings [4].

II INTERPRETING LINE DRAWINGS AS THREE-DIMENSIONAL SURFACES

The fundamental problem in interpreting line drawings is the massive ambiguity of the two-dimensional stimulus. Any line admits, in principle, an infinity of interpretations as a three-dimensional boundary. Even given the reconstructed boundaries, an infinity of surfaces may, in principle, be interpolated between them. Resolving this ambiguity requires the application of constraints that are powerful enough to determine unique solutions--or, at most, a small set of alternative solutions--and, at the same time, are able to yield at least qualitatively correct interpretations for most natural scenes. In addition to these computational requirements, three-dimensional perception by humans of line drawings provides important insights into the problem.

To recover the three-dimensional conformation of a surface discontinuity boundary from its image, we invoke two assumptions: smoothness and general position. The smoothness assumption implies that the space curve bounding a smooth surface will also be smooth. The assumption that the scene is viewed from a general position implies that a smooth curve in the image results from a smooth curve in the scene. The residual problem is to determine which smooth space curve is most likely. For the special case of a wire curved in space, we conjectured that, of all projectively equivalent space curves, humans perceive the one with the most uniform curvature and the least torsion [5]; they also perceive the space curve that is smoothest and most planar. Boundary conformation was determined by minimizing an integral measure of curvature and torsion within the projective constraints.

Given constraints on orientation along extremal and discontinuity boundaries, the next task is to interpolate smooth surfaces consistent with these boundary conditions. The problem of surface interpolation is not peculiar to contour interpretation; it is fundamental to surface reconstruction since data are generally not available at every point in the image. We have implemented a solution for an important case: the interpolation of approximately uniformly curved surfaces from initial orientation values and constraints on orientation. Our approach to the interpolation computation utilizes an observation that the components of the unit normal vector normally vary linearly across the images of surfaces of uniform curvature. The interpolation process was applied to several test cases for which essentially exact reconstructions were obtained, even when boundary values were extremely sparse or only partially constrained. For a full description of the reconstruction model and complete references, see Reference 1.

III RECOVERING SURFACE SHAPE AND ORIENTATION FROM TEXTURE

As with boundary reconstruction, the problem of inferring shape from the geometry of projected surface markings is fundamentally one of ambiguity; given only the constraints imposed by projective geometry, an infinity of solutions is theoretically possible. To achieve a unique solution, additional valid constraints must be imposed.

The inference of surface shape was treated as a problem of statistical estimation, combining constraints from projective geometry with simple statistical models of the processes by which surface markings are formed. The distortion imposed by projection was treated, quite literally, as a signal, the shapes themselves as noise. Both the signal and the noise contribute to the geometry of the image; statistical models of the noise permit the projective component to be isolated.

Texture geometry was described in terms of the distribution of tangent directions. With no prior knowledge of the expected distribution, it is natural to assume that all tangent directions on the surface are equally likely. Together with geometric constraints, this simple statistical model defines a probability density function for surface orientation, given a texture observed in the image.

Intuitively, the image texture is inversely projected onto that plane that yields the most uniform reconstructed texture. Curved surfaces are recovered by applying this planar technique to local regions.

This estimator was applied to geographic contours and surface markings extracted from natural images. It proved an effective estimator of surface shape and orientation, both by objective criteria, and by its concurrence of agreement with human perception.

The original formulation and implementation of the model were developed by A. Witkin at M.I.T. However, the statistical basis of the model was refined and expanded under NASA support at SRI, where a more extensive report was compiled for publication [3]. For a full description of the method, along with complete references, see Reference 3.

IV LINE CLASSIFICATION

Of central importance to the interpretation of line drawings is line sorting--the classification of the various lines according to the type of surface boundary they represent (i.e., extremal boundaries, shadow edges, surface orientation discontinuities, texture edges, and the like). Because each type imposes different constraints on three-dimensional interpretation, misclassification can lead to serious errors of interpretation. Line classification has been undertaken within the line drawing domain in relation to junction constraints, such global structural cues as parallelism and symmetry, and global optimization criteria for three-dimensional interpretation. The difficulty with performing a sorting operation on idealized line drawings is that all the lines in a drawing look fundamentally alike. The application of junction constraints requires perfect line data, and global cues such as symmetry are often inapplicable.

An alternative approach to line sorting is to return to the original image data, utilizing the intensity and spectral information in the vicinity of the edge. Horn [6] has suggested that the intensity profiles across edges (such as peak versus step) may provide signatures for some edge types. However, that technique has never been shown to work for complex imagery. This section reports an intensity-based, line-sorting technique that distinguishes line types by statistical comparison of the intensity variations along opposite sides of the edge. We have focused on two line types--extremal edges and cast-shadow boundaries--but extensions to other edge types have also been explored. We examined the possibility that the same technique might be directly applicable to the detection and correction of some flaws in remotely sensed imagery, such as variations in sensor gain and bad scan lines.

IV.1 DEFINING THE PROBLEM

Because line types are defined in terms of the scene events they denote, any method for line sorting must provide some basis for discriminating those events by their appearance in the image. We therefore begin by characterizing the distinctive properties of extremal boundaries and cast-shadow edges, and define the computational problem of identifying those edges.

Extremal Boundaries--Projective mapping from image to scene tends to be continuous because physical surfaces tend to be continuous. Almost everywhere in a typical image, therefore, nearby points in the image correspond to nearby points in the scene. This adjacency is maintained regardless of any change in viewpoint or scene configuration that does not actually sever the connected surfaces of which the scene is composed. The distinguishing property of extremal boundaries (which can be defined as discontinuities in the projective mapping) is their systematic violation of this rule: the apparent juxtaposition of two surfaces across an extremal edge represents no fixed property of either surface, but is subject to the vagaries of viewpoint and scene configuration. For example, if you position your finger to coincide with a particular feature on the wall or outside the window, a small change in the position of your head or hand can drastically affect their apparent relationship. Because the false appearance of proximity is the hallmark of extremal edges, the problem in identifying those edges is to distinguish in the image the actual proximity of nearby points on connected surfaces in the image from the pseudoproximity imposed by projection.

Cast Shadows--Cast shadows in outdoor scenes represent transitions from direct to scattered illumination that are caused by the interposition of an occluding body between the Sun and the viewed surface. The problem in identifying cast shadows is to distinguish these transitions in incident illumination from changes, for example, in albedo or surface orientation. This kind of discrimination presents a problem because the effects of all these parameters are confounded in the image data--a change in image brightness can result from a change in albedo or surface orientation, as well as from incident illumination. Because the interrelationship of illumination, reflectivity, orientation, and image

> irradiance is well known, the presence of shadows in an image could be readily detected if a constant reference pattern could be placed in the scene. As the apparent brightness of a constant pattern then varied with location, the change in brightness would by elimination have to be attributed to a change in illumination. Of course, such active intervention is generally impractical. Nevertheless, our goal is to achieve the effect of viewing a constant pattern across the shadow edge without actually placing such a pattern in the scene. This could be done if some fixed relationship were known to exist between the surface strips on either side of the shadow edge.

In short, extremal boundaries are curves across which distant points in space are placed in apparent juxtaposition by projection, thus violating the continuity of the projective mapping that is valid over most of the image. Therefore, the identification of extremal boundaries requires that actual proximity be distinguished from the pseudoproximity imposed by projection. Cast-shadow edges are contours across which the pattern of surface reflectance has been systematically transformed by an abrupt change in illumination. To identify cast-shadow edges, the effects of illumination must be distinguished from those of albedo and surface orientation—as if a constant reference pattern had been placed across the edge.

IV.2 COMPUTATIONAL THEORY

Our solution to line drawing interpretation rests on the simple principle that coherence in the image denotes real coherence in the scene, rather than a structural coincidence and a fortuitous alignment of distinct scene constituents. We measure coherence in the neighborhood of an edge by performing a normalized correlation on intensity values at corresponding points across the edge.

A high correlation implies that the edge and its vicinity correspond to a connected surface strip. Therefore, the edge is not an extremal boundary; furthermore, the regions on either side can be regarded as instances of a (statistically) constant pattern. In that case, the presence of a shadow can be detected by constructing a regression equation whose parameters signal any systematic distortion of the pattern across the edge. Ideally, this distortion is linear, but nonlinearites are introduced in practice by complex lighting effects, film or sensor response, and so forth.

A low correlation does not necessarily signal an extremal boundary, but could reflect low contrast or fragmented surface structure. However, a larger neighborhood of the image can be explored to establish a baseline for the surface properties. The given edge is consequently embedded in a family of parallel curves and a sequence of regressions performed from one curve to the next. In terms of this regression sequence, the various edge types display distinctive "signatures" that can be computed from the image data; extremal boundaries display a sharp downward spike in correlation where the fabric of the projective mapping is torn by the boundary. Cast-shadow boundaries display sustained high correlations, but exhibit abrupt spikes in the regression parameters, where the surface structure is systematically transformed by the illumination transition. A low correlation throughout implies that either the contrast is too low or the surface structure too fragmented for any positive conclusion to be drawn.

The motivation for this strategy follows from some simple observations regarding the character of natural scenes and images. First, as mentioned above, it follows from the fact that surfaces tend to be continuous that nearby points in the image usually correspond to nearby points in the scene (i.e., the projective mapping, as a rule, is continuous). Second, because the structure of surfaces tends to be coherent, such properties as reflectance and orientation at a given point on a connected surface are (statistically) good predictors of the properties at nearby points. Third, because scenes are made up of distinct objects whose structures and spatial configuration are governed by extremely complex factors, the properties of widely separated surface points, or of points on surfaces of distinct objects, can usually be regarded as unrelated and independent.

Because of these three principles--surface continuity, coherence, and independence--we can expect intensity values at nearby image points to be highly correlated. (That this is true for most images is easily verified.) In other words, a small step in the image usually corresponds to a small step on some connected surface, so that surface coherence imposes a statistical relation on the properties of nearby points. Thus, when we place in correspondence with one another, the points on either side of an arbitrary image curve, we should often expect to see a high correlation between the intensity values at those points. However, when that small step happens to cross an extremal boundary, the corresponding surface points generally belong to distinct objects, and may be widely separated in space. In that case, the properties of the points are independent. Thus, when the points on either side of an extremal boundary are placed in correspondence, we would never observe a high correlation unless the surfaces meeting at the boundary possessed identical structure, and, in addition, happen to lie in perfect alignment from the observer's viewpoint. The likelihood of this ideal confluence of curcumstances is vanishingly small.

Thus, we may confidently conclude that coherence of structure across an image curve (as measured by correlation) denotes true coherence of scene structure rather than an accidental phenomenon of scene configuration.

IV.3 IMPLEMENTATION

Our implementation assumes that an edge has been located by edge-finding techniques. In practice, edges were traced by hand, although zero-crossing edges were tried as inputs. We construct a parallel family of curves around the edge by imposing a new coordinate system on the image as follows. Arc length on the edge is taken as the y-coordinate, and orthogonal distance from the edge (right-handed) as the x-coordinate. This technique amounts to coercing a strip around the edge into a rectangular region whose central column corresponds to the original edge. The surrounding columns correspond to parallel curves using bilinear interpolation of intensity values, to reduce the artifacts of quantization. Figure 2 shows an image with an edge superimposed and the corresponding rectified strip image.

Once the rectified strip is constructed, a sequence of linear regressions is performed between columns. To avoid the imposition of spurious correlation by the imaging and digitizing process, regressions are computed between the ith column and the (i + 2)th. The outcome of this computation is normalized correlation, an additive regression term, and a multiplicative regression term, each a function of column position. The midpoints of these plots represent the regression across the original edge. See Figure 1 for idealized plots.

No attempt has yet been made to classify the edge-type signatures automatically; however, the computation was performed on a number of edges in both aerial and ground imagery. The results (see Figures 2 through 6 for examples) show that contour types can in many cases be clearly distinguished. Let us also recall that the overall correlation level provides the basis for a confidence measure, permitting a reduction in the mislabeling rate--but at the cost of increased conservatism.

IV.4 SUMMARY

The line-sorting method presented herein was derived from the basic properties of visual scenes. It shows promise as a useful technique, particularly in connection with established line junction techniques. Moreover, the method demonstrates the potential benefit of interplay between line-drawing and raw-image levels of representation.

We also considered the intriguing possibility that this technique might be directly applicable to the detection and correction of certain sensing and transmission flaws in satellite imagery:

* Stripes due to variation in sensor gain bear a formal resemblance to shadows, in that the underlying data are systematically transformed by a more or less constant gain factor. A line-to-line regression could prove useful in detecting these errors automatically, and the regression equation might be used to correct the data. We tried this out successfully on synthetic stripes generated by imposing a constant gain factor on some scan lines.

* Bad scan lines formally resemble occluding contours in that they are uncorrelated with adjacent lines. The line-to-

line regression technique may be useful for their automatic detection. Conservative automatic correction by interpolation might also be accomplished if the lines on either side of the bad line are highly correlated; it is then probably safe to replace the bad line by interpolation.

REFERENCES

1. H. G. Barrow and J. M. Tenenbaum, "Interpreting Line Drawings as Three-Dimensional Surfaces," Artificial Intelligence 17, No. AI256, pp. 75--116 (1981).

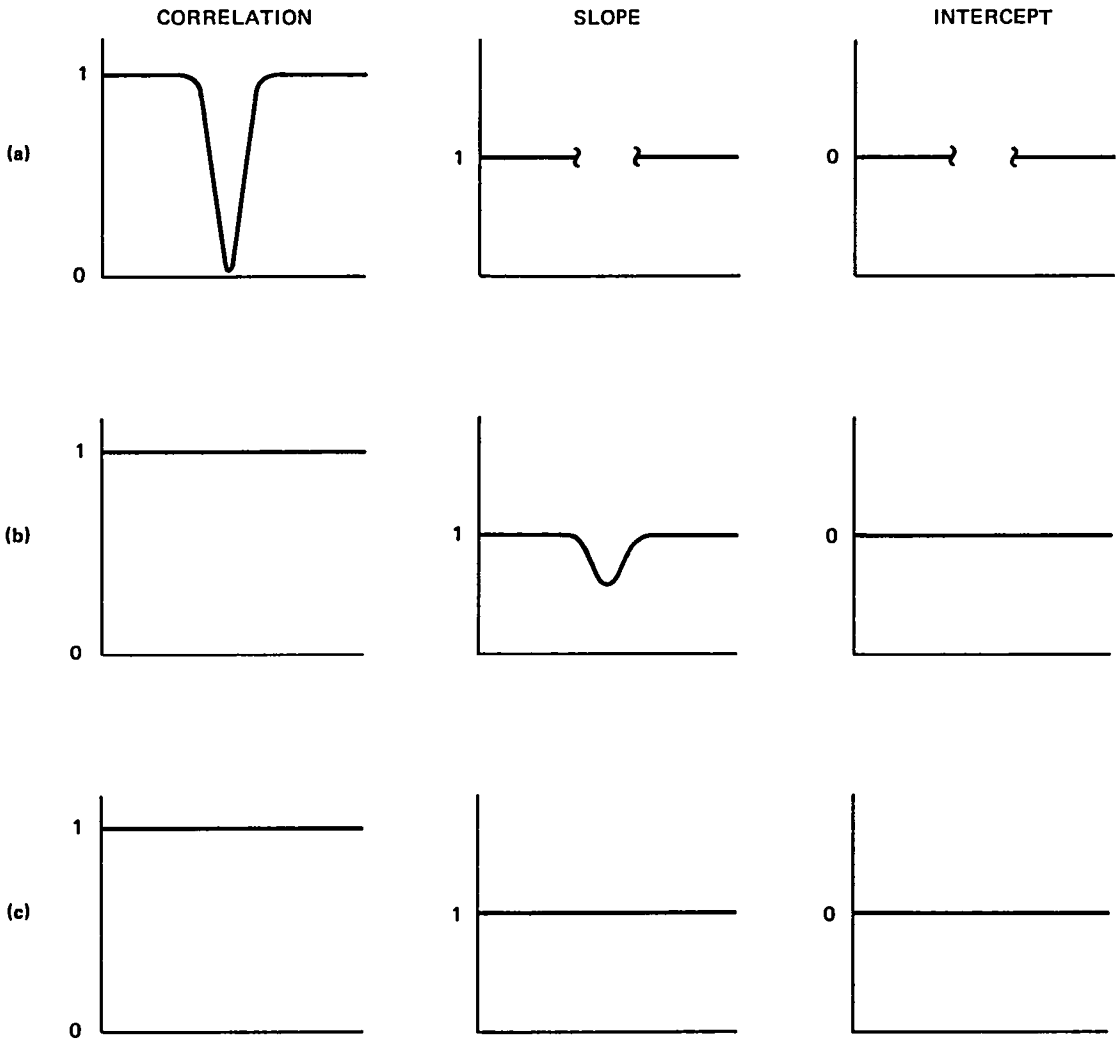
2. A. P. Witkin, "Shape from Contour," Artificial Intelligence Laboratory Technical Report 589, Massachusetts Institute of Technology, Cambridge, Massachusetts (November 1980).

3. A. P. Witkin, "Recovering Surface Shape and Orientation from Texture," Artificial Intelligence 17, No. AI258, pp. 17-45 (1981).

4. H. G. Barrow and J. M. Tenenbaum, "Computational Vision," Proceedings of the IEEE, Vol. 69, No. 5, pp. 572-595 (May 1981).

5. H. G. Barrow and J. M. Tenenbaum, "Recovering Intrinsic Scene Characteristics from Images," in Computer Vision Systems, A. Hanson and E. Riseman eds., pp. 3-26 (Academic Press, New York, New York, 1978).

6. B. K. P. Horn, "Understanding Image Intensities," Artificial Intelligence 21, pp. 201-231 (1977).



(a) **Extremal Boundary** — notch in correlation across the edge Slope and intercept in the low-correlation area are meaningless

(b) **Cast Shadow** — sustained high correlation across the edge, with disturbance of one or both regression parameters The nature of this disturbance depends on the sense of the edge (i e whether the shadow lies on the left or right), and on details of the imaging and digitizing process In practice, nonlinearities perturb the correlation slightly

(c) **No Edge Present** — sustained high correlation, no disturbance in regression parameters

FIGURE 1 IDEALIZED REGRESSION PLOTS

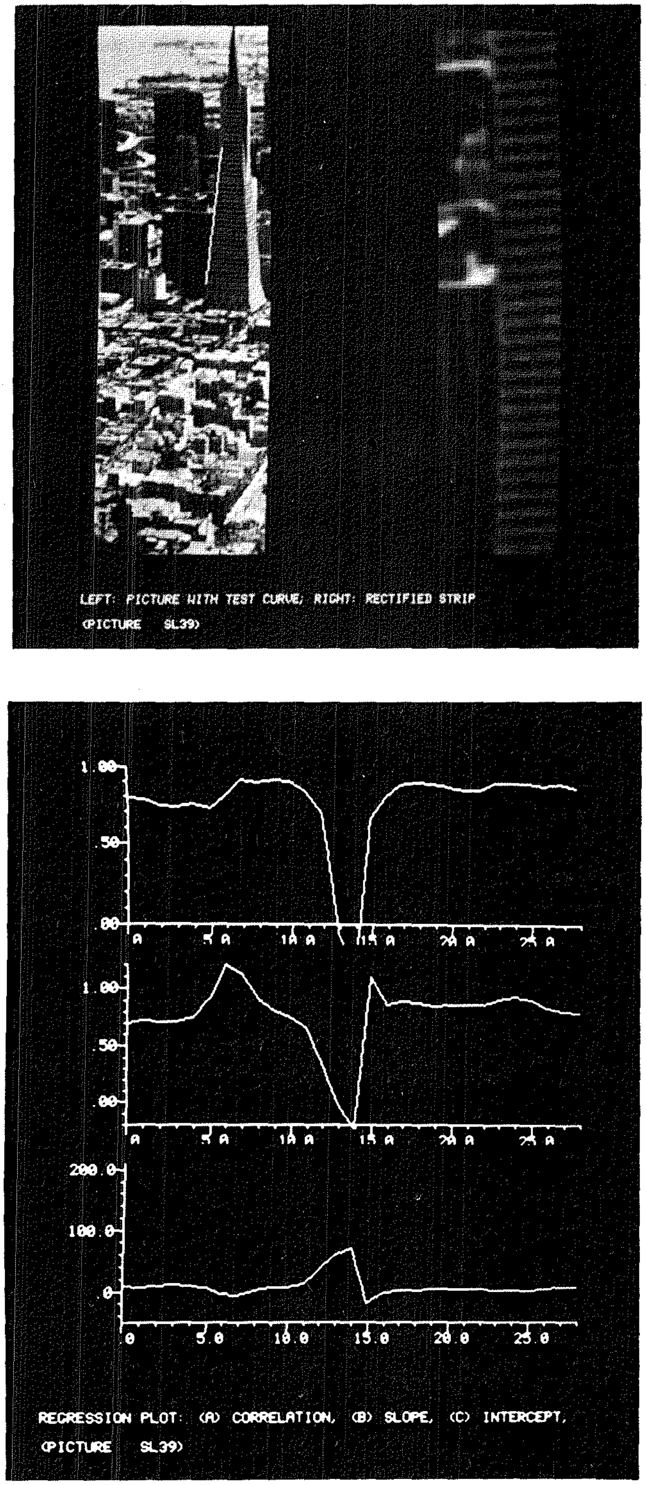
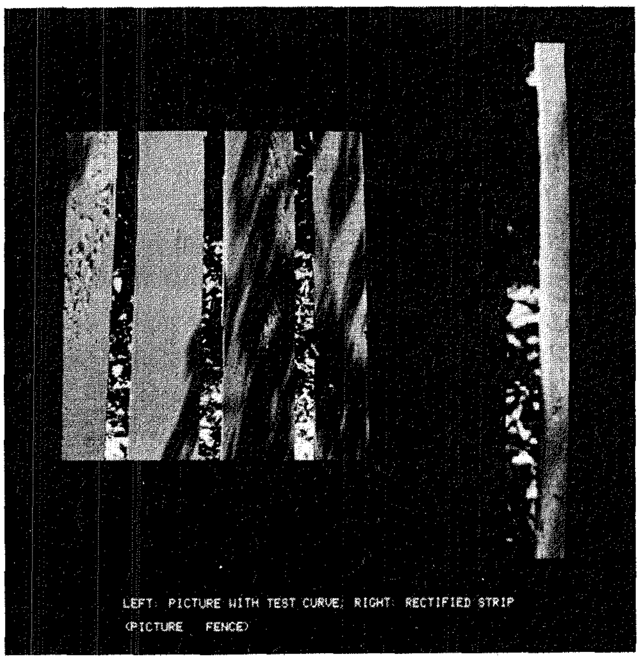


FIGURE 2 EXAMPLE OF EXTREMAL EDGE



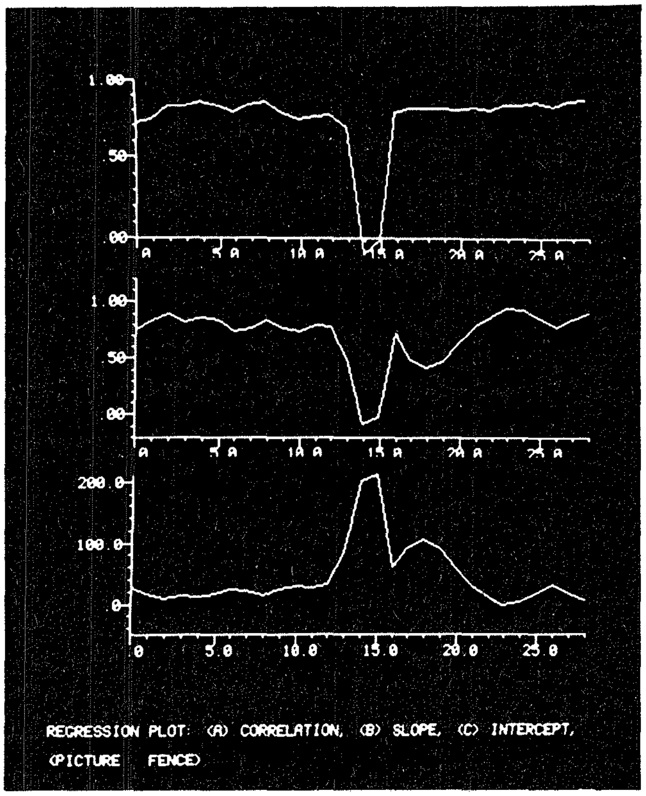


FIGURE 3 EXAMPLE OF EXTREMAL EDGE

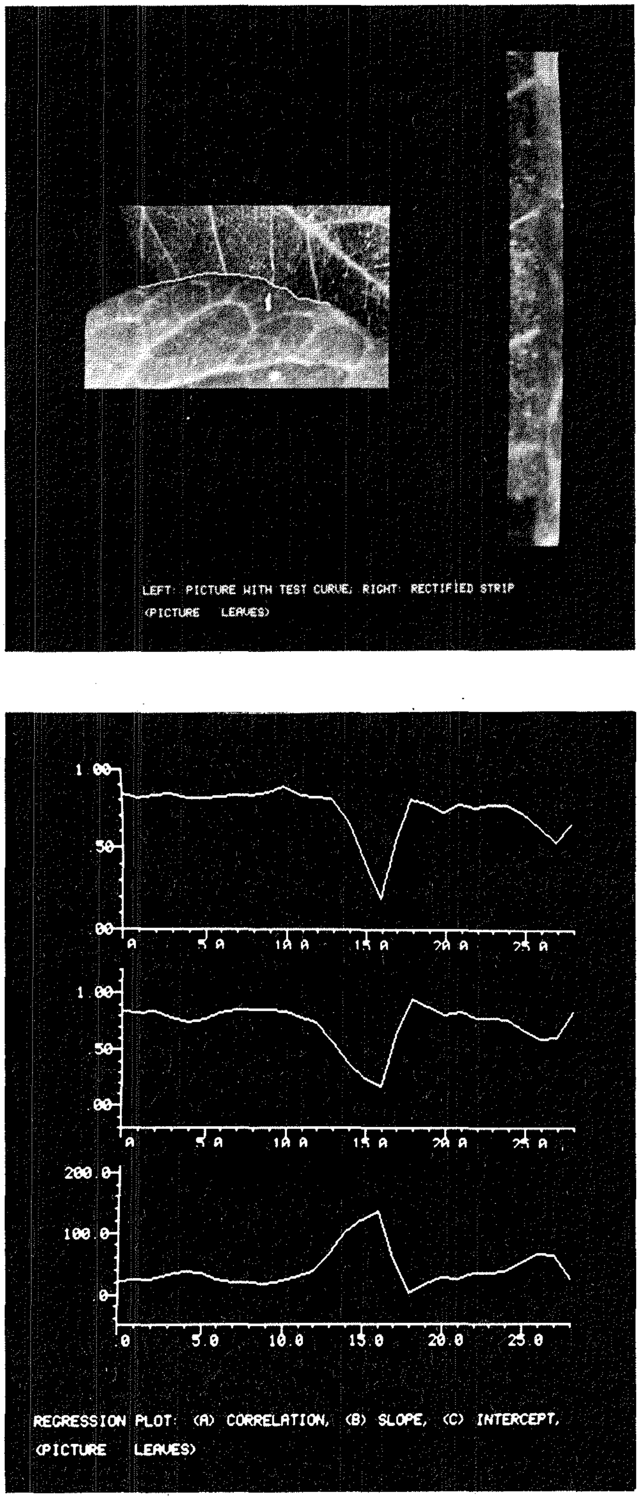


FIGURE 4 EXAMPLE OF LOW-CONTRAST EXTREMAL EDGE

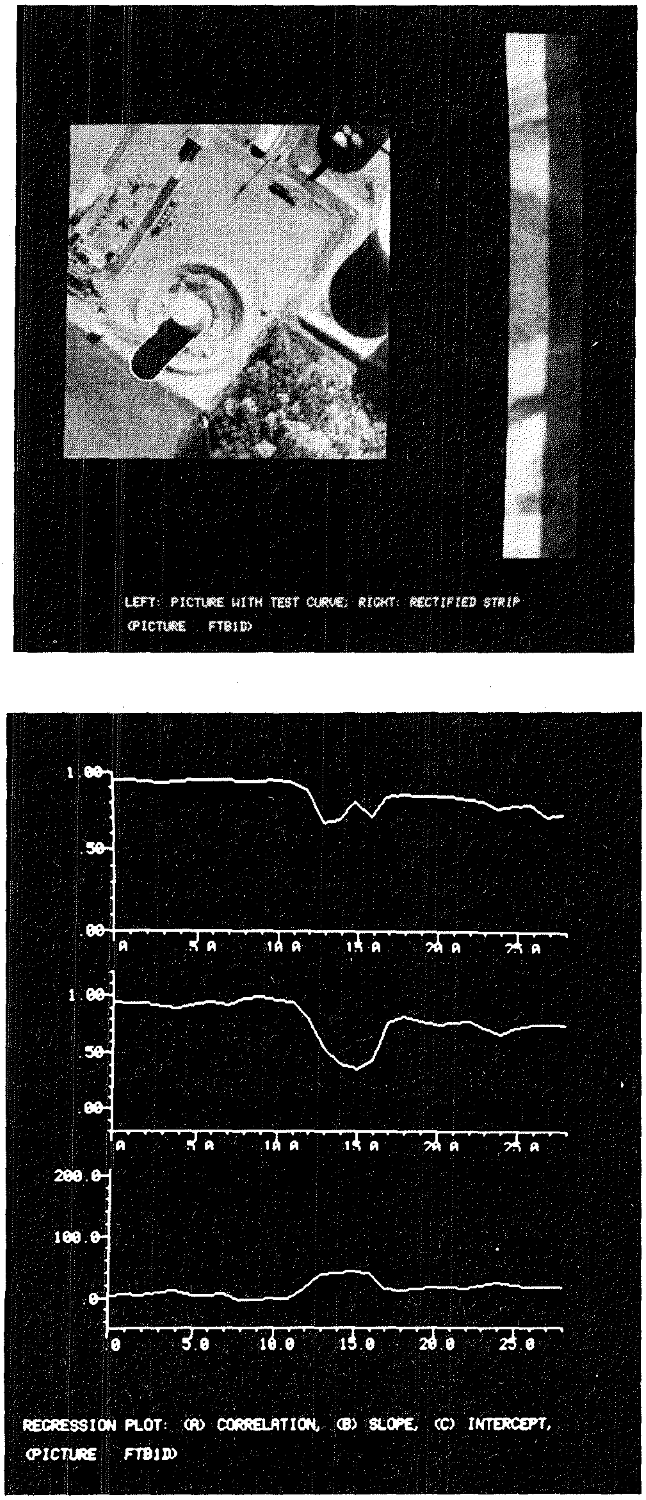


FIGURE 5 EXAMPLE OF CAST-SHADOW EDGE

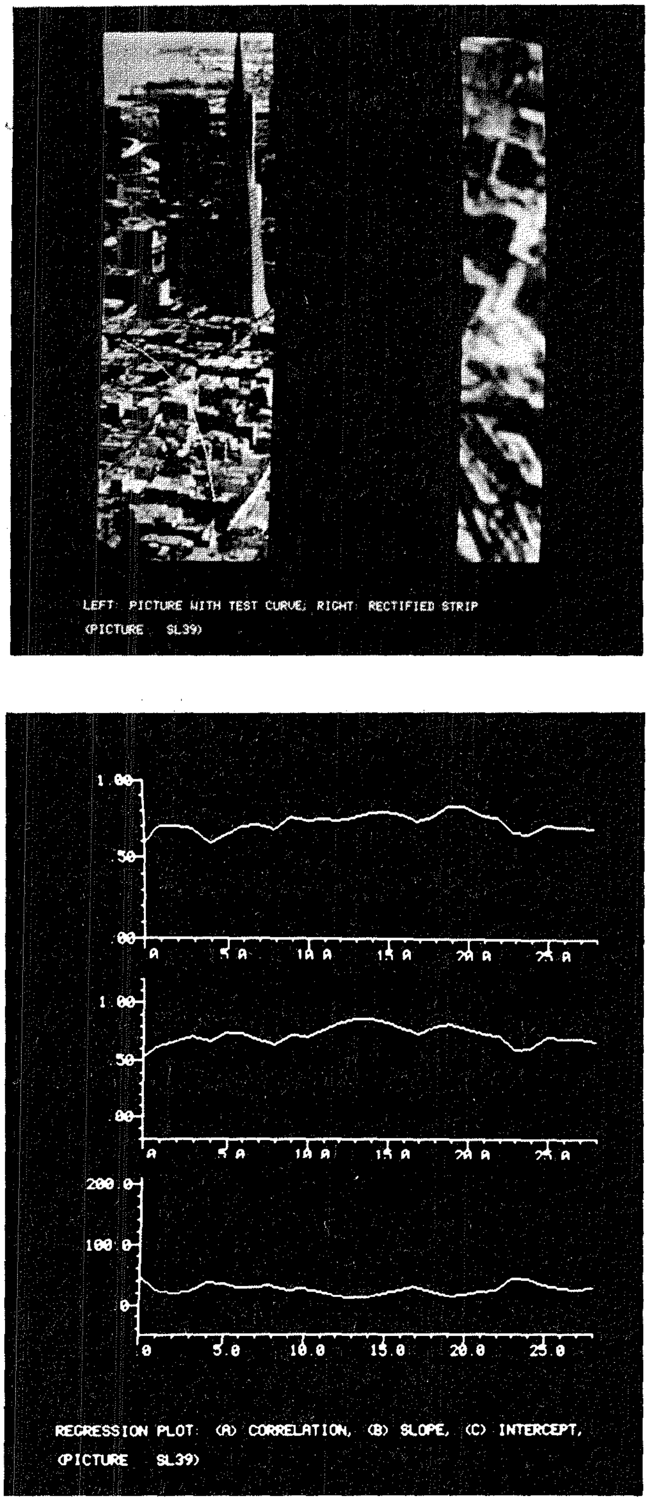


FIGURE 6 EXAMPLE OF REGRESSION PLOTS WHERE NO EDGE **IS PRESENT**

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR 165804		

4. Title and Subtitle	5. Report Date
COMPUTATIONAL THEORY OF LINE DRAWING INTERPRETATION	December 1981
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Andrew P. Witkin	
9. Performing Organization Name and Address	10. Work Unit No.
SRI International 333 Ravenswood Avenue Menlo Park, California 94025	11. Contract or Grant No. NAS1-16282
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
National Aeronautics and Space Administration Langley Research Center Hampton, Virginia 23665	Contractor Report
	14. Sponsoring Agency Code

15. Supplementary Notes

Langley technical representative: Dr. L. Keith Barker
Langley new technology representative: Mr. John Samos
Final Report

16. Abstract

Our research has focused on the recovery of the three-dimensional structure of visible surfaces depicted in an image by emphasizing the role of geometric cues present in line drawings, rather than by relying on analytic photometry and detailed lighting models. Three key components of our approach are line classification, line interpretation, and surface interpolation.

A model for three-dimensional line interpretation and surface orientation has been refined and implemented, and a theory for the recovery of surface shape from surface-marking geometry was developed.

A new approach to the classification of edges was developed and implemented--signatures were deduced for each of several edge types, expressed in terms of correlational properties of the image intensities in the vicinity of the edge. A computer program was developed that evaluates image edges as compared with these prototype signatures. We believe that this technique may also be directly applicable to the detection and correction of some common flaws in satellite imagery.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
	Unclassified-Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	30	

End of Document