USCIPI REPORT #720

Semiannual Technical Report

by

Harry C. Andrews

September 1976

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
Powell Hall of Engineering
University Park/MC-0272
Los Angeles, CA 90089 U.S.A.

ABSTRACT

This technical report summarizes the image understanding and image processing research activities performed by the Image Processing Institute at the University of Southern California during the period of 1 April 1976 to September 30, 1976 under contract number F-33615-76-C-1203 with the Advanced Research Projects Agency Information Processing Techniques Office.

The research program has as its primary purpose the development of techniques and systems for processing, transmitting, and analyzing images and two dimensional data arrays. Four tasks are reported: Image Understanding Projects; Image Processing Projects; Smart Sensor Projects; Recent Ph. D. Dissertations. The Image Understanding project has concentrated on edge detectors, image segmentors, clustering techniques and higher level symbolic descriptors from segmented imagery. The Image Processing projects have concentrated on a posteriori restoration, degree of freedom analyses of imaging systems, phase coding for optical image processing and artificial stereo algorithms. The Smart Sensor projects include the development of adaptive CCD image processing chips and circuitry. The Ph. D. dissertations completed in this time period are abstracted in the next task.

Table of Contents

1. RESEARCH PROJECT OVERVIEW

This report describes the progress and results of the University of Southern California image understanding and image processing research study for the period 1 April 1976 to 30 September 1976. The research study has been subdivided into five projects:

Image Understanding Projects

Image Processing Projects

Smart Sensors Projects

Recent Ph.D. Dissertations

Publications

The image understanding projects involve research directed toward the goal of developing generalized processing systems capable of analyzing images and extracting salient information. Specific studies include feature extraction, symbolic description, interpretation, clustering techniques, edge detectors, segmentors, and systems analysis. The image processing projects include research on novel image coding techniques based upon results of the image understanding study, advanced image restoration methods, including a posteriori phase computation techniques, degree of freedom, modelling in support of image understanding, and studies of phase coding for optical filtering techniques and artificial stereo algorithms. The smart sensors projects comprise investigations of electronic and optical processing methods which can be integrated with imaging sensors to perform low level image enhancement and feature extraction within the sensor. An adaptive CCD processor has been developed for first stage image

2. IMAGE UNDERSTANDING PROJECTS

The results of the image understanding projects described below are directed toward the eventual development of processing systems for generalized image analysis. Applications of such systems include photointerpretation, vehicle guidance, visual tracking, cartography, and many other areas of military interest.

Our research study is organized according to the structure of a conceptual image understanding system. This system consists of a feature extraction stage which detects and measures primitive features such as edges and texture regions from an input image. Next the primitive features are clustered into meaningful symbols such as object boundaries, segments of some common attribute, or basic shape structures. Such clustering then is used to completely segment the scene of interest. Finally image symbols are interpreted in terms of their semantic relationships to produce a concise quantitative description of the original image. Some extra-image knowledge base is assumed available for guidance of all elements of the image understanding system.

The philosophic approach to the study has been to bring together a research team skilled in digital signal processing and concepts of artificial intelligence. Although the research effort at USC involves all stages of the overall image understanding system, emphasis has been placed on the feature extraction, clustering, and symbolic description subsystem elements. Rapid progress is being made in these areas to assist the research in the higher stages of the image

2.1 EVALUATION OF A SIMPLIFIED HUECKEL OPERATOR

Ramakant Nevatia

Introduction

An edge detector due to Hueckel <1,2> has been in wide use. Though no definitive studies have been made, its performance is considered to be superior to that of other edge detectors (e.g. see <3>, see <4> for a contrary conclusion). However, computation required for a Hueckel edge detector is considerably larger than for other types of operators. The central part of the Hueckel edge detector involves approximation of a two dimensional image intensity pattern by expansion into a truncated Fourier series. Mero and Vassy described a method of simplifying the approximation by using fewer bases functions for expansion <5>. In this report, we examine a generalization of this approach and effect on performance of the reduction in the number of bases functions. The main experimental conclusion is that such reduction in computation is achieved at a significant loss in performance, if the images in use are noisy.

Operation of the Hueckel Edge-Line Detector

The basic process of Hueckel edge operator consists of optimally fitting an ideal edge-line to the image intensity values in a small circular neighborhood. The ideal edge-line is determined by a 7-tuple of parameters, three parameters determining the brightness levels (b_-, t_- and t_+, as shown in figure 1) and the other four parameters determine the position, orientation and the width of the line. The

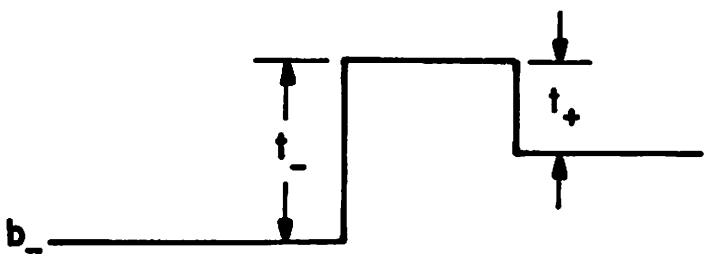


Figure 1. Intensity profile of an ideal edge-line.

fitting process consists of determining the values of the seven parameters for a best fit with the image intensities. Ideally a tuple of parameters is to be computed such that

$$N = \| I - S(\text{tuple}) \| \qquad (1)$$

is a minimum, where I is a vector representing image intensities and S is an ideal edge-line.

This minimization process is approximated by expansion of both the input image disk and the edge-line in an orthogonal Fourier series. Let a_i be the coefficients of expansion for the image and s_i be the coefficients for an ideal edge-line. The minimization of (1) is then approximated by choosing a tuple such that

$$N = \sum_{i=0}^{8} (a_i - s_i)^2 \qquad (2)$$

is minimized. Hueckel gives arguments for the optimality of the chosen series <2>. A crucial decision is to use only the first nine terms of expansion. Two reasons are given for the choice of this number:

(a) Higher order terms correspond to noise in the image and should be ignored.

(b) An analytical solution to the minimization problem is found using nine terms.

Decision as to the presence of an edge is based on the amplitude

The important question is the effect of ignoring some terms in the orthogonal expansion of the signal. It is clear, that if the signal indeed consists of an ideal edge element, the fewer terms of expansion are sufficient to characterize the signal exactly. Results of experimental evaluation signals containing noise are presented next.

Experimental Results

The input for these experiments consists of a picture with two vertical edges; the intensity profile is as illustrated in figure 2. A varying amount of random, Gaussian noise is added to this image. Results are presented for three images, figures 3a, 4a and 5a, with step size to noise variance ratios of: 10, 4 and 2.

The three images were processed by three versions of the edge detector:

1. Original Hueckel edge detector (operator 1)

2. Edge line detector using six coefficients (operator 2)

3. Edge detector using four coefficients (operator 3)

Determining the presence of an edge requires the use of a threshold on the computed step amplitude. Use of a higher threshold will allow fewer noise edges but also lose some of the desired edges. For the results presented, the thresholds were determined for the best subjective performance. The differences in the performance of three operators are so marked that a more careful control of thresholds

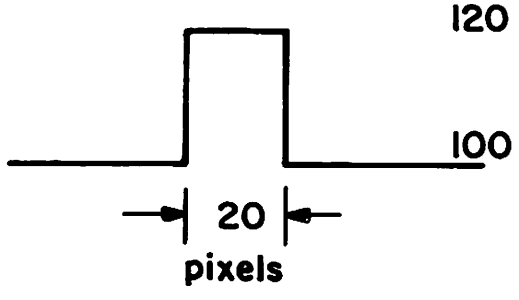


Figure 2. Intensity profile of a test step (without noise)

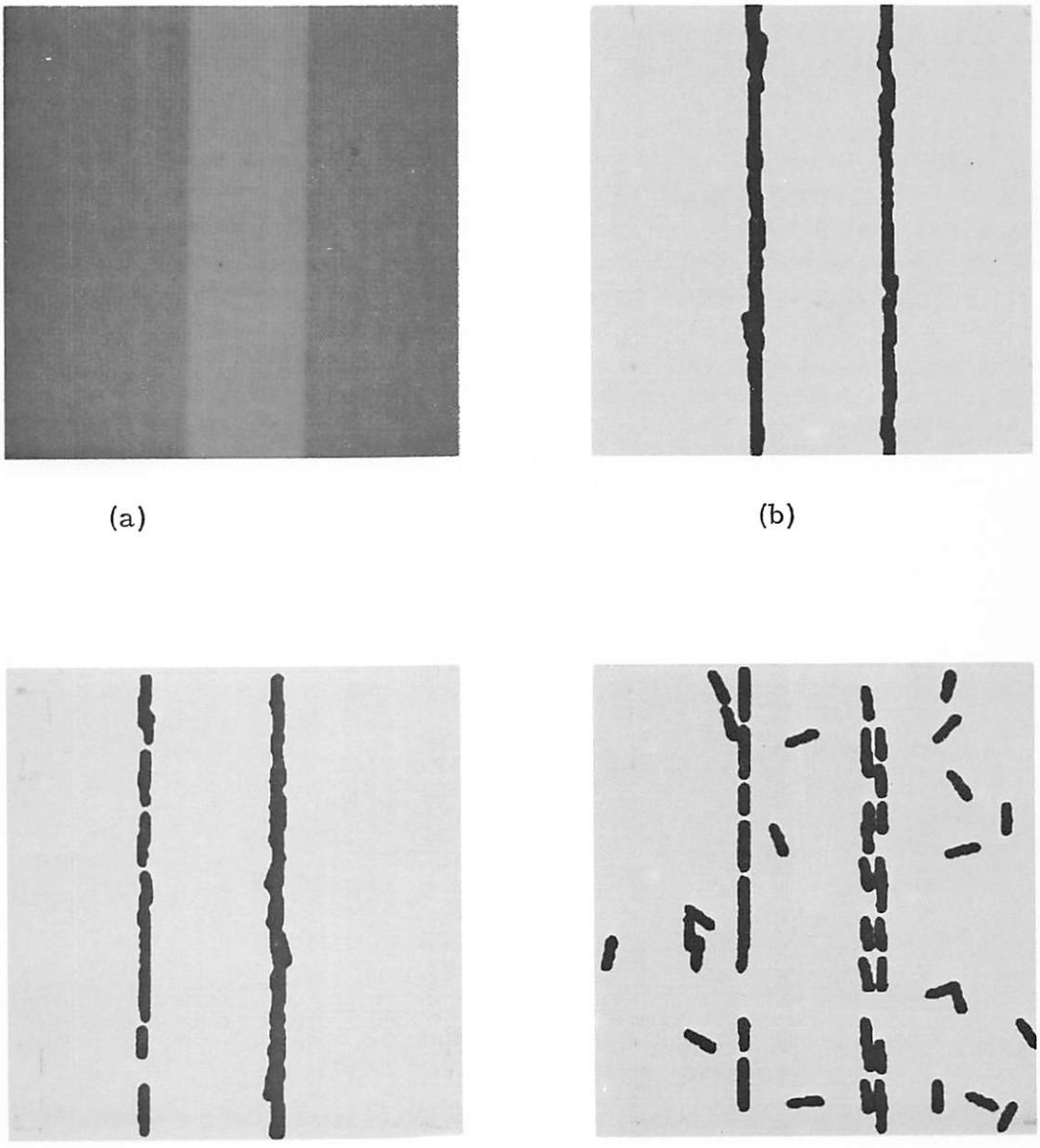
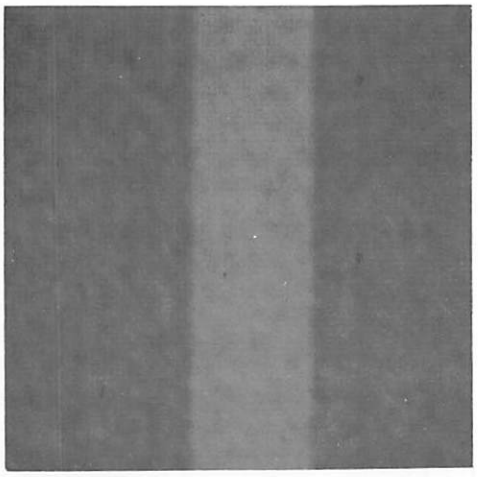
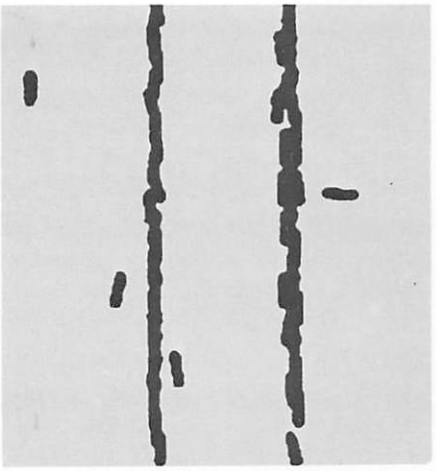


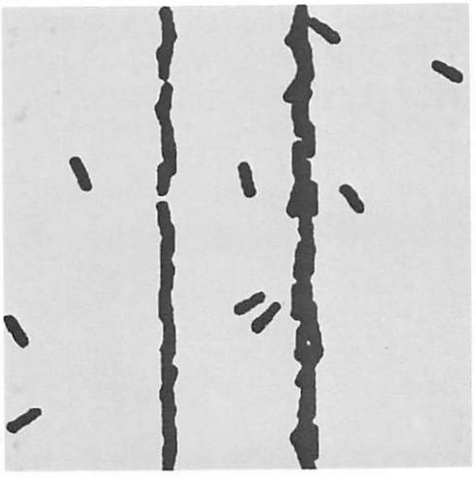
Figure 3. Low noise image and detected edges



(a)



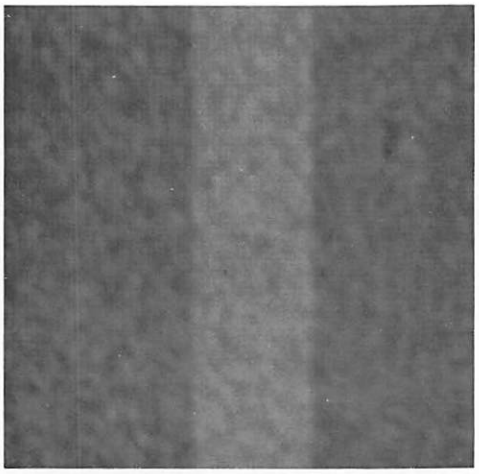
(b)



(c)

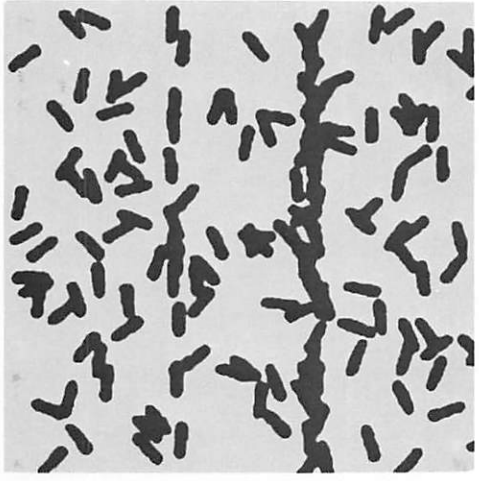
(d)

Figure 4. Medium noise image and detected edges

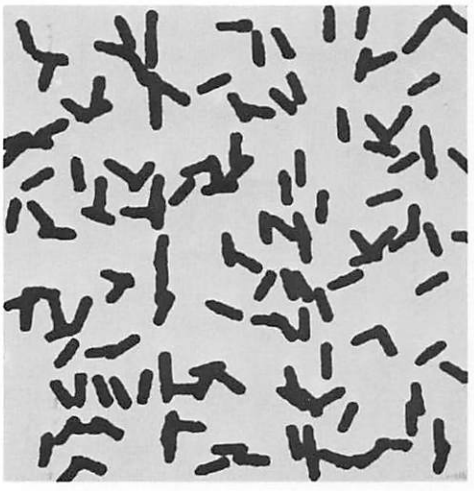


(a)

(b)



(c)



(d)

Figure 5. High noise image and detected edges

(e.g. by requiring the total number of detected edges to be the same) was not considered necessary.

Figures 3,4 and 5 show the results of the three operators on three images. (The edge detector was applied every three pixels along every third row. Each application looks at a neighborhood approximately circular and eight pixels in diameter. There is considerable overlap of regions in two neighboring applications, resulting in multiple edges for the same step. The phenomenon is pronounced in figures 3, 4 and 5 for the right edge only, because of accidental placements.)

It is clear that the performance of operator 1 declines gradually with the increase in added noise. Operators 2 and 3 on the other hand, perform reasonably well for high signal to noise ratio, but very poorly for images with high noise. Further, the performance of operator 2, using more coefficients of expansion, is superior to that of operator 3. (In absence of any noise, all three operators perform perfectly.)

Similar results are obtained for detection of line edges (i.e. when the width of step is small, say two or three pixels wide).

Conclusions

The results shown above are somewhat surprising. It is tempting to think that as the ideal step is completely determined by four parameters, four bases vectors can be found to span the space of ideal step elements. Unfortunately, the space spanned by the ideal steps is

$$s_4 = 2^{\frac{1}{2}}(\lambda_- r_- + \lambda_+ r_+)(c_x^2 - c_y^2) \quad (7)$$

$$s_5 = 8^{\frac{1}{2}}(\lambda_- r_- + \lambda_+ r_+)\, c_x c_y \quad (8)$$

$$s_6 = 5^{\frac{1}{2}}(\lambda_- r_-^2 + \lambda_+ r_+^2)\, c_x \quad (9)$$

$$s_7 = 5^{\frac{1}{2}}(\lambda_- r_-^2 + \lambda_+ r_+^2)\, c_y \quad (10)$$

$$s_8 = \lambda_- r_-(0.5 - 2.5r_-^2) + \lambda_+ r_+(0.5 - 2.5r_+^2) \quad (11)$$

where r_-, r_+ stand for the positions of brightness transitions, b_-, t_- and t_+ are as shown in figure 1, $\lambda_{\pm} = (3\pi)^{\frac{1}{2}} t_{\pm} (1-r_{\pm}^2)^2/4$, and c_x and c_y are the direction cosines of the edge orientation (see <1> for details). Consider two cases separately:

1. Step edges only:

   The ideal edge-line can be converted to a step by setting $t_+ = 0$ (hence $\lambda_+ = 0$). We are now interested in determining λ_-, r_-, c_x and c_y. We need use only eqs. (4), (5) and (6) and hence need only three coefficients.

2. Step and line edges:

   Six parameters need to be determined now and we need six equations. Several subsets of the above nine equations will suffice. In particular, eqs. (4) through (9) are sufficient. The solution can proceed by determining c_x and c_y from eqs. (5) and

Operators," in University of Southern California USCIPI Report 660, March 1976, 22-40.

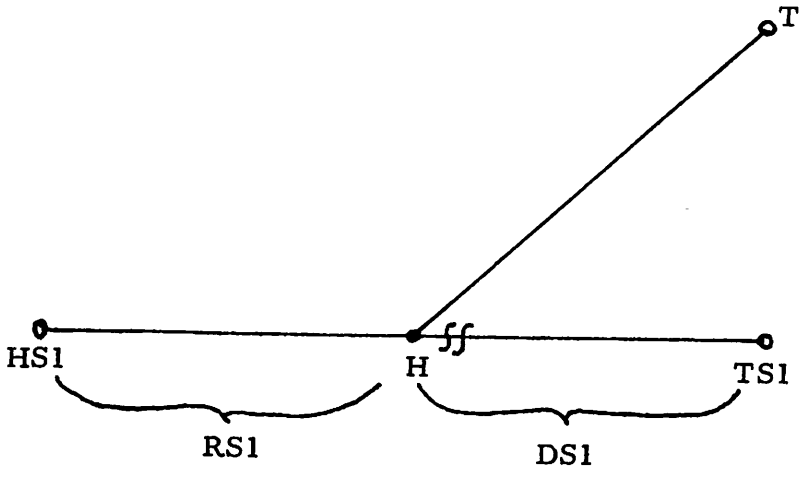
7. K. Paton, "Picture Description Using Legendre Polynomials," Computer Graphics and Image Processing, Vol. 4, No. 1, 1975, 40-54.

2.2 APPENDING OF LINEAR EDGE SEGMENTS WHICH CONTAIN COMMON EDGE POINTS

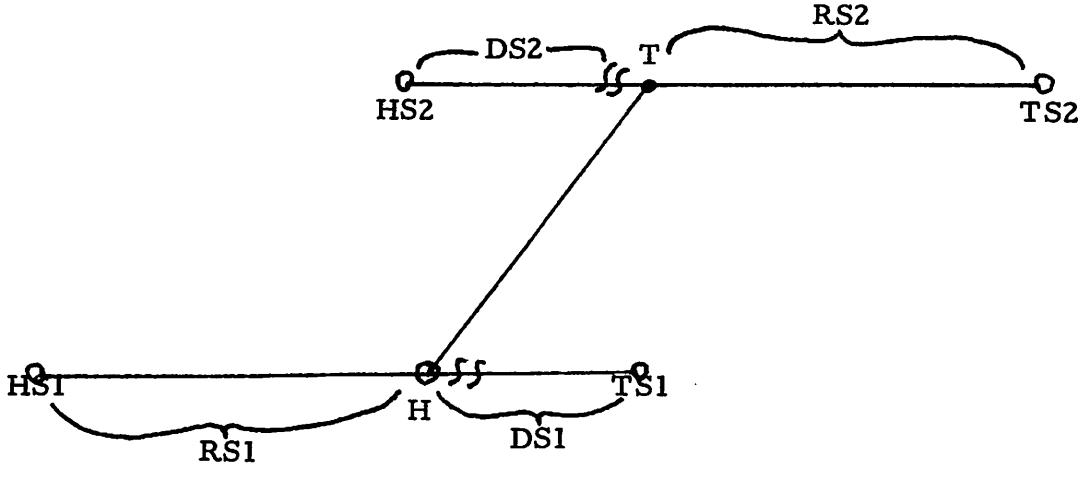
Lee W. Martin

In an image, there are generally sub-areas which stand out from the background. They may be classifiable as straight lines, curves, boundaries, etc. and they may be given the attributes of a feature, i.e. something that can be used in describing the image scene, using a set of image pixels as a single element. That element, or feature, has a more global information content than the sum of the pixels because of a recognizable interpixel relationship. This report considers the appending of linear segments to create longer, and possibly more identifiable edge segments to be used as features.

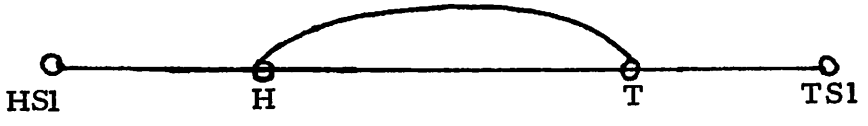
Elongated edge segments generally correspond to areas or objects of interest in a scene, e.g. partial boundaries of objects, identifiable segments along an edge (perhaps functions, or particular occurrence of an edge pattern). As such this feature would be useful for stereo analysis (small angle stereo) or scene segmentation. In the latter determine areas of the image where further (second level) processing would be most fruitful, e.g. determine segment continuability criteria determined by the feature segment. In small



(1a) Typical head appendage.



(1b) Typical connection, when H and T are in different PS's.



(1c) Typical case where H, T are in same PS. Then CS must be absorbed into S1.

Figure 1. Append Schematics

References

1. L.W. Martin and R. Nevatia, "Extraction of Prominent Features from Aerial Photographs," University of Southern California, Image Processing Institute Semiannual Technical Report 660, March 31, 1976, pp. 33-100.

2. M.H. Hueckel, "A Local Visual Operator which Recognizes Edges and Lines," Journal ACM, Vol. 20, No. 4, October 1973, pp. 634-647.

3. R. Nevatia, "Object Boundary Determination in Textured Environment," Proceedings of ACM 1975 Conference, Minneapolis, Minnesota, October 1975.

2.3 CONTRIBUTION OF EDGE ORIENTATION OF EDGE DETECTION

Peter Chuan

The purpose of this investigation is to find the class of edge detectors that could find edges with greater resolution. One of the properties that could be looked into is the orientation of edges in the image. The pictures we are most concerned with are pictures of natural images, most of which contain structured objects like trucks and vehicles on a natural background. Figure 1 shows the picture of a tank with a background of marsh. This background, for our purposes of extracting only the tank, is considered noise. However, it should be noticed that this very background which we condemn to be noise actually contains considerable information in itself. Therefore, any

Figure 1. Tank on a noisy Background

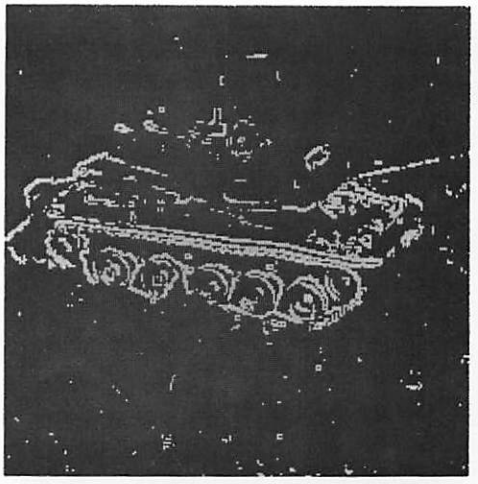


Figure 2. Kirsch thresholded to display 4000 most significant

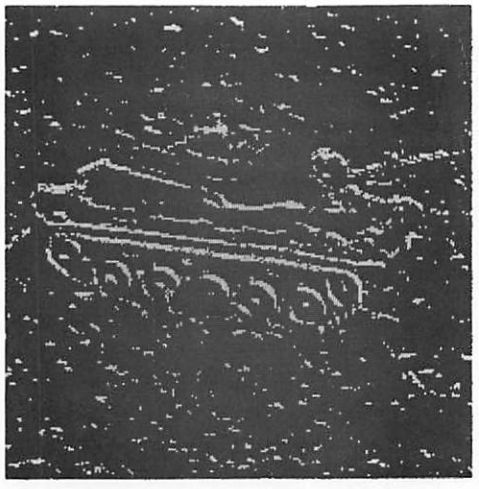


Figure 3. Kirsch with variance modifying function

Figure 4. Kirsch with continuity modifying function

Figures 5a, 5b, and 5c show the same three procedures applied to a noisy straight line edge and evaluated with Pratt's figure of merit <1>. The original picture of the line has an S/N ratio of five defined in the convention given by the above figure of merit.

The original Kirsch output has a figure of merit F = 68%. The Kirsch output implemented with directional variance measure has F = 51% and the Kirsch output implemented with continuity modifying function has F = 75%.

From figures 2, 3, and 4 it is obvious that the continuity and variance modifying functions create more background edges than the original edge magnitude map. However, figure 4 shows that more edges on the side of the tank were detected by the continuity modifying function while the other edges are more or less preserved. Figure 3 shows edges that have less random edge directions and therefore most detected edges appear as part of a chain of edge points. It is not exactly clear as to what extent the orientation information could contribute to detecting fine edges.

References

1. W.K. Pratt, "Figure of Merit for Edge Location," USCIPI Semiannual TR 660, 1976, pp. 85-93.

2. L.S. Davis, "A Survey of Edge Detection Techniques," Computer Graphics and Image Processing 4, 1975, pp 248-270.

3. G.S. Robinson, "Detection and Coding of Edges Using Directional

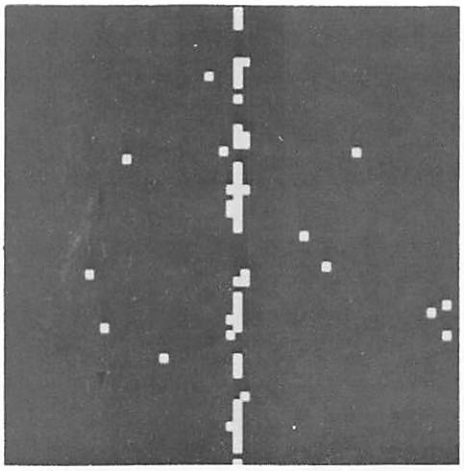


Figure 5a. Kirsch output
F = 68%

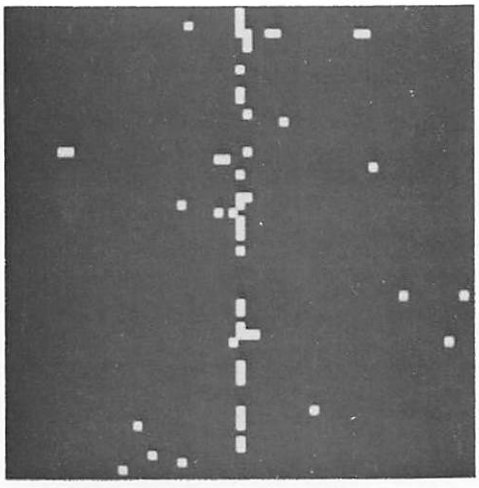


Figure 5b. Kirsch with variance modifying function
F = 0.51%

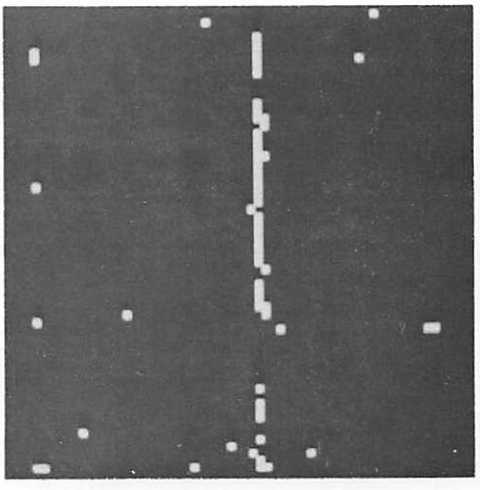


Figure 5c. Kirsch with continuity modifying function
F = 75%

Masks," USCIPI Semiannual TR 660, 1976, pp. 41-56.

2.4 SCENE SEGMENTATION BY CLUSTERING

Guy Coleman

The major portion of the effort during the period of this report has been directed at examination of scene features and their relationship to the overall clustering problem. The only intrinsic features of a pixel in a monochrome image are its position (x and y coordinates) and its brightness. All other features are of higher order, that is, they are based upon the relaticnship of the pixel to other pixels in the scene.

A number of different features were computed for overlapping blocks in an aerial photograph (see figure 1). As expected, the larger block sizes virtually destroy the scene detail. These features (see figures 2 through 6 for examples) were computed by averaging the value of the feature over the block.

A different approach was adopted for the tank scene (figure 7) and the armored personnel carrier scene (figure 8). In these scenes, features were computed for several block sizes centered on each pixel. While this approach requires more computation, the intrinsic resolution of the picture is maintained (see figures 9 through 22 for examples). Some of the features (mean, for example) still cause substantial blurring and loss of detail.

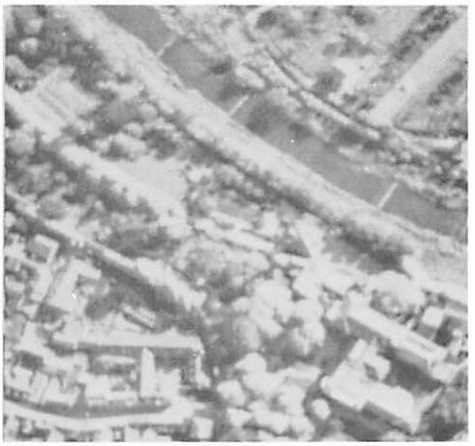


Figure 1. Aerial photograph original

Figure 2. Mean 8 x 8

Figure 3. Variance 8 x 8

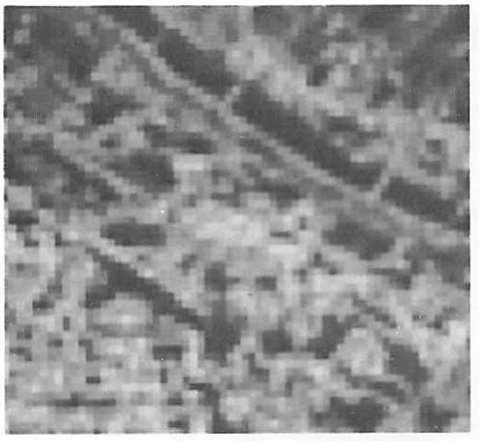


Figure 4. Texture 8 x 8

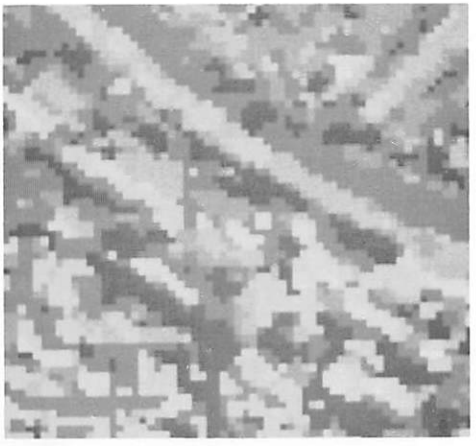


Figure 5. Mode 8 x 8

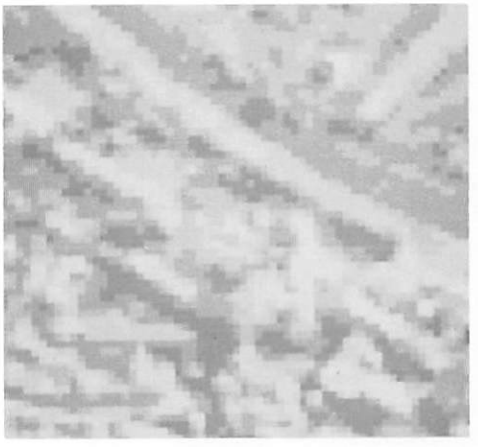


Figure 6. Median 8 x 8

Figure 7. Tank original

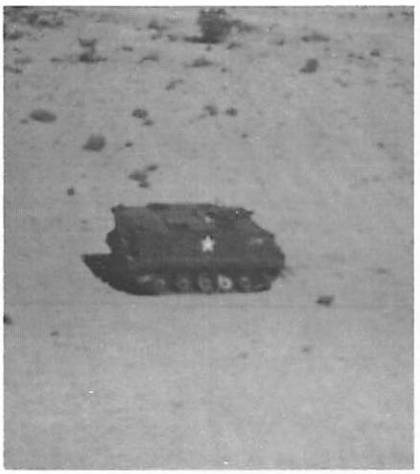


Figure 8. APC original

Figure 9. Tank mean 7 x 7

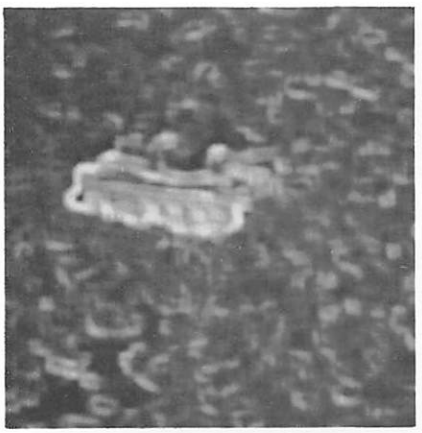


Figure 10. Tank variance

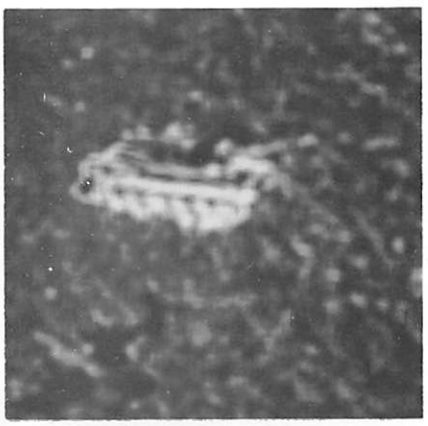


Figure 11. Tank texture 7 x 7

Figure 12. Tank mode 7 x 7

Figure 13. Tank mode 15 x 15

Figure 14. Tank median 7 x 7

Figure 15. Tank median 15 x 15

Figure 16. APC mean 7 x 7

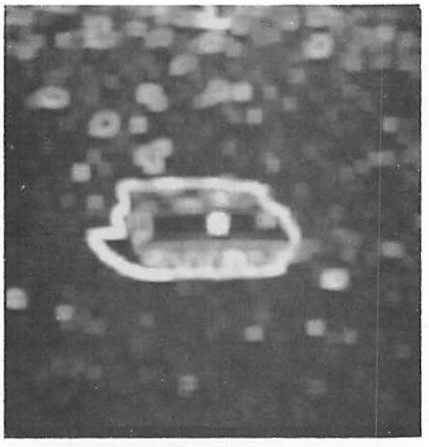


Figure 17. APC Variance 7 x 7

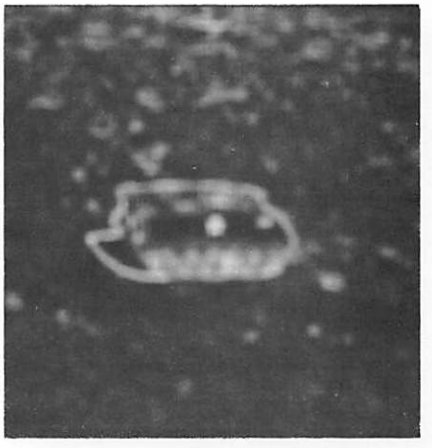


Figure 18. APC Texture 7 x 7

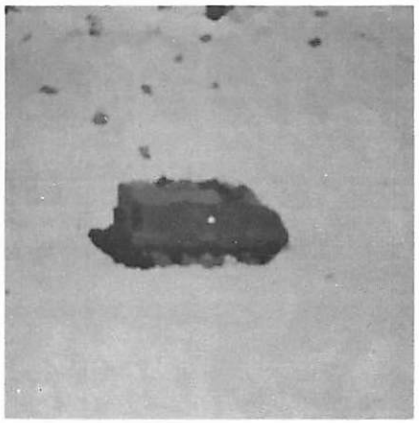


Figure 19. APC mode 7 x 7

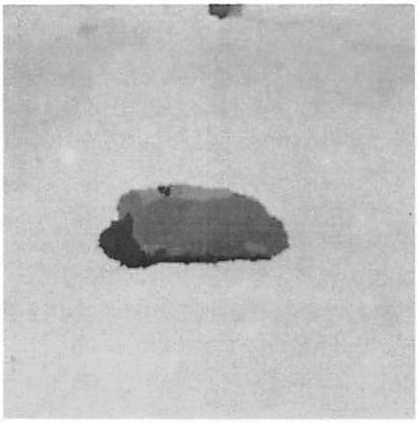


Figure 20. APC mode 15 x 15

Figure 21. APC median 7 x 7

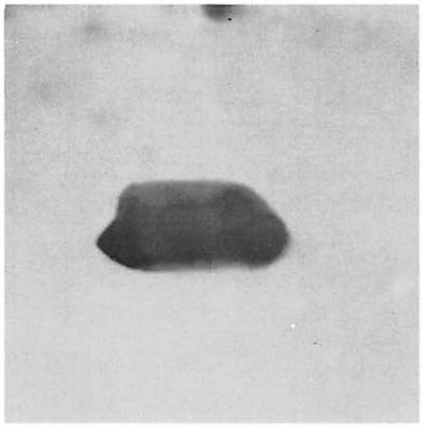


Figure 22. APC median 15 x 15

Some of the features are highly influenced by the value of a few pixels within the block. Examples of this are variance and texture (averaged modified Roberts Cross operator). These features always have large values at edge boundaries. Segmentations based on these features would therefore be expected to identify "edge" as a distinct region in the segmented image.

The median and mode features are formed by computing the grey scale histogram median and mode in a block centered on each pixel in the scene. These features (mode in particular) do not cause substantial degradation of the scene resolution. The median feature does cause some blurring while the mode feature fills in some concave areas and clips some highly convex areas. Nevertheless, features subjected to modal filtering are expected to substantially improve the classifier performance since they are based on the most frequently occurring values within the block centered on each pixel. Texture features having these properties are being investigated.

The activity in progress at present is concerned with feature thinning as a preprocess to clustering. The feature covariance matrix was computed and subjected to Karhunen-Loeve transformation. The resultant features are linear combinations of the original features and are statistically uncorrelated. The highly correlated set of original features will be replaced with an uncorrelated set that is substantially smaller. Reduction of the feature space dimension in this manner is expected to reduce the amount of computation required in the clustering algorithm.

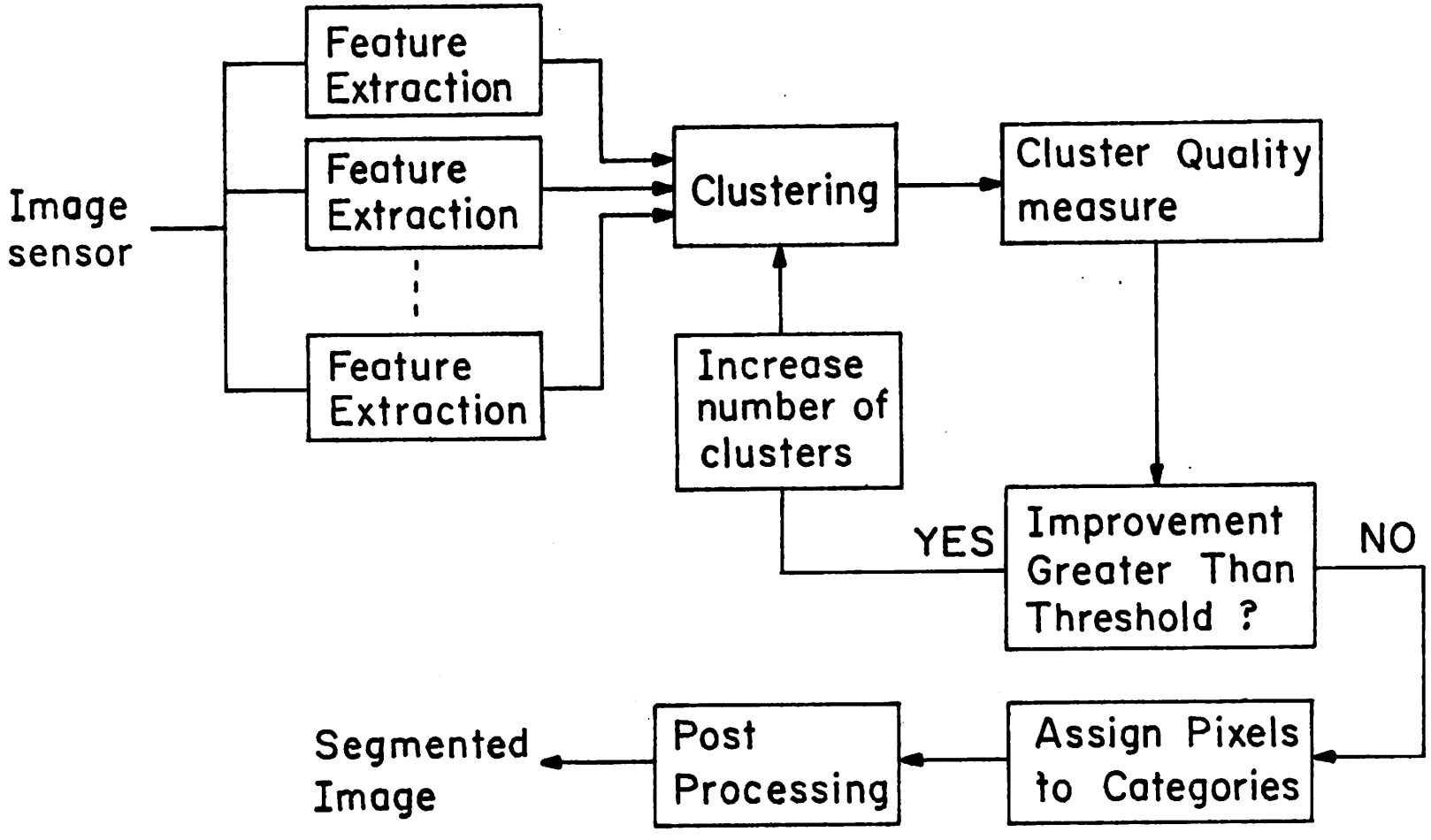


Figure 23. Clustering Procedure

$$w(k+1) = w(k) \qquad (m-1)N+1 \le k \le m\,N-1$$
$$c(k+1) = c(k) \qquad m = 1, \ldots, N \tag{3}$$

Since the object width is a positive quantity and the object must be within the image, $w(k)$ and $c(k)$ satisfy

$$w(k) \ge 0$$
$$w(k)/2 + |c(k)| \le N/2 \tag{4}$$

In terms of the widths and centers, the chosen representation for $\lambda(k)$ is

$$\lambda(k) = \begin{cases} 1 & \text{for } (m-1)N+N/2 + c(k) - w(k)/2 \le k \le \\ 0 & \text{else } (m-1)N+N/2 + c(k) + w(k)/2 \end{cases} \tag{5}$$

The quantities $w(k)$ and $c(k)$ are random sequences with known first and second order statistics. Let

$$w(k) = \tilde{w}(k) + \bar{w}$$
$$c(k) = \tilde{c}(k) + \bar{c} \tag{6}$$

where $\bar{w}$ and $\bar{c}$ are the mean values, respectively. As a first order approximation, the pair $r(k) \triangleq [\tilde{w}(k)\ \tilde{c}(k)]'$ is assumed to represent a first order Markov sequence,

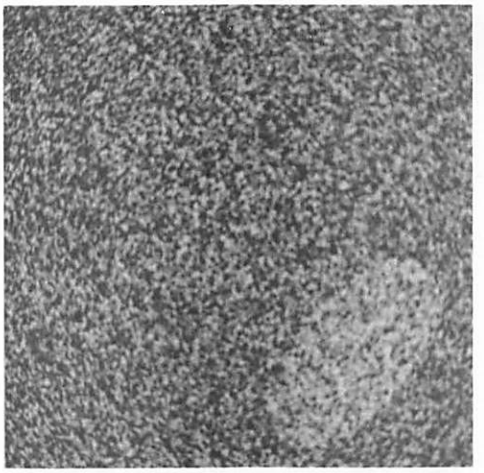
$$r(k+1) = A_o(k)\, r(k) + B_o(k)\, u_o(k) \tag{7}$$

where $u_o(k) = [u_1\ u_2]'$ is a zero mean white normal sequence with unit covariance, $A_o(k)$ and $B_o(k)$ are matrices of appropriate dimensions. As a result of eq.(3)

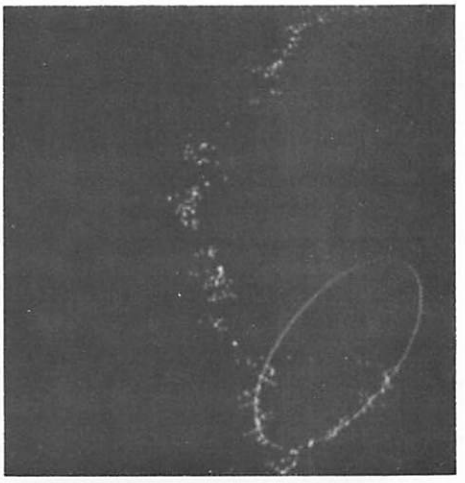
$$A_o(k) = I \qquad \text{(identity matrix)}$$
$$B_o(k) = 0 \tag{8}$$

for all k except $k = mN$, $m = 1, \ldots, N$.

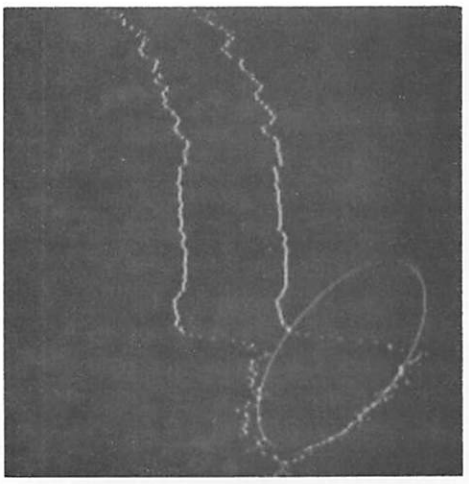
By augmenting eq.(7) with dynamic equations representing the statistics of the object and background it is possible to state a nonlinear MMS estimation problem that can be solved by the method



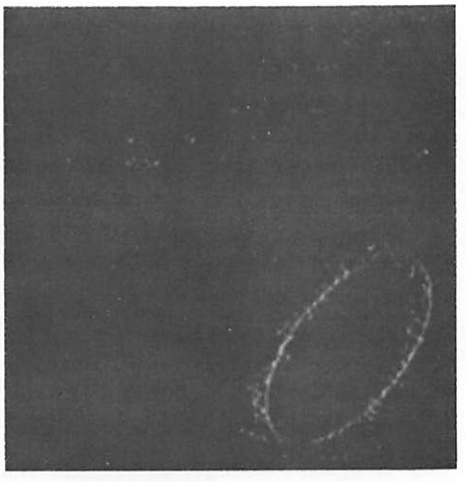
(a) the observation, $\rho = .25$



(b) estimate



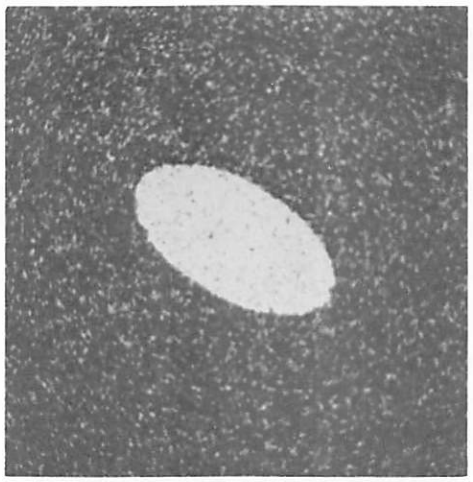
(d) estimate



(c) refined estimate

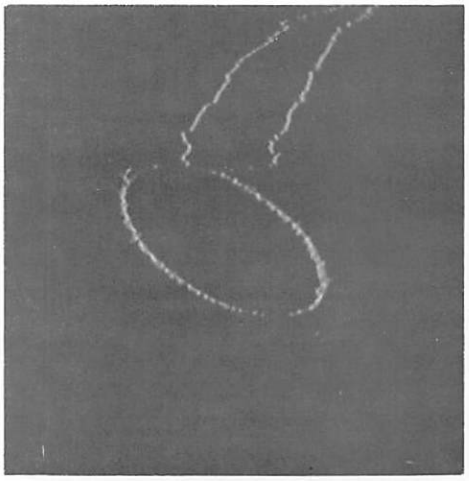
(e) refined estimate

Fig. 2. Estimated and original boundaries for known (b, c) and unknown (d, e) object luminance. ρ indicates signal-to-noise ratio.

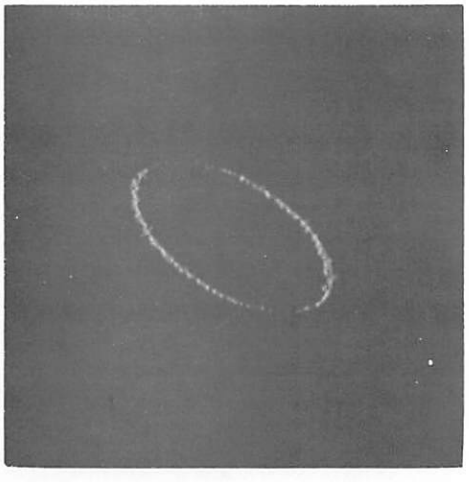


(a) the observation, $\rho = 1.0$

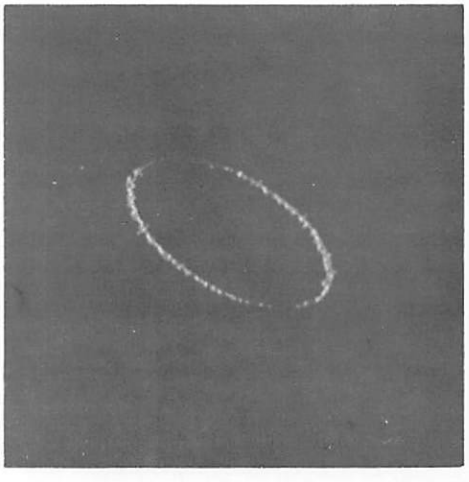
(b) estimate



(d) estimate



(c) refined estimate



(e) refined estimate

Fig. 1. Estimated and original boundaries for known (b, c) and unknown (d, e) object luminance. ρ indicates signal-to-noise ratio.

$$A(mN) = \begin{pmatrix} .926 & 0 & 0 \\ 0 & .97 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad B(mN) = \begin{pmatrix} 7.78 & 0 & 0 \\ 0 & 2.91 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$E\, s_o = 0.5 \qquad E(s_o - E\, s_o)^2 = 0.05$$

Although the estimator procedure is applicable to non-binary pictures, the way to reduce the possible increase in computation is still under investigation.

Reference

1. N. Nahi and M. Naraghi, "A General Image Estimation Algorithm Applicable to Multiplicative and Non-Gaussian Noise," 18th Midwest Symposium on Circuits and Systems, Montreal, Canada, August 11-12, 1975.

2.6 SHAPE ANALYSIS AND DESCRIPTION

Erica M. Rounds

This research is directed toward constructing region representations which are useful at the interpretation level. A good representation should be largely invariant with respect to location, orientation, and size to facilitate matching with object descriptions in the knowledge base <3>. The process of extracting global features from region descriptions is called here "shape analysis."

Shape analysis can be considered as an intermediate level between the segmentor and the interpreter. Each of these levels deals with image descriptions on a conceptually different basis. For example,

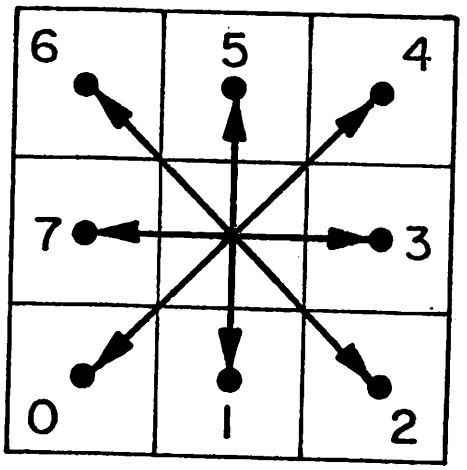


Figure 1. 8 - neighborhood.

knowledge base. Given the resolution of images it is often extremely difficult to manually extract straight continuous boundary segments.

Later portions of this report present an algorithm for generating figures or regions from their boundary coordinates.

References

1. R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, Wiley & Sons, New York, 1973.

2. H. Freeman, "Analysis of Line Drawings," Proceedings NATO Advanced Study Institute, France, June 1976.

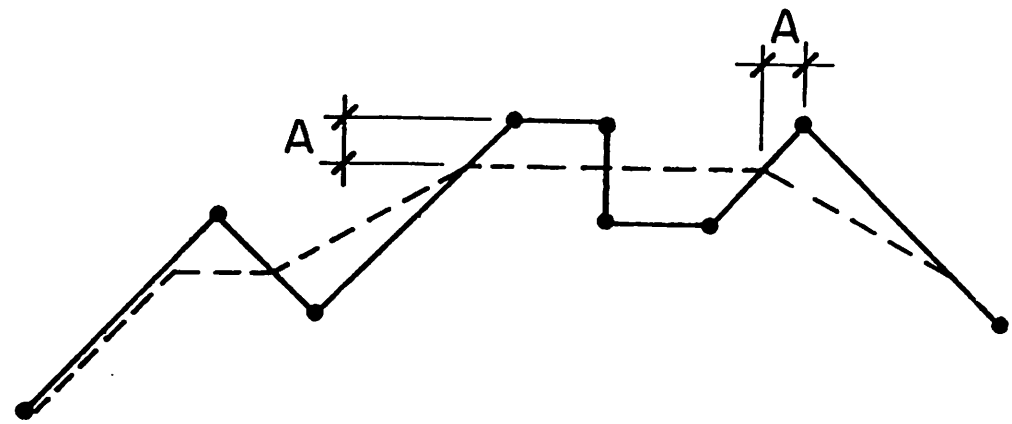
3. E.M. Rounds, "Interpretation and World Knowledge," USCIPI Report 660, Spring 1976, pp. 105-114.

4. K.J. Turner, "Computer Perception of Curved Objects Using a Television Camera," Ph.D. Thesis, University of Edinburgh, 1974.

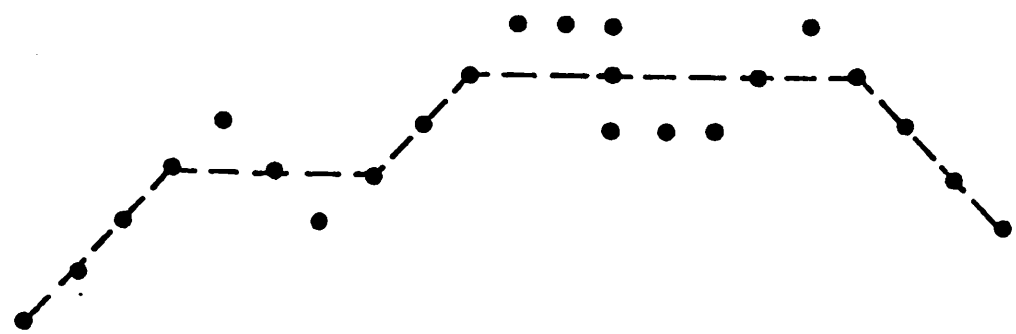
2.7 LOCAL SMOOTHING OF DIGITAL CONTOURS

Erica M. Rounds

In extracting shape features from region boundaries, a large amount of information is contained in the global discontinuities of curvature, i.e., at the extreme points where a significant change in direction occurs. It is frequently desirable to perform a prior smoothing operation on the boundary points to reduce the number of small local extrema. These may be due to noise or to the particular



(a)



(b)

Figure 1. Hysteresis smoothing

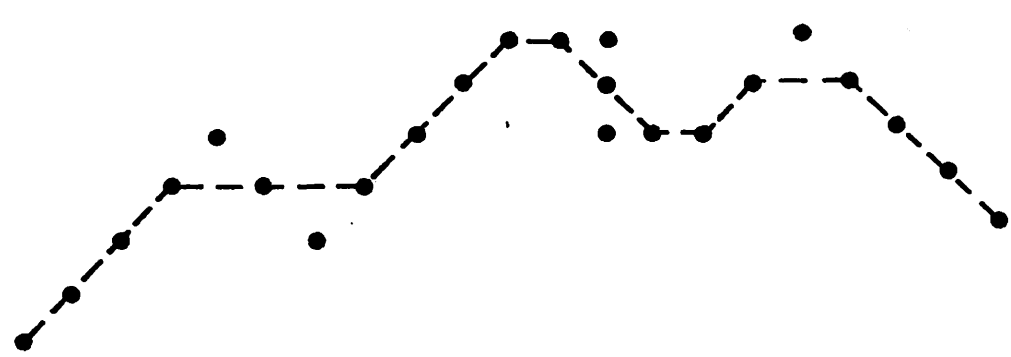


Figure 2. P-S pair smoothing

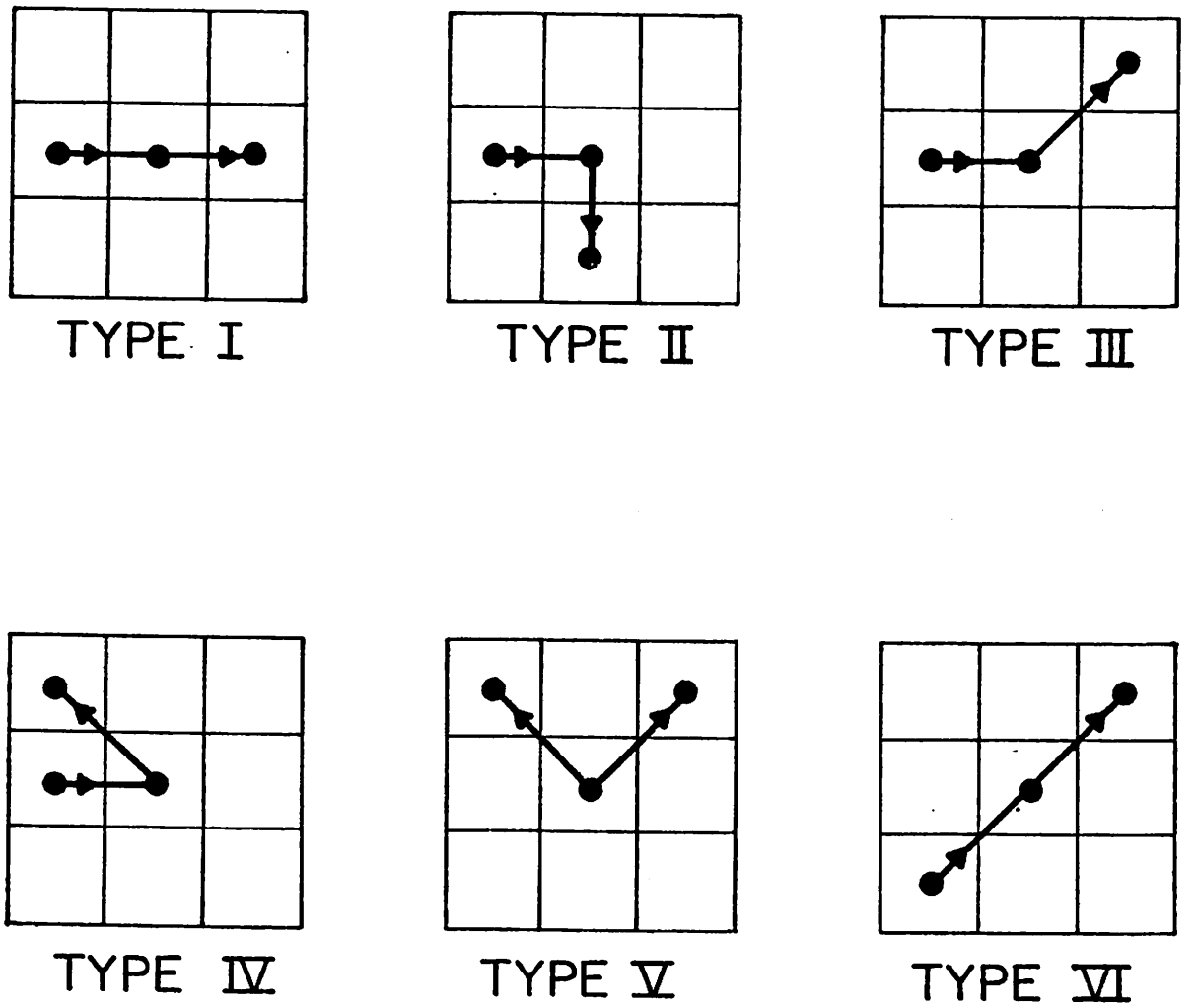


Figure 3. Distinct P-S pairs.

(0) Initialize: Scan the data points until a point p_i is found such that p_i has type I, III, or IV. This is the first smoothed point $\bar{p}_{j-1}$. Set i to i+1.

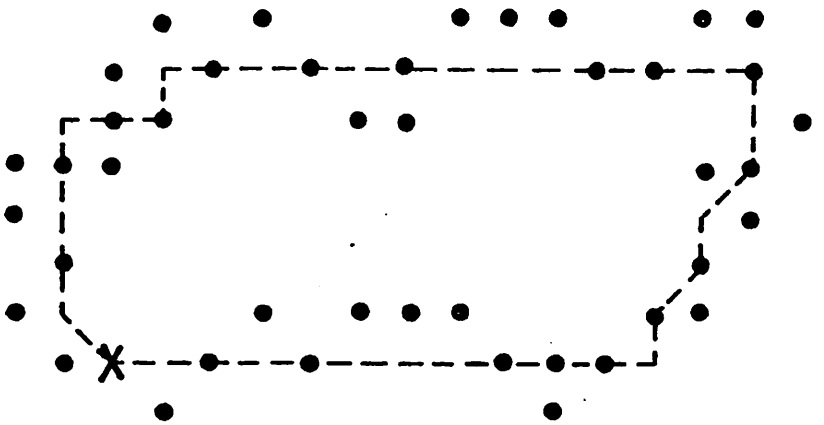
(1) Compute $\alpha = \alpha(\bar{p}_{j-1}, p_{i+1})$ and assign a type to p_i.

(2) If type = I, III, or IV then $\bar{p}_i = p_i$. Go to (3). If type = II or IV then set i to i+1 and go to (1). (This checks the type of the new reference point before assigning the smoothed point.) If type = V then compute $\bar{p}_j$ such that $\bar{p}_j$ has type I. Go to (3).

(3) Set i to i+1, j to j+1 and continue with remaining data points at (1).

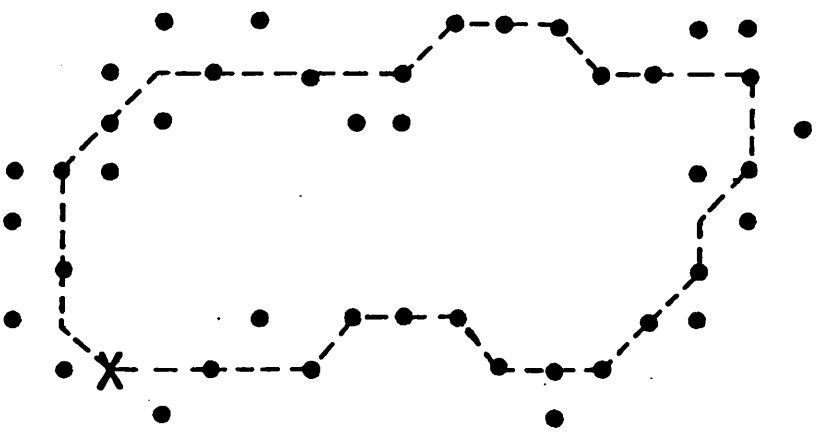
The performance of the algorithm is demonstrated on two digitized contours shown in figures 5a and 5b. Figures 6 and 7 compare the effects of Hysteresis smoothing with P-S pair smoothing. In general, P-S pair smoothing tends to produce a curve which has a close resemblance to the original one. Since this method is not direction-dependent it will preserve symmetry. It is also insensitive to gaps in the boundary, since successive points separated by more than 2A are not changed. The choice of starting point (marked in figures 6 and 7 by X) assures that the algorithm will produce a closed contour.

This smoothing algorithm can serve as a preprocessor before applying more sophisticated techniques for curve fitting. Although the 3 x 3 neighborhood is usually not large enough to assess the global significance of some corners, it is felt that the economy of the algorithm outweighs these disadvantages. The large number of



(a)

Hysteresis smoothing



(b)

P-S pair smoothing

Figure 6. Smoothed points of Figure 5. a.

possible configurations which must be considered in a 5 x 5 or 7 x 7 neighborhood would add considerable computational complexity without producing much better results.

Reference

1. S.J. Mason and J.K. Clemens, "Character Recognition in an Experimental Reading Machine for the Blind," in Recognizing Patterns, P.A. Kolers and M. Eden, eds., M.I.T. Press, Cambridge, Massachusetts, 1968.

2.8 FIGURE CONSTRUCTION FROM ITS CONTOURS

Erica M. Rounds

A method is presented for constructing a planar, digitized figure from its closed boundary. This can be useful for extracting a desired region from a given scene once its boundary is determined. Another potential application is image synthesis where the image is constructed from various regions to which different intensity, color, or texture values have been assigned. Image synthesis could be a valuable tool for comparing the processed picture information with the original input scene.

In the following we assume a digitized contour using eight-neighborhood adjacency and counter-clockwise traversal. The algorithm permits any number of "holes" in the interior of the figure or region. Hole boundary lists are simply appended to the region

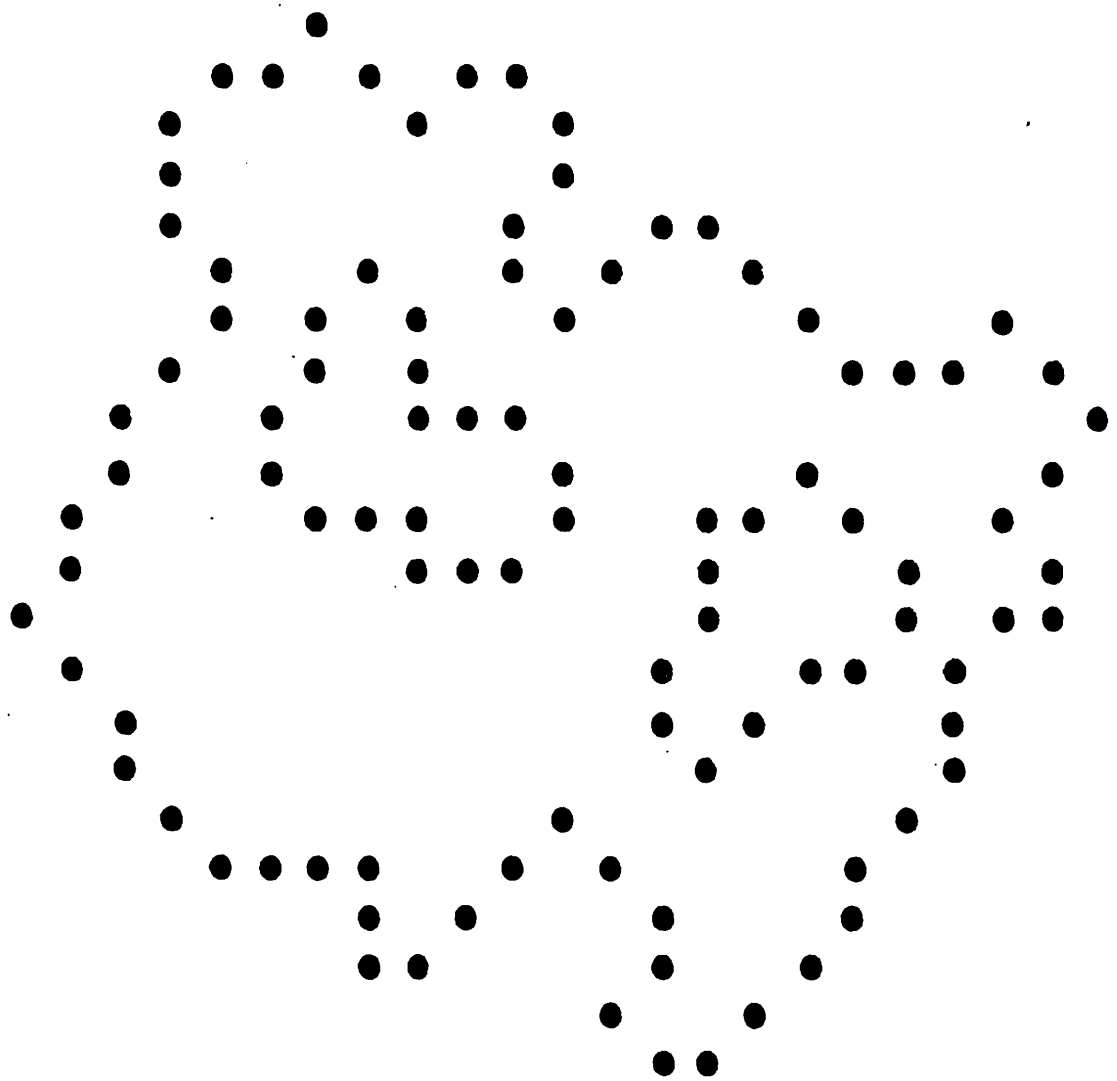


Figure 1. Contour points of regions with holes.

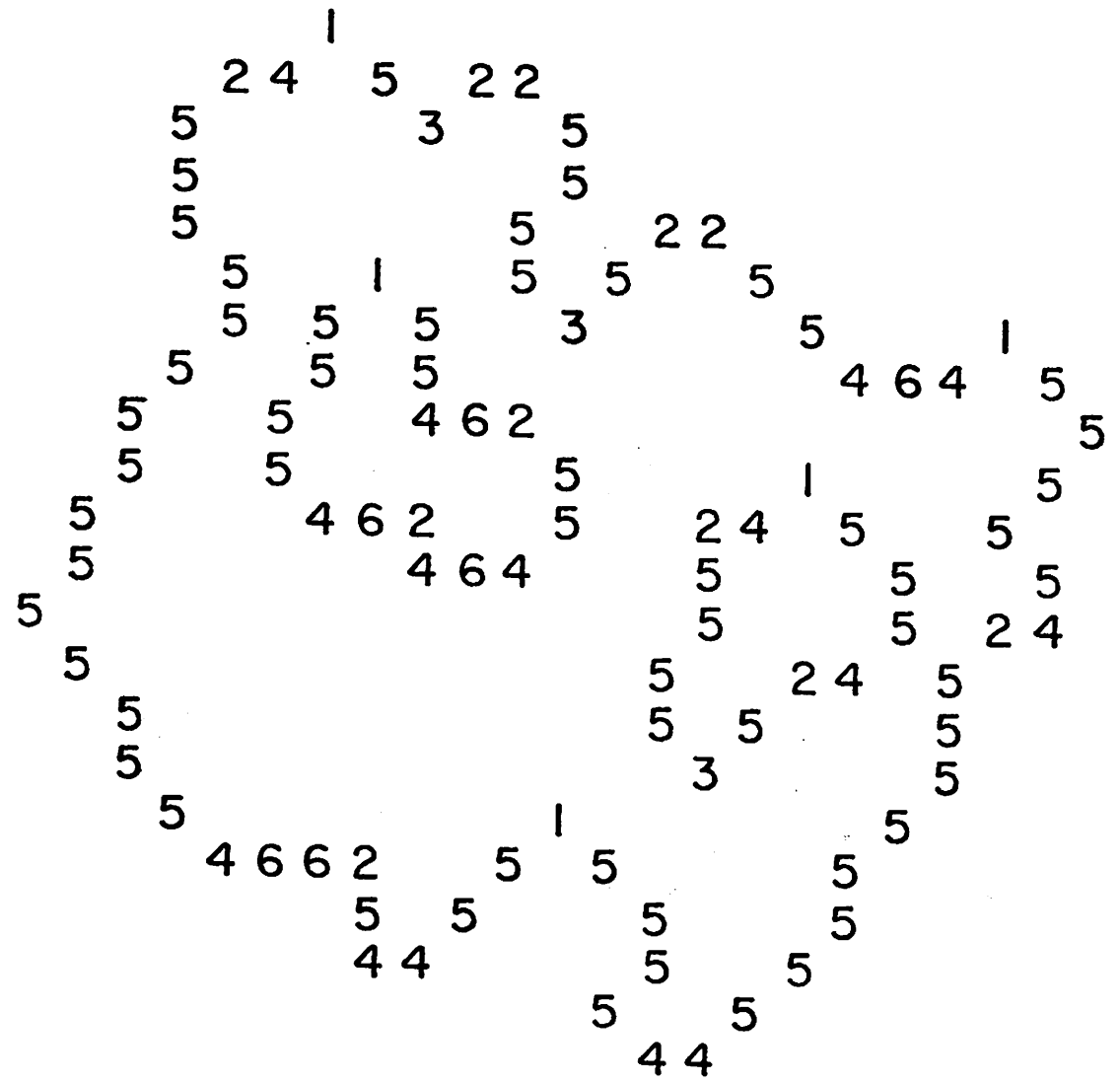


Figure 3. Assignment of boundary types.

(a) types 1 or 3 (figure 4f) and matched pairs of types 2 and 4 (figure 4a) are convex boundary segments which are ignored;
(b) type 5 (figure 4g) and unmatched pairs of 2 and 4 (figure 4c) are the beginning of an interior row.

(2) If it is approached from the interior, then
(a) types 1 or 3 (figure 4e) and matched pairs of types 2 and 4 (figure 4b) are concave boundary segments which signify the continuation of an interior row;
(b) type 5 (figure 4h) and unmatched pairs of 2 and 4 (figure 4d) are the end of an interior row.

These cases are easily distinguished by using two parameters, one for the interior/exterior condition and one for remembering when a horizontal boundary segment is scanned. Figure 5 shows the result of applying step 2 to the boundary types in figure 3.

The preceding algorithm requires only the list of boundary coordinates. This could be the output of a region analyzer. It is computationally quite efficient since each boundary point is examined only once during each of the two steps. Storage requirements are minimal and are determined by the region size. The resulting picture matrix can be mapped into any desired area for the purpose of image synthesis.

The silhouette generator by Dudani <1> is similar in many respects. It operates on a list of boundary points together with a list of predecessor/successor types (for a discussion of these, see

section 2.6 of this report). The notion of a "region crossing point" includes the above boundary type 5 as well as some other cases. The algorithm requires one additional step and cannot be used directly on regions with holes.

Reference

1. S.A. Dudani, "Region Extraction Using Boundary Following," to appear in Proceedings 1976 IEEE Workshop on Pattern Recognition and Artificial Intelligence.

2.9 AERIAL PICTURE SYNTHESIS

Benham Ashjari

Image synthesis involves the derivation by computer of "basic information" from an image and applying it to recreate the picture from a set of available pictorial data bases.

The research currently being conducted involves digitized aerial photographs. In some cases, the synthesized version of a picture and the picture itself are close enough to describe various applications of image synthesis. In general, the synthesized picture has statistical and visual similarities to the original one. The "basic information" derived from the original image suffices to describe and form the synthesized version, and therefore this "basic information is the information which represents the picture of interest and can be stored or transmitted instead of the original picture. The technique

Figure 1. Original Picture

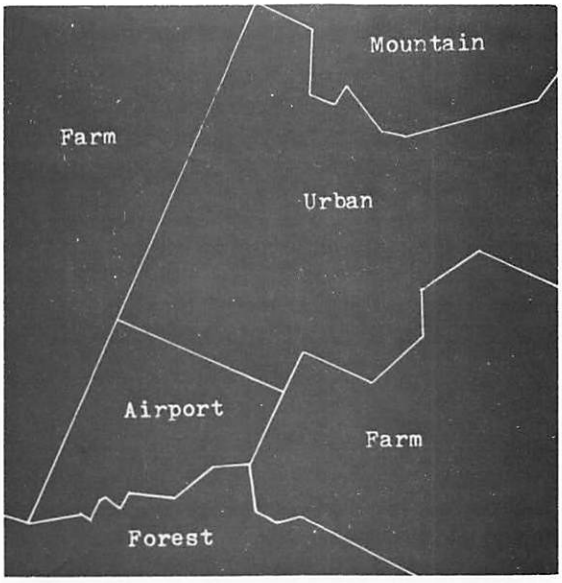


Figure 2. Segmented picture

Urban file (09)

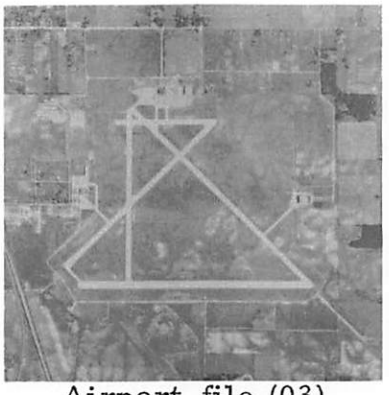
Urban file (13)

Farm file (37)

Farm file (48)

Forest file (54)

Mountain file (65)



Airport file (03)

Figure 3. Aerial data base used for synthesis of figure 1.

example, the urban area does not include a river passing through it. Figure 3 shows some of the 128 x 128 data base pictures. Ideally we need 32 pictures of each aerial scene (for a 5 bit code), and that would make the total number in the aerial data base 288. However, we have a total of 91 now available.

Statistical Analysis on the Data Base

In order to have a measure of closeness between the data base pictures and the corresponding segmented area of the original picture, we need to set some mathematical criteria. A variety of measures have been considered for coarseness, dissimilarity, homogeneity, and orientation which all fit into the area of texture analysis.

Gray-Level Co-occurrence Matrix

Before analyzing different criteria, we need to discuss the concept of gray-level co-occurrence matrix. (In the literature this is sometimes referred to as the grey tone spatial dependence matrix <3>.)

Consider the (x,y) plane and a picture f(x,y) defined over this plane. Let $\Delta = (\Delta x, \Delta y)$ be a vector in the plane. We can compute the joint probability density of any pair of gray-levels separated by the vector Δ. For the discrete case, $\Delta x, \Delta y$ are integers. By counting the number of times that each pair of gray levels occur at separation Δ, we can form an array of m x m, where m is the number of gray levels present in the picture. In this case, we count pairs of gray levels at separation Δ or $-\Delta$ indifferently. Therefore a symmetric matrix is

If a texture is coarse, and Δ is small compared to the sizes of the texture elements, the pairs of points at separation Δ should usually have similar gray levels. This means that the high values in the matrix C_Δ should be concentrated on or near its main diagonal. Conversely, for a fine texture, if Δ is comparable to the texture element size, then the gray levels of points separated by Δ should often be quite different, so that the values in C_Δ should be spread out relatively uniformly. Thus a good way to analyze texture coarseness would be to compute, for various values of the magnitude of Δ, some measure of the scatter of the C_Δ values around the main diagonal. In what follows P(i,J) is the (i,J)th element of the given matrix (which has size m x m = 32) divided by a normalizing factor R which is equal to the sum of all of the matrix entries <4>. P(i,J) is actually a measure of joint probability density that the pairs of gray levels (i,J)'s occur at pairs of points separated by $\Delta = (\Delta x, \Delta y)$.

Measure of correlation:

$$f_1 = \sum_{i=1}^{m} \sum_{J=1}^{m} i \cdot J \left(\frac{P(i \cdot J)}{R} \right)$$

f_1 is a measure of coarseness <2>. f_1 can be normalized using means and standard deviations associated with the marginal density functions $P_x^{(i,J)}/R$, $P_y^{(i,J)}/R$.

Homogeneity measures:

$$f_2 = \sum_{i=1}^{m} \sum_{J=1}^{m} \left(\frac{P(i,J)}{R} \right)^2$$

(sum of squares of all matrix entries) f_2 is the angular second moment and is smallest when the P(i,J) are all as equal as possible. It is

$$f_4 = \sum_{i=1}^{m} \sum_{J=1}^{m} (i-J)^2 \, P(i,J) / R$$

This is essentially the moment of inertia of the matrix around its main diagonal; it is a measure of the degree of spread of the matrix values. It is expected to have the highest value at the direction perpendicular to the directionality of the texture.

Also the "angular" distribution of values of Fourier Power Spectrum is sensitive to the directionality of the texture in f. It will have its high values concentrated around the perpendicular direction to the directionality of the texture. However, at this moment it is not being considered for our measurements.

Feature Vectors

A 32 x 32 window is taken from each data base picture and a gray level co-occurrence matrix is computed for each of the four directions 0°, 45°, 90° and 135°. f_1, f_2, f_3 and f_4 are computed for each matrix. There are, therefore, 16 numbers associated with each data base picture. These 16 numbers can be assumed to be the coordinates of a 16-dimensional vector. Therefore, there are 32 vectors for each class (scene) of data base ($D_{i1}, D_{i2}, \ldots D_{i32}$) and we have nine classes in the data base ($1 \leq i \leq 9$). These numbers will be saved on tape for future use.

Segmentation

The first stage in synthesizing a picture is delineating the distinct areas and sectioning them according to their texture and

Once the different areas have been isolated (either manually or automatically), each must be matched against a wide selection of standard data bases. For example, there are 32 types of mountains in the data base. The mountainous area in the original picture must be replaced with the closest of these 32. For this purpose, we have to take a window out of the scene to be synthesized and compute the gray level co-occurrence matrix in four directions and basically do the same type of calculations as described earlier and come up with 16 numbers which constitute a feature vector Z. Suppose the feature vector is from the mountain class. We have to compare each vector $D_{mountain1}$, $D_{mountain2}$, ..., $D_{mountain32}$ with Z and pick the closest. No scheme has been developed yet for this purpose. However, the Fisher linear discriminant technique has been discussed in the literature and may be considered. We can use a voting scheme to determine the closest data base for a given measurement Z.

Synthesis

The receiving computer has the following information: the segmentation grid and the data base numbers with their respective positions. Now the appropriate data base is selected for each section. If the computer finds the data base smaller than the space it is to fill, it automatically magnifies it 2 x 2, using an interpolative scheme (see figure 5).

The computer then matches the grid section to the data base, cutting off marginal areas, and pieces all of the segments together. What results is a composite picture with great similarities to the

Figure 4. Synthesized Picture of Figure 1

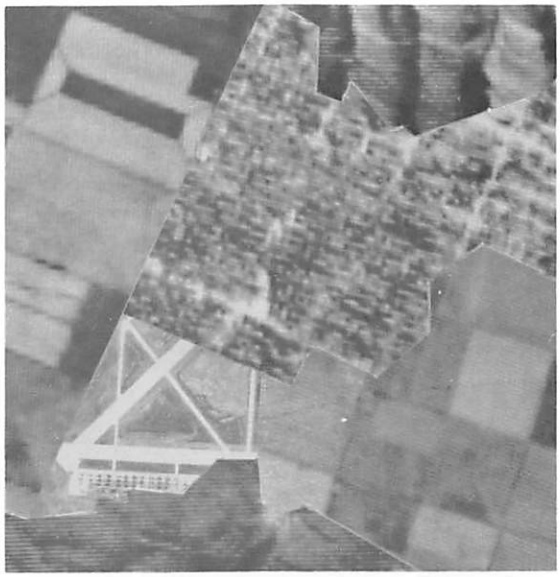


Figure 5. Synthesized with 2 x 2 magnification

The above figures (4 and 5) represent the simplest form of Image Synthesis for figure 1.

2.10 SOME EDGE SEGMENTATION RESULTS

Ramakant Nevatia

Satisfactory segmentation of objects in a scene is a key part of most scene analysis systems. The strategies employed for segmentation vary along many dimensions, often determined by the particular domain of the problem. Here, we restrict to the class of problems where no prototype specific information about the scenes is available, i.e., the objects present are not restricted to a small set of objects, nor any information known about their spatial relationships.

Results for four scenes, shown in figure 1 are presented here. All four pictures contain only a few (incidentally only one each) objects of interest and the objects are large (occupy a significant proportion of the picture). The first three pictures are of military vehicles against a desert background. The chosen pictures represent perhaps the simplest class of pictures, useful for practical applications. Yet, by the current state of the art, segmentation of these pictures is a difficult task (at least, without using prior knowledge). The complicating factors are the presence of texture and lack of multi-sensory information, such as color or range.

Partial segmentation results using an edge based approach are presented here. Figures 2a through 2d show the edges detected in the corresponding pictures by the application of a Hueckel Edge Operator <1>. The operator was applied at every second pixel to every other row. A small vector along the direction of the edge is displayed for each detected edge. Note that same thresholds were used for all four

these, image synthesis will never be useful. But for a great many uses, the time and cost saved in transmission of image data compensates for the loss of originality of the picture. The diagram shows various stages of synthesis and transmission.

References

1. H.C. Andrews, "Clustering for Image Segmentation," USCIPI Report 660, March 1976.

2. R.P. Kruger, "Image Textural Discrimination," USCIPI Report 444, February 1973.

3. R.M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, November 1973.

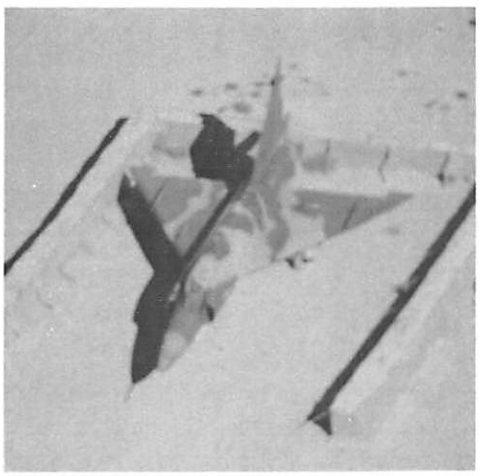
4. J.S. Weszka, C.R. Dyer, A. Rosenfeld, "A Comparative Study of Texture Measures for Terrain Classification," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-6, No. 4, April 1976.

5. E.B Troy, E.S. Deutsch, and A. Rosenfeld, "Gray-Level Manipulation Experiments for Texture Analysis," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, January 1973.

6. D.N. Graham, "Image Transmission by Two-Dimensional Contour Coding," IEEE Proceedings, Vol. 55, No. 3, March 1967.

(a)

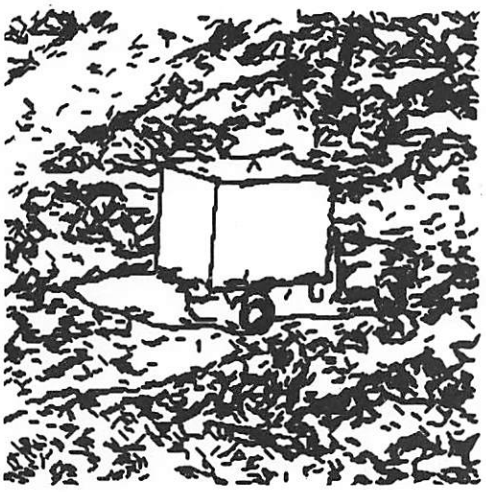
(b)



(c)

(d)

Figure 1. Four selected scenes.



(a)

(b)

(c)

(d)

Figure 2. Edges detected in pictures of Figure 1.

pictures and no manual tuning for optimum edge detection was attempted.

The edge pictures contain a large number of edges belonging to the textured background but the objects are clearly visible to the human eye. It is hypothesized that edges belonging to objects (particularly man-made objects) tend to occur along elongated segments, whereas most background texture edges tend to be distributed randomly. An edge linking procedure that links edges with similar orientations in straight line segments along chosen directions has been described previously <2>. Results of applying this linking procedure to edges shown in figure 2 are shown in figure 3. Here, only those edge segments which contain at least seven edge elements are retained.

Of course, such linked edge segments do not constitute a complete segmenation of the scene. However, they offer promising starting points for further segmentation. The next step, for example, may be a completion of the partial boundaries of the objects. For many applications, such as recognition, the partial boundaries may suffice in many cases. Research is in progress for further extending this approach.

Results shown here indicate that the techniques used are fairly successful in isolating extended edges of objects of interest. Many of the difficulties arise in the process of edge detection itself. Hueckel edge detector is unable to detect fine, high resolution edges because of the large size of the operator and also fails in presence

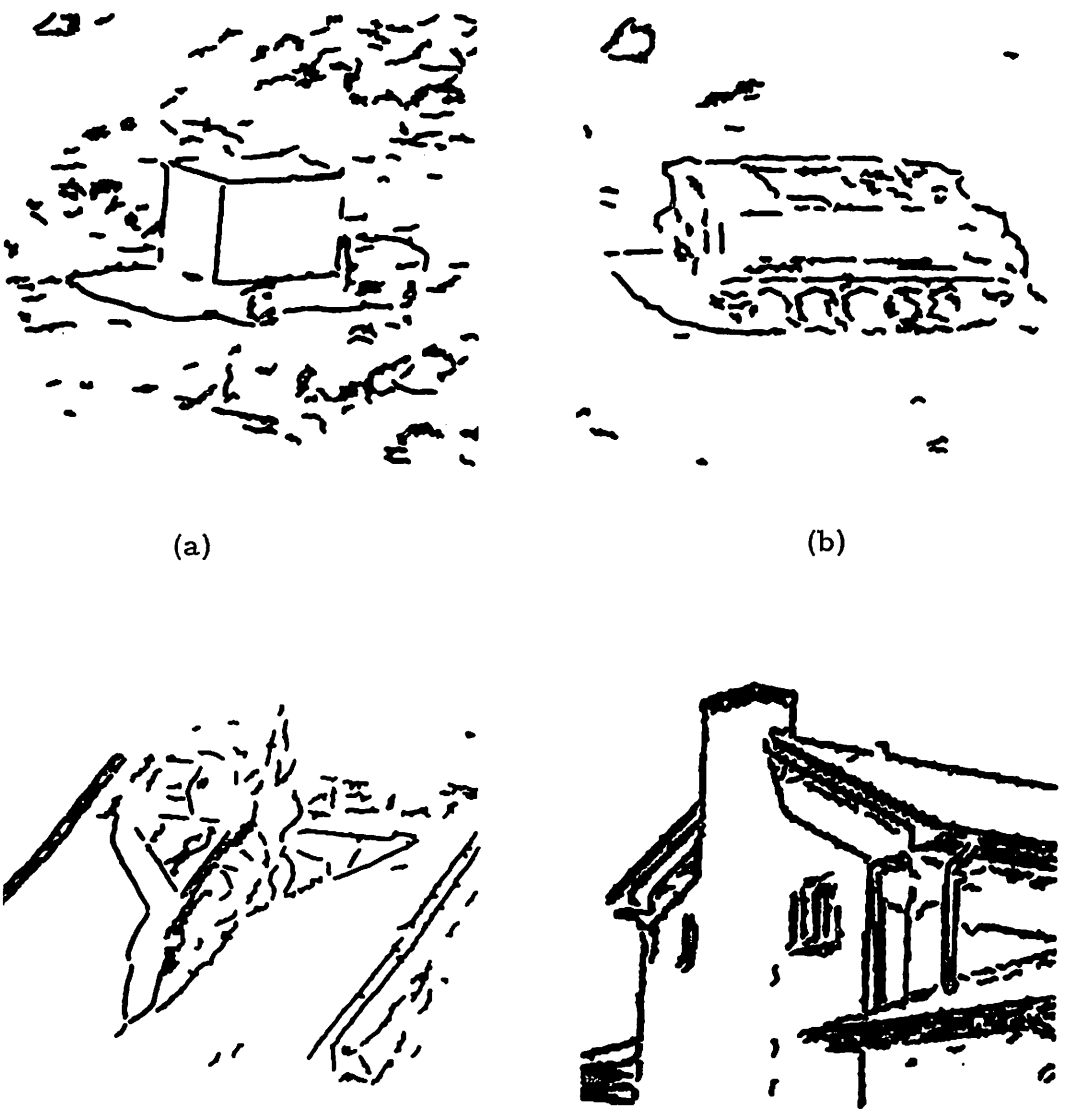


Figure 3. Linked edge segments from edges of Figure 2.

of fine texture. Local high resolution edge detectors, and also larger, texture edge operators may be used to supplement the edges detected by Hueckel operator.

A major motivation of presenting these intermediate results is to stimulate comparison of techniques used by various groups working on the ARPA Image Understanding program. It is felt that the selected scenes are good test vehicles for segmentation techniques. (A larger number of similar pictures is now available in the data base prepared by USC.)

The picture used in figure 1d is a sub-picture of the house scene used in Ohlander's work at CMU <3>. Ohlander's technique relies on recursive segmentation based on histogramming of various image attributes. For the house scene, the most effective attributes seem to be those based on color, such as hue and saturation. Note that the results presented here make no use of color information at all. (Use of color in segmentation using techniques similar to those described here is discussed in <4>.)

References

1. M. Hueckel, "An Operator which Locates Edges in Digitized Pictures," Journal of the ACM, Vol. 18, No. 1, January 1971, pp. 113-125.

2. R. Nevatia, "Object Boundary Determination in a Textured Environment," Proceedings of the ACM '75 Conference, Minneapolis, Minnesota, October 1975, pp. 32-36.

Another approach to the problem will be introduced here. In this approach a finite set of edges in a small block is considered. The corresponding Fourier coefficients are computed. Edge detection is implemented by comparing these coefficients to suitable thresholds.

The computation of Fourier coefficients is a simple problem, and can be easily generalized to shifted or rotated edges. This is not the case for the detection problem, where a statistical model is needed to obtain an optimum detection strategy. In fact the work described in <4> can be used in solving this problem.

The set of edges considered in this paper is shown in figure 1. A function $f(x,y)$ defined on an N x N block has corresponding Fourier coefficients $F(u,v)$ given by

$$F(u,v) = \sum_{y=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{x=-\frac{N-1}{2}}^{\frac{N-1}{2}} f(x,y)\, W^{xu+yv} \quad * \qquad (1)$$

where

$$W = \exp\left(-\frac{2\pi\gamma}{N}\right) \qquad (2)$$

In many cases closed forms for the Fourier coefficients can be derived. As an example, for the case of a horizontal edge given by

$$f_H(x,y) = b \qquad \text{for } y < 0 \qquad (3a)$$

$$= b + \frac{h}{2} \qquad \text{for } y = 0 \qquad (3b)$$

$$= b + h \qquad \text{for } y > 0 \qquad (3c)$$

The corresponding Fourier coefficients are

* $\frac{1}{N^2}$ is dropped for simplification.

$$F_H(0,0) = N^2(b + \frac{h}{2}) \qquad (4)$$

Otherwise

$$F_H(o,v) = Nh \left[\frac{1}{2} + \frac{W^{\frac{N+1}{2}v} - W^v}{W^v - 1}\right] \qquad (5)$$

$$F_H(u,v) = 0 \qquad (\text{if } u \neq 0) \qquad (6)$$

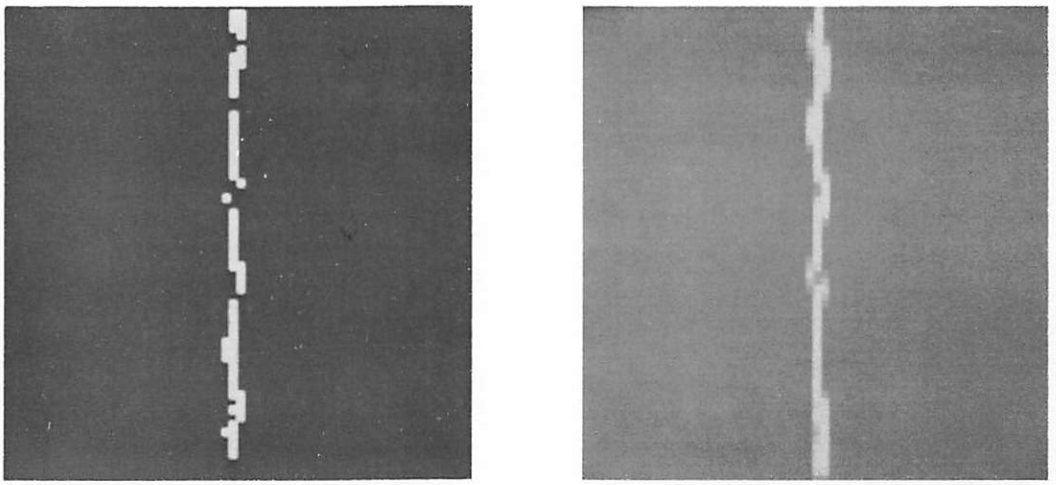
Fourier coefficients are calculated for all the cases shown in figure 1. The results for N = 5 are shown in figure 2.

From these results it is obvious that the edge direction can be determined from the Fourier coefficients. This technique is used with the test image given in <5>. The Fourier coefficients corresponding to each block are compared to different thresholds to decide the presence or absence of an edge. The direction of the edge, if present, is also determined. Edge maps for SNR = 10 and 1 are shown in figures 3 and 4, respectively. In figures 3a and 4a only the information about edge location is used, while in figures 3b and 4b information about the edge direction is used to enhance the edge map. In both cases the resulting edge maps show some degree of improvement over those obtained by the Sobel operator <5>.

References

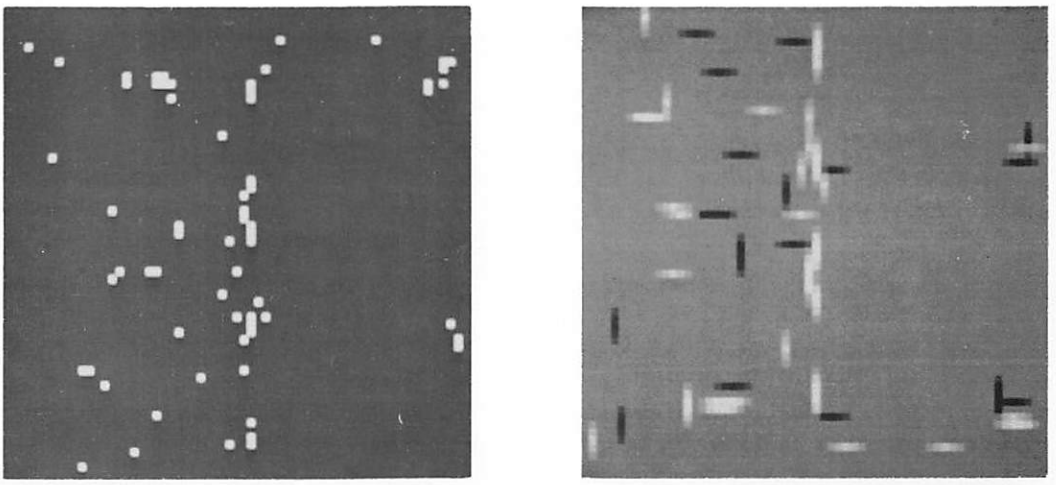
1. L.S. Davis, "A Survey of Edge Detection Techniques," University of Maryland Computer Science Center, November 1973.

2. M.H. Hueckel, "An Operator which Locates Edges in Digitized Pictures," Journal of the ACM, Vol. 18, No. 1, January 1971, pp.



(a) (b)

Figure 3. Edge maps for SNR = 10.



(a) (b)

Figure 4. Edge maps for SNR = 1.

113-125.

3. "A Local Visual Operator which Recognizes Edges and Lines," Journal of the ACM, Vol. 20, No. 4, October 1973, pp. 634-647.

4. I. Abdou and W.K. Pratt, "Statistical Design of Edge Detectors," see present report.

5. W.K. Pratt, "Figure of Merit for Edge Location," USCIPI Semiannual Technical Report, March 1976, pp. 85-93.

2.12 STATISTICAL DESIGN OF EDGE DETECTORS

Ikram E. Abdou and William K. Pratt

A wide variety of edge detection methods have been devised <1-3>. Only recently has there been any attempt to devise a figure of merit for edge detector performance <4>. This report summarizes an ongoing effort to develop a statistical model for edge detection that can be used in the design and a priori evaluation of edge detectors <5>.

Edge Model: Figure 1 illustrates the basic model for a local ideal edge defined over a 3 x 3 block of pixels. These ideal edges are assumed to be subject to additive white Gaussian noise with standard deviation σ. The probability density at each 3 x 3 point in an observation array is therefore a Gaussian density of the form

$$G(a,\sigma) = [2\pi\sigma^2]^{-\frac{1}{2}} \exp\left\{\frac{-(x-a)^2}{2\sigma^2}\right\} \qquad (1)$$

where X represents the observed pixel amplitude and a is its mean

value as specified by figure 1.

Linear Edge Detection: In a linear edge detector the observed pixel array F(j,k) is multiplied by an enhancement mask M(j,k) over an I x L window to produce an edge gradient function

$$G = \sum_{j=1}^{L} \sum_{k=1}^{L} F(j,k)\, M(j,k) \tag{2}$$

Since the gradient operation is linear, the probability density of the gradient $p(G) = \mathcal{L}_E(u,o)$ will also be Gaussian with mean and standard deviation given by

$$u = \sum_{j=1}^{3} \sum_{k=1}^{3} F(j,k)\, M(j,k) \tag{3a}$$

$$\sigma = \sigma \left[\sum_{j=1}^{3} \sum_{k=1}^{3} M^2(j,k) \right]^{\frac{1}{2}} \tag{3b}$$

where the overbar indicates an ensemble average. Edge detection is performed by thresholding the magnitude of the edge gradient function G such that $|G| \geq T$ implies the presence of an edge, and $|G| < T$ implies the absence of an edge. The probability density of the gradient magnitude is

$$p(|G|) = [\mathcal{L}_E(u,\sigma) + \mathcal{L}_E(-u,\sigma)] \tag{4}$$

for $G \geq 0$.

The detection procedure outlined above is entirely analogous to signal detection in radar and communication systems. The performance of such systems is commonly measured in terms cf the probability of detection when the signal is present, and the probability of false detection (false alarm) when no signal is present. Applying this

$$X \equiv [F(1,3) + 2F(2,3) + F(3,3)] - [F(1,1) + 2F(2,1) + F(3,1)] \quad (8a)$$

$$Y \equiv [F(1,1) + 2F(1,2) + F(1,3)] - [F(3,1) + 2F(3,2) + F(3,3)] \quad (8b)$$

The probability densities of X and Y are given by the Gaussian densities

$$p(X) = \mathcal{N}(u_X, \sqrt{12}\,\sigma) \quad (9a)$$

$$p(Y) = \mathcal{N}(u_Y, \sqrt{12}\,\sigma) \quad (9b)$$

where u_X and u_Y are the means of X and Y, respectively, for a particular edge orientation. The probability density of the edge gradient is then found to be

$$p(G) = \frac{G}{12\sigma^2} \exp\left\{-\left(\frac{G^2 + u^2}{24\sigma^2}\right)\right\} I_o\left(\frac{Gu}{12\sigma^2}\right) \quad (10)$$

where $u^2 = u_X^2 + u_Y^2$ and $I_o(\cdot)$ is the modified Bessel function of zero order.

With this statistical model the probability of false detection for no edge present is

$$P_F = \exp\left\{-\frac{T^2}{24\sigma^2}\right\} \quad (11)$$

where T is the gradient edge threshold. The probability of detection becomes

$$P_D = Q\left\{\sqrt{\frac{3}{2}}\,\frac{h}{\sigma}, \frac{T}{\sqrt{12}\,\sigma}\right\} \quad (12)$$

where Q(a,b) is Marcum's Q function <6>.

Conclusions: The statistical design procedure summarized in this report appears to hold promise for the design of edge detectors in the presence of noise. Further study is underway to evaluate the procedure.

3. IMAGE PROCESSING PROJECTS

The image processing projects comprise an ongoing research activity directed toward image coding, image restoration, vision modelling, and the implementation of image processing systems. In image coding novel ideas based upon the results of the image understanding study are being explored as a means of achieving significantly higher compression ratios than obtainable by conventional coding methods. The image restoration studies are directed toward the solution of two major problems: blind restoration in which a priori information about image degradation is unavailable or incomplete; and constrained restoration which involves the use of luminance bounds and smoothness criteria to improved image restoration. Vision modelling research activities include the extension of previously developed models of the human visual system to encompass higher levels of visual perception in support of the image understanding program. Implementation studies are underway on techniques of nonlinear, two dimensional optical filtering which can be utilized as a form of sensor based image processing.

Both of these methods deal with sampled versions of the image on some linear transformation of it, and are thus affected by the sampling method used. In the SVD case this is readily apparent by considering an image $f(x,y)$ that can be written as the product of two functions $f_1(x)$ and $f_2(y)$ as $f(xy) = f_1(x)\, f_2(y)$. If this image is sampled on a Cartesian grid (x_i, y_k), $i = 1,2,\ldots,N$, $k = 1,2,\ldots,N$. The image matrix can be written as the outer product of the two vectors $[f_1(x_1)\ldots f_1(x_N)]$ and $[f_2(y_1)\ldots f_2(y_N)]$ which is a rank one matrix for all such separable images f. The point is that the degrees of freedom should be a characteristic of the original image and reflected in the sampled image only by our inability to collect an uncountably infinite number of samples for application on a computer.

This brings up some conceptual difficulties since the number of degrees of freedom of a function defined on a continuum is countably infinite at best viz the space of all square integrable functions on $[-1,1]$, $L^2(-1,1)$, where any $f(x) \in L^2(-1,1)$ can be written as

$$f(x) = \sum_{i=1}^{\infty} f_i \phi_i(x)$$

$\{\phi\}$ being a complete orthonormal set and

$$f_i = \int_{-1}^{1} f(x)\phi^*(x)dx.$$

Here to exactly specify f in an L^2 sense a countably infinite set $\{f_i\}$ is required. In general if we are willing to accept an approximation with some error say, δ, then there exists a number $N(\delta,\varphi)$ which is a function of δ and $\{\varphi\}$, such that

$$\left[\int_{-1}^{1} \left|f(x) - \sum_{i=1}^{N(\delta,\varphi)} f_i \varphi_i(x)\right|^2 dx\right]^{\frac{1}{2}} < \delta.$$

We could then define the degrees of freedom of that f at level δ, DoF_δ as the $\inf_{\{\varphi\}} N(\delta,\varphi)$. In this case where $f \in L^2(-1,1)$ for any complete

waveforms and first described by Slepian and Pollack <6>. While this result is known, the determination and utilization of these waveforms for large space bandwidth problems has met with little success. Therefore to arrive at any meaningful and useful results we will have to restrict the search to functions that are computationally feasible and possess the approximating properties required. One such set is the set of polynomial splines $S_{k,n}$ of degree k with n free knots and defined as follows

$S_{k,n} = \{s(t) \mid$ there exists $0 = x_0 \leq x_1 \leq \ldots \leq x_{r+1} = 1$ and integers $m_1 \ldots m_r$ with $1 \leq m_i \leq k+1$ for each i and $m = n$ such that s(t) Π in each of the intervals (x_i, x_{i+1}) while $s \in C^{k-m_i}$ in an open neighborhood of x_i, $i = 0,\ldots,r\}$.

Here Π_k denotes the class of all polynomials of degree less than k. $S_{k,n}$ thus is the class of polynomial splines allowing multiplicities up to order k+1 at each knot and every $s(t) \in S_{k,n}$ possess the representation

$$s(t) = \sum_{i=0}^{k} a_i t^i + \sum_{i=0}^{n} \sum_{j=0}^{m_i} c_{ij} (t - x_i)_t^{n-j+1}$$

where

$$(x)_+^n = \begin{cases} x^n & x \geq 0 \\ 0 & 0^W \end{cases} .$$

Clearly the method of fitting a different polynomial in different subintervals of [0,1] as done by Pavlidis <7> can be classified as a polynomial spline with appropriate knot multiplicities at the

level epsilon in the L^2 metric, $DoF(k,\beta,\|\cdot\|_2)$, will be taken as

$$DoF(k,\beta,\|\cdot\|_2) = \min\{n \mid s_i(t) \in S_{k,n} \text{ where } s_i(t) \text{ is a best } L^2 \text{ approximation to } f \in L^2[0,1]\}.$$

The degrees of freedom at level epsilon can be similarly defined in two dimensions if we do the following. Using the B-spline bases of order k we can define the bilinear spline approximation $S_{k,n_x,n_y}(xy)$ to f(xy) as

$$\hat{f}(xy) = \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} b_{ij}B_i^{(k)}(x)B_j^{(k)}(y) \equiv S_{k,n_x n_y}(x,y)$$

where the knots have been defined on the Cartesian grid (x_i, y_k) for $i = 1,2,\ldots,n_x$, $k=1,2,\ldots,n_y$. The degrees of freedom can then be taken as:

$$DoF(k,\beta,\|\cdot\|_2) = \min\{n_x + n_y \mid S_{k,n_x n_y}(x,y) \text{ is a best } L^2 \text{ approximation to } f \in L^2(0,1)\}.$$

For our preliminary work we obtained the bicubic spline fits from the following equations

$$\hat{f}(x_k, y_k) = \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} b_{ij}B_i^{(3)}(x_k)B_j^{(3)}(y_\ell)$$

$$k = 1,2,\ldots,N \quad \text{and} \quad \ell = 1,2,\ldots,N$$

In matrix notation this becomes

$$[f]_{N\times N} = [B_x^{(3)}][b_{ij}][B_y^{(3)}]$$

where $[B_x^{(3)}]$ is an N x n_x matrix, $[b_{ij}]$ is an n_x x n_y matrix and $[B_j^{(3)}]$ is an n_y x N matrix. The solution minimizing the residual, $\sum_{ij} |f(x_i,y_j) - \hat{f}(x_i,y_j)|^2$ is given by

$$[b_{ij}] = \left[[B_x^{(3)}]^T[B_x^{(3)}]\right]^{-1}[B_x^{(3)}]^T[f][B_y^{(3)}]^T\left[[B_y^{(3)}][B_y^{(3)}]^T\right]^{-1}.$$

For each experiment N was taken to be 128 so that the effects of the quadrature error implicit in this should not be a significant factor,

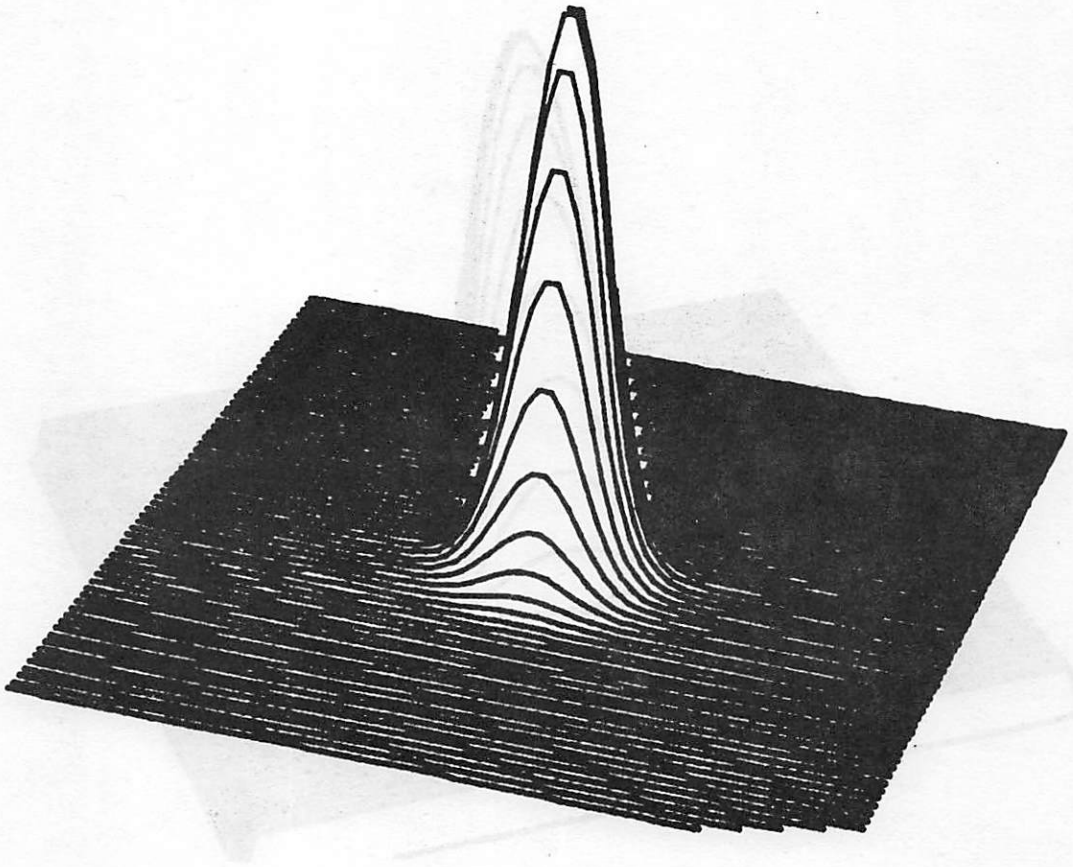


Figure 1. Bicubic Spline Fit with 10 Knots in X and Y Directions Determined from 4th Partials of X and Y Projections

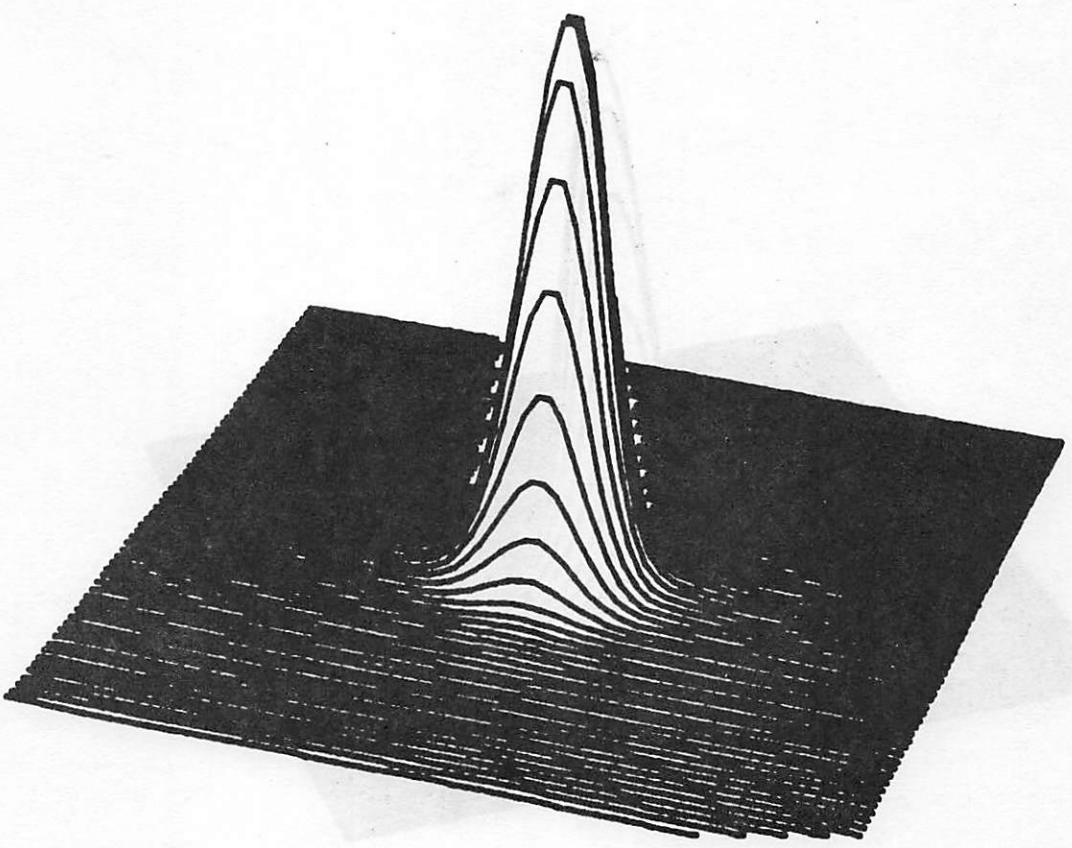


Figure 3. Bicubic Spline Fit with 20 Uniform Knots in X and Y Directions

NUMBER OF KNOTS	PLACEMENT MODE	RESIDUAL SUM OF SQUARES
10	UNIFORM	2.158
10	4th DIFFERENCES	1.29×10^{-3}
20	UNIFORM	3.24×10^{-3}

Table 1. Residual Sum of Squares for Knot Placement on Function $e^{-50(x^2+y^2)}$

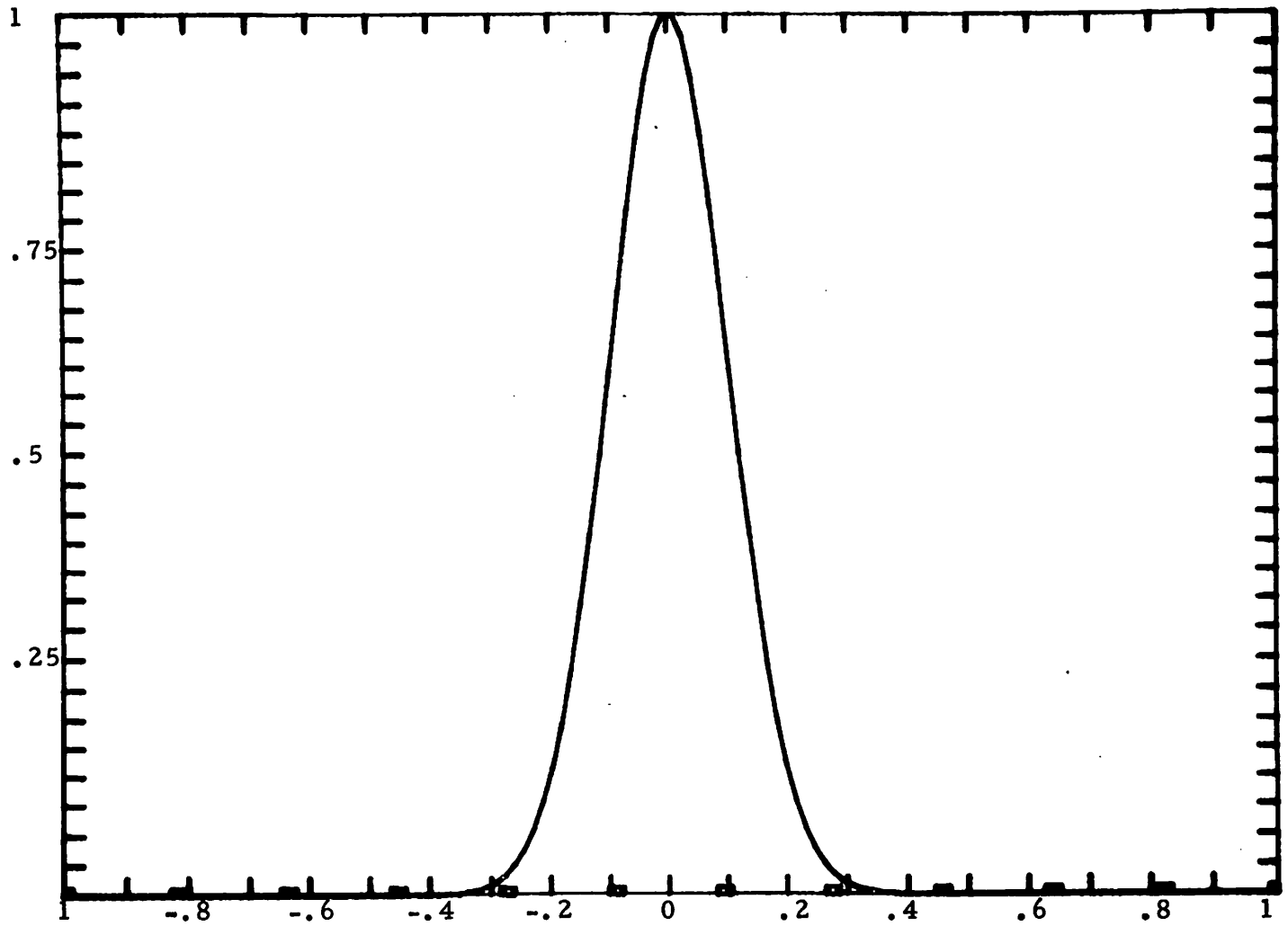


Figure 6. Y Projection and X-Knot Placement for 10 Uniform Knots (Y-Knot Placement Identical)

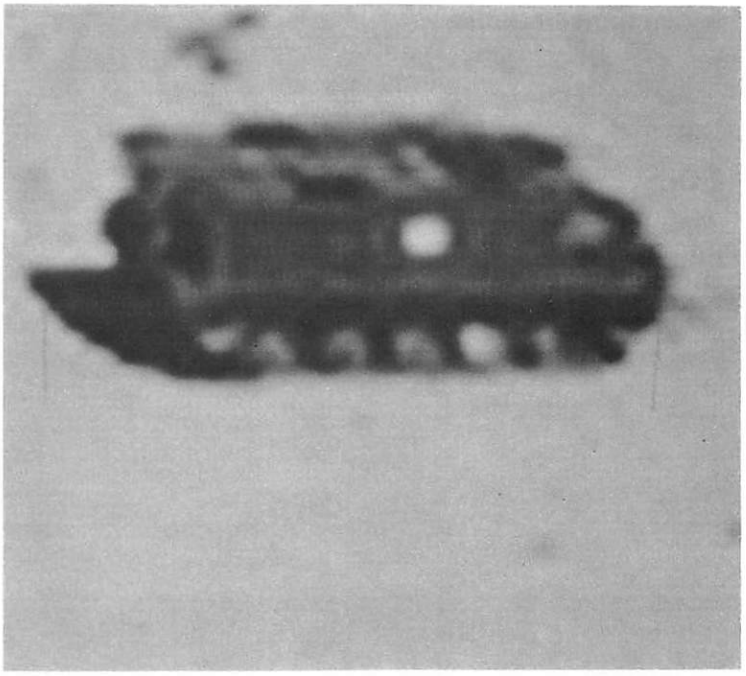


Figure 7. Bicubic Spline Approximation for APC with 40 Knots Uniformly Placed in X and Y Direction

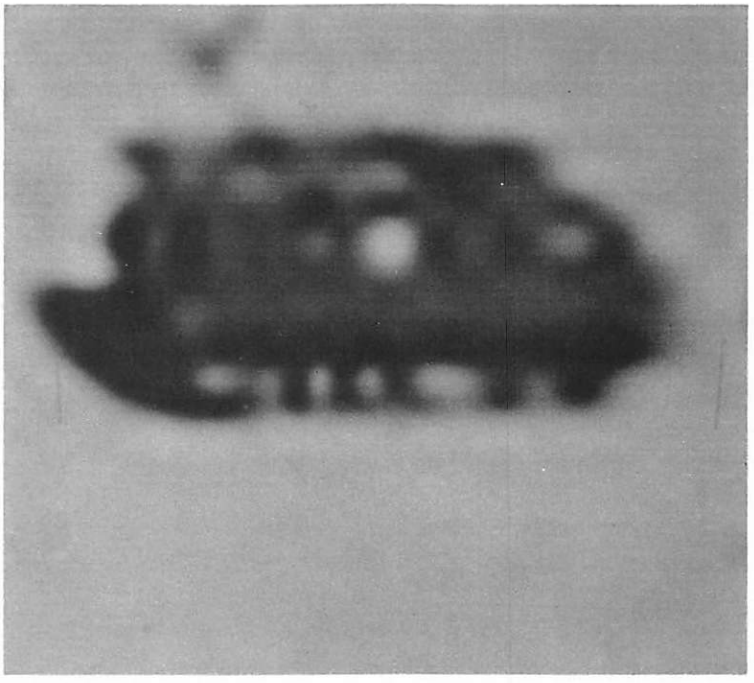


Figure 8. Bicubic Spline Approximation for APC with 40 Knots in X and Y Directions Placed by 4th Differences on X and Y Projections

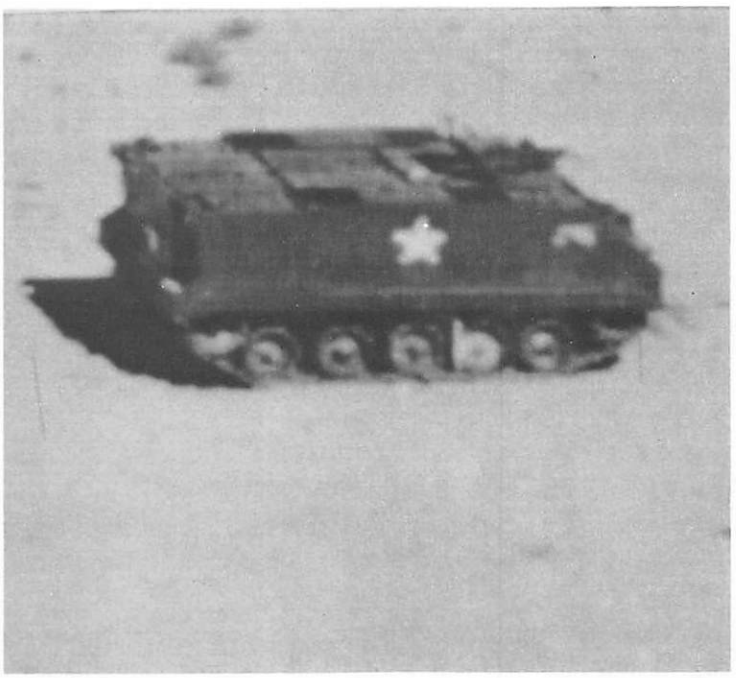


Figure 9. Original APC Image

knots placed according to the fourth difference of the projection data in figure 8. Figure 9 is the original. Table 2 contains the respective residuals. The respective knot placements along with the row and column projections are shown in figures 10 and 11.

Conclusions

In this paper the attempt has been to define the concept of an "epsilon degrees of freedom" of an image and to estimate that degrees of freedom by the number of variable knot bicubic splines necessary to approximate the image of an error level epsilon. Considerable success was achieved for the analytical image $e^{-50(x^2+y^2)}$ where 10 knots, each in the x and y directions were placed by an algorithm dependent on the fourth partials of the projection, and provided an error reduction of three orders of magnitude over the situation where each of the knot sets were uniformly spaced in the x and y directions. The error was such that 10 knots placed in the above manner provided an error lower than that achieved by placing 20 knots uniformly in the x and y directions. For an actual image the results were not so good - mainly a result of the difficulties in numerically obtaining the fourth derivatives. However, while the residual sum of squares is greater for the variable knot case than the uniform knot case more detail is perceptable in the variable knot bicubic spline approximation.

Splines were chosen as the expansion functions owing to their relatively attractive computation and properties and it is felt that with further research an efficient and accurate knot placement algorithm can be developed which would provide better results for real

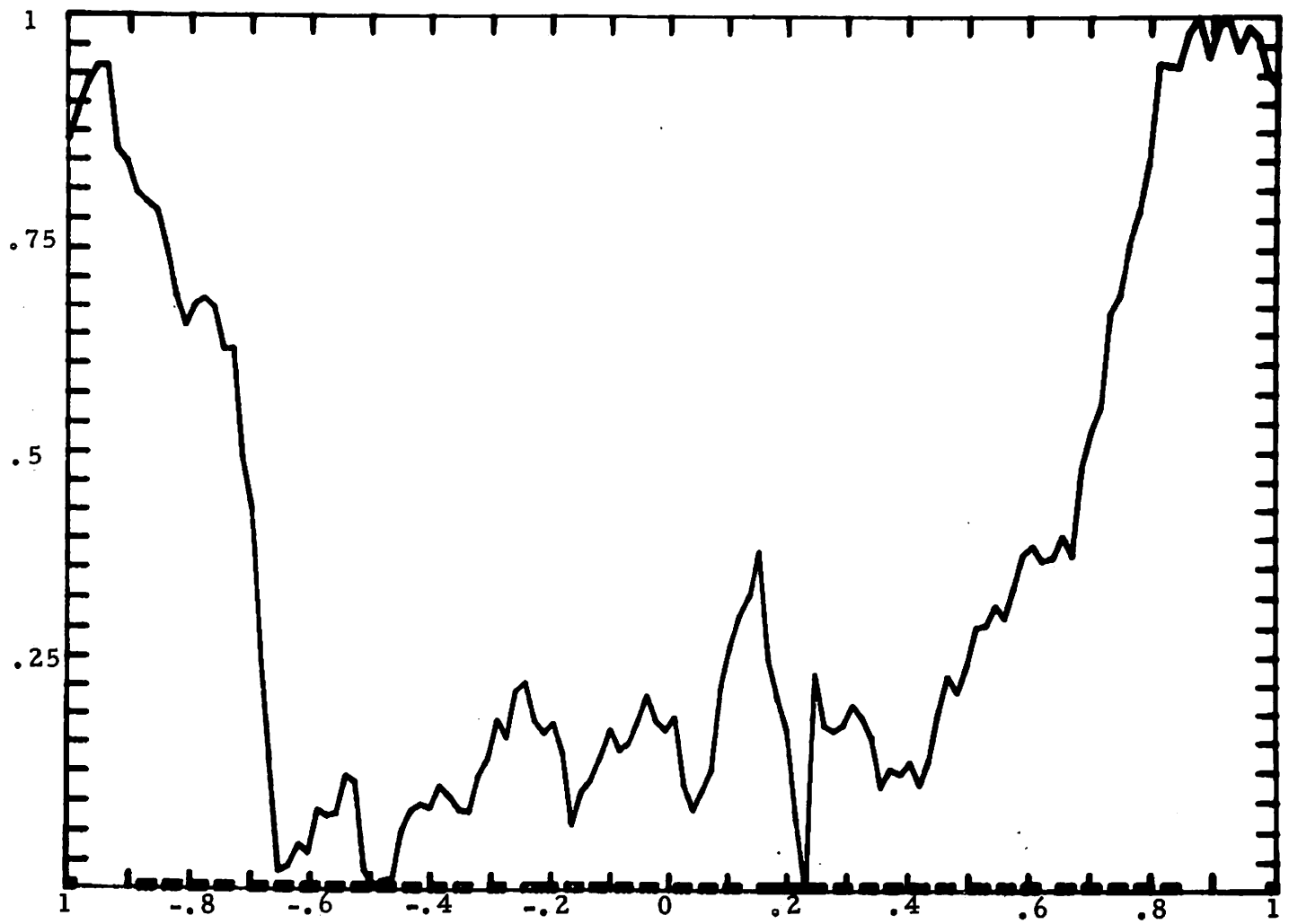


Figure 10a. Knot Placement from 4th Differences on Column Projections, 40 Knots

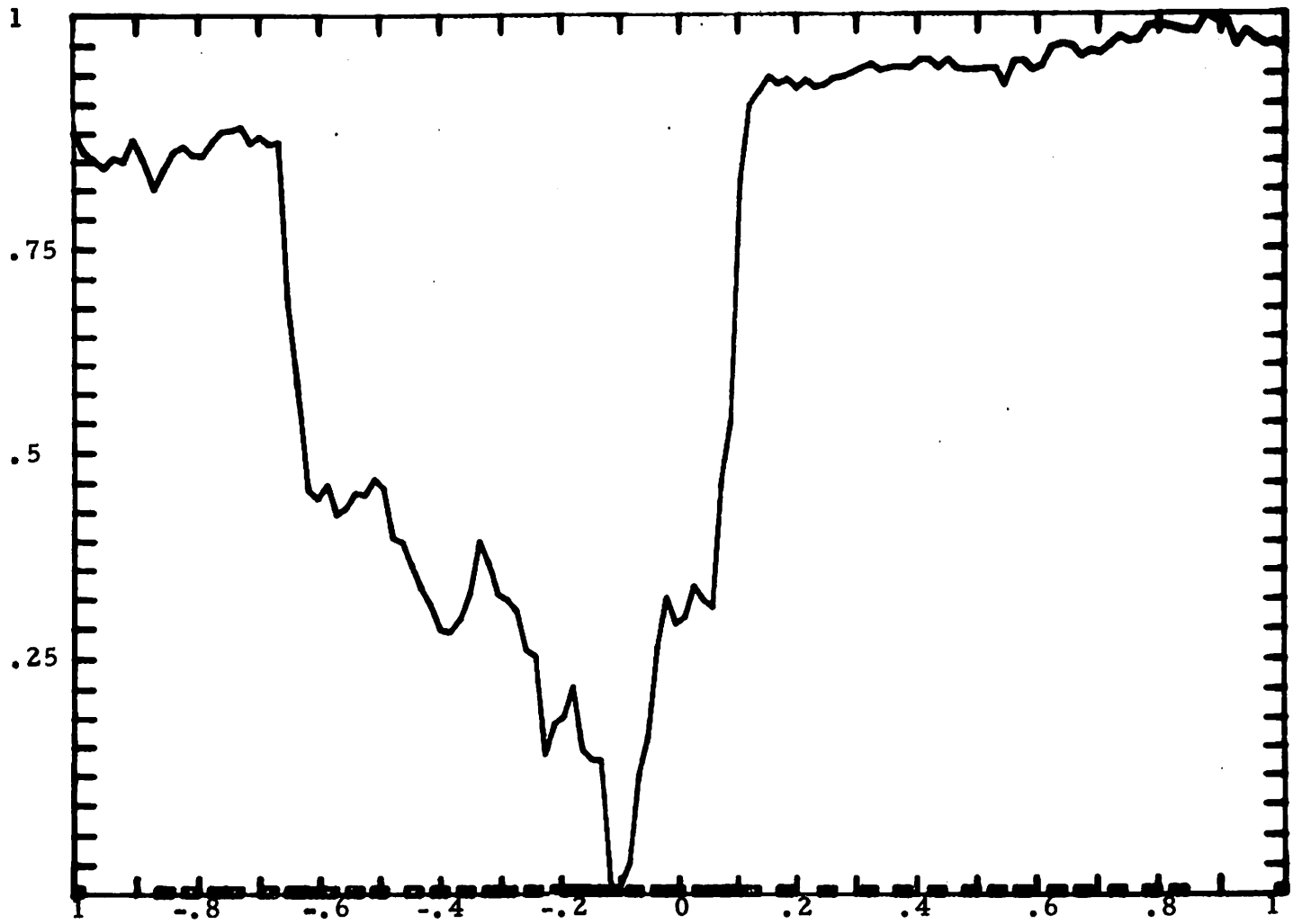


Figure 11a. Knot Placement from 4th Differences on Row Projections, 40 Knots

images than would be obtained for uniform knot bispline approximations.

References

1. S. Twomey, "The Application of Numerical Filtering to the Solution of Integral Equations Encountered in Indirect Sensing Measurements," Journal of the Franklin Institute, Vol. 279, 1965, pp. 95-109.

2. D. McCaughey and H.C. Andrews, "Degrees of Freedom and Projection Imaging," to appear in IEEE Transactions on Acoustics, Speech, and Signal Processing.

3. M. Cannon, "Blind Deconvolution of Spatially Invariant Image Blurs with Phase," IEEE Transactions on Acoustis, Speech, and Signal Processing, Vol. ASSP-24, No. 1, 1976, pp. 58-62.

4. H.C. Andrews and C.L. Patterson, "Singular Value Decompositions and Digital Image Processing," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-24, No. 1, 1976, pp. 26-53.

5. H.J. Landau and H.O. Pollak, "Prolate Spheroidal Waveforms, Fourier Analysis, and Uncertainty - III: The Dimension of the Space of Essentially Time-and-Band-Limited Signals," Bell Systems Technical Journal, Vol. 41, No. 4, 1962, pp. 1295-1336.

6. D. Slepian and H.O. Pollak, "Prolate Spheroidal Waveforms, Fourier Analysis, and Uncertainty - I," Bell Systems Technical Journal, Vol. 40, No. 1, 1961, pp. 43-64.

The majority of restoration techniques assume knowledge of $h(x,y,\alpha,\beta)$ and estimate $f(x,y)$. Yet, in most cases one would not have knowledge of $h(x,y,\alpha,\beta)$. Thus, the technique under study estimates both $h(x,y,\alpha,\beta)$ and $f(x,y)$ solely from $g(x,y)$ while assuming the following:

1) the PSF is spatially invariant, i.e., $h(x,y,\alpha,\beta) = h(x-\alpha,y-\beta)$,

2) the extent of the PSF is small compared to the extent of the image, and

3) the image is not so severely blurred such that one cannot tell the general class, for example, building, outdoor scene, etc., to which the blurred image belongs.

Consider the Fourier transform of h

$$F\{h\} = H(u,v)$$

Note in general that H is complex and can be represented in magnitude phase form as

$$H(u,v) = |H(u,v)|\, e^{;\theta(u,v)}$$

Thus, if one knows $|H(u,v)|$ and $\theta(u,v)$, one in effect knows $h(x,y)$ via the inverse Fourier transform relationship. Previous work in this area was termed "blind deconvolution" <1-4> and has had some moderate successes. The work of Cole <1> estimated $|H(u,v)|$. Cannon <2-4> extended this work by estimating $|H(u,v)|$ and estimating $\theta(u,v)$ for a set of three common blurs. The estimate of $\theta(u,v)$ was accomplished by assuming the blur was one of three possible blurs. Using a pattern recognition technique Cannon could discriminate between the three possible blurs and effectively estimate $\theta(u,v)$. The present work does

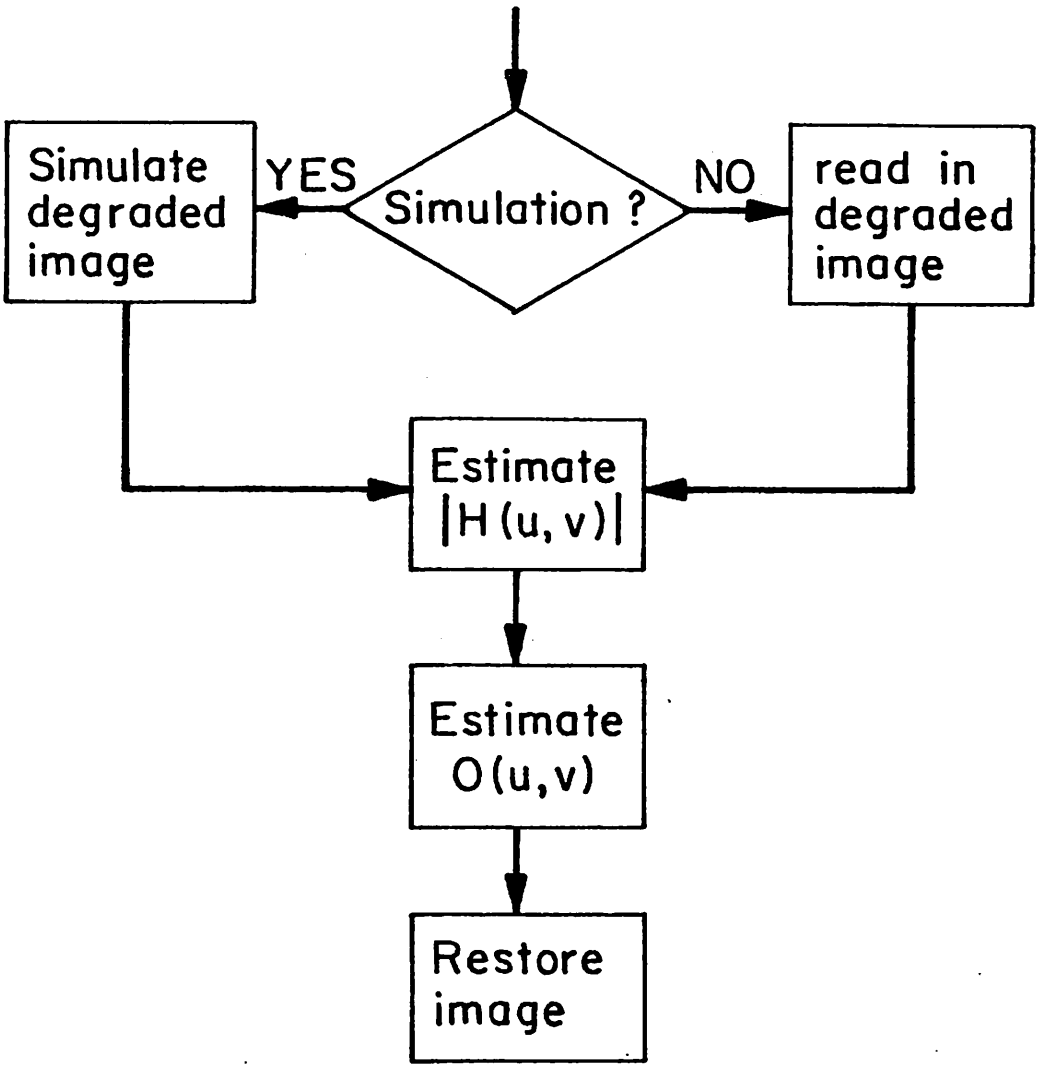


Figure 1. A posteriori restoration schematic.

Substituting $G_i(u,v)$ from equation (1) into equation (2)

$$R_G(u,v,\Delta u,\Delta v) \approx \langle H(u,v)F_i(u,v)H^*(u+\Delta u,v+\Delta v)F_i^*(u+\Delta u,v+\Delta v)\rangle$$

$$= H(u,v)H^*(u+\Delta u,v+\Delta v)\langle F_i(u,v)F_i^*(u+\Delta u,v+\Delta v)\rangle \qquad (3)$$

$$= H(u,v)H^*(u+\Delta u,v+\Delta v)R_F(u,v,\Delta u,\Delta v);[\theta(u,v)-\theta(u+\Delta u,v+\Delta v)]$$

$$= |H(u,v)|\,|H^*(u+\Delta u,v+\Delta v)|\,e \qquad R_F(u,v,\Delta u,\Delta v)$$

From equation (1)

$$|H(u,v)|\,|H^*(u+\Delta u,v+\Delta v)| \approx \frac{|R_G(u,v,\Delta u,\Delta v)|}{R_F(u,v,\Delta u,\Delta v)|} \qquad (4)$$

Combining equations (3) and (4) and rearranging results,

$$e^{;[\theta(u,v)-\theta(u+\Delta u,v+\Delta v)]} \approx \frac{R_G(u,v,\Delta u,\Delta v)}{|R_G(u,v,\Delta u,\Delta v)|}\,\frac{|R_F(u,v,\Delta u,\Delta v)|}{R_F(u,v,\Delta u,\Delta v)}$$

and it follows that

$$\theta(u+\Delta u,v+\Delta v) \approx \theta(u,v)-\tan^{-1}\left\{\frac{R_G(u,v,\Delta u,\Delta v)}{|R_G(u,v,\Delta u,\Delta v)|}\,\frac{|R_F(u,v,\Delta u,\Delta v)|}{R_F(u,v,\Delta u,\Delta v)}\right\}$$

Assuming one knows R_G and R_F and letting $\theta(0,0) = 0$, one may obtain $\theta(u,v)$ by the above recursive relationship. Note, however, that one can calculate R from the degraded image but one does not know R a priori. Thus to estimate (u,v) one must estimate R . It has been assumed that the image is not so severely degraded that one cannot tell the general class to which it belongs. As a result, the estimate of R will be calculated by using a similar clear image. Thus, the success of the phase estimation technique under study is contingent on the similarity of R for different images within a given class.

Progress to Date

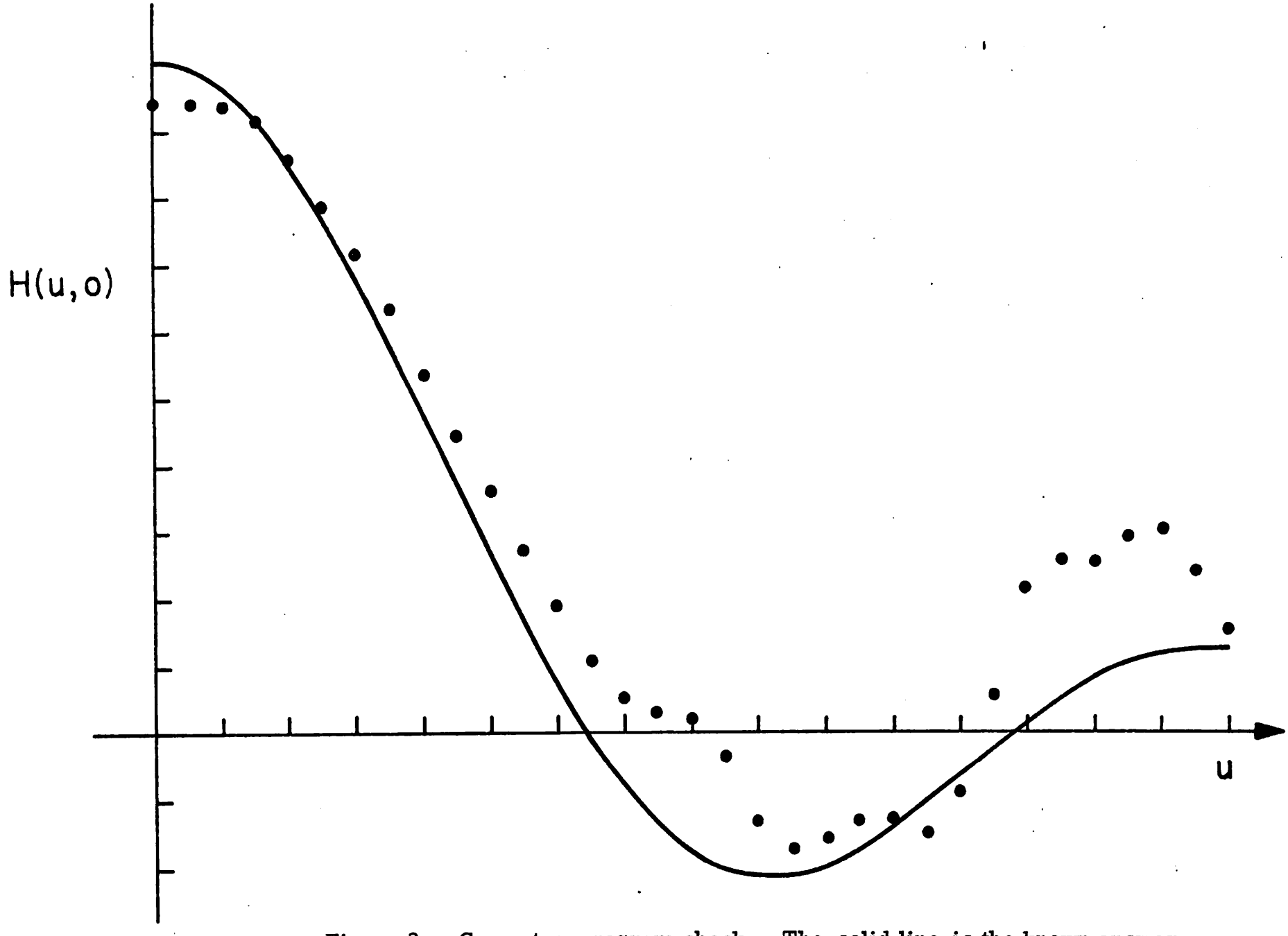


Figure 2. Computer program check. The solid line is the known answer. The dots are the results obtained via the computer programs.

respectively. Note in both figures 2 and 3 that the estimated values are approximately correct and in general the error increases for increasing u and v. It is believed that the error is due to the estimation procedure and not the computer programs. Sources of error include the discretization of a continuous process, limitations inherent in using eight bit/pixel images, and inaccuracies in using the discrete Fourier transform to approximate the continuous Fourier transform.

References

1. E.R. Cole, "The Removal of Unknown Image Blurs by Homomorphic Filtering," Department of Computer Science, University of Utah, ARPA Technical Report UTEC-CSc-74-029, June 1973.

2. T.M. Cannon, "Digital Image Deblurring by Nonlinear Homomorphic Filtering," Department of Computer Science, University of Utah, ARPA Technical Report UTEC-CSc-74-091, August 1974.

3. T.G. Stockham, T.M. Cannon, and R.B. Ingebritsen, "Blind Deconvolution through Digital Signal Processing," Proceedings IEEE, Vol. 63, No. 4, April 1975, pp. 678-692.

4. T.M. Cannon, "Blind Deconvolution of Spatially Invariant Image Blurs with Phase," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-24, No. 1, February 1976, pp. 58-63.

5. K.T. Knox and B.J. Thompson, "Recovery of Images from Atmospherically Degraded Short-Exposure Photographs," The

investigated. Comparison with FFT has shown a substantial improvement in efficiency and accuracy. Variations on the basic number theoretic transforms are also being developed.

It is the purpose of the on-going experiment to point out that the two-dimensional number theoretic transform may be defined as a simple extension of the one-dimensional number theoretic transform. For practical cases, two-dimensional finite impulse response filters involve impulse response of a lot fewer than ordinary picture size 256 x 256 or 512 x 512 in almost all cases. Since these number theoretic transforms have the difficulty that the word length required is proportional to the length of the sequence being convolved, it is conceivable that sectioning techniques may be of value when sequences must be convolved exactly, such as when multiplication is very costly, or when special-purpose hardware for modest length convolution is being considered. For the large picture mentioned above the convolution can be done by sectioning the picture by efficient overlap-add sectioning deleting and end effect. For $y(n) = x(n) * h(n)$, we assume that $x(n)$ is of long duration or that the number of samples representing $x(n)$ exceeds the size of the $h(n)$. Then, since the shift of the convolution $y(n)$ is simply the sum of the shifts of the functions being convolved, no information is lost if we shift each sectioned function to the origin prior to convolution. As a result, it is necessary to decompose $x(n)$ into sections and compute the discrete convolution as many smaller convclutions to avoid the problem of dynamic range of the number theoretic transform. At the same time, if N is small, then we would have the advantage that

systems. With the use of general amplitude and phase optical elements, many useful image processing operations can be done very easily with analog incoherent systems. If general optical elements combining amplitude and phase variation can be made, the only major restriction on incoherent systems is that the impulse response must be non-negative.

A very useful function suitable for an incoherent optical processing system is a B-spline interpolator to give a continuous desampled output from the discrete spots on a CRT or other display system. Other useful operations include Wiener filters, high spatial frequency emphasis filters, edge or directionally sensitive filters, or restoration filters.

The filtering operation of the optical system can be described by a complex amplitude point-spread function $\underline{h}(x)$ or non-negative intensity response equal to the magnitude-squared $|\underline{h}(x)|^2$. In the frequency or pupil domain, omitting scaling factors for simplicity, the desired complex pupil function $\underline{H}(f)$ is given by

$$\underline{H}(f) = \mathcal{F}\{\underline{h}(x)\} \tag{1}$$

where $\mathcal{F}\{\cdot\}$ denotes the Fourier transform operation. The incoherent optical transfer function $\mathcal{H}(f)$ is then described by

$$\mathcal{H}(f) = \underline{H}(f) * \underline{H}(f) = \mathcal{F}\{|\underline{h}(x)|^2\} \tag{2}$$

using transform theory.

With these definitions, the problems and tradeoffs of optical processing can be defined. Given a desired $\underline{h}(x)$ or $|\underline{h}(x)|^2$, it has

recursive searching routine to reduce the dynamic range, complexity, and object error. Although this numerical searching is a time-consuming process in general, it does show better results than using a diffuser alone. In the application of analog filtering, the hologram is computed only once and the computation time is no longer a serious problem as long as good impulse response has been plotted.

One phase coding method was first developed at IBM <1> for the coding of kinoforms. Kinoform errors due to truncation and quantization are discussed by Gallagher and Liu <2>. In this method, the phases of object are assigned randomly and the object is then Fourier transformed. The transform is set to have a constant amplitude and is then inversely transformed. The resultant phase is used as the new phase for the object. This procedure is repeated until an error criterion is met.

Although this method generally gives good results, it has been found <3> that the error at some points fails to decrease, stabilizing at a high level. This prevents the error from being arbitrarily small. To overcome this deficiency, Fienup <3> adopts an Input-Output Approach, which allows the amplitude of the input as well as the phase to be modified. This gives better control of the reconstruction. In this model if certain output is used as input then the same output is obtained. Hence modification can be either added to the previous input or present output. We will refer to these two methods as Fienup I and Fienup II, respectively.

root-mean squared intensity error <4> is used. Since the actual quantity measured is the intensity of the image, the latter one might be a better criterion although it is larger in general. We can also argue that since human eyes are sensitive to relative intensity other than absolute intensity, it would be better tc normalize the above quantities so that these quantities are not affected by the absolute intensity and, hopefully, the specific object used. The error criterion adopted in this study is the following for an N x N picture

$$E = \left[\frac{1}{N^2} \sum_{M,N=0}^{N-1} \left(|a_{mn}|^2 - \tau_{mn}^2 \right)^2 \right]^{\frac{1}{2}} / \tau_{max}^2 \qquad (3)$$

where $\hat{a}_{mn}$, τ_{mn} are the resultant image and the desired image, respectively and

$$\tau_{max} = \operatorname*{Max}_{m,n} \tau_{mn} \quad . \qquad (4)$$

One dimensional objects are used in the experiment for simplicity. Only one of them will be illustrated here. Figure 1 shows one line of a picture. This is a 256 point array imbedded in a 512 points array in order to smooth the transform and thus the kinoform. Figures 2, 3, and 4 show the resultant images and kinoforms after 34 iterations by the methods of Hirsch <1>, Fienup I and Fienup II <4>. Compared with figure 1, we found that Fienup II is slightly better but almost the same as the first method. Fienup I, however, gets nearly perfect reconstruction in the region of the object while more noise is introduced elsewhere. Since intensity is measured in practice, this noise in amplitude contributes less in intensity.

The corresponding errors evaluated by eq. (1) are given by

$$E_1 = .05222 \quad E_2 = .01297 \quad E_3 = .04449$$

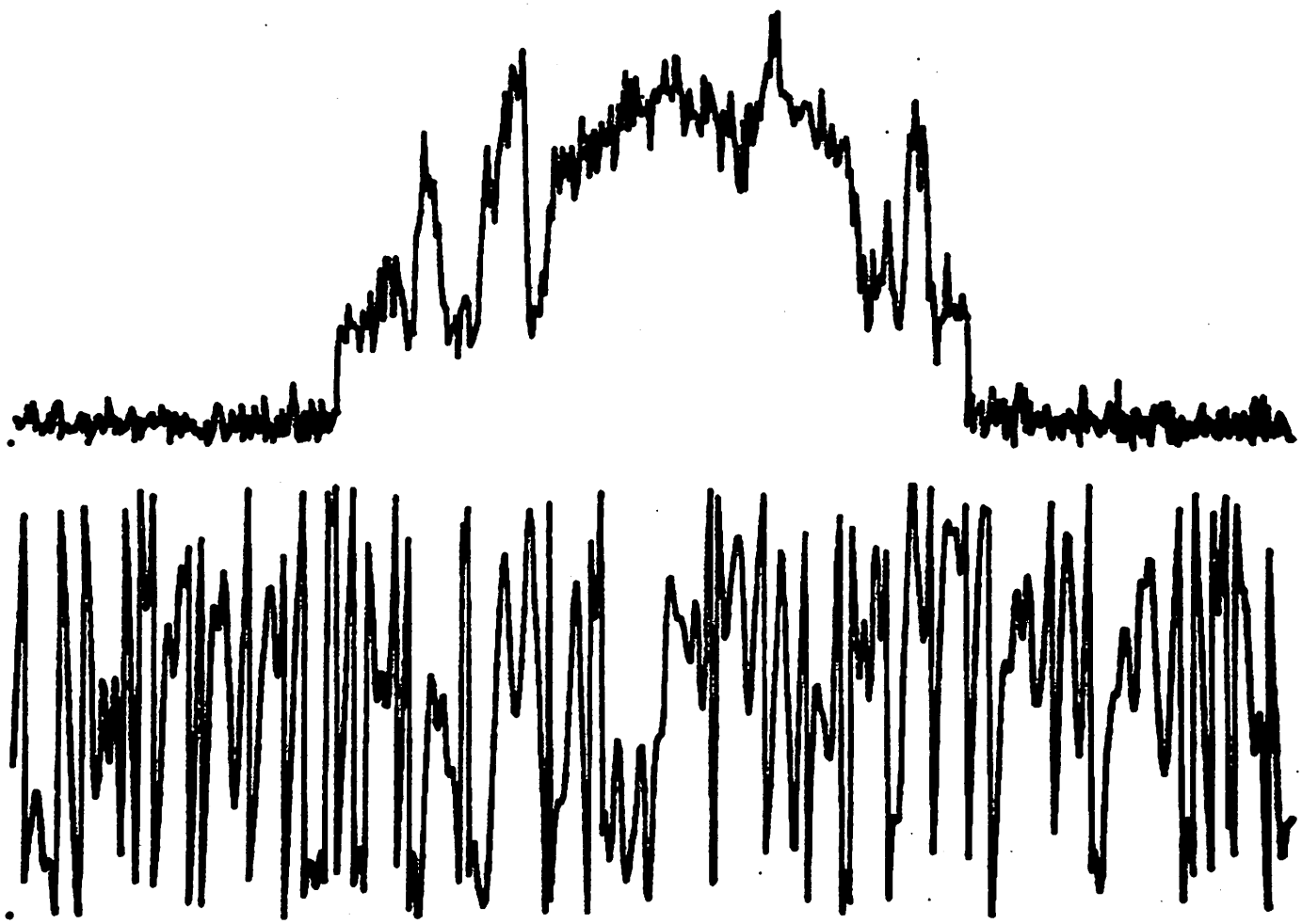


Figure 2. Image and Kinoform by using Hirsch's method

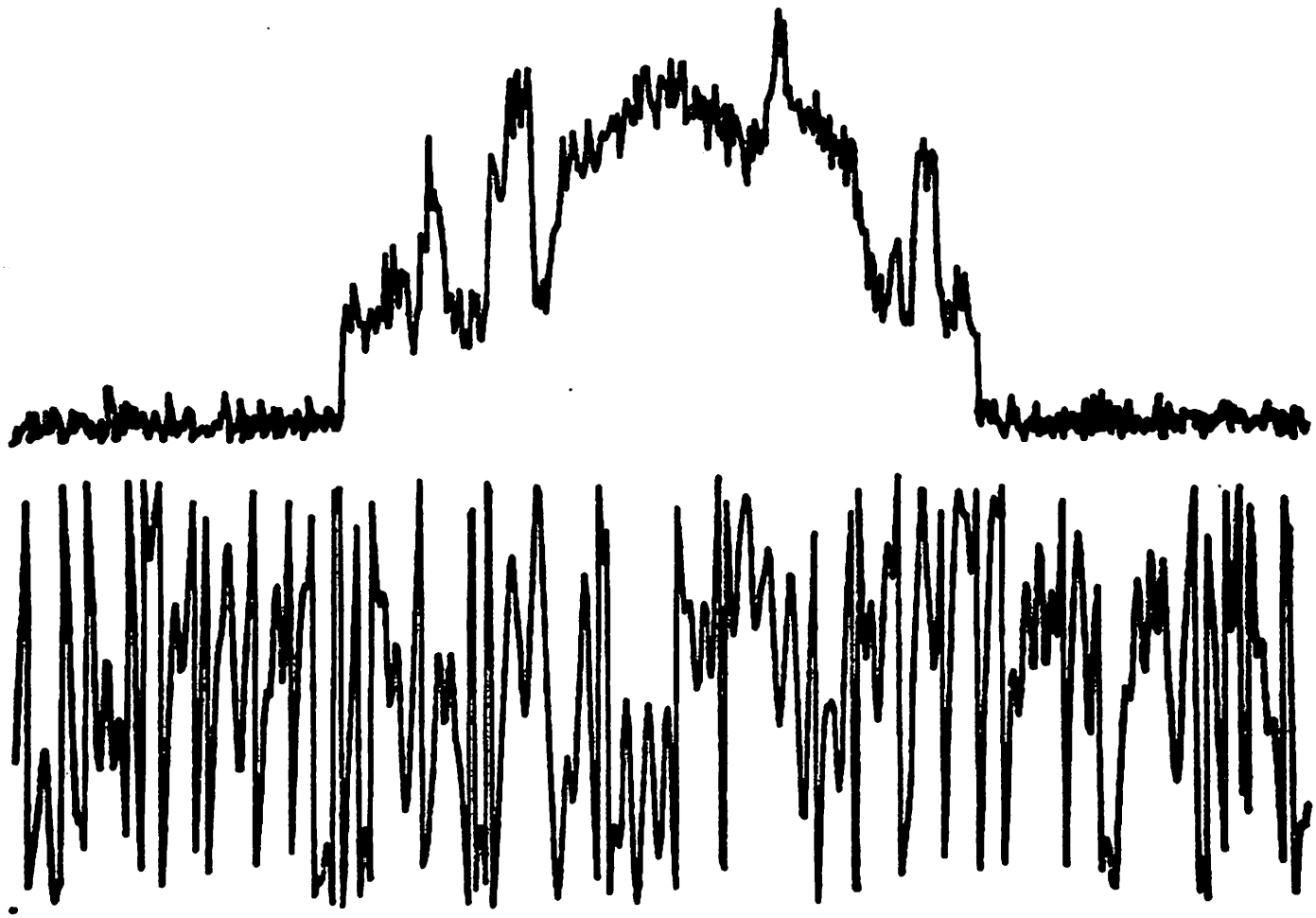


Figure 4. Image and Kinoform by using Fienup II

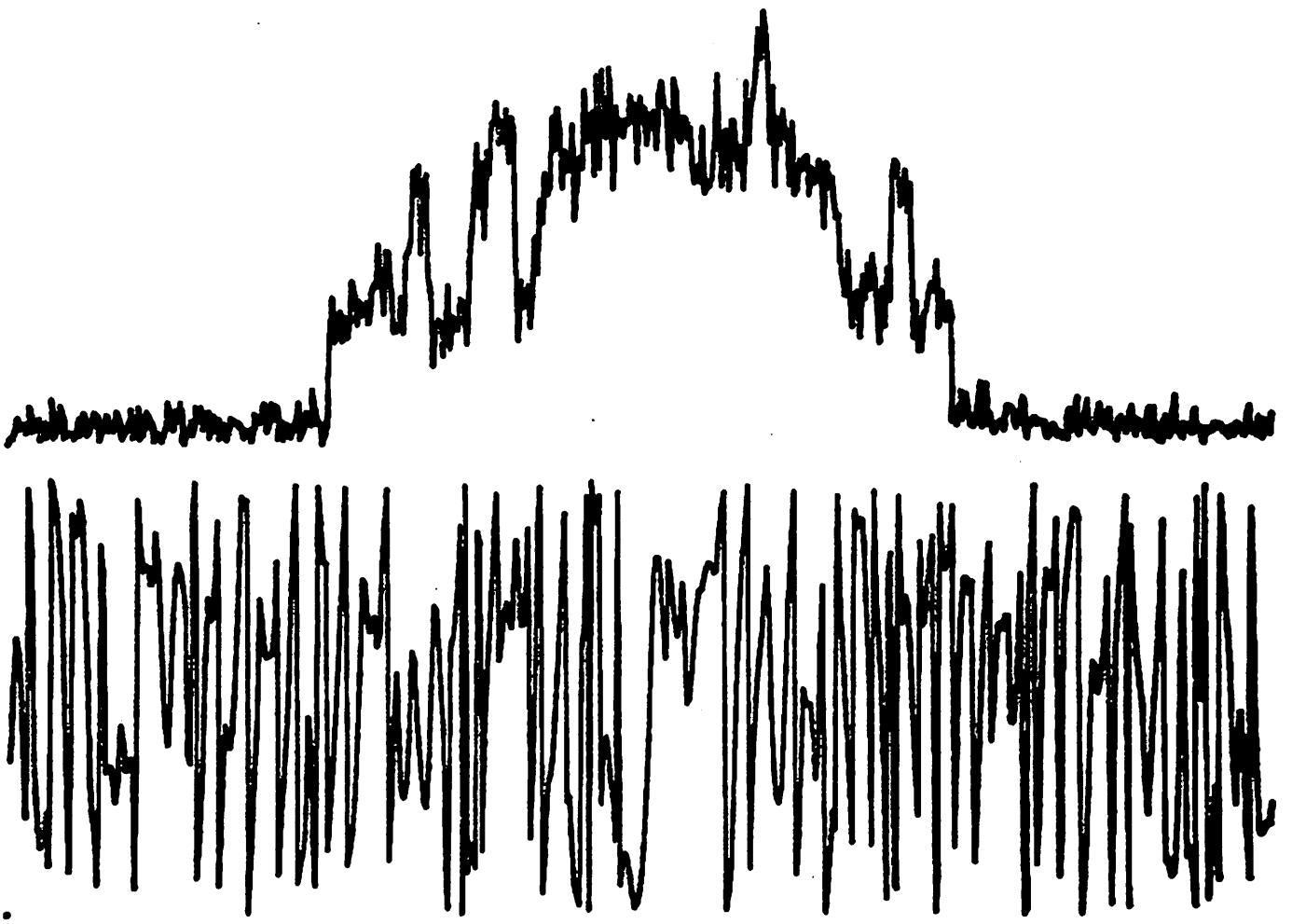


Figure 5. Image and Kinoform by using cosine function as the amplitude of the Kinoform

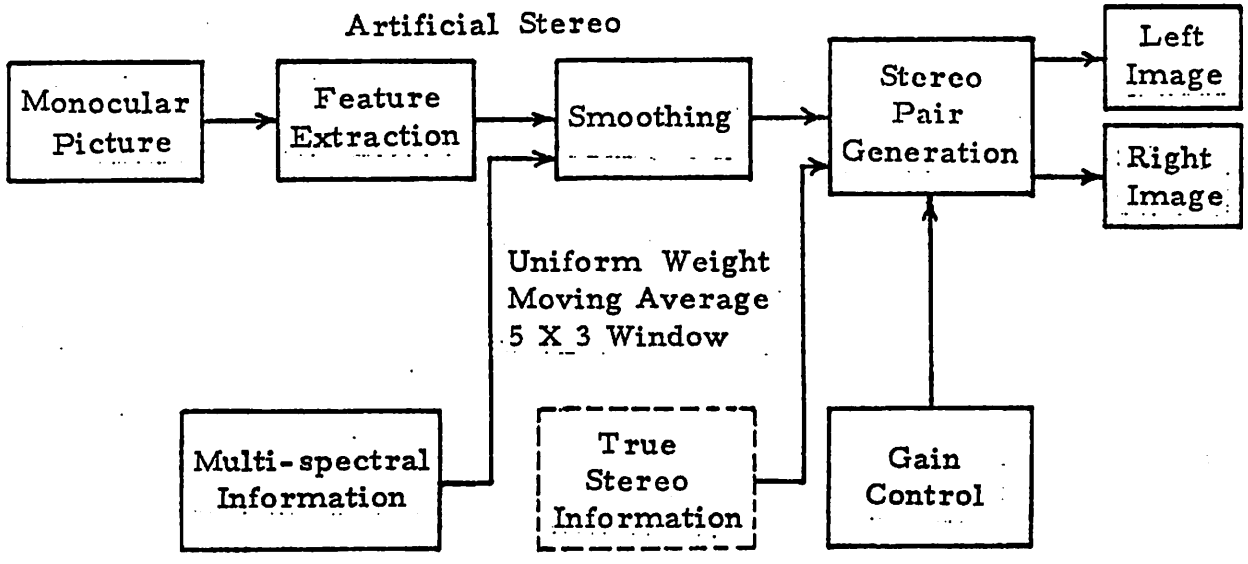


Figure 1. Artificial stereo block diagram.

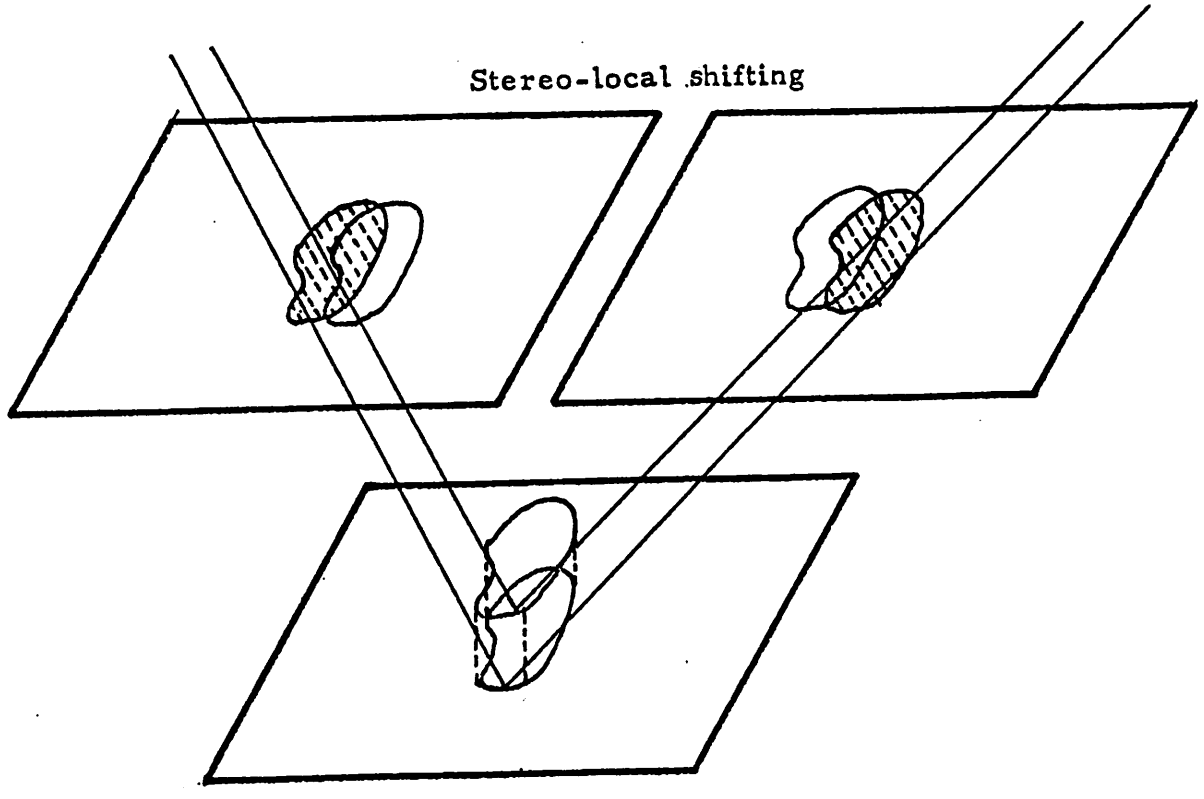


Figure 2. Local shifiting for depth illusion.

The second graph is then made monotonic by a curve-following algorithm, as shown in the bottom graph of figure 3. In the input regions of zero slope, the pixels are obliterated by adjacent ones moved in. In regions of large positive slope, certain output areas may be missed and left blank. When this happens, pixels are moved in, beginning from left to right, to fill in the blank areas. Thus, a local distortion takes place to create the local shifting. The directions for the right eye are simply opposite.

Future work on artificial stereo will concentrate on the case of different features and algorithm improvement. It has been recently found that the addition of a brightness histogram equalizing step in the block diagram of figure 1 just before the stereo pair generation gives a much better distribution of heights and a more pleasing stereo illusion.

Reference

1. A.A. Sawchuk and H.C. Andrews, "Artificial Stereo from Image Features," USC Image Processing Institute Semiannual Technical Report, 1 September 1975 - 31 March 1976, USCIPI Report 660, March 1976, pp. 195-199.

4.0 SMART SENSOR PROJECTS

This section presents the results of the industrial liaison between USCIPI and Hughes Research Laboratories. The objectives of the past year's subcontract were twofold: first, to develop real time nonlinear optical data processing elements and second, to develop a CCD chip for smart sensor image understanding signal processing. Both objectives have been met, in part, with quite exciting results promised in the latter project. All projects are reported herein. The construction of the CCD chip is on schedule with completion, test, and evaluation expected within six months.

which is designed to detect vertical edges and then designed to detect horizontal edges, the two outputs of the vertical and horizontal detectors then being combined by an absolute value operation to form the Sobel edge output. The absolute value operation is in lieu of a square rooting operation. This edge detector is a fairly useful one, requires only a small kernel; and in addition, contains nonlinear devices (i.e. the absolute value operations). The reason for selecting this edge operator is because of the nonlinearities involved as well as the possibilities for future implementation by looking at the ratio of the horizontal to the vertical edge energy. This will provide edge direction which also becomes a useful feature technique to be utilized in higher level image understanding operations. The Sobel operator can be viewed in image form and should provide a fairly clean edge detected image. References on this operation are readily available in the literature.

The mean operator is essentially nothing more than a low pass filter. A 3 x 3 matrix was suggested for implementation as a convolving kernel. The mean operator measures the average brightness in a 3 x 3 region surrounding the center pixel. This operator, as it stands alone, will tend to blur the image (remove high frequency noise as well as removing edge detail). The operator is suggested as possibly a signal processing low pass filter but more significantly as a switching signal to allow for adaptive signal processing in subsequent operations. Because the mean operator tends to track the low frequency variation in an image, it becomes very useful for discriminating between regions of high brightness and regions of large

convolving kernel, to either completely on or completely off based upon whether that center pixel brightness is greater than or less than the mean average of its surrounds. This particular operation has the advantage of providing a binary video signal in the context of certain processes or has the advantage of providing binary adaptive thresholded signals as illustrated in adaptive binarization for fingerprint identification. The motivation in the latter application is in use of the a priori knowledge that the original fingerprint is best viewed under binary circumstances. However, latent fingerprints, when digitized, have a great many shades cf gray and the adaptive binarizer has the property of tracking low frequency smudge contents and allowing edge and fingerprint information to be enhanced by subtracting off or essentially switching out the low frequency information. I might also suggest that another application for this particular operator is in the generation of binary video. As I understand it, in a lot of the Hughes Aircraft forward-looking infrared work, binary video becomes a very useful signal in two dimensions upon which to subsequently correlate. This particular type of binarization is adaptive in the sense that the local mean is detected by the mean operator and then controls binarization by tracking that mean in the bright region by rising high, and in low regions by lowering the amplitude values. The possible drawback of this particular binarizing filter lies in the fact that the filter is only 3 x 3 in extent. Possibly, the 5 x 5 and even 10 x 10 adaptive binarizing filters might be more useful. This would require a larger array on the chip but possibly because all of the weights are

and whether the processed images themselves are pictorially and cosmetically more pleasing or not, is really not the point of the development of the chip under consideration. Rather, the selected five functions that have been suggested, themselves represent a large variety and great breadth of signal processing operations that conceivably could become quite useful on the front end of a sensor device. Therefore, it is possible that all the implementations will require larger kernel arrays or modifications of the signal processing operation described thus far. However, the important thing to emphasize is the fact that if these signal processing operations can indeed be implemented, and with the linear and nonlinear implications of the implementations so described, such implementation is a success even if the applications results are not striking and I feel that success in itself will motivate further investigation.

Appendix

I. There shall be two 3 x 3 convolving kernels on a chip.

kernel #1 shall be the Sobel operator, f

kernel 2 shall be the mean operator, f

II. There shall be various combinations of the outputs of the above kernels. These shall form five different output lines, and therefore output images.

1. Sobel operator = edge detection = $f_s(j,k)$
2. mean operator = low pass filter = $f_m(j,k)$
3. Sobel + α mean = unsharp masking = fu (j,k)

5. $$f_a(j,k) = \begin{cases} 2 \text{ Min } \left\{f(j,k), \frac{r}{2}\right\} & f_m(j,k) \leq \frac{r}{2} \\ 2 \text{ Max } \left\{f(j,k) - \frac{r}{2}\right), 0\right\} & f_m(j,k) > \frac{r}{2} \end{cases}$$

$f(j,k)$ = center pixel in 3 x 3 array to be processed

1 **Sobel operator = edge detection**

$$\begin{matrix} & j \longrightarrow \\ k \downarrow & \end{matrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$f_s(j,k) = \Big[\big| (a + 2b + c) - (g + 2h + i) \big| + \big| (c + 2f + i) - (a + 2d + g) \big| \Big] 1/8$$

output is center position given by $f_s(j,k)$

i.e. $$f_s(j,k) = \Big[\big| (f(j-1,k+1) + 2\, f(j,k-1) + f(j+1,k-1)) - (f(j-1,k+1) + 2\, f(j,k+1) + f(j+1,k+1) \big| + \big| (f(j+1,k-1) + 2 f(j+1,k) + f(j+1,k+1)) - (f(j-1,k-1) + 2f(j-1,k) + f(j-1,k+1)) \big| \Big] 1/8$$

2 **mean operator = low pass filter**

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

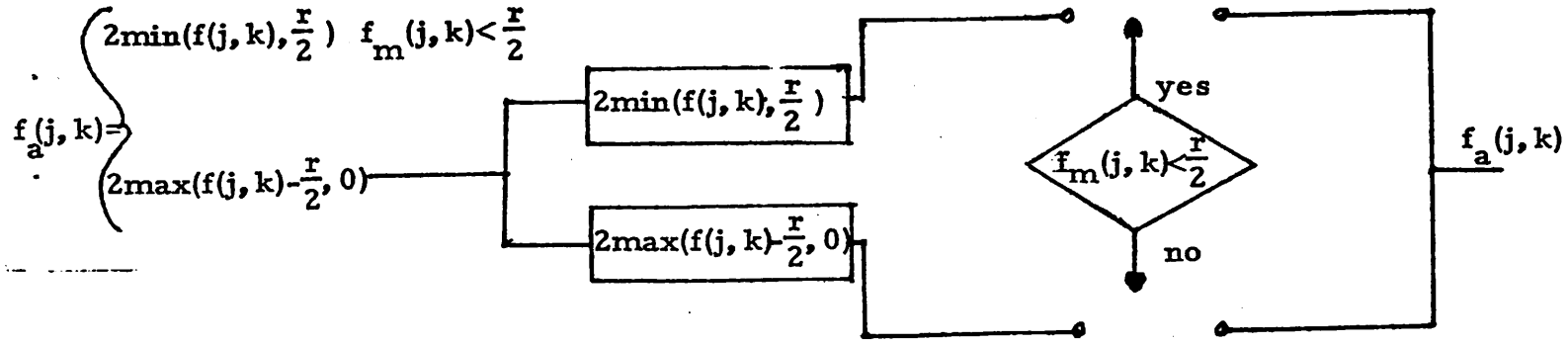
r = input dynamic range

$\frac{r}{2}$ = half input dynamic range

The above transfer curves are to be switched between - depending on whether

$$\frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix} = f_m(j,k) \quad \text{is}$$

greater or less than r/2. The center pixel is to be the input



4.2 CCD IMAGE PROCESSING CIRCUITRY

Graham Nudd

During the past year we have investigated the possibility of performing the following five algorithms with CCD/MOS integrated circuitry:

1. Chirp transformation
2. Robert's cross
3. Sobel operator
4. Hueckel operator
5. Histogram operation.

We have developed circuit concepts for each of the above (apart from the Hueckel operator) and include here block schematics together with a brief functional description. Details of the circuit design and layout are in general omitted, and only factors which directly affect the overall concept such as speed, dynamic range, etc. are discussed. We have also initiated work on the detailed design and layout of two

test circuits, one of which (the Sobel operator) is currently being fabricated. Details of these are also included.

At the present time we have concentrated our efforts on charge transfer technology and compatible MOS circuitry, because of its inherently low power-delay product 10^{-2} pJ (versus 50 pJ for typical bipolar) and its application to both image detection and processing. This will be particularly advantageous for the development of "Smart Sensors." A concept of an integrated detector and pre-processor is shown in figure 1. Here a CCD imaging array is shown as detecting the image and creating charge carriers which are then clocked through a matrix of CCD/MOS processing circuits, (which might, for example, perform a two dimension convolution of the Sobel type) and the processed data stored in a CCD analog store.

Chirp Transformation

A one dimensional CCD implemented chirp transform has been described in the literature <1>. Factors affecting the accuracy and speed of this process are described in the previous semi-annual report. The major issues are in the area of (a) integrating the peripheral circuitry such as clocking circuits and input/output devices on the chip, (b) providing integrated accurate ($\geq$6 bit) multipliers. Both these areas are being pursued by other contractors, and hence we have devoted little effort in this area. The major impact Hughes Research Laboratories could have in this area is to employ its high resolution electron beam exposure techniques to fabricate a high resolution ($\geq$8 bits) CCD multiplying digital to

analog converter <2> as previously discussed. Two dimensional Fourier Transform performed by concatenation of two one dimension transforms requires a permuting memory to transform the data. Work in this area is currently being sponsored by the Naval Undersea Center.

Robert's Cross Operator

A circuit which will perform the Robert's cross operator is shown schematically in figure 2. As shown it consists of two parallel CCD channels into which adjacent rows of picture elements are fed. The inputs (not shown in figure 2) might typically consist of the Tompsett surface potential equilibration circuits <3> or the CCD circuit might also be used as the sensor. The charge corresponding to the four adjacent pixels (a through d) used in the Robert's cross are simultaneously sensed by floating gate electrodes <4>. These outputs are used to drive a nonlinear CCD circuit which performs the magnitude operation $|a-d| + |b-c|$, required for the Robert's cross operation.

A schematic of the circuit is shown in figure 3. The two signals corresponding to pixels "a" and "d" are connected to parallel gates, as shown, and create potential wells beneath the oxide proportional to their magnitude. The source diffusions are then pulsed such that charge flows across the barrier ϕ_1 and into both wells. For surface channel devices the surface potential ϕ_s can be written as

$$\phi_s = V_o + V - \sqrt{V_o^2 + 2VV_o} \qquad (1)$$

where $V = V_g - V_{FB} - Q/Co$,

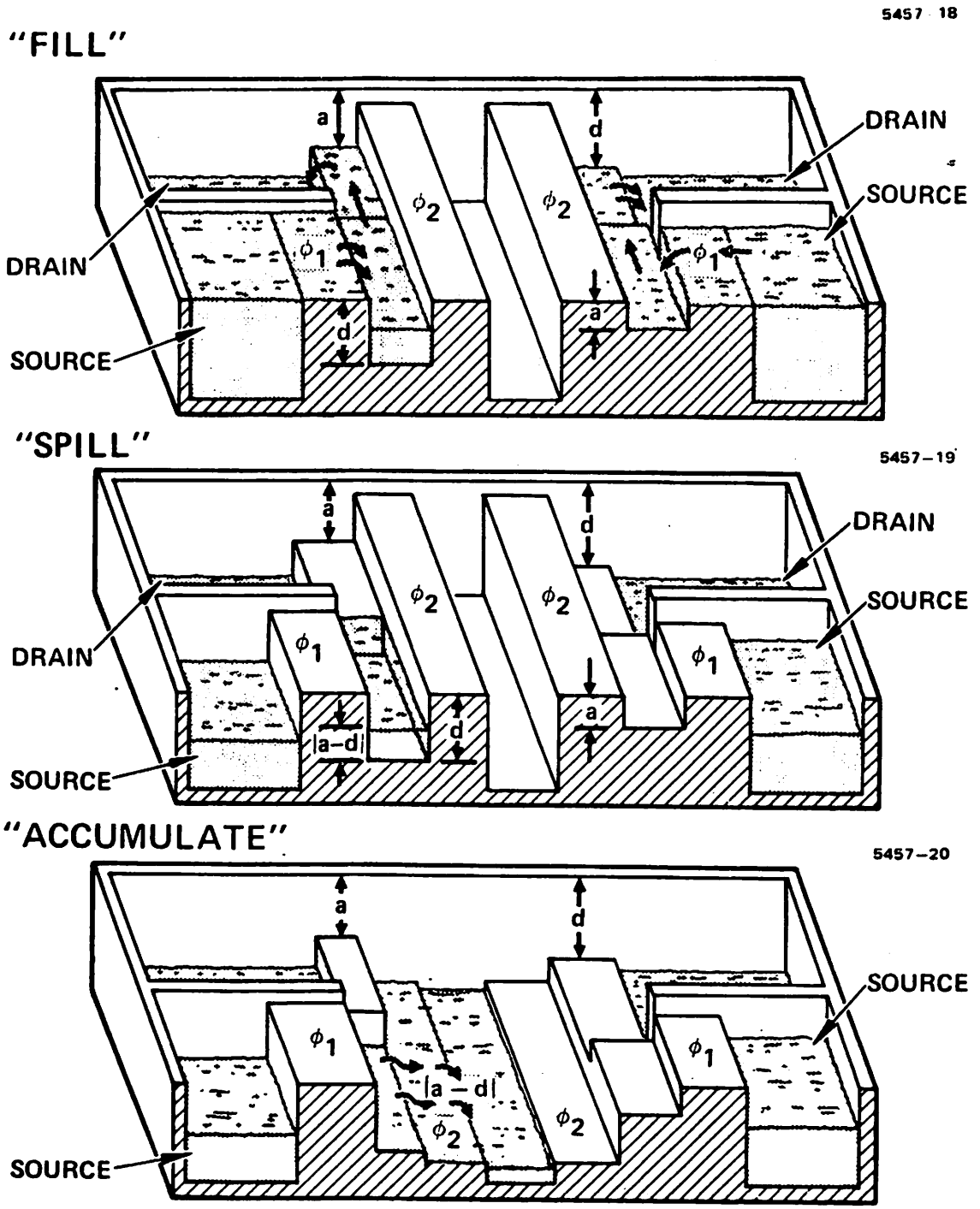


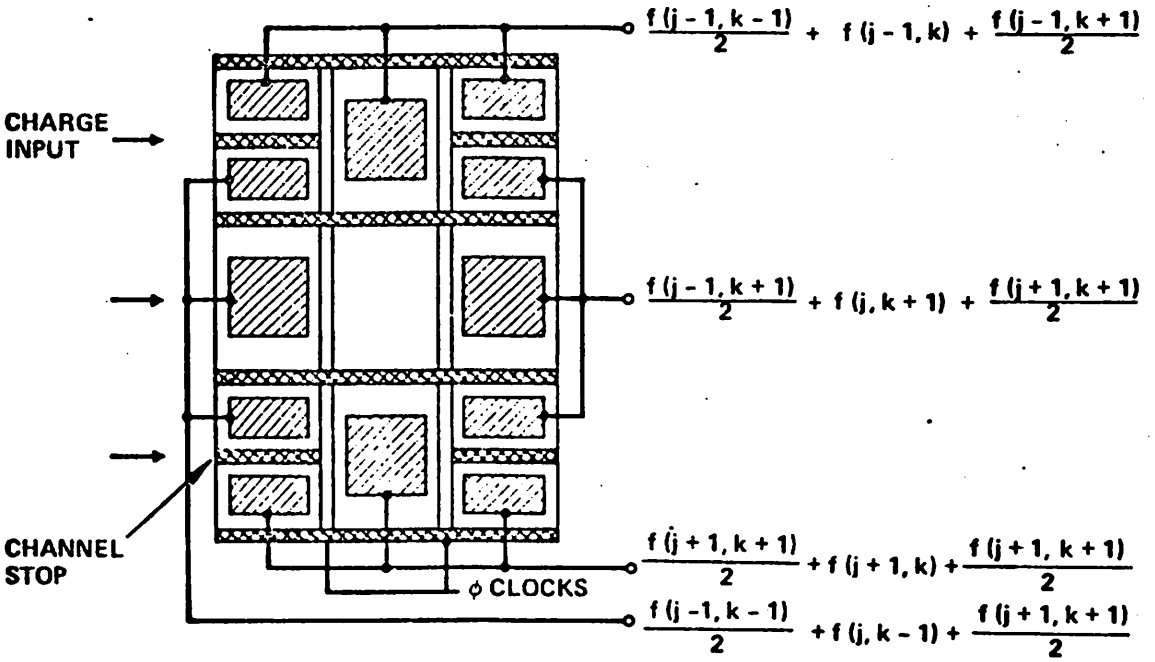
Figure 3. Schematic of nonlinear CCD circuit to evaluate $|a-d|$.

5457–15

(a)

f(j−1, k−1)	f(j−1, k)	f(j−1, k+1)
f(j, k−1)	f(j, k)	f(j, k+1)
f(j+1, k−1)	f(j+1, k)	f(j+1, k+1)

3 x 3 ARRAY OF PIXELS



(b) WEIGHTED SUMMATION TECHNIQUE REQUIRED FOR SOBEL OPERATOR

Figure 4. Weighting technique for Sobel Operator.

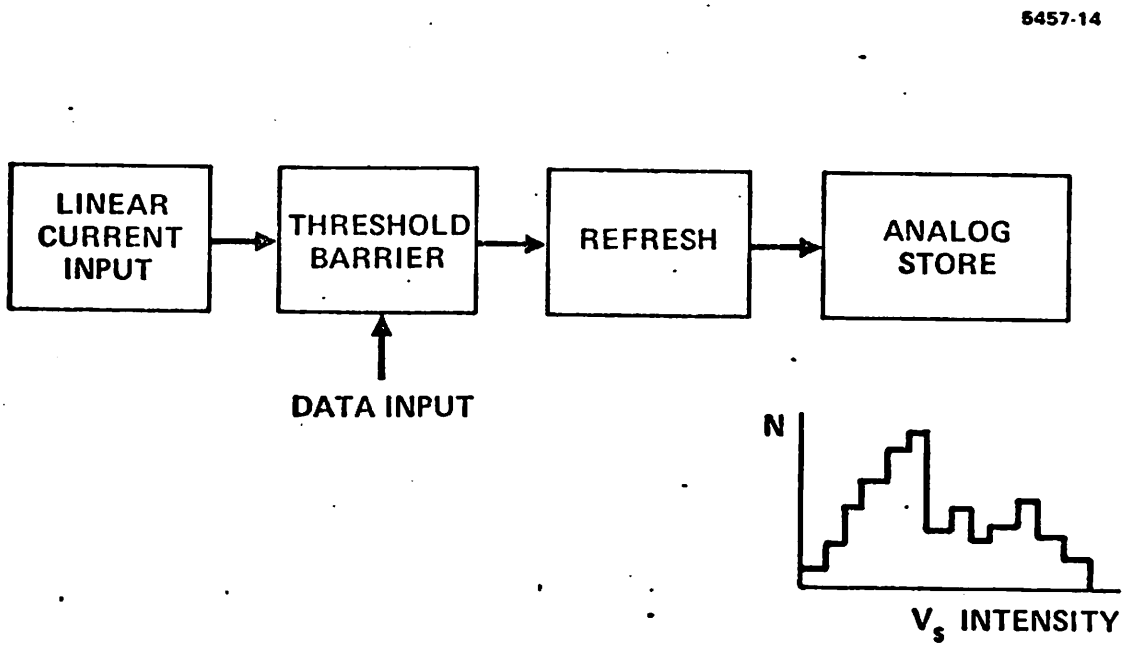


Figure 5. Block schematic of histogram circuit.

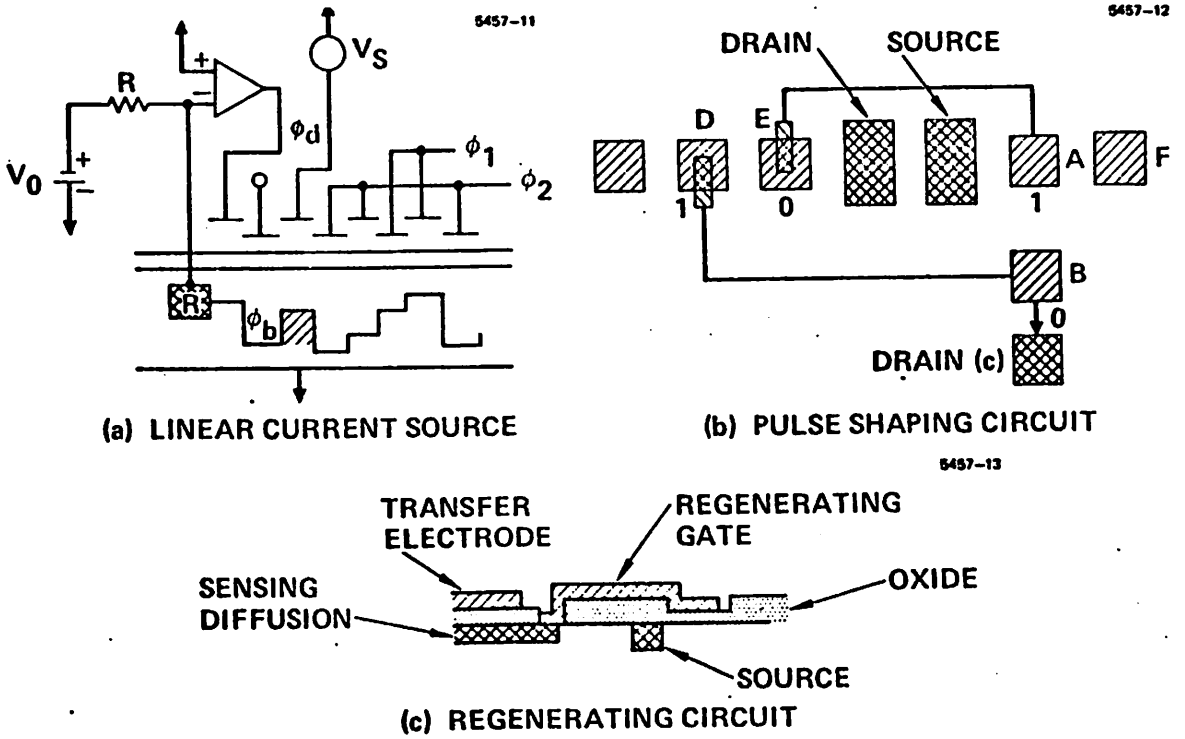


Figure 6. Circuit elements required for histogram circuit.

operate at approximately 25 MHz, which is about the present state of the art. Data from this register is shifted out in parallel to the analog store (figure 7) which accepts data at the pixels rate, 100 kHz. The total storage time is practically limited to about 0.1 seconds by dark current, limiting the operation to a total of about 10^4 samples.

The pulse-shaping circuit illustrated in figure 6b and 6c must be included to provide a logical one to the input of the digital shift register. As shown, the circuit is based on digital refresh described by Kosnocky and Carnes <5>. After charge flows over the barrier ϕ_b a string of logical ones with charge equivalent to approximately 10^5 electrons will flow. The circuit uses floating gate sensors to provide a single pulse output coincident with the start of this string. When data under gates D and E are both "zeros" the refresh circuit will create potential wells under both gates A and B. This will cause data to flow to drain C. When the input data changes (from zero to one) the potential well under gate B disappears and the charge drawn from the source is clocked through to F prior to entering the CCD digital delay. When both D and E are "ones" the potential under gate A rises and no charge flows thus providing a single data output at a time equivalent to the start of current flow. A cross section of one of the gate interconnects is shown in figure 6c together with the time constants involved. Present photolithographic techniques indicate that this circuit will operate with delays of less than a few nanoseconds.

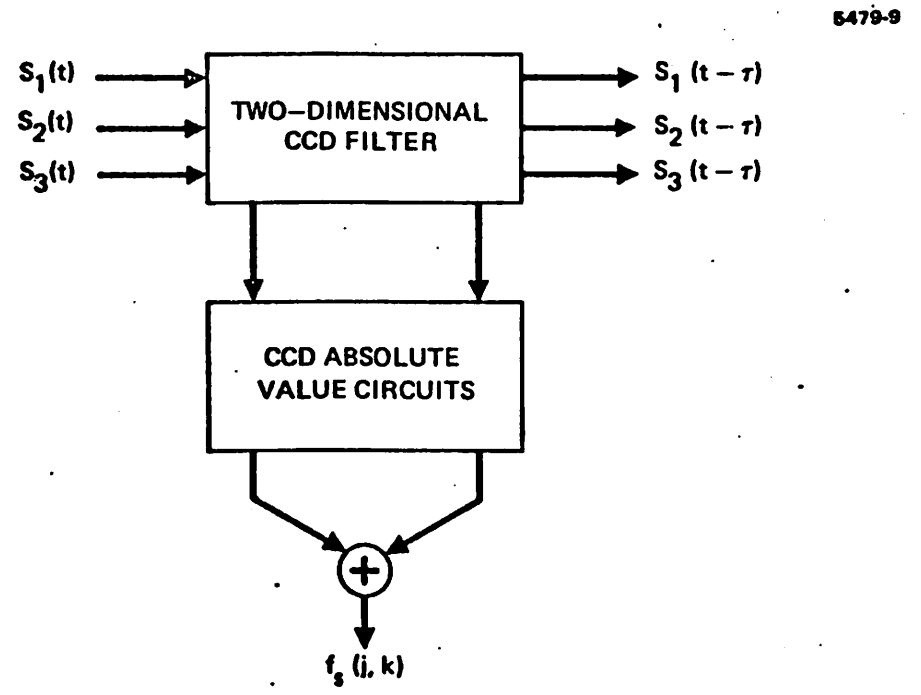


Figure 8. Schematic of Sobel Operator.

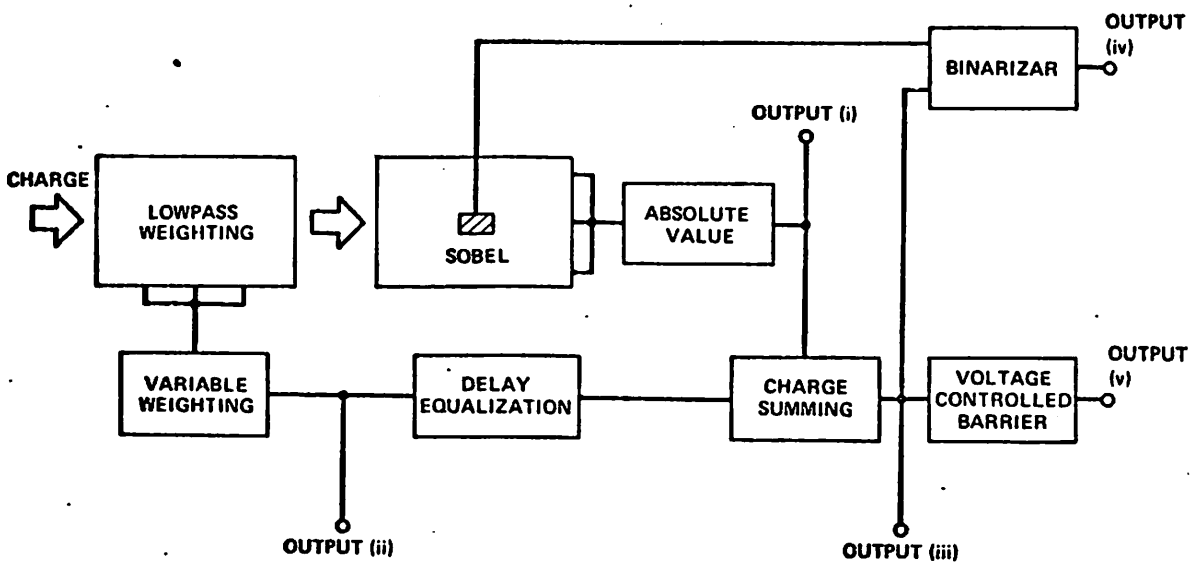


Figure 9(a). Block schematic for test Circuit II.

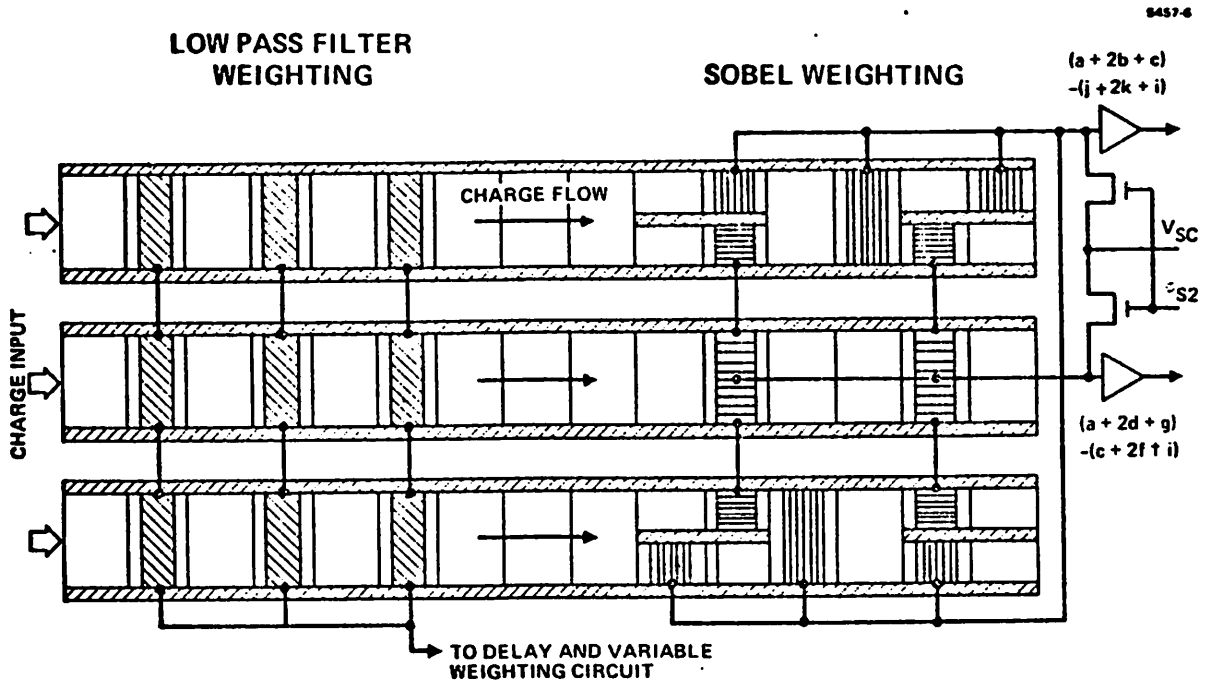


Figure 10. Weighting required for test Circuit II.

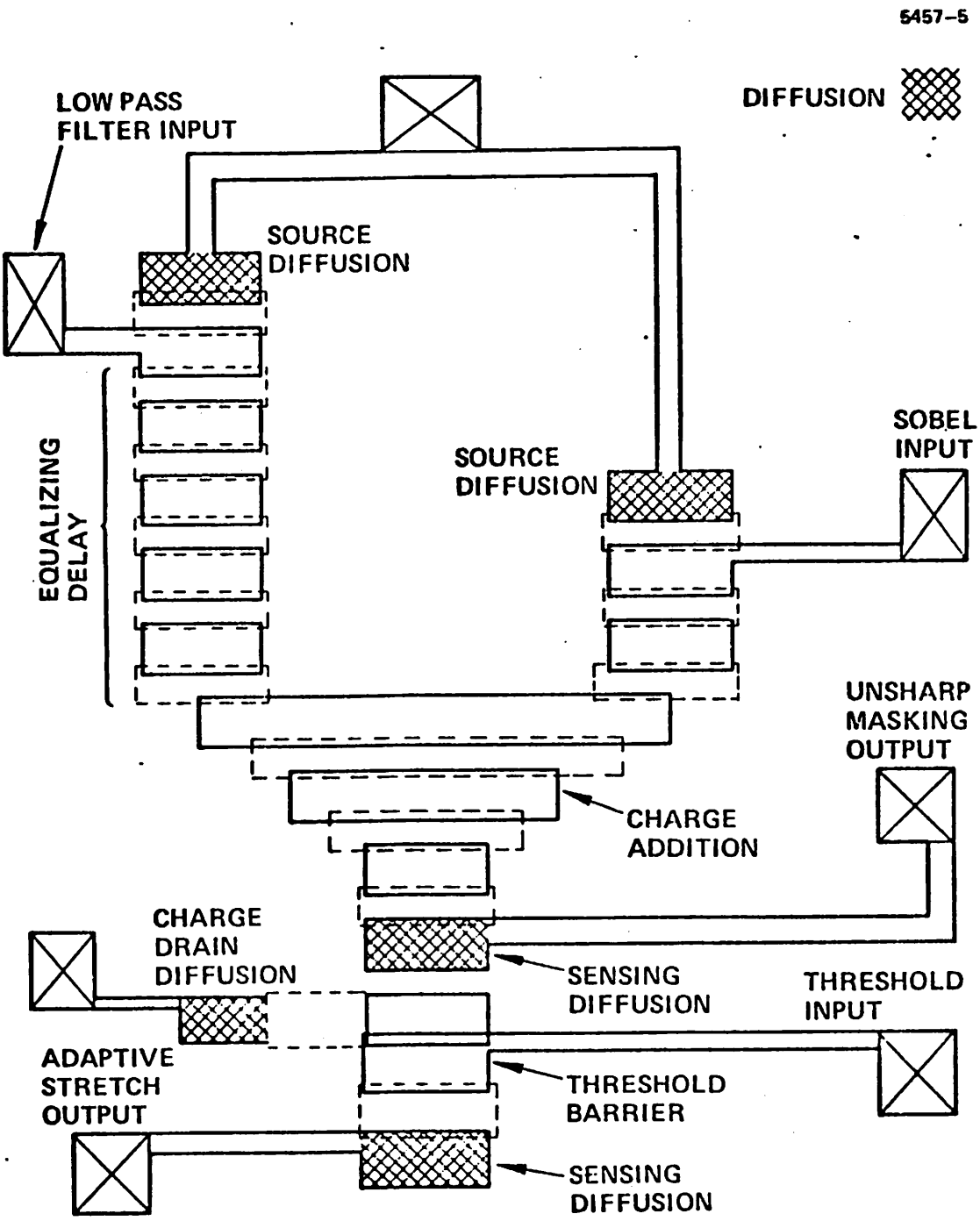


Figure 11. Illustration of layout for unsharp masking and adaptive stretching.

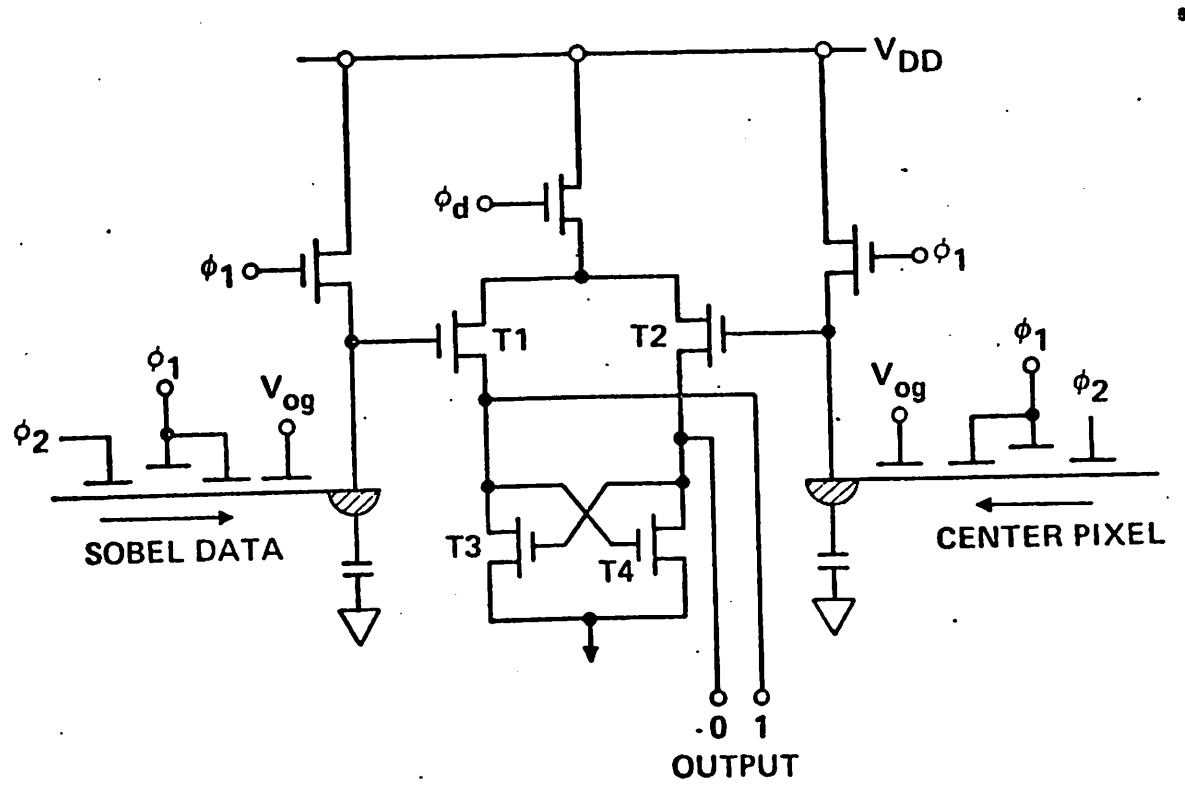


Figure 12. Concept of binarizer.

as part of the next phase of this program.

References

1. R.W. Brodersen, C.R. Hewes, and D.D. Buss, "A 500-Stage CCD Transversal Filter for Spectral Analysis," IEEE ED-23, No. 2, February 1976, p. 143.

2. J.F. Albaran and D.A. Hodges, "A Charge Transfer Multiplying D to A Converter," 1976 IEEE International Solid State Conference, Philadelphia, February 1976.

3. M.F. Tompsett, "Surface Potential Equilibrium Method for Setting Charge in CCDs," IEEE ED-22, No. 6, June 1976.

4. M.H. White, I.A. Mack, F.J. Kirk, D.R. Lampe, and J.L. Fagan, "An Analog CCD Transversal Filter with Floating Clock Electrode Sensor and Variable Tap," IEEE First Solid State Conference, Philadelphia, February 1976.

5. W.F. Kosnocky and J.E. Carnes, "Charge-Coupled Digital Circuits," IEEE, Vol. SC-6, No. 5, October 1971.

4.3 REAL TIME NONLINEAR OPTICAL DATA PROCESSING

J.L. Erickson, A. Au and J. Grinberg

During the past year we have been using the hybrid field effect liquid crystal light valve (LCLV) as an incoherent-to-coherent interface in optical data processing (ODP) experiments.

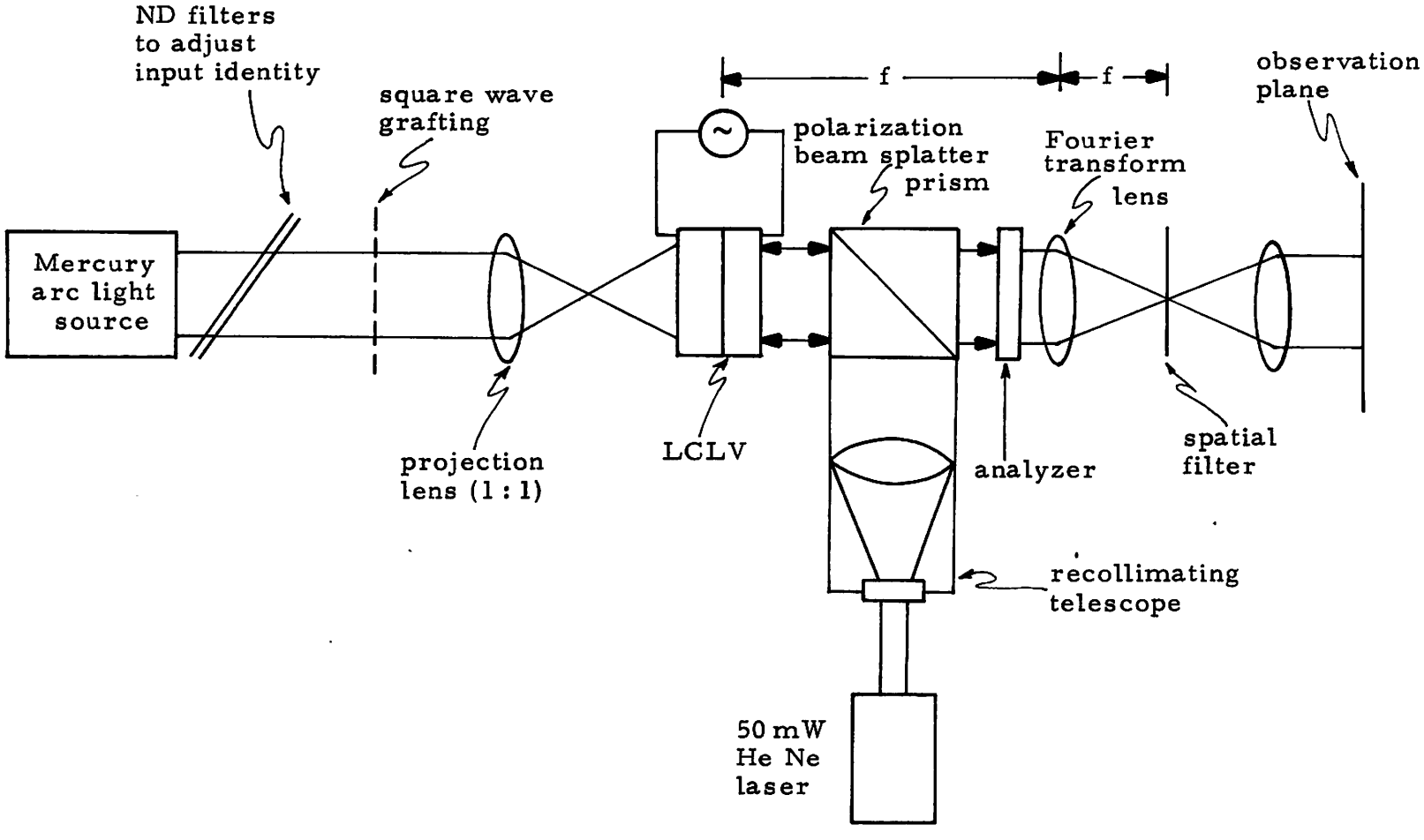


Figure 1. Experimental Setup

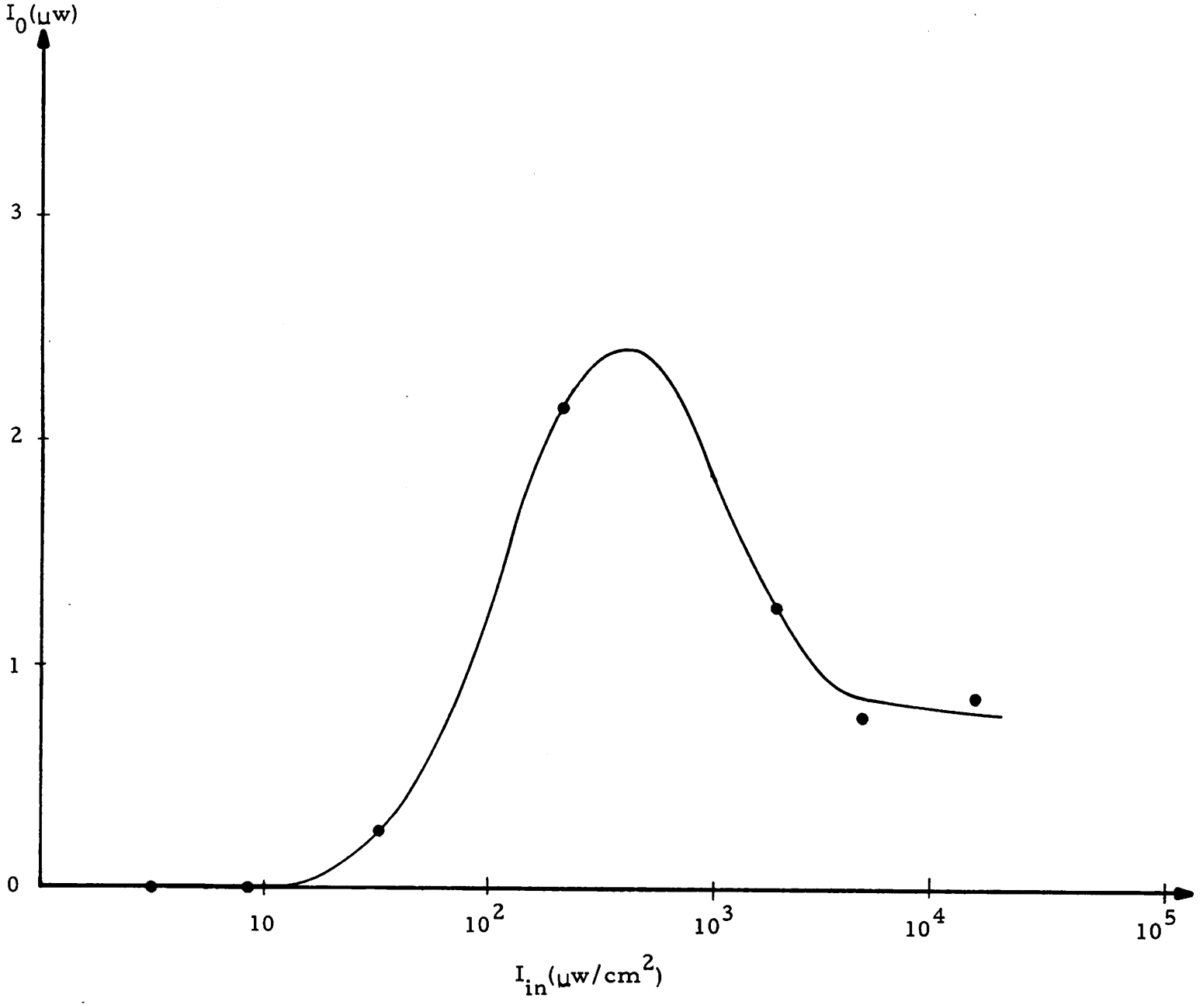


Figure 3. Level Slicing I/O Characteristic for V = 4.0 volts rms at 2KHz.

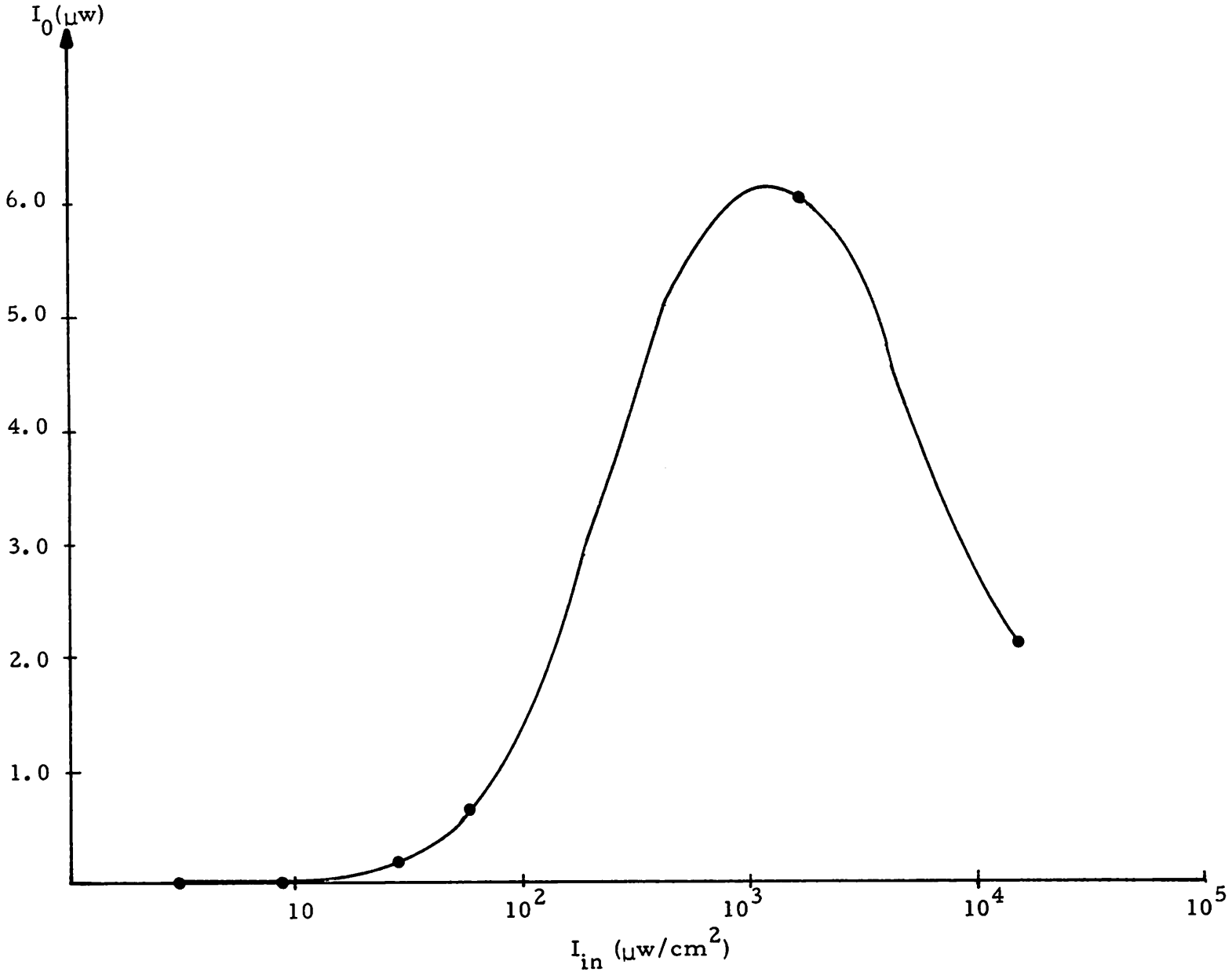


Figure 4. Level Slicing I/O Characteristic for V = 5.8 volst rms at SKHz

the grating and I_L is the incident lamp intensity. Replacing the inequalities by equalities, the gray scale (density) range that appears as a single "sliced" level is

$$\Delta D_S = \log \frac{I_M}{I_T} \frac{T_{max}}{T_{min}} .$$

The total number of levels that can be independently sliced out depends on the maximum lamp intensity, I_{Lmax}. The transparency density range covered is

$$D_{total} = \log \frac{T_{min} I_{Lmax}}{I_M} ,$$

and the number of levels, n, is

$$n = \frac{D_{total}}{\Delta D_S} .$$

These equations can be expressed in terms of quantities that are characteristic of the system components as follows. The differential density of the grating is

$$\Delta D_G = \log \frac{T_{max}}{T_{min}} .$$

The device "density" range is

$$\Delta D_D = \log \frac{I_M}{I_T} ,$$

which is, in a photographic analogy, the number of "log exposure" units required to go from threshold to saturation. The lamp "density" is

$$D_L = \log \frac{I_{Lmax}}{I_M} ,$$

i.e., the number of saturation "log exposure" units available from the lamp. Finally, the basic grating density is

$$D_G = - \log T_{min} .$$

With these definitions, we have

$$\Delta D_S = \Delta D_D + \Delta D_G$$

photoconductor LCLVs. The silicon characteristics indicate that up to 10 levels could be isolated with the same equipment and a silicon LCLV. It is unlikely, however, that a silicon device will be available for tests.

The most lucrative device modification is to find liquid crystal materials that give a smaller $I_M - I_T$. This would reduce ΔD_D and increase D_L, both of which would increase the number of levels n. There is a good probability that addition of some cholesteric liquid crystal to the device mixture would provide some improvement in I_M. However, the required liquid crystal chemistry is non-trivial and experiments of this kind are beyond the scope of the current program.

We are in the process of making some noise measurements on the LCLV, but the results are preliminary and will not be reported here.

5. RECENT PH.D DISSERTATIONS

This section includes the abstracts of recent Ph.D. dissertations completed over the past six-month research period. Due to the length of their reports, it was felt that abstracts presented in this semi-annual report was appropriate. For those individuals interested in further details, these dissertations appear in their entirety in the form of USC-Image Processing Institute reports. They are available upon request.

Experimental results with general halftone screens made on a plotting flatbed microdensitometer are presented. Logarithmic, exponential, and level slice characteristics have been achieved with monotonic cells. Intensity notch filter and quantization characteristics have been achieved with nonmonotonic cells. Other generalizations of the technique are discussed, including the possibility of real-time nonlinear processing with optical input transducers.

5.2 ESTIMATION AND DETECTION OF IMAGES DEGRADED BY FILM-GRAIN NOISE

Firouz Naderi

Film-grain noise is a term describing the intrinsic noise produced by a photographic emulsion during the process of recording an image on film. Although film-grain noise has recently been considered within the field of image processing, the nature of the noise is still somewhat misunderstood.

One goal of this study has been to use the theoretical and experimental results on film characteristics obtained by photographic scientists in order to define film-grain noise within the context of estimation theory. A detailed model for the overall photographic imaging system is presented. There are linear blurring effects at the initial and the final segments of this model to account for various optical and chemical degradations. The middle segment of the model represents signal dependence effects of film-grain noise and includes

5.3 IMAGE RESTORATION BY SPLINE FUNCTIONS

Mohammad J. Peyrovian

Spline functions, because of their highly desirable interpolating and approximating characteristics, are used as a potential alternative to the conventional pulse approximation method in digital image processing. For uniformly spaced knots, a class of spline functions called B-splines has the useful properties of shift invariance, positiveness, convolutional property and local basis property. These properties are exploited in image processing for linear incoherent imaging systems.

The problem of image degradation in a linear imaging system is described by a superposition integral. For simulation of degradation and restoration by means of a digital computer, the continuous imaging model must be discretized. Thus, a theoretical and experimental study of quadrature formulae, particularly monospline and best quadrature formulae in the sense of Sard, is presented. It is shown that a good choice of degree for a monospline highly depends on the frequency content of the integrand, and in most cases, a cubic monospline generates less error than the pulse approximation method and Newton-Cotes quadrature formulae.

In space-invariant imaging systems, the object and point-spread function are represented by B-splines of degrees m and n. Exploiting the convolutional property, the deterministic part of the blurred

are achieved by exploiting statistical redundancies within an image. This is combined by transmission of only those portions of the mathematical image representation which the human observer is most sensitive to. This dissertation describes research intended to extend current image coding techniques to the coding of sequences of digital images transmitted over a digital communications channel. The emphasis is directed towards definition of an image coding system that exploits temporal as well as spatial image redundancies.

A primary objective of this investigation is to develop a class of interframe hybrid transform/predictive coders having near optimum levels of performance. The interframe hybrid coder implementations considered employ two-dimensional unitary transforms in the spatial domain coupled with first-order DPCM predictive coding in the temporal domain. Based on a statistical image representation, a model is developed for the hybrid coder transform coefficient temporal difference variance matrix. With this model, theoretical MSE performance levels for the hybrid coder with zonal coding are determined as a function of spatial subblock size.

Implementations of the interframe hybrid coder using discrete cosine and Fourier transforms are experimentally evaluated. High quality image reconstruction are demonstrated for reductions of 32:1 in average pixel bit rate. Operational considerations investigated for the hybrid interframe coder include initial conditions, spatial and temporal adaptation, reinitialization, and total transmission bit rate. Also, sensitivity of the interframe hybrid coder to channel

6. PUBLICATIONS

1. H.C. Andrews, "A Potpourri of Mathematical Techniques Applied to Image Science," Proceedings of the First Annual Symposium on Mathematical Techniques in Image Science, USNPGS, Monterey, California, November 1976.

2. H.C. Andrews, "Image Processing at the University of Southern California," CalTech/JPL Conference on Image Processing, November 1976.

3. H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice Hall, New Jersey, 1977.

4. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Analysis and Synthesis," submitted to Applied Optics.

5. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Effects of Input Medium and Precompensation," submitted to Applied Optics.

6. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Non-monotonic Halftone Cells and Phase Halftones," submitted to Applied Optics.

7. W. Frei, "Image Enhancement by Histogram Hyperbolization," accepted for publication in Computer Graphics and Image Processing.

8. W. Frei, "Rate-Distortion Coding Simulation for Color Images," Proceedings of the SPIE, San Diego, California, August 1976.

17. R. Nevatia and T.O. Binford, "Description and Recognition of Curved Objects," accepted for publication in the Artificial Intelligence Journal.

18. W.K. Pratt, "Digital Image Enhancement and Restoration Display Techniques," Society for Information Display, 1976 SID International Symposium, Los Angeles, California, May 1976.

19. W.K. Pratt, "Pseudoinverse Image Restoration Computational Algorithms," presented at 1976 Joint Soviet-USA Optical Information Processing Symposium, Novasibirsk Siberia: to be published in Optical Information Processing, Vol. II, Plenum Press, 1977.

20. W.K. Pratt, "Two-Dimensional Unitary Transforms," presented at NATO ASI Symposium on Digital Image Processing: to be published in Proceedings NATO ASI Symposium on Digital Image Processing, Plenum Press, 1977.

21. J.A. Roese and W.K. Pratt, "Theoretical Performance Models for Interframe Transform and Hybrid Transform/DPCM Coders," Proceedings SPIE Symposium on Advances in Image Transmission Techniques, San Diego, California, August 1976.

22. A.A. Sawchuk, "Signal-Dependent Noise Models and Nonlinear Nonstationary Image Restoration," NATO Advanced Study Institute on Digital Image Processing and Analysis, Bonas, France, June 1976.

23. A.A. Sawchuk, "Nonlinear Nonmonotonic Optical Processing," Gordon Research Conference on Coherent Optics and Holography, Santa