LOW LEVEL AND INTERMEDIATE LEVEL VISION IN AERIAL IMAGES

by

Oscar A. Zuniga

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

APPROVED:

Roger W. Ehrich, Chairman

Robert M. Haralick, Co-chairman

Layne T. Watson

Richard W. Conners

F. G. Gray

September 1988

Blacksburg, Virginia

LOW LEVEL AND INTERMEDIATE LEVEL VISION IN AERIAL IMAGES

by

Oscar A. Zuniga

Roger W. Ehrich, Chairman

Electrical Engineering

(ABSTRACT)

Low-level and intermediate-level computer vision tasks are regarded as transformations from lower to higher level representations of the image information. An edge-based representation that makes explicit linear features and their spatial relationships is developed. Examples are presented in the scene domain of aerial images of urban scenes containing man-made structures.

The techniques used are based on a common structural and statistical model of the image data. This model assumes that the image data is adequately represented locally by a bivariate cubic polynomial plus additive independent Gaussian noise. This model, although simple, is shown to be useful for the design of effective computer vision solving tasks.

Four low-level computer vision modules are developed. First, a gradient operator which reduces sharply the gradient direction estimate bias that plagues current operators while also reducing sensitivity to noise. Secondly, a Bayes decision procedure for automatic gradient threshold selection that produces results which are superior to those obtained by the best subjective threshold. Thirdly, the new gradient operator and automatic gradient threshold selection are used in Haralick's directional zero-crossing edge operator resulting in improved performance. Finally, a graytone corner detector with significantly better probability of correct corner assignment than other corner detectors available in the literature.

Intermediate-level modules are developed for the construction of a number of intermediate-level units from linear features. Among these is a linear segment extraction method that uses both, zero-crossing positional and angular information together with their distributional characteristics to accomplish optimal linear segment fitting. Methods for hypothesizing corners and relations of parallelism and collinearity among pairs of linear segments are developed. These relations are used to build higher level groupings of linear segments that are likely to correspond to cultural objects.

Acknowledgements

Many individuals have helped and contributed to the completion of this dissertation. I thank my dissertation advisor Dr. Robert M. Haralick. He has taught me by example the value of hard work, inquisitiveness, and enthusiasm. He originated many of the ideas in this dissertation and has patiently guided me, and constantly challenged me to do my best.

Dr. Layne T. Watson was very generous with his time in discussing and reviewing my dissertation. He offered much helpful advice and corrections that improved the original manuscript. Dr. Roger W. Ehrich and Dr. F. G. Gray were always very supportive. I thank Dr. Richard W. Conners for his encouragement and for providing an excellent working environment at the Spatial Data Analysis laboratory.

Many fellow graduate students and staff have made my work at the Spatial Data Analysis laboratory a very pleasant experience. Thanks to and . I enjoyed many estimulating discussions with , , , and .

Thanks to the family and my extended family that gathers at their home every Tuesday evening. Their constant encouragement and prayers have made this dissertation a reality. helped with some of the drawings and provided much moral support.

My parents never lost faith. I thank the Lord for them. Thanks to my grandmother , my brother , and my uncles. This dissertation is also theirs.

To my wife : I love you. She has been a true friend and I could always count on her for encouragement and help. She has given much of herself and sacrificed to support me during the difficult times. She also provided constructive criticism of my writing which improved the clarity of the presentation. I dedicate this dissertation to her.

Finally, this dissertation would not have been possible without the Lord's help and guidance. To Him be the glory.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

INTRODUCTION

1.1 Computer Vision

Vision is a biological perceptual task that has been the subject of much study and interest. Early vision research in this century was carried out mostly by investigators in the field of Psychology. Some of the main results of these studies are the Gestalt principles of organization (Wertheimer (1912), Kochler (1929), Koffka (1935)) and Gibson's theory of space perception (1950), and ecological approach (1979). With the increasing use of digital computers during the last two decades, vision has also become a subject of study in Computer Science and Engineering. It is hoped that this combined effort will bring additional insight into the vision problem.

Computer vision systems typically start with a **digital image.** A digital image is an array of picture elements (henceforth denoted as **pixels**), whose values represent visual sensory information obtained from a video camera or other such imaging device, and

digitized to a certain number of quantization levels. The task of the system is to identify and describe the nature of the real world objects which give rise to the digital image. The extent and exact nature of that description depends of course upon the goals and expectations of the observer. This task has proved to be, in general, an extremely complex one.

There are typically three levels of information processing in computer vision. **Low-level vision** or early processing extracts primitive features such as change of intensity and orientation of edge elements from the original intensity array. In the next level of processing, called **intermediate-level vision**, higher level features such as lines and regions, and intrinsic scene characteristics such as surface orientation and occlusion are extracted. **High-level vision** or late processing deals with objects and relies on domain-specific knowledge to construct description of scenes.

This dissertation deals with some problems at the first two levels. The techniques used to solve these problems are based on a common structural and statistical model of the image data. We believe much work in computer vision proceeds with the underlying model being either unstated or completely non-existent. The structural and statistical model used assumes that the image data is adequately represented locally by a bivariate cubic polynomial plus additive independent Gaussian noise. This model, although simple, is shown to be useful for the design of effective computer vision solving tasks.

Low-level and intermediate-level computer vision tasks are regarded as transformations from lower to higher level representations of the image information. The explicit models used enables us to make effective judgements about the significance of the units or primitives extracted at each representational level.

1.2 Problems and Scope

Two problems in low-level computer vision and two problems in intermediate-level vision are tackled in this dissertation: Local edge detection, Gray level corner detection, Linear feature extraction, and Linear feature organization. Solutions to these problems are sought for the purpose of constructing an edge-based representation that makes explicit linear features and their spatial relationships. This representation is specially useful in the recognition of man-made structures in aerial images.

The Local edge detection problem is to transform the array of image gray level intensities into a set of primitives, ie. **edgels**, that represent significant local spatial discontinuities in gray level intensity. These edgels are described by their strength, orientation, and position. The technique used is based on Haralick's zero-crossing of second directional derivative of a cubic facet (1984). Three subproblems are investigated: Function approximation with continuous and discrete orthogonal polynomials, accurate estimates of edge orientation, and automatic selection of gradient strength threshold in order to distinguish zero-crossings due to actual variation of a physical characteristic of the scene from zero-crossings due to noise.

The Gray level corner detection problem is to locate significant local spatial discontinuities in image intensity that take place in two near orthogonal directions.

The Linear feature extraction problem consists in transforming the set of edgels into sets of **linear segments**. Each linear segment is a maximally connected set of edgels having similar local and global orientation. As with edgels, linear segments are described by their strength, orientation, and position.

The Linear feature organization problem is to organize the set of linear segments into groups such that the segments on each group are related by some perceptually significant characteristic, such as end-point proximity, collinearity, and parallelism.

1.3 Assumptions and models

A digital image encodes much information about the scene. This information is manifested primarily through spatial or temporal intensity variations which correspond to variations in some physical characteristic of the scene such as surface reflectance, orientation, and lighting source characteristics. Our ability to decode these intensity variations and to produce a correct interpretation of the scene depends critically upon our use of:

1. Models of the scene domain, the illumination, and the imaging process.

2. Adequate representations of knowledge. This knowledge can be apriori knowledge about the scene domain or knowledge gained at different stages of the recognition process.

3. Extraction of the required digital image features upon which the higher levels of image information are built.

Models are needed to effectively encode the apriori information. The extent to which the interpretation and understanding succeeds depends on the success of the low level image feature extraction.

In what follows we make explicit the assumptions and models used.

1.3.1 Scene Domain

Although many of the methods developed in this dissertation are domain independent, they will mainly be applied to the scene domain of monocular, monochromatic **aerial images** of urban scenes. Examples are presented toward the recognition of **man-made structures**.

1.3.2 Image and Noise Model

All low level operations which manipulate the image data make extensive use of an image model referred to as the **facet model** (Haralick (1980), Haralick and Watson (1981)). This model has been used succesfully in the past to perform many low-level and intermediate-level tasks such as edge detection (Haralick, 1982), region growing (Pong et al., 1981), the topographic primal sketch representation of the digital image (Haralick, Watson, and Laffey (1983)), determining intrinsic scene characteristics from images (Pong, 1984), and optic flow and rigid body motion (Lee, 1985).

The local facet model is the result of a particular interpretation of the image data. This interpretation says that the image data consists of observed noisy samples from an underlying continuous graytone intensity function. For our purposes we assume that the noise is additive, independent, and Gaussian with zero mean. To be precise, the basic philosophy of the facet model derives from recognizing that the discrete set of values which form the digital image are the result of sampling and quantizing a real-valued

function f defined on the domain of the image which is a bounded and connected subset of the real plane. Thus, any property associated with a pixel or a neighborhood of pixel values should be evaluated by relating it to the property of the corresponding graytone suface f which underlies the neighborhood. This involves estimating the surface funtion f locally, from the neighborhood samples available to us. The most natural way of accomplishing this is by assuming a parametric form for f and then estimating its associated parameters. The choice of parametric form, number of parameters, and neighborhood size is not without consequences, since they affect the estimation of the surface function f.

The parametric form we use is a bivariate cubic. We assume that the neighborhood around each pixel is suitably fit by the bivariate cubic:

$$\begin{aligned} f(r,c) = & K_1 \\ & + K_2 r + K_3 c \\ & + K_4 r^2 + K_5 rc + K_6 c^2 \\ & + K_7 r^3 + K_8 r^2 c + K_9 rc^2 + K_{10} c^3 \end{aligned}$$

where r and c are the image now and column coordinates of an arbitrary pixel.

The K parameters are estimated by a least squares fit. This computation is both efficient and numerically stable although, in general, not statistically the best. Once this surface has been estimated around each pixel, the first and second directional derivatives are easily obtained by analytical means. Tasks such as topographic classification and detection of edges and corners are totally based on the first and second directional derivatives.

1.3.3 Imaging System Model

The imaging system is shown in Figure 1 on page 8.

The following assumptions are made:

1. The camera is a pinhole camera situated at the origin of the world coordinate system with the z-axis pointing towards the image plane.

2. The image plane is in front of the camera at a distance f, the focal length of the camera lens.

3. The image plane is parallel to the ground plane, that is the camera points straight down to the center of the image.

The relationship between a world coordinate point (x,y,z) and its corresponding image point (x_i, y_i) is governed by the following two perspective projection equations:

$$x_i = \frac{f}{z} x \ , \quad y_i = \frac{f}{z} y$$

1.3.4 Illumination Model

We assume a single point light source located infinitely far away. This means that all the light rays emanating from the light source are parallel. The light source angles θ and ϕ are known and are illustrated in Figure 2 on page 9.

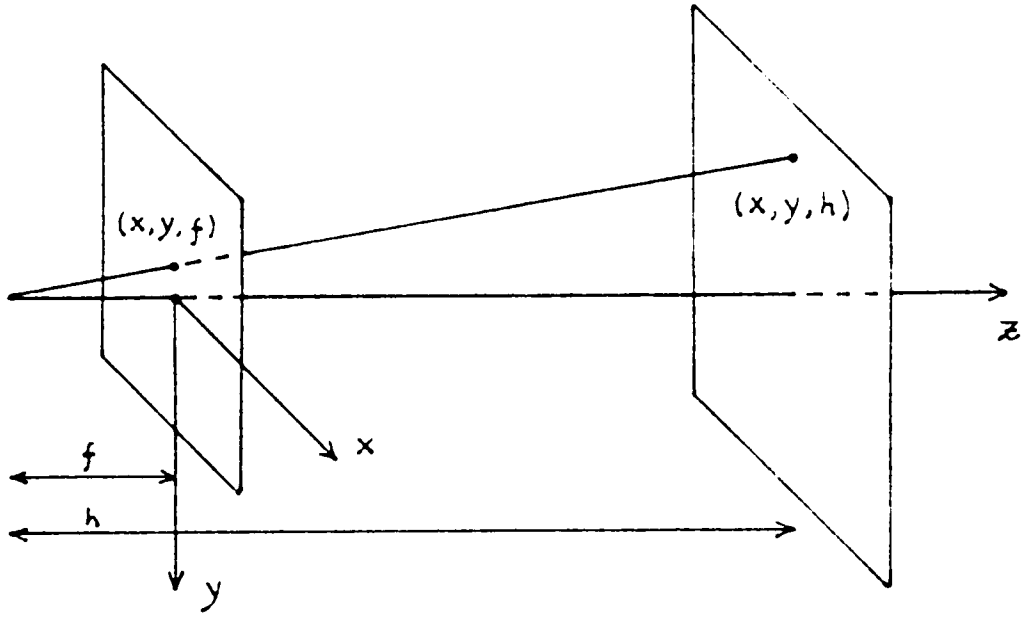


Figure 1. shows the image system model.

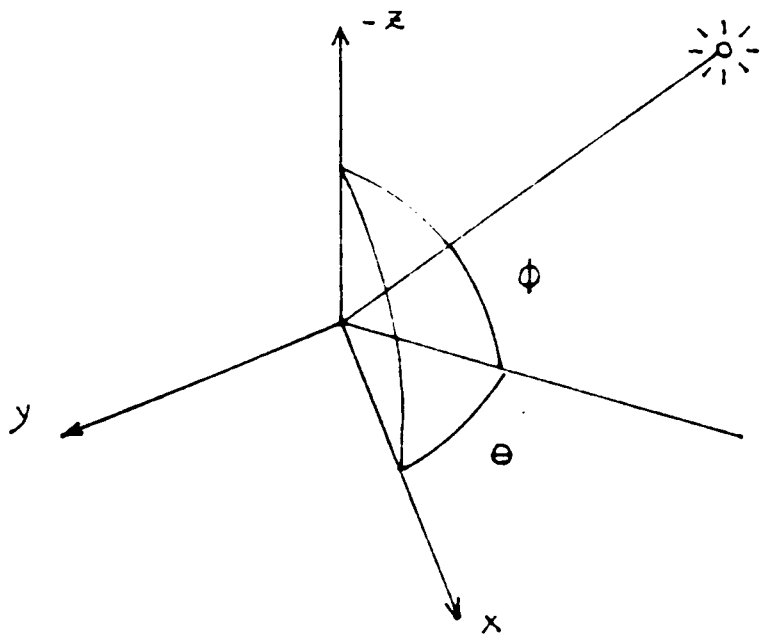


Figure 2. shows the illumination model.

1.4 Contributions

This dissertation makes a number of contributions to low-level and intermediate-level computer vision research. These include:

1. A gradient operator based on the integrated directional derivative which reduces sharply the gradient direction bias that plagues current operators while also reducing sensitivity to noise. Furthermore, the new operator is relatively insensitive to edge or line orientation and its step edge response decreases smoothly to zero as edge displacement increases, preventing the creation of some false zero-crossings.

2. A Bayes decision procedure for automatic gradient threshold selection that produces results which are superior to those obtained by the best subjective threshold.

3. An edge detector based on Haralick's zero-crossing of second directional derivative of a cubic facet. Integration of the new gradient operator and Bayesian gradient threshold selection into the edge detector results in improved detection and localization of zero-crossings.

4. A graytone corner detector with significantly better probability of correct corner assignment than other corner detectors available in the literature.

5. A linear segment extraction method that uses both, zero-crossing positional and angular information together with their distributional characteristics to accomplish optimal linear segment fitting.

6. Methods for hypothesizing corners and relations of parallelism and collinearity among pairs of linear segments and evaluating their significance. These relations are used to construct higher level groupings of linear segments possibly corresponding to cultural objects. Illustrations are presented in the domain of aerial images of urban scenes where the objects of interest are modeled as being rectangular or composed of rectangular components.

1.5 Outline

Chapter 2 presents the background and literature review on the subject of the construction of an **edge-based representation** that makes explicit linear features and their spatial relationships for application to the recognition of man-made structures in aerial images. The chapter is organized from the point of view of transformations that take place on lower level **units** or primitives to organize them into higher level units, and the low-level and intermediate-level tasks which bring about those transformations.

Chapter 3 develops the concept of the integrated first directional derivative gradient operator, and compares its performance to other existent gradient operators with both simulated and real data.

Chapter 4 deals with some statistical methods to select a suitable gradient threshold to distinguish zero-crossings due to edges arising from a real variation of a physical characteristic in the scene from false zero-crossings due to noise.

Chapter 5 reviews function approximation with continuous and discrete orthogonal polynomials and presents some results that relate derivatives of the approximated surfaces to some currently prevalent edge detection methods. The integrated first directional

derivative technique developed in chapter 3 and the gradient threshold selection method developed in chapter 4 are used in Haralick's zero-crossing of second-directional derivative edge operator. Analysis of its performance with simulated and real data is presented.

Chapter 6 develops a number of techniques to detect gray level corners, and present a comparison of their performance with other existent corner detectors.

Chapter 7 presents an analysis of the distributional characteristics of zero-crossing position and orientation and uses this information in the design of a method for linear segment extraction. The distributional properties of the extracted segments are then derived and the application of the method to aerial images of urban scenes is illustrated.

Chapter 8 develops methods for the organization of linear segments based on proximity, collinearity, and parallelism. Illustrations are presented with aerial images of urban scenes.

Finally chapter 9 presents a summary and conclusions.

The various low-level and intermediate-level computer vision modules developed in this dissertation have been implemented within the GIPSY programming environment. GIPSY is an acronym for "General Image Processing System", a coordinated system of commands for image processing and image understanding developed by Professor Robert M. Haralick and students at Virginia Polytechnic Institute and State University. Each vision module is implemented as one or several commands, each of which operates on one or more input data files and produces one or more output data files. Data files contain image intensity information or information derived from image analysis tasks. Appendix F illustrates the various commands used in generating the results shown in this dissertation.

Chapter 2

BACKGROUND AND LITERATURE REVIEW

2.1 Units and Representations

Many low-level and intermediate-level vision tasks can be regarded as transformations from lower to higher level **representations** of the image information. A representation makes explicit the information needed at a given level. A representation is composed of **units** or primitives and their associated features or properties. The units of a representation are often formed by grouping or aggregating units of lower level representations.

The lowest representational level corresponds to the image array of gray level intensities. The units in this representation are pixels and their associated features are gray level intensities. What can be computed with the information at this level mostly determines the structure of the representation at the next level. Examples are Marr's primal sketch (Marr, 1976); Haralick, Watson, and Laffey's topographic primal sketch

(Haralick, Watson, and Laffey, 1983); and Ehrich and Foith's relational trees (Ehrich and Foith, 1978).

Marr's **primal sketch** is obtained by grouping units of similar type and feature obtained by comparing operators of several sizes. The units in this representation include edges, shading-edges, extended-edges, lines, and blobs. Unit features include position, termination points, contrast, size, and fuzziness.

Haralick et al. use the local facet model to describe locally the shape of each surface patch of gray level intensities. Each pixel is thus topograghically classified in units such as peak, pit, ridge, ravine, saddle, flat, and hillside. Edges are included as inflection points in hillsides. Pixels can then be grouped according to their topographic classification to ideally obtain a rich, hierarchical, and structurally complete representation of the fundamental structure of the digital image. This representation is called the **topographic primal sketch**.

Ehrich and Foith represent gray level intensity profiles by relational trees. A **relational tree** is a tree data structure that is a comprehensive description of the one-dimensional contextual relationships defined by the peaks and valleys of an intensity profile. Pointers from the tree vertices link the structure elements to attribute lists from which the original scene can be completely reconstructed. The units chosen in this representation are the peaks of the intensity profile, and algorithms operating on this representation are developed for region growing and edge detection.

The information explicitly provided by these representations is, in general, not sufficient for segmenting the image into regions of homogeneous properties corresponding to three-dimensional surfaces. Two closely related representations have been proposed at this level. Barrow and Tenenbaum (1978) suggest a number of transformations acting on arrays of image features (such as those provided by the previously mentioned representation) to recover **intrinsic scene characteristics** at each point in the image. These

intrinsic scene characteristics are physical characteristics of the surface element visible at each point in the image, such as: surface reflectance, surface orientation, incident illumination, range, etc. This information makes possible meaningful segmentation with a minimum of apriori knowledge about the scene. Marr (1977) proposes a viewer centered representation of surface orientation and range which is called the 2½-D sketch.

The construction of these representations is the goal of low-level and intermediate-level tasks. Low-level vision tasks include, for instance, the detection and description of edges, and topographic features in the topographic primal sketch. Intermediate-level tasks include the grouping processes in the topographic primal sketch to form one-dimensional and two-dimensional units, and also the use of this information to obtain intrinsic images.

Lowe (1987) argues that the need for depth reconstruction is the exception rather than the rule. He proposes the use of a process of **perceptual organization** to bridge the gap between the two-dimensional image and knowledge of three-dimensional objects. This process of perceptual organization allows the formation of groupings and structures in the image that are likely to be invariant over a wide range of viewpoints. This argument is supported by the fact that for a number of domain specific situations, reasonable success in recognition has been achieved without full use of the representations we have considered. An example is the domain of aerial images containing man-made structures, where the presence of **linear features** and their spatial relationships provide a strong clue for recognition.

According to the types of units, representations used in aerial image understanding fall broadly in three classes: edge-based representations, region-based representations, and combined or integrated edge-region representations. Notice that edges can be found from a region-based representation by declaring edges at the boundaries between regions. In this work, while discussing edge-based representations, we will only describe

methods that locate edges without necessarily finding regions first. Edge-based representations are specially suited for recognizing and describing the linear features present in man-made structures, and we will only discuss in this chapter the construction of this kind of representation. Some examples of the use of region-based representations for aerial image understanding are found in Nagao et al. (1980), Selfridge and Sloan (1982), Price and Reddy (1979), Stockman et al. (1982), Brooks (1981), and Besl (1986). Examples of the integration of edge-based and region-based representations are: Reynolds et al. (1984), Clark et al. (1979), Kestner et al. (1980), Milgram (1979), Nevatia and Price (1982), and Fua and Hanson (1987).

In the following sections, we discuss the construction of an edge-based representation that makes explicit linear features and their spatial relationships for their application to the recognition of man-made structures in aerial images.

2.2 *From Pixels to Edges*

Edges are local discontinuities in some image attribute. We discuss mainly techniques to detect local discontinuities in gray level intensity, i.e., **step edges**, although many of these techniques are easily extendable to other edge types.

Edge detection can be regarded as a transformation from an array of gray level intensities to an array of edge elements described by their contrast and orientation. This transformation usually consists of two steps. In the first step, an edge operator is applied to the image, and measures of the strength and orientation of the edge candidate are determined. In the second step, edges are declared only for those candidates whose edge strength exceeds a threshold. This second step is necessary to distinguish true edges

arising from a variation of a physical characteristic of the scene, from false edges arising through random fluctuations in the gray level intensity due to noise.

2.2.1 Edge Operators

Edge operators are broadly classified as gradient operators, edge template operators, surface fitting operators, and zero-crossing operators. Due to the abundance of different edge operators in the literature, this classification is not meant to be exhaustive.

2.2.1.1 Gradient Operators

Given the two-dimensional graytone intensity function f, it is intuitive that at an edge the rate of change of f is maximum across the edge. This direction corresponds to the gradient direction, and the rate of change of f measured in this direction corresponds to the gradient magnitude. This intuition has led to the use of the gradient as one of the earliest edge operators.

In a row-column coordinate system centered around a given neighborhood, the **gradient** of f is given by:

$$\nabla f = \left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial c}\right)$$

The partial derivatives $\partial f/\partial r$ and $\partial f/\partial c$ measure the rate of change of f in the row and column directions respectively. The magnitude of ∇f is given by:

$$|\nabla f| = \sqrt{(\partial f/\partial r)^2 + (\partial f/\partial c)^2}$$

The direction θ of ∇f is obtained from:

$$\sin\theta = \frac{\partial f / \partial r}{|\nabla f|}$$

$$\cos\theta = \frac{\partial f / \partial c}{|\nabla f|}$$

where the gradient direction is measured as a clockwise angle from the column axis.

A number of operators have been proposed to measure $\partial f / \partial r$ and $\partial f / \partial c$. Some of them are shown in Figure 3 on page 19: Roberts cross operator (Roberts, 1965), Prewitt operator (Prewitt, 1970), Sobel operator (Duda and Hart, 1973), and Cubic Facet operator (Haralick, 1982). Notice that the Roberts operator measures the rate of change of f in a coordinate system that is rotated by 45 degrees from the row and column coordinate system.

Most gradient operators face two major problems. One problem is the inherent bias in their estimate of edge direction. Another problem is their sensitivity to the presence of noise in the image data. The second problem can be alleviated by an increase in the processing neighborhood size but usually at the expense of an increase in estimate bias and also errors in the processing of small or thin objects. Chapter 3 discusses an operator based on the cubic facet model which sharply reduces both estimate bias and noise sensitivity, with no increase in computational complexity.

2.2.1.2 Edge Templates

Gradient operators estimate edge magnitude and direction by using derivative operators in two perpendicular directions. Another approach is to detect edges by match-

-1	0
0	1

0	1
-1	0

(a)

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

(b)

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

(c)

31	-5	-17	-5	31
-44	-62	-68	-62	-44
0	0	0	0	0
44	62	68	62	44
-31	5	17	5	-31

31	-44	0	44	-31
-5	-62	0	62	5
-17	-68	0	68	17
-5	-62	0	62	5
31	-44	0	44	-31

(d)

Figure 3. shows the row and column derivative masks for four gradient operators: (a) Roberts operator, (b) Prewitt operator, (c) Sobel operator, (d) Cubic Facet operator.

ing with templates of desired edges, also called edge masks in a number of orientations. Edge magnitude and direction are determined by the magnitude and direction associated with the maximum match.

Figure 4 on page 21 shows edge templates proposed by Prewitt (1970) and Kirsch (1971). Further examples of edge templates can be found in Nevatia and Babu (1980), Robinson (1977), and Canny (1986).

2.2.1.3 Surface Fitting

Two approaches have been proposed: general surface fitting and step fitting.

In general surface fitting, a parametric function is fit to the graytone intensity values inside a local neighborhood centered around a pixel, and the gradient of the estimated surface at the neighborhood center is used as an estimate of the gradient of the image at that pixel. Prewitt (1970) uses a least squares fit with a two-dimensional quadratic polynomial in a 3x3 neighborhood. The resulting edge operators are shown in Figure 3 on page 19. Haralick (1982) uses a least squares fit with a two-dimensional cubic polynomial. The resulting edge operators for a 5x5 neighborhood are shown in Figure 3 on page 19. Brooks (1978) and Beaudet (1978) also use the polynomial fitting approach.

In step fitting, an ideal step edge s(r,c) is fit to the graytone intensity values f(r,c) inside a local neighborhood. The step edge is described by its angle, distance from the neighborhood center, and graytone values on each side of the edge. Hueckel's approach (1971) to finding the best-fitting step edge, was to expand both f(r,c) and s(r,c) in terms of a set of eight orthogonal polar form Fourier basis functions defined on a disk, and to minimize an error measure defined as the sum of the square differences between corresponding coefficients. Similar approaches can be found in O'Gorman (1978) who uses

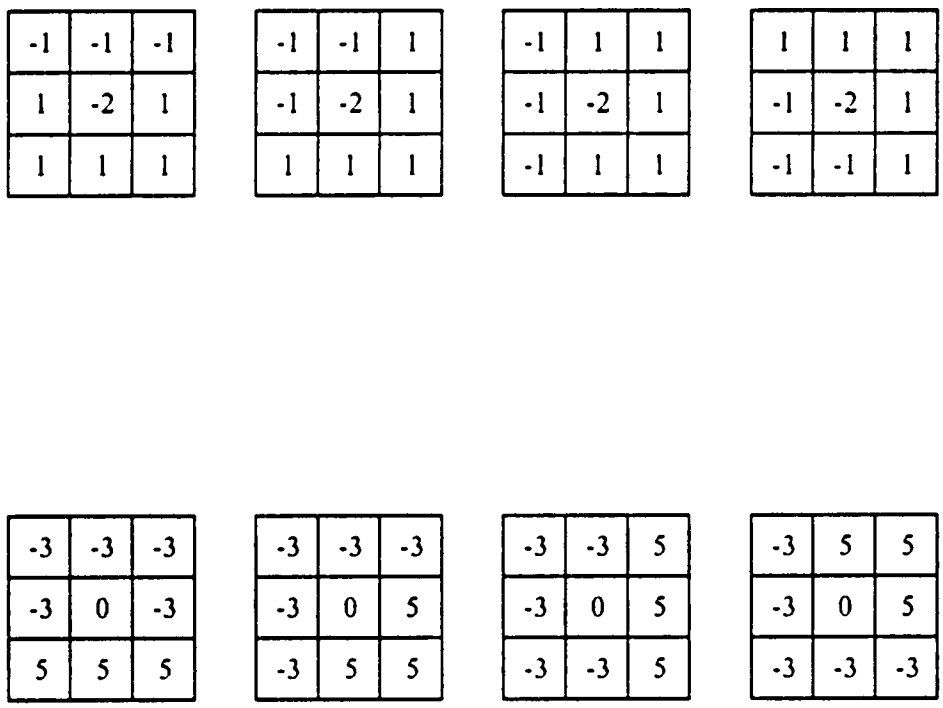


Figure 4. **shows from top to bottom the Prewitt and Kirsh edge templates:** Only four of eight possible directions are shown for each operator.

a set of two-dimensional Walsh functions defined on a square; Hummel (1979) who uses a set of optimal basis functions derived from the Karhunen-Loeve expansions of the local graytone intensity values; and Nalwa and Binford (1986) who use tanh functions.

2.2.1.4 Zero Crossing Operators

If we consider the one-dimensional profile of the edge shown in Figure 5 on page 23, we can say that edges occur at extreme values of the first derivative of the graytone intensity function, or at zero-crossings of the second derivative. The natural extension in two dimensions is to compute the derivatives in the direction of the gradient. Haralick (1984) has proposed this **directional** edge operator. Haralick computes derivatives of the graytone intensity function, using a local facet model in which the graytone intensity values in a local neighborhood are fitted to a two-dimensional cubic polynomial by a least squares error procedure. Once the polynomial coefficients have been estimated, the computation of the gradient direction and second directional derivative are straightforward.

Marr and Hildreth (1980) propose an **isotropic** edge operator in which edges are declared at zero-crossings of the Laplacian of a two-dimensional Gaussian filtered version of the gray tone intensity values.

Comparisons have been made about the performances, advantages, and disadvantages of directional and isotropic zero-crossing edge operators (Haralick (1984), Grimson and Hildreth (1985), Haralick (1985), Torre and Poggio (1986)).

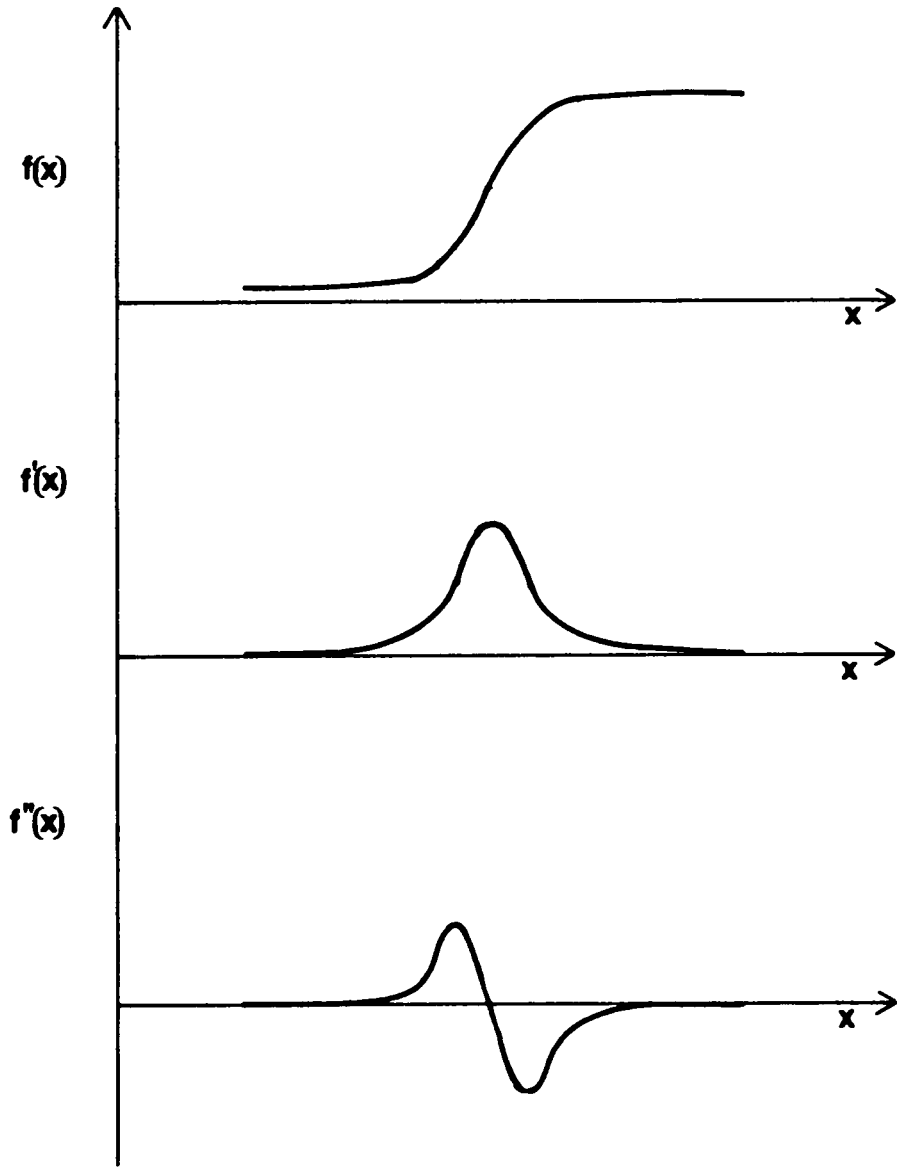


Figure 5. shows a step edge and its first and second derivatives.

2.2.2 Gradient Threshold Selection

As mentioned earlier, the second step of many edge dectectors consists of thresholding the edge or gradient magnitudes produced by the first step. Although it might seem that zero-crossing edge operators do not require a threshold, in practice, it is necessary to distinguish zero-crossings caused by true edges from zero-crossings caused by variations in the gray tone intensity surface due to noise. The latter are more likely to have lower gradient values than the former.

Methods for automatic gradient threshold selection often use the histogram of the gradient image. If all edges are of approximately the same strength, it is clear that the histogram of the gradient image will be bimodal, one mode corresponding to the distribution of edge points and another to the distribution of non-edge points. The valley of the histogram is an obvious choice for the gradient threshold. Unfortunately, for many real images and for noisy images the distributions of edge and non-edge points overlap considerably, and consequently, the histogram of the gradient values is no longer bimodal. The choice of a gradient threshold thus becomes a tradeoff between the missing of valid edges and the creation of noise-induced false edges.

Jacobus and Chien (1978) fit the gradient distribution with two Gaussians, one representing valid edges and another representing non-edge points or noise induced false edges. The threshold chosen is that which makes the noise edges some fraction of the real edges. Robinson (1977) suggests using the gradient mean as a threshold. Ohlander (1975) thresholds at some fraction of the standard deviation above the gradient mean. Baird (1977) proposes thresholding at the peak of the second derivative of the gradient distribution. A related method based on change of curvature of the distribution is reported by Perkins (1980)

Abdou and Pratt (1979) suggest a statistical design procedure in which the gradient threshold A is chosen such that the edge hypothesis is accepted if

$$\frac{p(A \mid edge)}{p(A \mid noedge)} \geq \frac{P(noedge)}{P(edge)}$$

and the no-edge hypothesis is accepted otherwise. Here p(A|edge), and p(A|noedge) represent the conditional probability densities of the gradient image and P(edge) and P(noedge) represent the apriori edge probabilities. This equation defines the well known maximum likelihood ratio test associated with the Bayes minimum error decision rule of classical decision theory. Abdou and Pratt compute the conditional probability densities for the Roberts, Sobel, and Prewitt gradient operators under the assumption of additive independent Gaussian noise. When the noise does not meet these assumptions they suggest a pattern classification design procedure which uses the linear discriminant function method in which the edge hypothesis is selected if $W'x \geq 0$ and rejected if $W'x < 0$, where w(1) and w(2) are weighting factors of the weighting vector $W = [w(1), w(2)]'$ and $x = [A, 1]'$. The weight factors are related to the decision threshold by

$$t = -\frac{w(2)}{w(1)}$$

Components of W are determined by a training procedure using a set of prototype pixel regions containing edges or no-edges.

Herskovits (1970) devised a technique for choosing a threshold for scenes of polyhedra viewed through a sensor with known characteristics to maximize the sensitivity of edge detection while minimizing the response to noise and smooth illumination gradients.

A single threshold for the whole image does not often give good results especially if the noise variance varies from area to area in the image or in images made under conditions of uneven illumination. Robinson (1977) suggests a locally adaptive threshold obtained by comparing the analog gradient image with a blurred version of the original image. Rosenfeld and Kak (1982) suggest dividing the image into blocks and applying threshold selection techniques to each block.

In chapter 4 the problem of automatic gradient threshold selection is posed within a statistical framework based on a cubic facet model for the image data and a Gaussian model for the noise. Under these assumptions two statistics which are functions of the gradient strength are derived. Experiments show that thresholds on these statistics produce results which are superior to those obtained by the best subjective threshold on the gradient image. A Bayes decision procedure that makes gradient selection automatic is developed.

2.3 From Edges to Segments

Section 2.2 described edge detection as a transformation from an array of gray level intensities to an array of edge elements described by their contrast or strength and their orientation. The result of this transformation is a set of unrelated edge elements. This section describes the processes or transformations acting on the array of edge elements to group them into one-dimensional edge units, i.e., **edge segments**. We review first, general local and global linking methods and then the extraction of linear segments.

2.3.1 Local Linking

Linking edge elements based on local criteria is a simple yet effective method of obtaining edge segments. Examples are the work of McKee and Aggarwal (1975), Perkins (1978), and Nevatia and Babu (1980). Prior to the linking process, edges are usually thinned by suppressing non-maxima edge magnitudes in the direction accross the edge. Linking of edge elements then takes place based on proximity and orientation.

Typically each edge element has two neighbors along the boundary, known as a predecessor and a successor, except for end-elements, isolated elements, and where two boundary segments join or intersect. For each edge element, its neighbors on a 3x3 grid are examined to determine its predecessors and successors. The decision as to what neighbors are acceptable predecessors or successors is usually based on local criteria such as directional continuity, proximity, and continuity of intensity on both sides of the edges.

Given a starting point an edge segment can be traced by following the successor elements until an element with no successors or an element included in a previous edge segment is reached.

2.3.2 Global Linking

Of special interest among the global methods are search techniques. Edge elements are interpreted as nodes in a graph where an arc exists between spatially adjacent edge elements. Weights or costs are assigned to each arc, usually a function of edge strength and orientation . The problem is then posed as that of finding a minimun cost path, or an optimal path with respect to a given evaluation function, between two given nodes

in the graph. Starting with the heuristic search of Martelli (1972) and the dynammic programming approach of Montanari (1971) a body of literature on graph searching techniques for edge following has developed.

The problem of finding all 'best' or 'optimal' paths when initial and goal nodes are not predefined has been investigated by Ramer (1975) and more recently by Mero (1981). In Ramer's approach the image array is scanned repeatedly for edges suitable to use as initial nodes. Once an initial node has been found, an optimal path is generated by searching in front and behind the edge for possible successor and predecessor edges so as to maximize a figure of merit defined as the difference between the maximum arc cost times the path length, and the cost of the path. Edge elements in the path are marked and are not reused again. The process is repeated until all edge elements are contained in an optimal path of at least three elements. Edges in smaller paths are described. A number of pruning techniques are then applied to eliminate parasitic paths. In Mero's technique, junctions or bifurcation points are detected first on the basis of variances in edge orientations. Optimal paths are then found between junctions by determining minimum cost paths that meet a set of well defined constraints.

Although global search techniques perform in general better than local techniques, they have the disadvantage of being computationally expensive, requiring many passes through the image array.

2.3.3 Extracting Linear Segments

The presence of linear features and their interrelationships provide a strong clue for the recognition of man-made structures in aerial images. Methods for extracting **linear segments** can be classified as: piecewise linear approximation methods, Hough transform

methods, and connected components methods. Each of these methods fits a straight line or a simple curve to groups of edge elements. They differ in the way the grouping is performed.

2.3.3.1 Piecewise Linear Approximation

This method starts with an arbitrary edge segment, i.e., a list of edge elements produced by either local or global linking.

A simple technique for dividing the edge segment into a set of approximate straight segments is iterative end point fitting (Duda and Hart, 1973). This technique operates by connecting the end points of the edge segment by a straight line and searching for the point in the segment that is farthest from this line. If this distance is unacceptably large, the edge segment is divided in two segments at the point of maximum excursion and the process is iterated for the two segments. Nevatia and Babu (1980) have applied this technique effectively for the extraction of linear features in aerial images.

Another approach is to divide the edge segment at the points of abrupt change in curvature. Perkins (1978) uses this approach and fits the resulting groups of edge elements to straight lines and circular arcs. Shirai (1977) fits the groups to straight lines and elliptic arcs. Chu (1983) and Nalwa and Pauchon (1987) use straight lines and conics.

2.3.3.2 Hough Transform

This technique is applicable to the extraction of straight lines and any parametric curve. Unlike piecewise linear approximation, it does not require a linking step to obtain initial edge segments but it can operate on the whole array of edge elements.

Duda and Hart (1972) introduced this technique for the detection of straight lines. The basic idea is to transform the set of edge elements described by their row and column positions (r,c) into Hough space described by the pairs (θ, ρ). This transformation is described by the following relationship.

$$\rho = -r \cos\theta + c \sin\theta$$

where θ is the orientation of the edge element at position (r,c) and ρ is the distance from the origin to the line with slope θ passing through the edge element. The parameters θ and ρ are quantized into a number of bins with each bin reporting the number of contributions from edge elements with parameters (θ, ρ) in the corresponding interval. Thus, groups of edge elements belonging to the same line are manifested as local maxima of clusters in Hough space.

O'Gorman and Clowes (1976) weight the contribution of each edge element in Hough space by its gradient strength. Dudani and Luk (1977) compute an smoothed histogram of the θ values and carry out the division at points of local minima in the histogram. Then for each θ-group, a smoothed histogram of the ρ values is computed and further division is carried out at points of local minima. Finally for each θ-ρ group lines are fitted to each cluster of edge elements within a given distance from each other.

2.3.3.3 Connected Components

Another approach to the extraction of linear features is to fit straight lines to connected components of edge elements having similar orientation. We describe two ways in which the connected components are obtained and two ways of fitting lines to them.

Tavakoli and Rosenfeld (1982) obtain connected components by scanning the edge image row by row. On each row a running average of edge orientations is computed and a point is accepted into the run only if its orientation is very close to the current average. This yields runs of points all having approximately the same orientation. Adjacent runs on consecutive rows are linked if their averages are very close to one another. A least squares error line fit is performed on each resulting connected component.

Burns, Hanson, and Riseman (1986) partition the 360 degree range of gradient directions into a small set of regular intervals. Connected components of pixels which belong to the same partition are then computed to serve as **edge-support regions.** All pixels participate in this process irregardless of their gradient strengths. To avoid problems of improper merging of two distinct linear features, or of fragmentation of a single distinct feature, caused by a fixed partition size, two overlapping sets of histogram partitions are used, one set rotated a half-partition interval. Thus, two alternative edge-support regions are computed and then merged. A plane is fit to the intensity surface of the set of pixels in each support region. The fit is weighted by the gradient magnitude associated with the pixels so that the important intensity changes will dominate. Finally a straight line is extracted from the edge-support region by intersecting the fitted plane with a horizontal plane representing the weighted average intensity.

2.4 Linear Segment Organization

A characteristic of man-made structures appearing in aerial images is the presence of a large number of linear segments. These segments are usually highly organized and this organization is manifested primarily in three forms: parallelism, as in the opposite

sides of buildings and roads; collinearity, as in a row of houses; and adjacency, as in the corners formed by adjacent sides of buildings.

The highly organized nature of the linear segments in aerial images of urban scenes leads to the construction of a number of intermediate-level units described in this section. The succesful extraction of these primitives requires in many situations little or no top-down flow of information. Succesful matching of aerial images to models of man-made structures can usually take place at this level.

2.4.1 Parallelism - From Segments to PARs and APARs

Using this property on a pair of linear segments, two useful intermediate-level units can be constructed: parallel pair, abbreviated **PAR**, and antiparallel pair, abbreviated **APAR**. The segments of a PAR have the same direction while the segments of an APAR have opposite directions (Figure 6 on page 33). They can be described by the distance between their members and the amount of overlap. Of these two units APARs have been more widely used.

Nevatia and Babu (1980) search for candidate pairs for APARs among pairs of overlapping segments whose directions are 180 degrees apart within a given tolerance, and that are within a certain distance from each other. When there is more than one possible pairing, the closest pair is chosen. Tavakoli and Rosenfeld (1982) use essentially the same geometrical constraints as Nevatia and Babu while adding a gray level constraint. They require that the average gray levels of the object sides of each member of the pair do not differ by more than ten percent of the gray scale. The average gray level is computed in a strip of width four pixels parallel to the segment.

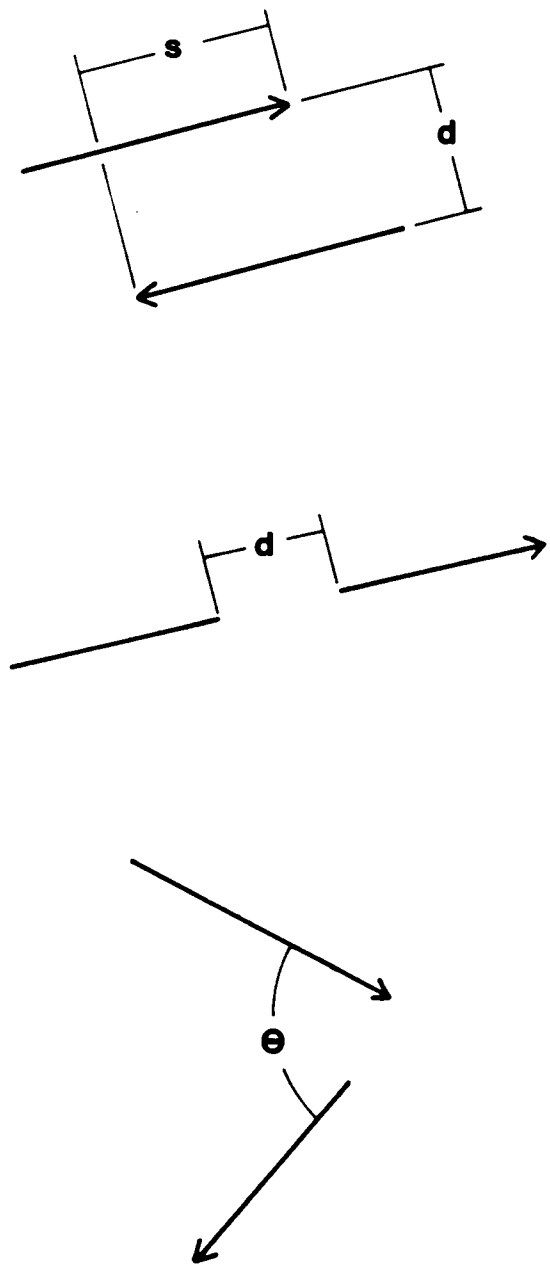


Figure 6. shows the organization of linear features.

Scher et al. (1982a) assign rankings to candidate pairs based on the amount of overlap, the distance between the segments, and their lengths. Ambiguities resulting from segments that compete for the same antiparallel partner are resolved based on best mutual matches.

Further examples of the use of PARs and APARs as intermediate-level units for the recognition of man-made structures in aerial images can be found in Cheng and Huang (1981,1984), Medioni (1984), and Loew (1982). Closely related to APARs is the work of Brooks (1981) with **ribbons**. Ribbons are two-dimensional specializations of generalized cylinders (Agin and Binford, 1973) and are more general than APARs in that a ribbon has a straight axis, but may be bounded by straight lines that are not antiparallel.

2.4.2 Collinearity - From Segments to COLs and ACOLs

As with parallel and antiparallel pairs, two additional intermediate-level units can be defined: collinear pair, abbreviated **COL** and anticollinear pair, abbreviated **ACOL**. Collinear segments have the same direction while anticollinear segments have opposite directions. They can be described by the distance between their members (Figure 6 on page 33).

Nevatia and Babu (1980) use collinearity to merge pairs of APARs into longer APARs. Scher et. al. (1982b) suggest a number of merging methods and splitting methods for finding clusters of collinear collections of line segments. Weiss and Boldt (1986) merge pairs of straight lines based on collinearity, proximity, and similarity of contrast.

Cheng and Huang (1984) use the binary relations parallel, antiparallel, collinear, and adjacent to register two aerial photographs of an airport by matching relational structures.

2.4.3 Adjacency - From Segments to ADJ pairs

Two adjacent segments, i.e., and **ADJ** pair, can be described by their angle (Figure 6 on page 33). They can be required to meet a vertex and form ninety degree **corners** ((Huertas, 1981), (Huertas and Nevatia, 1983)) or they can be said to be **compatible** if they meet certain geometrical and gray level constraints (Tavakoli and Rosenfeld, 1982).

Compatibility implies that a given pair of segments could be consecutive parts of an object boundary. Tavakoli and Rosenfeld (1982) define three conditions that have to be met for two segments to be compatible. Firstly, the average gray level in a strip along the object side of one segment should be close (within a given tolerance) to the average gray level of the other segment and also close to the average gray level in a strip connecting the near ends of the two segments. Secondly, the near ends should not be farther apart than two-thirds the length of the smaller segment or than two-thirds the distance between their far ends. Thirdly, the angle between the two segments should lie in some range [$-180+\alpha$, $180-\alpha$] (a typical value for α is 25 degrees). In other words compatible segments should not meet at a sharp acute or obtuse angle.

Compatible segments can be organized into **closed groups**, that is cyclically closed sequences of compatible segments, and **semiclosed groups**, sequences of compatible segments having a gap, between the first and last segment, shorter than the longest line connecting the ends of a compatible pair in the group. Closed groups and semiclosed groups are regarded as having a high probability of belonging to a building. Closely related are the **connectivity groups** of Clark et al. (1979).

Cheng and Huang (1981) use the adjacency relation in matching relational structures for the recognition of airplanes in aerial images. Like Tavakoli and Rosenfeld, their

definition of adjacency means that the tail of one segment is close to, but not necessarily in touch with, the head of the other edge segment.

Huertas and Nevatia (1983,1988) classify corner pairs as strongly compatible, weakly compatible, and incompatible. Corner compatibility is defined as the positional relationship between every pair of detected corners. Strong compatibility occurs when two corners share a common segment. Two corners are said to be weakly compatible if they belong to one of a number of arrangements that satisfy certain geometrical and shadow casting constraints. Strong and weak compatibility determine the initial likelihood for every pair of corners of being part of the same building outline. Strongly compatible corners are grouped into **corner chains**.

Further examples of the use of adjacency to build intermediate-level primitives can be found in Jacobus et al. (1980,1981), Dudani and Luk (1977), and Herman and Kanade (1986).

2.5 *Systems*

Figure 7 on page 37 illustrates the material covered in this chapter, describing Edge-Based Representations for the recognition of man-made structures in aerial images. The tree structure on the right illustrates the transformations that take place to build higher level units from lower level units, starting at the pixel level up to units formed by groups of linear features. The block diagram on the left illustrates the low-level and intermediate-level computer vision tasks which carry out the transformations illustrated in the tree structure.

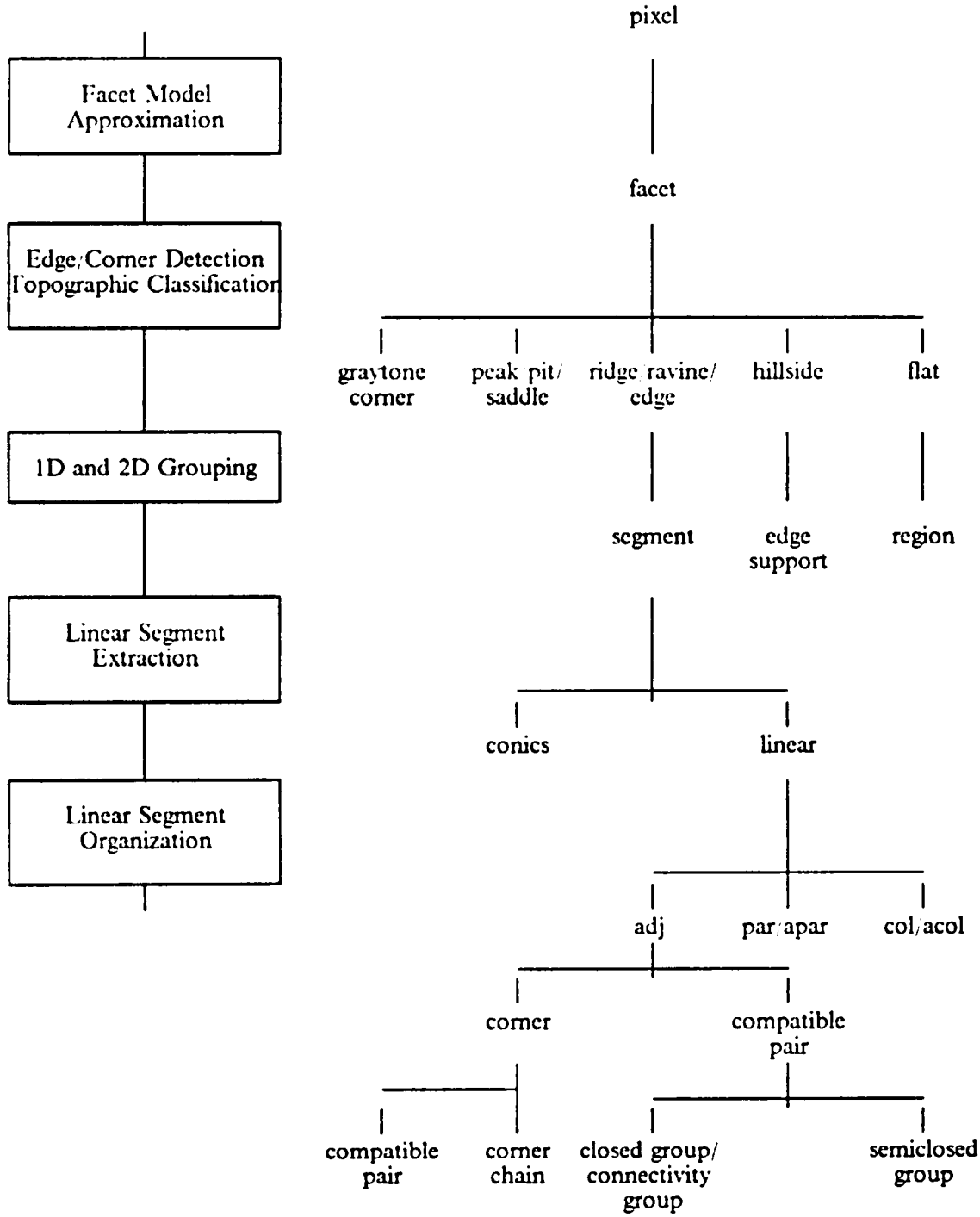


Figure 7. shows transformations from lower to higher level units.

SYSTEM	EDGE DETECTION	1D GROUPING	LINEAR SEGMENT EXTRACTION	LINEAR SEGMENT ORGANIZATION
Nevatia and Babu	5x5 templates, fixed threshold	thinning, local linking	piecewise linear approximation, iterative end point	apars, collinearity
Tavakoli and Rosenfeld	Sobel, Relaxation	Connected Components by angle	LSE line fit	apars, compatible segments, closed groups
Cheng and Huang	Sobel, fixed threshold	thinning, local linking	piecewise linear approximation	par, apar, col, adj
Burns et al.	Roberts, no threshold	Connected Components by angle	edge support line fit	parallellism
Dudani and Luk	Hueckel, fixed threshold	Hough Transform	Principal Components line fit	parallelism, vertex model
Perkins	Hueckel, fixed threshold	thinning, local linking	piecewise linear approximation, curvature based partitioning	list of lines and circular arcs
Huertas and Nevatia	Zero-crossing of Laplacian of Gaussian	thinning, local linking	piecewise linear approximation, iterative end point	corners, compatible corners, corner chain
Clark et al.	5x5 templates, fixed threshold	thinning, local linking	piecewise linear approximation, iterative end point	adj, connectivity groups
Zuniga and Haralick	Zero-crossing of 2nd directional derivative, Integrated gradient, automatic threshold	thinning, local linking	piecewise linear approximation, iterative end point, positional and angular line fit	parallelism, collinearity, corners, segment chains, cultural objects

Figure 8. shows a number of systems that utilize edge-based representations to extract linear features.

A number of systems, comprising a fairly good sample of the research work in this area, is shown in Figure 8 on page 38. They include Nevatia and Babu (1980), Tavakoli and Rosenfeld (1982), Cheng and Huang (1981,1984), Burns, Hanson, and Riseman (1986), Dudani and Luk (1977), Perkins (1978), Huertas and Nevatia (1983,1988), and Clark et al. (1979). Also shown is the system developed in this dissertation. Each system is described according to the organization of this chapter, that is, the type of edge operator and gradient threshold selection, grouping methods, linear segment extraction, and the ways these segments are organized.

Chapter 3

INTEGRATED DIRECTIONAL DERIVATIVE GRADIENT OPERATOR

3.1 Introduction

The computation of edge direction is a required step in many image processing tasks. Hough transformation techniques (Duda and Hart, 1972) have been used extensively together with edge direction information to detect lines (Gorman and Clowes, 1976), circles (Kimme, Ballard, and Sklansky, 1975), and arbitrary shapes (Ballard, 1981). Martelli (1972) and Ramer (1975) each use edge direction information to perform edge linking. Kitchen and Rosenfeld (1980), and Zuniga and Haralick (1983) use edge direction information in schemes to detect corners.

A variety of operators for computing local edge direction have been proposed, many of them estimating a kind of gradient of the picture function. Local edge direction is then

estimated as that direction which is orthogonal to the gradient direction. Knowledge of the directional derivatives D_1 and D_2 in any two orthogonal directions is sufficient to compute the directional derivative in any arbitrary direction. The gradient magnitude which is defined as the maximum such directional derivative is computed as $\sqrt{D_1^2 + D_2^2}$ and its direction as $\tan^{-1} D_2/D_1$. A number of local operators have been utilized to estimate these directional derivatives. Examples are the Roberts operator (Roberts, 1965), the Prewitt operator (Prewitt, 1970), the Sobel operator (Duda and Hart, 1973), and the Hueckel operator (Hueckel, 1971). Of special interest are those operators which result from a local surface fit to the graytone image data and subsequent computation of directional derivatives in two orthogonal directions from the analytic description of the estimated surface. Thus the original Roberts operator has been shown to result from a linear surface fit within a 2x2 pixel neighborhood (Haralick, 1980). The original Prewitt operator was derived from a quadratic surface fit within a 3x3 pixel neighborhood. Haralick (1982), and Zuniga and Haralick (1983) compute edge direction information from a cubic polynomial surface fit. This fitting idea can also be extended to arbitrary sized neighborhoods (Brooks, 1978), (Haralick, 1980).

These gradient operators face two major problems. One problem is that their estimate of edge direction is inherently biased as a function of true edge direction and displacement of the true edge from the pixel's center. The second major problem is the sensitivity of these operators to data noise. An obvious approach to decrease noise sensitivity is to increase the neighborhood size. However, this creates problems with small or thin objects. Also for all the gradient operators mentioned above an increase in neighborhood size causes an increase in edge direction estimate bias.

We use an operator based on a cubic polynomial local surface fit. This cubic facet model has been successfully used in the past to detect edges (Haralick, 1982), topographic features (Haralick, Watson, and Laffey, 1983), and corners (Zuniga and Har-

alick, 1983). Instead of computing directional derivatives directly from the fitted surface as in the case of the standard cubic facet gradient operator mentioned earlier, the new operator measures gradient strength as the maximum value of the integral of first directional derivative taken over lines going in all possible directions. The line direction which maximizes the integral defines the new estimate of gradient direction. Experiments show that this operator posseses two main characteristics:

1. Edge direction estimate bias is sharply reduced as compared with the bias of the standard cubic facet, Sobel, and Prewitt gradient operators. Noise sensitivity is comparable to that of the Sobel and Prewitt operators and much better than the standard cubic facet operator.

2. Unlike the standard cubic facet, Sobel, and Prewitt operators increasing the neighborhood size decreases both estimate bias and noise sensitivity. For ramp edges, the integrated operator is very nearly unbiased. The worst bias for the 7x7 operator is less than 0.09 degrees and the worst bias for the 5x5 operator is less than 0.26 degrees.

Section 3.2 describes the standard cubic facet gradient operator. Section 3.3 describes the mathematical analysis necessary to derive the new gradient estimate. Section 3.4 presents experimental results and provides a comparison of the integrated directional derivative gradient operator against the standard cubic facet gradient operator, Prewitt operator, and Sobel operator, for step edges and ramp edges contaminated by zero mean Gaussian noise. Appendix A describes the mathematical analysis of the cubic facet.

3.2 *The standard cubic facet gradient operator*

Given a graytone intensity function f defined in the row and column coordinate system of a given pixel neighborhood, the gradient vector function ∇f is given by

$$\nabla f = \left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial c}\right) \tag{3.1}$$

Under the cubic facet model each surface facet centered about a given pixel is approximated by the bivariate cubic form:

$$\begin{aligned} f(r,c) = & K_1 \\ & + K_2 r + K_3 c \\ & + K_4 r^2 + K_5 rc + K_6 c^2 \\ & + K_7 r^3 + K_8 r^2 c + K_9 rc^2 + K_{10} c^3 \end{aligned} \tag{3.2}$$

where the K coefficients change from neighborhood to neighborhood and are estimated using a least square error surface fit. A detailed analysis of this estimation procedure is provided in appendix A. Evaluating the first row and column partial derivatives of equation (3.2) at the neighborhood center (0,0) and replacing their values in equation (3.1) the gradient vector at the neighborhood center becomes

$$\nabla f = (K_2, K_3)$$

The magnitude of the standard cubic facet gradient operator is therefore given by $\sqrt{K_2^2 + K_3^2}$ and the gradient direction θ, measured as a clockwise angle from the column axis, satisfies

$$\sin\theta = K_2 / \sqrt{K_2^2 + K_3^2}$$

$$\cos\theta = K_3/\sqrt{K_2^2 + K_3^2}$$

3.3 *The integrated directional derivative gradient operator*

For a given direction vector ($\sin\theta, \cos\theta$) , it is well known that the first directional derivative $f'_\theta(r, c)$ of f in the direction θ can be evaluated as the component of the gradient ∇f along the direction vector, that is

$$f'_\theta(r,c) = \frac{\partial f}{\partial r}\sin\theta + \frac{\partial f}{\partial c}\cos\theta \tag{3.3}$$

Let F_θ be defined as:

$$F_\theta = \frac{1}{4LW}\int_{-W}^{W}\int_{-L}^{L} f'_\theta(\rho\cos\theta + \omega\sin\theta, -\rho\sin\theta + \omega\cos\theta)d\rho d\omega \tag{3.4}$$

for a given N x N neighborhood.

F_θ represents the integrated first directional derivative along lines orthogonal to the direction θ forming a rectangle of length 2L and width 2W centered at the origin of the coordinate system. The proposed gradient estimate is:

$$G = F_{\theta_{MAX}}\, u_{\theta_{MAX}} \tag{3.5}$$

where $F_{\theta_{MAX}} = \max_\theta F_\theta$ and $u_{\theta_{MAX}}$ is a unit vector in the direction that maximizes F_θ.

Using equations (3.2) and (3.3) $f'_\theta(\rho\cos\theta + \omega\sin\theta, -\rho\sin\theta + \omega\cos\theta)$ reduces to:

$$\begin{aligned} f'_\theta&(\rho\cos\theta + \omega\sin\theta, -\rho\sin\theta + \omega\cos\theta) \\ &= [3(K_9 - K_7)\sin^3\theta + 3(K_8 - K_{10})\cos^3\theta \\ &+ (3K_7 - 2K_9)\sin\theta + (3K_{10} - 2K_8)\cos\theta]\rho^2 \\ &+ 2[-K_8\sin^3\theta + (3K_7 - 2K_9)\sin^2\theta\cos\theta \\ &+ (2K_8 - 3K_{10})\sin\theta\cos^2\theta + K_9\cos^3\theta]\rho\omega \\ &+ [3(K_7 - K_9)\sin^3\theta + 3(K_{10} - K_8)\cos^3\theta \\ &+ 3K_9\sin\theta + 3K_8\cos\theta]\omega^2 \\ &+ [-K_5\sin^2\theta + 2(K_4 - K_6)\sin\theta\cos\theta + K_5\cos^2\theta]\rho \\ &+ 2[K_4\sin^2\theta + K_5\sin\theta\cos\theta + K_6\cos^2\theta]\omega \\ &+ K_2\sin\theta + K_3\cos\theta \end{aligned} \tag{3.6}$$

Substituting equation (3.6) in equation (3.4) results in:

$$F_\theta = \frac{1}{4LW}\int_{-W}^{W}\int_{-L}^{L}(A\rho^2 + B\rho\omega + C\omega^2 + D\rho + E\omega + F)d\rho d\omega$$

where A, B, C, D, E and F are the coefficients of the quadratic equation (3.6). Evaluating this integral results in:

$$F_\theta = \frac{1}{3}AL^2 + \frac{1}{3}CW^3 + F$$

Finally,

$$\begin{aligned} F_\theta &= (K_9 - K_7)(L^2 - W^2)\sin^3\theta + (K_8 - K_{10})(L^2 - W^2)\cos^3\theta \\ &+ [L^2K_7 - \tfrac{1}{3}(2L^2 - 3W^2)K_9 + K_2]\sin\theta \\ &+ [L^2K_{10} - \tfrac{1}{3}(2L^2 - 3W^2)K_8 + K_3]\cos\theta \end{aligned} \tag{3.7}$$

Thus, F_θ reduces to a trigonometric expression in $\sin\theta$ and $\cos\theta$. Notice that if L = W, the cubic terms $\sin^3\theta$ and $\cos^3\theta$ drop and F_θ is maximized when

$$\sin\theta = D_1/\sqrt{D_1^2 + D_2^2} \tag{3.8}$$

$$\cos\theta = D_2/\sqrt{D_1^2 + D_2^2}$$

where

$$D_1 = L^2K_7 + \frac{1}{3}L^2K_9 + K_2$$

$$D_2 = L^2K_{10} + \frac{1}{3}L^2K_8 + K_3$$

Then

$$F_{\theta_{MAX}} = \sqrt{D_1^2 + D_2^2} \tag{3.9}$$

The computation of the gradient strength given by equation (3.9) and gradient direction given by equation (3.8) can be simplified by precomputing 'row derivative' and 'column derivative' masks from the numerator and denominator of equation (3.8) respectively using the masks for the K parameters derived in appendix A. As a special case when L = W = 0, then equations (3.8) and (3.9) reduce to the standard cubic facet operator defined in section 3.2.

If $L \neq W$, then the maximization of F_θ to obtain the estimated gradient strength $F_{\theta_{MAX}}$ and estimated gradient direction θ_{MAX} can be carried out by the following procedure:

From equation (3.7) F_θ can be expressed as:

$$F_\theta = u \sin^3\theta + v \cos^3\theta + w \sin\theta + z \cos\theta \tag{3.10}$$

where u, v, w, and z are functions of the K parameters. Then:

$$F'_\theta = 3u \sin^2\theta \cos\theta - 3v \cos^2\theta \sin\theta + w \cos\theta - z \sin\theta$$

Equating F'_θ to zero to obtain an extremum results in:

$$(3u(1 - \cos^2\theta) + w) \cos\theta = (3v \cos^2\theta + z) \sin\theta \tag{3.11}$$

Let $x = \cos^2\theta$, then equation (3.11) becomes:

$$(3u(1 - x) + w)\sqrt{x} = (3vx + z)\sqrt{1 - x}$$

Squaring this expression and after some algebraic manipulation we finally obtain the cubic expression in x:

$$9(u^2 + v^2)x^3 - (18u^2 + 6uw + 9v - 6vz)x^2 + (9u^2 + 6uw + w^2 - 6vz + z^2)x - z^2 = 0 \tag{3.12}$$

Then, $F_{\theta_{MAX}}$ and θ_{MAX} can be obtained by the following steps:

1. Find all the roots x of equation (3.12) between 0 and 1. These roots must exist since the original equation is trigonometric and therefore periodic, with period 2π.

2. Convert these x's to all possible $\theta's$ from $x = \cos^2\theta$.

3. Evaluate F_θ at all these $\theta's$.

4. Choose maximum value $F_{\theta_{MAX}}$ and θ_{MAX} .

3.4 *Experimental Results*

The experiments were performed with step and ramp edges contaminated by zero mean Gaussian noise. Step edges were generated in a rectangular grid with orientations θ from zero to ninety degrees, and with random displacement from the grid's center uniformly distributed within the range (-D,D) with the maximum displacement D given by

$$D = \begin{cases} 0.5 \cos\theta, & 0 \le \theta < 45 \\ 0.5 \sin\theta, & 45 \le \theta \le 90 \end{cases}$$

assuming a unit distance between two 4-neighbor pixels in the grid. A step edge passing through a pixel divides it into two parts having areas A_1 and A_2, with $A_1 + A_2 = 1$. Let the corresponding graytone intensities to each side of the edge be I_1 and I_2. The pixel is then assigned a graytone intensity I according to the rule

$$I = I_1 A_1 + I_2 A_2$$

The experiments were performed with values for I_1 and I_2 equal to 100 and 200 respectively, that is the edge contrast was set to 100.

Ramp edges were generated by defocusing step edges with a 3x3 averaging filter. Finally, both step and ramp edges were contaminated by adding zero mean Gaussian noise with a given standard deviation.

3.4.1 Tuning up the integrated directional derivative gradient operator.

The performance of the integrated directional derivative gradient operator depends on the choice of integration limits L and W. As seen in section 3.3 this operator is equivalent to the standard cubic facet gradient operator when L = W = 0. We expect its performance to improve as L and W move away from zero and to reach a maximum for some value of $L \neq 0$, and some value of $W \neq 0$. We are primarily interested in improving the edge direction estimate. We use two performance measurements, edge direction estimate bias and edge direction estimate standard deviation. The latter measures noise sensitivity. The estimate bias is defined as the difference between the estimate mean and the true edge direction. A single performance measurement to compare two sets of values of L and W is obtained by combining the previous two measurements into the root mean square error of the estimate which can be shown to be equal to the square root of the sum of the square bias and the estimate variance. It was observed that the operator achieved best performance in the root mean square error sense when L = W = 1.8, for a 5x5 neighborhood size, and L = W = 2.5, for a 7x7 neighborhood size, for both step and ramp edges, and for a variety of noise levels.

3.4.2 Comparing different gradient operators.

The following gradient operators were tested: 5x5 extended Sobel (Iannino and Shapiro (1979)), 5x5 and 7x7 Prewitt, 5x5 and 7x7 standard cubic facet, and 5x5 and 7x7 integrated directional derivative. Figure 9 on page 55 shows the 5x5 row derivative masks for each of the operators tested and Figure 10 on page 56 shows the 7x7 row

derivative mask for the integrated directional derivative gradient operator. The column derivative masks can be obtained from the row masks by transposition.

For a step or ramp edge of a given orientation and noise standard deviation each operator was applied to the grid's center ten thousand times, each time with a different noise sample, and a different edge displacement from the grid's center. Under zero noise conditions the operators were applied only one hundred times. Edge orientations varied from zero to ninety degrees and noise standard deviation from zero to one hundred. Edge contrast was set to one hundred. The edge direction estimates produced by each operator were plotted as follows. For a fixed noise standard deviation, estimate bias against true edge direction, and estimate standard deviation against true edge direction. For a fixed edge direction, estimate bias against noise standard deviation, and estimate standard deviation against noise standard deviation.

We compare first the standard cubic facet gradient operator and the integrated directional derivative gradient operator. Under zero noise conditions Figure 11 on page 57 and Figure 12 on page 58 show estimate bias against true edge direction for step and ramp edges. Three things can be observed. First, the integrated operator is clearly superior to the standard cubic facet gradient operator. Under zero noise conditions the 7x7 integrated directional derivative gradient operator has a worst bias of less than 0.09 degrees and the 5x5 integrated directional derivative gradient operator has a worst bias of less than 0.26 degrees on ramp edges. For comparison purposes the 7x7 standard cubic facet gradient operator has a worst bias of about 1.2 degrees and the 5x5 standard cubic facet gradient operator has a worst bias of 0.5 degrees. This improvement in worst bias stays with the contamination of edges by additive independent zero mean Gaussian noise. Second, for the integrated operator, estimate bias decreases as the neighborhood size increases while the opposite happens with the standard cubic facet gradient operator. Third, both operators perform better with ramp edges than with step edges.

Figure 13 on page 59 through Figure 16 on page 62 show estimate bias and estimate standard deviation against noise standard deviation for a fixed edge orientation of 22.5 degrees and additive independent Gaussian noise. Again, the integrated operator is uniformly superior to the standard cubic facet gradient operator for both step and ramp edges.

Next, we compare the integrated directional derivative operator with the Prewitt and extended Sobel operator. Under zero noise conditions Figure 17 on page 63 and Figure 18 on page 64 show estimate bias as a function of true edge direction for step and ramp edges. The 7x7 integrated operator has the smallest bias followed by the 5x5 integrated operator, the 5x5 extended Sobel, and the 5x5 and 7x7 Prewitt operators. Notice that for ramp edges the response of the integrated operator is nearly flat about zero, that is the operator is nearly unbiased. For the 7x7 integrated operator the worst bias is less than 0.09 degrees, and for the 5x5 integrated operator the worst bias is less than 0.26 degrees. For comparison purposes the worst bias in the 7x7 Prewitt operator is about 5 degrees and the worst bias in the 5x5 Prewitt operator is about 4 degrees. Again, the integrated operator is the only one for which bias decreases as neighborhood size increases. Only the 5x5 Sobel operator is shown, but as previously demonstrated by Iannino and Shapiro (1979), the 3x3 Sobel operator has a smaller bias than the 5x5 extended Sobel, but it is still significantly larger than the bias of the integrated operator and with a much worse noise sensitivity. Iannino and Shapiro (1979) also show results with 3x3 and 5x5 iterated Sobel operators. The bias for these iterated operators is still larger than the bias of the integrated operator, and they are more expensive computationally.

Figure 19 on page 65 through Figure 22 on page 68 show estimate bias and estimate standard deviation as a function of true edge direction for step and ramp edges, when the noise standard deviation is equal to 25. The bias for all the operators shown

is nearly identical to the bias under zero noise conditions. It can be seen from the plots of estimate standard deviation that as expected the 7x7 operators are less sensitive to noise than the 5x5 operators. The estimate standard deviations for the integrated operator and the Prewitt operator are about the same. The Sobel operator has a slightly larger estimate standard deviation.

Figure 23 on page 69 through Figure 26 on page 72 show estimate bias and estimate standard deviation as a function of noise standard deviation for a fixed edge orientation of 22.5 degrees. Several things can be observed from these plots. First, estimate bias for all the operators remains nearly flat as the noise level increases up to about a standard deviation of 90. Some of the operators show an increase in estimate bias at this point. The smallest bias corresponds to the integrated operator followed by the Sobel. The Prewitt operator shows the largest bias. Second, the estimate standard deviation for all operators increases linearly with an increase in noise standard deviation. The Prewitt operator has the smallest estimate standard deviation followed closely by the integrated operator and the Sobel operator. The 7x7 operators have a much smaller standard deviation than the 5x5 operators. For all the operators, ramp edges produce smaller estimate bias than step edges, while step edges produce smaller estimate standard deviation.

Finally we show the gradient strength response of each of the 5x5 operators on the two images shown in Figure 27 on page 73. The first image is a 64x64 synthetic image consisting of a bright triangle on a dark background with zero mean Gaussian noise added to it. The background gray level is 50, the object gray level is 150 and the standard deviation of the noise is 30. The second image is a 64x64 aerial scene. Figure 28 on page 74 and Figure 29 on page 75 show the gradient strength response for each of the 5x5 operators. The integrated directional derivative gradient operator and the Sobel operators yield edges with similar amounts of blur and both operators produce a good performance in the presence of noise. The Prewitt operator has also a good performance in

the presence of noise but produces a larger amount of edge blur. The standard cubic facet gradient operator yield edges with the least amount of blur but has a poor performance in the presence of noise.

3.5 Conclusions

A gradient operator based on an integrated directional derivative on a cubic facet has been investigated. Experimental results with step and ramp edges contaminated by zero mean Gaussian noise show that this operator posseses the following characteristics:

1. Edge direction estimate bias is sharply reduced as compared with the bias of the standard cubic facet, Sobel, and Prewitt gradient operators. Noise sensitivity is comparable to that of the Sobel and Prewitt operators and much better than the standard cubic facet gradient operator.

2. Unlike the standard cubic facet, Sobel, and Prewitt gradient operators increasing the neighborhood size decreases both estimate bias and noise sensitivity. For ramp edges, the integrated operator is very nearly unbiased. The worst bias for the 7x7 operator is less than 0.09 degrees and the worst bias for the 5x5 operator is less than 0.26 degrees. In comparison the worst bias for the 7x7 Prewitt operator is about 5 degrees and the worst bias for the 5x5 Prewitt operator is about 4 degrees.

3. Edge strength response in the presence of noise is as good as that of the Sobel operator and better than the response of the Prewitt and standard cubic facet gradient operators.

4. Computational complexity is the same as the complexity of the Sobel, Prewitt, and standard cubic facet gradient operators since it only involves the application of precomputed row and column derivative masks.

1/10500

-116	-530	-668	-530	-116
-128	-335	-404	-335	-128
0	0	0	0	0
128	335	404	335	128
116	530	668	530	116

1/420

31	-5	-17	-5	31
-44	-62	-68	-62	-44
0	0	0	0	0
44	62	68	62	44
-31	5	17	5	-31

1/50

-2	-2	-2	-2	-2
-1	-1	-1	-1	-1
0	0	0	0	0
1	1	1	1	1
2	2	2	2	2

1/240

-5	-8	-10	-8	-5
-4	-10	-20	-10	-4
0	0	0	0	0
4	10	20	10	4
5	8	10	8	5

Figure 9. shows the row derivative masks for 5x5 gradient operators clockwise from top left corner: Integrated Directional Derivative, Standard Cubic Facet, Extended Sobel, and Prewitt

1/28224

-3	-348	-555	-624	-555	-348	-3
-142	-372	-510	-556	-510	-372	-142
-113	-228	-297	-320	-297	-228	-113
0	0	0	0	0	0	0
113	228	297	320	297	228	113
142	372	510	556	510	372	142
3	348	555	624	555	348	3

Figure 10. shows the row derivative mask for the 7x7 Integrated Directional Derivative gradient operator.

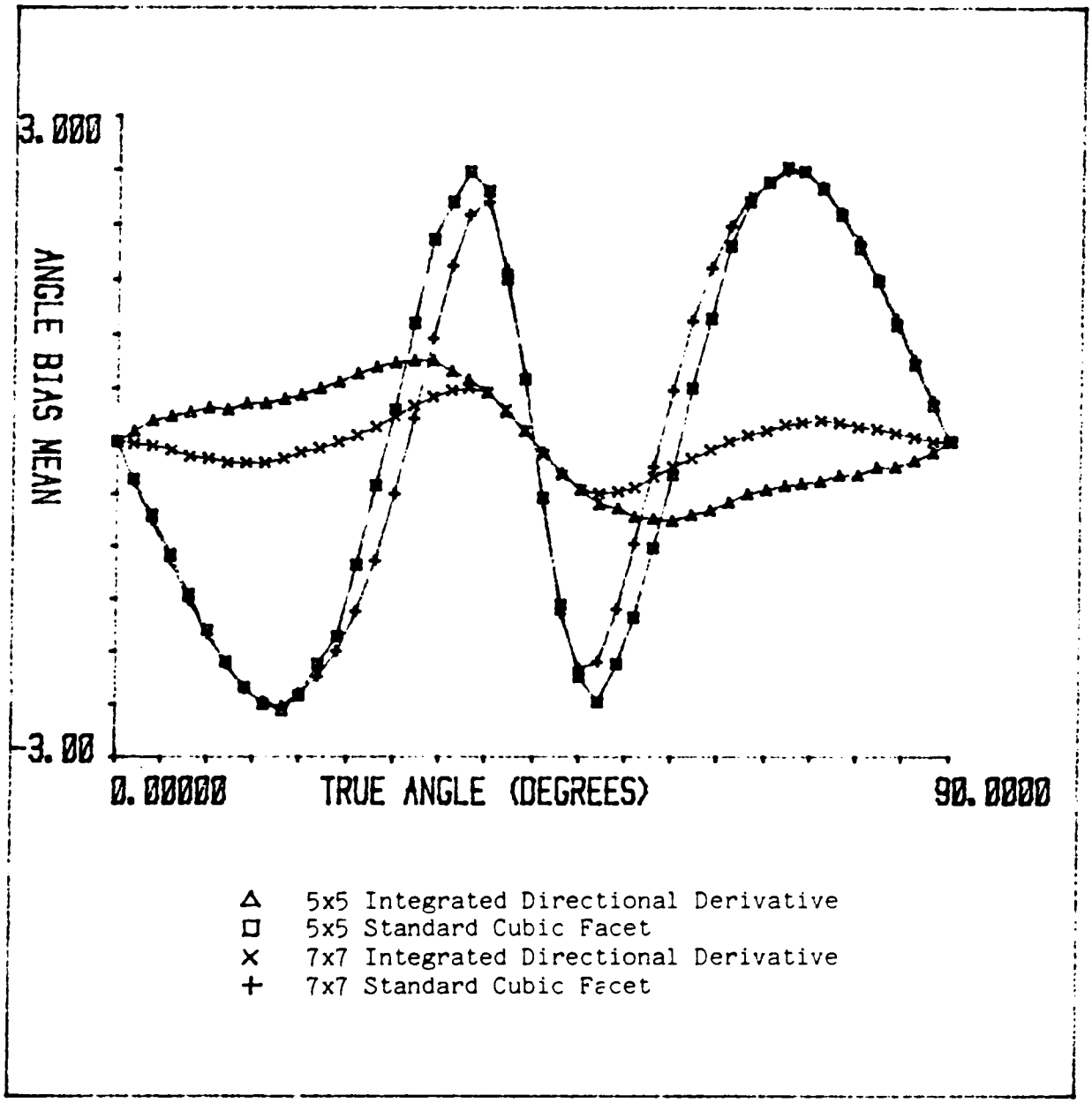


Figure 11. shows estimate bias as a function of true edge direction for step edges under zero noise conditions.

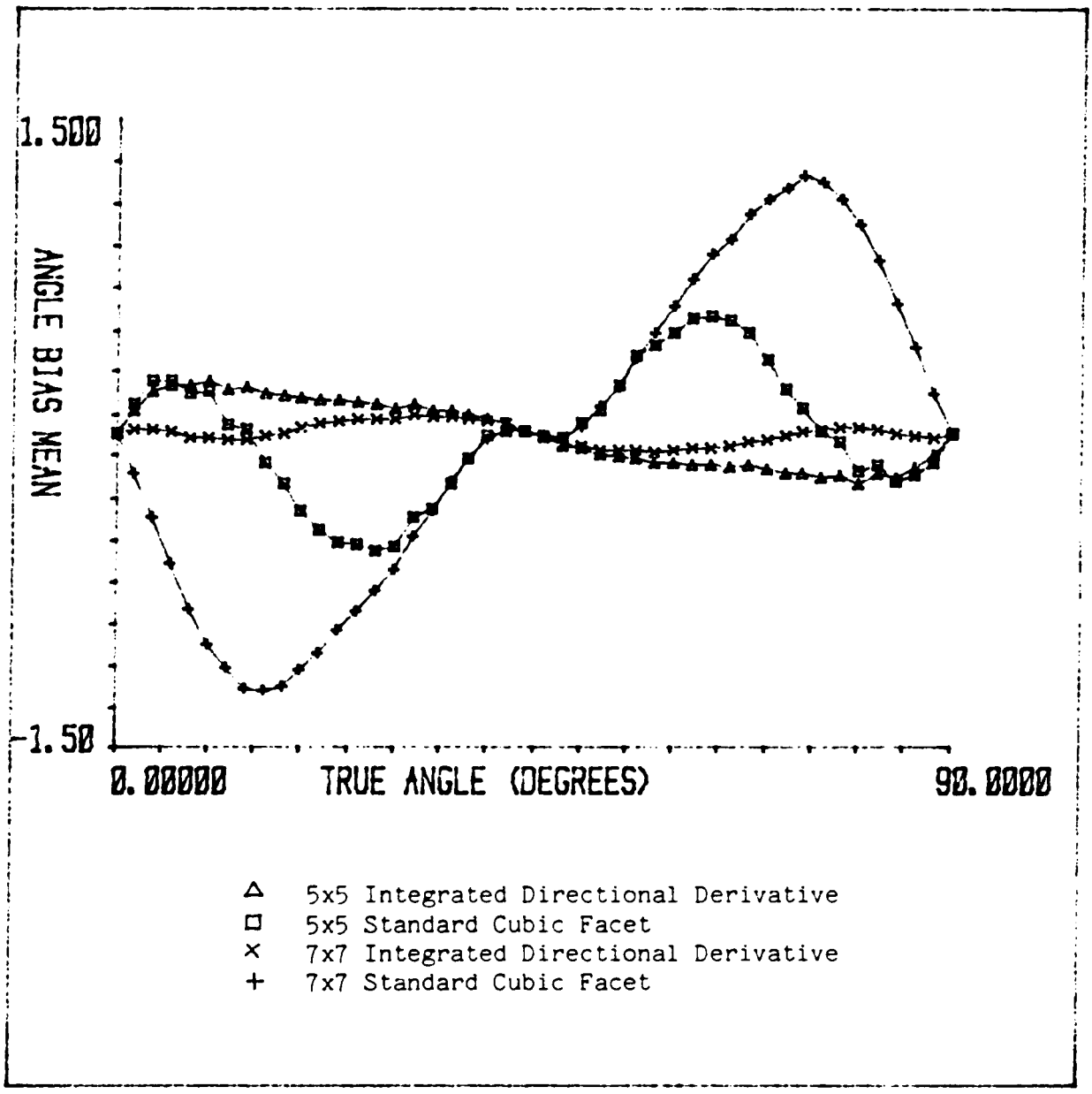


Figure 12. shows estimate bias as a function of true edge direction for ramp edges under zero noise conditions.

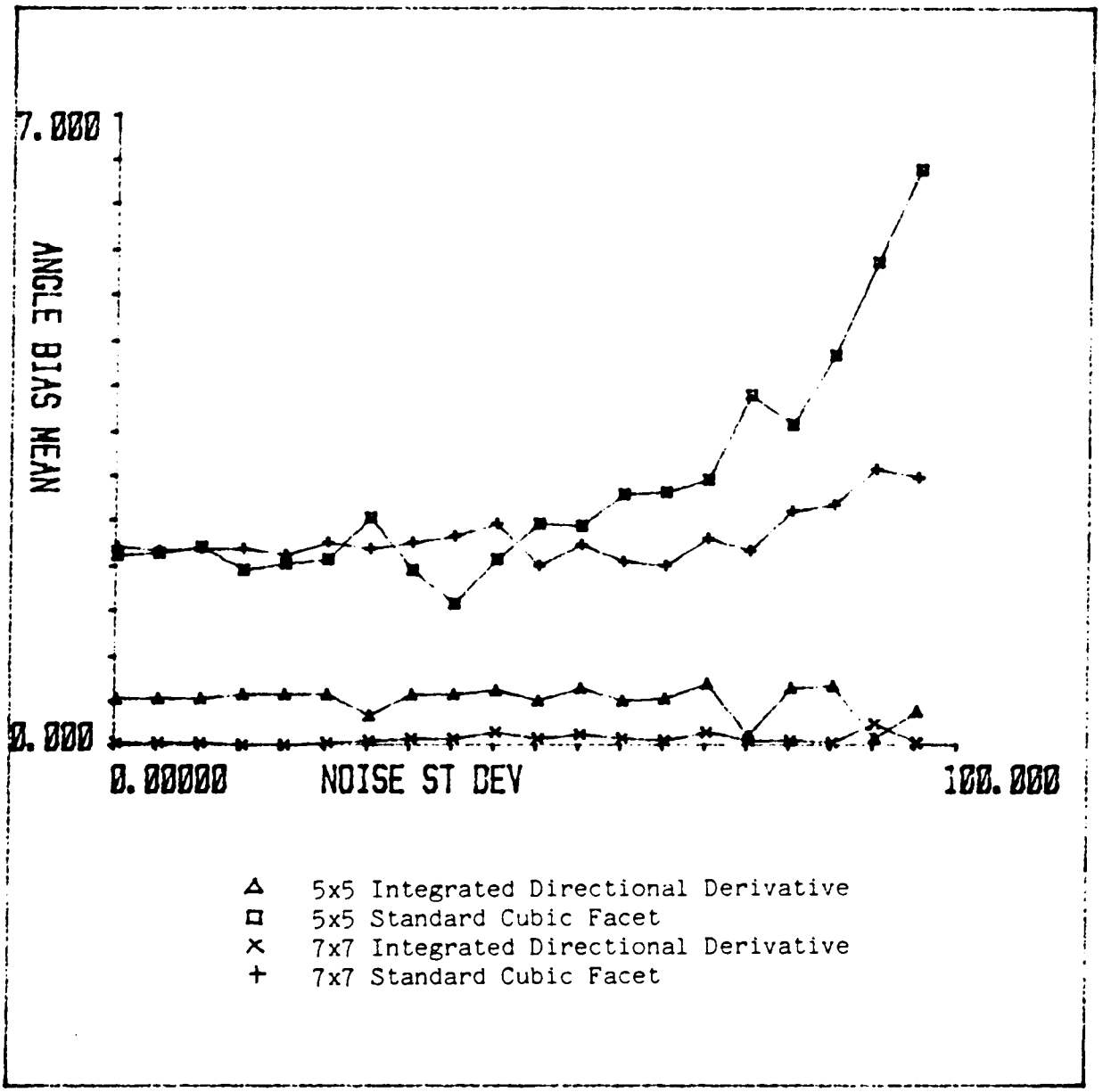


Figure 13. shows estimate bias as a function of noise standard deviation for a step edge: Edge orientation is 22.5 degrees. Edge contrast is 100.

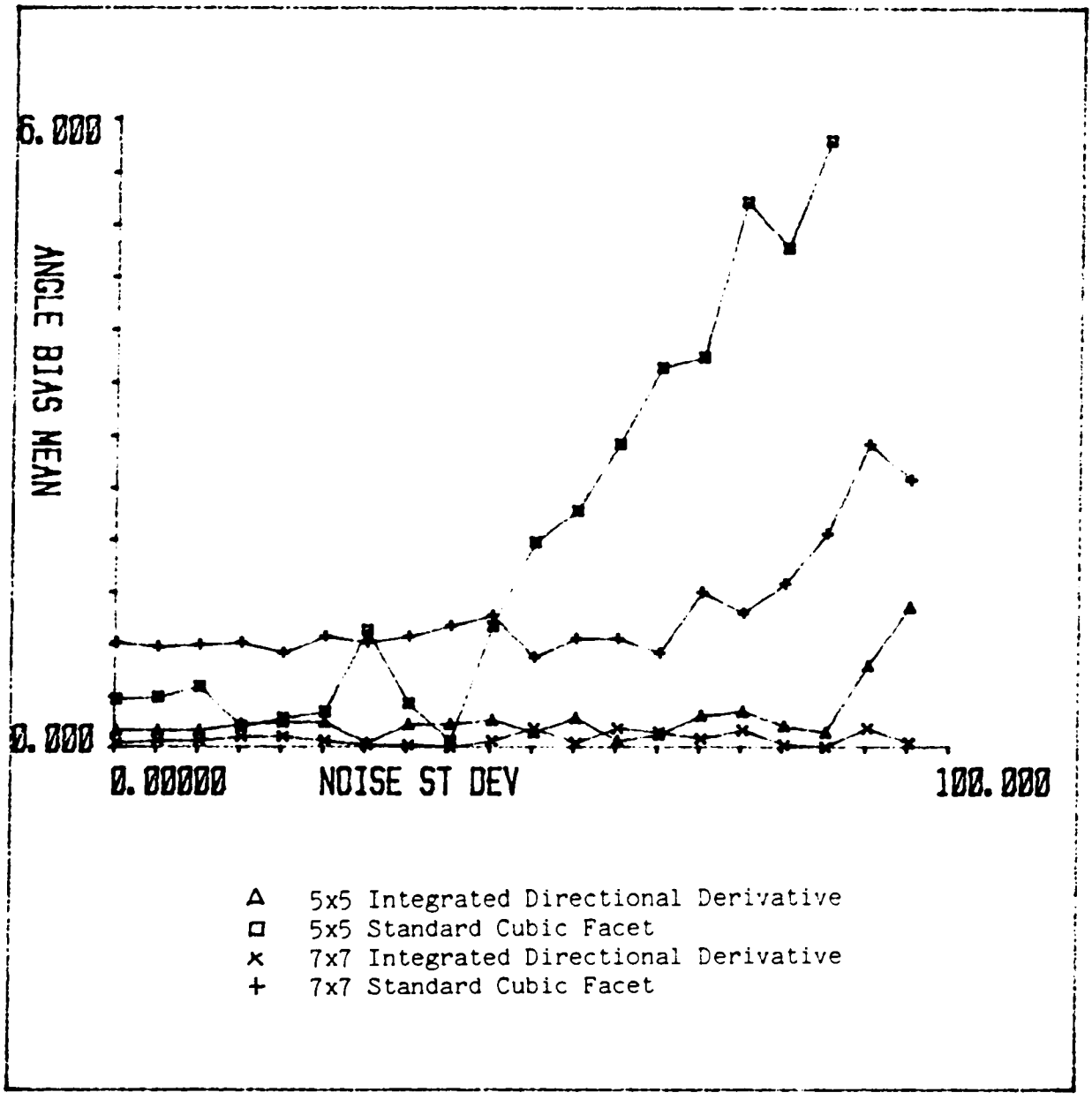


Figure 14. shows estimate bias as a function of noise standard deviation for a ramp edge: Edge orientation is 22.5 degrees. Edge contrast is 100.

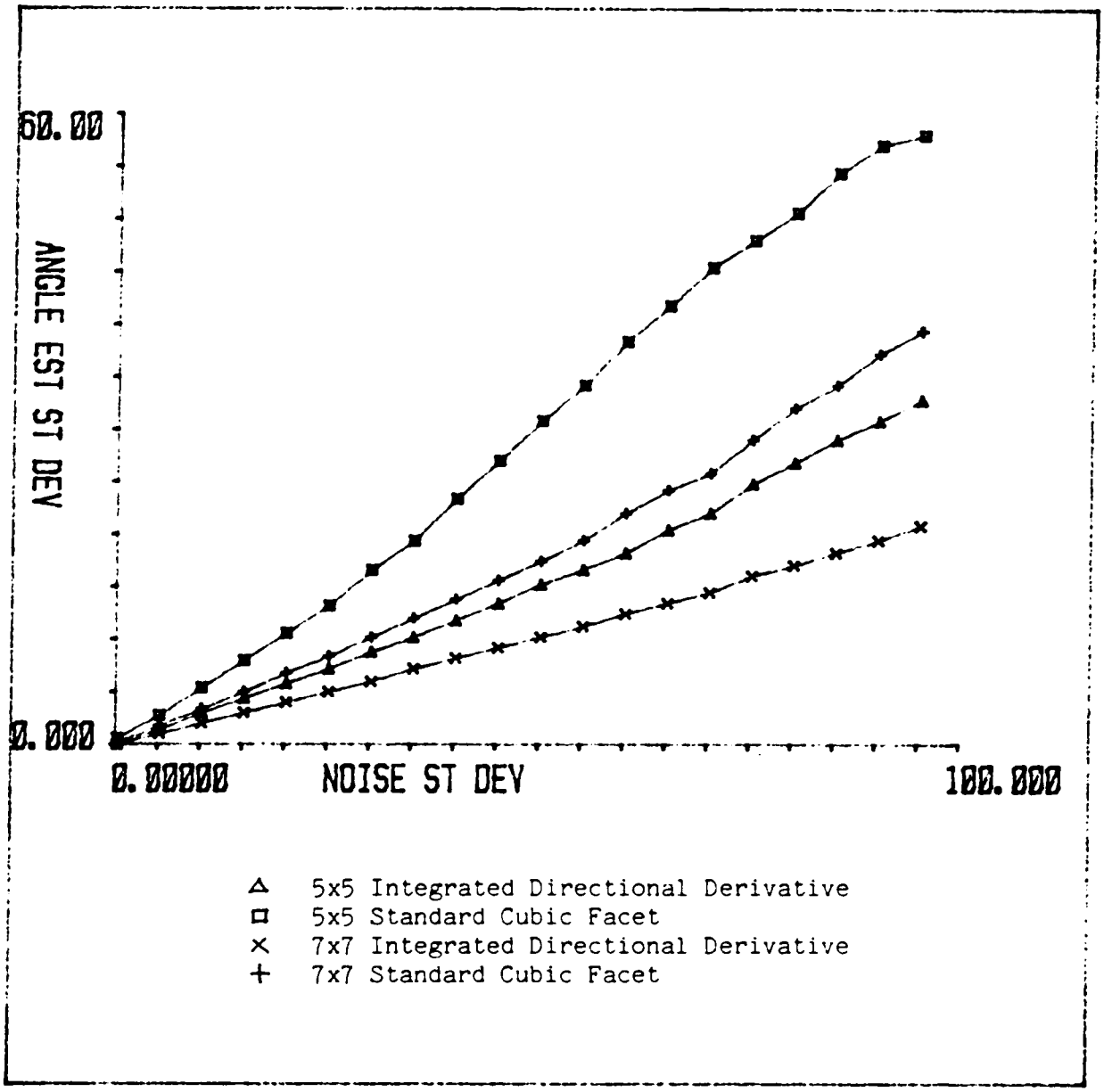


Figure 15. shows estimate standard deviation as a function of noise standard deviation for a step edge: Edge orientation is 22.5 degrees. Edge contrast is 100.

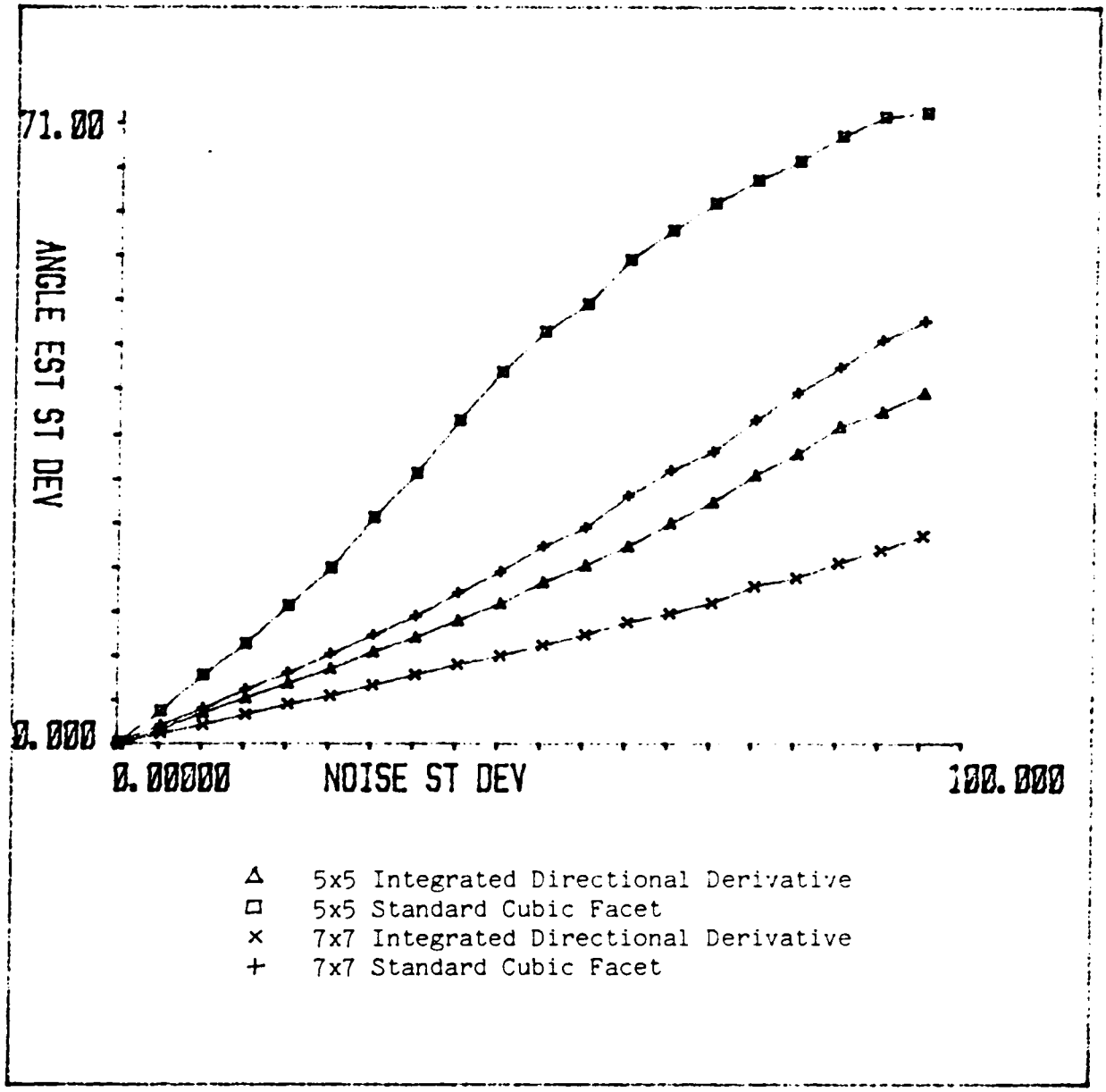


Figure 16. shows estimate standard deviation as a function of noise standard deviation for a ramp edge: Edge orientation is 22.5 degrees. Edge contrast is 100.

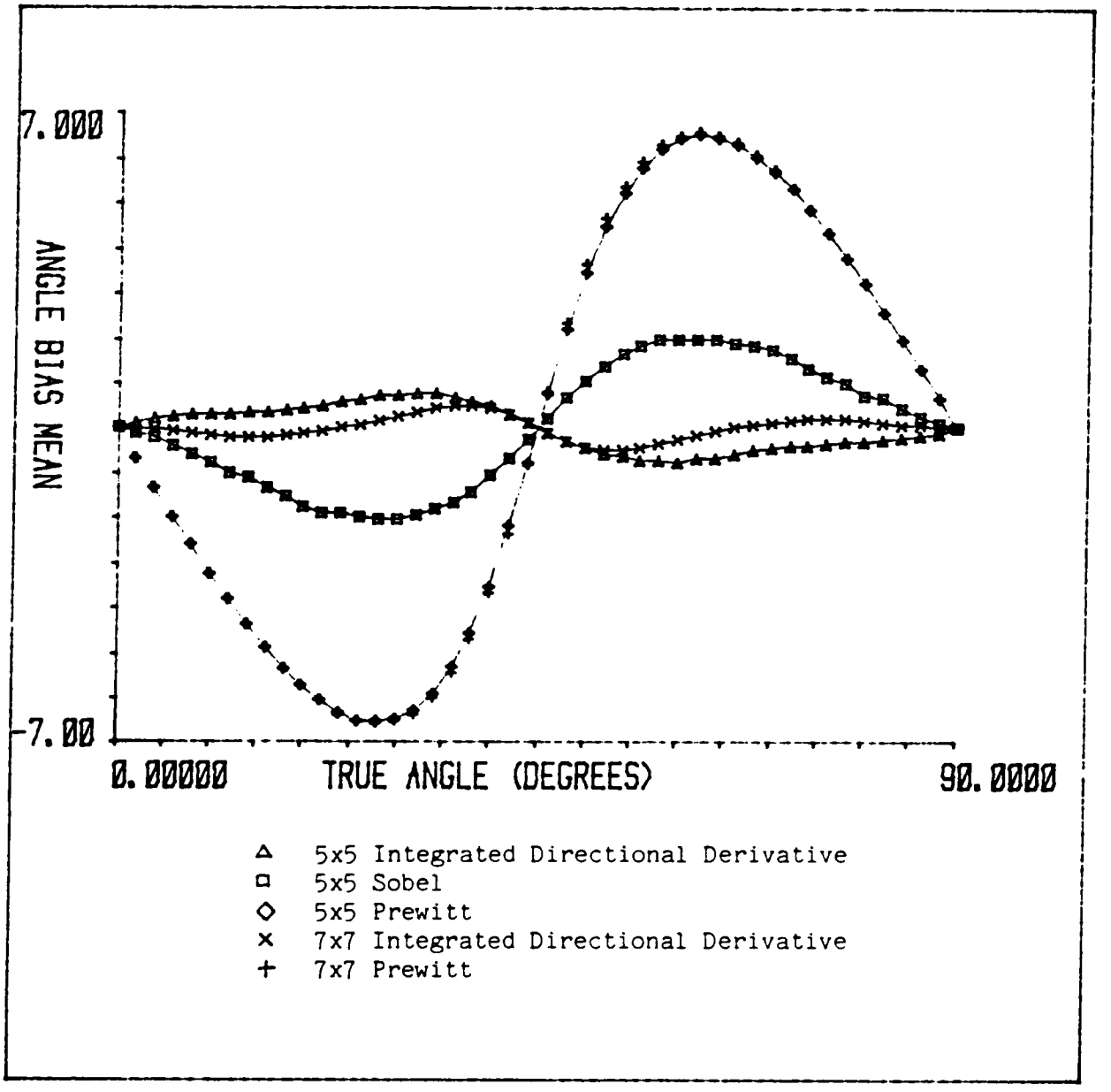


Figure 17. shows estimate bias as a function of true edge direction for step edges under zero noise conditions.

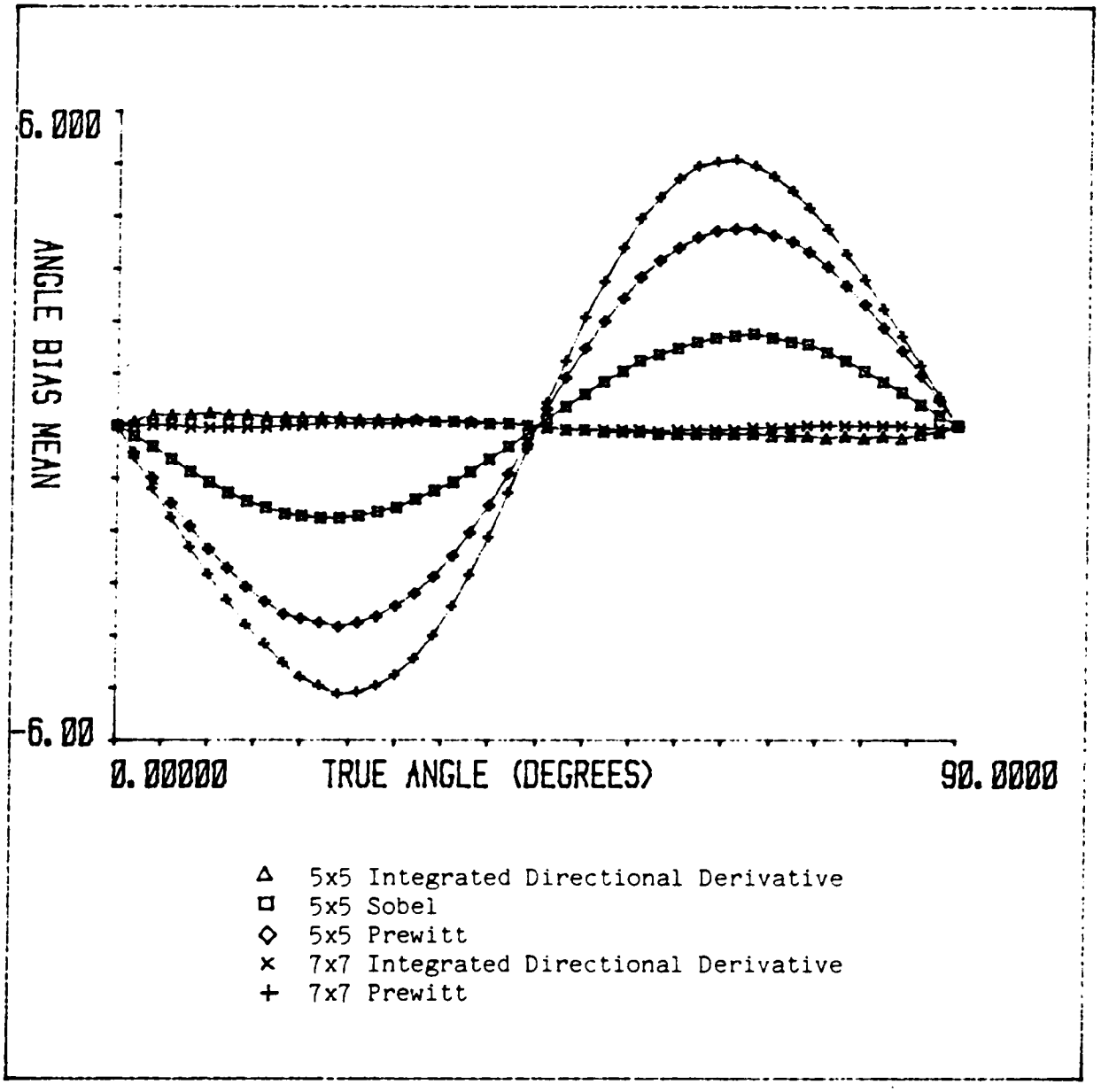


Figure 18. shows estimate bias as a function of true edge direction for ramp edges under zero noise conditions.

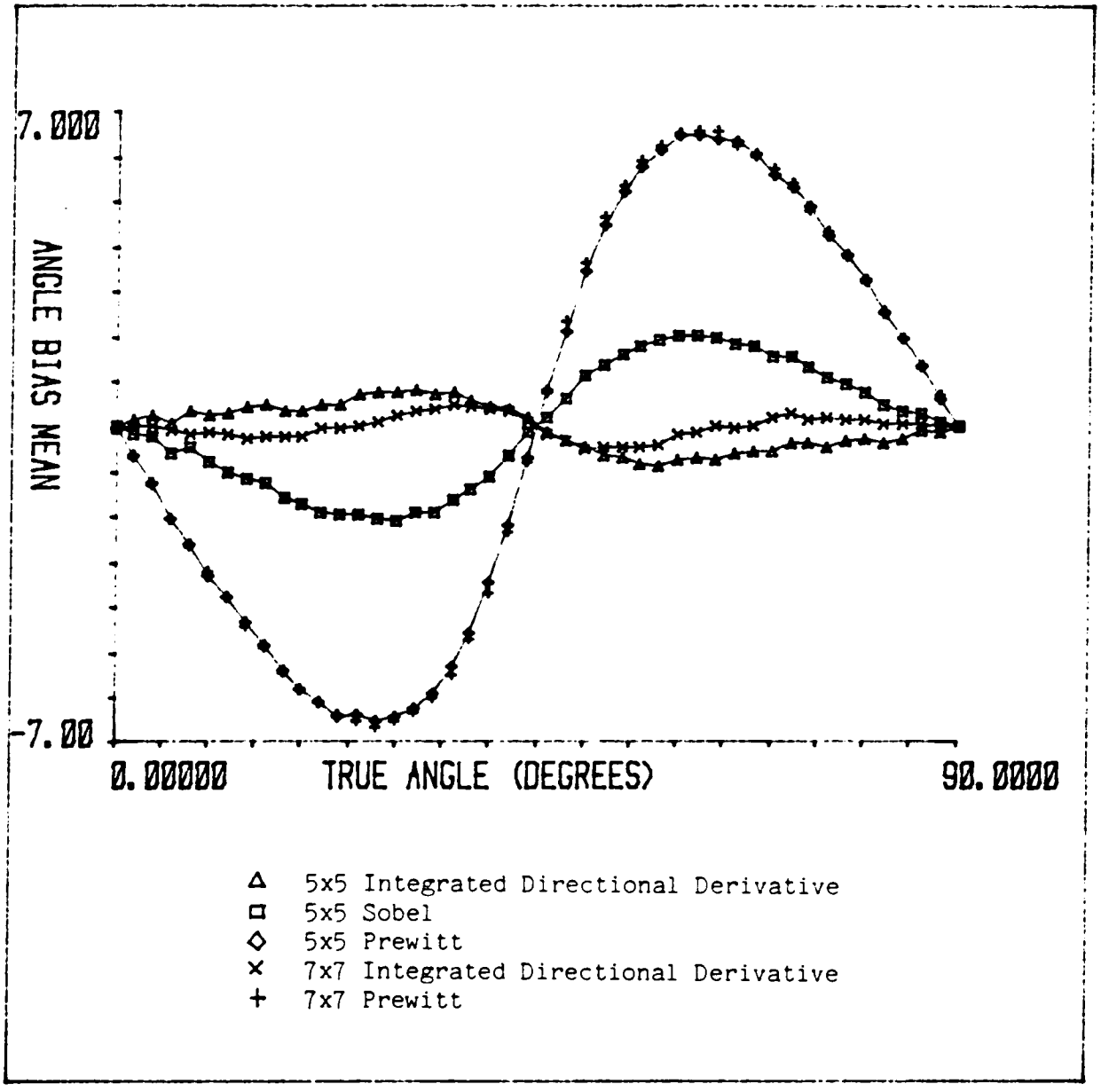


Figure 19. shows estimate bias as a function of true edge direction for a step edge: Noise standard deviation is 25. Edge contrast is 100.

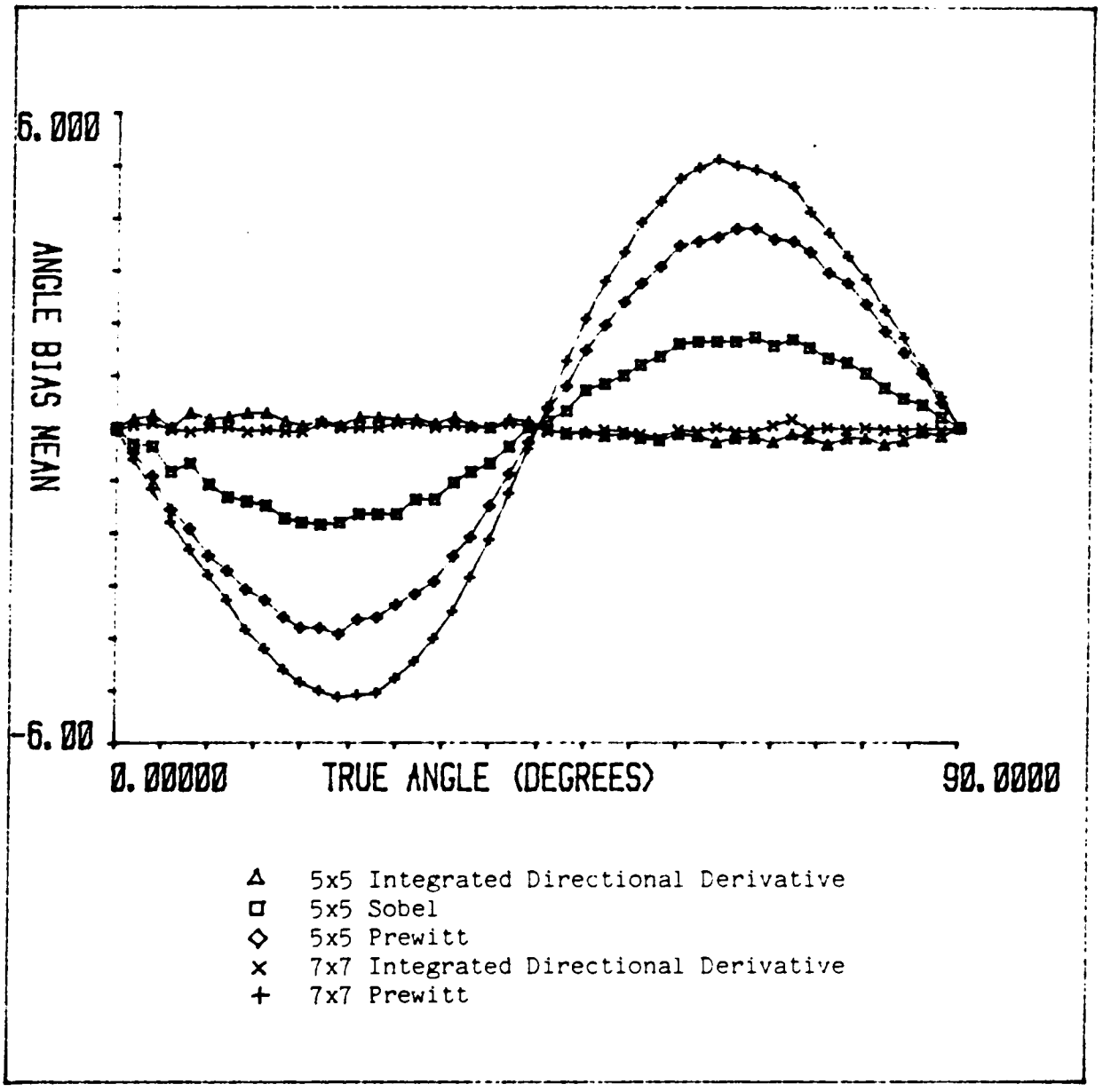


Figure 20. shows estimate bias as a function of true edge direction for a ramp edge: Noise standard deviation is 25. Edge contrast is 100.

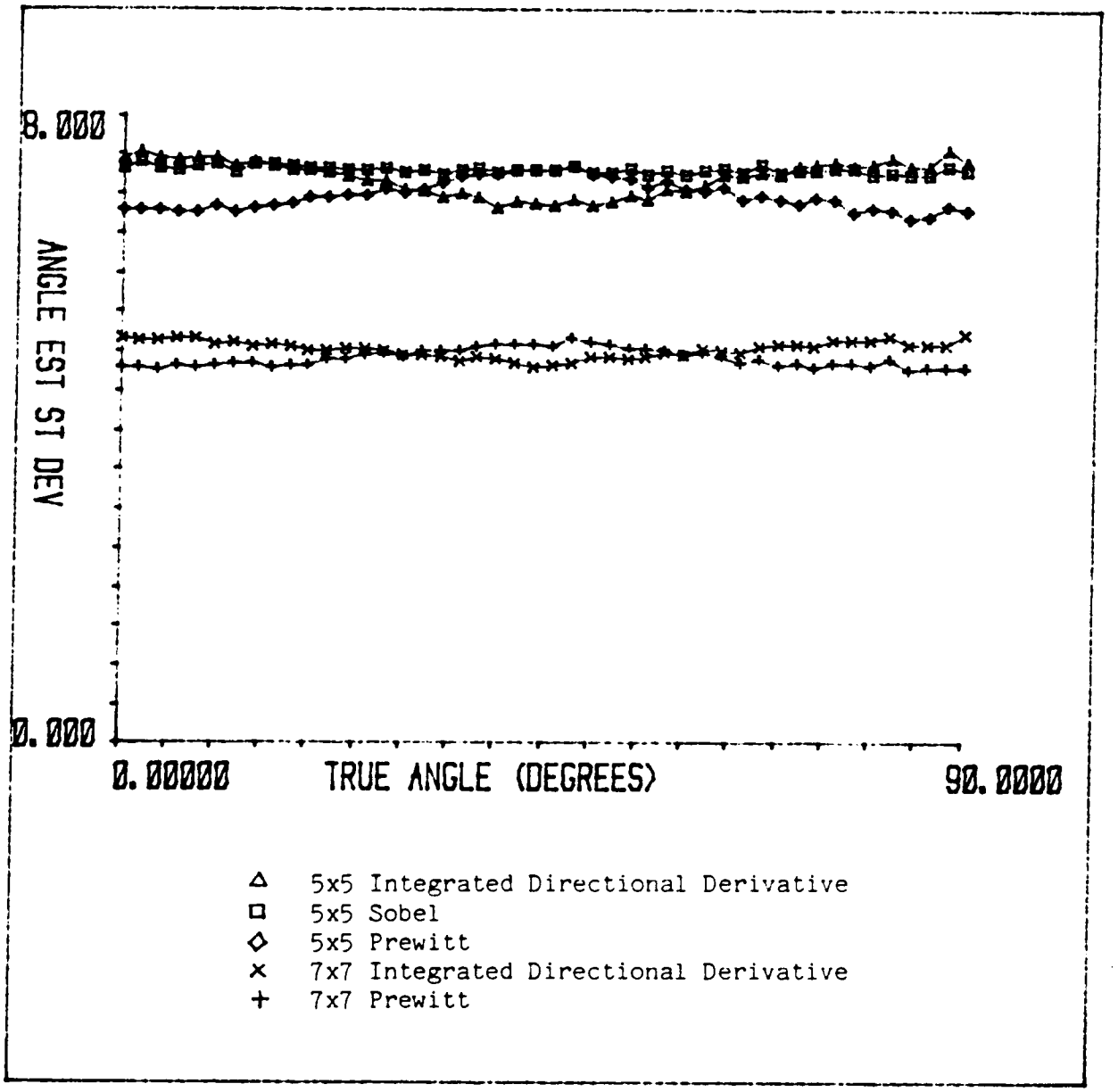


Figure 21. shows estimate standard deviation as a function of true edge direction for a step edge: Noise standard deviation is 25. Edge contrast is 100.

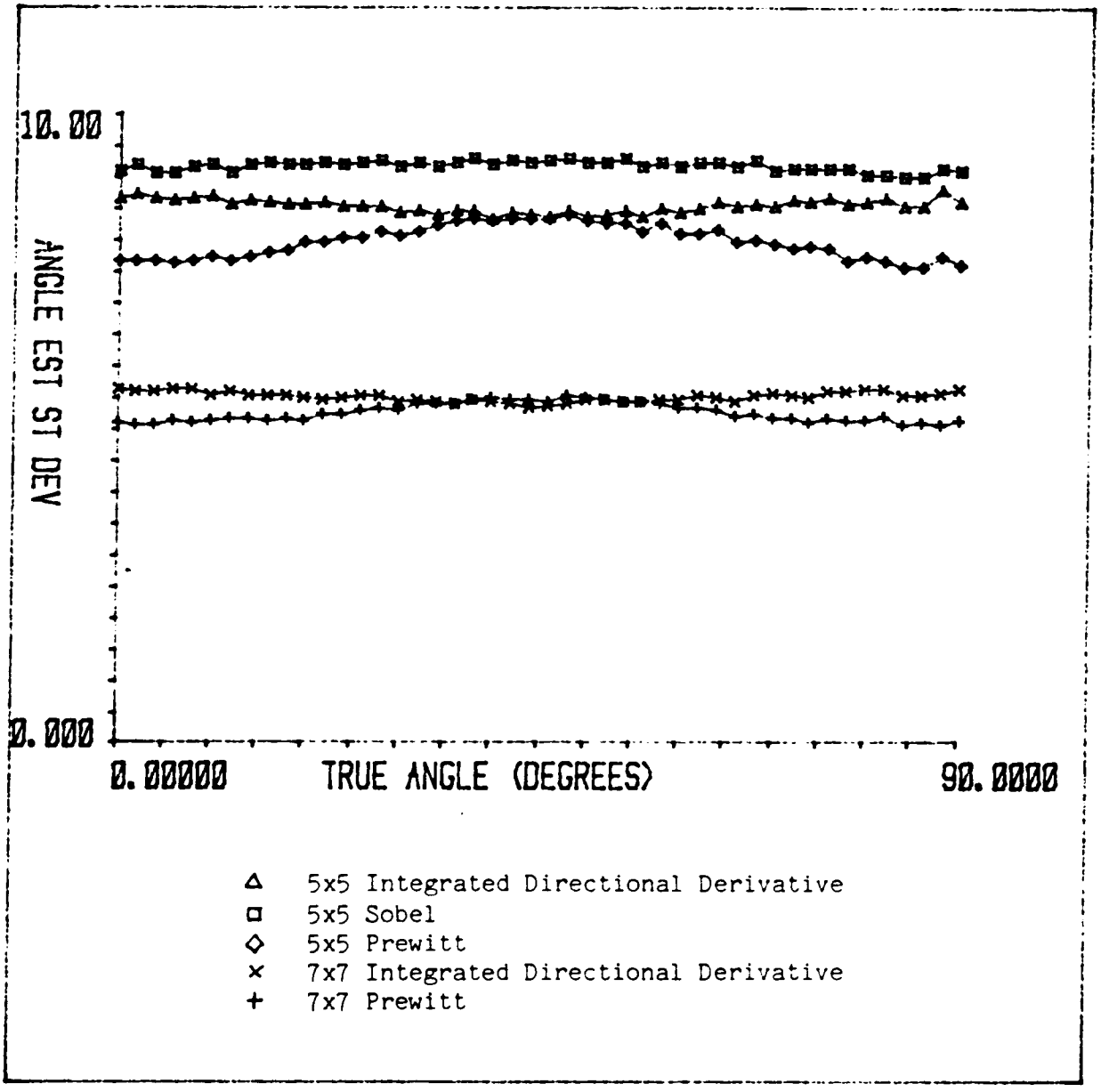


Figure 22. shows estimate standard deviation as a function of true edge direction for a ramp edge: Noise standard deviation is 25. Edge contrast is 100.

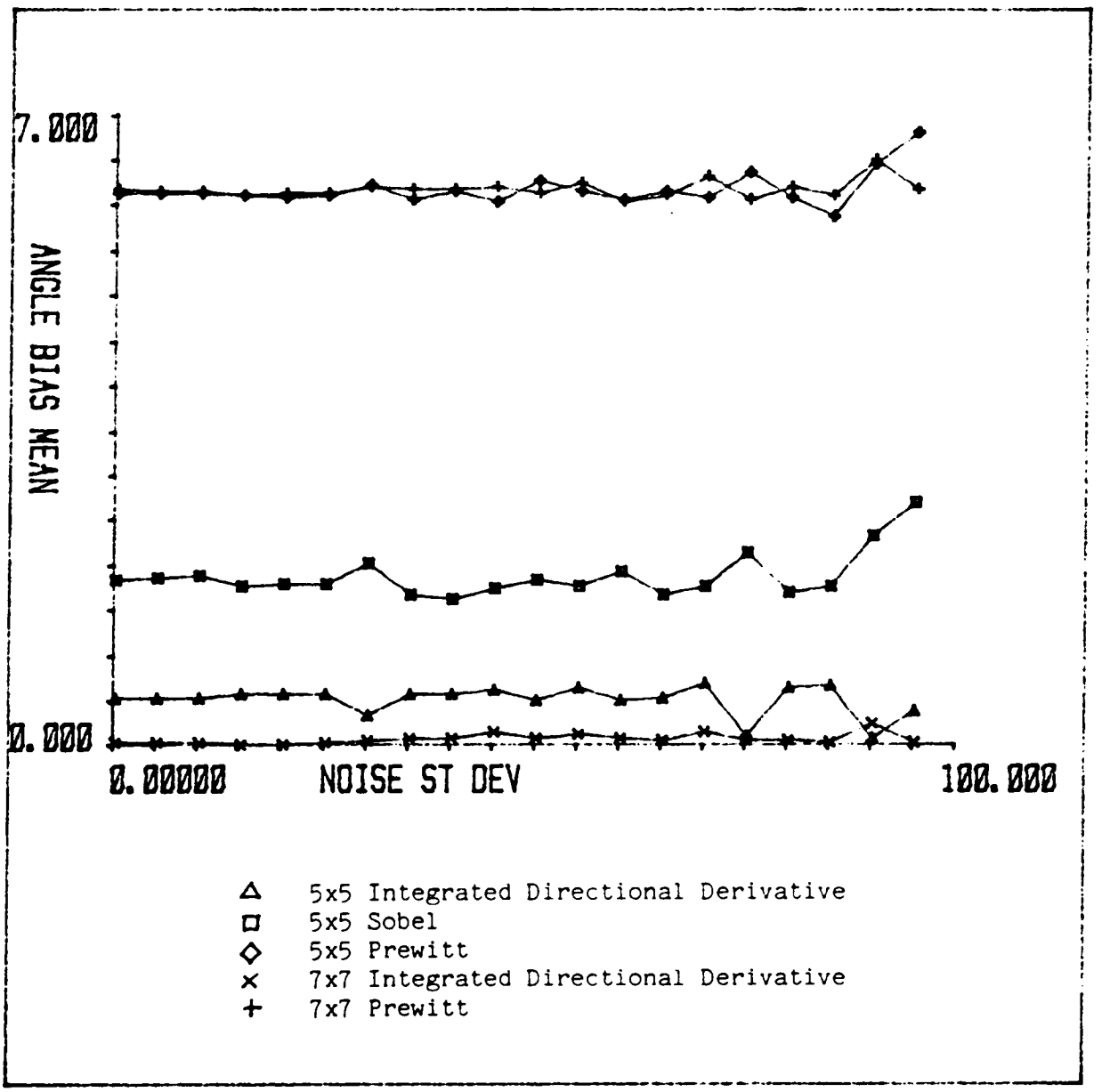


Figure 23. shows estimate bias as a function of noise standard deviation for a step edge: Edge orientation is 22.5 degrees. Edge contrast is 100.

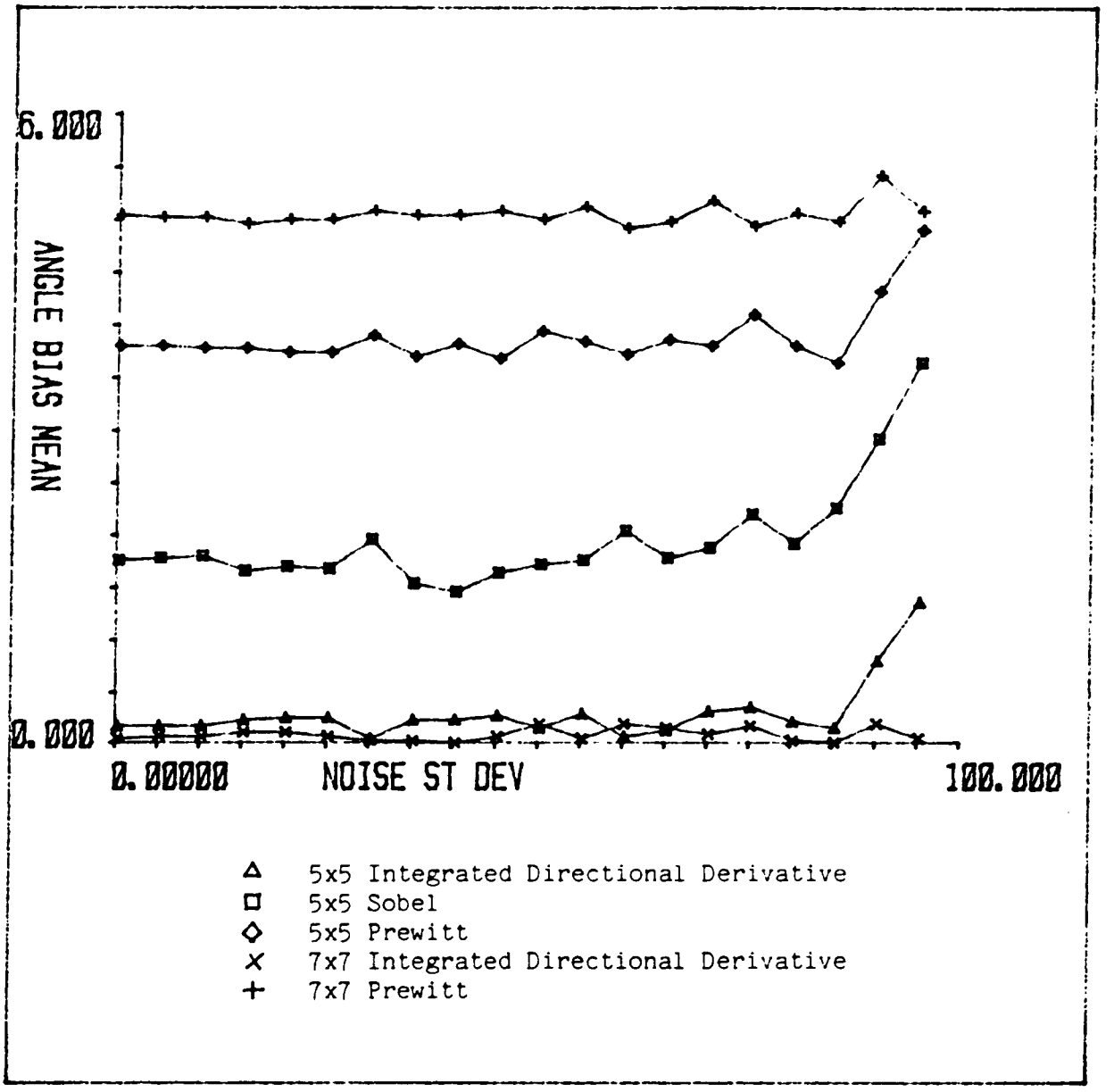


Figure 24. shows estimate bias as a function of noise standard deviation for a ramp edge: Edge orientation is 22.5 degrees. Edge contrast is 100.

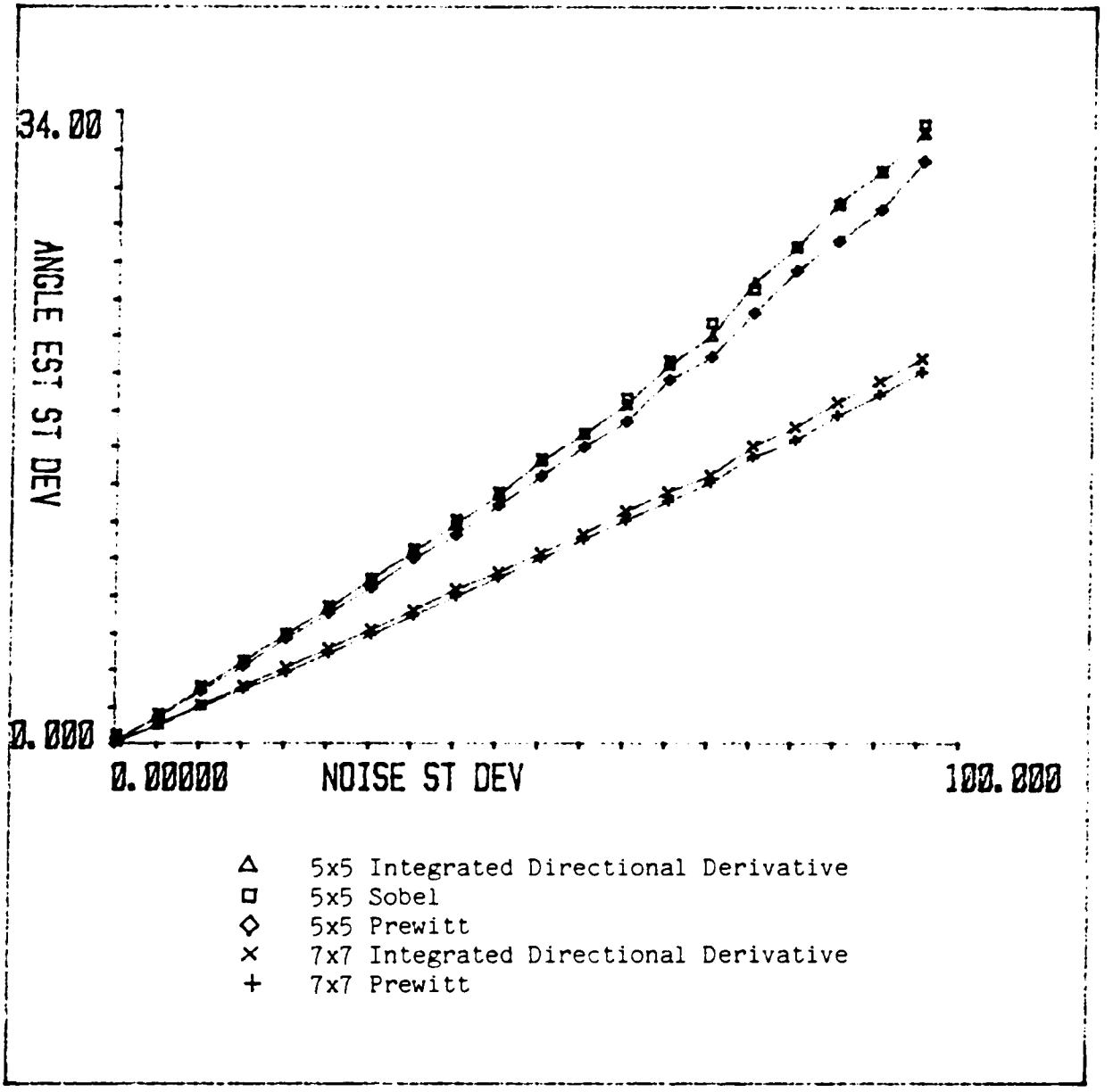


Figure 25. **shows estimate standard deviation as a function of noise standard deviation for a step edge:** Edge orientation is 22.5 degrees. Edge contrast is 100.

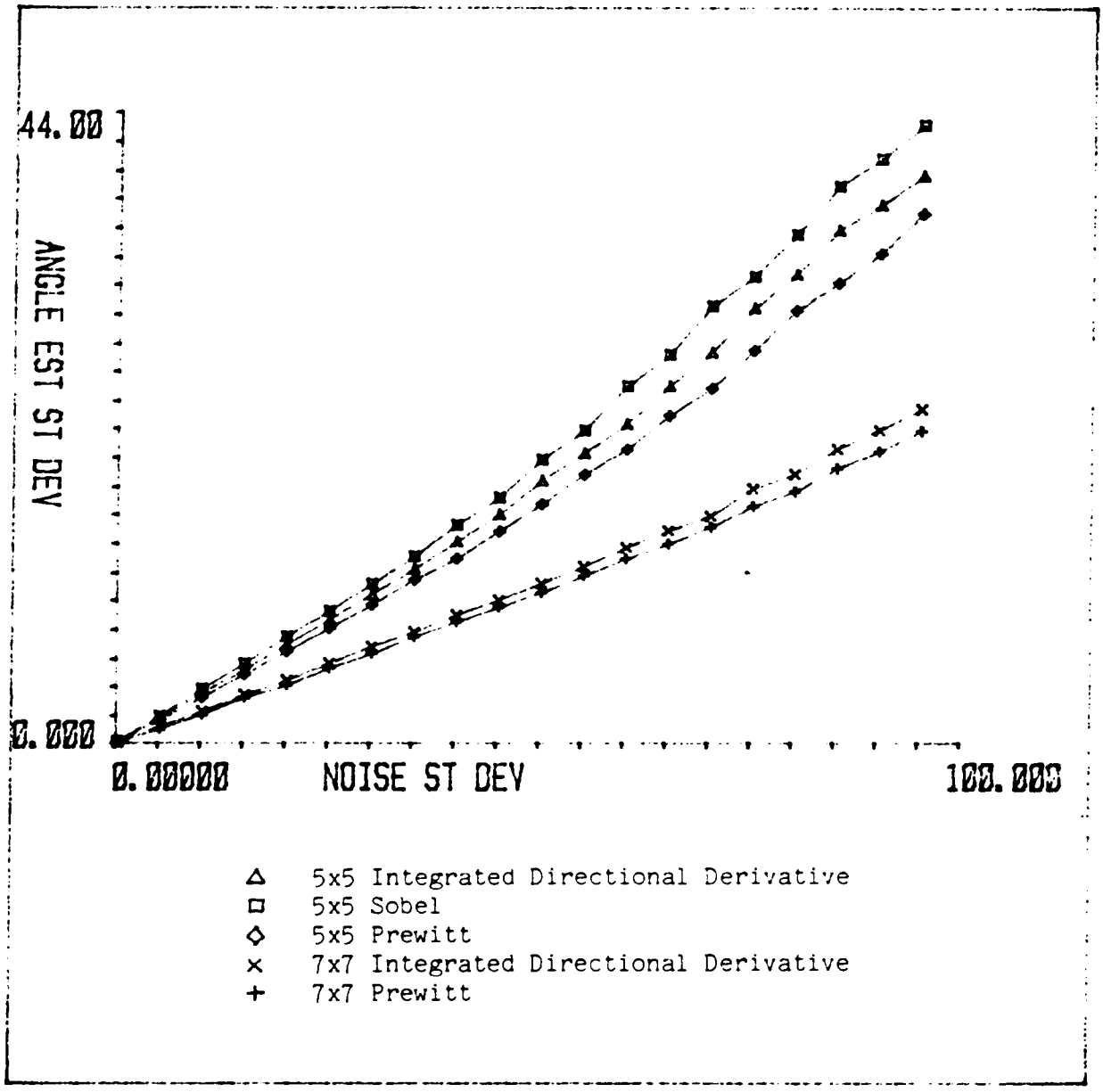


Figure 26. shows estimate standard deviation as a function of noise standard deviation for a ramp edge: Edge orientation is 22.5 degrees. Edge contrast is 100.

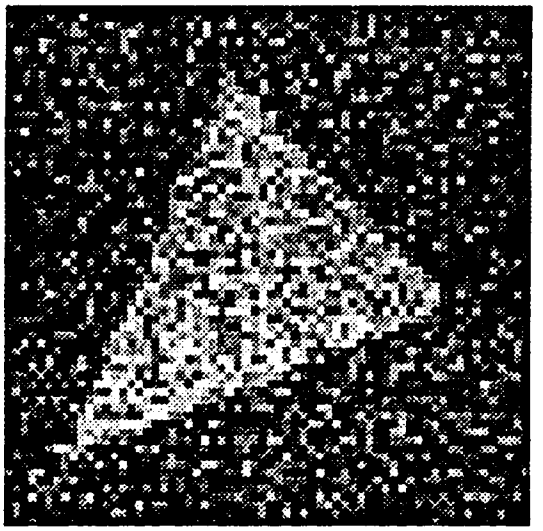


Figure 27. **shows the images used to compare the gradient operators:** Top: Synthetic image, bottom: aerial scene

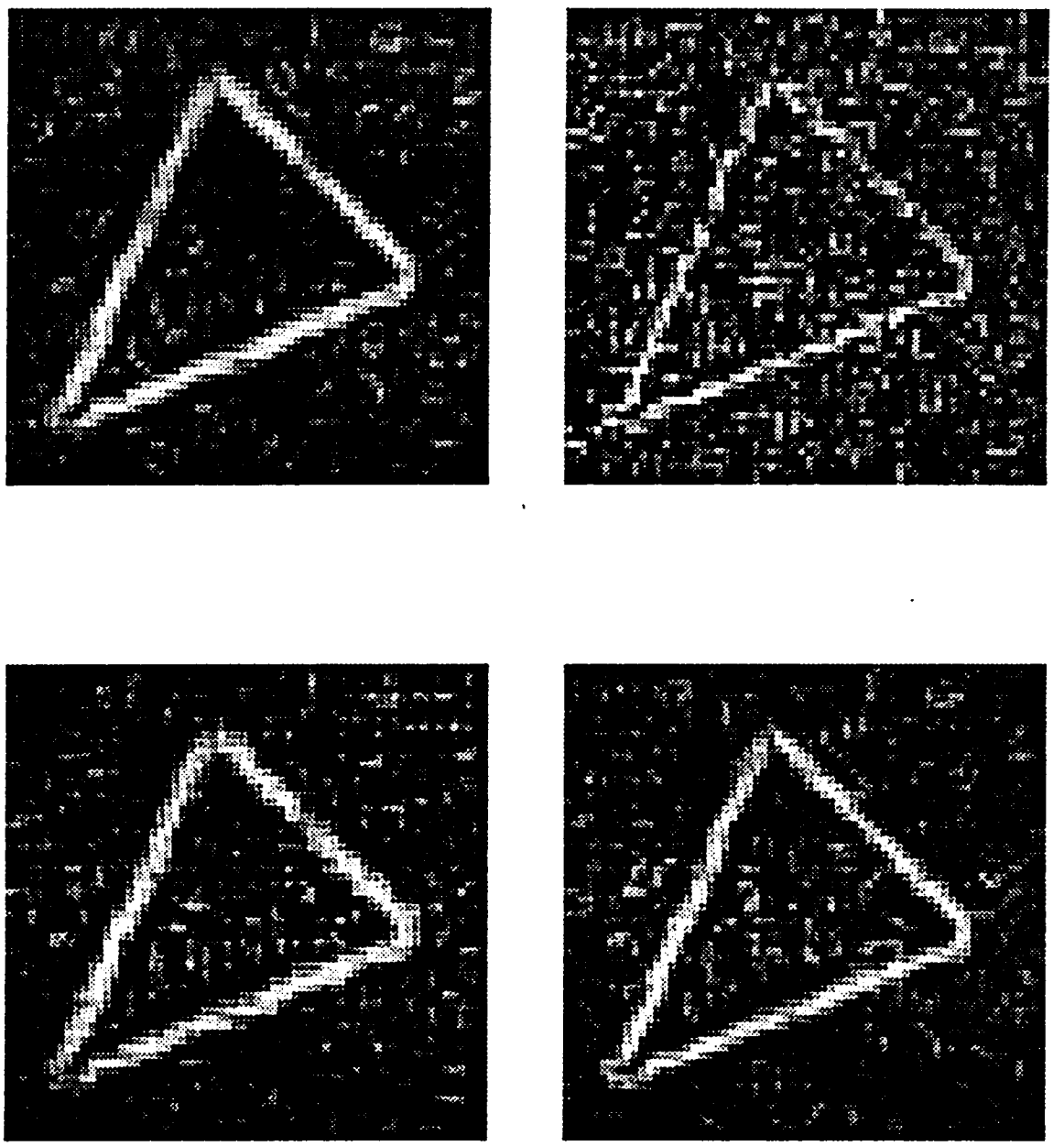


Figure 28. shows gradient strength response for the 5x5 operators on the synthetic image. Clockwise from top-left: Integrated Directional Derivative, Standard Cubic Facet, Extended Sobel, and Prewitt.

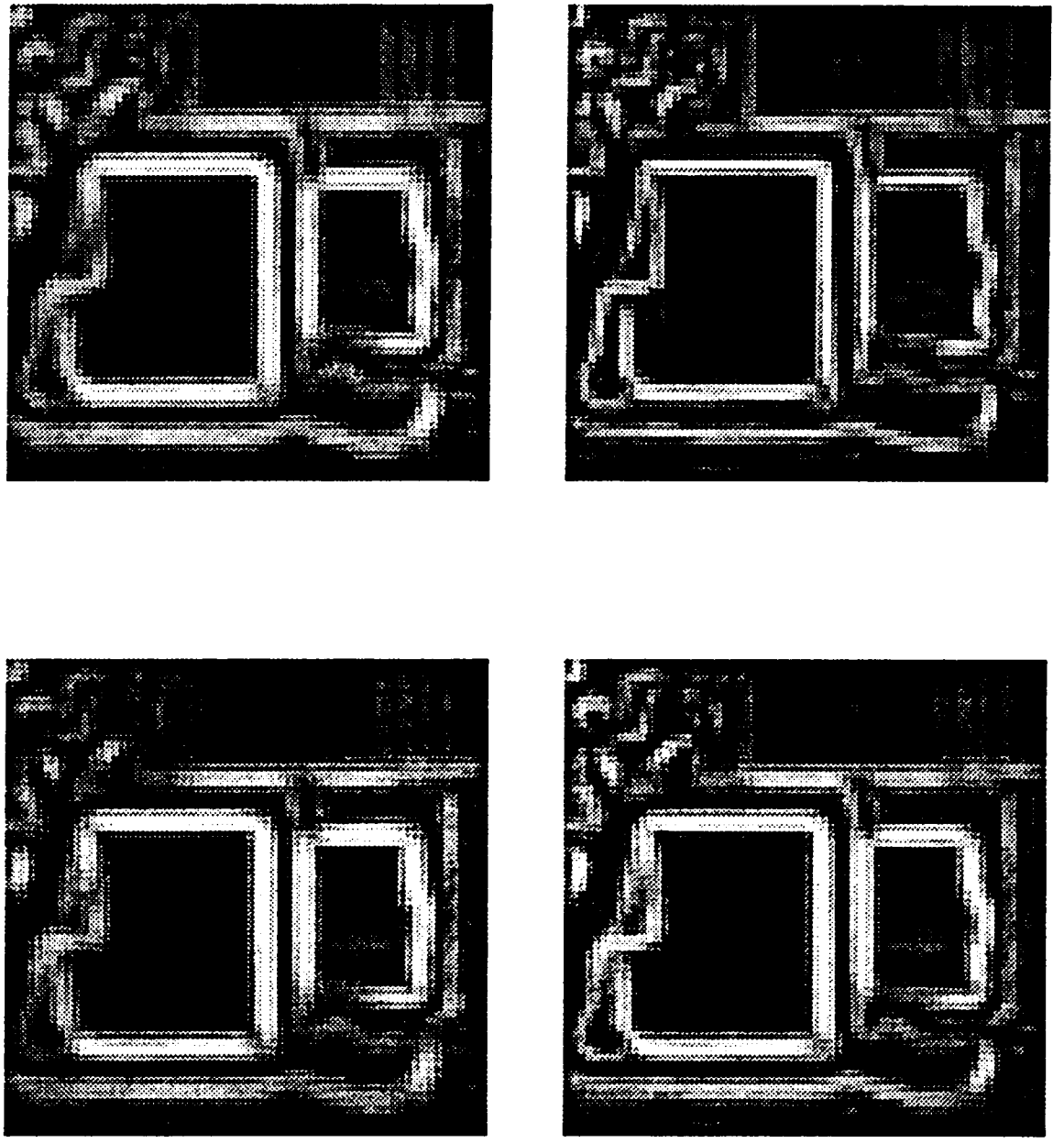


Figure 29. shows gradient strength response for the 5x5 operators on the aerial scene. Clockwise from top-left: Integrated Directional Derivative, Standard Cubic Facet, Extended Sobel, and Prewitt.

Chapter 4

GRADIENT THRESHOLD SELECTION

4.1 Introduction

The computation of the gradient of a digital image is usually one step of image processing tasks such as edge or corner detection. Typically edge candidates are taken from pixels whose gradient strength is significally different from zero. One of the most effective ways of detecting edges and corners in digital image data is by finding zero-crossings of the second directional derivative of the underlying graytone intensity surface (Haralick, 1982), (Zuniga and Haralick, 1983). These zero-crossings have associated with them non-zero gradient strengths. The presence of noise in real image data causes the occurrence of spurious zero-crossings. We are faced then with the problem of distinguishing between zero-crossings which arise due to the presence of a true edge or corner and zero-crossings which arise due to the presence of noise. In practice this

problem is handled by considering only zero-crossings whose gradient strength is significantly different from zero.

Our approach to this problem is based on the facet model for digital images (Haralick (1980), Haralick and Watson (1981)). The basic philosophy of this model derives from recognizing that the discrete set of values which form the digital image are the result of sampling and quantizing a real-valued function f defined on the domain of the image which is a bounded and connected subset of the real plane. Thus any property associated with a pixel or a neighborhood of pixel values should be evaluated by relating it to the property of the corresponding gray tone surface f which underlies the neighborhood. This involves estimating the surface function f locally, from the neighborhood samples available to us. The most natural way of accomplishing this is by assumming a parametric form for f and then estimating its associated parameters.

We are concerned with that property of a pixel called gradient strength which is defined as the Euclidean norm of the first partial derivatives of the graytone intensity surface evaluated at the pixel position. An assumption about the nature of the noise enables us to put the problem of choosing a suitable gradient threshold in a statistical framework. We assume the noise to be Gaussian with zero mean and variance σ^2. We derive two statistics which are functions of the gradient strength and the facet residual error of fit. We show that:

1. Thresholds on these statistics produce results which are superior to those obtained by the best subjective threshold on the gradient image.

2. Threshold selection can be made automatic by applying a Bayes decision method.

The analysis presented in the following sections is valid for a facet model of arbitrary order although the experiments were performed with a cubic facet model. Section

4.2 shows how the statistics are derived and poses the problem of selecting a suitable threshold as a hypothesis test problem. Section 4.3 shows how the thresholding can be made automatic by posing the problem as a Bayesian decision problem. Section 4.4 presents the experimental results. Appendix B provides the mathematical analysis required to derive the statistical distributions of the Euclidean norm of any subset of facet parameters or partial derivatives, and the distribution of the total facet residual error.

4.2 *Gradient threshold selection as a hypothesis test.*

Let μ_r and μ_c denote the true but unknown values of the first row and column partial derivatives of the underlying graytone intensity surface at a pixel position. Let $\hat{\mu}_r$ and $\hat{\mu}_c$ denote their estimates based upon a neighborhood of K values. According to the results of appendix B $\hat{\mu}_r$ and $\hat{\mu}_c$ are normally distributed and:

$$\begin{aligned} E[\hat{\mu}_r] &= \mu_r \\ E[\hat{\mu}_c] &= \mu_c \\ V[\hat{\mu}_r] &= \sigma^2 k \\ V[\hat{\mu}_c] &= \sigma^2 k \\ E[\hat{\mu}_r \hat{\mu}_c] &= \mu_r \mu_c \end{aligned} \tag{4.1}$$

where k is a constant whose value depends on the neighborhood size and the basis functions used to estimate the graytone intensity surface.

Consider testing the hypothesis that $\mu_r = \mu_c = 0$ (zero gradient). This hypothesis must be rejected if there is to be a zero-crossing of second directional derivative. Under this hypothesis:

$$\frac{\hat{\mu}_r^2 + \hat{\mu}_c^2}{k\sigma^2}$$

. has a χ_2^2 distribution.

Also, from the results of appendix B, the total residual error S^2 normalized by the noise variance S^2/σ^2 has a χ_{K-N}^2 distribution, where N is the number of basis functions or the number of facet parameters. Hence,

$$x_1 = \frac{(\hat{\mu}_r^2 + \hat{\mu}_c^2)/2}{kS^2/(K-N)} \tag{4.2}$$

has a $F_{2,K-N}$ distribution and the hypothesis for zero gradient would be rejected for suitable large values of x_1. The value of the threshold for x_1 is chosen to correspond to a given significance level of test.

Notice that the statistic defined by equation (4.2) may be regarded as a significance or reliability measure associated with the existence of a non-zero gradient. It is essentially proportional to the square gradient normalized by $S^2/(K-N)$ which is a random variable whose expected value is σ^2, the variance of the noise. This scaling of the gradient by a local estimate of the image noise makes optimum selection possible by a fix threshold procedure.

If the noise variance is known to be constant everywhere throughout the image domain a better estimate for it is possible by averaging the total residual errors S^2 over M non-overlapping neighborhoods. Let this average be denoted by E^2, then ME^2/σ^2 has a $\chi_{M(K-N)}^2$ distribution and

$$x_2 = \frac{(\hat{\mu}_r^2 + \hat{\mu}_c^2)/2}{kE^2/(K-N)} \tag{4.3}$$

has a $F_{2,M(K-M)}$ distribution and the hypothesis for zero gradient would be rejected for suitable large values of x_2 . Again, the threshold on x_2 is chosen to correspond to a given significance level of test.

4.3 Gradient threshold selection as a Bayesian decision problem.

This method derives from considering any image point to be in one of two states: zero gradient strength or non-zero gradient strength. We denote these two states by z and nz respectively. Let x be a continuous random variable test statistic whose distribution depends on the pixels' state. We define x as:

$$x = \frac{(\hat{\mu}_r^2 + \hat{\mu}_c^2)/2}{kS^2/(K-N)} \tag{4.4}$$

As previously seen, for pixels with zero gradient strength x has a $F_{2,K-N}$ distribution. A simple Bayes decision rule with a unity gain function is:

$$Decide\begin{cases}\text{zero gradient if} & P(z\,|\,x) > P(nz\,|\,x)\\ \text{non-zero gradient,} & otherwise\end{cases} \tag{4.5}$$

and the probability of error associated with this decision is:

$$P(error\,|\,x) = \begin{cases}P(nz\,|\,x) & if \;\; P(z\,|\,x) > P(nz\,|\,x)\\ P(z\,|\,x), & otherwise\end{cases}$$

or equivalently

$$P(error \mid x) = \min\{P(z \mid x), P(nz \mid x)\} \tag{4.6}$$

We can express the decision rule and probability of error in terms of the distributions of x by using Bayes rule.

$$P(z \mid x) = \frac{P(x \mid z)P(z)}{P(x)}$$
$$P(nz \mid x) = \frac{P(x \mid nz)P(nz)}{P(x)}$$

The decision rule then becomes

$$Decide \begin{cases} \text{zero gradient if} & P(x \mid z)P(z) > P(x \mid nz)P(nz) \\ \text{non-zero gradient,} & otherwise \end{cases} \tag{4.7}$$

and the probability of error becomes

$$P(error \mid x) = \frac{\min\{P(x \mid z)P(z), P(x \mid nz)P(nz)\}}{P(x)} \tag{4.8}$$

The density function $P(x \mid z)$ is known to be $F_{2,K-N}$, P(x) is the mixture distribution and can be estimated from a histogram of the gradient image. Then $P(x \mid nz)$ can be obtained as follows:

$$\begin{aligned} P(x) &= P(x,z) + P(x,nz) \\ &= P(x \mid z)P(z) + P(x \mid nz)P(nz) \end{aligned} \tag{4.9}$$

and from here

$$P(x \mid nz) = \frac{[P(x) - P(x \mid z)P(z)]}{P(nz)}$$

$$P(x \mid nz) = \frac{[P(x) - P(x \mid z)P(z)]}{[1 - P(z)]} \tag{4.10}$$

The total probability of error is given by:

$$P(error) = \sum_x P(error \mid x)P(x)$$

Using equations (4.8) and (4.10)

$$P(error) = \sum_x \min\{P(x \mid z)P(z),\ P(x \mid nz)P(nz)\}$$

$$P(error) = \sum_x \min\{P(x \mid z)P(z),\ P(x) - P(x \mid z)P(z)\} \tag{4.11}$$

Since the prior probability P(z) of zero gradient is not known it must be user specified. For many images values of .9 to .95 are reasonable. Another method of choosing P(z) is obtained by observing from equation (4.9) that

$$\text{for all } x \qquad P(x) \geq P(x \mid z)P(z)$$

Then

$$\text{for all } x,\ P(x \mid z) > 0 \qquad P(z) \leq \frac{P(x)}{P(x \mid z)}$$

A suitable value for P(z) is therefore

$$P(z) = \min_{x,\, P(x|z)>0} \frac{P(x)}{P(x|z)} \tag{4.12}$$

Once P(z) has been specified, a suitable threshold x_{th} can be chosen from rule (4.7) by solving

$$P(x_{th}|z)P(z) = P(x_{th}|nz)P(nz) \tag{4.13}$$

We assume that only one intersection point x_{th} exists. The same analysis can be carried out with the statistic defined by equation (4.3) if the noise variance is known to be constant throughout the image domain.

4.4 Experimental Results

4.4.1 Test images.

Two artificially generated images TEST1 and TEST2 were tested. They are shown in Figure 30 on page 88. Their size is 100 X 100 pixels and they consist of four bright circles on a dark background. The diameter of the circles is 25. The edge contrasts for each of the four circles are 9, 18, 27, and 50. The edges are ramp edges whose width is five pixels. The noise is additive Gaussian with zero mean. The standard deviation of the noise for TEST1 is 5. The standard deviation of the noise for TEST2 changes from circle to circle and their values are 5, 10, 15, and 28 which keeps the signal to noise ratio constant at about 20 log (9/5) or approximately 5 DB.

The first partial derivatives of the graytone intensity function at every pixel position were computed by fitting a cubic polynomial surface defined in the row and column coordinates of a 9 x 9 neighborhood centered about the pixel. Figure 31 on page 89 shows the gradient images computed from the Euclidean norm of the first partial derivatives and Figure 32 on page 90 shows the result of applying a constant threshold to the gradient images. The thresholds for each image were found interactively with subjective quality as the selection criteria. Notice the inability of this procedure to perform well everywhere in TEST2 due to the changing nature of the noise variance.

4.4.2 The $F_{2,M(K-N)}$ test statistic.

In this case the test statistic defined by equation (4.3) is used for both the hypothesis test method and the Bayesian decision method of automatic gradient threshold selection. This statistic was derived under the assumption that the variance of the noise remains constant throughout the image.

Figure 33 on page 91 shows images of the test statistic x and Figure 34 on page 92 shows the results of a 1 percent and 5 percent significant level hypothesis test. Figure 35 on page 93 shows the histograms of the mixture distribution P(x), and the conditional probability distributions given zero gradient P(x|z) and non-zero gradient P(x|nz). The prior probability of non-zero gradient P(nz) was computed automatically using equation (4.12). These prior probabilities were found to be 0.22 for image TEST1 and 0.45 for TEST2. We can compare these values with the true prior probability of non-zero gradient computed as 4(3.14)D(W)/10000 where D is the diameter of the circles and W the width of the non-zero gradient region. This yields a value of 0.157 for P(nz). Notice that since TEST2 does not meet our assumption of constant noise variance, the

error in its estimate of P(nz) is much bigger than that in TEST1. The values of the computed threshold for the test statistic x were found to be 3.94 for TEST1 and 3.4 for TEST2. Finally Figure 36 on page 94 shows the images of the threshold of the test statistics.

4.4.3 The $F_{2,K-N}$ test statistic

In this case the test statistic defined by equation (4.2) is used for both the hypothesis test method and the Bayesian decision method of automatic gradient threshold selection. This statistic is valid under both constant or changing noise variance conditions.

Figure 37 on page 95 shows images of the test statistic x and Figure 38 on page 96 shows the results of a 1 percent and 5 percent significant level hypothesis test. Figure 39 on page 97 shows the histograms of the mixture distribution P(x), and the conditional probability distributions given zero gradient P(x|z) and non-zero gradient P(x|nz). The prior probability of non-zero gradient P(nz) was computed automatically using equation (4.12). These prior probabilities were found to be 0.24 for image TEST1 and 0.23 for TEST2. Notice than in this case since TEST2 does meet the assumption on the nature of the noise variance its estimate of P(nz) is in closer agreement with the true value of P(nz). The values of the computed threshold for the test statistic x were found to be 3.93 for TEST1 and 3.21 for TEST2. Finally Figure 40 on page 98 shows the images of the threshold of the test statistics.

4.4.4 Comparison with other methods.

As a comparison with other methods of thresholding the gradient image we used the running mean method. Each pixel in the gradient image was thresholded in proportion to the mean of the gradient values on a local 20 x 20 neighborhood around the given pixel. The results are shown in Figure 41 on page 99 for a variable threshold of 1.5 times the running mean. The value 1.5 was found interactively for best subjective results. Notice that this method performs as well as the methods that use the $F_{2,K-N}$ statistic, although user interaction is required to select the best value for the constant of proportionality (1.5 for the test images). The Bayesian decision method requires no interaction.

4.5 Conclusions

Several conclusions can be drawn from the experimental results of the previous section:

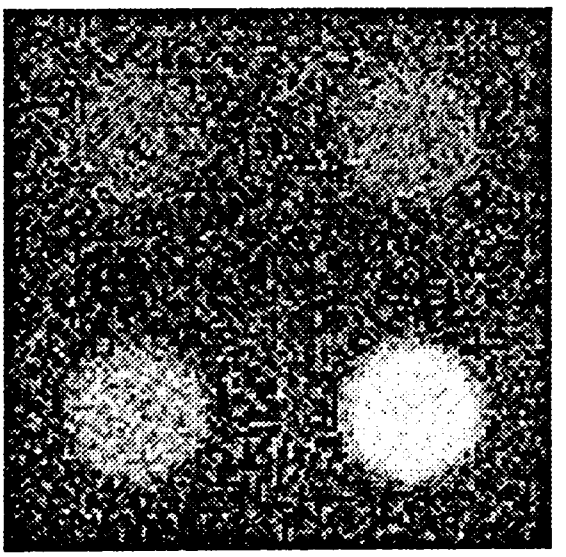
1. Thresholds on the $F_{2,K-N}$ and $F_{2,M(K-N)}$ test statistics produce similar results to that obtained by a best subjective user interactive threshold on the gradient image, when the noise variance remains constant throughout the image.

2. A threshold on the $F_{2,K-N}$ statistic produces a clearly superior result to either a threshold on the $F_{2,M(K-N)}$ statistic or to the best subjective user interactive threshold

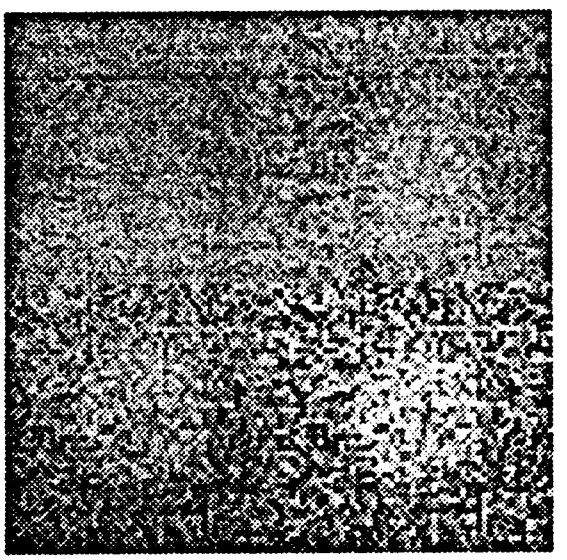
on the gradient image, when the noise variance changes from region to region in the image.

3. The Hypothesis test method of gradient threshold selection is inferior to the Bayesian decision method in the sense that the optimal value of significance level of the test cannot be known in advance and thus user interaction is required. No user interaction is required by the Bayesian decision method, except the choice of neighborhood size which is required in all the methods tested.

These conclusions are valid under the controlled nature of our experiments where the test images meet the image and noise model assumptions. Further research is necessary to determine the sensitivity of these methods to departures from the model assumptions.

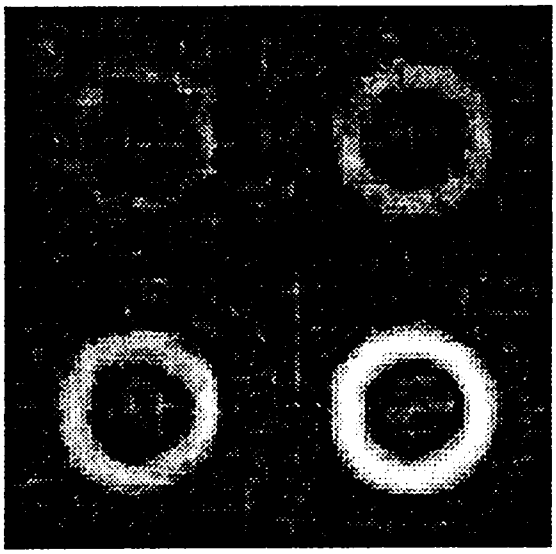


(a)

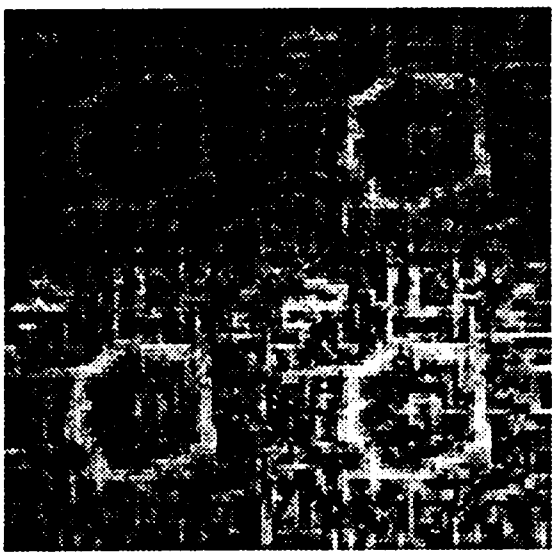


(b)

Figure 30. shows the original test images: (a) TEST1, (b) TEST2.

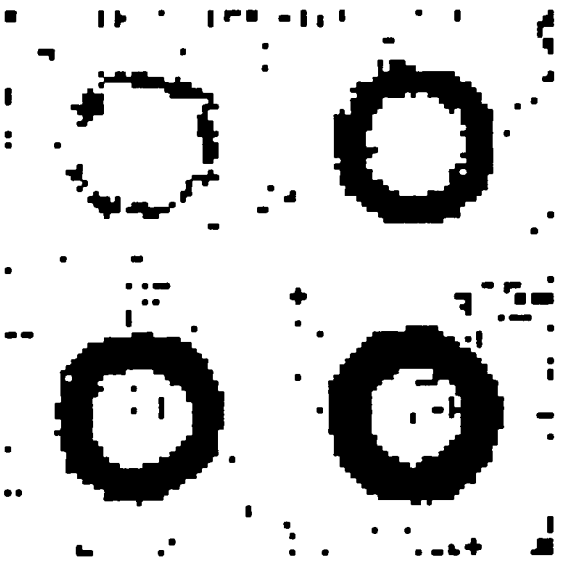


(a)



(b)

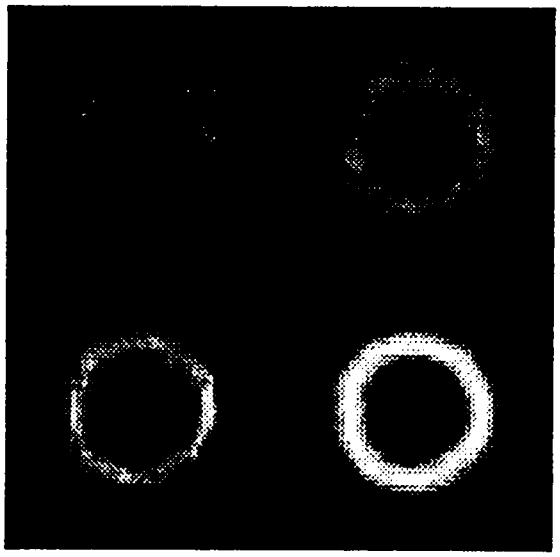
Figure 31. shows the gradient images computed from a 9x9 cubic facet applied to the test images: (a) TEST1, (b) TEST2.



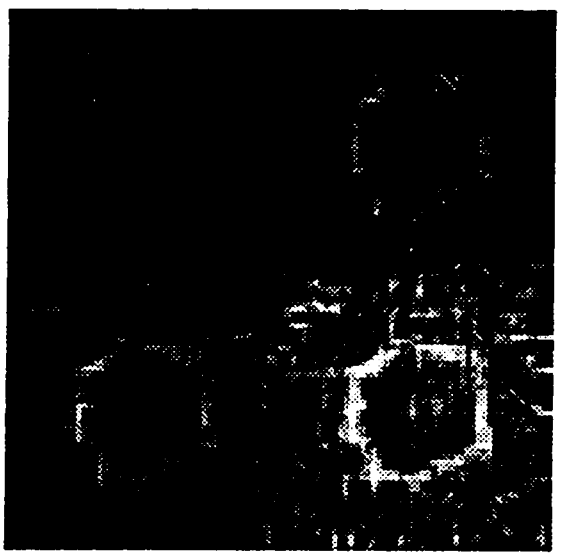
(a)

(b)

Figure 32. shows a user selected subjective threshold applied to the gradient images: (a) TEST1, (b) TEST2.



(a)



(b)

Figure 33. shows the values of the test statistic $F_{2,M(K-N)}$ for each of the test images: (a) TEST1, (b) TEST2.

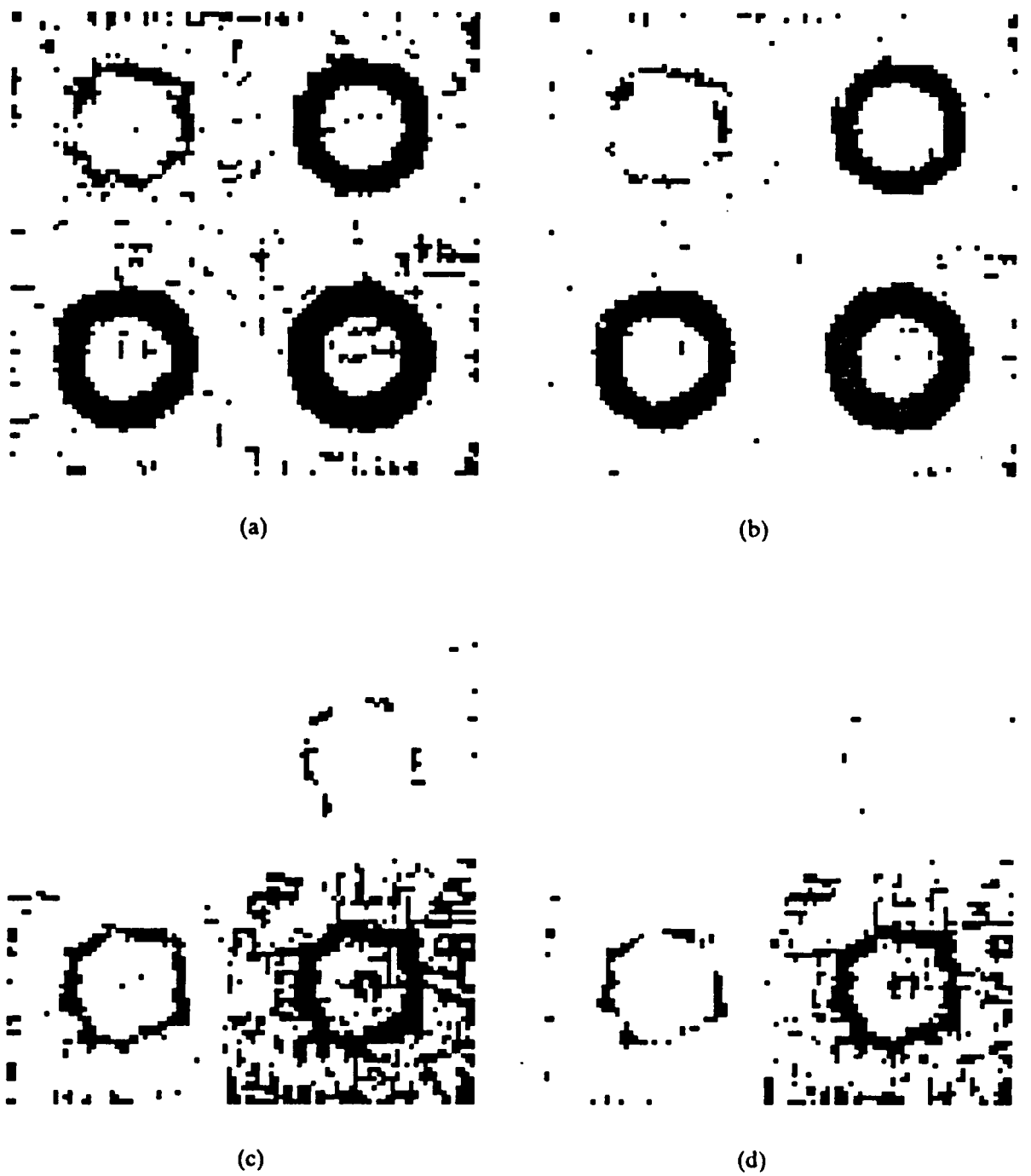
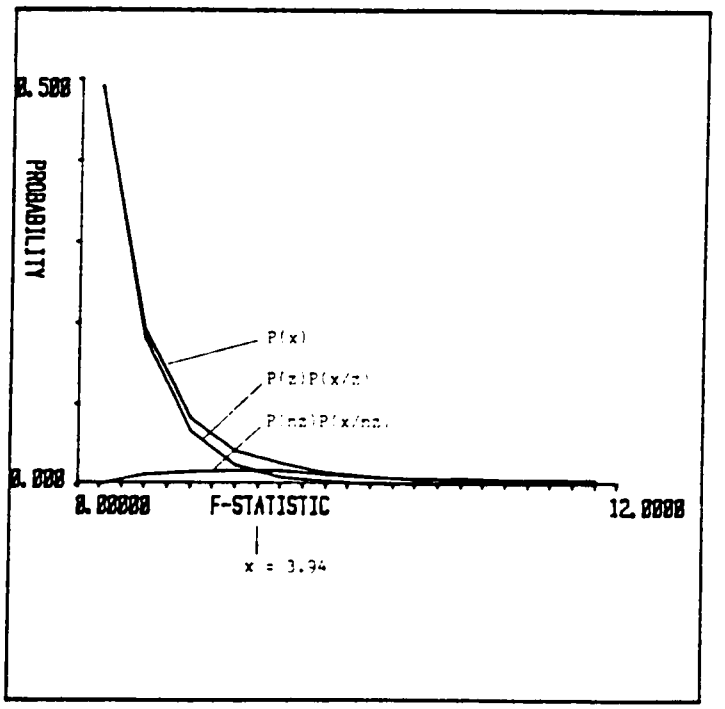


Figure 34. shows the results of a 5 and 1 percent significant level test on the $F_{2,M(K-N)}$ statistics: (a) TEST1, 5%, (b) TEST1, 1%, (c) TEST2, 5%, (d) TEST2, 1%.



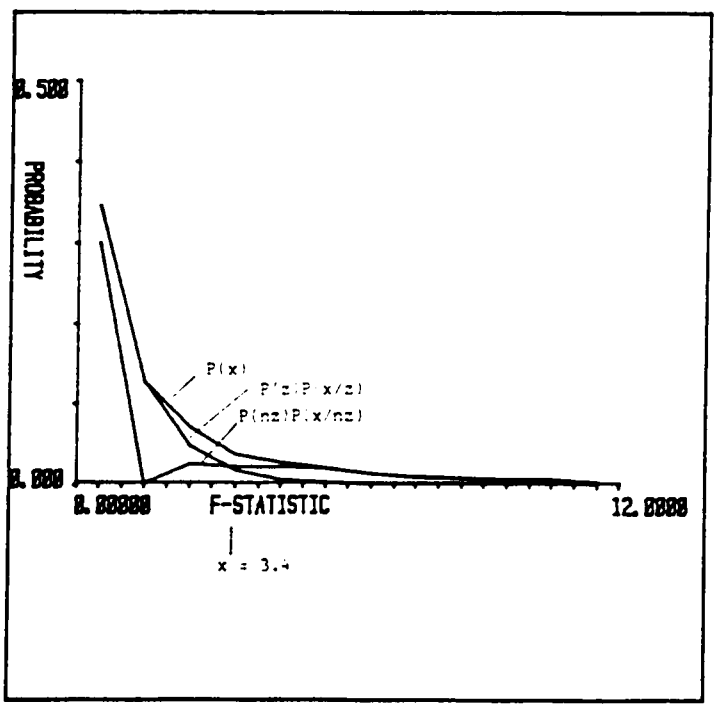
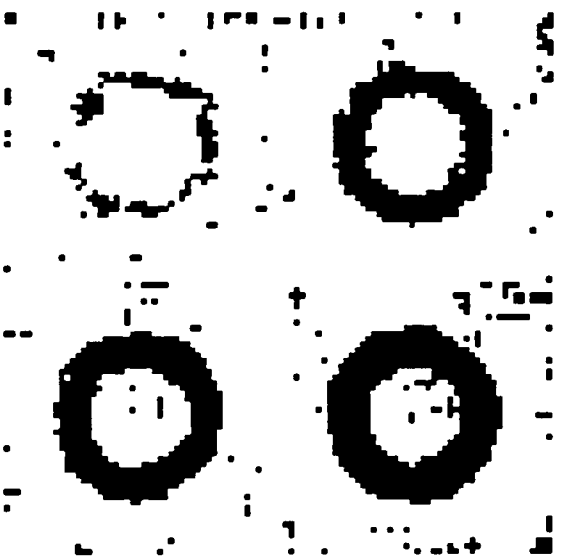


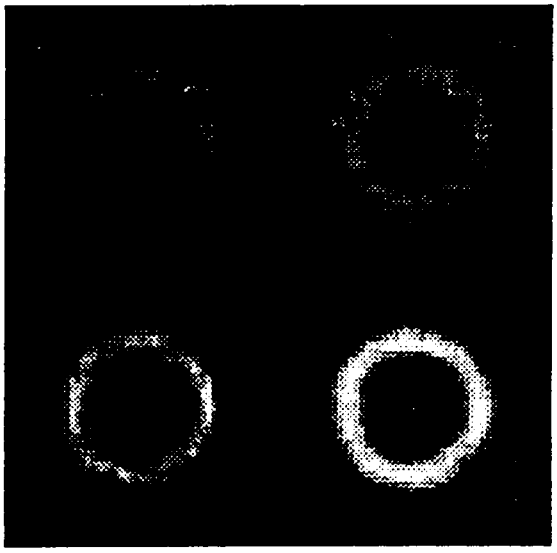
Figure 35. shows the histograms P(x), P(x|z), and P(x|nz) for the $F_{2,M(K-N)}$ statistic: (a) TEST1, (b) TEST2.



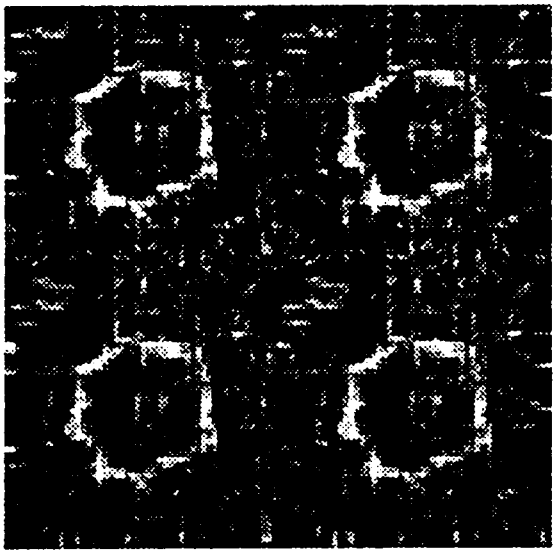
(a)

(b)

Figure 36. shows the thresholds on the $F_{2,M(K-N)}$ statistic from the Bayesian method: (a) TEST1, (b) TEST2.



(a)



(b)

Figure 37. shows the values of the test statistics $F_{2,K\text{-}N}$ for each of the test images: (a) TEST1, (b) TEST2.

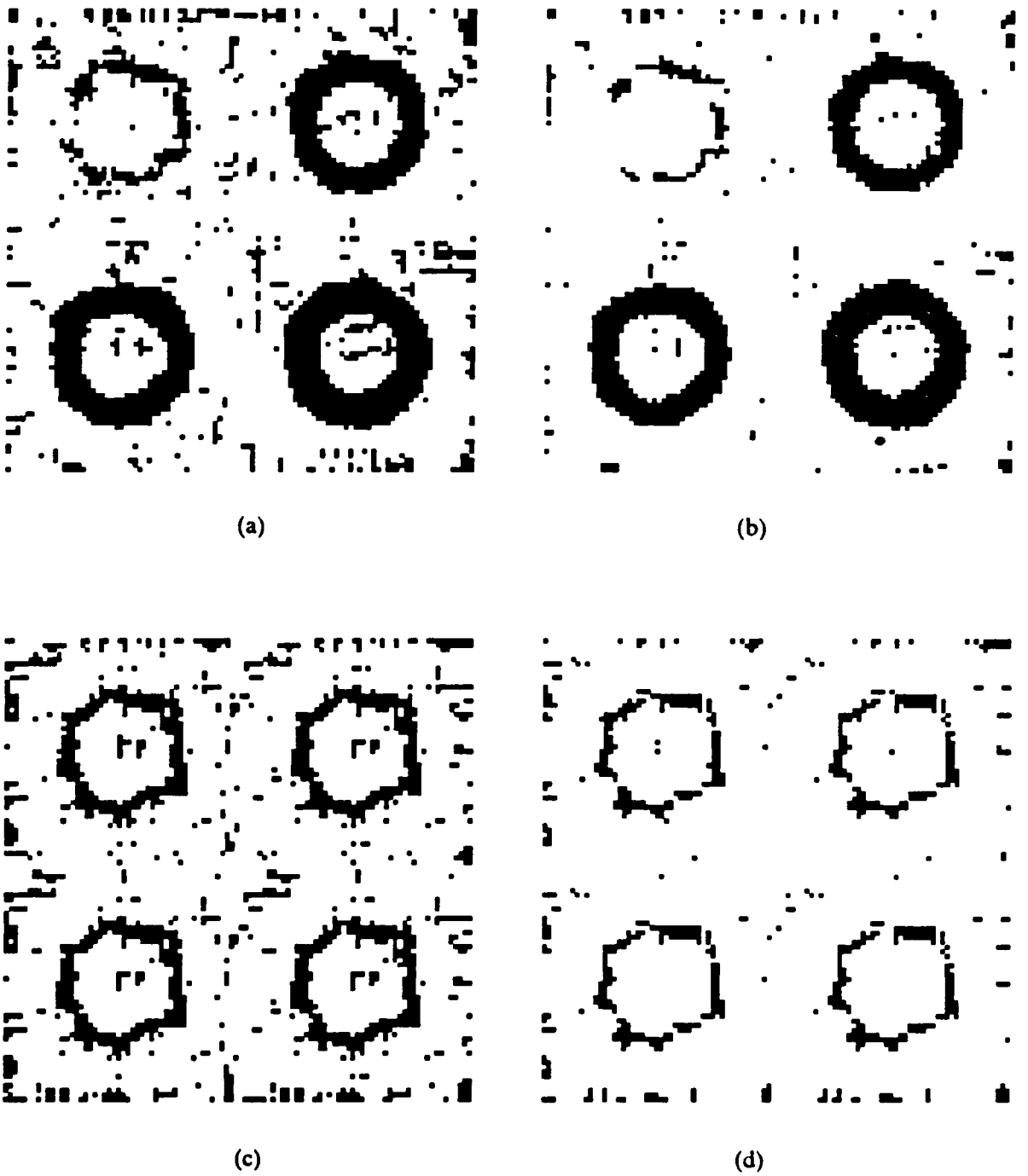
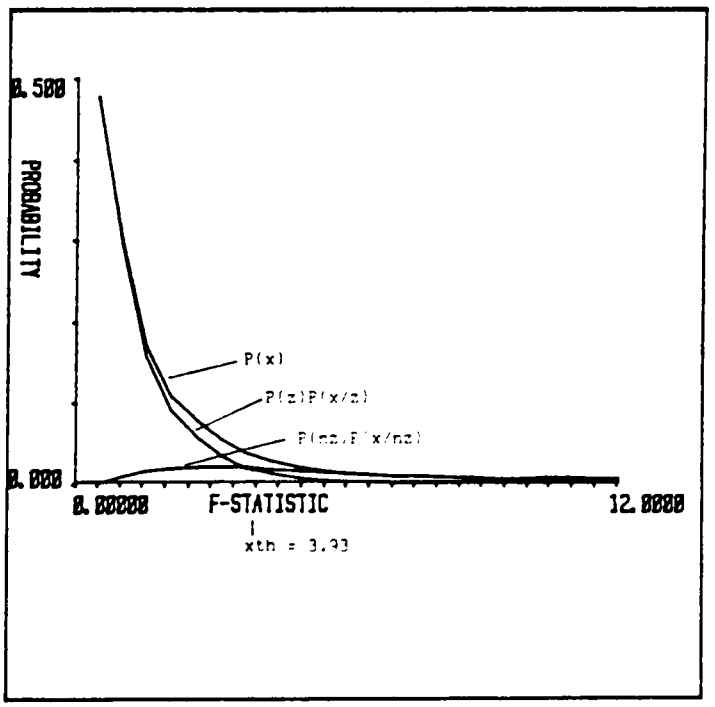


Figure 38. shows the results of a 5 and 1 percent significant level test on the $F_{2,K-N}$ statistics: (a) TEST1, 5%, (b) TEST1, 1%, (c) TEST2, 5%, (d) TEST2, 1%.



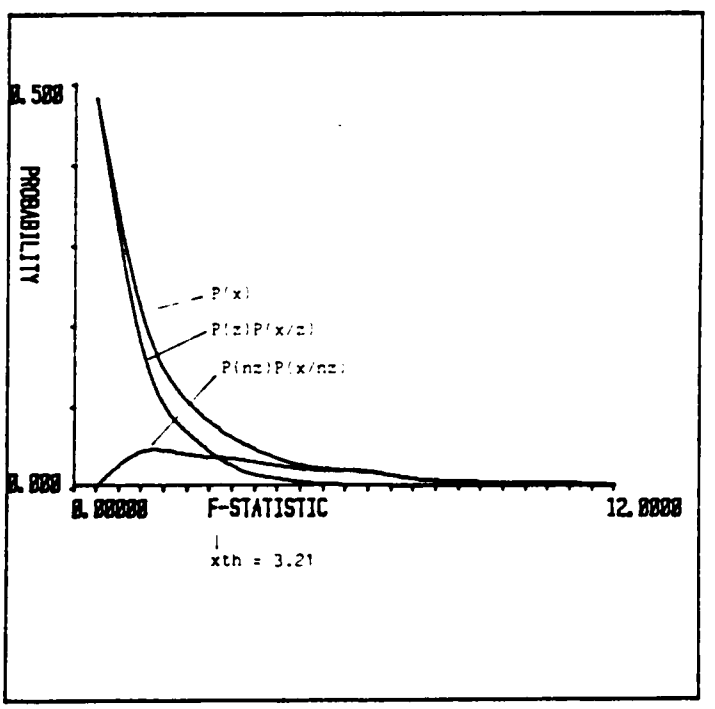
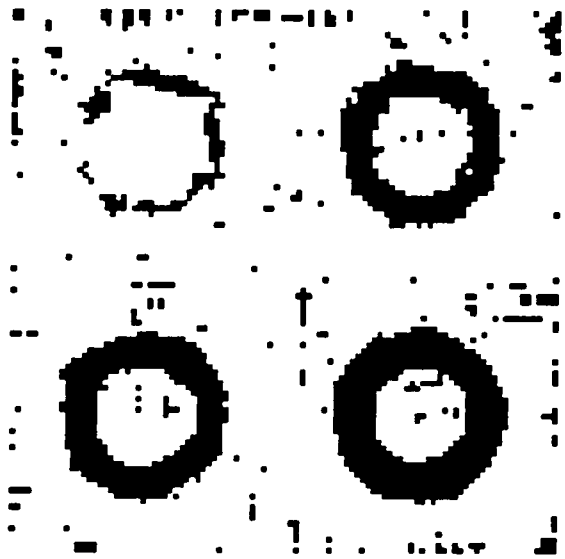
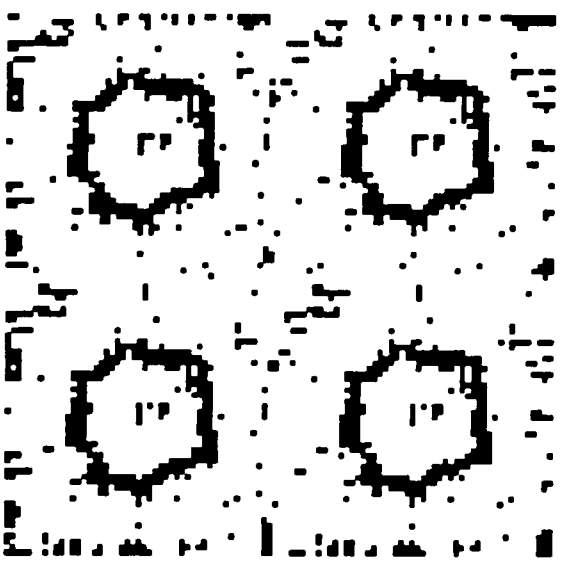


Figure 39. shows the histograms P(x), P(x|z), and P(x|nz) for the $F_{2,K-N}$ statistic: (a) TEST1, (b) TEST2.

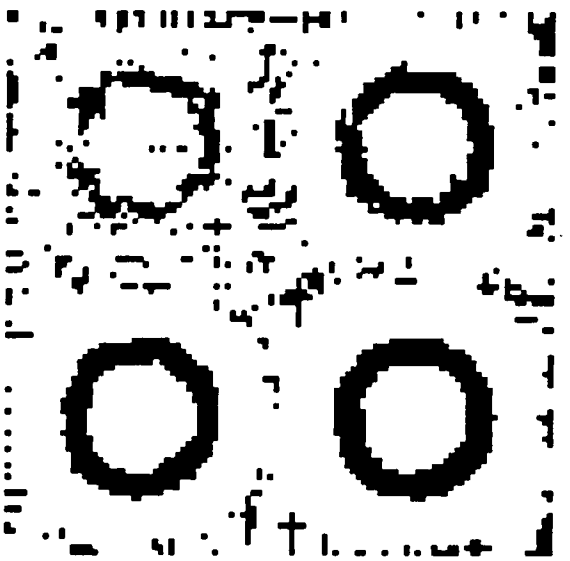


(a)

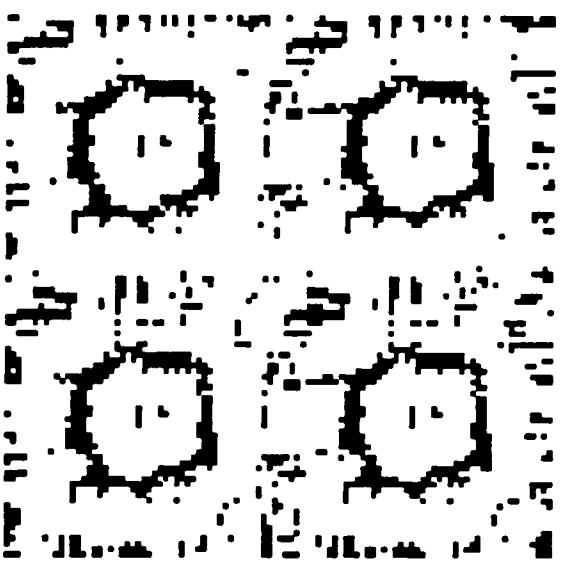


(b)

Figure 40. shows the thresholds on the $F_{2,K-N}$ statistics from the Bayesian method: (a) TEST1, (b) TEST2.



(a)



(b)

Figure 41. shows the running mean threshold applied to the gradient images: (a) TEST1, (b) TEST2.

Chapter 5

LOCAL EDGE DETECTION

5.1 Introduction

Edges are local discontinuities in some image attribute. We are concerned in this chapter with the detection of significant local spatial discontinuities in gray level intensity, i.e., step edges, as an essential initial stage in computer vision which provides the primitives necessary for later processing.

Local edge detection can be regarded as a transformation from an array of gray level intensities to an array of edge elements described by their strength and orientation. This transformation usually consists of two steps. In the first step, a local edge operator, applied to the image, selects suitable edge candidates and measures of the strength and orientation of the edge candidates are determined. In the second step, edges are declared only for those candidates whose edge strength exceeds a threshold. This second step is necessary to distinguish true edges arising from a variation of a physical characteristic

of the scene, from false edges arising through random fluctuations in the gray level intensity due to noise.

Natural intuition about edges suggests that edge candidates should be taken from those pixels having an extrema of the first directional derivative or a zero-crossing of the second directional derivative of the gray level intensity function along the gradient. Edge detection schemes have thus been proposed by Haralick (1984), Canny (1986), and Torre and Poggio (1986).

A convenient way to compute derivatives of the gray level intensity function is to approximate the intensity function locally, within a given sized neighborhood, with a suitable low order polynomial. The order of the approximating polynomial should be high enough to capture the essential structure of the local gray level intensity changes, but low enough to prevent fitting the noise unavoidably present in the image data. Determining with subpixel resolution the position and orientation of zero-crossings of the second directional derivative requires the approximating polynomial to be at least a cubic (Haralick (1984)).

Function approximation is not only convenient but provides also needed regularization of the image data to make the differentiation problem well-posed. Torre and Poggio (1986) have shown that one solution to the regularization problem is to approximate the data with a cubic spline. Function approximation, in order to estimate derivatives of the image, has been used by Prewitt (1970), Brooks(1978), Beaudet(1978), Hashimoto and Sklansky (1987), and it is explicitly used in the facet model (Haralick (1980), Haralick and Watson (1981)).

Several issues are addressed in this chapter. Section 5.2 provides a review of several least-squares function approximation methods using polynomials orthogonal over finite and infinite intervals. It is shown that these approximation methods yield polynomial coefficients that can be computed through convolution with derivatives of Gaussian or

roughly Gaussian functions. Section 5.3 generalizes Haralick's zero-crossing of second directional derivative edge operator, based on a discrete Chebyshev cubic facet, to a number of other approximation types. The relationship between a third-order approximation and a first-order approximation is illustrated. In particular, a Hermite, ie., Gaussian-weighted, least-squares approximation of the image f with a first-order polynomial yields an estimate of first directional derivative $\partial f/\partial u$ equal to $f * (\partial G/\partial u)$. This type of operator has been used by Young (1985), Canny (1986) and Torre and Poggio (1986). Image approximation with a third-order polynomial, however, produces an estimate of $\partial f/\partial u$ equal to $f * \frac{\partial}{\partial u}(G - \frac{\sigma^2}{2}\nabla^2 G)$ and estimates of second directional derivative $\partial^2 f/\partial u^2$ and third directional derivative $\partial^3 f/\partial u^3$ given by $f * (\partial^2 G/\partial u^2)$ and $f * (\partial^3 G/\partial u^3)$ respectively. The order of approximation, thus, affects the functional form of the estimated derivatives. These results are of particular significance since a number of edge detection filters satisfying different optimality criteria have been shown to be approximated by first or second derivatives of Gaussian functions (Marr and Hildredth (1980), Canny (1986), Torre and Poggio (1986), Shanmugan et. al. (1979), Lunscher and Beddoes (1986)). We show how the integrated directional derivative technique (Zuniga and Haralick (1987)), and Bayesian gradient threshold selection are used to improve on the detection and angular localization of zero-crossings. The performance of the edge operator under noise is investigated in section 5.4 and results for a number of real images are illustrated. Finally, a summary and conclusions are given in section 5.5.

5.2 *Function Approximation with Orthogonal Polynomials.*

5.2.1 Continuous Domain.

The problem in the one-dimensional case can be stated as follows. Approximate the observed (noisy) function f(x) by a polynomial y(x) of degree n over a continuous interval domain D such that the integrand $\int_D w(x)(f(x)-y(x))^2dx$ is minimized, where w(x) is a continuous, nonnegative function that provides a measure of the relative precision of the observed value f(x) at each point x in the domain D. The statistically oriented reader will recognize that the value 1/w(x) could be a measure of the variance of the noisy observation f(x).

A computationally efficient and numerically stable solution to this least-squares problem is obtained when y(x) is expressed as a linear combination of orthogonal polynomials, that is

$$y(x) = \sum_{k=0}^{n} a_k P_k(x) \tag{5.1}$$

where $\{P_k(x), k = 0,...,n\}$ is an orthogonal set relative to the weighting function w(x) over the domain D, i.e.,

$$\int_D P_i(x)P_j(x)w(x)dx \neq 0 \Leftrightarrow i = j. \tag{5.2}$$

This problem has been studied extensively and for several choices of the weighting function w(x), it leads to the classic orthogonal polynomials (Szego (1975), Hildebrand (1956)).

The Fourier coefficients a_k are given by

$$a_k = \frac{\int_D f(x)P_k(x)w(x)dx}{\int_D P_k^2(x)w(x)dx} . \tag{5.3}$$

The extension to two dimensions is easily accomplished by constructing the two-dimensional orthogonal set $\{P_i(r)P_j(c), \forall i,j \ni 0 \le i+j \le n\}$ relative to the weighting function $w(r,c) = w(r)w(c)$, over the domain D^2 , from the orthogonal set $\{P_k(x), k = 0,...,n\}$ relative to the weighting function w(x) over the domain D.

The Fourier coefficients a_{ij} in the least-squares orthogonal polynomial approximation y(r,c), of degree n,

$$y(r,c) = \sum_{\substack{i \\ 0 \le i+j \le n}} \sum_{j} a_{ij}P_i(r)P_j(c) \ , \ (r,c) \in D^2 \tag{5.4}$$

to the gray level intensity function f(r,c), that minimize the integrand

$$\int_D \int_D w(r,c)(f(r,c) - y(r,c))^2 drdc \tag{5.5}$$

are given by

$$a_{ij} = \frac{\int_D \int_D f(r,c) P_i(r) P_j(c) w(r,c) dr dc}{\int_D \int_D P_i^2(r) P_j^2(c) w(r,c) dr dc}. \tag{5.6}$$

We consider next the cases of approximation with Legendre and with Hermite polynomials.

5.2.1.1 Legendre Approximation.

Legendre polynomials are orthogonal over the interval (-1,1) relative to the weighting function $w(x) = 1$. A change of variables enables us to express the polynomials over the interval (-N,N):

$$P_k(x/N) = \frac{(-1)^k N^k}{2^k k!} \frac{d^k}{dx^k} (1 - (x/N)^2)^k. \tag{5.7}$$

The first four polynomials are given in appendix C.

In two dimensions $w(r,c) = w(r)w(c) = 1$ and from equation (5.6) the coefficients a_{ij} are given by

$$a_{ij} = \frac{(2i+1)(2j+1)}{4N^2} \int_{-N}^{N} \int_{-N}^{N} f(r,c) P_i(r/N) P_j(c/N) dr dc. \tag{5.8}$$

Let

$$g_k(x) = \frac{(2k+1)!}{(k!)^2 2^{2k+1} N} (1 - (x/N)^2)^k \tag{5.9}$$

and

$$g_{ij}(r,c) = g_i(r)g_j(c). \tag{5.10}$$

Let us further define for notational convenience

$$g_{ij}^{(i,j)}(r,c) = \frac{\partial^{i+j}}{\partial r^i \partial c^j} g_{ij}(r,c).$$

Then using equations (5.7), (5.9) and (5.10) in equation (5.8) a_{ij} can be expressed as

$$a_{ij} = \frac{(-1)^{i+j}(2N)^{i+j} i!\, j!}{(2i)!(2j)!} \int_{-N}^{N}\int_{-N}^{N} f(r,c) g_{ij}^{(i,j)}(r,c)\, dr\, dc. \tag{5.11}$$

If this function approximation is done at every (2N+1)x(2N+1) neighborhood in the image f then the coefficients a_{ij} at every neighborhood can be expressed as a convolution

$$a_{ij} = \frac{(2N)^{i+j} i!\, j!}{(2i)!(2j)!} f * g_{ij}^{(i,j)}. \tag{5.12}$$

Notice that $g_k(x)$ is an even function with a maximum at the origin and falling off to zero at the interval ends. It roughly approximates a Gaussian for k > 1 (see Figure 42 on page 132). It can also be verified that

$$\int_{-N}^{N}\int_{-N}^{N} g_{ij}(r,c)\, dr\, dc = 1.$$

From equations (5.4), (5.7) and (5.11), the partial derivatives of order (n-1) and n of the approximating polynomial y(r,c) at the neighborhood center are

$$\begin{aligned} y^{(i,j)}(0,0) &= a_{ij} P_i^{(i)}(0) P_j^{(j)}(0) \\ &= (-1)^{i+j} \int_{-N}^{N} \int_{-N}^{N} f(r,c) g_{ij}^{(i,j)}(r,c) dr dc. \end{aligned} \tag{5.13}$$

These partial derivatives at every pixel in the image can thus be expressed as

$$y^{(i,j)} = f * g_{ij}^{(i,j)} \quad , \quad n-1 \le i+j \le n. \tag{5.14}$$

5.2.1.2 Hermite Approximation.

Hermite polynomials $H(x)$ are orthogonal over the interval $(-\infty \le x \le \infty)$ relative to the weighting function $w(x) = \exp(-x^2)$. A change of variables enables us to express Hermite polynomials $H(x/\sqrt{2}\,\sigma)$ orthogonal over the interval $(-\infty \le x \le \infty)$ relative to the weighting function $w(x) = \exp(-x^2/2\sigma^2)$

$$H_k(x/\sqrt{2}\,\sigma) = (-1)^k (\sqrt{2}\,\sigma)^k \exp(x^2/2\sigma^2) \frac{d^k}{dx^k} \exp(-x^2/2\sigma^2). \tag{5.15}$$

The first four polynomials are shown in appendix C.

In two dimensions $w(r,c) = w(r)w(c) = \exp(-(r^2 + c^2)/2\sigma^2)$, and using equation (5.6) the coefficients a_{ij} of the polynomial expansion are given by

$$a_{ij} = \frac{1}{2^{i+j} i!\, j!} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(r,c) H_i(r/\sqrt{2}\,\sigma) H_j(c/\sqrt{2}\,\sigma) f(r,c) dr dc \tag{5.16}$$

where G(r,c) is the two-dimensional Gaussian function

$$G(r,c) = \frac{1}{2\pi\sigma^2} \exp(-(r^2 + c^2)/2\sigma^2). \tag{5.17}$$

Then using equation (5.15) in equation (5.16) results in

$$a_{ij} = \frac{(-1)^{i+j}(\sqrt{2}\,\sigma)^{i+j}}{2^{i+j}\, i!\, j!} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(r,c) G^{(i,j)}(r,c)\, dr dc. \tag{5.18}$$

Since $(-1)^{i+j} G^{(i,j)}(r,c) = G^{(i,j)}(-r,-c)$, the coefficients a_{ij} at every neighborhood can be expressed as

$$a_{ij} = \frac{(\sqrt{2}\,\sigma)^{i+j}}{2^{i+j}\, i!\, j!} f * G^{(i,j)}. \tag{5.19}$$

Thus, the coefficients of the least-squares local expansion of the image f in terms of Hermite polynomials can be obtained by convolving the image with the partial derivatives of a two-dimensional Gaussian.

The partial derivatives of order (n-1) and n of the approximating polynomial y(r,c) at the neighborhood center are

$$\begin{aligned} y^{(i,j)}(0,0) &= a_{ij} H_i^{(i)}(0) H_j^{(j)}(0) \\ &= (-1)^{i+j} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(r,c) G^{(i,j)}(r,c)\, dr dc. \end{aligned} \tag{5.20}$$

These partial derivatives at every pixel in the image can thus be estimated by

$$y^{(i,j)} = f * G^{(i,j)} \quad , \quad n-1 \le i+j \le n. \tag{5.21}$$

Similar results have been reported by Young (1985) and Hashimoto and Sklansky (1987).

5.2.2 Discrete domain.

The analysis in the continuous domain performed in the previous section enables us to obtain valuable insight into the nature and properties of function approximation with orthogonal polynomials. Unfortunately, the observed gray level intensity function f(r,c) is typically available only at discrete values of the coordinates (r,c). One method often used to approximate in the discrete domain the coefficients a_{ij} of the function approximation in the continuous domain, given by equations (5.12) and (5.19), is to perform discrete convolution with the sampled values of the continuous domain kernels $g_{ij}^{(i,j)}(r,c)$ and $G^{(i,j)}(r,c)$. In this case the kernel values must be manipulated by a small amount to ensure the values add up to zero. Another approach is to find sets of discrete orthogonal polynomials which asymptotically approach the properties of the continuous orthogonal polynomials considered in the previous section. For the Legendre and Hermite polynomials, the corresponding discrete orthogonal polynomials are the Gram polynomials (also called discrete Chebyshev polynomials (Hildebrand (1956))) and the discrete Krawtchouk polynomials (Szego (1939)).

The discrete domain function approximation problem can be stated as follows. Let R and C be discrete index sets. Approximate the observed noisy samples of the image function f(r,c) by the polynomial y(r,c) of degree n over the discrete domain R X C such that the weighted error $\sum_{r \in R} \sum_{c \in C} w(r,c)(f(r,c) - y(r,c))^2$ is minimized. This is the essence of the facet model.

It is convenient to express y(r,c) as a linear combination of discrete orthogonal polynomials

$$y(r,c) = \sum_{\substack{i \\ 0 \le i+j \le n}} \sum_{j} a_{ij} p_i(r) p_j(c) \ , \ (r,c) \in RXC \tag{5.22}$$

where $\{p_i(r)p_j(c), \forall i,j \ni 0 \le i+j \le n\}$ is an orthogonal set relative to the weighting function $w(r,c) = w(r)w(c)$ over the discrete domain RXC. This two-dimensional orthogonal set is easily obtained by tensor products from the one-dimensional orthogonal set $\{p_k(x), 0 \le k \le n\}$ relative to the weighting function w(x), that is

$$\sum_{x} w(x) p_i(x) p_j(x) \neq 0 \Leftrightarrow i = j. \tag{5.23}$$

The Fourier coefficients a_{ij} are given by

$$a_{ij} = \frac{\sum_{r \in R} \sum_{c \in C} f(r,c) p_i(r) p_j(c) w(r,c)}{\sum_{r \in R} \sum_{c \in C} p_i^2(r) p_j^2(c) w(r,c)}. \tag{5.24}$$

We discuss next the Gram polynomials and the Krawtchouk polynomials. Both of these polynomials are usually defined over $\{0, ..., N\}$. For convenience in our applications we will define these polynomials over the symmetric range $\{-N, ..., 0, ..., N\}$.

5.2.2.1 Gram approximation.

The Gram polynomials are orthogonal over the discrete range $\{-N, ..., 0, ..., N\}$ relative to the weighting function $w(x) = 1$. They are given by

$$p_m(x) = \frac{(-1)^m}{\binom{2N}{m}} \sum_{i=0}^{m} (-1)^i \binom{m+i}{m} \binom{2N-i}{m-i} \binom{N+x}{i}. \tag{5.25}$$

The first four polynomials are given in appendix C.

It can be verified that $p_m(Nx)$ tends to the mth Legendre polynomial (Hildebrand (1956), Szego (1939))

$$\lim_{N \to \infty} p_m(Nx) = P_m(x). \tag{5.26}$$

From equation (5.24) the coefficients a_{ij} are given by

$$a_{ij} = \frac{(2i+1)(2j+1)\binom{2N}{i}\binom{2N}{j}}{(i+1)(j+1)\binom{2N+i+1}{i+1}\binom{2N+j+1}{j+1}} \sum_{r=-N}^{N} \sum_{c=-N}^{N} f(r,c) p_i(r) p_j(c). \tag{5.27}$$

Haralick (1984) has used function approximation with Gram polynomials in the facet model to detect edges.

5.2.2.2 *Discrete Krawtchouk approximation.*

The discrete Krawtchouk polynomials are orthogonal over the discrete range $\{-N, ..., 0, ..., N\}$ relative to the binomial weighting function $w(x) = \binom{2N}{N+x} p^{N+x} q^{N-x}$, $p > 0, q > 0, p + q = 1$. With $p = q = 1/2$, $w(x) = \binom{2N}{N+x}/4^N$, and the Krawtchouk polynomials can then be computed by (Szego (1939))

$$k_m(x) = \frac{(-1)^m}{2^m} \sum_{i=0}^{m} (-1)^i \binom{N-x}{m-i} \binom{N+x}{i} . \quad . \tag{5.28}$$

The first four polynomials are given in appendix C.

It can be verified that $k_m(x)$ tends to the mth Hermite polynomial (Szego (1939)).

$$\lim_{N \to \infty} (2^{3m} m!)^{1/2} \binom{2N}{m}^{-1/2} k_m(\sqrt{N}\, x) = H_m(x). \tag{5.29}$$

Likewise, it is well known that the binomial weighting function $w(x) = \binom{2N}{N+x}/4^N$, $x \in \{-N, ..., 0, ..., N\}$ tends to the Gaussian function with variance $\sigma^2 = N/2$.

From equation (5.24), the coefficients a_{ij} are given by

$$a_{ij} = \frac{4^{i+j-2N}}{\binom{N}{i}\binom{N}{j}} \sum_{r=-N}^{N} \sum_{c=-N}^{N} f(r,c) \binom{2N}{N+r} \binom{2N}{N+c} k_i(r) k_j(c). \tag{5.30}$$

Hashimoto and Sklansky (1987) use a Krawtchouk approximation of order n to estimate the nth derivative of the image function in a multiple-order derivative scheme to detect step edges. Young (1985) also uses Krawtchouk approximation in edge detection.

5.3 *Edge Detection Using the Cubic Facet*

As previously mentioned, determining with subpixel resolution the position and orientation of zero-crossings of the second directional derivative along the gradient requires the approximating polynomial to be at least a bivariate cubic. This cubic facet model for edge detection has been proposed by Haralick (1984) using the Gram polynomials and its performance under noise has been shown to compare favorably against other existing edge detectors.

The cubic facet zero-crossing edge detector is generalized here using least-squares approximation with various types of continuous and discrete orthogonal polynomials. Once zero-crossings of the second directional derivative have been determined, they are assigned the strength and orientation of the first directional derivative. We show how the integrated first directional derivative technique (Zuniga and Haralick (1987)) can be used to improve on the estimation of these edge descriptors. Further, a decision based on first directional derivative strength must be made as to which zero-crossings constitute valid edges. We show how a Bayesian gradient threshold technique can be used in performing such a decision.

5.3.1 From Pixels to Edges

For each (2N+1)X(2N+1) neighborhood in the image, a bivariate cubic y(r,c) which best approximates the gray level intensity function f(r,c) in some weighted least-squares sense is determined. The directional derivatives of the approximating polynomial along the gradient provide local estimates for the directional derivatives of the image f.

Edge candidates are taken from pixels having an extrema of first directional derivative along the gradient. If the first directional derivative is to be a relative maximum, then the second directional derivative must be zero and the third directional derivative must be negative if the edge is crossed from the lower value to the higher value of the gray level intensity function.

The simplest way to think about directional derivatives is to cut the surface y(r,c) with a plane which is oriented in the desired direction and which is orthogonal to the row-column plane. By convention, we take the angle to be measured clockwise from the column axis. To cut the surface y(r,c) with a plane in the direction θ we just require that $r = \rho \sin\theta$ and $c = \rho \cos\theta$ where ρ is the independent variable. This requirement produces the curve $y_\theta(\rho)$ which can be expressed as

$$y_\theta(\rho) = y_\theta(0) + y_\theta^{(1)}(0)\rho + \frac{1}{2}y_\theta^{(2)}(0)\rho^2 + \frac{1}{6}y_\theta^{(3)}(0)\rho^3 \tag{5.31}$$

Therefore the first, second, and third directional derivatives are given by

$$y_\theta^{(1)}(\rho) = y_\theta^{(1)}(0) + y_\theta^{(2)}(0)\rho + \frac{1}{2}y_\theta^{(3)}(0)\rho^2 \tag{5.32}$$

$$y_\theta^{(2)}(\rho) = y_\theta^{(2)}(0) + y_\theta^{(3)}(0)\rho \tag{5.33}$$

$$y_\theta^{(3)}(\rho) = y_\theta^{(3)}(0) \tag{5.34}$$

We define the desired direction to be the gradient direction at the center of the given pixel. For a pixel to be an edge candidate, the second directional derivative must have a negatively sloped zero crossing sufficiently near the center of the pixel. In this case, with the origin taken as the center of the pixel, there must be a ρ sufficiently small in magnitude satisfying $y_\theta^{(2)}(\rho) = 0$ and $y_\theta^{(3)}(\rho) < 0$. If $y_\theta^{(3)}(\rho) < 0$ then from equation (5.33) a ρ having value $-y_\theta^{(2)}(0)/y_\theta^{(3)}(0)$ exists which makes $y_\theta^{(2)}(\rho) = 0$. If $|y_\theta^{(2)}(0)/y_\theta^{(3)}(0)| < \rho_0$, where we take ρ_0 to be somewhat less than a pixel length, then the pixel qualifies as an edge candidate. Edges are then declared at those edge candidates for which $y_\theta^{(1)}(0)$ exceeds a given threshold. In essence, this is the procedure given by Haralick (1984).

5.3.2 Directional Derivatives

Implementation of the cubic facet zero-crossing edge detector requires the computation of directional derivatives of the approximating polynomial y(r,c) up to third order, evaluated at the neighborhood center.

Directional derivatives of arbitrary order k can be computed by

$$y_\theta^{(k)} = \sum_{i=0}^{k} \binom{k}{i} y^{(k-i,i)} \sin^{k-i}\theta \cos^{i}\theta \tag{5.35}$$

The first three derivatives are

$$y_\theta^{(1)} = y^{(1,0)} \sin\theta + y^{(0,1)} \cos\theta \tag{5.36}$$

$$y_\theta^{(2)} = y^{(2,0)} \sin^2\theta + 2y^{(1,1)} \sin\theta \cos\theta + y^{(0,2)} \cos^2\theta$$

$$y_\theta^{(3)} = y^{(3,0)} \sin^3\theta + 3y^{(2,1)} \sin^2\theta \cos\theta + 3y^{(1,2)} \sin\theta \cos^2\theta + y^{(0,3)} \cos^3\theta$$

Thus, these directional derivatives can be readily computed from the partial derivatives $y^{(i,j)}$ of the approximating polynomial y(r,c) using equation (5.4) in the continuous domain and equation (5.22) in the discrete domain. These partial derivatives can, in turn, be expressed in terms of the least-squares fitting coefficients a_{ij} given by equations (5.12), (5.19), (5.27), and (5.30).

For Hermite approximation, further simplification of the directional derivatives is possible. Using equation (5.21) in equation (5.35) results in

$$y_\theta^{(k)}(0) = f * \sum_{i=0}^{k} \binom{k}{i} G^{(k-i,i)} \sin^{k-i}\theta \cos^i\theta \ , \quad 2 \le k \le 3$$

or equivalently

$$\partial^k y / \partial\rho^k = f * (\partial^k G / \partial\rho^k) \tag{5.37}$$

Therefore, the second and third directional derivatives $\partial^2 y/\partial\rho^2$ and $\partial^3 y/\partial\rho^3$ of the Hermite approximation polynomial of third order, evaluated at the neighborhood center, for every neighborhood in the image are given by $f * (\partial^2 G/\partial\rho^2)$ and $f * (\partial^3 G/\partial\rho^3)$ respectively.

In order to compute the first directional derivative, we proceed as follows. From equation (5.4), the first partial derivatives at the origin are

$$y^{(1,0)}(0,0) = a_{10} H_1^{(1,0)}(0) H_0(0) + a_{30} H_3^{(1,0)}(0) H_0(0) + a_{12} H_1^{(1,0)}(0) H_2(0)$$

$$y^{(0,1)}(0,0) = a_{01}H_0(0)H_1^{(0,1)}(0) + a_{03}H_0(0)H_3^{(0,1)}(0) + a_{21}H_2(0)H_1^{(0,1)}(0)$$

Using equation (5.19), these partial derivatives, at every neighborhood in the image, become

$$y^{(1,0)} = f * (G^{(1,0)} - \frac{\sigma^2}{2}(G^{(3,0)} + G^{(1,2)}))$$

$$y^{(0,1)} = f * (G^{(0,1)} - \frac{\sigma^2}{2}(G^{(0,3)} + G^{(2,1)}))$$

or equivalently

$$\partial y/\partial r = f * \frac{\partial}{\partial r}(G - \frac{\sigma^2}{2}\nabla^2 G) \tag{5.38}$$

$$\partial y/\partial c = f * \frac{\partial}{\partial c}(G - \frac{\sigma^2}{2}\nabla^2 G) \tag{5.39}$$

Finally, from equation (5.36) the first directional derivative of the Hermite approximation bivariate cubic polynomial, evaluated at the origin, for every neighborhood in the image, can be expressed in the convolution form

$$\partial y/\partial \rho = f * \frac{\partial}{\partial \rho}(G - \frac{\sigma^2}{2}\nabla^2 G) \tag{5.40}$$

Using equation (5.21) in equation (5.40), we can establish a relationship between the first derivatives of the approximation polynomial of first and third order, evaluated at the origin.

$$\frac{\partial y_3}{\partial \rho}(0) = \frac{\partial}{\partial \rho}(y_1 - \frac{\sigma^2}{2}\nabla^2 y_3)(0) \tag{5.41}$$

where y_1 and y_3 represent the approximating polynomials of first and third order respectively. Thus, while a first order Hermite approximation yields an estimate of first directional derivative $\partial y_1/\partial\rho$ equal to $f * (\partial G/\partial\rho)$, a third order approximation yields an estimate of first directional derivative $\partial y_3/\partial\rho$ equal to $f * \frac{\partial}{\partial\rho}(G - \frac{\sigma^2}{2}\nabla^2 G)$.

It can be verified that $f * (G - \frac{\sigma^2}{2}\nabla^2 G)$ is the value of the Hermite approximation polynomial evaluated at the origin, i.e., y(0,0), for every neighborhood, since from equation (5.4)

$$y(0,0) = a_{00}H_0^2(0) + a_{20}H_2(0)H_0(0) + a_{02}H_0(0)H_2(0)$$

and using equation (5.19)

$$y = f * (G - \frac{\sigma^2}{2}(G^{(2,0)} + G^{(0,2)}))$$

Finally

$$y = f * (G - \frac{\sigma^2}{2}\nabla^2 G) \tag{5.42}$$

Equations analog to equation (5.41), relating the first derivative of the approximating polynomial of third order to the first derivative of the approximating polynomial of first order, evaluated at the origin, can also be obtained for other approximation types.

For Legendre approximation

$$y_3^{(1,0)} = y_1^{(1,0)} - \frac{N^2}{10}y_3^{(3,0)} - \frac{N^2}{6}y_3^{(1,2)} \tag{5.43}$$

$$y_3^{(0,1)} = y_1^{(0,1)} - \frac{N^2}{10}y_3^{(0,3)} - \frac{N^2}{6}y_3^{(2,1)}$$

For Gram approximation

$$y_3^{(1,0)} = y_1^{(1,0)} - \frac{3N^2 + 3N - 1}{30} y_3^{(3,0)} - \frac{N(N+1)}{6} y_3^{(1,2)} \tag{5.44}$$

$$y_3^{(0,1)} = y_1^{(0,1)} - \frac{3N^2 + 3N - 1}{30} y_3^{(0,3)} - \frac{N(N+1)}{6} y_3^{(2,1)}$$

and for Discrete Krawtchouk approximation

$$y_3^{(1,0)} = y_1^{(1,0)} - \frac{3N - 1}{12} y_3^{(3,0)} - \frac{N}{4} y_3^{(1,2)} \tag{5.45}$$

$$y_3^{(0,1)} = y_1^{(0,1)} - \frac{3N - 1}{12} y_3^{(0,3)} - \frac{N}{4} y_3^{(2,1)}$$

It is evident from these equations that as $N \to \infty$, the Gram approximation tends to the Legendre approximation, and the discrete Krawtchouk approximation tends to the Hermite approximation with variance $\sigma^2 = N/2$.

5.3.3 The Integrated First Directional Derivative

Given the bivariate function y(r,c), let the row-column coordinate system r-c be rotated clockwise by θ degrees measured from the column axis. Let $y_\theta(\omega, \rho)$ be the expression of y(r,c) in the rotated coordinate system $\omega - \rho$. The integrated first directional derivative Δy_θ is defined by

$$\Delta y_\theta = \frac{1}{4L^2} \int_{-L}^{L} \int_{-L}^{L} \frac{\partial}{\partial \rho} y_\theta(\omega, \rho) d\rho d\omega \tag{5.46}$$

or equivalently

$$\Delta y_\theta = \frac{1}{2L}\left(\frac{1}{2L}\int_{-L}^{L} y_\theta(\omega, L)d\omega - \frac{1}{2L}\int_{-L}^{L} y_\theta(\omega, -L)d\omega\right)$$

It has been shown (Zuniga and Haralick (1987)) that if y(r,c) is the bivariate cubic in canonical form

$$\begin{aligned} f(r,c) = & K_1 \\ & + K_2 r + K_3 c \\ & + K_4 r^2 + K_5 rc + K_6 c^2 \\ & + K_7 r^3 + K_8 r^2 c + K_9 rc^2 + K_{10} c^3 \end{aligned} \tag{5.47}$$

then Δy_θ is given by

$$\Delta y_\theta = (K_2 + L^2 K_7 + \frac{L^2}{3} K_9)\sin\theta + (K_3 + L^2 K_{10} + \frac{L^2}{3} K_8)\cos\theta \tag{5.48}$$

Since the coefficients K_i are related to the partial derivatives of y(r,c) evaluated at the origin, equation (5.48) can be rewritten as

$$\Delta y_\theta = (y^{(1,0)} + \frac{L^2}{6}(y^{(3,0)} + y^{(1,2)}))\sin\theta + (y^{(0,1)} + \frac{L^2}{6}(y^{(0,3)} + y^{(2,1)}))\cos\theta$$

or equivalently

$$\Delta y_\theta = \frac{\partial}{\partial r}(y + \frac{L^2}{6}\nabla^2 y)\sin\theta + \frac{\partial}{\partial c}(y + \frac{L^2}{6}\nabla^2 y)\cos\theta \tag{5.49}$$

Finally

$$\Delta y_\theta = \frac{\partial}{\partial \rho}(y + \frac{L^2}{6}\nabla^2 y)(0,0) \tag{5.50}$$

Further, it can be verified that the value of $(y + \frac{L^2}{6}\nabla^2 y)$ at the origin is the the result of the integral $\frac{1}{4L^2}\int_{-L}^{L}\int_{-L}^{L} y(r,c)drdc$.

If y(r,c) is the Hermite approximation bivariate cubic, then use of equation (5.40) in equation (5.50) yields

$$\Delta y_\theta = f * \frac{\partial}{\partial \rho}(G - (\frac{\sigma^2}{2} - \frac{L^2}{6})\nabla^2 G) \tag{5.51}$$

5.3.4 The Standard Cubic Facet Gradient vs the Integrated First Directional Derivative Gradient

Experiments have been reported on the accuracy of estimates of the orientation of step and ramp edges using a number of gradient operators (Zuniga and Haralick (1987)). In particular, it has been shown that an implementation of the integrated first directional derivative gradient operator $\nabla(y + \frac{L^2}{6}\nabla^2 y)$ with a Gram approximation cubic polynomial y(r,c), yields a very significant reduction in edge orientation estimate bias as compared with other gradient operators, including the standard cubic facet gradient operator ∇y, also implemented with Gram polynomials.

In addition to improved zero-crossing angular localization, it is shown in this section that the magnitude of the integrated operator $\nabla(y + \frac{L^2}{6}\nabla^2 y)$ posseses as well a number of desirable characteristics and advantages when compared with the magnitude of the standard operator ∇y . These advantages become clear when comparing the step edge response and the unit width line response of the operators. These include:

1. Relative insensitivity of the magnitude response to edge or line orientation.

2. Some situations that give rise to false zero-crossings when using the Standard operator ∇y are avoided when using the integrated operator $\nabla(y + \frac{L^2}{6}\nabla^2 y)$

3. Reduced sensitivity to noise (shown in section 5.4).

Figure 43 on page 133 and Figure 44 on page 134 illustrate how edge and line orientation affect the value of the response of the standard directional derivative $\partial y/\partial \rho$ across the edge or line. Figure 45 on page 135 and Figure 46 on page 136 illustrate the same for the integrated directional derivative $\frac{\partial}{\partial \rho}(y + \frac{L^2}{6}\nabla^2 y)$. The size of the operators is 5 x 5 pixels. It is clear from observing the edge and line response of the operators, that the integrated operator is much less sensitive to edge and line orientation than the standard operator. This is consistent with earlier results (Zuniga and Haralick (1987)) showing that edge orientation estimate bias of the integrated operator $\nabla(y + \frac{L^2}{6}\nabla^2 y)$ is significantly lower than the bias of the standard gradient operator ∇y . We have verified that these observations apply as well for larger operators. The value of L used in the $(2N+1)$x$(2N+1)$ Integrated operator is $L = 0.36(2N + 1)$. This value of L minimizes the edge orientation estimate bias of the Integrated operator $\nabla(y + \frac{L^2}{6}\nabla^2 y)$.

Figure 47 on page 137 shows the responses of the first directional derivative $\partial y/\partial \rho$, and the third directional derivative $\partial^3 y/\partial \rho^3$ for a 45 degree step edge. Figure 48 on page 138 shows the responses of the integrated first directional derivative $\frac{\partial}{\partial \rho}(y + \frac{L^2}{6}\nabla^2 y)$ and third directional derivative $\partial^3 y/\partial \rho^3$. Figure 49 on page 139 shows zero-crossing displacement from the neighborhood center as a function of edge displacement. The operator size was 5x5 pixels. According to the discussion in section 5.3.1 an edge candidate pixel should have a zero-crossing sufficiently near to the pixel's center and such that the first directional derivative and third directional derivative have opposite signs. From Figure 47 on page 137, Figure 48 on page 138, and Figure 49 on page 139 we can ob-

serve that as expected a zero-crossing meeting the above conditions occurs when the edge passes within a pixel's boundaries, but when using the standard first directional derivative $\partial y/\partial \rho$ a zero-crossing meeting the conditions also occurs when the edge is displaced by 2 pixels. This problem is caused by the change of sign of the first derivative $\partial y/\partial \rho$ that occurs when the edge is displaced by 2 pixels (see Figure 47 on page 137). This problem does not occur when using the integrated first derivative $\frac{\partial}{\partial \rho}(y + \frac{L^2}{6}\nabla^2 y)$ since its step edge response decreases smoothly to zero as edge displacement increases as shown in Figure 48 on page 138. Again, we have verified the same behavior with larger neighborhood sizes and different edge orientations. In general, with a $(2N+1)$ by $(2N+1)$ neighborhood, the change in sign in $\partial y/\partial \rho$ occurs when the edge is displaced by about N pixels.

5.3.5 Gradient Threshold Selection

Distinguishing zero-crossings of the second directional derivative, arising from a real physical event on the scene, from zero-crossings produced by image noise requires that a decision be made as to what minimum value of zero-crossing strength (gradient strength) should be considered as indicative that a valid edge has occurred. The presence of random noise in the image data causes the approximating polynomial and its derivatives to become random variables. Thus the gradient threshold selection problem becomes statistical in nature.

As indicated in the previous sections, the observed gray level intensity function f(r,c) is approximated locally in a $(2N+1)X(2N+1)$ neighborhood by a bivariate cubic polynomial y(r,c) which minimizes the weighted residual error $S^2 = \sum_{r=-N}^{N}\sum_{c=-N}^{N} w(r,c)(f(r,c) - y(r,c))^2$. If image noise is modeled as additive independent

Gaussian with variance $V(r,c) = \sigma^2/w(r,c)$, then y(r,c) is a minimum variance unbiased estimate of the true but unknown gray level intensity function of which f(r,c) is the observed value.

Let $K = (2N + 1)^2$ be the neighborhood size. Then according to the derivations in appendix B, the weighted residual error S^2 has a $\chi^2(K - 10)$ distribution. Furthermore, under the hypothesis of zero gradient (no edge present) the estimated squared gradient $|\nabla y|^2$ (or $|\nabla(y + \frac{L^2}{6}\nabla^2 y)|^2$ if the integrated directional derivative is used) evaluated at the origin divided by its variance $k\sigma^2$ has a $\chi^2(2)$ distribution, where k is given by

$$k = \sum_{r=-N}^{N}\sum_{c=-N}^{N} M_r^2(r,c)/w(r,c) = \sum_{r=-N}^{N}\sum_{c=-N}^{N} M_c^2(r,c)/w(r,c) \tag{5.52}$$

and $M_r(r,c)$ and $M_c(r,c)$ are the convolution kernels used to compute the partial derivatives $\partial y/\partial r$ and $\partial y/\partial c$ (or $\frac{\partial}{\partial r}(y + \frac{L^2}{6}\nabla^2 y)$ and $\frac{\partial}{\partial c}(y + \frac{L^2}{6}\nabla^2 y)$ if the integrated directional derivative is used) at the origin. Then, under the hypothesis of zero gradient, the statistics

$$x_1 = \frac{|\nabla y|^2/2}{k_1 S^2/(K - 10)} \tag{5.53}$$

and

$$x_2 = \frac{|\nabla(y + \frac{L^2}{6}\nabla^2 y)|^2/2}{k_2 S^2/(K - 10)} \tag{5.54}$$

have $F_{2,K-10}$ distributions.

If σ^2 is assumed to remain constant from neighborhood to neighborhood, then better results can often be obtained by replacing S^2 by E^2, the average of the residuals S^2 over M non-overlapping neighborhoods. In this case, the above statistics have $F_{2,M(K-10)}$ distributions.

In chapter 3, a Bayesian decision procedure was developed in order to select a threshold on the statistic x. This procedure selects a threshold $x = t$ that equalizes the aposteriori probabilities of zero gradient $P(z \mid x = t)$, and non-zero gradient $P(nz \mid x = t)$

.

5.4 *Experiments*

We present in this section a number of experiments to study the effectiveness of the integrated directional derivative technique and the Bayesian gradient threshold selection technique, in improving the detection of zero-crossings, when used with Haralick's zero-crossing of second directional derivative edge detector, under noisy conditions in a computer generated test image, and with real images.

5.4.1 Test image

We use a test image of size 128 by 128 pixels, similar to that used by Kitchen and Rosenfeld (1981) for edge detector evaluation. This image, shown in Figure 50 on page 140, consists of four concentric bright rings on a dark background. The rings gray level is 140 and the background gray level is 115. The central dark circle diameter is 16 pixels

and it is surrounded by four bright rings of width 8 pixels, separated from each other by 8 pixels width of background. This image was originally generated as a 512 by 512 image and then reduced to size 128 by 128, by replacing each 4 by 4 block with a single pixel having the average gray level of the block.

Independent zero-mean Gaussian noise was added to the test image at four different signal to noise ratios: 1, 5, 20, and 100. The signal to noise ratio is defined using Pratt's measure (1979)

$$SNR = \left(\frac{h}{\sigma}\right)^2$$

where h is the edge contrast and σ is the noise standard deviation. The contrast h for the test image is 25 and the standard deviation σ is adjusted to give the selected values of signal to noise ratio. The noisy images are shown in Figure 51 on page 141.

5.4.2 Edge operator parameters

The zero-crossing of second directional derivative edge detector, described in section 5.3, has a number of parameters that must be set when applying the operator to an image. These parameters are the neighborhood size, zero-crossing radius ρ_0 , gradient threshold, and choice of standard gradient or integrated directional derivative gradient.

The experiments reported in this section show how the choice of parameter values affect the performance of the edge detector. Results are shown with neighborhood sizes of 5x5 and 7x7 pixels, zero-crossing radius ρ_0 ranging from 0.3 to 1.0, in 0.1 increments, and 15 gradient thresholds from 0.1t to 10t, where t is the threshold on the $F_{M(K-10)}$ sta-

tistic obtained by the Bayesian Gradient threshold selection technique explained in section 5.3, with both Standard gradient and integrated directional derivative gradient.

5.4.3 Performance Measures

The performance measures used are the probability of the edge operator assigning an edge given that an edge is present P(AE|E), the probability of assigning no-edge when no edge is present P(ANE|NE), and the average of these two probabilities which we will refer to as the probability of correct detection.

In order to compute these probabilities, edge images produced by the edge detector are compared against the ideal edge image shown in Figure 52 on page 142. This ideal edge image is obtained by setting a pixel as an edge pixel if an edge passes through it at a distance no farther than 0.5 from the pixel's center. The generation of the ideal edge image is made possible by our complete knowledge about the geometry of the test image.

5.4.4 Performance Under Noise

Figure 53 on page 143 and Figure 54 on page 144 display curves in the P(AE|E) vs P(ANE|NE) plane, obtained by varying the gradient threshold while keeping constant the zero-crossing radius ρ_0 and the signal to noise ratio SNR. Figure 53 on page 143 corresponds to the 5x5 operators and Figure 54 on page 144 corresponds to the 7x7 operators. Each curve is marked at the place where the gradient threshold coincides with that obtained by the Bayesian gradient threshold technique. Notice the Bayesian threshold occurs near the desired place at the elbow of the curves, where the probability of

correct detection is maximized. As expected the probability of correct detection increases as the SNR increases. Also notice the improvement in detection when the integrated directional derivative is used instead of the Standard derivative.

Figure 55 on page 145 and Figure 56 on page 146 display curves in the P(AE|E) vs P(ANE|NE) plane obtained by varying the zero-crossing radius ρ_0 while keeping constant the gradient threshold and the SNR. Figure 55 on page 145 corresponds to the 5x5 Integrated operators and Figure 56 on page 146 corresponds to the 7x7 Integrated operators. Several curves are shown for gradient thresholds corresponding to 0.25t, 0.5t, t, and 2t, where t is the Bayesian threshold, and SNR values of 1, 5, 20, and 100. Each curve is marked at the place where the zero-crossing radius is 0.7.

Figure 57 on page 147 and Figure 58 on page 148 show probability of correct detection as a function of F-statistic threshold, which is expressed as a multiple of the Bayesian threshold. Figure 57 on page 147 corresponds to the 5x5 operators and Figure 58 on page 148 corresponds to the 7x7 operators. We notice, as with Figure 53 on page 143 and Figure 54 on page 144, that the probability of correct detection is maximized at a place near the Bayesian threshold. This is more pronounced when using the Integrated derivative. Again, probability of correct detection increases as SNR increases and when the Integrated derivative is used instead of the Standard derivative. Also, the 7x7 operators yield larger probability of correct detection than the smaller 5x5 operators. In a real image situation, however, scene objects size must be taken in consideration when selecting the size of the operator.

Figure 59 on page 149 and Figure 60 on page 150 show probability of correct detection as a function of zero-crossing radius ρ_0 for the 5x5 and 7x7 operators. We observe that a radius of 0.5 is near optimal for low values of SNR, but a radius of 0.7 is better for high values of SNR.

The results of applying the 5x5 and 7x7 operators to the noisy ring images are shown in Figure 61 on page 151 and Figure 62 on page 152. The threshold used was the Bayesian threshold, and the zero-crossing radius was 0.7.

5.4.5 Real Images

We conclude by presenting in Figure 66 on page 156 through Figure 71 on page 161 the results of applying 5x5 and 7x7 edge operators to the aerial images shown in Figure 63 on page 153, Figure 64 on page 154, and Figure 65 on page 155. Each image is 8 bits and the sizes are 256x256, 256x256, and 300x300 pixels respectively. The first image corresponds to an urban scene from Baton Rouge, Louisiana. Houses in this image appear as dark rectangular objects and car ports appear as bright rectangular objects. The last two images are urban scenes from Sunnyvale, California containing several bright buildings. The gradient threshold used was the Bayesian threshold and the zero-crossing radius was 0.5.

Two things are worth noticing. Firstly, for the Baton Rouge scene some houses in the lower right quarter of the image are missed by the standard cubic facet edge operator while they are correctly detected by the integrated operator. Secondly, for the Sunnyvale images the automatically selected Bayesian gradient threshold, working with the standard cubic edge operator, tends to be too low resulting in an excessive number of edges. This is more pronounced for the 5x5 operators. The results are significantly better for the integrated operator.

5.5 *Conclusions*

Several least-squares function approximation methods using polynomials orthogonal over finite and infinite intervals have been analized. It has been shown that these approximation methods yield polynomial coefficients that can be computed through convolution with derivatives of Gaussian or roughly Gaussian functions. Haralick's edge operator based on zero-crossing of second directional derivatives of a discrete Chebyshev cubic facet can be readily generalized to other approximation types.

It was shown that the order of approximation affects the functional form of the estimated derivatives. In particular, a Hermite, ie., Gaussian-weighted, least-squares approximation of the image f with a first-order polynomial yields an estimate of first directional derivative $\partial f/\partial u$ equal to $f * (\partial G/\partial u)$. Image approximation with a third-order polynomial, however, produces an estimate of $\partial f/\partial u$ equal to $f * \frac{\partial}{\partial u}(G - \frac{\sigma^2}{2}\nabla^2 G)$ and estimates of second directional derivative $\partial^2 f/\partial u^2$ and third directional derivative $\partial^3 f/\partial u^3$ given by $f * (\partial^2 G/\partial u^2)$ and $f * (\partial^3 G/\partial u^3)$ respectively.

Finally, we demonstrated that the use in Haralick's edge operator of the integrated directional derivative technique developed in chapter 3 and the Bayesian gradient threshold selection technique developed in chapter 4, result in improved detection and localization of zero-crossings. Advantages of the Integrated directional derivative over the Standard derivative include improved angular localization, reduced sensitivity to noise, relative insensitivity to edge or line orientation, and smooth approach to zero of the step edge response as edge displacement increases, which prevents creation of some false zero-crossings. Experiments with artificially generated noisy images show that filtering the zero-crossings of Haralick's edge operator using the automatically generated Baye-

sian threshold on the integrated directional derivative gradient produces edge images with near optimal probability of correct edge detection.

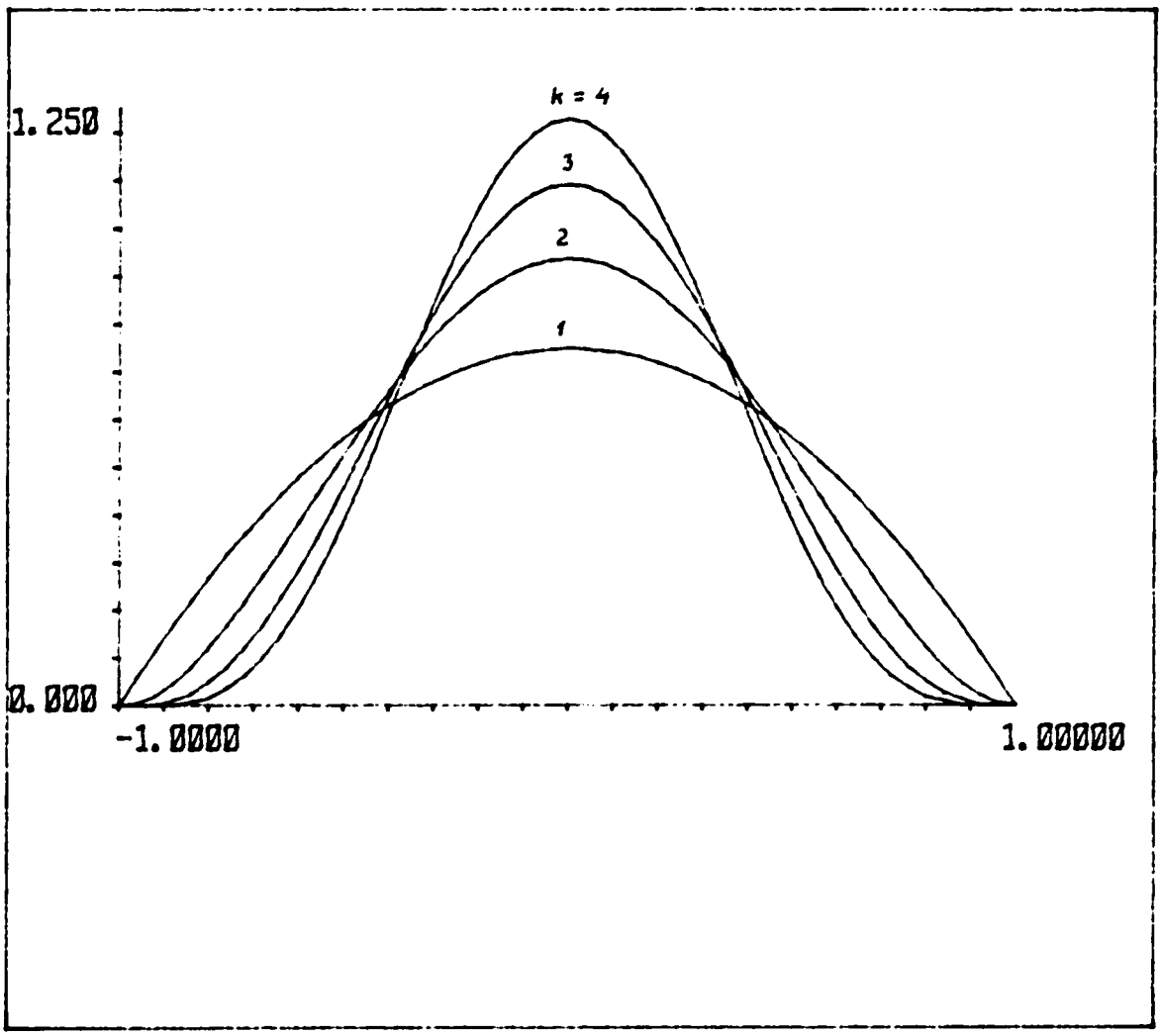


Figure 42. shows the function gk(x) for N = 1 and values of k of 1, 2, 3, and 4.

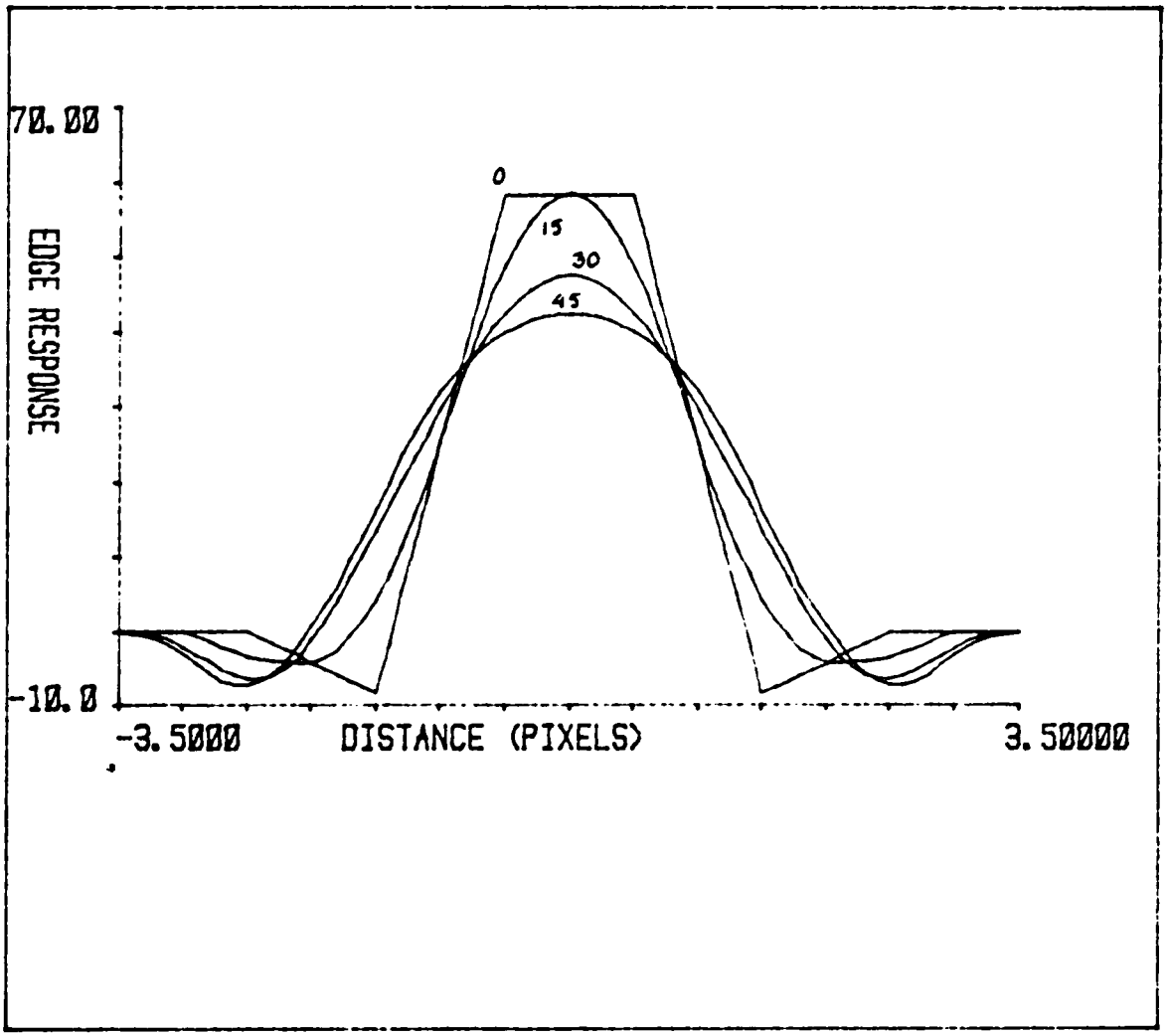


Figure 43. shows the edge response of the standard directional derivative operator accross the edge: Edge orientations are 0, 15, 30, and 45 degrees.

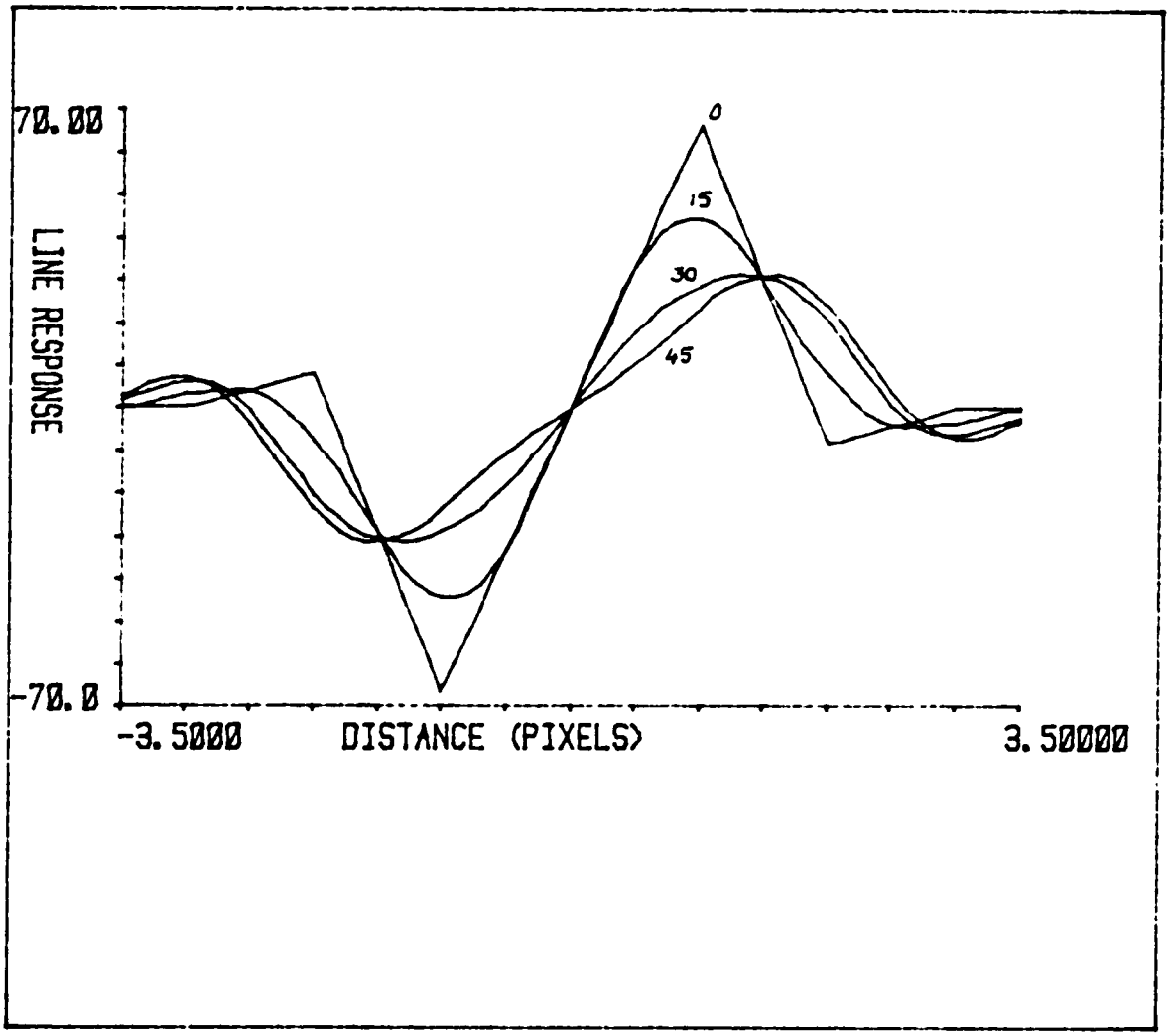


Figure 44. shows unit width line response of the standard directional derivative operator accross the line: Line orientations are 0, 15, 30, and 45 degrees.

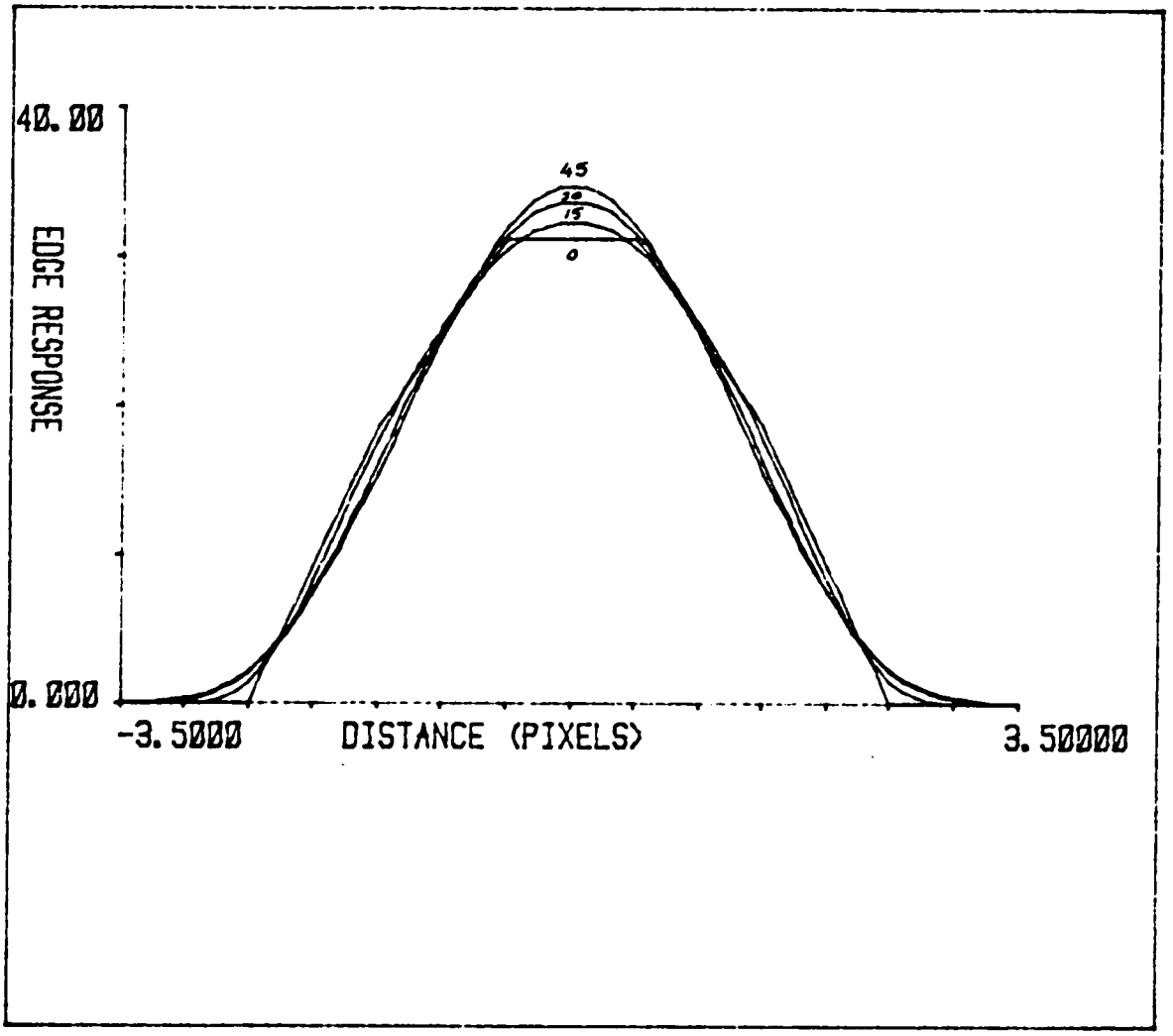


Figure 45. shows the edge response of the integrated directional derivative operator accross the edge: Edge orientations are 0, 15, 30, and 45 degrees.

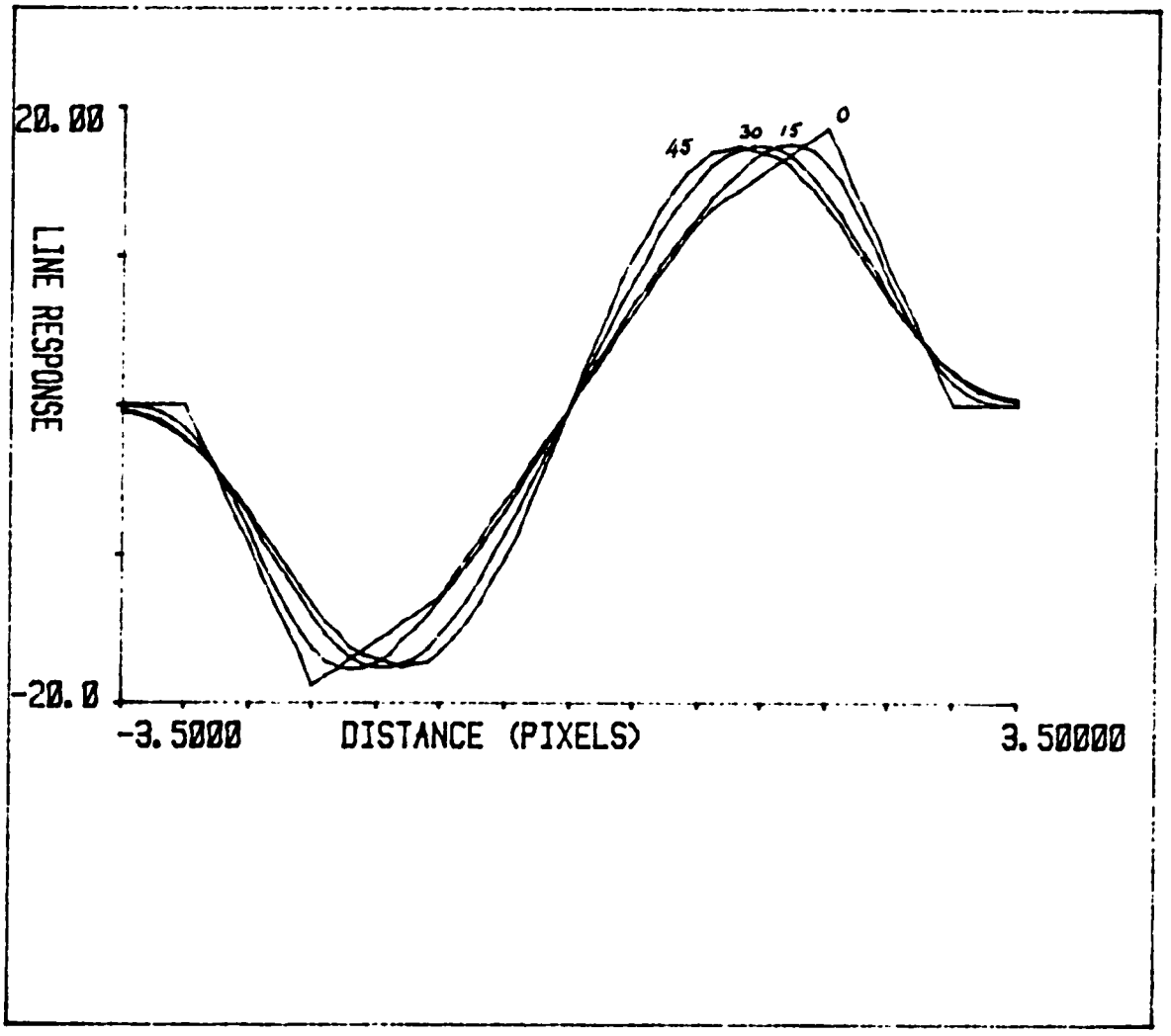


Figure 46. shows the unit width line response of the integrated directional derivative accross the line: Edge orientations are 0, 15, 30, and 45 degrees.

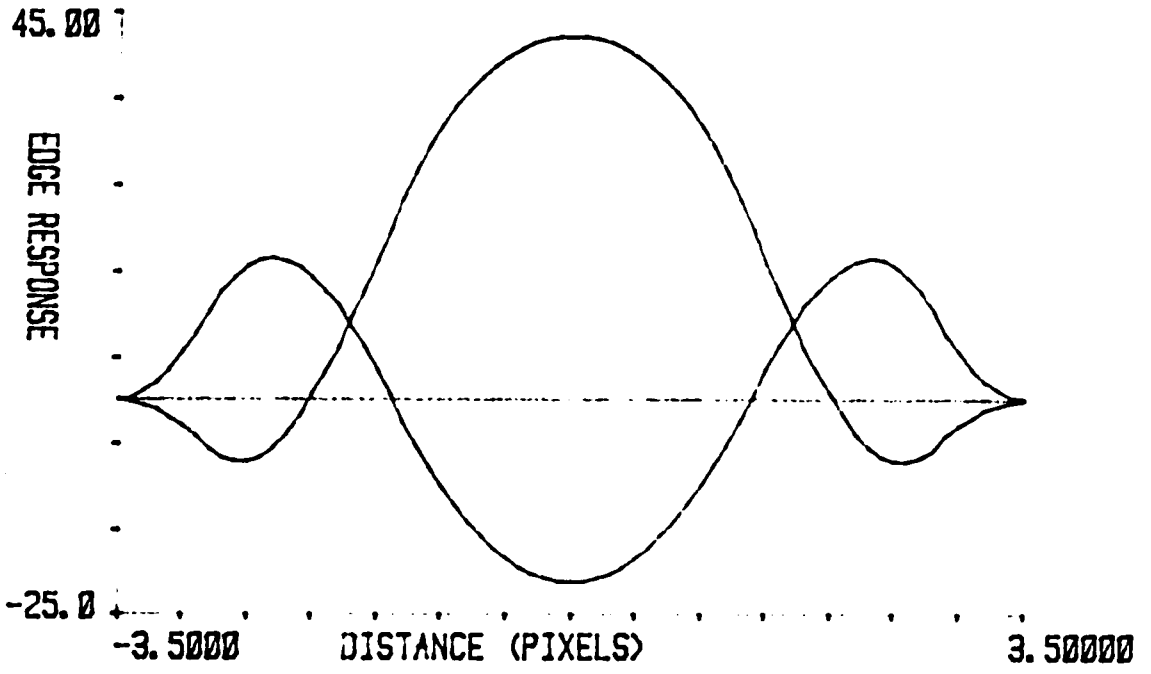


Figure 47. shows the responses of the first directional derivative, and the third directional derivative: Step edge at 45 degrees.

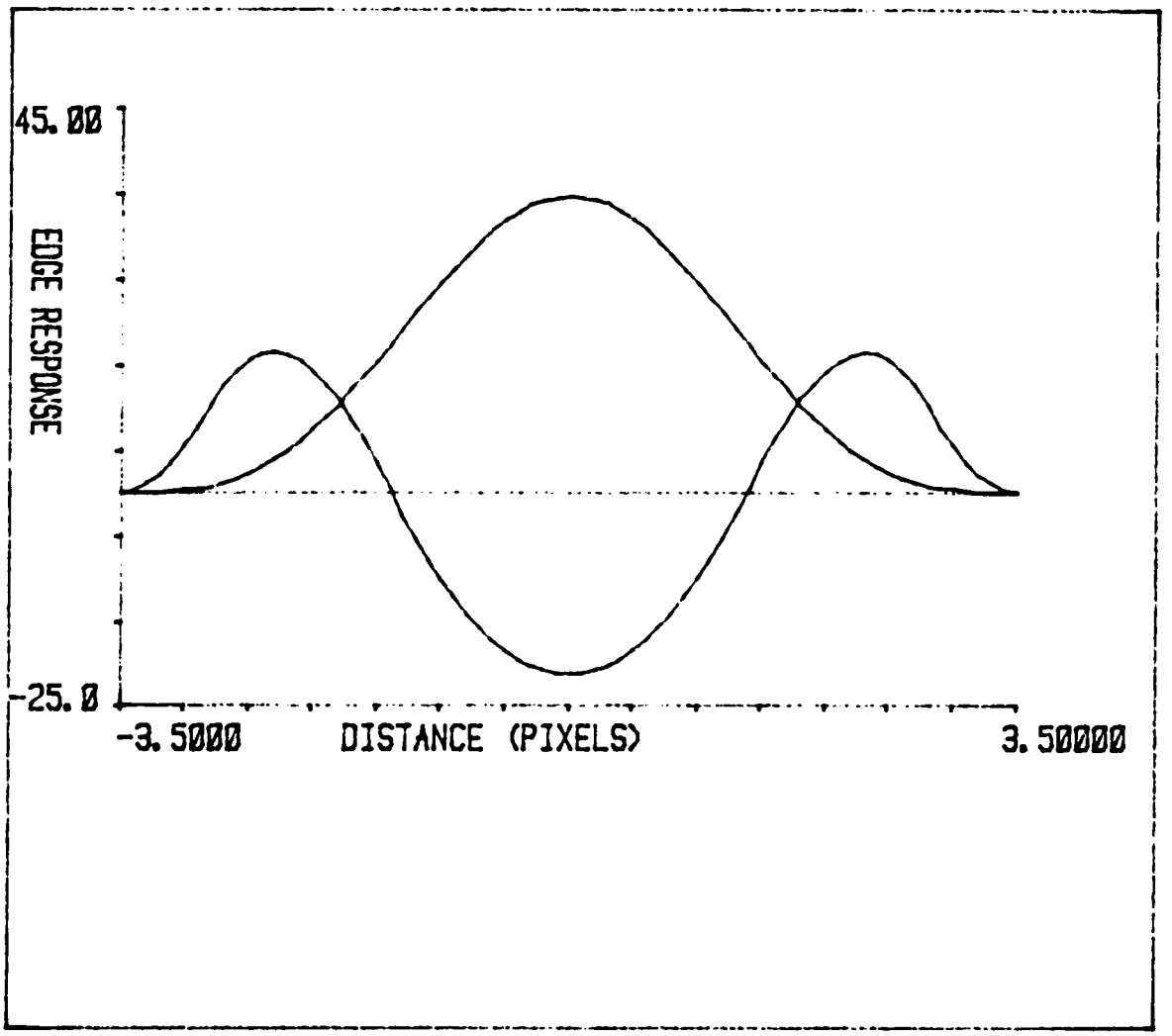


Figure 48. shows the responses of the integrated first directional derivative, and the third directional derivative: Step edge at 45 degrees.

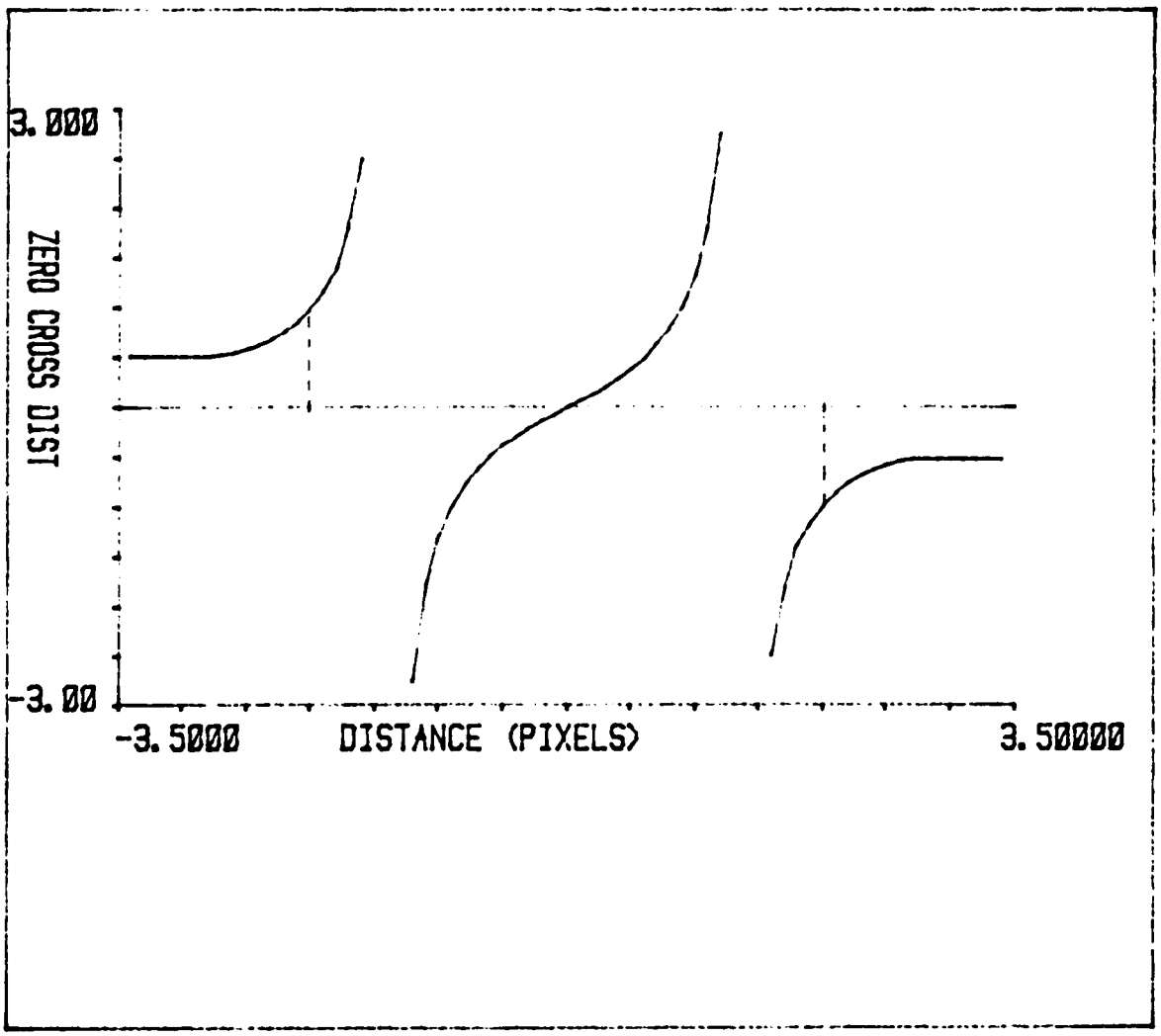


Figure 49. shows zero-crossing displacement from the neighborhood center as a function of edge displacement.

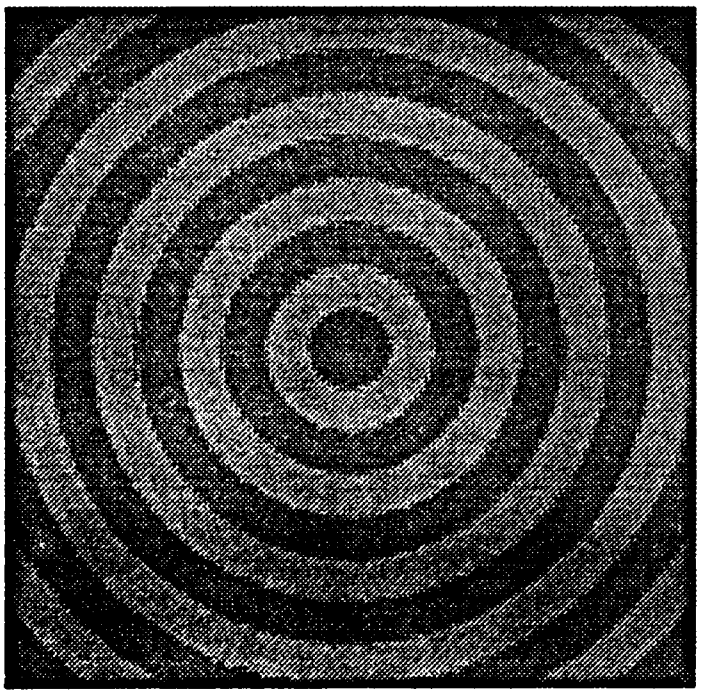


Figure 50. shows the computer generated test image.

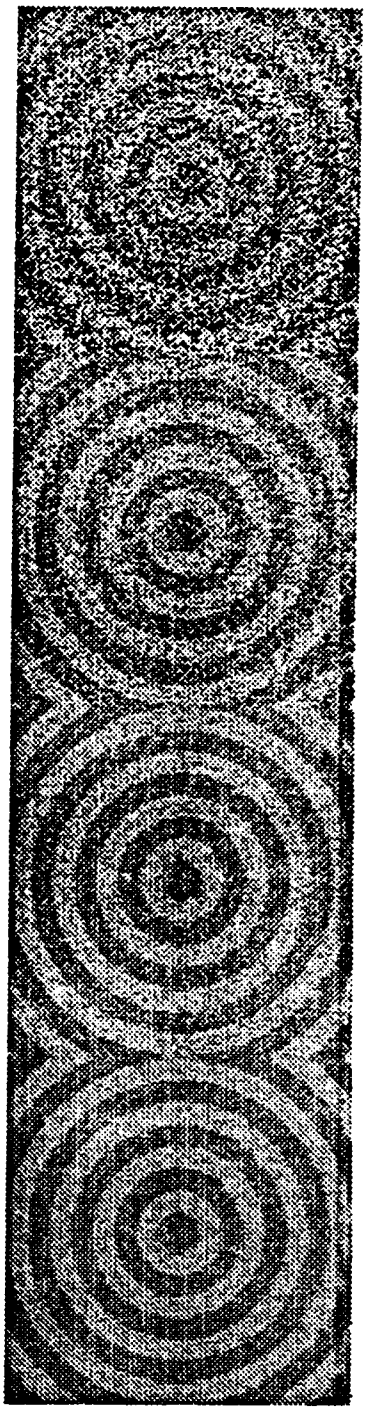


Figure 51. shows the noisy test images: SNR values are 1, 5, 20, and 100.

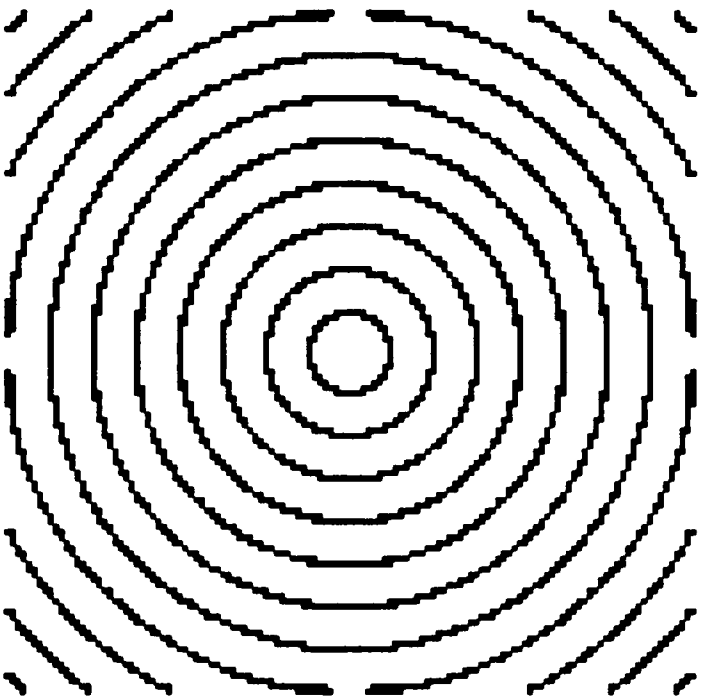


Figure 52. shows the ideal test edge image.

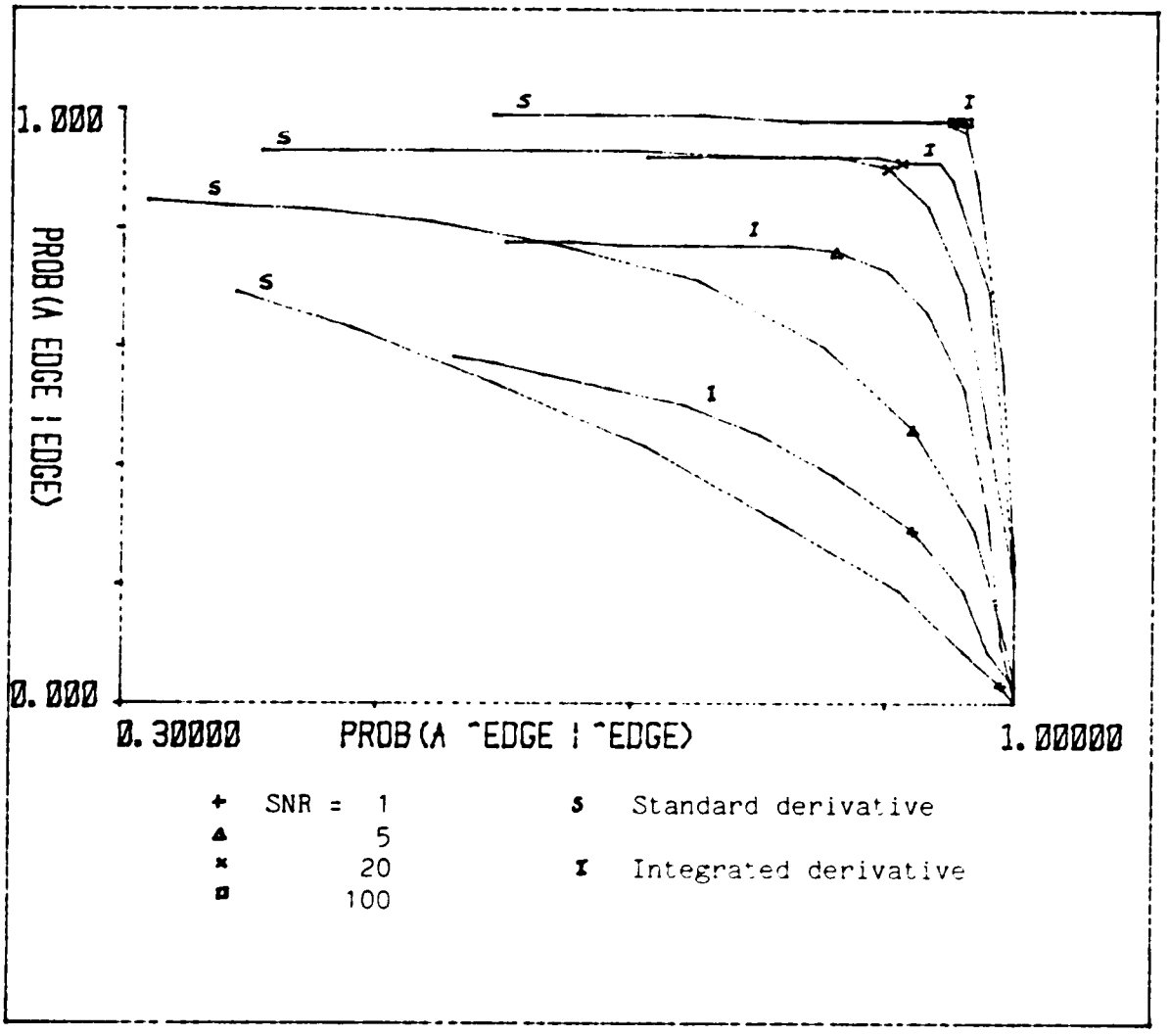


Figure 53. shows curves of constant zero-crossing radius and SNR, and varying gradient threshold for the 5x5 operators.

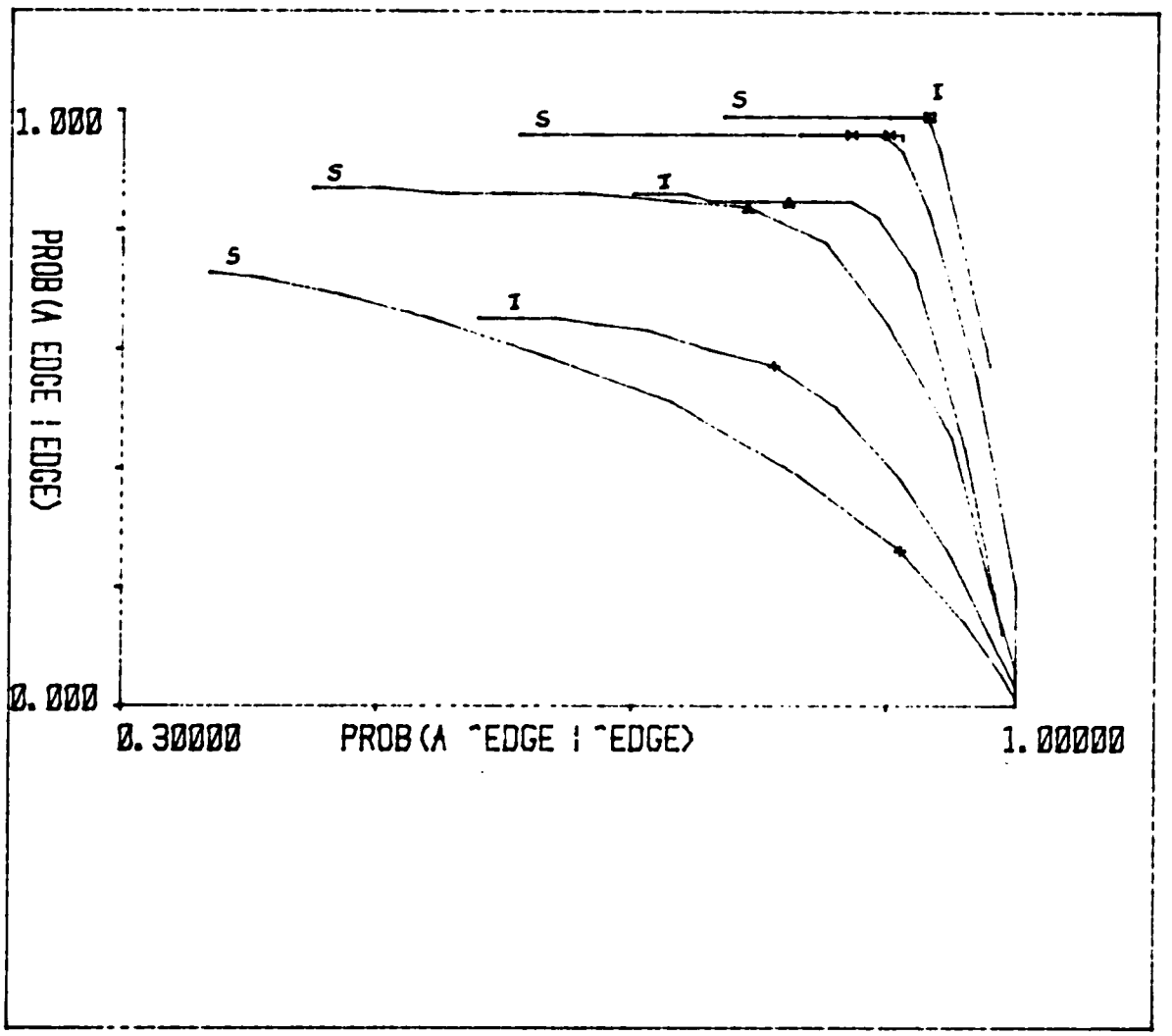


Figure 54. shows curves of constant zero-crossing radius and SNR, and varying gradient threshold for the 7x7 operators.

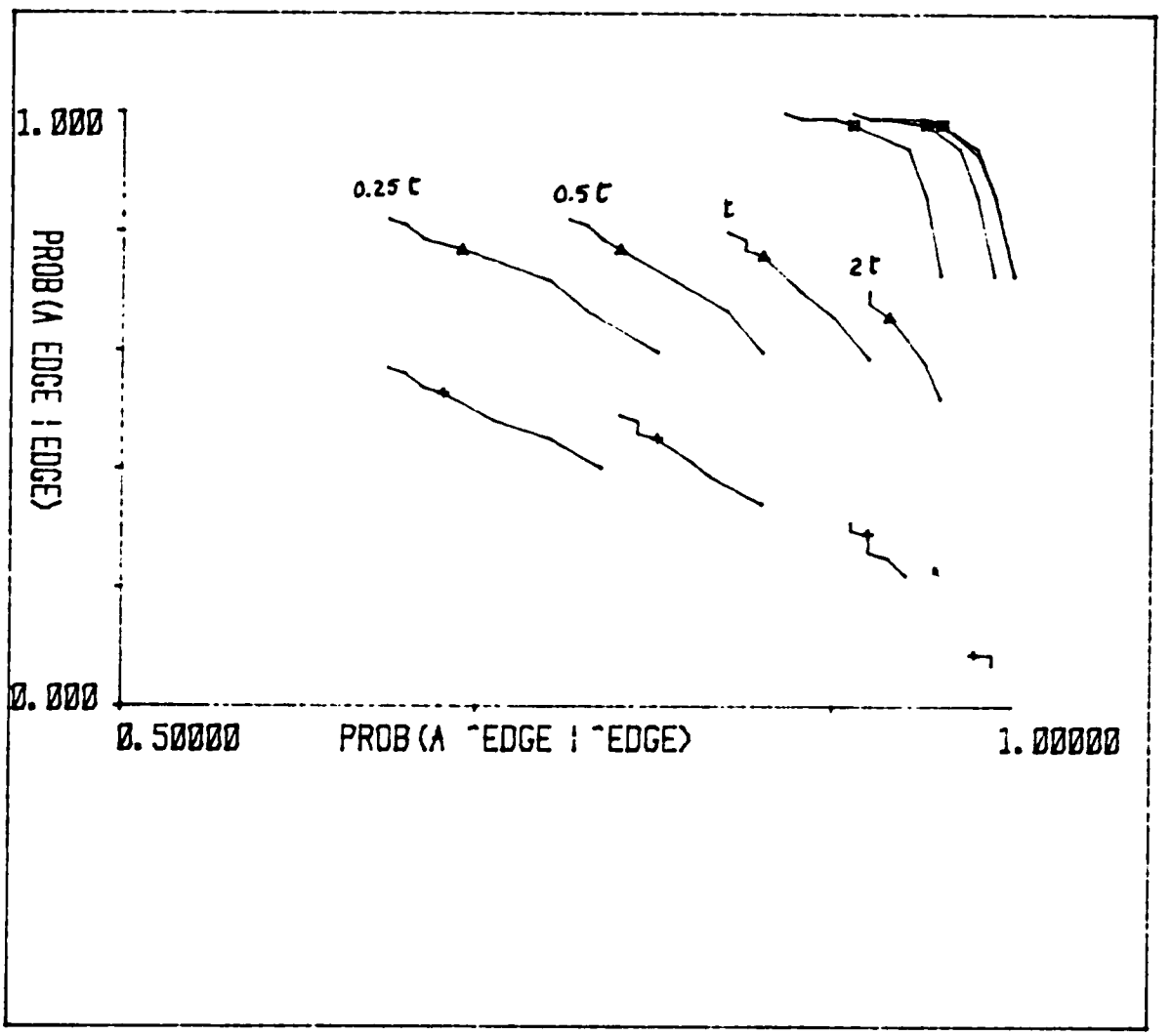


Figure 55. shows curves of constant gradient threshold and SNR, and varying zero-crossing radius for the 5x5 integrated operators.

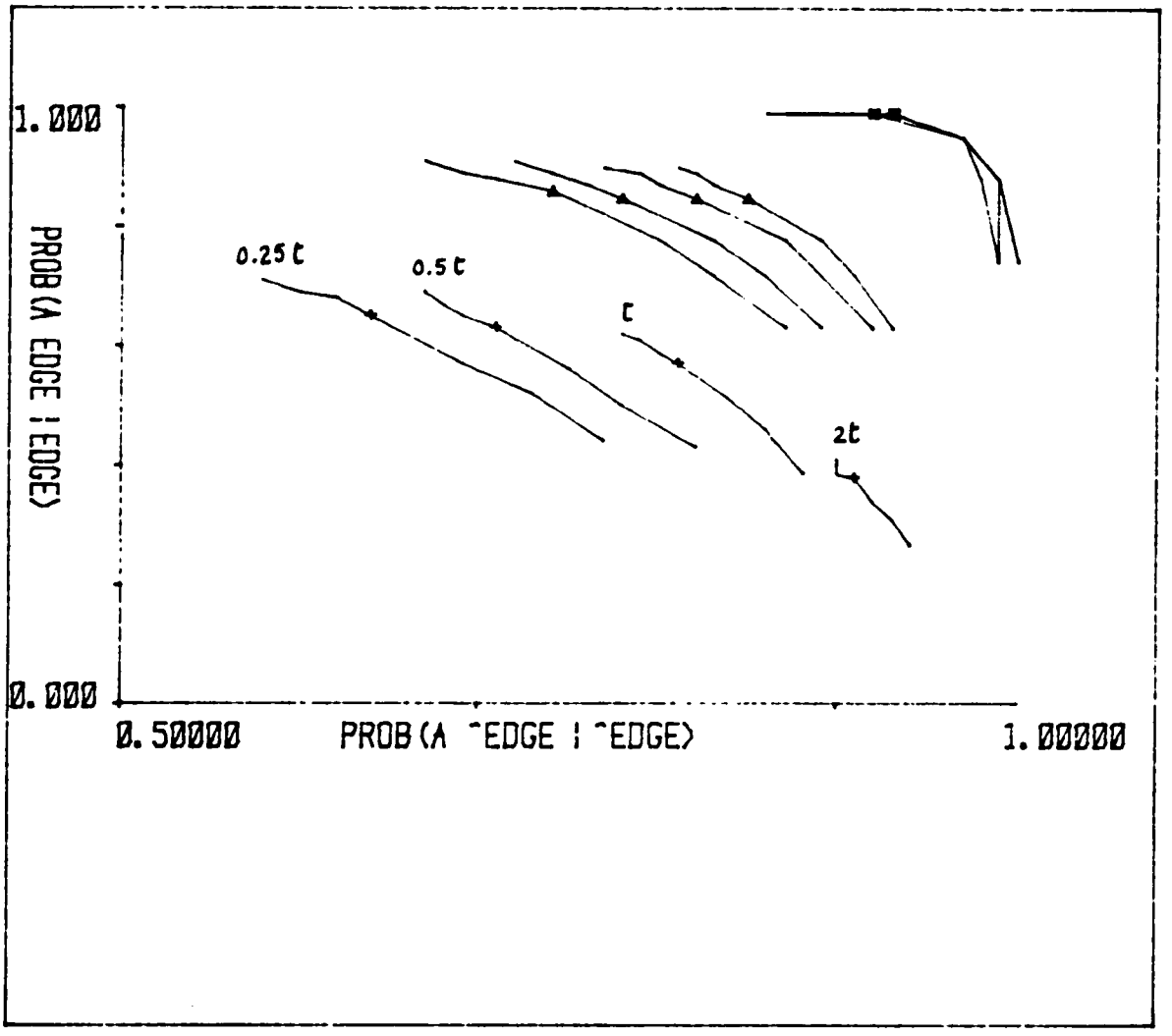


Figure 56. shows curves of constant gradient threshold and SNR, and varying zero-crossing radius for the 7x7 integrated operators.

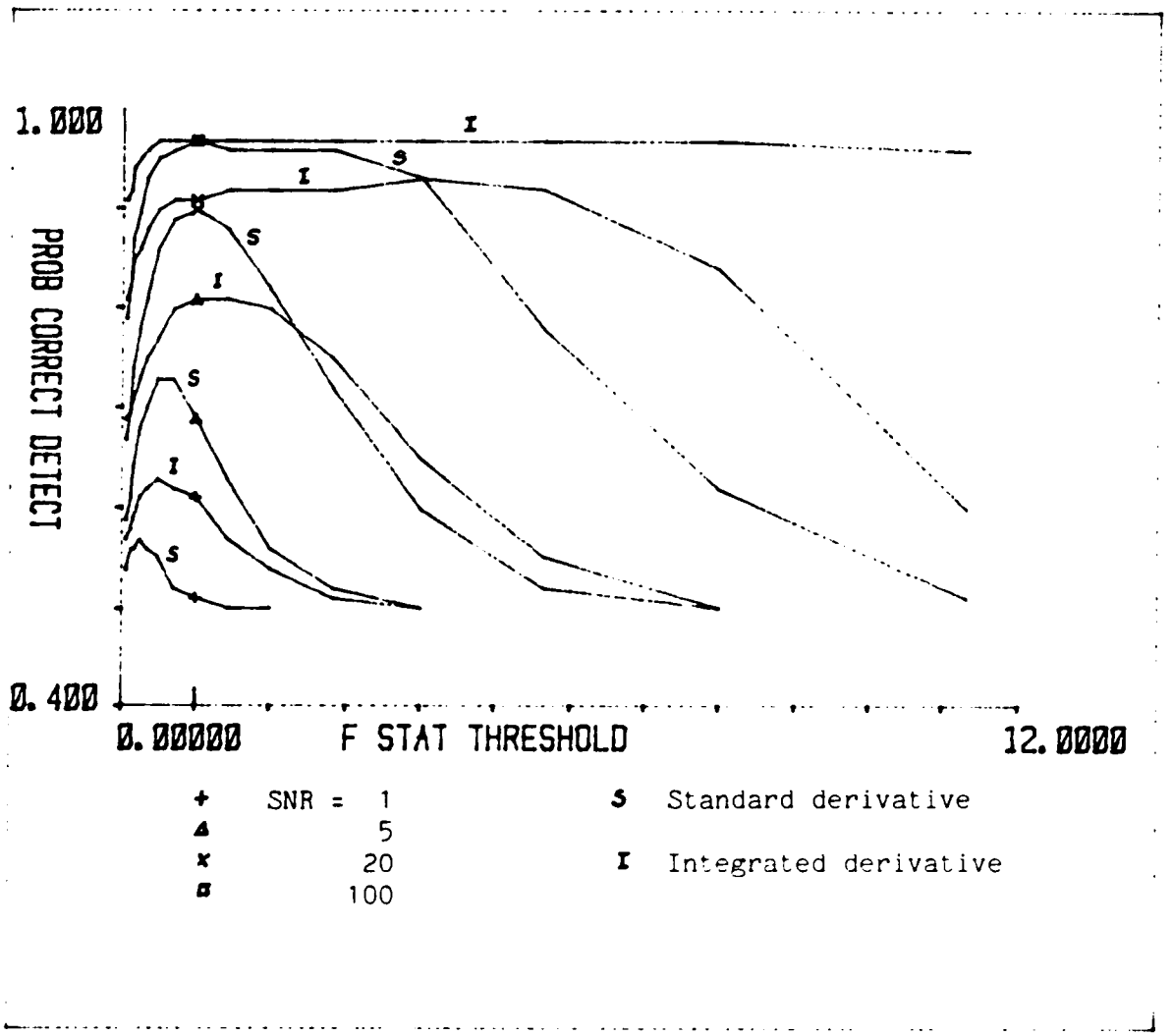


Figure 57. shows probability of correct detection as a function of F-statistic threshold for the 5x5 operators.

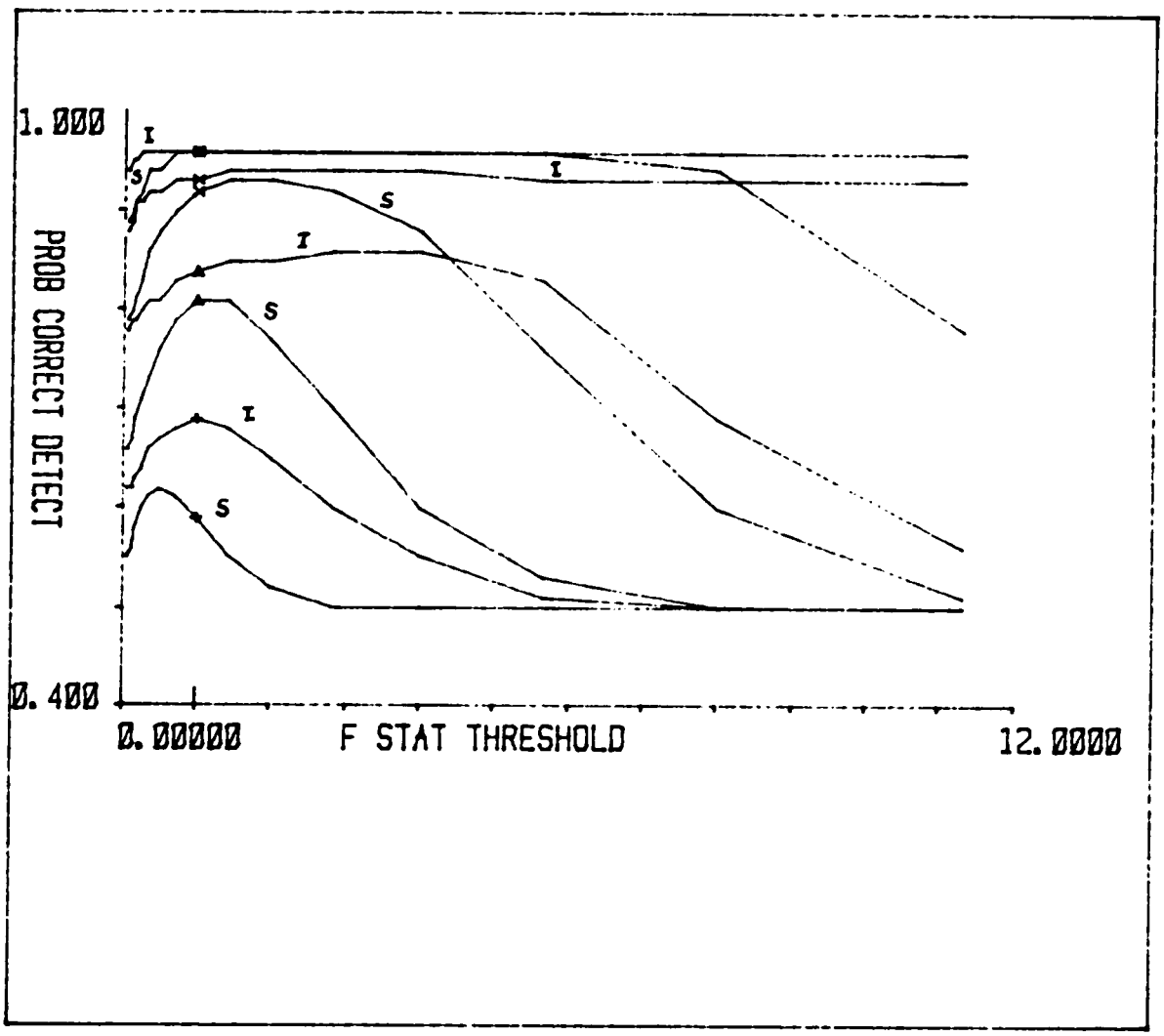


Figure 58. shows probability of correct detection as a function of F-statistic threshold for the 7x7 operators.

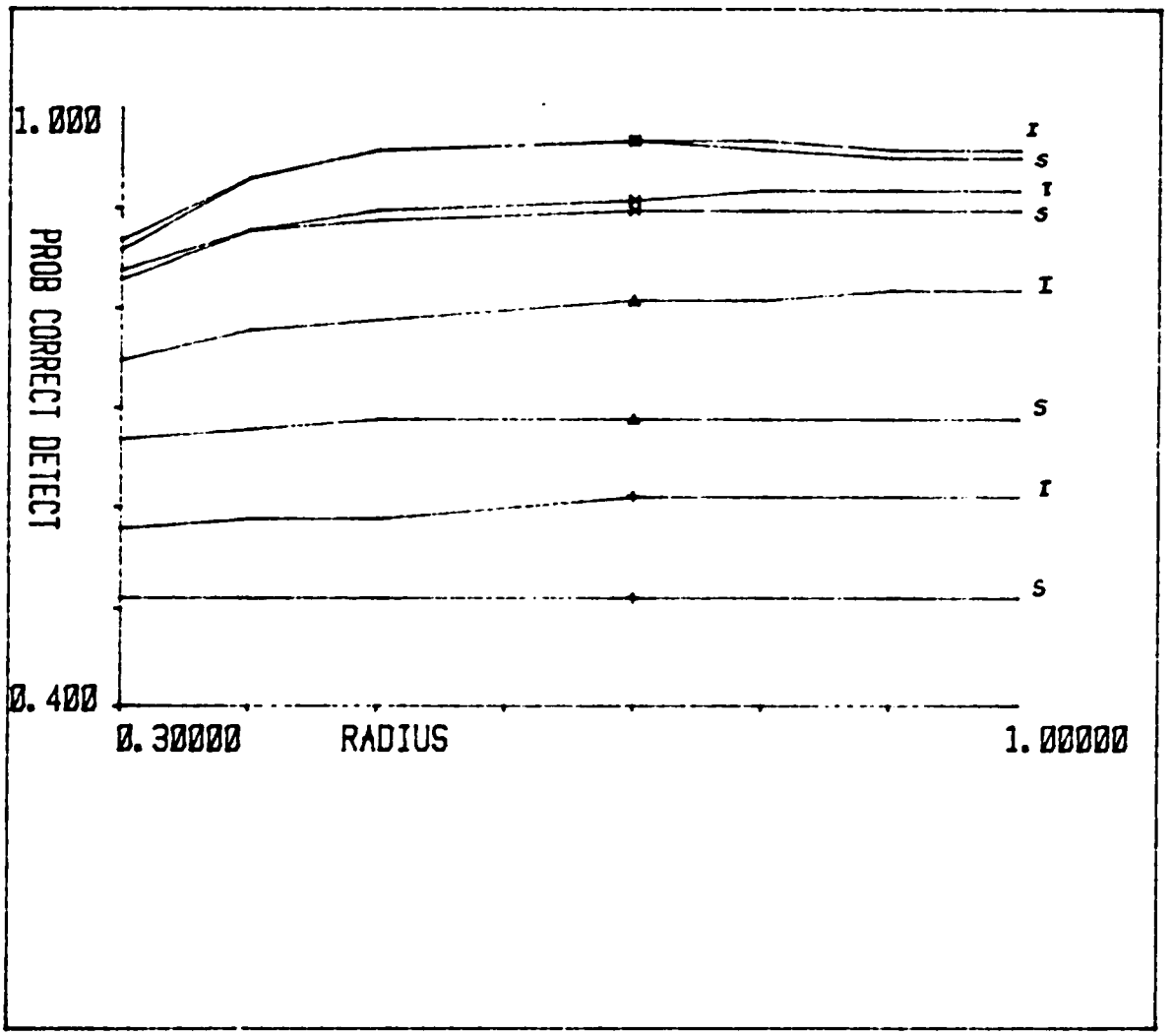


Figure 59. shows probability of correct detection as a function of zero-crossing radius for the 5x5 operators.

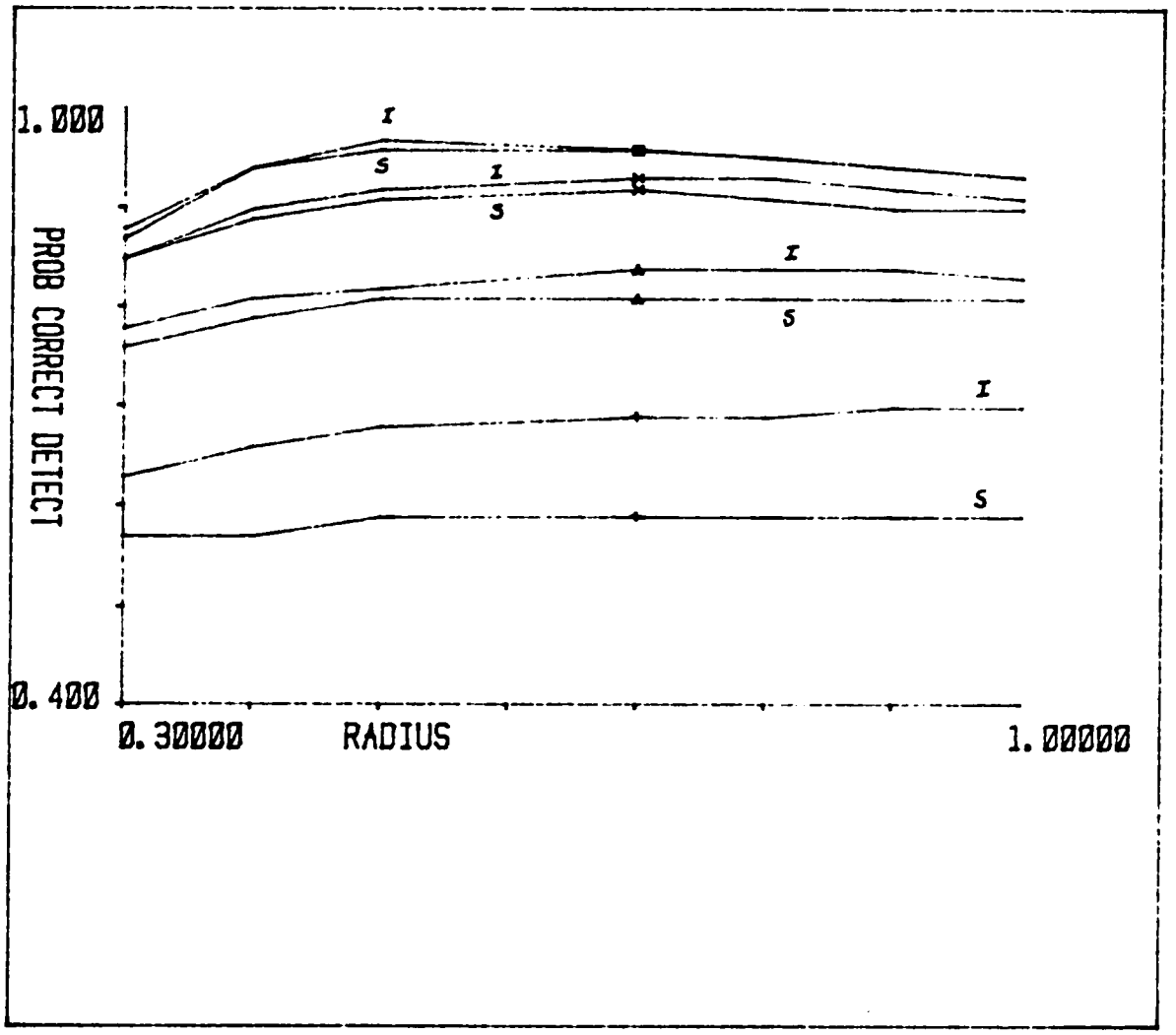


Figure 60. shows probability of correct detection as a function of zero-crossing radius for the 7x7 operators.

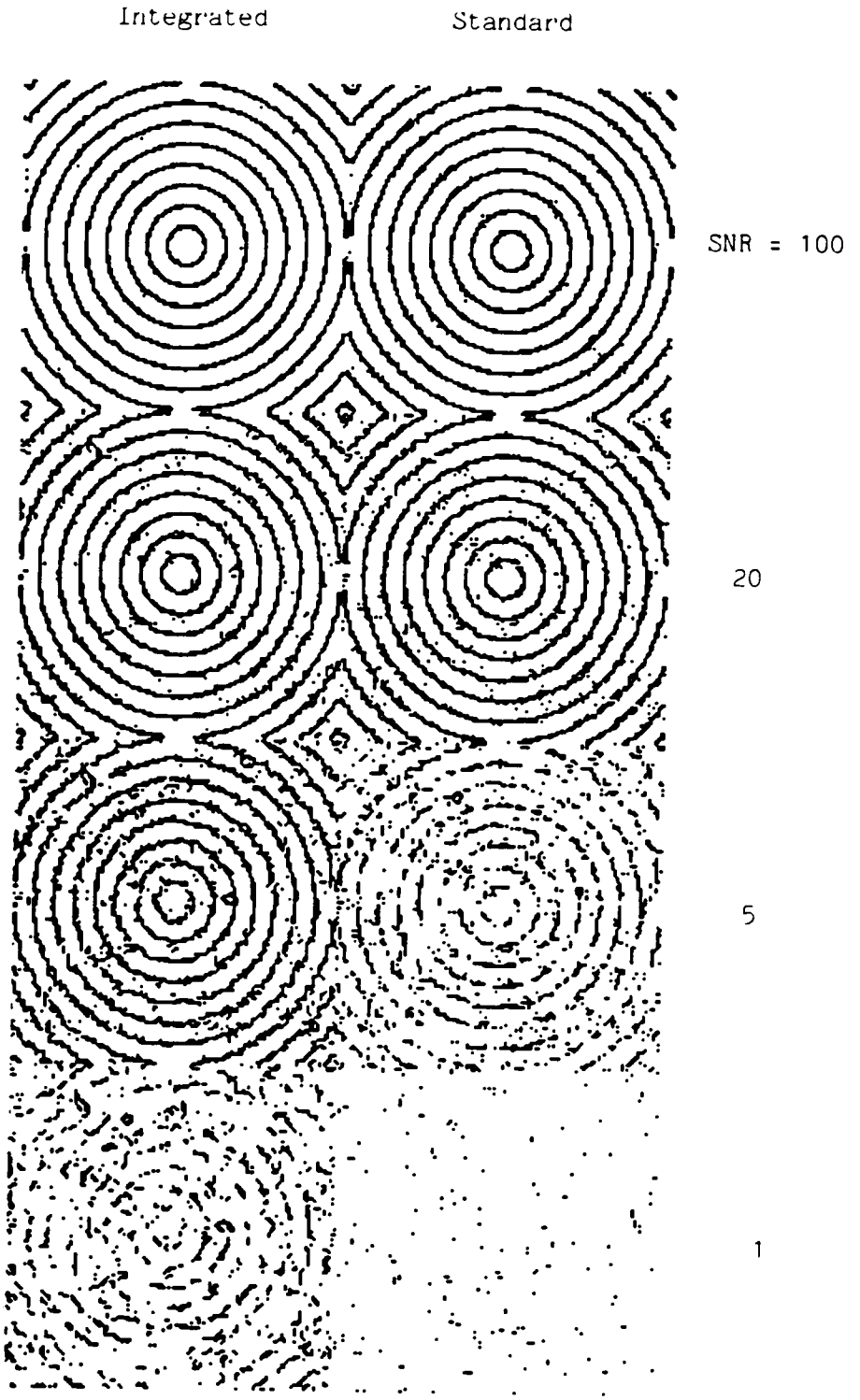


Figure 61. shows the 5x5 operators applied to the noisy test images.

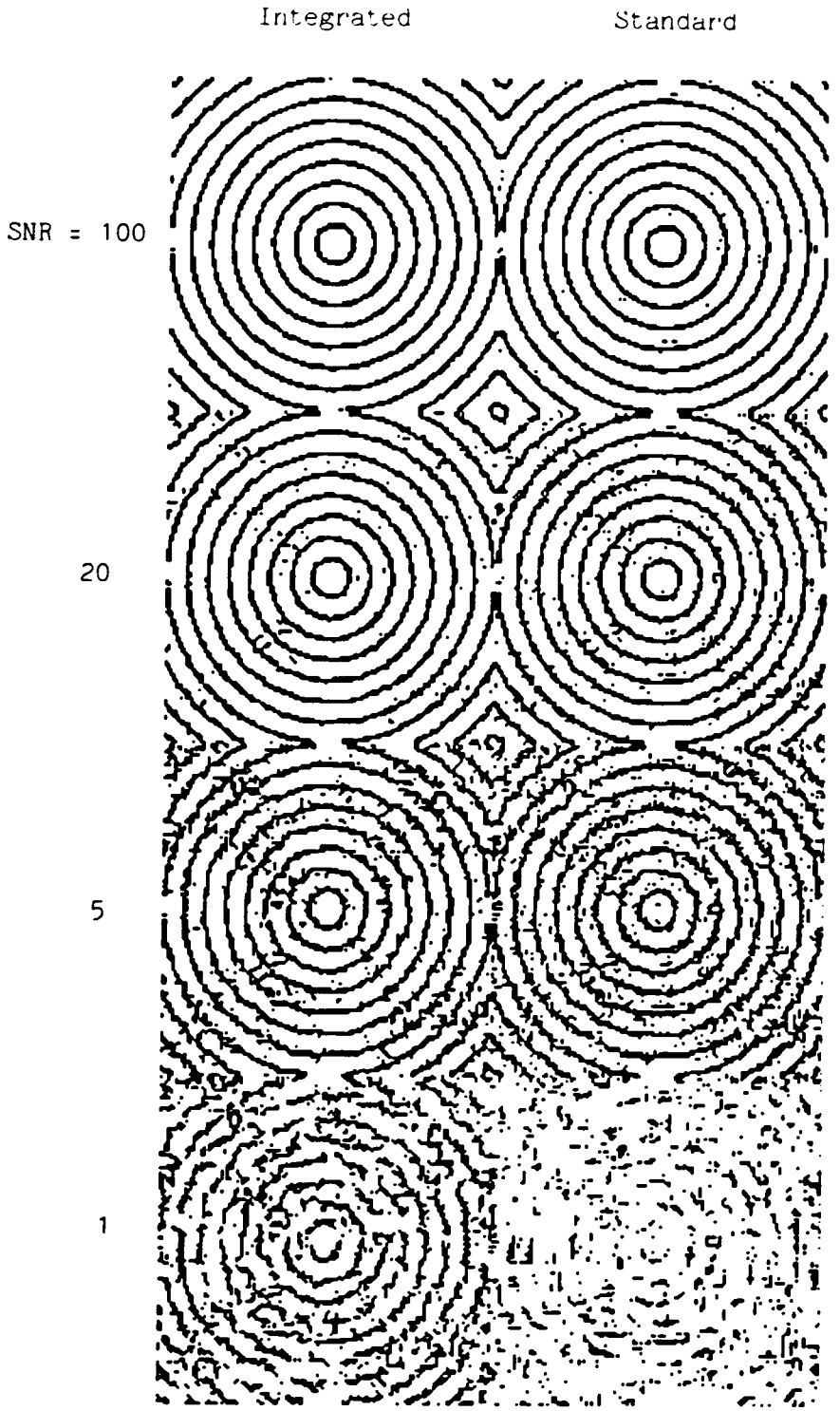


Figure 62. shows the 7x7 operators applied to the noisy test images.

Figure 63. shows a Baton Rouge, Louisiana, urban scene.

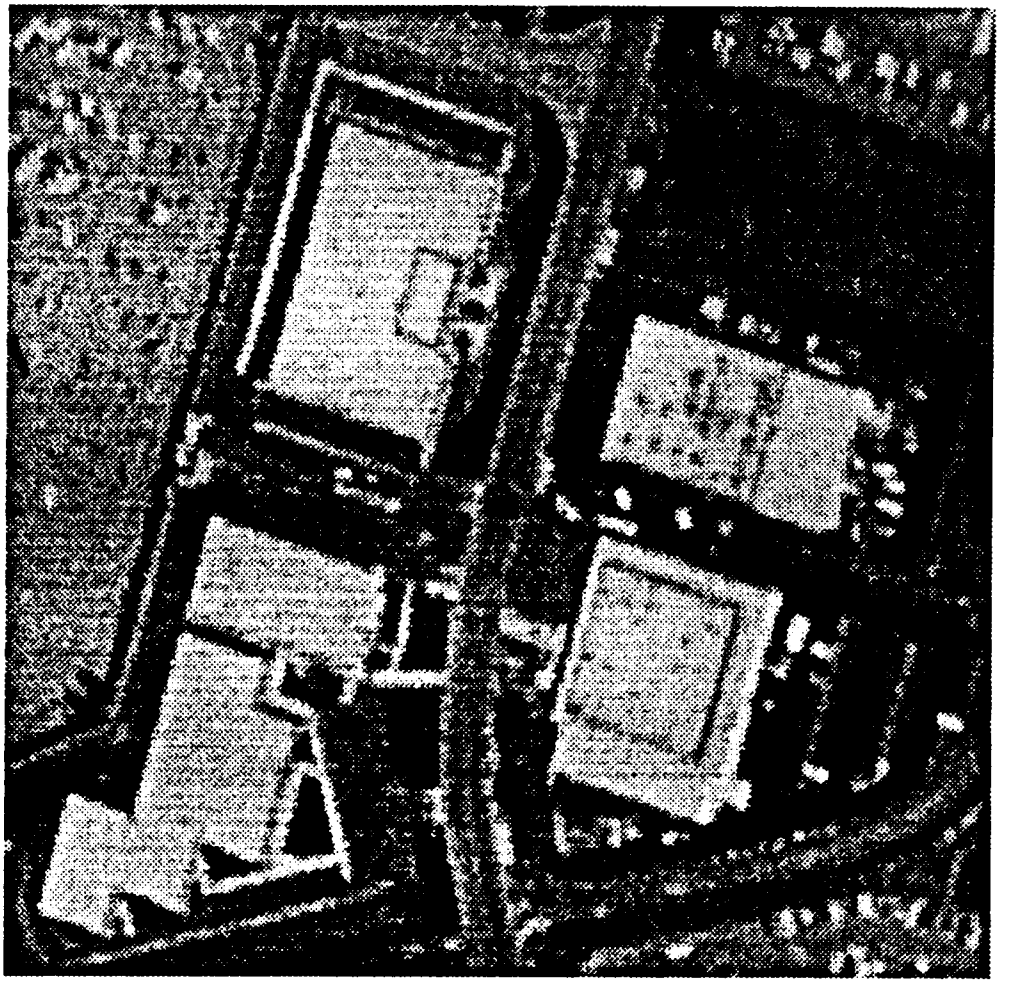


Figure 64. shows a Sunnyvale, California, urban scene.

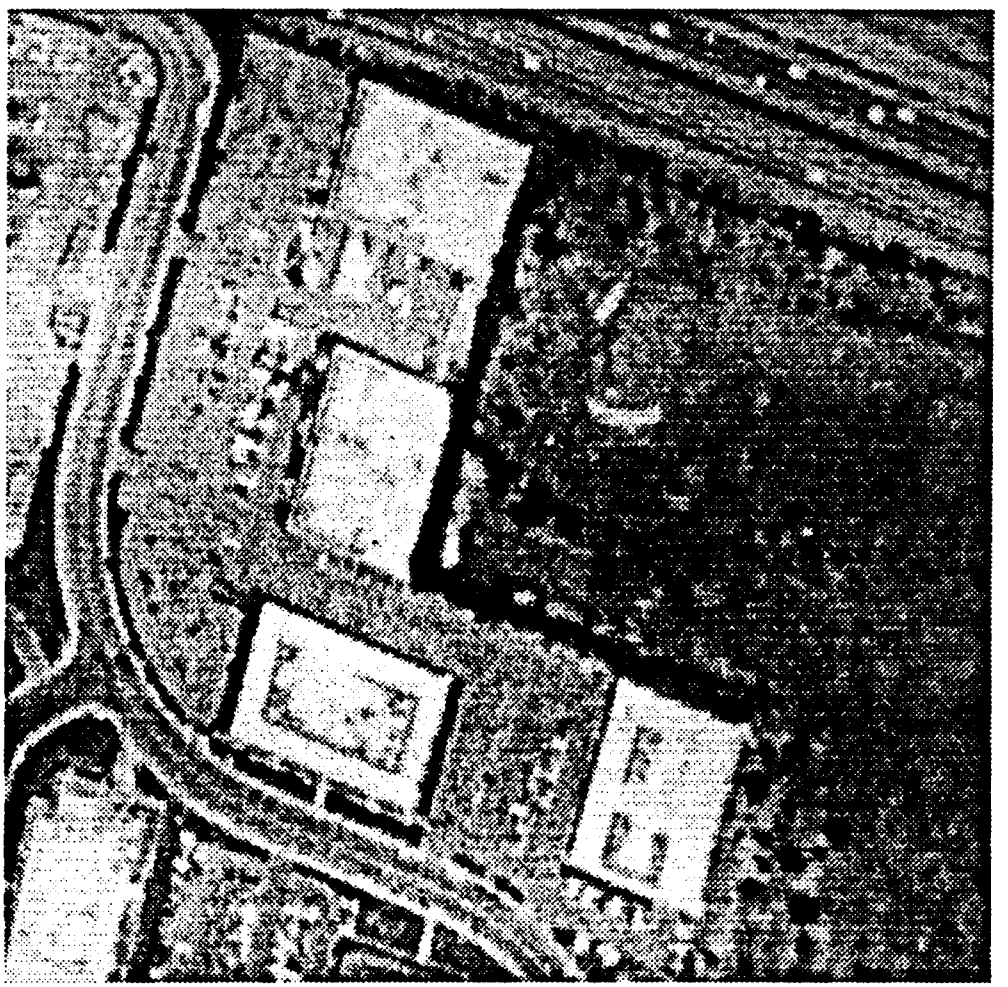


Figure 65. shows a Sunnyvale, California, urban scene.

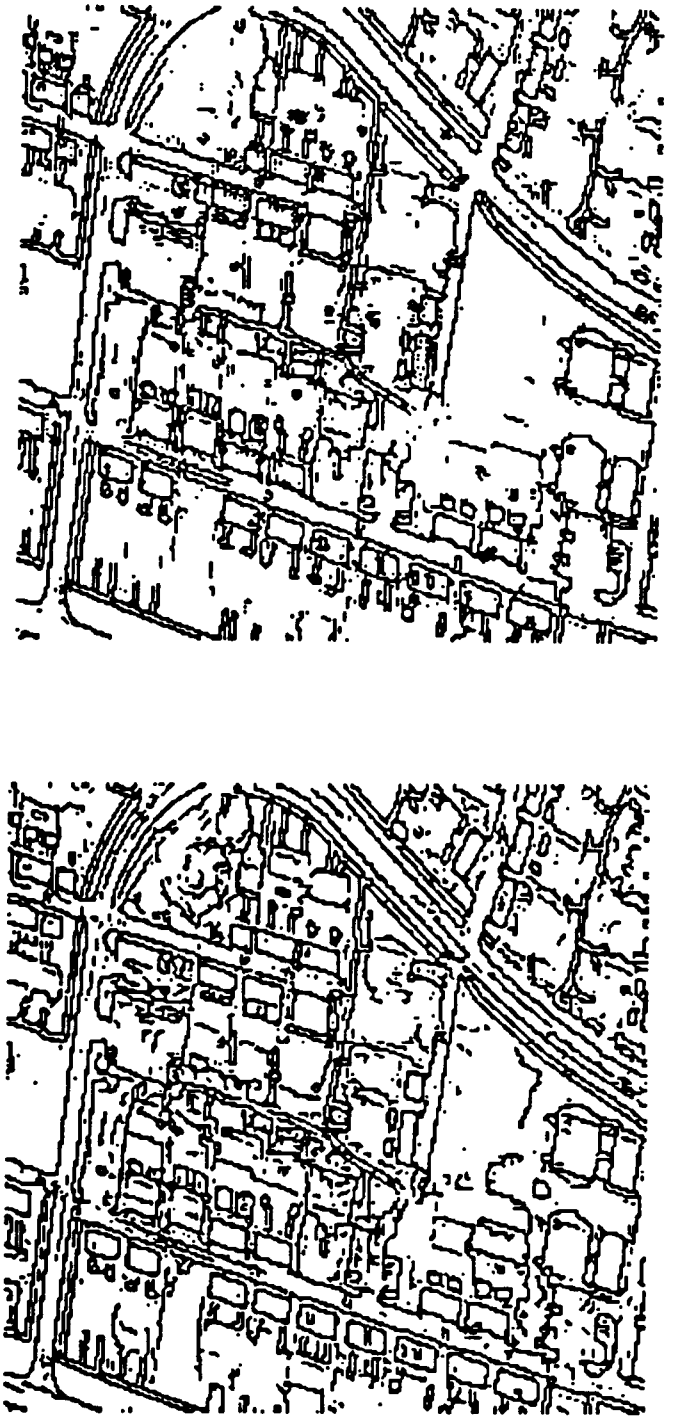


Figure 66. shows the 5x5 operators applied to the Baton Rouge scene: (a) Standard Cubic (b) Integrated Directional Derivative

Figure 67. shows the 7x7 operators applied to the Baton Rouge scene: (a) Standard Cubic (b) Integrated Directional Derivative

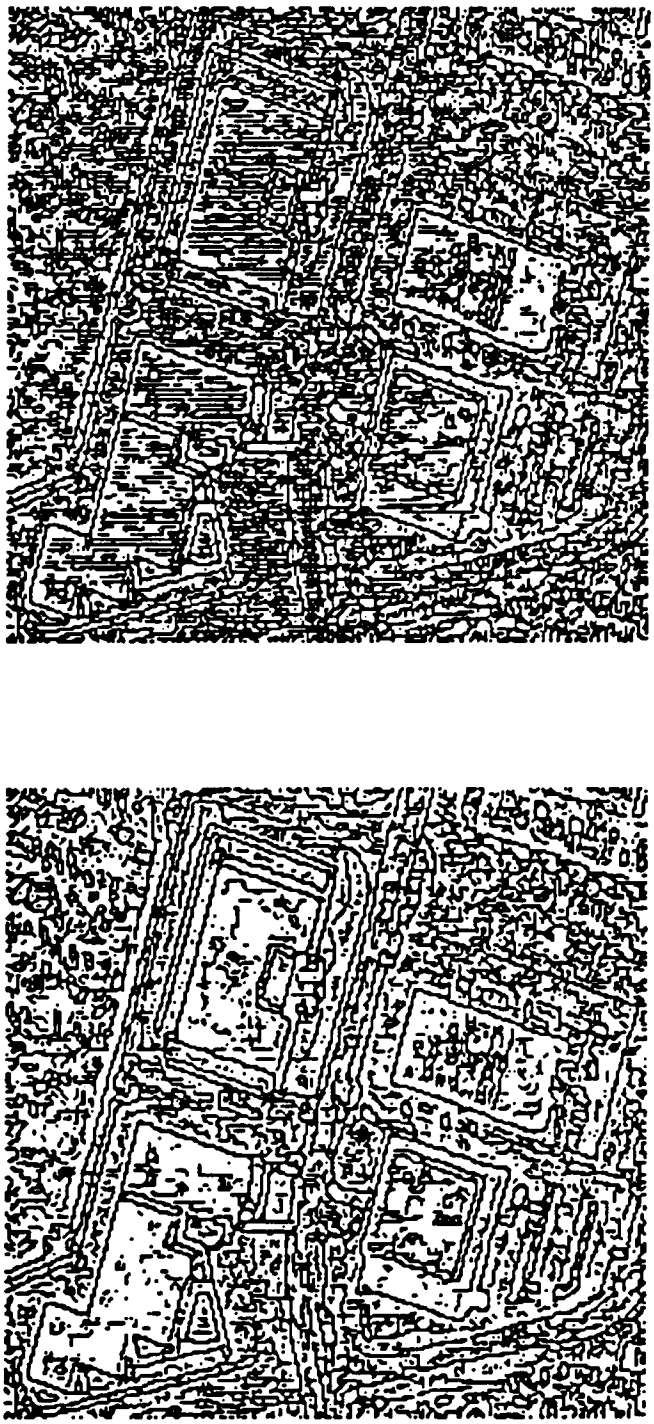


Figure 68. shows the 5x5 operators applied to the Sunnyvale scene: (a) Standard Cubic (b) Integrated Directional Derivative

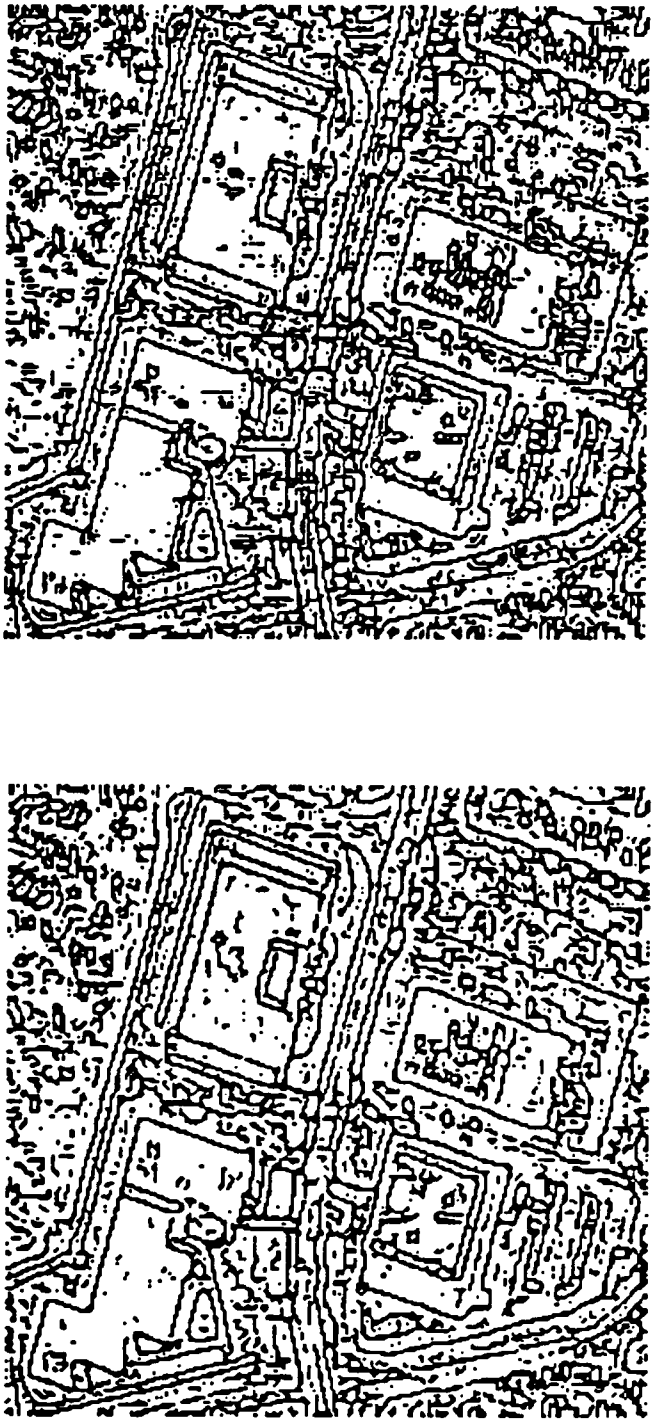


Figure 69. shows the 7x7 operators applied to the Sunnyvale scene: (a) Standard Cubic (b) Integrated Directional Derivative

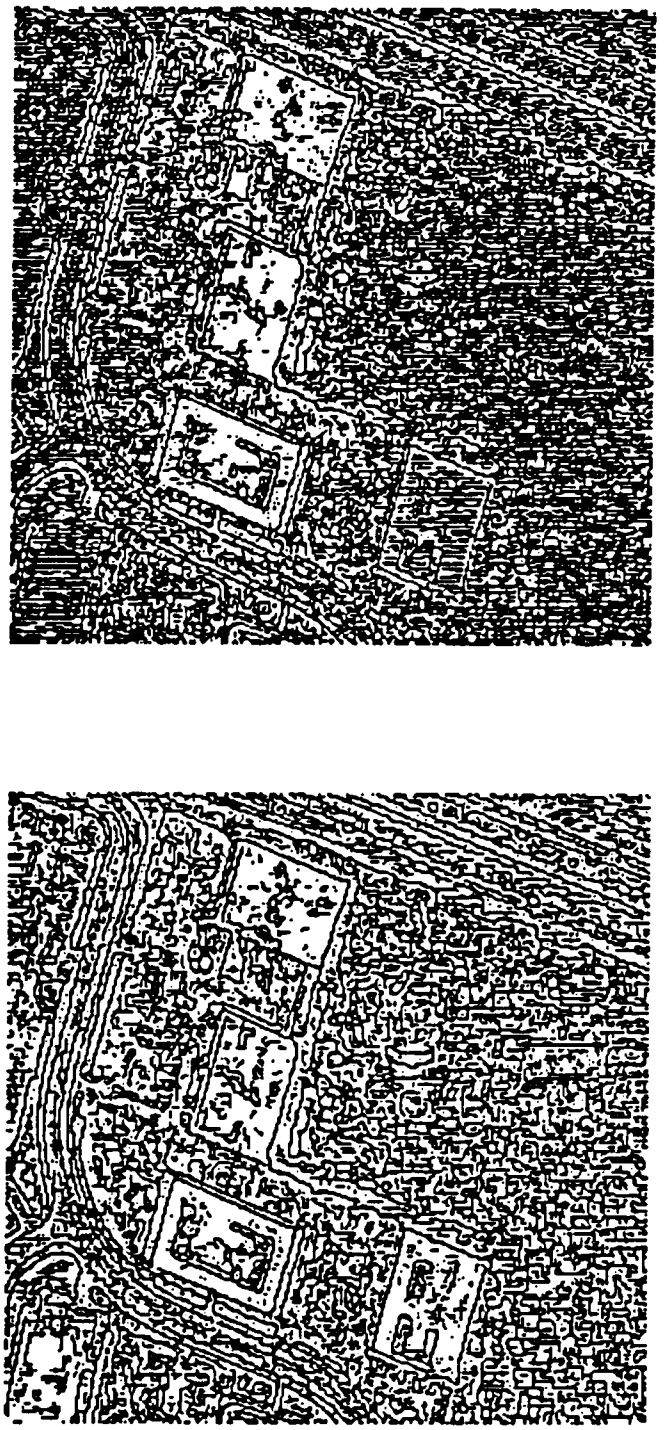


Figure 70. shows the 5x5 operators applied to the Sunnyvale scene: (a) Standard Cubic (b) Integrated Directional Derivative

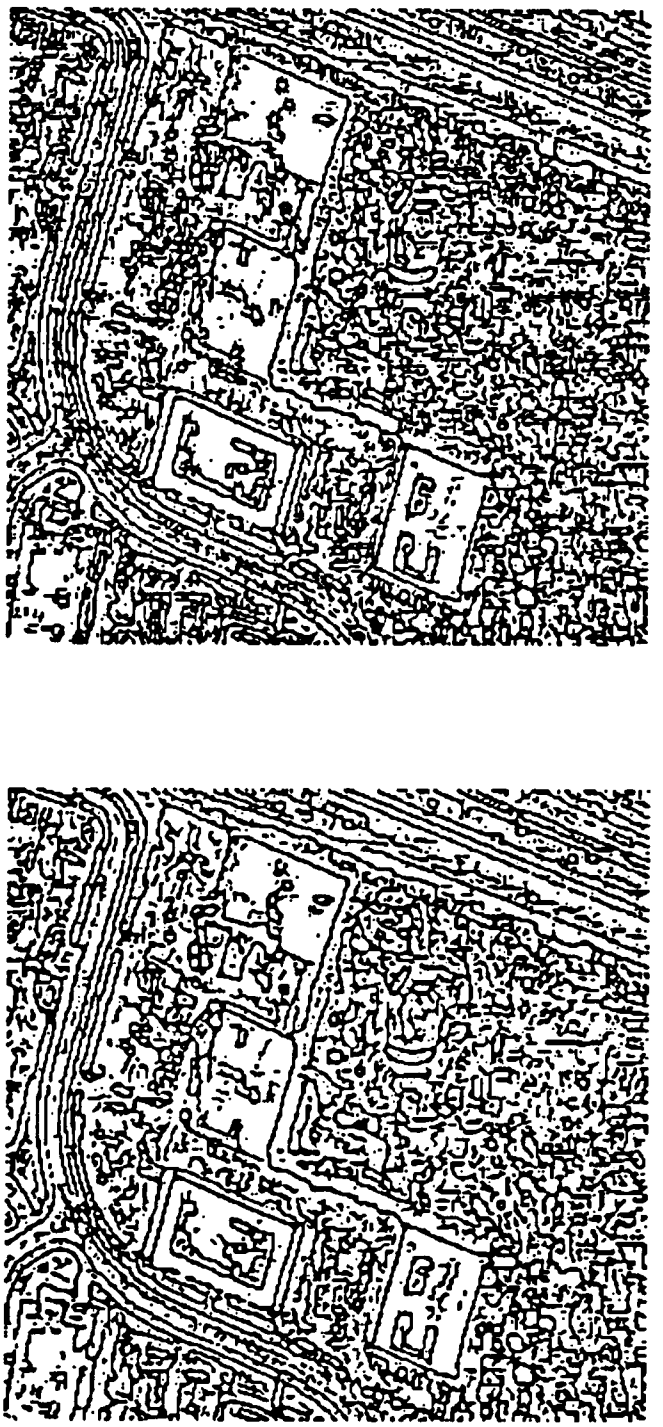


Figure 71. shows the 7x7 operators applied to the Sunnyvale scene: (a) Standard Cubic (b) Integrated Directional Derivative

Chapter 6

CORNER DETECTION

6.1 Introduction

The detection of corners in images has been shown to be extremely useful for computer vision tasks. Huertas (1981) uses corners to detect buildings in aerial images. Nagel and Enkelmann (1982) use corner points to determine displacement vectors from a pair of consecutive images taken in time sequence. Much of the past research in corner detection has relied on prior segmentation of the image and subsequent analysis of region boundaries. Rutkowski and Rosenfeld (1978) provide a comparison of several corner detection techniques along those lines.

More recent research has focused on developing "gray tone corner detectors" which detect corners by operating directly on the gray tone image. As Kitchen and Rosenfeld (1980) point out, the main advantage of such corner detectors is that their performance is not dependent on the success or failure of a prior segmentation step. Among the

earliest such corner detectors is Beaudet's DET operator (1978) which responds significantly near corner and saddle points and Moravec's Interest Operator (1977) which returns regions which are local maxima of a directional variance measure. Kitchen and Rosenfeld report results using several operators which measure cornerness by the product of gradient magnitude and rate of change of gradient direction. Dreschler and Nagel (1981) investigate points lying between extrema of Gaussian curvature as suitable candidates for corner points.

Our approach to corner detection is based on the facet model for digital images (Haralick (1980), Haralick and Watson (1981)). The basic philosophy of this model derives from recognizing that the discrete set of values which form the digital image are the result of sampling and quantizing a real-valued function f defined on the domain of the image which is a bounded and connected subset of the real plane. Thus, any property associated with a pixel or a neighborhood of pixel values should be evaluated by relating it to the property of the corresponding gray tone surface f which underlies the neighborhood. This involves estimating the surface function f locally, from the neighborhood samples available to us. The most natural way of accomplishing this is by assumming a parametric form for f and then estimating its associated parameters. In this dissertation we have chosen to approximate the gray tone surface f by a cubic polynomial in the row and column coordinates of the neighborhood array. More precisely for each neighborhood, f is assumed to be of the form:

$$\begin{aligned} f(r,c) = {} & K_1 \\ & + K_2 r + K_3 c \\ & + K_4 r^2 + K_5 rc + K_6 c^2 \\ & + K_7 r^3 + K_8 r^2 c + K_9 rc^2 + K_{10} c^3 \end{aligned} \tag{6.1}$$

We investigate that property of a pixel which we will call "cornerness". Suppose for instance that we are shown an aerial photograph of a city and are asked to identify corners in it. Our attention would be likely drawn first to objects such as buildings where the concept of corners expresses itself very clearly in a natural way. Two walls of a building meet usually at 90 degrees and we would declare the intersection point to be a corner. But there are other places where we would see corners as well, where the shadows due to two adjacent walls meet, where two roads intersect, in a patch of field of a certain shape, etc. In general what we are usually inclined to call a corner occurs where two edge boundaries meet at a certain angle or where the direction of an edge boundary is changing very rapidly. We associate corners therefore with two things: the occurence of an edge and significant changes in edge direction.

These two concepts have a very straight-forward and clear meaning under the facet model. Edges under this model have been investigated by Haralick (1980). In particular a step edge operator based on zero-crossing of second directional derivatives has been developed by Haralick (1982) with very encouraging results even in such difficult imagery as aerial scenes. Edge direction is most naturally expressed as a direction orthogonal to the gradient direction at the point of occurrence of an edge.

Section 6.2 describes the essentials of Haralick's zero-crossing of second directional derivative step edge detector which we will use as part of the corner detection process. In section 6.3 we describe several corner detectors based on the facet model and in section 6.4 we present our experimental results and comparison against the Kitchen-Rosenfeld and Dreschler-Nagel corner detectors.

6.2 *Zero-Crossing of Second Directional Derivative Edge Detection.*

Under the facet model step edges occur at relative extrema in first directional derivative of the continuous gray tone function f underlying a given neighborhood of pixel values. Relative extrema in first directional derivative can reveal themselves as zero-crossings of the second directional derivative. More precisely, a pixel is marked as an edge pixel if in the pixel's immediate area there is a negatively sloped zero crossing of the second directional derivative taken in the direction of the gradient (Haralick (1982)). It has been shown that this kind of edge detector can respond to weak but spatially peaked gradients.

In the discussion that follows we consider a local row and column coordinate system centered at each pixel neighborhood with the positive row axis pointing down and the positive column axis pointing to the right. Thus for a $(2N+1)$ by $(2N+1)$ neighborhood the row and column index set is $[\ -N, -(N-1),..., -1, 0, 1,..., N-1, N\ ]$.

Given a graytone intensity function f defined in the row and column coordinate system of a given pixel neighborhood, the gradient vector function ∇f is given by

$$\nabla f = \left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial c}\right) \tag{6.2}$$

Consider the direction vector $(\sin\theta, \cos\theta)$, which makes an angle θ measured clockwise with the column axis. It is well known that the directional derivative $f'_\theta(r, c)$ of f in the direction θ can be evaluated as the component of the gradient ∇f along the direction vector, that is

$$f'_\theta(r,c) = \frac{\partial f}{\partial r}\sin\theta + \frac{\partial f}{\partial c}\cos\theta \tag{6.3}$$

Likewise, the second directional derivative $f''_\theta(r, c)$ of f in the direction θ can be readily evaluated as:

$$f''_\theta(r, c) = \frac{\partial^2 f}{\partial r^2}\sin^2\theta + \frac{2\partial^2 f}{\partial r \partial c}\sin\theta\cos\theta + \frac{\partial^2 f}{\partial c^2}\cos^2\theta \tag{6.4}$$

Let us consider points (r,c) on a line passing through the point (r_0,c_0) in the direction θ. Then:

$$r = r_0 + \rho\sin\theta, \quad c = c_0 + \rho\cos\theta \tag{6.5}$$

We take θ to be the gradient angle at (r_0,c_0). Hence

$$\theta = \tan^{-1}\frac{\frac{\partial f}{\partial r}(r_0,c_0)}{\frac{\partial f}{\partial c}(r_0,c_0)} \tag{6.6}$$

Using the cubic polynomial approximation of f given in section 6.1 the gradient angle θ becomes:

$$\theta = \tan^{-1}\frac{K_2 + 2K_4 r_0 + K_5 c_0 + 3K_7 r_0^2 + 2K_8 r_0 c_0 + K_9 c_0^2}{K_3 + K_5 r_0 + 2K_6 c_0 + K_8 r_0^2 + 2K_9 r_0 c_0 + 3K_{10} c_0} \tag{6.7}$$

and

$$
\begin{aligned}
f''_{\theta}(r,c) &= f''_{\theta}(r_0,c_0,\rho) \\
&= 2(3K_7 \sin^2\theta + 2K_8 \sin\theta\cos\theta + K_9\cos^2\theta)r_0 \\
&+ 2(K_8 \sin^2\theta + 2K_9 \sin\theta\cos\theta + 3K_{10}\cos^2\theta)c_0 \\
&+ 2(K_4 \sin^2\theta + K_5 \sin\theta\cos\theta + K_6\cos^2\theta) \\
&+ 6(K_7 \sin^3\theta + K_8 \sin^2\theta\cos\theta + K_9 \sin\theta\cos^2\theta + K_{10}\cos^3\theta)\rho
\end{aligned}
\tag{6.8}
$$

Hence, we declare the point (r_0,c_0) to be an edge point if for some ρ, $|\rho| < \rho_0$, where ρ_0 is slightly smaller than the length of the side of a pixel, $f'_{\theta}(r_0,c_0,\rho) \neq 0$, $f''_{\theta}(r_0,c_0,\rho) = 0$ and $f'''_{\theta}(r_0,c_0,\rho) < 0$, the negatively sloped zero crossing of f''.

6.3 *Corner Detectors*

As previously discussed, under the facet model corners occur at edge points where a significant change in gradient direction takes place. Now, this change in gradient direction should ideally be measured as an incremental change along the edge boundary. We do not desire, however, to perform boundary following since that would require a prior segmentation step. There are several ways in which we have tried to circumvent this problem based on the realization that according to our model the direction of an edge point, that is the tangent to the edge boundary at that point, is orthogonal to the gradient vector at that same point. The simplest approach is to compute the incremental change in gradient direction along the tangent line to the edge at the point which is a corner candidate. The second approach is to evaluate the incremental change along the contour line which passes through the corner candidate. Finally we can compute the instantaneous rate of change in gradient direction in the direction of the tangent line.

In what follows we investigate each of these approaches. In all of them the analysis is based on a continuous surface f obtained by least squares fitting an NxN square neighborhood centered around the corner candidate pixel to the cubic polynomial described in equation (6.1).

The properties of those points away from the neighborhood center and possibly outside the pixel itself have been computed by two different methods:

1. Using the surface fit from the central neighborhood.

2. Using the surface fit from the neighborhood centered around the pixel closest to the tested point.

Although the first method is computationally less expensive than the second one, the possibility of better accuracy exists in the second one.

6.3.1 Incremental Change Along Tangent Line

Consider a row-column coordinate system centered at the corner candidate point. Let $\theta(r,c)$ be the gradient direction at coordinates (r,c) and let $\theta_0 = \theta(0,0)$. Then $(\sin\theta_0, \cos\theta_0)$ is a unit vector in the direction of the gradient at the origin. If the origin is an edge point, the tangent line to the edge boundary which passes through it has direction given by $(-\cos\theta_0, \sin\theta_0)$ and an arbitrary point lying on that line is $\rho(-\cos\theta_0, \sin\theta_0)$.

Consider two points $P_1 = (r_1, c_1)$, $P_2 = (r_2, c_2)$, equidistant to the origin and lying on the tangent line. (See Figure 72 on page 177). P_1 and P_2 are given by $-R(-\cos\theta_0, \sin\theta_0)$ and $R(-\cos\theta_0, \sin\theta_0)$ respectively where R is the distance from

each point to the origin. If R is not too large we can expect the true boundary to lie not too far away from either P_1 or P_2 . In this case a suitable test to decide whether the origin (0,0) is a corner point would involve meeting the following two conditions:

(1) (0,0), (r_1, c_1), (r_2, c_2) are edge points

(2) For a given threshold Ω, $|\theta(r_1, c_1) - \theta(r_2, c_2)| > \Omega$

6.3.2 Incremental Change Along Contour Line

It is reasonable to assume that points on the edge boundary to each side of the corner point and close to it are likely to have similar gray tone intensities. This motivates us to approximate the edge boundary by the contour line f(r,c) = f(0,0) which passes through the corner candidate point at the origin of the coordinate system.

We consider two points $P_1 = (r_1, c_1)$ and $P_2 = (r_2, c_2)$ equidistant to the origin and lying on the contour line instead of the tangent line as in section III.1 (See Figure 73 on page 177). Let θ(r,c) be the gradient direction at coordinates (r,c).

The test to decide whether the origin (0,0) is a corner point is similar to the one used in the previous approach. That is, (0,0) is declared to be a corner point if the following two conditions are satisfied:

(1) (0,0), (r_1, c_1), (r_2, c_2) are edge points

(2) For a given threshold Ω, $|\theta(r_1, c_1) - \theta(r_2, c_2)| > \Omega$

This approach is computationally more expensive than the previous one due to the need of intersecting the cubic curve f(r,c) = f(0,0) (the contour line) with the quadratic curve $r^2 + c^2 = R^2$ in order to determine the points P_1 and P_2 a distance R from the origin.

6.3.3 Instantaneous Rate of Change

Let θ(r,c) be the gradient direction at coordinates (r,c) and let $\theta'_\alpha(r, c)$ be the first directional derivative of θ(r,c) in the direction α. We can compute $\theta'_\alpha(r, c)$ as follows.

Let f(r,c) be the surface function underlying the neighborhood of pixel values centered at the corner candidate pixel. Let $f_r(r,c)$ and $f_c(r,c)$ denote the row and column partial derivatives of f. Consider the line passing through the origin in the direction α. An arbitrary point in this line is given by $\rho(\sin\alpha, \cos\alpha)$ and the gradient direction at that point is given by

$$\theta(\rho \sin\alpha, \rho \cos\alpha) = \tan^{-1} \frac{f_r(\rho \sin\alpha, \rho \cos\alpha)}{f_c(\rho \sin\alpha, \rho \cos\alpha)} \tag{6.9}$$

which can be written as:

$$\theta(\rho) = \tan^{-1} \frac{f_r(\rho)}{f_c(\rho)} \tag{6.10}$$

Differentiating with respect to the parameter ρ results in

$$\theta'(\rho) = \frac{f_c(\rho)f'_r(\rho) - f_r(\rho)f'_c(\rho)}{f_r^2(\rho) + f_c^2(\rho)} \tag{6.11}$$

Using the cubic polynomial approximation for f given in section 6.1 we have:

$$f_r(\rho) = K_2 + (2K_4 \sin\alpha + K_5 \cos\alpha)\rho + (3K_7 \sin^2\alpha + 2K_8 \sin\alpha \cos\alpha + K_9 \cos^2\alpha)\rho^2$$

$$f_c(\rho) = K_3 + (K_5 \sin\alpha + 2K_6 \cos\alpha)\rho + (K_8 \sin^2\alpha + 2K_9 \sin\alpha \cos\alpha + 3K_{10} \cos\alpha)\rho^2$$

$$f'_r(\rho) = (2K_4 \sin\alpha + K_5 \cos\alpha) + 2(3K_7 \sin^2\alpha + 2K_8 \sin\alpha \cos\alpha + + K_9 \cos^2\alpha)\rho$$

$$f'_c(\rho) = (K_5 \sin\alpha + 2K_6 \cos\alpha) + 2(K_8 \sin^2\alpha + 2K_9 \sin\alpha \cos\alpha + 3K_{10} \cos\alpha)\rho$$

The rate of change of gradient direction in the direction α evaluated at the origin ($\rho = 0$) is then:

$$\theta'_\alpha(0) = \frac{K_3(2K_4 \sin\alpha + K_5 \cos\alpha) - K_2(K_5 \sin\alpha + 2K_6 \cos\alpha)}{K_2^2 + K_3^2} \tag{6.12}$$

We are interested in the value of $\theta'_\alpha(0)$ when the direction α is orthogonal to the gradient direction at the origin (the edge direction). Since (K_2, K_3) is the gradient vector at the origin, $(-K_3, K_2)$ is a vector orthogonal to it, and

$$\alpha = \tan^{-1}\frac{-K_3}{K_2} \tag{6.13}$$

Finally using equation (6.13) in equation (6.12) we obtain

$$\theta'_\alpha(0) = \frac{-2(K_2^2 K_6 - K_2 K_3 K_5 + K_3^2 K_4)}{\sqrt{(K_2^2 + K_3^2)^3}} \tag{6.14}$$

The test to decide whether the origin (0,0) is a corner point is as follows. We declare (0,0) to be a corner point if the following two conditions are satisfied:

(1) (0,0) is an edge point

(2) For a given threshold Ω, $|\theta'_{\perp}(0)| > \Omega$.

6.4 *Experimental Results*

We evaluate the performance of the various facet model based gray tone corner detectors by applying them to two digital images. The first one represents a set of artificially generated rectangular shapes at various orientations. The second one is a real aerial image of an urban scene. The first image is 90x90 pixels and contains rectangular shapes of 20x20 pixels with orientations ranging from 0 to 90 degrees in 10 degrees increments. The rectangles have gray tone intensity 175 and the background has gray tone intensity 75. Independent Gaussian noise with zero mean and standard deviation 10 has been added to this image. Defining the signal to noise ratio as 10 times the logarithm of the range of signal divided by the standard deviation of the noise, the artifically generated image has a 10 DB signal to noise ratio. The perfect and noisy versions are shown in Figure 74 on page 178.

Section 6.4.1 illustrates the performance of the various facet model based gray tone corner detectors. Section 6.4.2 compares the performance of the best facet model based gray tone corner detector against the performance of the best Kitchen-Rosenfeld gray tone corner detector and the Dreschler-Nagel corner detector. It is shown that the facet model based gray tone corner detector performs best on the basis of probability of correct assignment.

6.4.1 Facet Model Based Corner Detectors

Each of the corner detector techniques discussed in section III was applied to the artificially generated noisy image using a neighborhood size of 7x7 pixels and a gradient strength threshold for edge detection equal to 20. The choice of neighborhood size depends on the size of the objects in the scene, degree of fit, and noise level. The 7x7 neighborhood used in our experiments is small enough to adequately represent the edges and corners of the objects in the scene with a cubic fit, and large enough to effectively protect against graytone variations due to noise. A gradient strength threshold is needed to eliminate those edges and corners due to noise. The values of gradient threshold used in these experiments were obtained interactively. A procedure for automatic gradient threshold selection based on a cubic facet image model is described elsewhere (Zuniga and Haralick). If the gradient exceeds the threshold value and a zero-crossing occurs in a direction of $\pm$ 14.9 degrees of the gradient direction within a circle of radius one pixel length centered at the point of test then this point is declared to be an edge point. The value of these parameters are the result of experiments performed by Haralick with zero-crossing of second directional derivative edge operators (Haralick, (1982)).

Two conditional probabilities are computed in order to evaluate the performance of the corner detectors. The first one is the conditional probability of assigning a corner within a given distance d from the true corner given that there is a true corner. We denote this probability by P(A|T). The second one is the conditional probability of there being a true corner within a given distance d of an assigned corner when a corner is assigned. This probability is denoted by P(T|A). The experiments with the artificially generated image show that as the threshold for gradient direction change increases, P(A|T) decreases and P(T|A) increases. We select that threshold which equalizes as best as

possible P(A|T) and P(T|A). This choice of threshold results in a maximum assignment of true corners while keeping false assignments to a minimum.

A true corner is defined as the interior pixel in the rectangular shape where two adjacent sides meet. Table 1 shows the probability of correct corner assignment for each case for distances d=0 and d=1. This table shows that a very high percentage of the assigned corner points are guaranteed to lie within one pixel distance from the true corner point. The method which performs best is the one which measures changes in gradient direction as incremental changes along a contour line and which computes properties of tested points away from the neighborhood center using the surface fit from the neighborhood centered around the pixel closest to the tested point. Surprisingly the next best is the simplest one, which uses incremental changes along the tangent line and properties from the same corner candidate central neighborhood for all the tested points in the tangent line.

6.4.2 Comparison With Other Gray Tone Corner Detectors

The performance of the best facet model based corner detector according to Table 1 has been compared against the performance of two recently developed gray tone corner detectors: the Kitchen-Rosenfeld corner detector (1980) and the Dreschler-Nagel corner detector (1982).

Kitchen and Rosenfeld investigated several techniques for gray tone corner detection. Each one computed for every pixel in the image a measure of cornerness and then corners were obtained by thresholding. Their best results are obtained by measuring cornerness by the product of gradient magnitude and instantaneous rate of change in gradient direction evaluated from a quadratic polynomial gray tone surface fit.

Dreschler and Nagel detect corners by the following procedure: For each pixel in the image compute its Gaussian curvature. This is done by doing a local quadratic polynomial fit for each pixel and computing the Hessian matrix. The Gaussian curvature is the product of the main curvatures (eigenvalues of the Hessian matrix). Next, locations of maximum and minimum Gaussian curvatures are found. A pixel is declared to be a corner if the following conditions are satisfied:

1. It has the steepest slope along the line which connects the location of the maximum with the location of the minimum of Gaussian curvature. (This is done only for extrema lying within a given radius from the corner candidate pixel.).

2. The gray tone intensity at the location of maximum Gaussian curvature is larger than the gray tone intensity at the location of minimum Gaussian curvature.

3. The orientation of the main curvature which changes sign between the two extrema, points into the direction of the associated extremum.

Figure 75 on page 179 illustrates the results of applying the facet model based, Kitchen-Rosenfeld, and Dreschler-Nagel gray tone corner detectors to the artificially generated noisy image. In all cases we use a cubic polynomial fitting on a 7x7 neighborhood. A slight modification of the Kitchen-Rosenfeld corner detector is also reported which allows considering only points whose gradient exceeds a given threshold. This results in a substantial improvement over the original Kitchen-Rosenfeld method. The Dreschler-Nagel corner detector was the most sensitive to noise and also a gradient threshold had to be used. Since all three methods being compared use the same cubic polynomial surface fit and the same 7 x 7 neighborhood size, the same gradient threshold of 20 was used in each of them to minimize the effects of the noise. The search for

Gaussian curvature extrema was done in a 5x5 neighborhood. Table 2 shows the probability of correct corner assignment for each case. The best results according to this table are obtained by using the facet model based corner detector, next comes the Kitchen-Rosenfeld corner detector. The Dreschler-Nagel corner detector performs the worst.

Finally Figure 76 on page 180 illustrates the results obtained by applying each of these corner detectors to the aerial image. In all cases we use a cubic polynomial fitting on a 7 x 7 neighborhood. Gradient thresholds are equal to 16.

6.5 Conclusions

We have investigated various approaches for gray tone corner detection which are based on the facet model. We have compared their performance with the Kitchen-Rosenfeld and the Dreschler-Nagel gray tone corner detectors. For the artificially generated image, the facet model based gray tone corner detector performs better than both the Kitchen-Rosenfeld and the Dreschler-Nagel corner detectors. For the real image the facet model based corner detector and the Kitchen-Rosenfeld corner detector have similar performances. Both perform better than the Dreschler-Nagel corner detector.

Further work needs to be done. We need to explore the relationship of basis function kind (polynomial, trigonometric polynomial, etc.), order of fit, and neighborhood size to the goodness of fit. An statistical analysis of each of the techniques described needs to be developed.

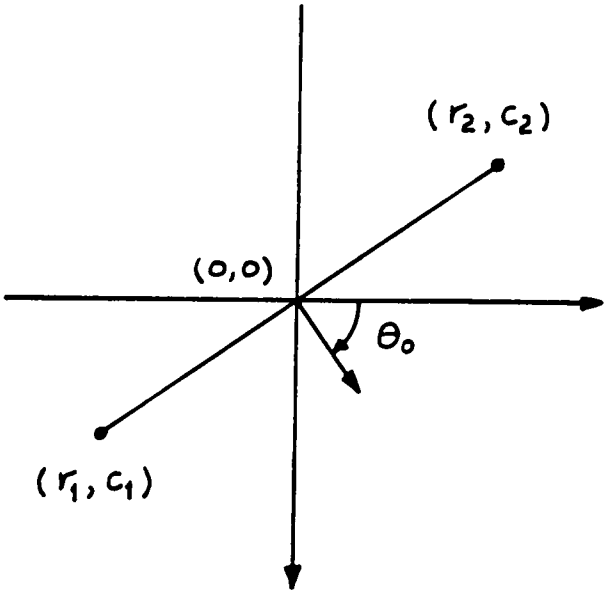


Figure 72. illustrates two points equidistant to the origin and lying on the tangent line to the edge boundary passing through it.

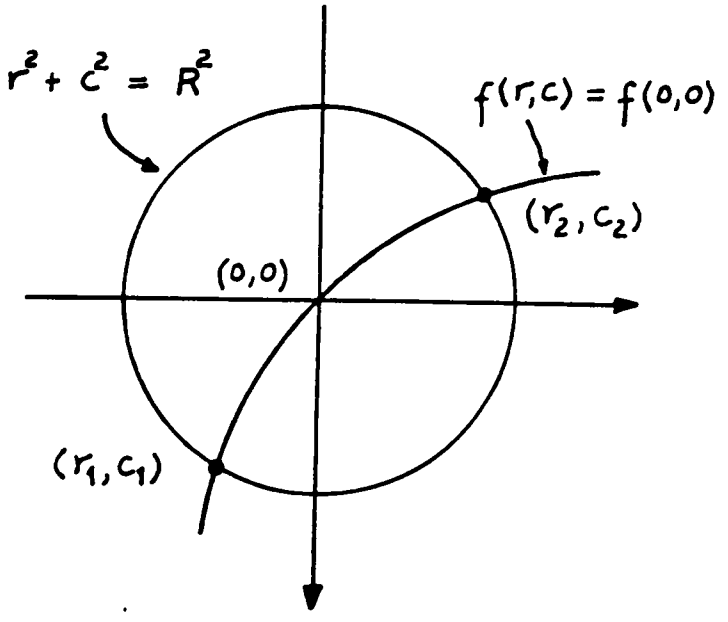


Figure 73. illustrates two points equidistant to the origin and lying on the contour line to the edge boundary passing through it.

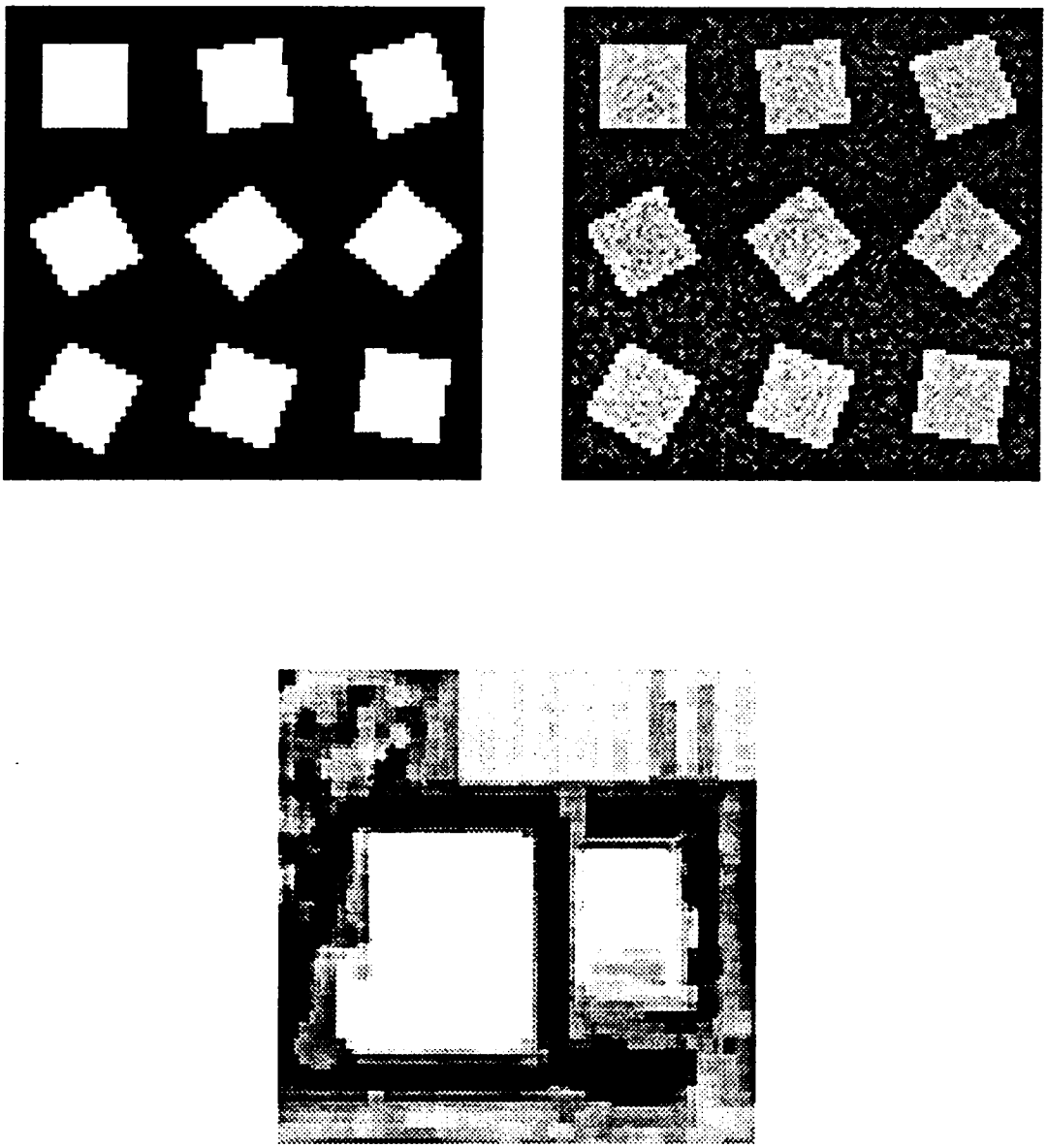


Figure 74. shows the perfect and noisy artificially generated image and the aerial scene.

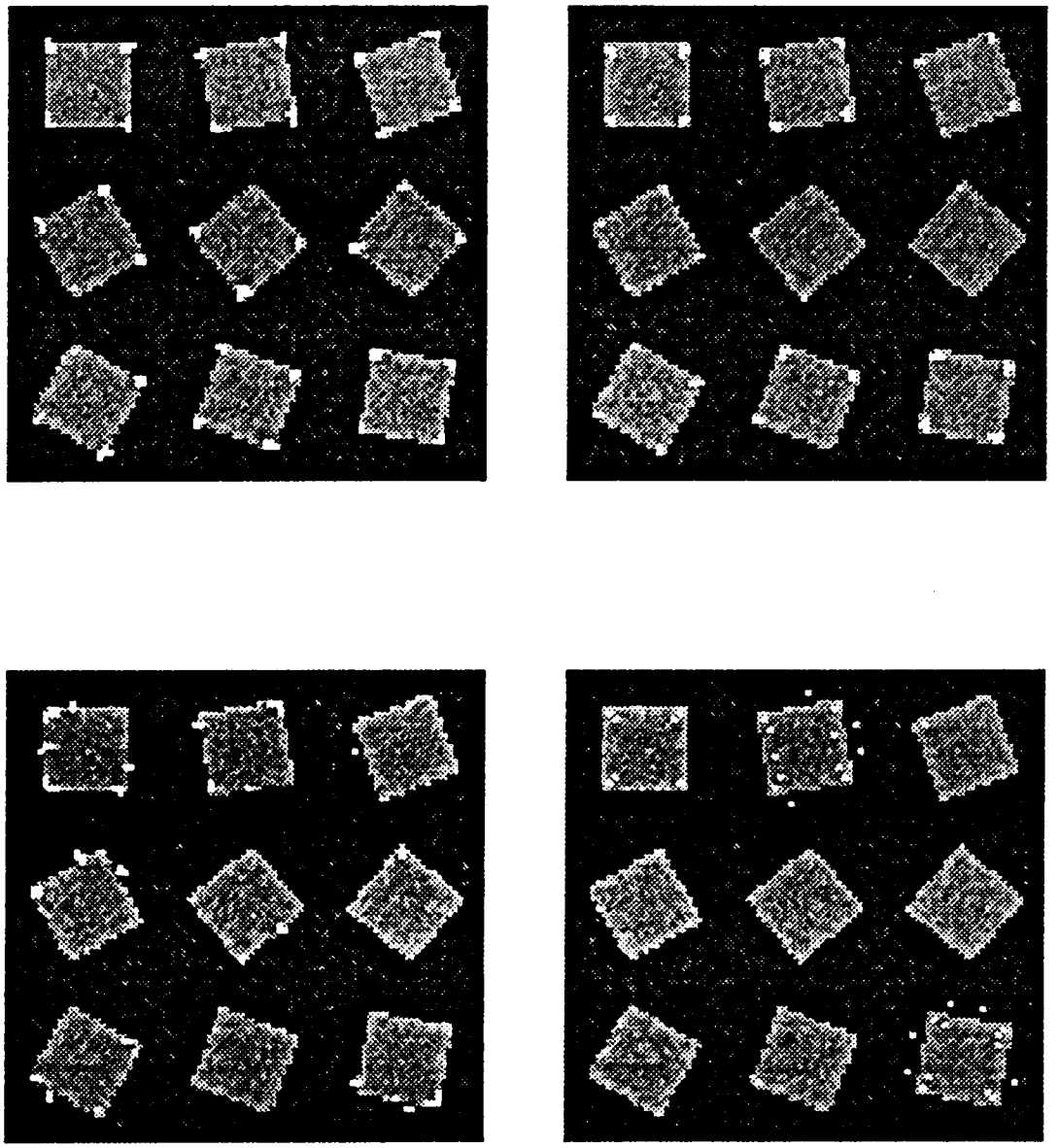


Figure 75. compares the corner assignments in the artificially generated image clockwise from top-left: best facet model, Kitchen-Rosenfeld (with and without gradient threshold) and Dreschler-Nagel corner detectors. Parameters are shown in Table 2 for d = 1.

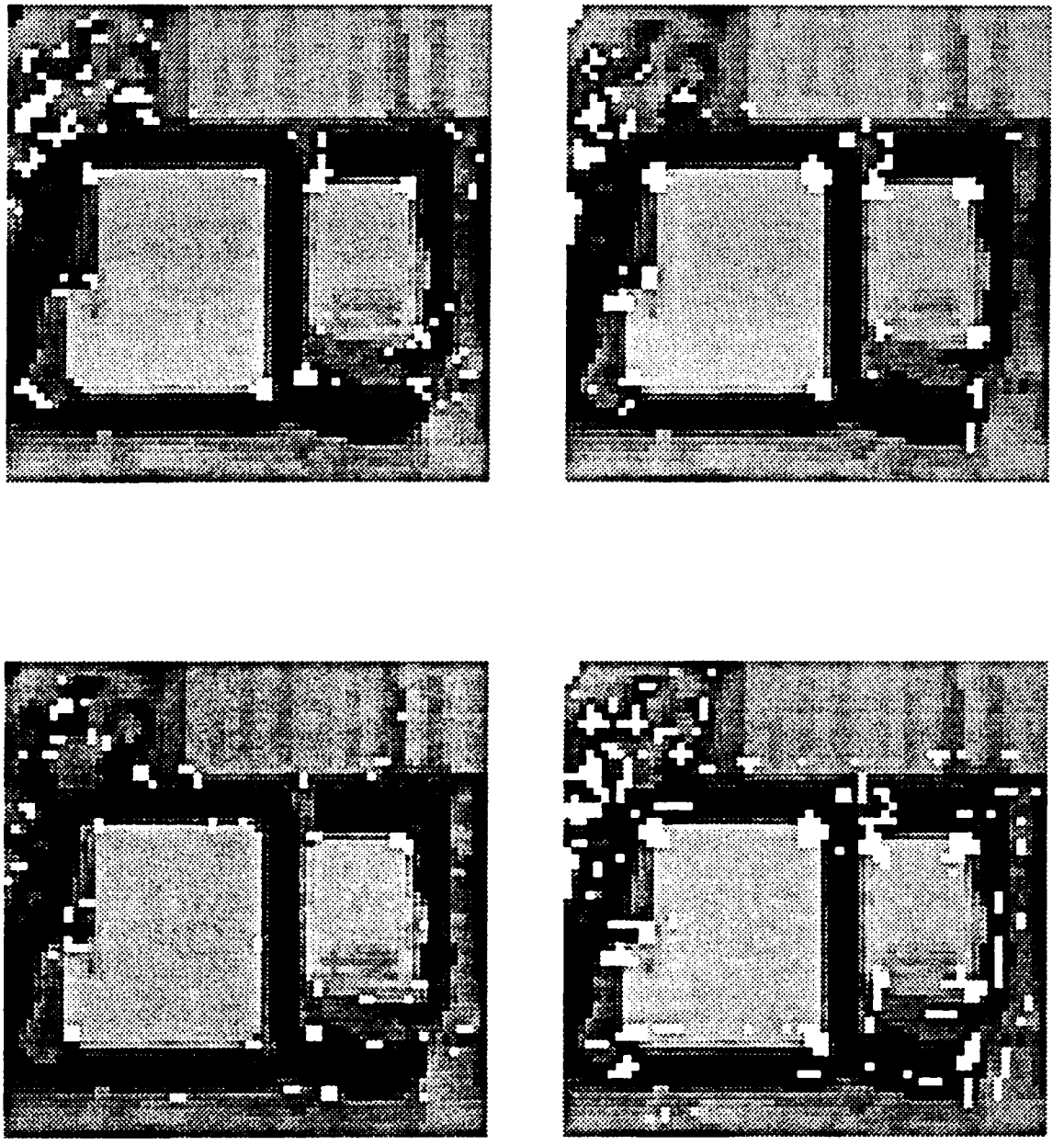


Figure 76. compares the corner assignments in the aerial scene clockwise from top-left: best facet model, Kitchen-Rosenfeld (with and without threshold) and Dreschler-Nagel corner detectors.

Grad Threshold = 20	d = 1			d = 0		
	P(AC|TC)	P(TC|AC)	Angle Threshold	P(AC|TC)	P(TC|AC)	Angle Threshold
Incremental change along tangent line. Central neighborhood. Increment = 3.50 pixels	0.97	0.99	47°	0.278	0.25	55°
Incremental change along tangent line. Nearest neighborhood. Increment = 3.5 pixels	0.97	0.97	67°	0.111	0.108	80°
Incremental change along contour line. Central neighborhood. Increment = 3.5 pixels	0.94	0.94	50°	0.278	0.294	63°
Incremental change along contour line. Nearest neighborhood. Increment = 4 pixels	0.97	0.97	76°	0.361	0.361	94°
Instantaneous rate of change	0.94	0.96	13°/pixel	0.083	0.075	16°/pixel

Table 1. Compares the performance of the facet model based corner detectors. P(AC|TC) is the conditional probability of assigning a corner given that there is a corner. P(TC|AC) is the conditional probability of there being a true corner when a corner is assigned. The parameter d is the maximum distance between assigned and true corners.

	d = 1		d = 0	
	P(AC|TC)	P(TC|AC)	P(AC|TC)	P(TC|AC)
Best facet model corner detector gradient threshold = 20	0.97	0.97	0.361	0.361
Kitchen-Rosenfeld No gradient threshold	0.36	0.36	0.055	0.021
Kitchen-Rosenfeld gradient threshold = 20	0.83	0.84	0.055	0.05
Dreschler-Nagel gradient threshold = 20	0.33	0.35	0.055	0.059

Table 2. **Compares the performance of the best facet model corner detector** with the Kitchen-Rosenfeld and the Dreschler-Nagel corner detectors.

Chapter 7

LINEAR FEATURE EXTRACTION AND DESCRIPTION

7.1 Introduction

The presence of linear features and their spatial relationships provides a strong clue for the recognition of man-made structures in aerial images. Given a low-level vision system that extracts from the image data a set of oriented primitives, ie. edges, the task of an intermediate-level vision system for extracting linear features can be decomposed into two subtasks. The first one consists on the partitioning of the set of primitives into perceptually significant groups. The second subtask consists in fitting each group to the ideal feature sought.

One approach for solving this problem is presented in sections 7.3 and 7.4. Optimal fitting of edge primitives to linear segments requires knowledge of the statistical prop-

erties of the position and orientation of these primitives. Section 7.2 presents a statistical analysis that provides the required knowledge based on a cubic facet model for the image data, and additive independent Gaussian noise. It is found that the joint distribution of edge position and orientation is approximately Normal.

In order to judge the significance of the extracted linear features, section 7.5 presents an statistical analysis of the position and orientation of the fitted linear segments. It is found that the joint distribution of linear segment position and orientation is also approximately Normal. These results suggest that some grouping operations on linear segments, for the purpose of extracting higher level primitives, can be carried out in the same fashion as those operations that manipulate edges. An illustration of this, is fitting a single segment to a number of collinear segments.

Section 7.6 describes the data structure used to represent and store the linear segments, and the descriptors associated with them. Section 7.7 illustrates the application of these methods to the extraction of linear features in aerial images and section 7.8 presents some conclusions.

7.2 Statistical Analysis of Edge position and orientation

We think of an edge as an oriented primitive (ρ, θ), in a coordinate system whose origin is at the pixel's center (see Figure 77 on page 185). Randomness in the image data introduces randomness in the estimate of position $\hat{\rho}$, and orientation $\hat{\theta}$. We study in this section the joint distributional properties of these estimates.

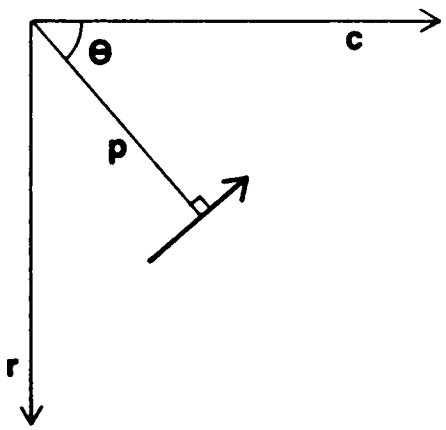


Figure 77. shows an edge as an oriented primitive with position relative to the pixel's center.

7.2.1 Edge position

As described in section 5.3.1, the estimate of edge position is given by the distance from the pixel's center of the zero-crossing of second directional derivative in the gradient direction. This distance is computed by the ratio of second to third directional derivative of the approximating cubic polynomial, where the derivatives are evaluated at the pixel's center. Thus,

$$\hat{\rho} = -\frac{y_\theta^{(2)}(0)}{y_\theta^{(3)}(0)} \tag{7.1}$$

where $y_\theta^{(3)}(0) < 0$, and θ is the gradient direction.

The second and third directional derivatives are

$$y_\theta^{(2)} = y^{(2,0)} \sin^2\theta + 2y^{(1,1)} \sin\theta \cos\theta + y^{(0,2)} \cos^2\theta \tag{7.2}$$

$$y_\theta^{(3)} = y^{(3,0)} \sin^3\theta + 3y^{(2,1)} \sin^2\theta \cos\theta + 3y^{(1,2)} \sin\theta \cos^2\theta + y^{(0,3)} \cos^3\theta$$

According to the results of Appendix B, the partial derivatives $y^{(i,j)}(0)$ are distributed according to a Normal distribution with mean $g^{(i,j)}(0)$ and variance $k_{ij}\sigma^2$, where g is the true but unknown underlying graylevel intensity function, and k_{ij} is given by

$$k_{ij} = \sum_{r=-N}^{N} \sum_{c=-N}^{N} M_{ij}^2(r,c)/w(r,c) \tag{7.3}$$

where $M_{ij}(r,c)$ is the convolution kernel used to compute the partial derivatives $y^{(i,j)}(0)$, based on a neighborhood of (2N+1)x(2N+1) pixels, σ^2 is the noise variance, and w(r,c)

is a weight function defined for each type of orthogonal polynomial approximation, e.g., Discrete Chebyshev, Krawtchouk, etc.

Let $A = y_\theta^{(3)}(0)$ and $B = y_\theta^{(2)}(0)$. From appendix B, the second-order partials are uncorrelated. Therefore, given the gradient direction θ, B is a Normal variable with mean and variance

$$\mu_B = E[y_\theta^{(2)}(0) \mid \theta] = g_\theta^{(2)}(0) \tag{7.4}$$

$$\sigma_B^2 = V[y_\theta^{(2)}(0) \mid \theta] = k_B \sigma^2$$

where k_B is

$$k_B = k_{20} \sin^4\theta + 4k_{11} \sin^2\theta \cos^2\theta + k_{02} \cos^4\theta \tag{7.5}$$

Likewise, the third-order partials are uncorrelated. Therefore, given the gradient direction θ, A is a Normal variable with mean and variance

$$\mu_A = E[y_\theta^{(3)}(0) \mid \theta] = g_\theta^{(3)}(0) \tag{7.6}$$

$$\sigma_A^2 = V[y_\theta^{(3)}(0) \mid \theta] = k_A \sigma^2$$

where k_A is

$$k_A = k_{30} \sin^6\theta + 9k_{21} \sin^4\theta \cos^2\theta + 9k_{12} \sin^2\theta \cos^4\theta + k_{03} \cos^6\theta \tag{7.7}$$

Let us assume, at present, that $\mu_B = g_\theta^{(2)}(0) = 0$, that is, the edge is passing through the pixel's center. Therefore, from the analysis of Appendix D, the conditional distribution of the edge position estimate $\hat{\rho}$ given the gradient direction θ is

$$p(\hat{\rho}\,|\,\theta) = \frac{\beta}{\pi(\beta^2+\hat{\rho}^2)}\, e^{-\alpha^2/2} + \frac{\alpha\beta^2}{\sqrt{2\pi}\,(\beta^2+\hat{\rho}^2)^{3/2}}\, e^{-\alpha^2\hat{\rho}^2/2(\beta^2+\hat{\rho}^2)} erf\left(\frac{\alpha\beta}{\sqrt{2}\sqrt{\beta^2+\hat{\rho}^2}}\right) \tag{7.8}$$

where $\alpha = \mu_A/\sigma_A$ and $\beta = \sigma_B/\sigma_A$

This conditional distribution is unimodal and symmetric about the mode $\hat{\rho} = 0$. It is the sum of a Cauchy distribution and some form of exponential distribution. Strictly speaking, the moments of any order of $p(\hat{\rho}\,|\,\theta)$ do not exist. Fortunately, if the signal to noise ratio is not excessively low, it can be shown that $p(\hat{\rho}\,|\,\theta)$ is approximately Normal. To see this, assume that $\alpha >> 1$, or equivalently $\mu_A >> \sigma_A$, that is, the mean of the estimated third directional derivative is large compared to its standard deviation. We should expect this condition to be met, unless the edge contrast is very poor. Under this condition, most of the density $p(\hat{\rho}\,|\,\theta)$ is concentrated at values of $\hat{\rho}$ that satisfy $\hat{\rho}^2 << \beta^2$. Therefore,

$$p(\hat{\rho}\,|\,\theta) \simeq \frac{1}{\sqrt{2\pi}\,\beta/\alpha}\, e^{-\hat{\rho}^2/2(\beta/\alpha)^2}$$

In general $\mu_B \neq 0$, and $\rho \simeq -\mu_B/\mu_A$. Thus, replacing the values for α and β in the previous expression, we have the approximate conditional distribution

$$p(\hat{\rho}\,|\,\theta) \simeq \frac{1}{\sqrt{2\pi}\,\sigma_B/\mu_A}\, e^{-(\hat{\rho}-\rho)^2/2(\sigma_B/\mu_A)^2} \tag{7.9}$$

Notice that this equation suggests that $\hat{\rho}$ is approximately independent of $\hat{\theta}$. The approximate conditional mean and variance of the edge position estimate are

$$\mu_\rho \simeq E[\hat{\rho}\,|\,\theta] = \rho \tag{7.10}$$

$$\sigma_\rho^2 \simeq V[\hat{\rho} \mid \theta] = \frac{\sigma_B^2}{\mu_A^2} = \frac{k_B}{\mu_A^2}\sigma^2$$

7.2.2 Edge orientation

Edge orientation is given by the gradient orientation θ. Let μ_r and μ_c be the true, but unknown partial derivatives of the graylevel intensity function. The estimated partial derivatives $\hat{\mu}_r$ and $\hat{\mu}_c$ at the pixel's center are

$$\hat{\mu}_r = \frac{\partial y^*}{\partial r}(0,0) \tag{7.11}$$

$$\hat{\mu}_c = \frac{\partial y^*}{\partial c}(0,0)$$

where $y^* = y$ if the Standard derivative is used, or $y^* = y + \frac{L^2}{6}\nabla^2 y$ if the Integrated derivative is used.

Let $G = \sqrt{\hat{\mu}_r^2 + \hat{\mu}_c^2}$ be the estimated gradient. Then

$$\sin\hat{\theta} = \frac{\hat{\mu}_r}{G} \tag{7.12}$$

$$\cos\hat{\theta} = \frac{\hat{\mu}_c}{G}$$

The estimated partials $\hat{\mu}_r$ and $\hat{\mu}_c$ are known to be distributed as $N(\mu_r, k_{10}\sigma^2)$ and $N(\mu_c, k_{01}\sigma^2)$ respectively, where $k_{10} = k_{01} = k$ for a square neighborhood, and it is given by equation 7.3. If we assume that the expected value of the estimated gradient is much greater than its variance, then $\sin\hat{\theta}$ and $\cos\hat{\theta}$ have approximate Normal distributions.

As previously stated, we expect this assumption to be correct unless the edge has very poor contrast. Thus, $\sin\hat{\theta}$ is approximately $N(\mu_s/G, k\sigma^2/G^2)$ and $\cos\hat{\theta}$ is approximately $N(\mu_c/G, k\sigma^2/G^2)$.

Therefore, $\sin(\hat{\theta}-\theta)$ is approximately distributed as $N(0, k\sigma^2/G^2)$. For small values of $\hat{\theta}-\theta$, $\sin(\hat{\theta}-\theta)\simeq\hat{\theta}-\theta$, and thus, $\hat{\theta}$ is approximately distributed as $N(\theta, k\sigma^2/G^2)$.

To verify that our assumptions are not unrealistic, the reader should compare the variance of the estimated gradient direction $\hat{\theta}$, given by $\sigma_{\hat{\theta}}^2\simeq k\sigma^2/G^2$ with the experimental results shown in Figure 15 on page 61, of chapter 3, for an edge orientation of 22.5 degrees. The predicted values of $\sigma_{\hat{\theta}}^2$ agree very closely with the experimental values. When using the Integrated directional derivative, this agreement persists even when the noise standard deviation σ is close to the value of edge contrast, suggesting robustness in our procedure.

7.2.3 Joint distribution of edge position and orientation

The analysis of the previous two sections suggests the following approximate joint probability density function for the estimates of position $\hat{\rho}$, and orientation $\hat{\theta}$ of an edge.

$$p(\hat{\rho},\hat{\theta}) = \frac{1}{2\pi\sigma_\rho\sigma_\theta}\, e^{-(\hat{\rho}-\rho)^2/2\sigma_\rho^2}\; e^{-\sin^2(\hat{\theta}-\theta)/2\sigma_\theta^2} \tag{7.13}$$

where the forms $\hat{\rho}$ and $\hat{\theta}$ are approximately independent, and

$$\sigma_\rho^2 = \frac{k_B}{\mu_A^2}\,\sigma^2 \tag{7.14}$$

$$\sigma_\theta^2 = \frac{k}{G^2}\,\sigma^2$$

7.3 *From Edges to Segments*

7.3.1 Local Linking

Edge labeled pixels forming one-dimensional spatial clusters are grouped or linked into edge segments. Prior to this linking operation, edges are thinned by suppressing non-maxima edge magnitudes accross the edge.

We define an edge segment to be a sequence of 8-connected pixels occuring continuously between two junctions, two end points, or one junction and one end point. We define a junction as the point where three or more edge segments meet. An end point is the pixel where only one edge segment originates or terminates. To include closed edge segments we allow an originating or terminating endpoint to be the same.

Any pixel in a edge segment which is not a junction or an end point is referred to as an arc. An arc has a successor and a predecessor pixel any of which can be either an arc, a junction, or an end point. Whether a pixel is interpreted as an arc, junction, or end point is determined by the following criteria.

1. An arc pixel together with its successor and predecessor pixels form an angle of at least ninety degrees.

2. A junction pixel together with any two of its neighboring arc pixels form an angle of at least ninety degrees

3. Whenever there is a choice between four-connectedness and eight-connectedness, four-connectedness is preferred.

Application of these criteria is illustrated in Figure 78 on page 196 and all possible interpretations are illustrated in Figure 79 on page 197.

The algorithm we use is based on a linked-list/hash-table data structure. Each edge segment is assigned an identification number which is used as a hashing key. Each edge segment has also a header node in the hash table which points to a linked list containing the current sequence of pixels which form the segment. Whenever a new pixel belonging to the segment is found, it is added either to the head or to the tail of the segment linked list according to the direction of the sequence. Several operations are implemented on this linked-list/hash-table data structure which include adding nodes to the head or to the tail of a linked list; merging two linked lists; reversing a linked list, and freeing the nodes of a linked list. The linked-list/hash-table data structure is kept in memory and contains the information for all the active edge segments. Whenever a segment becomes inactive; ie, the entire sequence is completed; the edge segment is written to a file and the nodes it occupies are freed. Only one pass through the image is required. A summary of the algorithm follows.

```
For each pixel in a top to bottom, left to right scan of the edge labeled image;
  {
  Determine if the pixel is an arc, a junction,
  or an end point;
  if arc;
    {
    if predecessor (successor) has already been
    added to a list;
      {
      add arc pixel to the tail (head) of list;
```

```
    }
  else;                see Figure 80 on page 198
    {
    start a new list;
    assign a new identification number to it;
    add arc pixel to the list;
    }
  if both predecessor and successor have already been
  added to different lists;    see Figure 81 on page 199
    {
    add arc pixel to the tail of predecessor list;
    merge predecessor and successor list;
    keep the lowest identification number;
    }
  }
if junction;
  {
  for each arc segment meeting at the junction;
    {
    if arc segment points to the east, south-east,
    south or south-west;        see Figure 80 on page 198
      {
      start a new list;
      assign a new identification number to it;
      add junction pixel to the list;
      }
```

```
    else
      {
      if junction pixel is a successor;
        add pixel to the tail of the arc list;
      else;
        add pixel to the head of the arc list;
      }
    }
}
if end point;
    {
    if end point points to the east, south-east,
    south or south-west;          see Figure 80 on page 198
      {
      start a new list;
      assign a new identification number to it;
      add end point to the list;
      }
    else;
      {
      if end point is the successor of the
      neighboring arc;
          add pixel to the tail of the arc list;
      else;
          add pixel to the head of the arc list;
      }
```

```
        }
    Write to a file the lists of those arc segments
    for which both end points and/or junctions have
    been found;
    Free the nodes for those lists;
    }
end
```

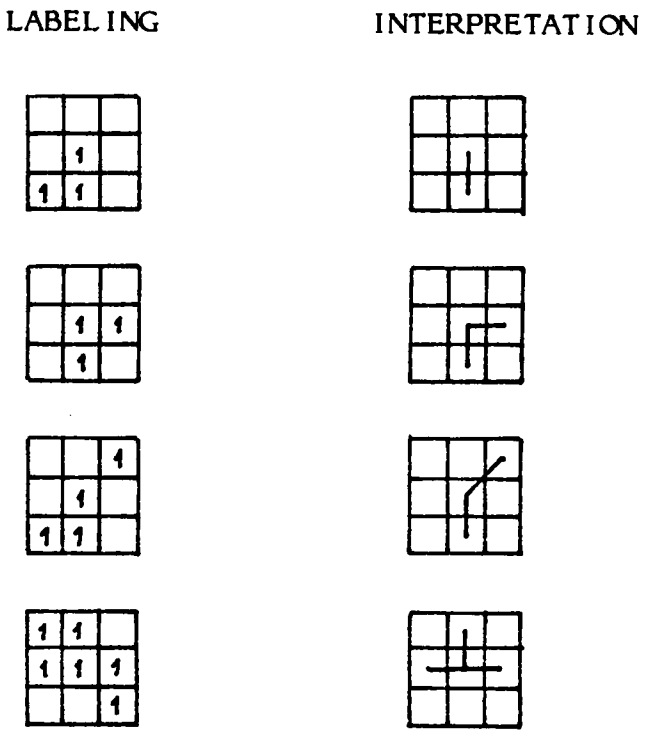


Figure 78. shows some edge labelings and their interpretation.

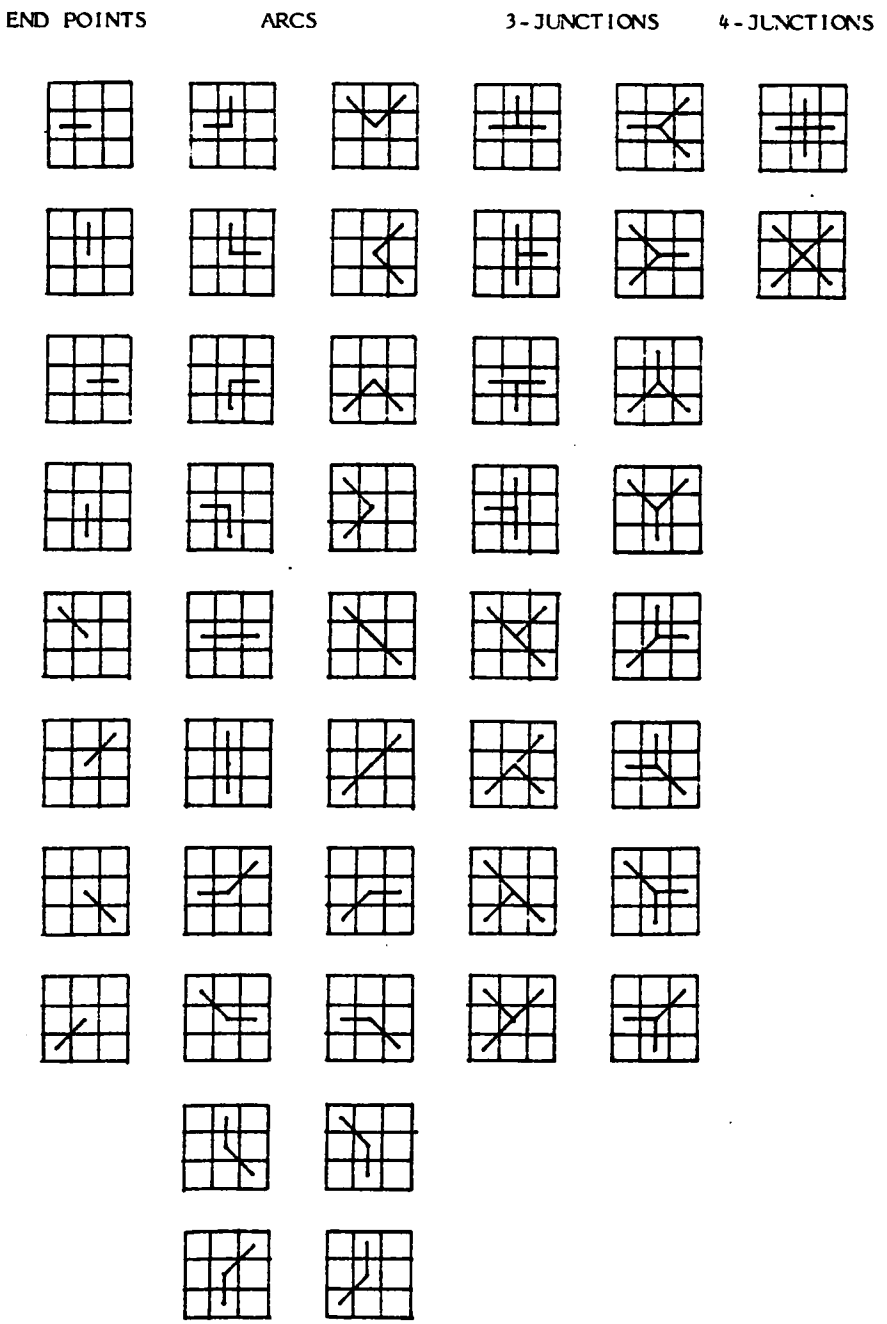


Figure 79. shows all possible interpretations for edge labelings.

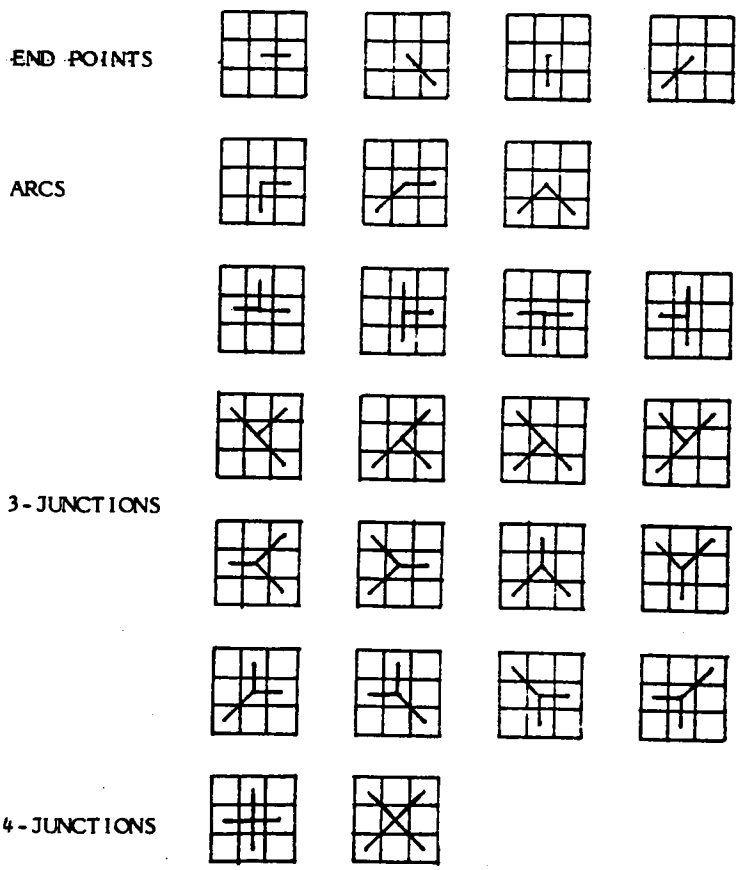


Figure 80. shows the endpoints, arcs, and junctions that originate new segment lists.

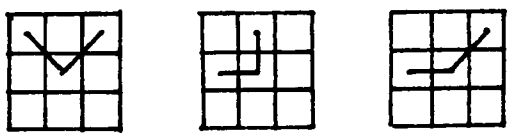


Figure 81. shows the arcs that give rise to the merging of two lists.

7.4 *Linear Segment Extraction*

The thinning and linking processes described in the previous section produce a set of edge segments. Each segment is a set of connected edge primitives. Relevant descriptors for each edge primitive are position, orientation, and first, and third directional derivatives in the gradient direction. At this stage, there are typically many small isolated segments, or small segments which do not continue other segments smoothly. In order to avoid excessive later processing, some pruning of segments is necessary. Segments below a user specified length are deleted unless they are linked to other segments in such a way that their combined length is above the specified length and they meet smooth continuation requirements. The result of this pruning operation is to retain the most perceptually significant segments.

The next step of interest is the partitioning of each segment into perceptually significant groups of primitives. The partitioning process is in general difficult, and several approaches have been described in the Literature Review of chapter 2. For our applications, these groups should correspond to straight segments. We have chosen to use the iterative end-point partitioning algorithm (Duda and Hart (1972)), which is specially suited for this case. Once this partitioning has been accomplished, the next step consists in fitting a linear segment to each group.

7.4.1 Fitting Linear Segments

The grouping process described in the previous section produces sets of primitives likely to correspond to straight line segments. The primitives in each set are described

by the position-orientation pair (ρ_i,θ_i). The parameter ρ_i in this pair is relative to a pixel centered coordinate system. For our purposes in this section, it is more convenient to express the position as the pair (r_i,c_i), in the image row-column coordinate system. The relationship between ρ_i and (r_i,c_i) is

$$r_i = r_i' + \rho_i \sin\theta_i$$

$$c_i = c_i' + \rho_i \cos\theta_i$$

where (r_i', c_i') represents the row-column coordinates at the center of the ith pixel in the segment under consideration.

We would like to assign to each set a straight line that fits best the primitives in a well defined sense. Let the ideal line (see Figure 82 on page 202) be described by the equation

$$r \sin\theta + c \cos\theta - \rho = 0 \tag{7.15}$$

In order to find the best estimates of the line parameters ρ and θ, we need a model for the errors in the position and orientation of the primitives we are fitting. This statistical model according to the analysis of section 7.2 is

$$p_{\theta,\rho}(r_i,c_i,\theta_i) = \frac{1}{2\pi\sigma_{\rho i}\sigma_{\theta i}} e^{-\frac{(r_i \sin\theta + c_i \cos\theta - \rho)^2}{2\sigma_{\rho i}^2}} e^{-\frac{\sin^2(\theta_i - \theta)}{2\sigma_{\theta i}^2}} \tag{7.16}$$

where $\sigma_{\rho i}^2$ and $\sigma_{\theta i}^2$ are given by

$$\sigma_{\rho i}^2 = \frac{k_B}{\mu_{Ai}^2}\sigma^2 \tag{7.17}$$

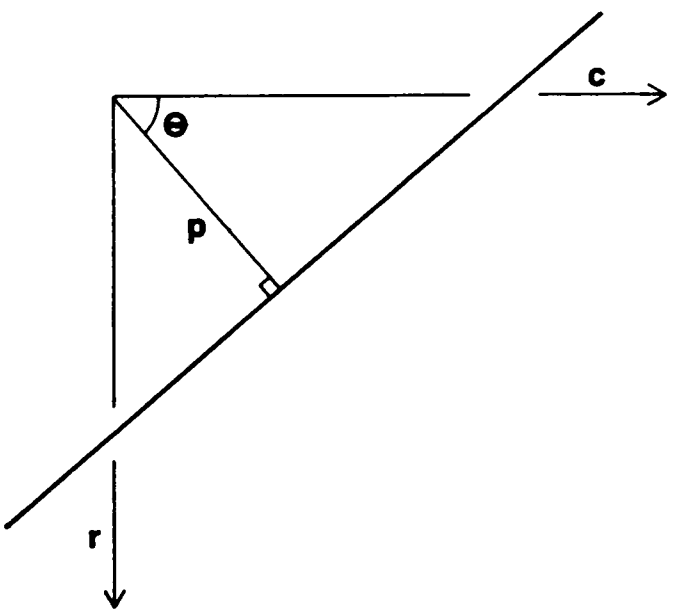


Figure 82. shows the ideal line to be fitted to a set of oriented edge primitives.

$$\sigma_{\theta l}^2 = \frac{k}{G_l^2}\sigma^2$$

Under this model, the results of Appendix E show that the minimum variance unbiased estimates of ρ and θ, obtained by minimizing the error

$$\varepsilon^2 = \sum_l \frac{(r_l \sin\theta + c_l \cos\theta - \rho)^2}{\sigma_{\rho l}^2} + \sum_l \frac{\sin^2(\theta_l - \theta)}{\sigma_{\theta l}^2} \tag{7.18}$$

are given by

$$\hat{\rho} = \bar{r}\sin\hat{\theta} + \bar{c}\cos\hat{\theta} \tag{7.19}$$

$$\hat{\theta} = \frac{1}{2}\tan^{-1}\frac{\overline{\sin 2\theta} - 2\frac{w_\rho}{w_\theta}S_{rc}}{\overline{\cos 2\theta} + \frac{w_\rho}{w_\theta}(S_r^2 - S_c^2)}$$

with $\bar{r}$, $\bar{c}$, $\overline{\sin 2\theta}$, $\overline{\cos 2\theta}$, S_r^2, S_c^2, S_{rc}, w_ρ, and w_θ as defined in Appendix E.

Substituting in equations (7.19) the expressions derived for $\sigma_{\rho i}^2$ and $\sigma_{\theta i}^2$, we have

$$\bar{r} = \frac{\sum_l r_l \mu_{Al}^2}{\sum_l \mu_{Al}^2} \tag{7.20}$$

$$\bar{c} = \frac{\sum_i c_i \mu_{Ai}^2}{\sum_i \mu_{Ai}^2}$$

$$\overline{\sin 2\theta} = \frac{\sum_i G_i^2 \sin 2\theta_i}{\sum_i G_i^2}$$

$$\overline{\cos 2\theta} = \frac{\sum_i G_i^2 \cos 2\theta_i}{\sum_i G_i^2}$$

$$S_r^2 = \frac{\sum_i \mu_{Ai}^2 (r_i - \bar{r})^2}{\sum_i \mu_{Ai}^2}$$

$$S_c^2 = \frac{\sum_l \mu_{Al}^2 (c_l - \bar{c})^2}{\sum_l \mu_{Al}^2}$$

$$S_{rc} = \frac{\sum_l \mu_{Al}^2 (r_l - \bar{r})(c_l - \bar{c})}{\sum_l \mu_{Al}^2}$$

$$\frac{w_\rho}{w_\theta} = \frac{k \sum_l \mu_{Al}^2}{k_B \sum_l G_l^2}$$

Further simplification of the set of equations (7.20) is possible since for a step edge we have, $\mu_{Al}^2 = ah_l^2$ and $G_l^2 = bh_l^2$, where h_l is the edge contrast, and a and b depend only on the edge orientation. Then equations (7.20) become

$$\bar{r} = \frac{\sum_l h_l^2 r_l}{\sum_l h_l^2} \tag{7.21}$$

$$\bar{c} = \frac{\sum_l h_l^2 c_l}{\sum_l h_l^2}$$

$$\overline{\sin 2\theta} = \frac{\sum_l h_l^2 \sin 2\theta_l}{\sum_l h_l^2}$$

$$\overline{\cos 2\theta} = \frac{\sum_l h_l^2 \cos 2\theta_l}{\sum_l h_l^2}$$

$$S_r^2 = \frac{\sum_l h_l^2 (r_l - \bar{r})^2}{\sum_l h_l^2}$$

$$S_c^2 = \frac{\sum_i h_i^2 (c_i - \bar{c})^2}{\sum_i h_i^2}$$

$$S_{rc} = \frac{\sum_i h_i^2 (r_i - \bar{r})(c_i - \bar{c})}{\sum_i h_i^2}$$

$$\frac{w_\rho}{w_\theta} = \frac{ka}{k_B b}$$

Notice that these equations allow for variable contrast along the edge segment being fitted.

7.5 *Statistical Analysis of Linear Segment position and orientation*

7.5.1 Linear Segment position

According to the statistical model for individual edge orientation and position in a linear segment, as expressed in equation (7.16), the conditional distribution of edge position ρ_i, given the linear segment orientation θ is

$$p(\rho_i \mid \theta) = p(r_i \sin\theta + c_i \cos\theta \mid \theta) \sim N(\rho, \sigma_{\rho i}^2) \tag{7.22}$$

Furthermore, given θ, the estimate of linear segment position, from equation (7.19) is

$$\hat{\rho} = \bar{r}\sin\theta + \bar{c}\cos\theta = \frac{\sum_i \rho_i/\sigma_{\rho i}^2}{\sum_i 1/\sigma_{\rho i}^2} \tag{7.23}$$

Therefore, from equations (7.22) and (7.23), the conditional distribution of $\hat{\rho}$, given θ is

$$p(\hat{\rho} \mid \theta) \sim N(\rho, 1/\sum_i 1/\sigma_{\rho i}^2) \tag{7.24}$$

7.5.2 Linear Segment orientation

Consider the transformation that takes every pair (r_i, c_i) to the pair (x_i, y_i), according to

$$x_i = -(r_i - \bar{r})\cos\theta + (c_i - \bar{c})\sin\theta \tag{7.25}$$

$$y_i = (r_i - \bar{r})\sin\theta + (c_i - \bar{c})\cos\theta$$

This transformation consists of a translation of the coordinate system to the center of mass of the samples, plus a rotation θ, such that the rotated axis x coincides with the ideal linear segment. We consider the x_is to be fixed numbers. From equations (7.22), and (7.24), y_i is a Normal random variable with zero mean. If $\sigma_{\rho i}^2 >> 1/\sum_i 1/\sigma_{\rho i}^2$, then the variance of y_i is approximately equal to $\sigma_{\rho i}^2$. Thus,

$$y_i \sim N(0, \sigma_{\rho i}^2) \tag{7.26}$$

Multiplying x_i and y_i in equation (7.25) an dividing by $\sigma_{\rho i}^2$, produces

$$\frac{x_i y_i}{\sigma_{\rho i}^2} = -\frac{1}{2}\sin 2\theta\left[\frac{(r_i - \bar{r})^2}{\sigma_{\rho i}^2} - \frac{(c_i - \bar{c})^2}{\sigma_{\rho i}^2}\right] - \cos 2\theta\left[\frac{(r_i - \bar{r})(c_i - \bar{c})}{\sigma_{\rho i}^2}\right]$$

Summation over i, and division by $\sum_i 1/\sigma_{\rho i}^2$, yields

$$\frac{2\sum_i x_i y_i/\sigma_{\rho i}^2}{\sum_i 1/\sigma_{\rho i}^2} = -\sin 2\theta(S_r^2 - S_c^2) - \cos 2\theta(2S_{rc})$$

Multiplying by w_ρ/w_θ and calling the result $\bar{u}$ produces

$$\bar{u} = \frac{2\sum_i x_i y_i/\sigma_{\rho i}^2}{\sum_i 1/\sigma_{\theta i}^2} = -\sin 2\theta \frac{w_\rho}{w_\theta}(S_r^2 - S_c^2) - \cos 2\theta(2\frac{w_\rho}{w_\theta} S_{rc}) \tag{7.27}$$

From equations (7.17) and (7.26), $\bar{u}$ is a Normal variable

$$\bar{u} \sim N(0, \sigma_u^2) \tag{7.28}$$

$$\sigma_u^2 = V[\bar{u}] = \frac{4\sum_i x_i^2/\sigma_{\rho i}^2}{(\sum_i 1/\sigma_{\theta i}^2)^2} = 4\frac{k^2}{k_B}\frac{\sum_i (x_i\mu_{Ai})^2}{(\sum_i G_i^2)^2}\sigma^2$$

Now, from the model of equation (7.16), we have

$$\sin(\theta_i - \theta) \sim N(0, \sigma_{\theta i}^2) \tag{7.29}$$

and from the analysis of section 7.2.2, we can assume that $\theta_i - \theta$ is small. Then approximately

$$v_i = \sin 2(\theta_i - \theta) \sim N(0, 4\sigma_{\theta i}^2) \tag{7.30}$$

Dividing by $\sigma_{\theta i}^2$,

$$\frac{v_i}{\sigma_{\theta i}^2} = \frac{\sin 2\theta_i}{\sigma_{\theta i}^2} \cos 2\theta - \frac{\cos 2\theta_i}{\sigma_{\theta i}^2} \sin 2\theta$$

Summation over i, and division by $\sum_i 1/\sigma_{\theta i}^2$ yields,

$$\bar{v} = \frac{\sum_i v_i/\sigma_{\theta i}^2}{\sum_i 1/\sigma_{\theta i}^2} = -\sin 2\theta\, \overline{\cos 2\theta} + \cos 2\theta\, \overline{\sin 2\theta} \tag{7.31}$$

From equations (7.30) and (7.17), $\bar{v}$ is a Normal variable

$$\bar{v} \sim N(0, \sigma_v^2) \tag{7.32}$$

$$\sigma_v^2 = V[\bar{v}] = \frac{4}{\sum_i 1/\sigma_{\theta i}^2} = \frac{4k}{\sum_i G_i^2}\sigma^2$$

From equations (7.27) and (7.31)

$$\bar{u} + \bar{v} = -\sin 2\theta[\overline{\cos 2\theta} + \frac{w_\rho}{w_\theta}(S_r^2 - S_c^2)] + \cos 2\theta[\overline{\sin 2\theta} - 2\frac{w_\rho}{w_\theta} S_{rc}] \tag{7.33}$$

Let Δ be defined as

$$\Delta = \sqrt{(\overline{\sin 2\theta} - 2\frac{w_\rho}{w_\theta} S_{rc})^2 + (\overline{\cos 2\theta} + \frac{w_\rho}{w_\theta}(S_r^2 - S_c^2))^2} \tag{7.34}$$

Then, from equations (7.19) and (7.33)

$$\frac{\bar{u}+\bar{v}}{\Delta} = -\sin 2\theta \cos 2\hat{\theta} + \cos 2\theta \sin 2\hat{\theta}$$
$$= \sin 2(\hat{\theta} - \theta)$$

We expect $\hat{\theta} - \theta$ to be small. Therefore

$$\frac{\bar{u}+\bar{v}}{2\Delta} \simeq \sin(\hat{\theta} - \theta) \tag{7.35}$$

Further, we assume that the variance of Δ is small compared to its mean. Therefore

$$\sin(\hat{\theta} - \theta) \sim N(0,(\sigma_u^2 + \sigma_v^2)/4\Delta^2) \tag{7.36}$$

7.5.3 Joint distribution of Linear Segment position and orientation

Equations (7.24) and (7.36) suggest the following joint probability density function for the estimates of position $\hat{\rho}$, and orientation $\hat{\theta}$ of a linear segment.

$$p(\hat{\rho}, \hat{\theta}) = \frac{1}{2\pi\sigma_\rho\sigma_\theta} e^{-(\hat{\rho}-\rho)^2/2\sigma_\rho^2} e^{-\sin^2(\hat{\theta}-\theta)/2\sigma_\theta^2} \tag{7.37}$$

where

$$\sigma_\rho^2 = 1/\sum_l 1/\sigma_{\rho l}^2 \tag{7.38}$$

$$\sigma_\theta^2 = (\sigma_u^2 + \sigma_v^2)/4\Delta^2$$

Notice that the joint p.d.f. of the parameters $\hat{\rho}$ and $\hat{\theta}$ of a linear segment has exactly the same form as the joint p.d.f of the estimates ρ_i and θ_i for individual edges. This implies that we can work with sets of linear segments in the same fashion we have worked with sets of edges. An example of this is fitting a set of collinear segments to a single segment.

7.6 Linear Segment Description

Linear segments are organized into two random access files, a property file, and a chain code file. Each record of the property file contains a linear segment identification number, a list of properties, a list of relationships with neighboring segments, and a pointer to a chain code file that stores the chain code of the corresponding linear segment.

One type of relationship among linear segments is established at the points where several segments meet (junctions or corner points). This relationship is described by the 'winged-edge' representation commonly used in boundary representations of polyhedrons (Baumgart, 1972). Thus each linear segment description includes the identification number of its next leftmost, next rightmost, previous leftmost, and previous rightmost segments.

The algorithm described in section 7.3 produces a property file and a chain code file. The property file has the following format:

- Record 1: Header array

- Record 2-8: Header records. Each record as follows:

  2 Maximum of each property

  3 Minimum of each property

  4-5 Unused

  6 property numbers

  7-8 Unused

- Record 9-n: Data records. Each column in a record as follows:

  1 Linear segment identification number

  2 Beginning junction identification number

  3-4 Beginning row and column

  5 Ending junction identification number

  6-7 Last row and column

  10 pointer to chain code file

  11 chain code length

  12 Next leftmost segment identification number

  13 Next rightmost segment identification number

  14 Previous leftmost segment identification number

  15 Previous rightmost segment identification number

  24 mean gray level on the segment's dark side

  25 mean gray level on the segment's bright side

  26 edge label

  39 length

43 linear segment orientation (in degrees)

44 linear segment orientation standard deviation (σ_θ)

50-51 row and column of the segment center of mass

52 linear segment position

53 linear segment position standard deviation (σ_ρ)

54 mean squared gradient strength

55 mean squared third directional derivative

56 orientation estimate significance ($1/\sigma_\theta$)

57 distance estimate significance ($1/\sigma_\rho$)

58 line segment significance ($1/\sqrt{\sigma_\rho \sigma_\theta}$)

Each record in the chain code file contains the chain code for the corresponding segment stored in the property file. As seen above each segment record in the property file contains a pointer to the record in the chain code file storing the segment chain code. Chain codes are assigned according to the following convention

direction	code
East	0
South-east	1
South	2
South-west	3
West	4
North-west	5
North	6
North-east	7

7.7 *Experimental Results*

This section illustrates the application of the linear segment extraction method to the urban scenes shown in figures Figure 63 on page 153, Figure 64 on page 154, and Figure 65 on page 155. We used the edge images obtained using Haralick's zero-crossing of second directional derivative, with the integrated directional derivative gradient and automatic Bayesian threshold, and shown in Figure 66 on page 156, Figure 68 on page 158, and Figure 70 on page 160. We discuss first the Baton Rouge aerial scene.

Figure 83 on page 218 shows the thinned edges after non-maxima suppression of edge magnitudes accross the edge. After linking the thinned edges into segments, a pruning operation is applied to delete small isolated segments, or small segments which do not continue other segments smoothly. Segments below a user specified length are deleted unless they are linked to other segments in such a way that their combined length is above the user specified length and they meet smooth continuation requirements. Figure 84 on page 219 shows the results of the pruning operation. The user specified length was 7. It can be seen that this pruning operation does a good job at retaining the most perceptually significant segments.

Figure 85 on page 220 shows the results of partitioning each segment into piecewise linear groups using the iterative end-point partitioning algorithm and fitting the resulting groups to straight lines using the method of section 7.4. It can be seen that this fitting operation succeeds in providing good estimates of the orientation and position of most of the building sides present in the scene.

As pointed out earlier, we can judge the significance of the extracted linear segments by examining the standard deviation of their orientation and position parameter estimates as derived in section 7.5. Figure 86 on page 221 and Figure 87 on page 222 show

images in which segment intensities have been made inversely proportional to their orientation and position parameter standard deviation respectively.

Finally Figure 88 on page 223 through Figure 97 on page 232 illustrate the results for the Sunnyvale images.

7.8 Conclusions

We have described a linear segment extraction method that uses both zero-crossing positional and angular information together with their distributional properties to accomplish optimal linear segment fitting.

We found that under a cubic facet model for the image data and additive independent Gaussian noise, the joint distribution of zero-crossing position and orientation is approximately Normal. We also investigated the joint distributional form of the estimated positional and angular parameters of the fitted linear segments. We demonstrated that this joint distribution is also approximately Normal. Experiments with real images showed that the standard deviation of the linear segment parameter estimates provides a very effective way of judging the perceptual significance of the extracted segments.

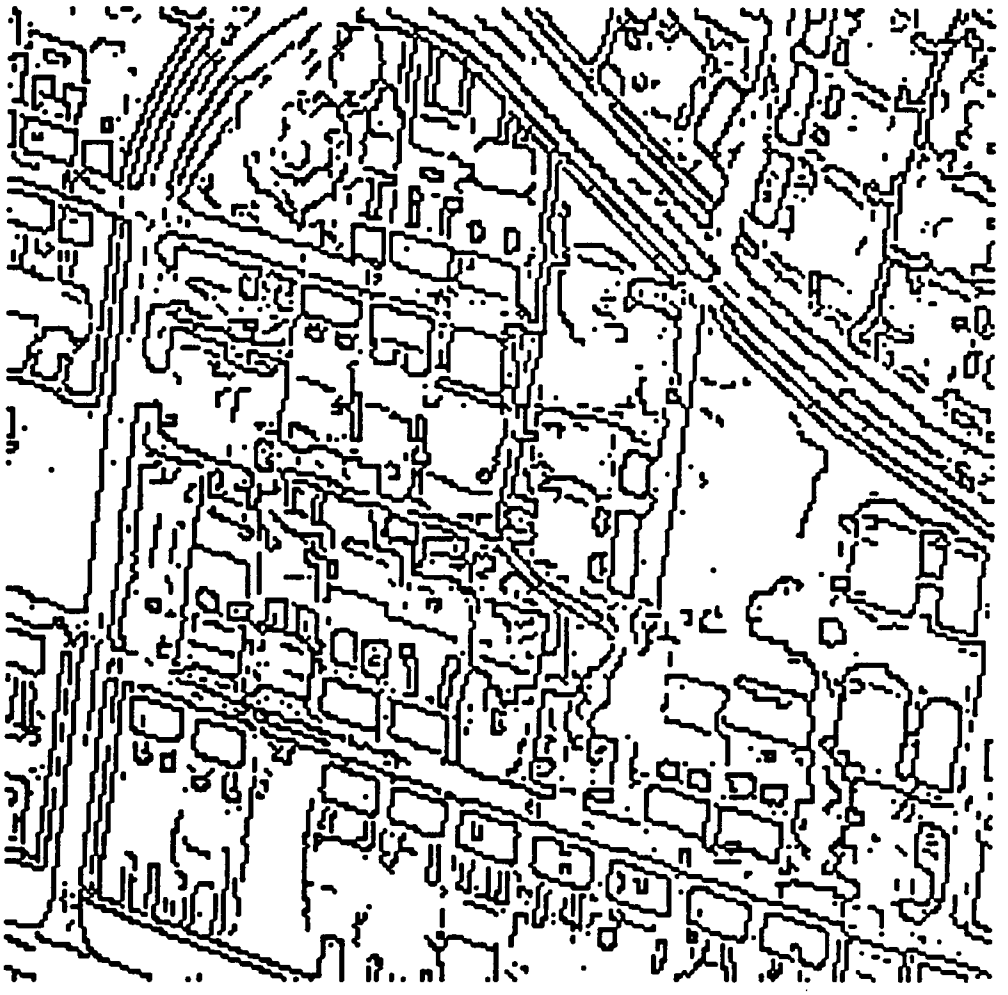


Figure 83. shows the thinned edges after non-maxima suppression.

Figure 84. shows the edge segments remaining after pruning.

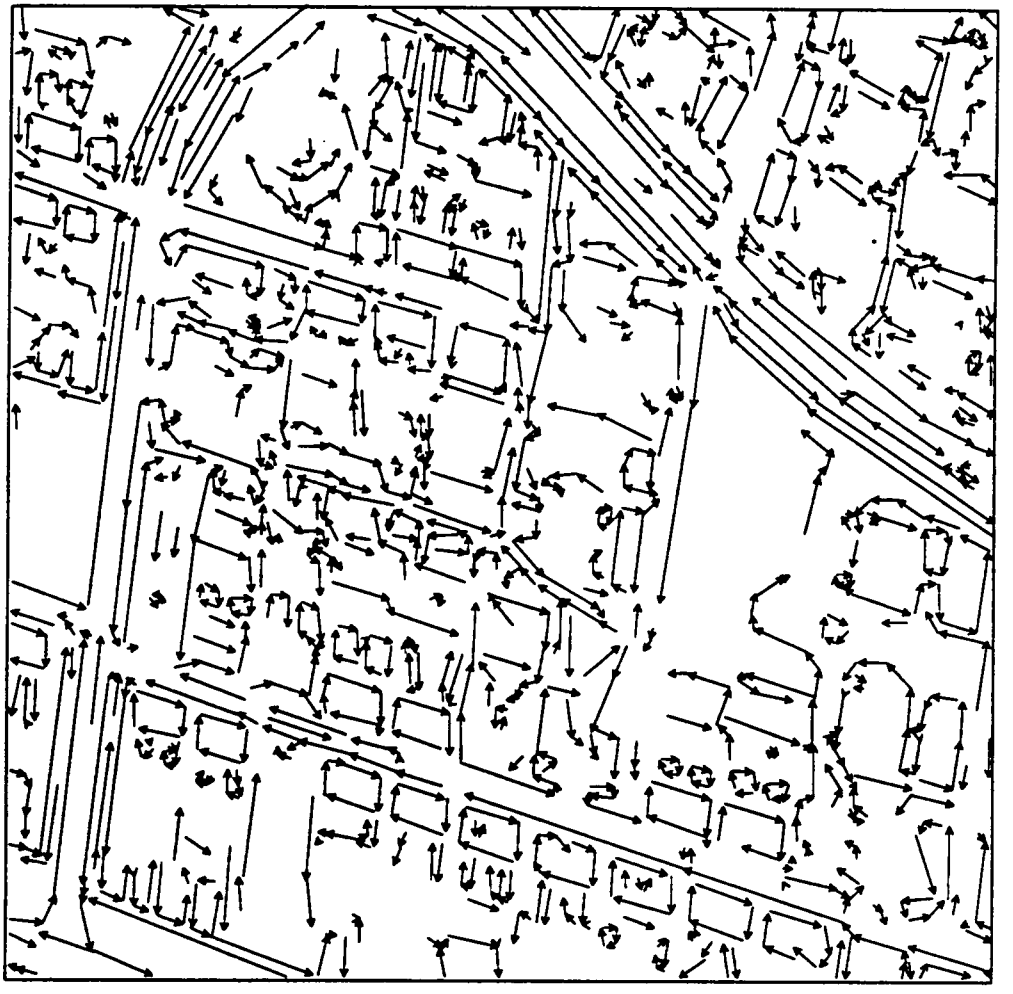


Figure 85. shows the result of iterated end-point partitioning followed by straight line fitting: For each extracted line, the immediate region to the left side of the arrow is darker than the region to the right.

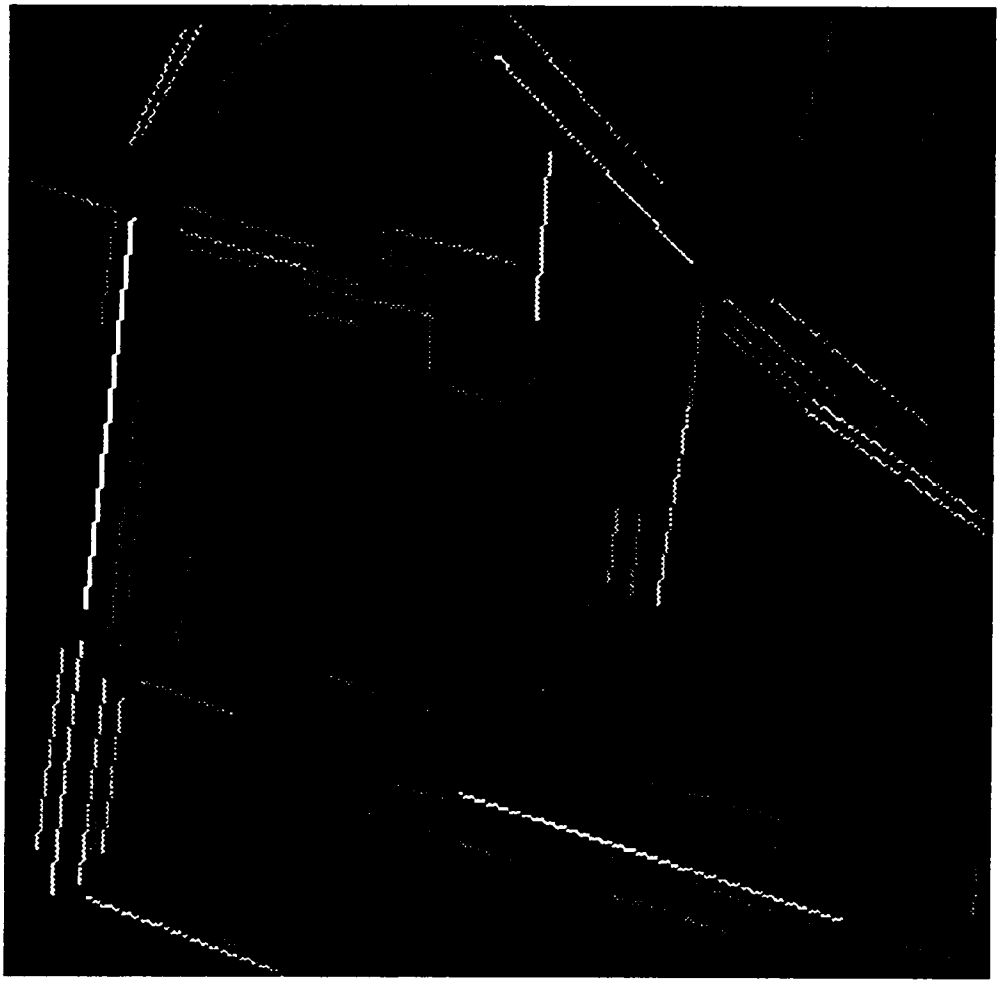


Figure 86. shows the segments with their brightness inversely proportional to their orientation standard deviation.

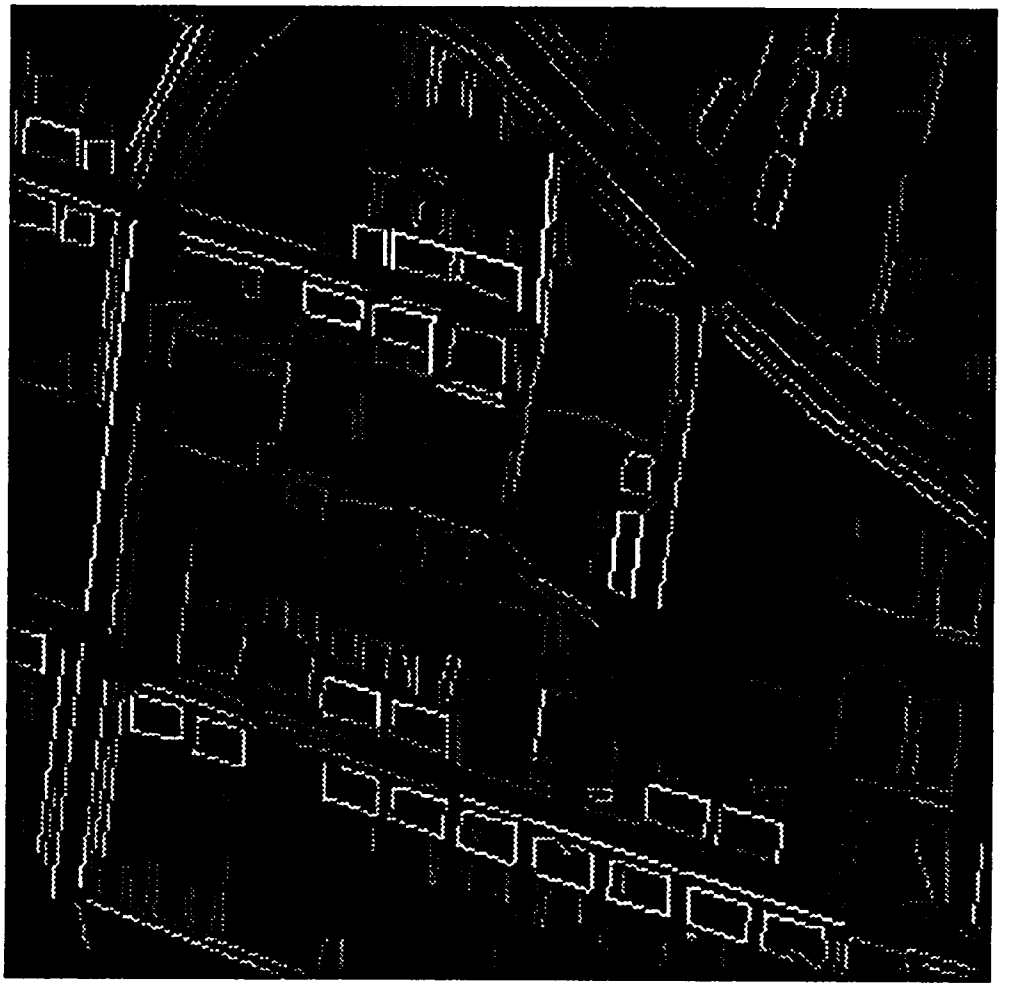


Figure 87. shows the segments with their brightness inversely proportional to their position standard deviation.

Figure 88. shows the thinned edges after non-maxima suppression.

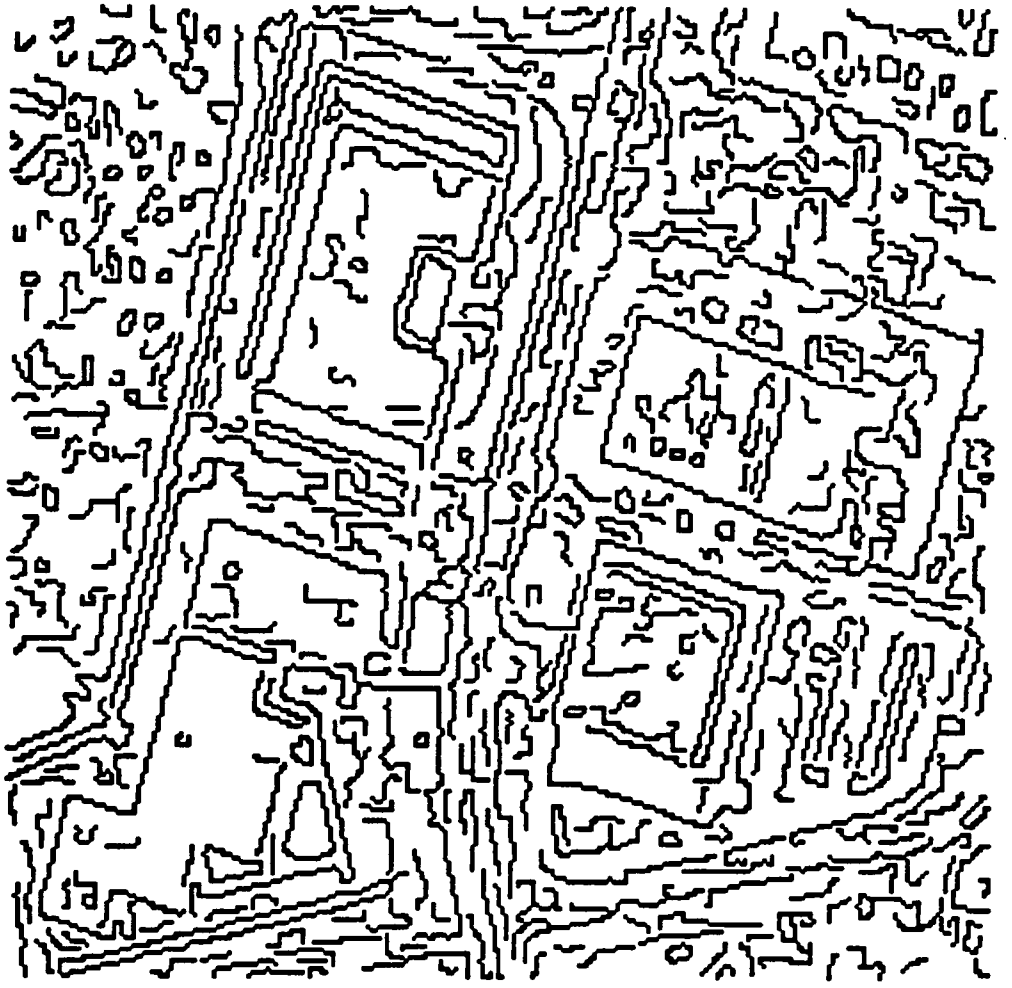


Figure 89. shows the edge segments remaining after pruning.

Figure 90. shows the result of iterated end-point partitioning followed by straight line fitting.

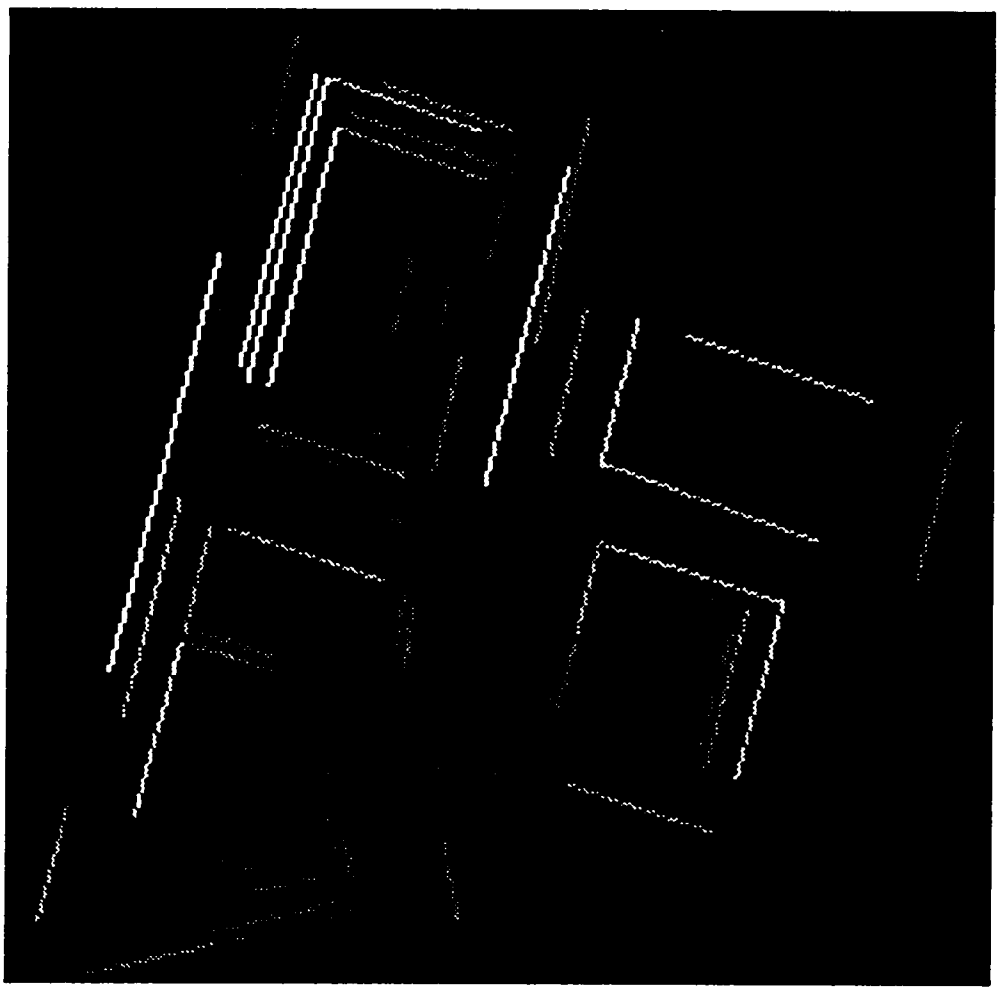


Figure 91. shows the segments with brightness inversely propportional to their orientation standard deviation.

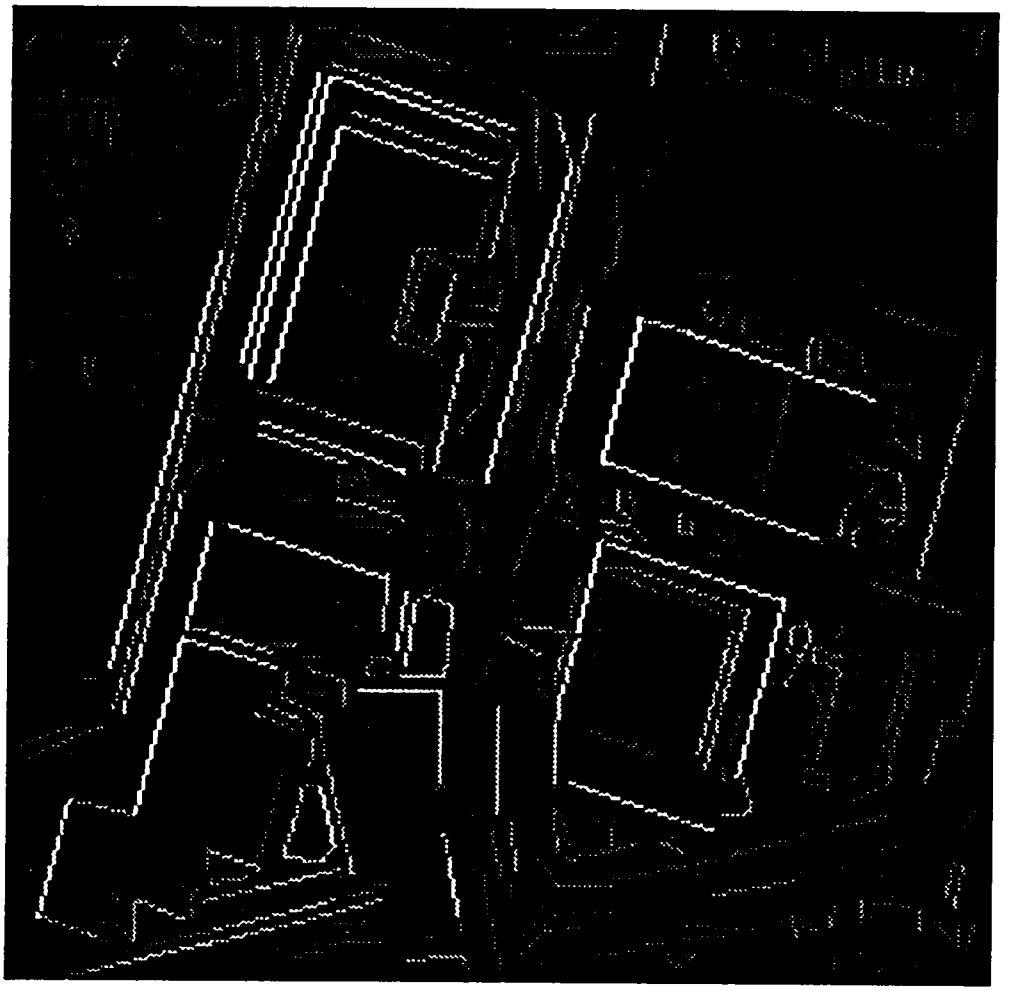


Figure 92. shows the segments with brightness inversely proportional to their position standard deviation.

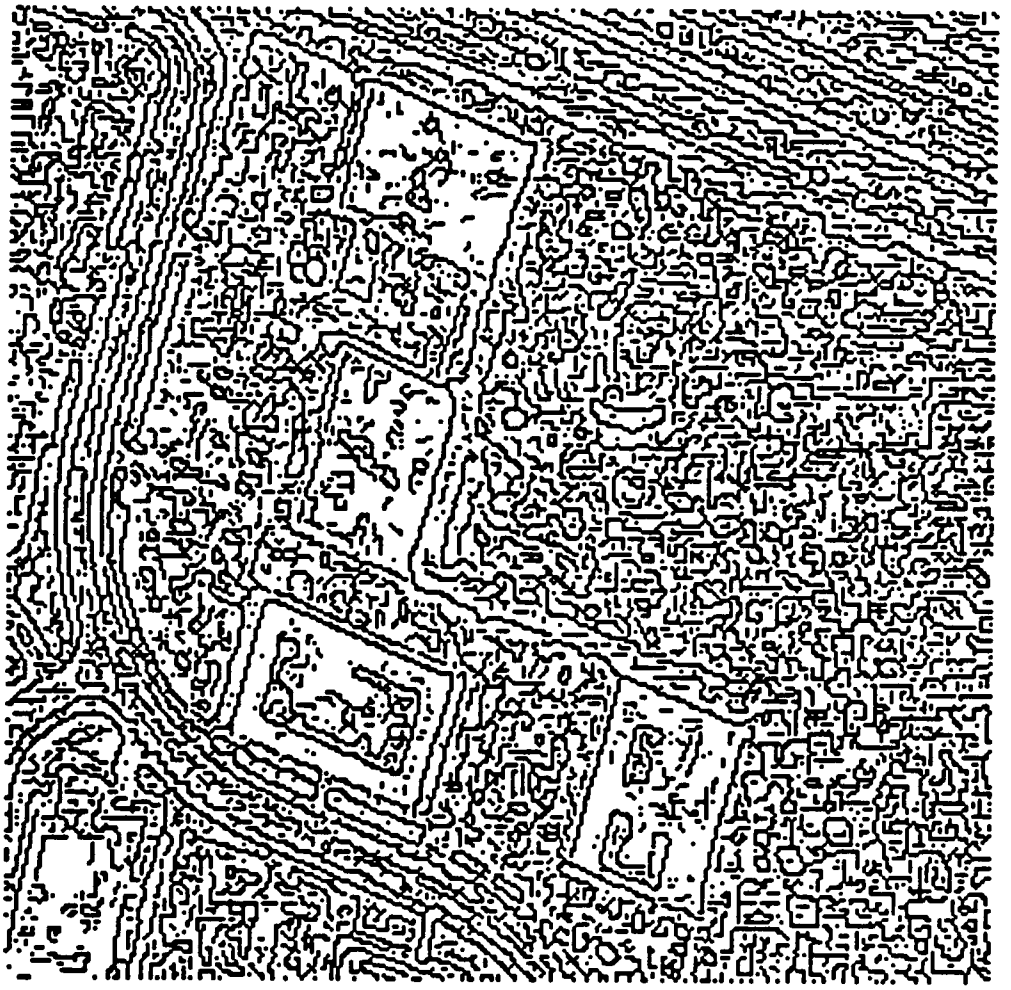


Figure 93. shows the thinned edges after non-maxima suppression.

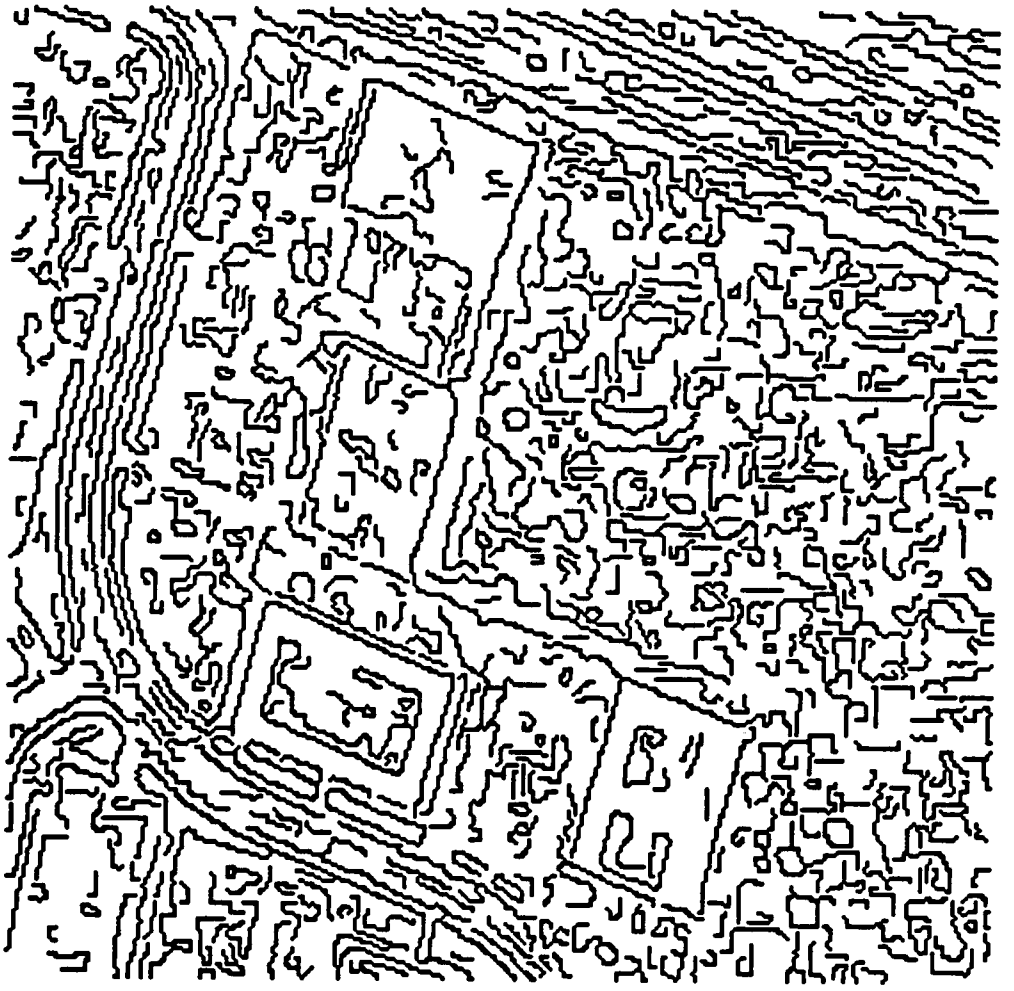


Figure 94. shows the edge segments remaining after pruning.

Figure 95. shows the result of iterated end-point partitioning followed by straight line fitting.

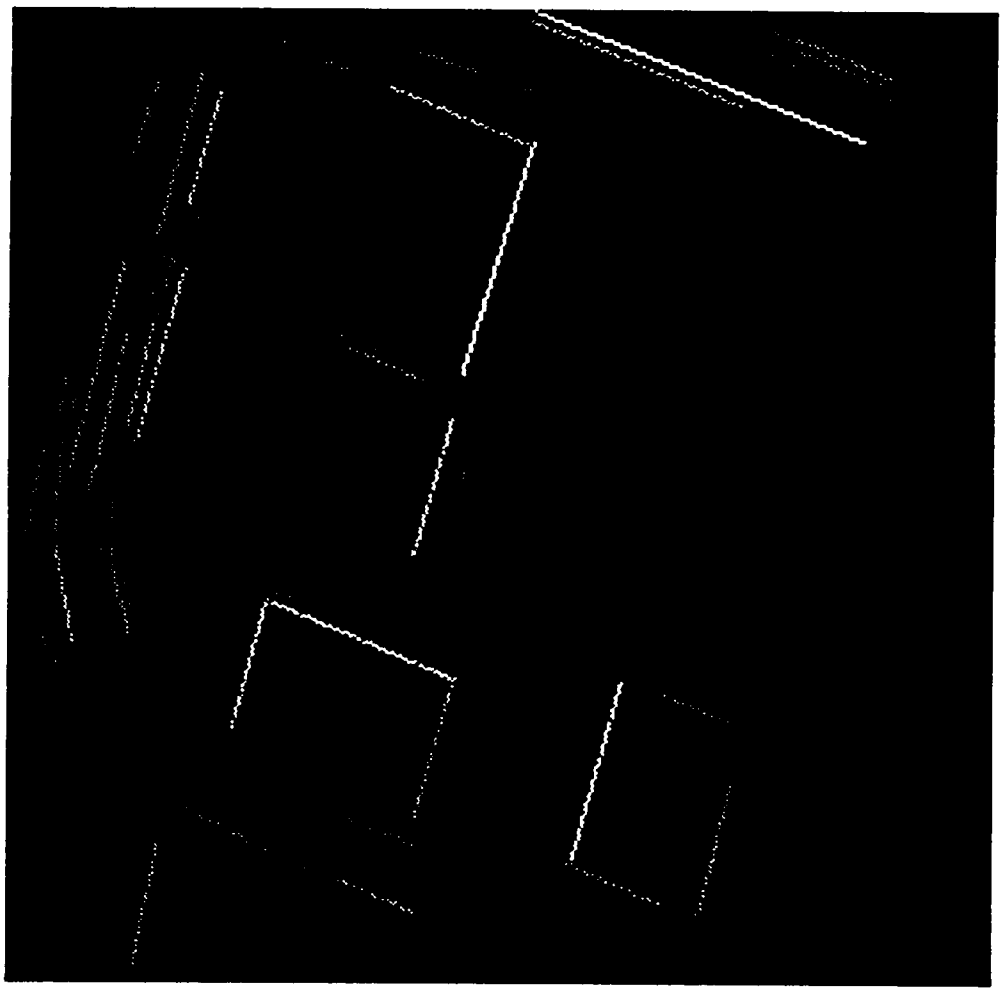


Figure 96. shows the segments with brightness inversely proportional to their orientation standard deviation.

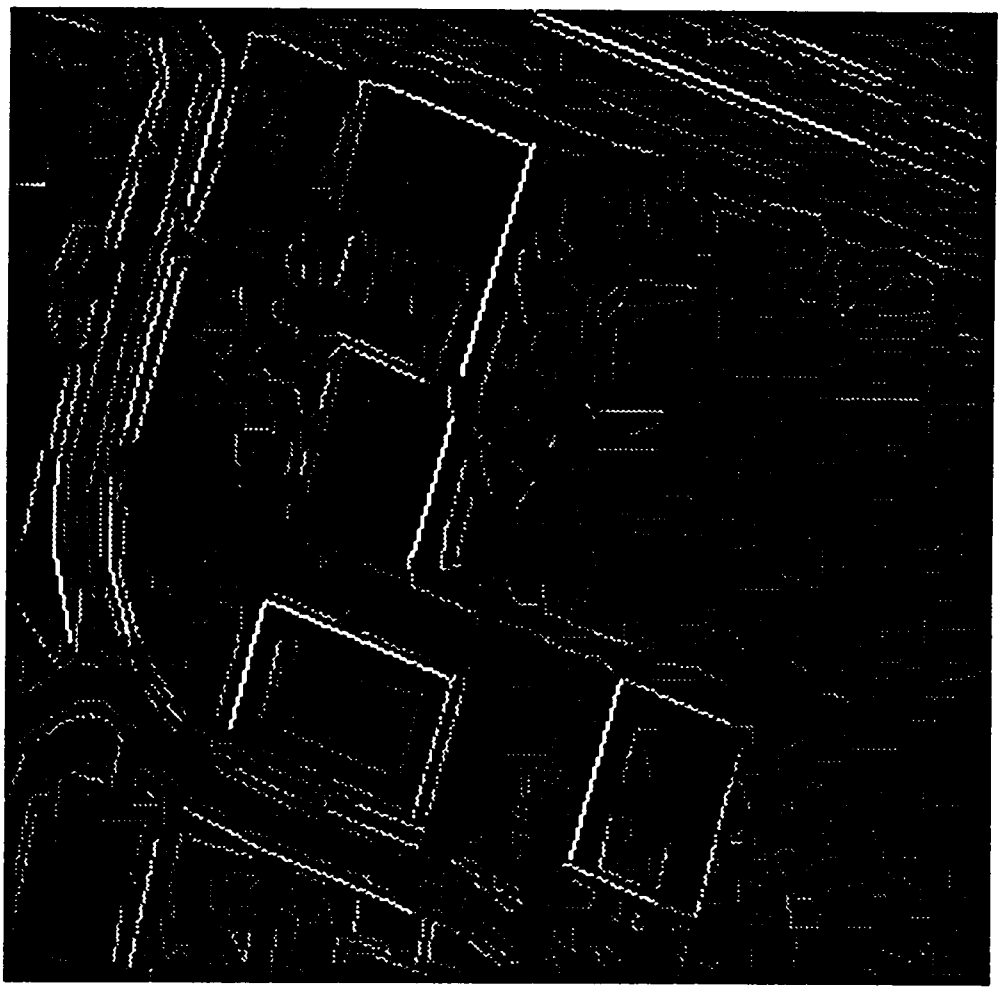


Figure 97. shows the segments with brightness inversely proportional to their position standard deviation.

Chapter 8

PERCEPTUAL ORGANIZATION

8.1 Introduction

A characteristic of man-made structures appearing in aerial images is the presence of a large number of linear segments. These segments are usually highly organized and this organization is manifested primarily in three forms: parallelism, as in the opposite sides of buildings and roads; collinearity, as in a row of houses; and adjacency, as in the corners formed by adjacent sides of buildings.

The highly organized nature of the linear segments in aerial images of urban scenes leads to the construction of a number of intermediate-level units described in this section. The succesful extraction of these primitives requires in many situations little or no top-down flow of information. Succesful matching of aerial images to models of man-made structures can usually take place at this level.

Section 8.2 illustrates methods to extract corners and to test for instances of parallelism and collinearity among pairs of linear segments. The extracted corners are used in section 8.3 as a beginning point for constructing segment chains, ie., ordered lists of linear segments that are possibly parts of the contour of cultural objects such as buildings. This section illustrates the application of these methods to the extraction of cultural features in aerial images. Section 8.4 presents some further examples. Finally section 8.5 presents some conclusions.

8.2 Linear Segment Organization

8.2.1 Corners

Many cultural objects, such as buildings, are characterized by the presence of near orthogonal L-junctions, ie. corners. The extraction of these features provides a useful starting point for the construction of higher level groupings of linear segments corresponding possibly to cultural objects.

Corners are extracted by examining pairs of linear segments satisfying the following constraints (see Figure 98 on page 235) :

1. The ending point of one segment must be within "r" pixels of the beginning point of the other segment.

2. The angle between the segments (module 180) must satisfy

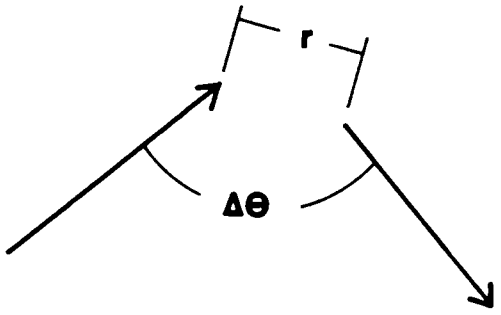


Figure 98. shows a corner relation among two linear segments

$$90 - \varepsilon - z_{\alpha/2}\sqrt{\sigma_{\theta 1}^2 + \sigma_{\theta 2}^2} \;\leq\; \Delta\theta \;\leq\; 90 + \varepsilon + z_{\alpha/2}\sqrt{\sigma_{\theta 1}^2 + \sigma_{\theta 2}^2} \tag{8.1}$$

where ε is typically less than 15 degrees, $\sigma_{\theta 1}^2$ and $\sigma_{\theta 2}^2$ are the variances of the normally distributed segment orientations θ_1 and θ_2, and $\Delta\theta$ is computed as

$$\Delta\theta = |(\theta_1 - \theta_2) \bmod 180| \tag{8.2}$$

The constant $z_{\alpha/2}$ is the upper 50α percentage point of the unit normal distribution. This contraint on the angle between the segments is equivalent to accepting the null hypothesis

$$H_0\colon\ |\Delta\theta - 90| \;\leq\; \varepsilon$$

against the alternative hypothesis

$$H_1\colon\ |\Delta\theta - 90| \;>\; \varepsilon$$

at the 100α percent significance level.

Not every possible pair of segments is tested. Given the ending point of a segment s_i, the search for the beginning point of a candidate segment s_j, forming a corner with s_i , is restricted to a square window centered at the ending point of s_i From contraint 1, the size of the search window is (2r + 1) by (2r + 1) pixels, with the value of r typically chosen between 1 to 3 pixels. A candidate corner pair $s_i s_j$ is assigned a significance equal to:

$$c_{ij} = \frac{1}{d\,|\Delta\theta + z_{\alpha/2}\sqrt{\sigma_{\theta i}^2 + \sigma_{\theta j}^2} - 90|} \tag{8.3}$$

where d is the distance between the segments end points. If more than one corner candidate segment exist, we select the one having the highest corner significance with segment s_i and having no prior corner assigment of higher significance outside the search window.

Following the notation of Huertas and Nevatia (1988), corners are classified as **bright** or **dark** according to whether the average intensity of the region in the concave side of the corner is higher or lower than the average intensity of the region in the convex side.

8.2.2 Parallelism

Deciding if two given segments are parallel can take the form of a hypothesis test. Let θ_1, $\sigma^2_{\theta 1}$, θ_2, $\sigma^2_{\theta 2}$, be the segments orientations and variances. Then testing for parallelism would involve testing the null hypothesis

$$H_0: \ |\Delta\theta| \ \leq \ \varepsilon$$

against the alternative hypothesis

$$H_1: \ |\Delta\theta| \ > \ \varepsilon$$

where $\Delta\theta = (\theta_1 - \theta_2) \bmod 180$ and ε is typically less than 15 degrees.

Since θ_1, and θ_2 are normally distributed, the null hypothesis is accepted if

$$|\Delta\theta| \ \leq \ \varepsilon + z_{\alpha/2}\sqrt{\sigma^2_{\theta 1} + \sigma^2_{\theta 2}} \tag{8.4}$$

at the 100α significance level.

8.2.3 Collinearity

Consider the two segments shown in Figure 99 on page 239 with angular and positional parameters (θ_1, ρ_1), and (θ_2, ρ_2), and parameter variances $\sigma^2_{\theta 1}$, $\sigma^2_{\rho 1}$, $\sigma^2_{\theta 2}$, $\sigma^2_{\rho 2}$. Let ρ_1, and ρ_2 be measured with respect to a line L passing through the combined center of mass of the segments. This center of mass $(\bar{r}, \bar{c})$ is given by

$$\bar{r} = \frac{\bar{r}_1/\sigma^2_{\rho 1} + \bar{r}_2/\sigma^2_{\rho 2}}{1/\sigma^2_{\rho 1} + 1/\sigma^2_{\rho 2}} \tag{8.5}$$

$$\bar{c} = \frac{\bar{c}_1/\sigma^2_{\rho 1} + \bar{c}_2/\sigma^2_{\rho 2}}{1/\sigma^2_{\rho 1} + 1/\sigma^2_{\rho 2}}$$

where $(\bar{r}_1, \bar{c}_1)$ and $(\bar{r}_2, \bar{c}_2)$ are the centers of mass for each segment. Let the orientation θ of the line L be given by

$$\theta = \tan^{-1} \frac{\sin\theta_1/\sigma^2_{\theta 1} + \sin\theta_2/\sigma^2_{\theta 2}}{\cos\theta_1/\sigma^2_{\theta 1} + \cos\theta_2/\sigma^2_{\theta 2}} \tag{8.6}$$

Then the separation s between the segments in a direction perpendicular to the line L is given by

$$s = \Delta\rho = |\rho_1 - \rho_2| \tag{8.7}$$

Furthermore, let the gap g between the segments in the direction along the line L be greater than zero.

Deciding if the two segments are collinear involves two hypothesis tests. One hypothesis test is used for deciding if the segments have similar orientations. This test is identical to the one described in the previous section. The other hypothesis test is used

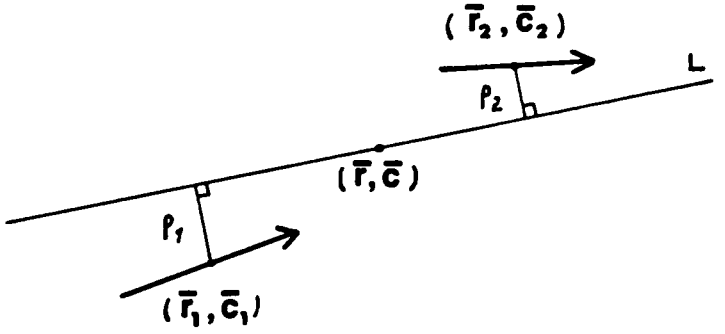


Figure 99. shows a collinearity relation among two linear segments

to decide if the lateral separation s between the segments is small enough. The null hypothesis is

$$H_0: s \leq \gamma$$

and the alternative hypothesis is

$$H_1: s > \gamma$$

where γ is typically less than 3 pixels.

The null hypothesis is accepted at the 100α significance level if

$$s \leq \gamma + z_{\alpha/2}\sqrt{\sigma_{\rho 1}^2 + \sigma_{\rho 2}^2} \tag{8.8}$$

8.3 *Detecting cultural features - an illustration*

The purpose of this section is to illustrate how the pairwise relationships among linear segments discussed in the previous section can be used to construct higher level primitives that are likely to correspond to cultural objects. The method used is based on the assumption that the objects of interest are rectangular or composed of rectangular components. Furthermore, any two adjacent object sides must meet at a near orthogonal angle.

The procedure consists of the following steps:

1. Extract corners between pairs of linear segments using the procedure described in section 8.2.1.

2. Extract segment chains, ie., ordered lists of linear segments representing possibly contours of cultural objects. Initially, the extracted chains are such that any two consecutive segments on a chain are joined by a corner.

3. Expand each open segment chain by adding segments at each end which are either collinear ot at 90 degrees with the respective end segments.

4. Continue expanding the chains until either we obtain a closed chain or no further expansion is possible.

5. Fit the resulting closed chains to the cultural object model.

This procedure, although simple, will be shown to be useful in detecting some cultural objects.

8.3.1 Corners and initial segment chains.

Figure 100 on page 249 shows a 256x256, 8 bits aerial image containing five buildings of varied shapes ansd sizes. Figure 101 on page 250 shows the linear segments extracted using the procedure described in chapter 7. The extracted corners are shown in Figure 102 on page 251. These corners were obtained using a 7x7 search window, and a hypothesis test on corner angle at the 5 percent significance level, with ε equal to 7 degrees. Figure 103 on page 252 shows the initial segment chains formed by these corners.

A segment chain is said to be **bright** if the number of bright corners in it is no less than the number of dark corners. Otherwise the segment chain is said to be **dark**. Bright (dark) segment chains are assumed to belong to bright (dark) cultural objects.

8.3.2 Expansion of chains

Cultural objects often appear in orientation clusters. A histogram of the orientation of segments belonging to segment chains reveals the orientation of the most prevalent clusters. For illustration purposes we will consider only the highest peak, and the peak nearly 90 degrees apart from the highest, in the smoothed histogram. In obtaining the histogram, the segment orientations are weighted by the inverse of their variances and orientations which are 180 degrees apart are clustered together. The histogram is smoothed using a moving average window of pre-specified odd width. Figure 104 on page 253 shows the smoothed histogram obtained using a window 5 degrees wide. The program correctly selects the peaks at -71 and 14 degrees. Since orientations 180 degrees apart were clustered together, the other two orientations of interest are -166 and 109 degrees. These four orientations correspond to the orientations of the contours of cultural objects in the most prevalent cluster, according to our model. Furthermore, these orientations guide the search for further segments in a contour starting from the initial segment chains.

Any segment s_i whose orientation θ_i falls within $\pm (\varepsilon + z_{\alpha/2}\sigma_{\theta i})$ degrees from a cluster orientation θ_A is assigned to that cluster and its positional parameter ρ_i is recomputed as

$$\rho_i = \bar{r}_i \sin \theta_A + \bar{c}_i \cos \theta_A \tag{8.9}$$

where $(\bar{r}_i, \bar{c}_i)$ represent the segment center of mass. Finding segments collinear to s_i becomes then straightforward. We search for segments in the same orientation cluster as s_i and whose lateral separation from s_i is small enough according to the hypothesis test of section 8.2.3. Notice that this lateral separation among segments is evaluated as the difference of their recomputed position parameters from equation (8.9). The search for collinear segments can be made efficient by sorting the segments belonging to the same orientation cluster according to their recomputed position parameters.

The initial segment chains are ranked according to the sum of the inverse variances of their segments parameter estimates. Chain expansion takes place in decreasing order of their ranks.

Expansion of a segment chain proceeds by examining each end of the chain and searching for candidate segments to continue the contour. According to our model for cultural objects, the candidate segment may appear in any of three orientations: one collinear with the end segment in the chain, one at a 90 degrees angle with it, and one at a -90 degrees angle. These orientations correspond to one of the four orientations of interest obtained from the smoothed segment orientation histogram.

Candidate segments are restricted to be within a certain distance from the end point of the segment being expanded. This distance is a function of the expected size of the cultural object. An estimate of this size is the current longest dimension of the chain and may be updated every time a new segment is added to it.

Candidate segments may compete for the right of expanding a chain. Which candidate is chosen is based on a number of simple heuristics. These include:

1. A candidate segment that closes the chain is favored against one who does not.

2. A collinear candidate segment is favored against a candidate segment that forms a corner.

3. For hypothesized bright cultural objects, a candidate segment forming a bright corner with the chain is favored against a candidate that forms a dark corner with it. The opposite holds for hypothesized dark cultural objects.

4. Among competing corner candidate segments, the one with highest corner significance as evaluated by equation (8.3) is chosen.

5. A bright corner candidate segment must not belong to a dark segment chain.

6. A candidate segment must not increase the size of the chain beyond a given threshold which is a function of the expected cultural object size.

7. A candidate segment must not belong to another chain unless it is the first segment in that chain if forward expansion is taking place, or the last segment in that chain in the case of backward expansion.

Forward and backward expansion continue until one of the following conditions exist:

1. The chain is closed.

2. No further candidates for expansion are found.

3. The sum of the interior angles in the chain is 360 degrees and the distance between the beginning and ending point of the chain is smaller than a given threshold, typically less than 10 pixels.

The above segment expansion heuristics tend to produce closed contours with few non-convexities. Figure 105 on page 254 shows the closed segment chains that result after expansion.

8.3.3 Fitting closed chains

Each side of a closed chain is fitted to a straight line. The center of mass $(\bar{r}, \bar{c})$ of the fitted line segment is computed using formulas similar to those given in appendix E:

$$\bar{r} = \frac{\sum_i \bar{r}_i / \sigma_{\rho i}^2}{\sum_i 1/\sigma_{\rho i}^2}$$

$$\bar{c} = \frac{\sum_i \bar{c}_i / \sigma_{\rho i}^2}{\sum_i 1/\sigma_{\rho i}^2}$$

In these formulas, $(\bar{r}_i, \bar{c}_i)$ represents the center of mass of an individual line segment that participates in the fitting of one side of the closed chain. The symbol $\sigma_{\rho i}^2$ represents

the variance of the position estimate of the line segment. The orientation of the fitted line is given by the cluster orientation.

Figure 106 on page 255 shows the fitted segment chains. Notice, that the cultural object detection algorithm succeeds in detecting all five of the buildings present on this scene. Furthermore the boundaries are well placed and they have the correct orientation.

8.4 Discussion and further examples

Figure 107 on page 256 shows another aerial scene containing four buildings. This image is 300 x 300 pixels and 8 bits. Figure 108 on page 257 through Figure 113 on page 262 show the results of processing this image. Three of four buildings have been correctly extracted.

Casual observation of the buildings present in the two aerial scenes shown in Figure 100 on page 249 and Figure 107 on page 256 produces a perception of bright "box" like objects, bounded by straight edges meeting at 90 degrees corners. Closer inspection, however, reveals some departures from these idealized expected shapes. Some building sides do not appear straight but jagged, some building sides merge with the surrounding background, many corners appear rounded.

The non-ideal building shapes are caused, in part, by limitations of image resolution, poor contrast, shading effects, and the presence of small objects nearby building sides. These inherent limitations on the quality of the image data, together with inadequacies of the low-level vision techniques, and line extraction algorithm conspire to produce a less than ideal line segmentation (see Figure 101 on page 250 and Figure 108 on page

257) in which building contours contain gaps, some corners are missing, and some of the extracted lines have inaccurate orientations.

In spite of these difficulties, the cultural object detection algorithm succeeds in detecting and accurately delineating eight out of nine buildings present in the two aerial scenes (see Figure 106 on page 255 and Figure 113 on page 262). Three factors contribute to this relatively good performance:

1. Use of a strong perceptual clue to initiate the tracing of building contours: the initial segment chains.

2. Full use of the statistical distributions of the parameter estimates of the extracted segments to test instances of corners, collinearity, and parallelism among the segments and to rank them so that contour tracing proceeds from the most significant to the least significant segment chains.

3. Use of a explicit model for cultural object shape to direct the contour finding process.

The cultural object detection algorithm, as it stands, has one major limitation that can restrict its applicability to more complex aerial scenes or to images of lower quality than the ones illustrated. This limitation resides on the fact that decisions about which edges to link into a contour, after initial segment chains have been extracted, are made on the basis of very local information and that these decisions are final (no backtracking is implemented). This explains in part the failure of the algorithm to correctly detect the top building in the second aerial scene illustrated (see Figure 113 on page 262). A more robust algorithm would involve implementing the contour tracing as a tree search process similar to that described by Brooks (1981). Since this tree search can

be computationally expensive, the same heuristics previously used for segment chain expansion can be used for tree pruning in order to reduce the search space.

8.5 Conclusions

We have illustrated methods to hypothesize pairwise relationships among linear segments, namely, corners, parallelism, and collinearity in aerial images containing cultural structures. The methods make use of our knowledge of the statistical distributions of the linear segments parameter estimates.

We have also illustrated how these relationships can be used to construct higher level groupings of linear segments that correspond to cultural objects modeled as being rectangular or composed of rectangular components.

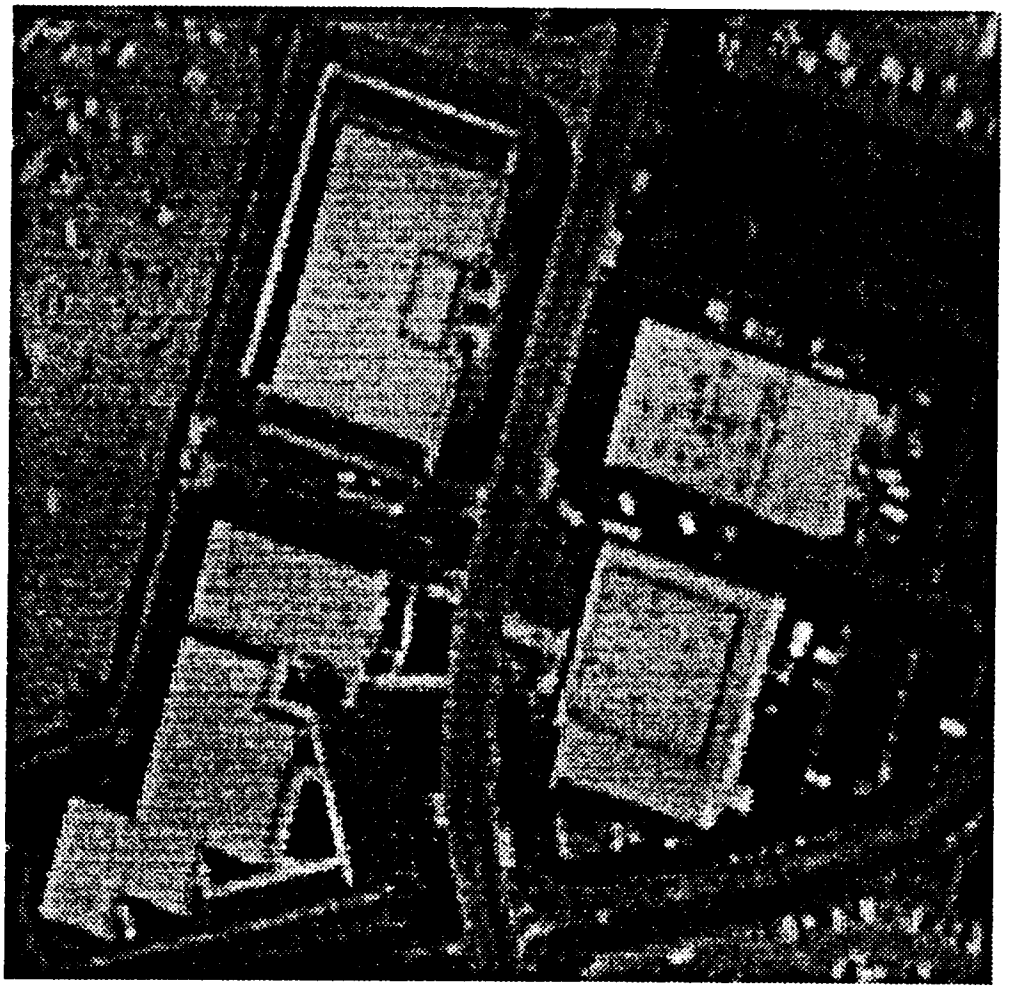


Figure 100. shows an aerial image.

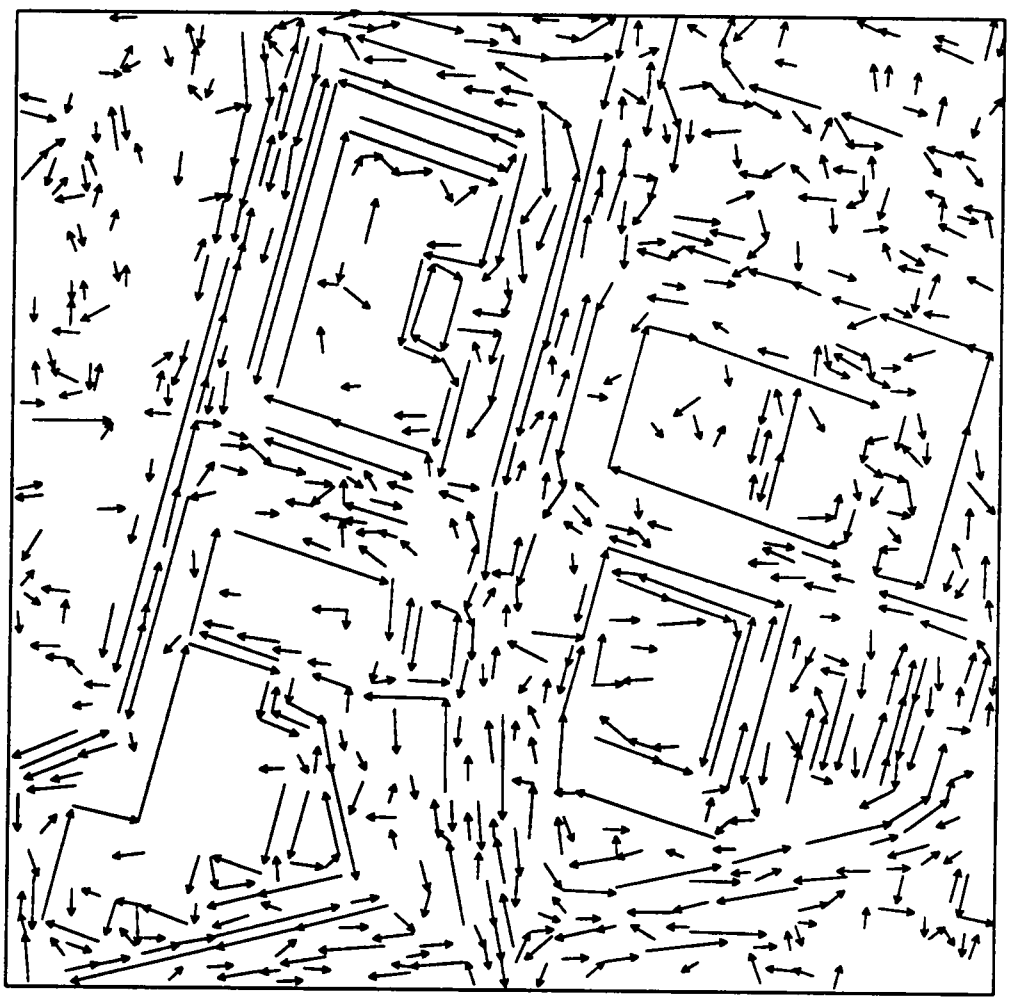


Figure 101. shows the extracted linear segments.

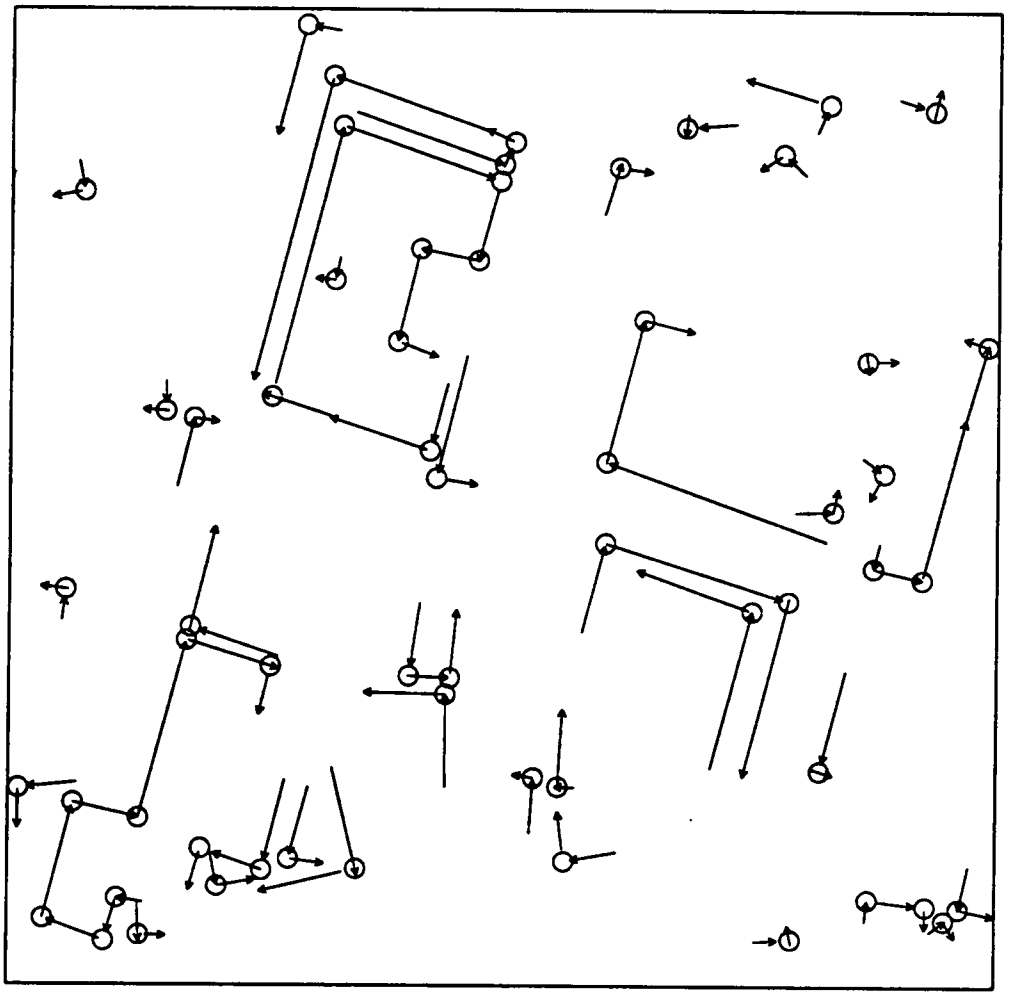


Figure 102. shows the extracted corners.

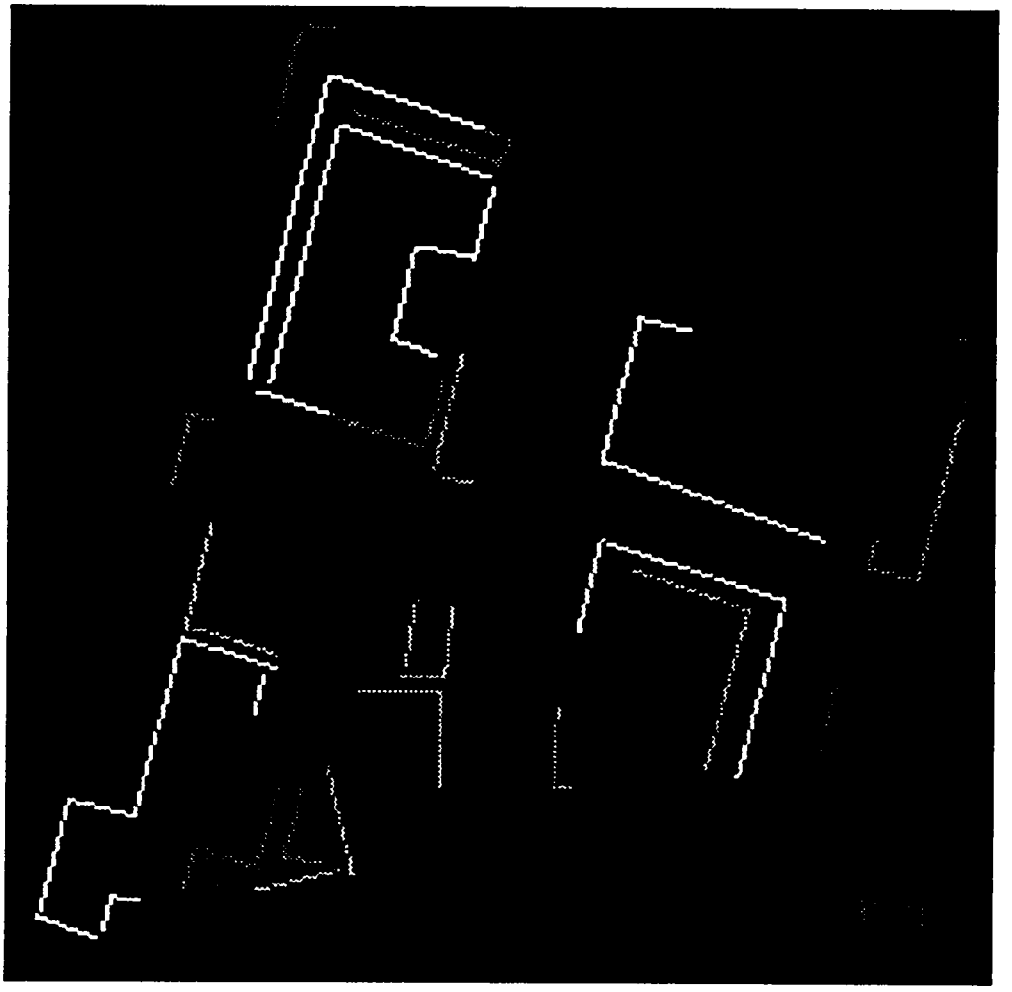


Figure 103. shows the initial segment chains with brightness proportional to their ranks.

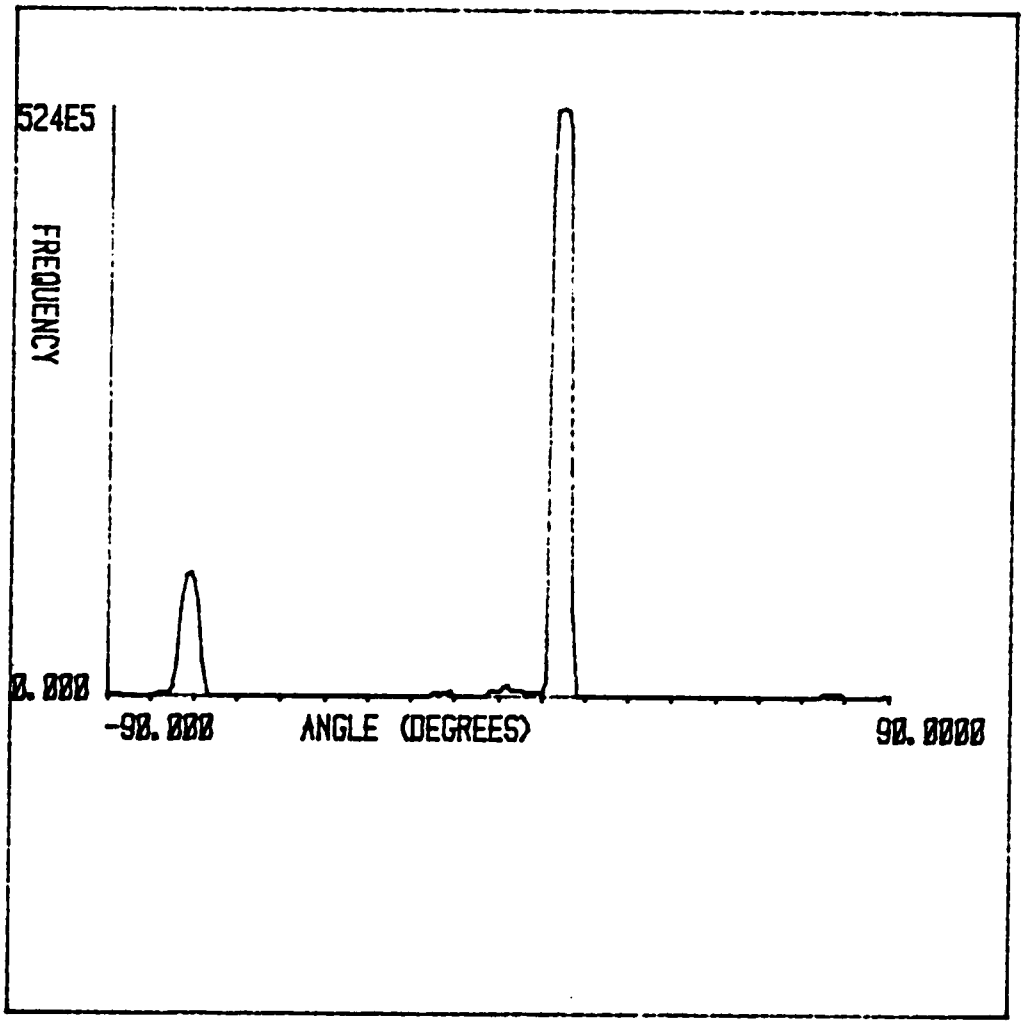


Figure 104. shows the smoothed orientation histogram of segments belonging to initial segment chains.

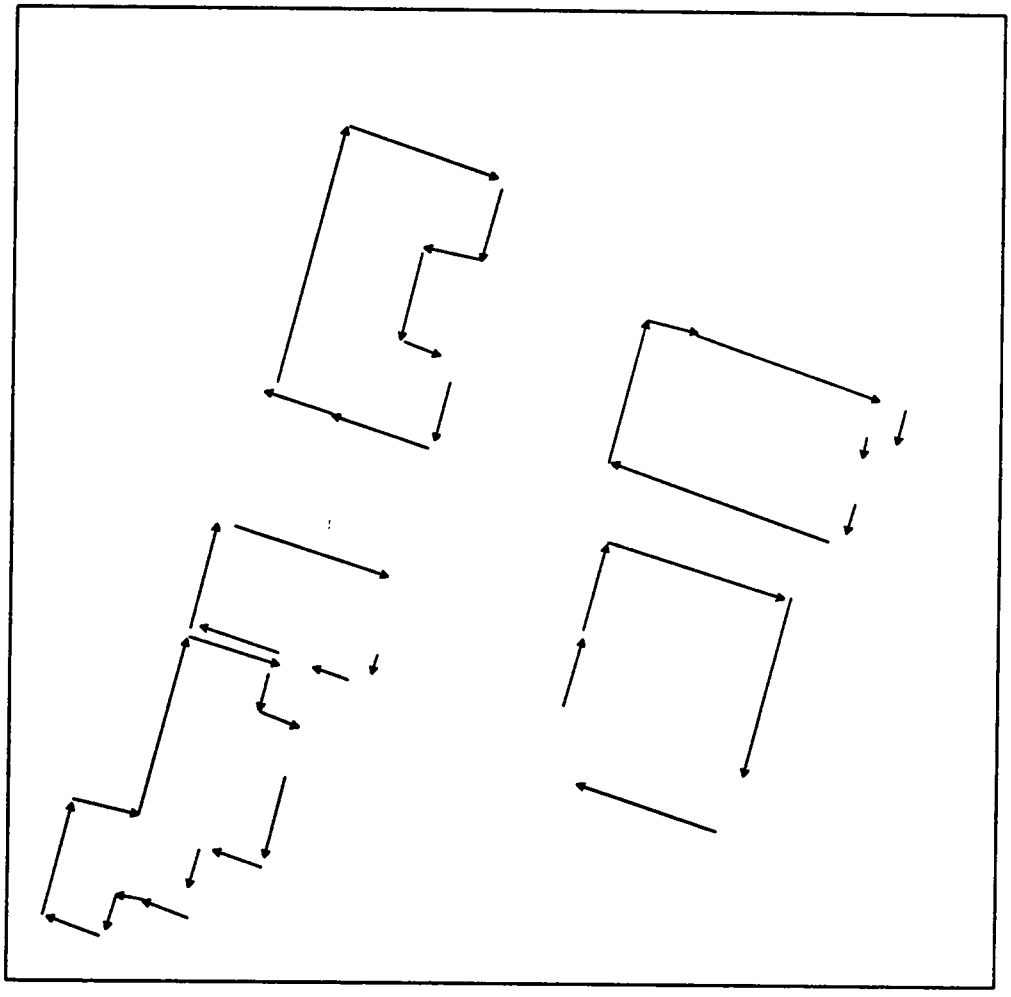


Figure 105. shows the expanded closed chains.

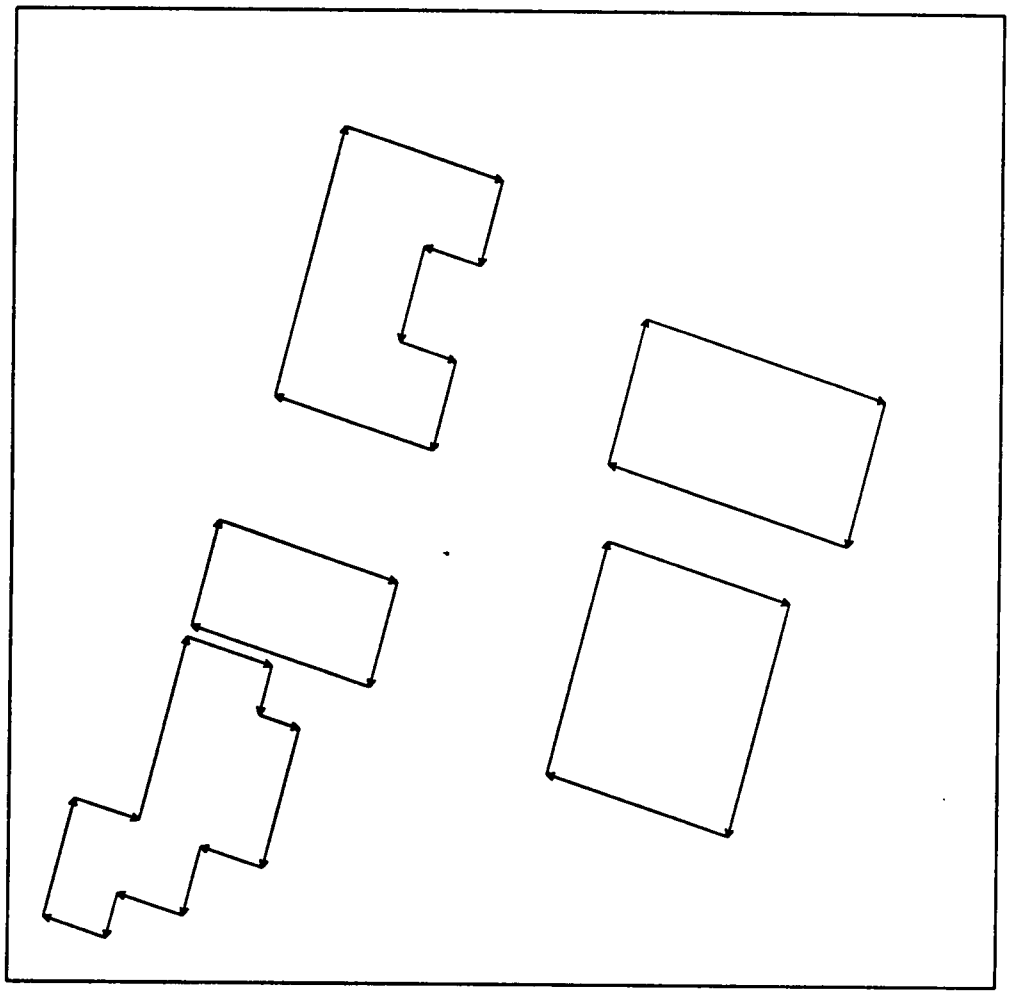


Figure 106. shows the fitted closed chains.

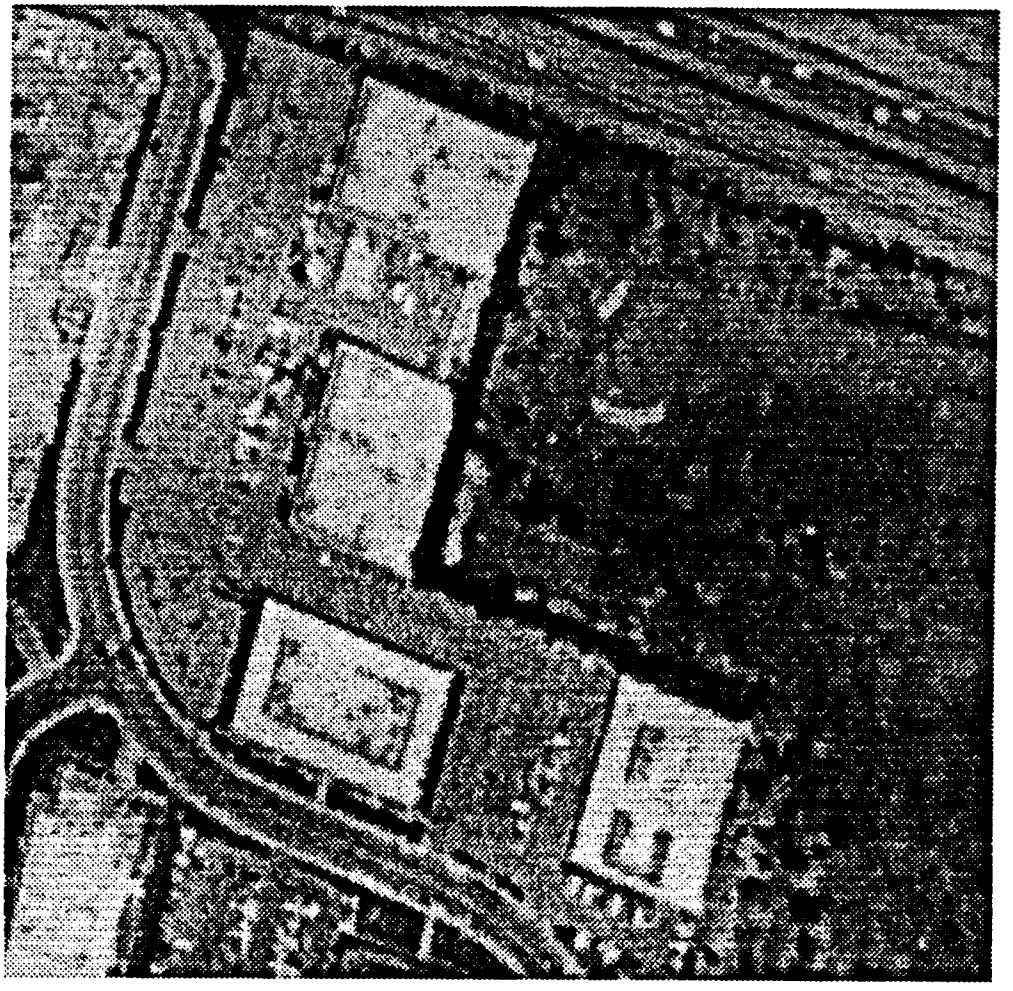


Figure 107. shows another aerial image.

Figure 108. shows the extracted linear segments.

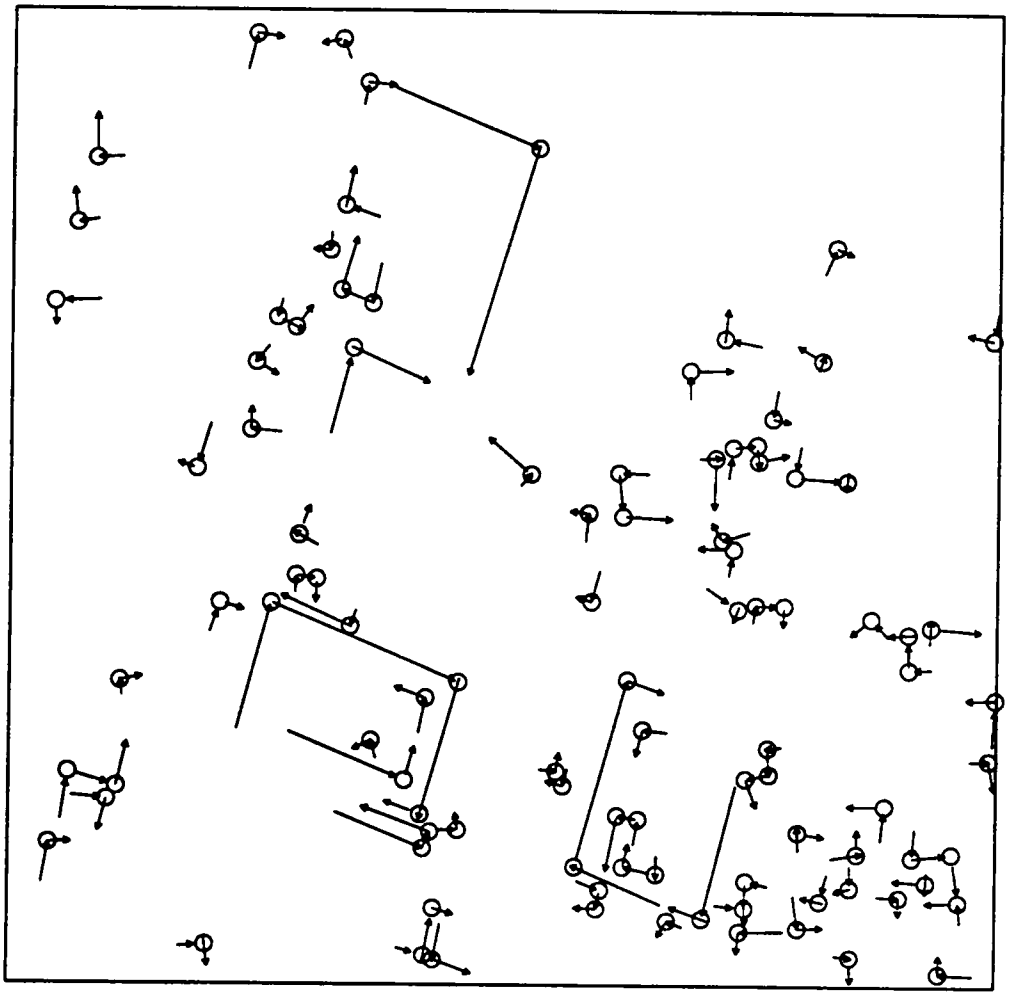


Figure 109. shows the extracted corners.

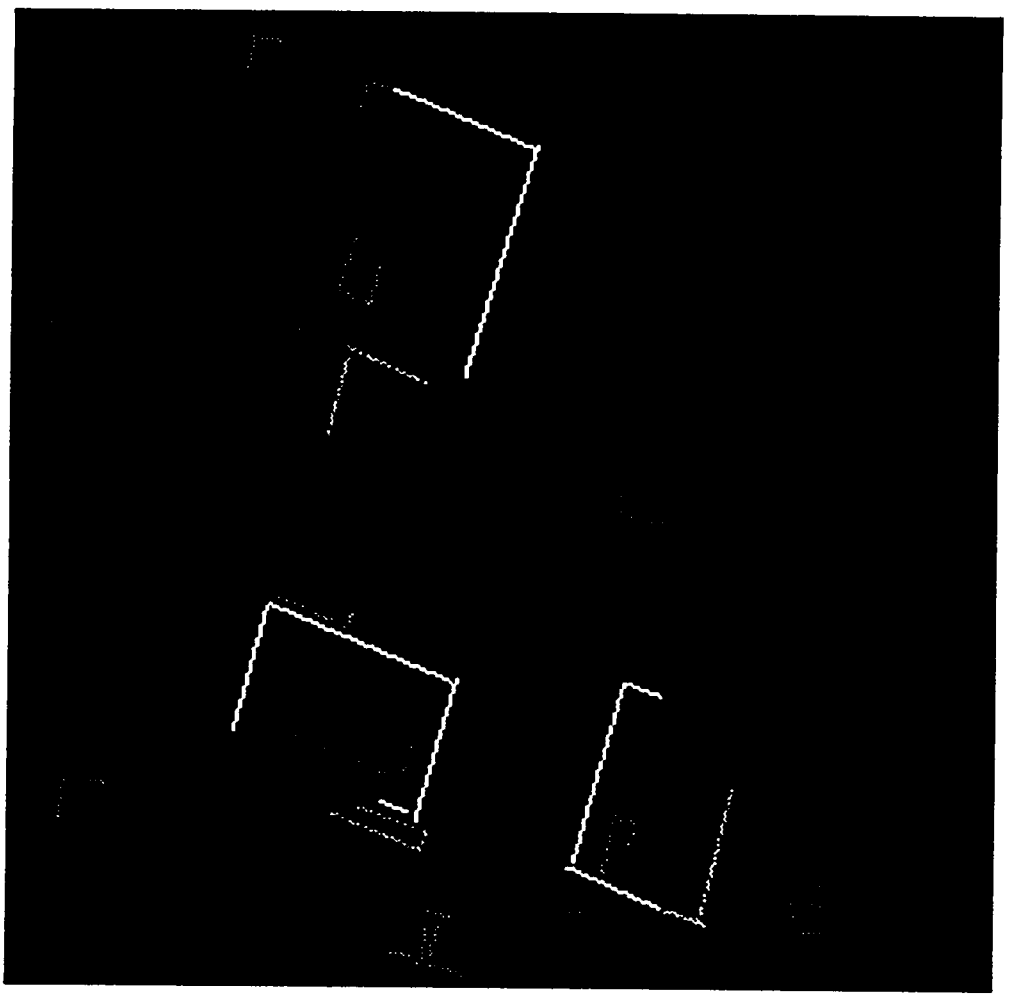


Figure 110. shows the initial segment chains with brightness proportional to their ranks.

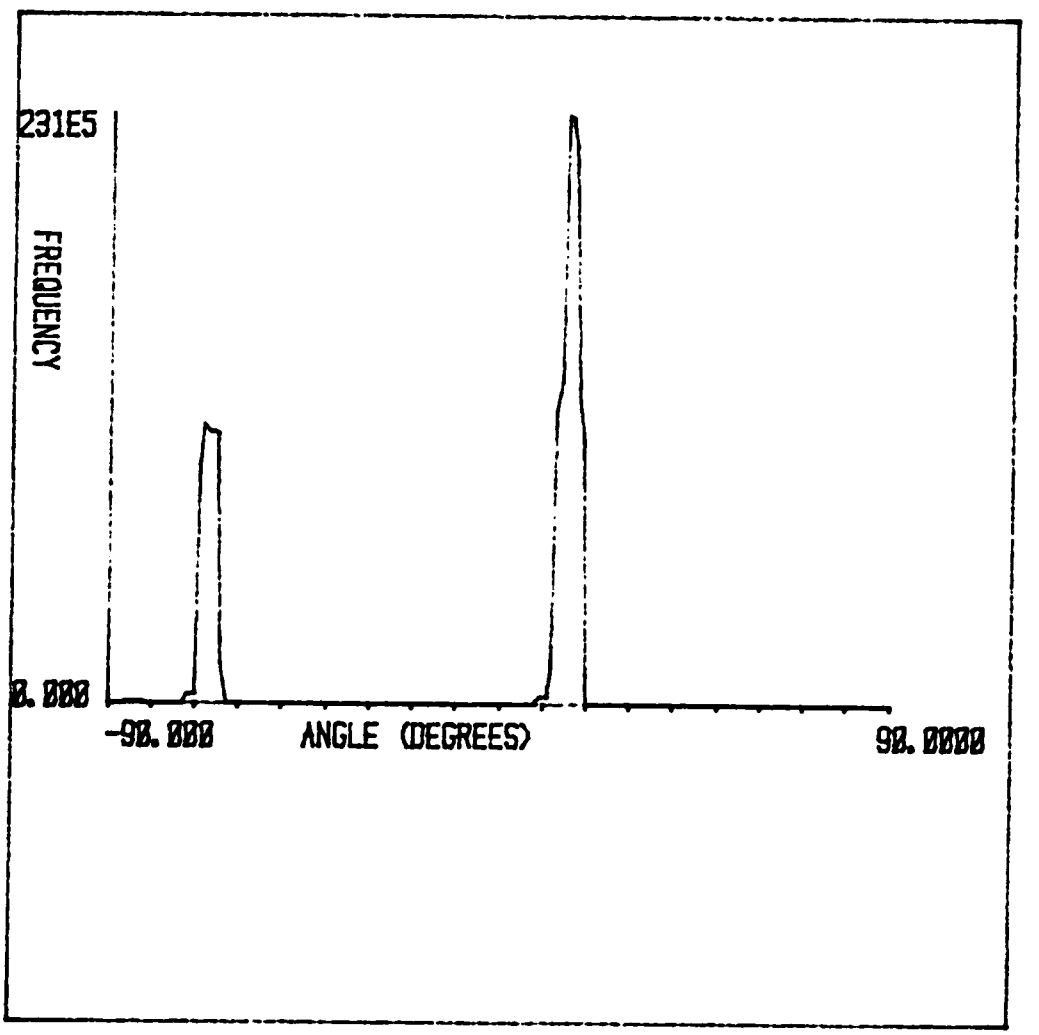


Figure 111. shows the smoothed orientation histogram of segments belonging to initial segment chains.

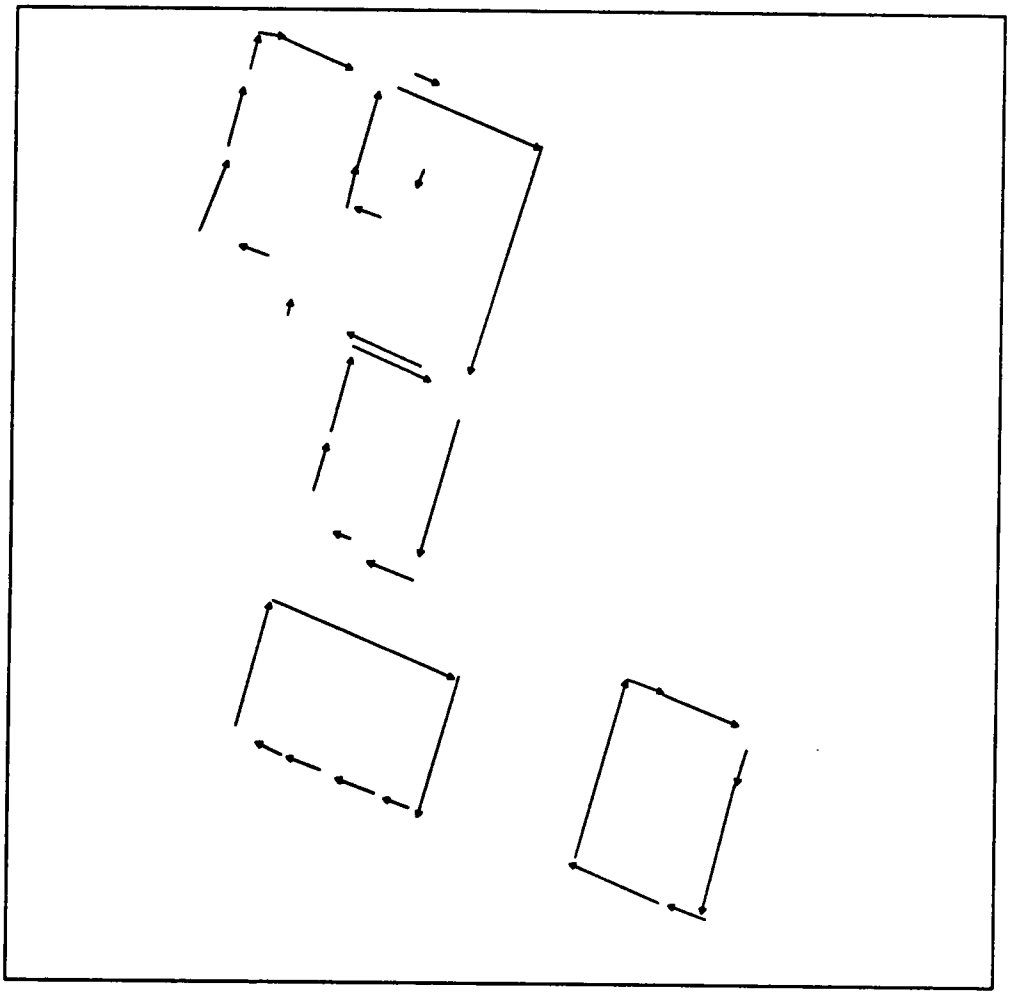


Figure 112. shows the expanded closed chains.

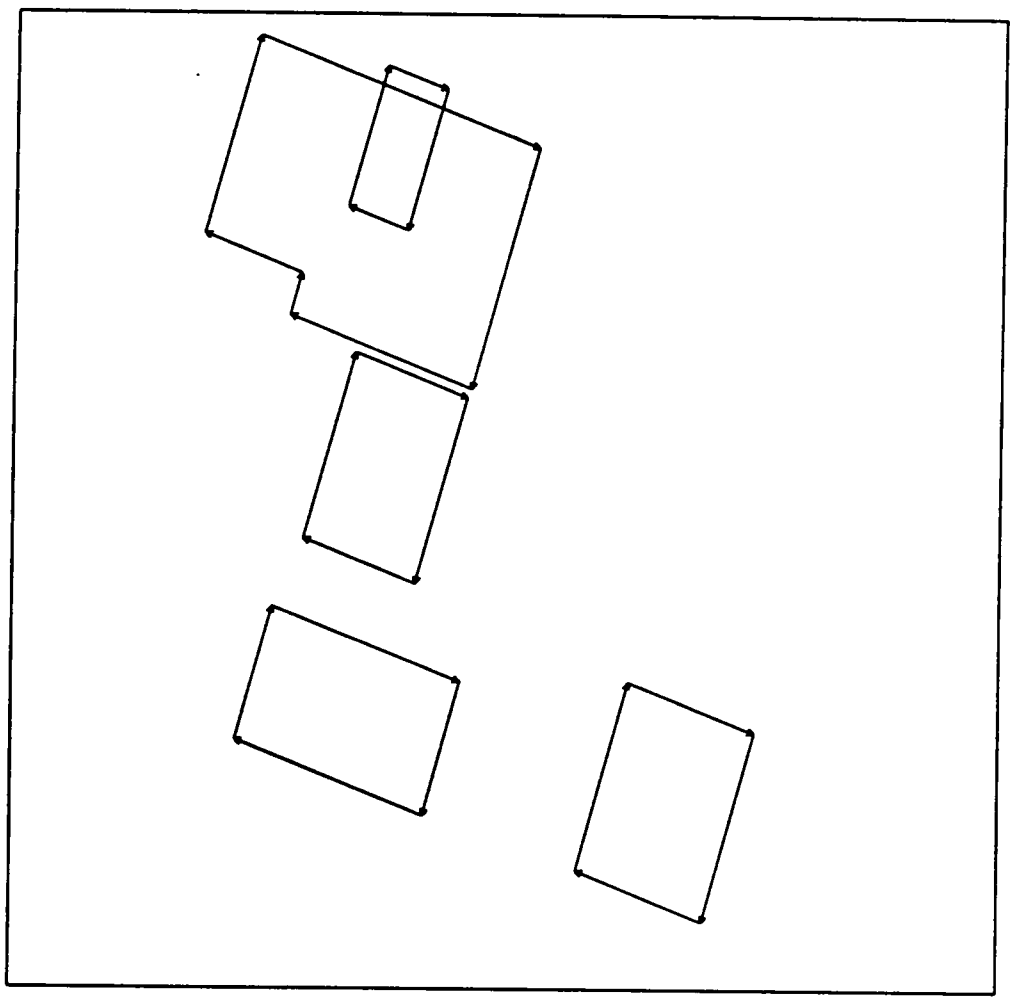


Figure 113. shows the fitted closed chains.

Chapter 9

CONCLUSIONS

A number of low-level and intermediate-level computer vision modules have been developed for the purpose of the construction of an Edge-based representation that makes explicit linear features and their spatial relationships. Although most of these methods are domain independent, their performance has been illustrated in aerial images of urban scenes containing man-made structures.

The techniques used are based on a common structural and statistical model of the image data. This model provides an effective way of designing the algorithms and judging the results.

This dissertation has made a number of contributions to low-level and intermediate-level computer vision research. These include:

1. A gradient operator based on the integrated directional derivative which reduces sharply the gradient direction bias that plagues current operators while reducing also sensitivity to noise. Furthermore the new operator is relatively insensitive to edge

or line orientation and its step edge response decreases smoothly to zero as edge displacement increases, preventing the creation of some false zero-crossings.

2. A Bayes decision procedure for automatic gradient threshold selection that produces results which are superior to those obtained by the best subjective threshold.

3. An edge detector based on Haralick's zero-crossing of second directional derivative of a cubic facet. Integration of the new gradient operator and Bayesian gradient threshold selection into the edge detector results in improved detection and localization of zero-crossings.

4. A graytone corner detector with significantly better probability of correct corner assignment than other corner detectors available in the literature.

5. A linear segment extraction method that uses both, zero-crossing positional and angular information together with their distributional characteristics to accomplish optimal linear segment fitting.

6. Methods for hypothesizing corners and relations of parallelism and collinearity among pairs of linear segments and evaluating their significance. Furthermore, these relations are used to construct higher level groupings of linear segments corresponding possibly to cultural objects. Illustrations are presented in the domain of aerial images of urban scenes where the objects of interest are modeled as being rectangular or composed of rectangular components.

BIBLIOGRAPHY

Abdou, I. and W. Pratt, "Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors", *Proceedings of the IEEE* **67**, May, 1979.

Agin, G. and T. Binford, "Computer Descriptions of Curved Objects", *Proceedings of the Third International Joint Conference on Artificial Intelligence*, pp.629-635, 1973.

Arcelli, C. and G. Sanniti de Baja, "A Thinning Algorithm Based on Prominence Detection", *Pattern Recognition* **13**, pp.225-235, 1981.

Baird, M.L., "Image Segmentation Technique for Locating Automotive Parts on Belt Conveyors", *Proc. 5th Int. Joint Conf. Artificial Intelligence*, pp.694-695, 1977.

Ballard, D.H., "Generalizing the Hough Transform to detect arbitrary shapes", *Pattern Recognition* **13**, pp.111-122, 1981.

Barnett, V., "Comparative Statistical Inference", *Wiley Series in Probability and Mathematical Statistics*, 1982.

Barrow, H. and J. Tenenbaum, "Recovery of Intrinsic Scene Characteristics from Images", *Computer Vision Systems, A. Hanson and E. Riseman (ed.)*, Academic Press, New York, pp.3-26, 1978.

Baumgart, B.G., "Winged-edge Polyhedron Representation", *STAN-CS-320, AIM-179, Standford AI lab*, Oct 1972.

Beaudet, P.R., "Rotationally Invariant Image Operators", *4th International Joint Conference on Pattern Recognition, Tokyo, Japan*, pp.579-583, Nov 1978.

Besl, P.J. and R.C. Jain, "Segmentation through Symbolic Surface Descriptons", *Computer Vision and Pattern Recognition Conference*, pp.77-85, 1986.

Bickel P., and K. Doksum, "Mathematical Statistics: Basic Ideas and Selected Topics", *Holden-Day Series in Probability and Statistics*, 1977.

Brown, C., "Advances in Computer Vision, vol 1 and vol 2", *Lawrence Erlbaum Associates, Publishers*, 1988.

Brooks, M., "Rationalizing Edge Detectors", *Computer Graphics and Image Processing* **8**, pp.277-285, 1978.

Brooks, R.A., "Goal Directed Edge Linking and Ribbon Finding", *Proceedings ARPA Image Understading Workshop, Menlo Park, CA*, pp.72-76, 1979.

Brooks, R.A., "Symbolic Reasoning Among 3D Models and 2D Images", *Artificial Intelligence* **17**, pp.285-348, 1981.

Burns, J., A. Hanson, and E. Riseman, "Extracting Straight Lines", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8**, pp.425-455, July 1986.

Bury, K., "Statistical Models in Applied Science", *Wiley Series in Probability and Mathematical Statistics*, 1975.

Canny J., "A Computational Approach to Edge Detection", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8**, pp.679-698, Nov. 1986.

Cheng, J. and T. Huang, "Image Recognition by Matching Relational Structures", *IEEE Pattern Recognition and Image Processing Conference*, pp.542-547, 1981.

Cheng, J. and T. Huang, "Image Registration by Matching Relational Structures", *Pattern Recognition* **17**, pp.149-160, 1984.

Chu, Y.H., "Analyzing Perspective Views of a Wire Frame Object Model", *Ph.D. Dissertation, Virginia Polytechnic Institute and State University*, 1983.

Clark, C., W. Eckhardt, C. McNary, R. Nevatia, K. Olin, and E. VanOrden, "High Accuracy Model Matching for Scenes containing Man-made Structures", *in Digital Processing of Aerial Images, Proc. SPIE* **186** pp.54-62, 1979.

Dreschler, L. and H. Nagel, "Volumetric Model and 3D Trajectory of a Moving Car Derived from Monocular TV-frame Sequences of a Street Scene", *International Joint Conference on Artificial Intelligence*, pp.692-697, 1981.

Duda, R., and P. Hart, "Use of Hough Transformation to Detect Lines and Curves in pictures", *Commun. Ass. Computer Machinery* **15**, pp.11, 1972.

Duda, R., and P. Hart, "Pattern Classification and Scene Analysis", Wiley, New York, 1973.

Dudani, S. and A. Luk, "Locating Straight-Line Edge Segments on Outdoor Scenes", *IEEE Pattern Recognition and Image Processing Conference*, 1977.

Dyer, C.R. and A. Rosenfeld, "Thinning Algorithms for Gray Scale Pictures", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1**, pp.88-89, Jan 1979.

Ehrich, R.W. and J.P. Foith, "Topology and Semantics of Intensity Arrays", *Computer Vision Systems*, Academic Press, New York, pp.111-128, 1978.

Fischler, M. and H. Wolf, "A General Approach to Machine Perception of Linear Structure in Imaged Data", *SRI Artificial Intelligence Center, Technical Note 276*, 1983.

Fua, P. and A. Hanson, "Resegmentation Using Generic Shape: Locating General Cultural Objects", *Pattern Recognition Letters* **5**, pp.243-252, 1987.

Gibson, J.J., "The Perception of the Visual World ", Boston, Massachusetts, Houghton, 1950.

Gibson, J.J., "The Ecological Approach to Visual Perception ", Boston, Massachusetts, Houghton, 1979.

Graybill, F.A., "Theory and Application of the Linear Model ", Duxbury Press, 1976

Grimson, W. and E. Hildreth, "Comments on Digital Step Edges from Zero Crossings of Second Directional Derivatives", *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-7**, pp.121-126, 1985.

Haralick, R.M., "Author's Reply" *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-7**, pp.127-129, 1985.

Haralick, R.M., "Second Directional Derivative Zero-crossing Detector Using the Cubic Facet Model", *IEEE Computer Vision and Pattern Recognition Conference* , pp.672-677, 1985.

Haralick, R.M., "Edge and Region Analysis for Digital Image Data", *Computer Graphics and Image Processing* **12**, pp.60-73, 1980.

Haralick, R.M. and L.T. Watson, "A Facet Model for Image Data", *Computer Graphics and Image Processing* **15**, pp.113-129, 1981.

Haralick, R.M., "Zero-Crossing of Second Directional Derivative Edge Operator", *Society of Photogrammetric Instrumentation Engineering Symposium on Robot Vision, Washington, D.C.*, May 1982.

Haralick, R.M., L.T. Watson, and T.J. Laffey, "The Topographical Primal Sketch", *The International Journal of Robotics Research* **2**, pp.50-72, 1983.

Haralick, R.M., "Digital Step Edge from Second Directional Derivatives", *IEEE Trans. Pattern Anal. Mach. Intell.* , **PAMI-6** , pp.58-68, 1984.

Hashimoto, M. and J. Sklansky, "Multiple Order Derivatives for Detecting Local Image Characteristics", *Computer Vision, Graphics, and Image Processing* **39** , pp.28-55, 1987.

Herman, M. and T. Kanade, "The 3-D Mosaic Scene Understanding System", *From Pixels to Predicates, A. Pentland, Editor,* 1986.

Herskovits, A., "On Boundary Detection", *MIT Project MAC Memo 183, Cambridge, Massachussetts,* 1970.

Hildebrand F. B., "Introduction to Numerical Analysis", *McGraw Hill, New York* , 1956.

Hueckel, M., "An Operator which Locates Edges in Digital Pictures", *Journal Ass. Computer Machinery* **18,** pp.113-125, 1971.

Huertas, A., "Corner Detection for Finding Buildings in Aerial Images", *USCIPI Technical Report 1050, University of Southern California,* pp.61-68, 1981.

Huertas, A. and R. Nevatia, "Detection of Buildings in Aerial Images Using Shape and Shadows", *Proc. 8th Int. Conf. Artificial Intelligence,* pp.1099-1103, 1983.

Huertas, A. and R. Nevatia, "Detecting Buildings in Aerial Images", *Computer Vision, Graphics, and Image Processing* **41,** pp.131-152, 1988.

Hummel, R., "Feature Detection Using Basis Functions", *Computer Graphics and Image Processing* **9,** pp.40-55, 1979.

Iannino, A., and S. D. Shapiro, "An Iterative Generalization of the Sobel Edge Detection Operator", *IEEE Pattern Recognition and Image Processing Conference,* pp.130-137, 1979.

Jacobus, C. and R. Chien, "Variable Neighborhood Computations in Scene Analysis", *Univ. of Illinois, Urbana, CSL Rep. T-60,* 1978

Jacobus, C., R. Chien, and J. Selander, "Motion Detection and Analysis of Matching Graphs of Intermediate Level Primitives", *IEEE Trans. Pattern Anal. Mach. Intell.* **2,** pp.495-510, Nov. 1980.

Jacobus, C. and R. Chien, "Intermediate Level Vision - Building Vertex-String-Surface (V-S-S) Graphs", *Computer Graphics and Image Processing* **15,** pp.339-363, 1981.

Kestner, W., M. Bohner, R. Scharf, and M. Sties, "Object Guided Segmentation of Aerial Images", *5th International Conference on Pattern Recognition,* pp.529-531, 1980.

Kimme, C., D. Ballard, and Sklansky, "Finding Circles by an Array of Accumulators", *Commun. Ass. Computer Machinery* **18,** pp.120-122, 1975.

Kirsh, R. "Computer Determination of the Constituent Structure of Biological Images", *Computer Biomedical Research* **4,** pp.315-328, 1971.

Kitchen, L. and A. Rosenfeld, "Gray Level Corner Detection", *Technical Report 887, Computer Science Center, University of Maryland,* 1980.

Kitchen, L. and A. Rosenfeld, "Edge Evaluation Using Local Edge Coherence", *IEEE Trans. Systems, Man, and Cybernetics* **SMC-11** pp.597-605, Sept 1981.

Koehler, W., "Gestalt Psychology", New York, Liveright, 1929.

Koffka, K., "Principles of Gestalt Psycholgy", New York, Harcourt, Brace, 1935.

Lee, J., "Facet Model Optic Flow and Rigid Body Motion", *Ph.D. Dissertation, Virginia Polytechnic Institute and State University*, 1985.

Loew, M.H., L. Goldman, and R. Pickholtz, "Classification of Discrete Man-Made Structures in Cartographic Images", *6th International Conference on Pattern Recognition*, pp.1217, 1982.

Lowe, D., "Three-Dimensional Object Recognition from Single Two-Dimensional Images", *Artificial Intelligence* **31**, pp.355-395, 1987.

Lumia, R., L. Shapiro and O. Zuniga, "A New Connected Component Algorithm for Virtual Memory Computers", *Computer Vision, Graphics, and Image Processing* **2**, pp.287-300, 1983.

Lunscher W., and M. Beddoes, "Optimal Edge Detector Design I: Parameter Selection and Noise Effects", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8** pp.164-177, March 1986.

Marr, D., "Early Processing of Visual Information", *Phil. Trans. R. Soc. London* **B 275**, pp.483-524, 1976.

Marr, D., "Representing Visual Information", *Machine Vision 77, Advanced Papers for the Workshop on Computer Vision Systems, vol II*, 1977.

Marr, D. and E. Hildreth, "Theory of Edge Detection", *Proc. Roy. Soc. London* **B 204**, pp.187-217, 1980.

Martelli, A., "Edge Detection Using Heuristic Search Methods", *Computer Graphics and Image Processing*, **1**, pp.169-182, 1972.

McKee, J. and J. Aggarwal, "Finding the Edges of the Surfaces of Three-Dimensional Curved Objects by Computer", *Pattern Recognition* **7**, pp.25-52, 1975.

Medioni, G., "Obtaining 3D Information from Shadows in Aerial Images", *IEEE Computer Vision and Pattern Recognition Conference*, pp.73-76, 1983.

Medioni, G. and R. Nevatia, "Matching Images using Linear Features", *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, pp.675-685, 1984.

Mero, L., "An Optimal Line Following Algorithm", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3**, Sept 1981.

Milgram, D., "Region Extraction Using Convergence Evidence", *Computer Graphics and Image Processing* **11**, pp.1-12, 1979.

Montanari, U., "On the Optimal Detection of Curves in Noisy Pictures", *Commun. Ass. Computer Machinery* **14**, pp.335-345, 1971.

Moravec, H.P., "Towards Automatic Obstacle Avoidance", *5th International Joint Conference on Artificial Intelligence*, pp.584, 1977.

Morrison, D. "Multivariate Statistical Methods", *McGraw-Hill*, 1967.

Nagao, M., T. Matsuyama, and Y. Ikeda, "Region Extraction and Shape Analysis of Aerial Photographs", *Computer Graphics and Image Processing* pp.195-223, 1979.

Nagel, H. and W. Enkelmann, "Investigation of Second Order Grayvalue Variations to Estimate Corner Point Displacements", *IEEE Pattern Recognition and Image Processing Conference, Munich*, 1982.

Nalwa, V. and T. Binford, "On Detecting Edges", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8**, pp.699-714, Nov. 1986.

Nalwa, V. and E. Pauchon, "Edgel Aggregation and Edge Description", *Computer Vision, Graphics, and Image Processing* **40**, pp.79-94, Oct 1987.

Nevatia, R. and K. Babu, "Linear Feature Extraction and Description", *Computer Graphics and Image Processing* **13**, pp.257-269, 1980.

Nevatia, R. and K. Price, "Locating Structures in Aerial Images", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-4**, Sept 1982.

O'Gorman, F., and M. B. Clowes, "Finding Picture Edges through Collinearity of Feature Points", *IEEE Trans. on Computers* **25**, pp. 449-456, 1976.

O'Gorman, F., "Edge Detection Using Walsh Functions", *Artificial Intelligence* **10**, pp.215-223, 1978.

Ohlander, R.B., "Analysis of Natural Scenes", *Ph.D. dissertation, Carnegie-Mellon Univ., Pittsburgh, PA* 1976.

Perkins, W., "A Model Based Vision System for Industrial Parts", *IEEE Trans. Computers* **27**, 1978.

Perkins, W., "Area Segmentation of Images Using Edge Points", *IEEE Trans. Pattern Analysis and Mach. Intell.* **PAMI-2**, pp.8-15, Jan 1980.

Pong, T.C., L.G. Shapiro, L.T. Watson, and R.M. Haralick, "Experiments in Segmentation Using a Facet Model Region Grower", *Computer Vision, Graphics, and Image Processing* , pp.1-23, Jan 1984.

Pong, T.C., "Determining Intrinsic Scene Characteristics from Images", *Ph.D. Dissertation, Virginia Polytechnic Institute and State University*, 1984.

Prewitt, J.M.S., "Object Enhancement and extraction", *in Picture Processing and Psychopictorics (B. S. Lipkin, and A. Rosenfeld, Eds.)*, Academic Press, New York, pp.75-149, 1970.

Price, K. and R. Reddy, "Matching Segments of Images", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1,** pp.110-116, 1979.

Ramer, E., "Transformation of Photographic Images into Stroke Arrays", *IEEE Trans. Circuits and Systems* **CAS-22,** pp.363-373, 1975.

Reynolds, G., N. Irwin, A. Hanson, and E. Riseman, "Hierarchical Knowledge-Directed Object Extraction using a Combined Region and Line Representation" *Workshop on Computer Vision, Representation and Control,* pp.238-247, 1984.

Roberts, L.G., "Machine Perception of Three Dimensional Solids", in *Optical and Electro-optical Information Processing (J. T. Tippet, et al. Eds.), MIT Press, Cambridge, Mass.* pp.159-197, 1965.

Robinson, G., "Edge Detection by Compass Gradient Masks", *Computer Graphics and Image Processing* **6,** pp.492-501, 1977.

Rosenfeld, A., "A Characterization of Parallel Thinning Algorithms", *Information Control* **29,** pp.286-291, 1975.

Rosenfeld, A. and L.S. Davis, "A Note on Thinning", *IEEE Trans. Systems, Man, and Cybernetics* **6,** pp.226-228, 1976.

Rosenfeld, A. and A. Kak, "Digital Picture Processing", *Academic Press,* 1982.

Rutkowski, W.S. and A. Rosenfeld, "A Comparison of Corner Detection Techniques for Chain-coded Curves", *Technical Report 623, Computer Science Center, University of Maryland,* Jan 1978.

Scher, A., M. Schneier, and A. Rosenfeld, "A Method for Finding Pairs of Antiparallel Straight Lines", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-4,** pp.41-50, 1982a.

Scher, A., M. Schneier, and A. Rosenfeld, "Clustering of Collinear Line Segments", *Pattern Recognition* **15,** pp.85-91, 1982b.

Selfridge, P. and K. Sloan, "Reasoning about Images: Using Meta-Knowledge in Aerial Image Understanding", *International Joint Conference on Artificial Intelligence,* pp.755-757, 1982.

Shanmugan, K., F. Dickey, and J. Green, "An Optimal Frequency Domain Filter for Edge Detection in Digital Pictures", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1** pp.37-49, Jan 1979.

Shirai, Y., "Recognition of Real World Objects Using Edge Cue", *Machine Vision 77, Advanced Papers for the Workshop on Computer Vision Systems, vol I,* 1977.

Shirai, Y., "Three-Dimensional Computer Vision", *Springer-Verlag Series, Symbolic Computation,* 1987

Stefanelli, R. and A. Rosenfeld., "Some Parallel Thinning Algorithms for Digital Pictures", *J. Ass. Comput. Mach.* **18,** pp.255-264, 1971.

Stockman, G., S. Kopstein, and S. Benett, "Matching Images to Models for Registration and Object Detection Via Clustering", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-4**, pp.229-241, 1982.

Szego G., "Orthogonal Polynomials", *American Mathematical Society, Providence, Rhode Island*, 1939

Tavakoli, M. and A. Rosenfeld, "Building and Road Extraction from Aerial Photographs", *Trans. Systems, Man, and Cybernetics* **12**, pp.84-91, 1982.

Tavakoli, M. and A. Rosenfeld "Edge Segment Linking Based on Gray Level and Geometric Compatibilities", *Pattern Recognition* **15**, pp.369-377, 1982.

Torre, V. and T. Poggio, "On Edge Detection", *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8**, pp.147-163, 1986.

Watson L.T., T.J. Laffey, and R.M. Haralick, "Topographic Classification of Digital Image Intensity Surfaces using Generalized Splines and the Discrete Cosine Transformation", *Computer Vision, Graphics, and Image Processing*, **29**, pp.143-167, 1985.

Weiss, R. and M. Boldt, "Geometric Grouping Applied to Straight Lines", *IEEE Computer Vision and Pattern Recognition Conference*, pp.489-495, 1986.

Wertheimer, M., "Experimentelle Studien ueber das Sehen von Bewegunger", *Z. Psychol.* **61**, pp.161-265, 1912.

Weszka, J.S. "A Survey of threshold selection techniques", *Computer Graphics and Image Processing* , **7** , pp.259-265, 1978.

Young, R.A., "The Gaussian derivative model for machine and biological image processing", *GMR Technical Report, GMR-5128, Warren, MI*, 1985

Zuniga, O.A. and R.M. Haralick, "Corner Detection using the Facet Model", *IEEE Computer Vision and Pattern Recognition Conference* , 30-37 (1983).

Zuniga, O.A., and R.M. Haralick, "Integrated Derivative Gradient Operator", *IEEE Trans. Systems, Man, and Cybernetics* **SMC-17** pp.508-516, May/June 1987.

Appendix A

The Cubic Facet

The cubic facet is described by the two-dimensional polynomial surface function:

$$\begin{aligned} f(r,c) = & K_1 \\ & + K_2 r + K_3 c \\ & + K_4 r^2 + K_5 rc + K_6 c^2 \\ & + K_7 r^3 + K_8 r^2 c + K_9 r c^2 + K_{10} c^3 \end{aligned}$$

This polynomial is fit to the graytones f(r,c) inside a given image neighborhood. We will assume a rectangular shaped neighborhood whose row index set is R and whose column index set is C, and whose center is taken to be at (0,0). Notice that for an even sized neighborhood the center falls at the point where the four center pixels meet. The squared fitting error over this neighborhood is given by

$$e^2 = \sum_{r \in R} \sum_{c \in C} (K_1 + K_2 r + K_3 c + K_4 r^2 + K_5 rc + K_6 c^2 + K_7 r^3 + K_8 r^2 c + K_9 rc^2 + K_{10} c^3 - f(r,c))^2$$

Taking the partial derivatives of e^2 with respect to the parameters $K_1, \ldots, K_{10}$ results in:

$$\begin{bmatrix} \partial e^2/\partial K_1 \\ \partial e^2/\partial K_2 \\ \partial e^2/\partial K_3 \\ \partial e^2/\partial K_4 \\ \partial e^2/\partial K_5 \\ \partial e^2/\partial K_6 \\ \partial e^2/\partial K_7 \\ \partial e^2/\partial K_8 \\ \partial e^2/\partial K_9 \\ \partial e^2/\partial K_{10} \end{bmatrix} = \sum_{r \in R} \sum_{c \in C} 2(K_1 + K_2 r + K_3 c + K_4 r^2 + K_5 rc + K_6 c^2 + K_7 r^3 + K_8 r^2 c + K_9 rc^2 + K_{10} c^3 - f(r,c)) \begin{bmatrix} 1 \\ r \\ c \\ r^2 \\ rc \\ c^2 \\ r^3 \\ r^2 c \\ rc^2 \\ c^3 \end{bmatrix}$$

Because the sum is between symmetric limits there is a considerable amount of cancellation when carrying out the summation.

$$\begin{bmatrix} \partial e^2/\partial K_1 \\ \partial e^2/\partial K_2 \\ \partial e^2/\partial K_3 \\ \partial e^2/\partial K_4 \\ \partial e^2/\partial K_5 \\ \partial e^2/\partial K_6 \\ \partial e^2/\partial K_7 \\ \partial e^2/\partial K_8 \\ \partial e^2/\partial K_9 \\ \partial e^2/\partial K_{10} \end{bmatrix} = 2 \sum_{r \in R} \sum_{c \in C} \begin{bmatrix} K_4 r^2 + K_6 c^2 + K_1 - f(r,c) \\ K_7 r^4 + K_9 r^2 c^2 + K_2 r^2 - r f(r,c) \\ K_8 r^2 c^2 + K_{10} c^4 + K_3 c^2 - c f(r,c) \\ K_4 r^4 + K_6 r^2 c^2 + K_1 r^2 - r^2 f(r,c) \\ K_5 r^2 c^2 - rc f(r,c) \\ K_4 r^2 c^2 + K_6 c^4 + K_1 c^2 - c^2 f(r,c) \\ K_7 r^6 + K_9 r^4 c^2 + K_2 r^4 - r^3 f(r,c) \\ K_8 r^4 c^2 + K_{10} r^2 c^4 + K_3 r^2 c^2 - r^2 c f(r,c) \\ K_7 r^4 c^2 + K_9 r^2 c^4 + K_2 r^2 c^2 - r c^2 f(r,c) \\ K_8 r^2 c^4 + K_{10} c^6 + K_3 c^4 - c^3 f(r,c) \end{bmatrix}$$

Setting the partials to zero and solving we immediately obtain the least squares estimate K'_5 of K_5.

$$K'_5 = \frac{\sum_{r \in R} \sum_{c \in C} rc f(r,c)}{\sum_{r \in R} \sum_{c \in C} r^2 c^2}$$

The least squares estimates for the remaining parameters are obtained by solving the following three systems of equations:

$$\begin{bmatrix} \sum_r \sum_c 1 & \sum_r \sum_c r^2 & \sum_r \sum_c c^2 \\ \sum_r \sum_c r^2 & \sum_r \sum_c r^4 & \sum_r \sum_c r^2c^2 \\ \sum_r \sum_c c^2 & \sum_r \sum_c r^2c^2 & \sum_r \sum_c c^4 \end{bmatrix} \begin{bmatrix} K_1 \\ K_4 \\ K_6 \end{bmatrix} = \begin{bmatrix} \sum_r \sum_c f(r,c) \\ \sum_r \sum_c r^2 f(r,c) \\ \sum_r \sum_c c^2 f(r,c) \end{bmatrix}$$

$$\begin{bmatrix} \sum_r \sum_c r^2 & \sum_r \sum_c r^4 & \sum_r \sum_c r^2c^2 \\ \sum_r \sum_c r^4 & \sum_r \sum_c r^6 & \sum_r \sum_c r^4c^2 \\ \sum_r \sum_c r^2c^2 & \sum_r \sum_c r^4c^2 & \sum_r \sum_c r^2c^4 \end{bmatrix} \begin{bmatrix} K_2 \\ K_7 \\ K_9 \end{bmatrix} = \begin{bmatrix} \sum_r \sum_c rf(r,c) \\ \sum_r \sum_c r^3 f(r,c) \\ \sum_r \sum_c rc^2 f(r,c) \end{bmatrix}$$

$$\begin{bmatrix} \sum_r \sum_c c^2 & \sum_r \sum_c r^2c^2 & \sum_r \sum_c c^4 \\ \sum_r \sum_c r^2c^2 & \sum_r \sum_c r^4c^2 & \sum_r \sum_c r^2c^4 \\ \sum_r \sum_c c^4 & \sum_r \sum_c r^2c^4 & \sum_r \sum_c c^6 \end{bmatrix} \begin{bmatrix} K_3 \\ K_8 \\ K_{10} \end{bmatrix} = \begin{bmatrix} \sum_r \sum_c cf(r,c) \\ \sum_r \sum_c r^2 cf(r,c) \\ \sum_r \sum_c c^3 f(r,c) \end{bmatrix}$$

For n = 0, 1, 2 or 3 let R_n and C_n be defined as:

$$R_n = \sum_{r \in R} r^{2n} \quad , C_n = \sum_{c \in C} c^{2n}$$

Furthermore let

$$G = R_0 R_2 C_0 C_2 - R_1^2 C_1^2$$
$$A = R_1 R_3 C_0 C_2 - R_2^2 C_1^2$$
$$B = R_0 R_2 C_1 C_3 - R_1^2 C_2^2$$
$$Q = C_0 (R_0 R_2 - R_1^2)$$
$$T = R_0 (C_0 C_2 - C_1^2)$$
$$U = C_0 (R_1 R_3 - R_2^2)$$
$$V = C_1 (R_0 R_2 - R_1^2)$$
$$W = R_1 (C_0 C_2 - C_1^2)$$
$$Z = R_0 (C_1 C_3 - C_2^2)$$

The solution of the above systems of equation is then given by:

$$K'_1 = \frac{1}{Q\,T} \sum_r \sum_c (G - T\,R_1 r^2 - Q\,C_1 c^2) f(r,c)$$

$$K'_2 = \frac{1}{U\,W} \sum_r \sum_c (A - W\,R_2 r^2 - U\,C_1 c^2)\, r f(r,c)$$

$$K'_3 = \frac{1}{V\,Z} \sum_r \sum_c (B - Z\,R_1 r^2 - V\,C_2 c^2)\, c f(r,c)$$

$$K'_4 = \frac{1}{Q} \sum_r \sum_c (R_0 r^2 - R_1) f(r,c)$$

$$K'_6 = \frac{1}{T} \sum_r \sum_c (C_0 c^2 - C_1) f(r,c)$$

$$K'_7 = \frac{1}{U} \sum_r \sum_c (R_1 r^2 - R_2)\, r f(r,c)$$

$$K'_8 = \frac{1}{V} \sum_r \sum_c (R_0 r^2 - R_1)\, c f(r,c)$$

$$K'_9 = \frac{1}{W} \sum_r \sum_c (C_0 c^2 - C_1)\, r f(r,c)$$

$$K'_{10} = \frac{1}{Z} \sum_r \sum_c (C_1 c^2 - C_2)\, c f(r,c)$$

Appendix B

Statistical Distributions of the norm of facet parameters, partial derivatives, and facet errors

In this appendix we derive the distributions of the norm of any subset of facet parameters or partial derivatives, and the distribution of the total facet residual error. We assume the noise to be Gaussian with zero mean and known covariance matrix Σ. In order to proceed we will make use of three theorems frequently used in multivariate statistics. Proofs of these theorems can be found in Graybill (1976).

Theorem 1

Let x be a K X 1 random vector with a $N(\mu, \Sigma)$ distribution, that is the elements of x have a multivariate normal (Gaussian) distribution with mean vector μ and covariance matrix Σ. Then $(x-\mu)' \Sigma^{-1}(x-\mu)$ is a chi-squared variate $\chi^2(K)$ with K degrees of freedom.

Theorem 2

Let x be a K X 1 random vector with distribution $N(\mu, \Sigma)$. Consider the m linear functions on the elements of x defined by y = Bx where y is an m X 1 vector, $m \leq K$ and B is an m X K real matrix of rank m. Then y has a distribution $N(B\mu, B\Sigma B')$.

Theorem 3

Let $Q = Q_1 + \cdots + Q_{L-1} + Q_L$ where $Q, Q_1, \ldots, Q_L$ are L+1 random variables that are quadratic forms in any multivariate normal variables. Let Q be a $\chi^2(r)$ variate, let Q_i be $\chi^2(r_i), i = 1, \ldots, L-1$ variates, and let Q_L be non-negative. Then the random variables $Q_1, \ldots, Q_L$ are mutually independent and hence Q_L is a $\chi^2(r_L = r - r_1 - \cdots - r_{L-1})$ variate.

The model

The image data model is described by

$$F\alpha + \eta = x \tag{1}$$

where F is a K X N basis matrix, α is an N X 1 parameter vector, x is a K X 1 observation vector, and η is a K X 1 noise vector. We assume the noise to be Gaussian with zero mean and known covariance matrix Σ, that is η has a multivariate $N(0,\Sigma)$ distribution.

The minimum variance unbiased estimate $\hat{\alpha}$ of the true parameter vector α (an estimate that minimizes $(x - \hat{x})' \Sigma^{-1}(x - \hat{x})$) is known to be (Graybill (1976)):

$$\hat{\alpha} = Px \tag{2}$$

where

$$P = (F' \Sigma^{-1} F)^{-1} F' \Sigma^{-1} \tag{3}$$

Distribution of the parameter vector

Replacing equation (1) in equation (2) we obtain

$$\begin{aligned}\hat{\alpha} &= P(F\alpha + \eta) \\ &= PF\alpha + P\eta \\ &= \alpha + P\eta\end{aligned}$$

therefore

$$\hat{\alpha} - \alpha = P\eta \tag{4}$$

Since η is N(0,Σ), applying Theorem 2 we conclude that $(\hat{\alpha} - \alpha)$ is $N(0, P\Sigma P')$.

Using (3) $P\Sigma P'$ reduces to

$$P\Sigma P' = (F'\Sigma^{-1}F)^{-1}$$

therefore

$$(\hat{\alpha} - \alpha) \quad is \quad N(0,(F'\Sigma^{-1}F)^{-1}) \tag{5}$$

that is the parameter vector $\hat{\alpha}$ has a multivariate normal distribution with mean α and covariance matrix $(F'\Sigma^{-1}F)^{-1}$.

Distribution of the norm of the parameter vector

Using (5) and applying theorem 1 it follows that

$$(\hat{\alpha} - \alpha)'(F'\Sigma^{-1}F)(\hat{\alpha} - \alpha) \quad is \quad \chi^2(N) \tag{6}$$

a chi-squared variate with N degrees of freedom.

This result also applies to the norm of any subset of elements of $(\hat{\alpha} - \alpha)$ as follows. Let α_m be an m X 1 vector obtained by selecting m elements of α, $1 \leq m \leq N$. Moreover let F_m be a K X m matrix containing the m basis vectors of the basis matrix F which correspond to the m elements selected from α. It then follows that

$$(\hat{\alpha}_m - \alpha_m)(F'_m \Sigma^{-1} F_m)(\hat{\alpha}_m - \alpha_m) \quad is \quad \chi^2(m) \tag{7}$$

a chi-squared variate with m degrees of freedom.

Distribution of the total residual error

The residual error vector e is given by

$$\begin{aligned} e &= x - \hat{x} \\ &= x - F\hat{\alpha} \\ &= F\alpha + \eta - F\hat{\alpha} \\ &= \eta - F(\hat{\alpha} - \alpha) \end{aligned}$$

The total residual error is therefore

$$e' \Sigma^{-1} e = (\eta - F(\hat{\alpha} - \alpha))' \Sigma^{-1} (\eta - F(\hat{\alpha} - \alpha))$$

$$e' \Sigma^{-1} e = \eta' \Sigma^{-1} \eta - 2(\hat{\alpha} - \alpha)' F' \Sigma^{-1} \eta + (\hat{\alpha} - \alpha)' F' \Sigma^{-1} F(\hat{\alpha} - \alpha) \tag{8}$$

From (3) and (4) we obtain

$$F' \Sigma^{-1} \eta = (F' \Sigma^{-1} F)(\hat{\alpha} - \alpha)$$

Substituting this expression in (8) yields

$$e' \Sigma^{-1} e = \eta' \Sigma^{-1} \eta - (\hat{\alpha} - \alpha)'(F' \Sigma^{-1} F)(\hat{\alpha} - \alpha) \tag{9}$$

Since by assumption η is N(0,Σ), applying theorem 1 it follows that

$$\eta' \Sigma^{-1} \eta \quad is \quad \chi^2(K)$$

a chi-squared variate with K degrees of freedom.

Finally using (6) and applying theorem 3 we obtain the result

$$e' \Sigma^{-1} e \quad is \quad \chi^2(K-N) \tag{10}$$

that is the total residual error is a chi-squared variate with K-N degrees of freedom.

Distribution of the partial derivatives

We assume that F is a polynomial basis matrix. It then follows that each partial derivative at (0,0) in the row and column directions is given as some linear combination of the elements of the parameter vector.

Let μ be an m X 1 vector containing any m partial derivatives, $1 \leq m \leq N-1$ Let B be the m X N linear combination matrix. Then

$$\mu = B\alpha \quad \wedge \quad \hat{\mu} = B\hat{\alpha}$$

Also

$$\hat{\mu} - \mu = B(\hat{\alpha} - \alpha) \tag{11}$$

Using (11), (5), and applying theorem 2 we find that

$$(\hat{\mu} - \mu) \quad is \quad N(0, B(F' \Sigma^{-1} F)^{-1} B') \tag{12}$$

Some special cases of interest

Independent, equally distributed noise

In this case η is $N(0, \sigma^2 I)$ and the minimum variance estimate of α becomes a least square estimate.

$$\hat{\alpha} = (F'F)^{-1}F'x$$

Our previous results (5), (7), (10), and (12) reduce to

$$(\hat{\alpha} - \alpha) \quad is \quad N(0, \sigma^2(F'F)^{-1})$$

$$(\hat{\alpha}_m - \alpha_m)'(F'_m F_m)(\hat{\alpha}_m - \alpha_m)/\sigma^2 \quad is \quad \chi^2(m)$$

$$e'e/\sigma^2 \quad is \quad \chi^2(K-N)$$

$$(\hat{\mu} - \mu) \quad is \quad N(0, \sigma^2 B(F'F)^{-1}B')$$

Independent noise, orthonormal basis

Our previous results further simplify to

$$\hat{\alpha} = F'x$$

$$(\hat{\alpha} - \alpha) \quad is \quad N(0, \sigma^2 I)$$

$$(\hat{\alpha}_m - \alpha_m)'(\hat{\alpha}_m - \alpha_m)/\sigma^2 \quad is \quad \chi^2(m)$$

$$e'e/\sigma^2 \quad is \quad \chi^2(K-N)$$

$$(\hat{\mu} - \mu) \quad is \quad N(0, \sigma^2 BB')$$

Appendix C

Orthogonal Polynomials

C.1 Legendre polynomials

The first four Legendre polynomials $P_k(x/N)$, $k = 1,...,4$ are:

$$1$$

$$\frac{1}{N}x$$

$$\frac{3}{2N^2}(x^2 - \frac{1}{3}N^2)$$

$$\frac{5}{2N^3}(x^3 - \frac{3}{5}N^2x)$$

C.2 Hermite polynomials

The first four Hermite polynomials $H_k(x/\sqrt{2}\,\sigma)$, $k = 1,...,4$ are:

$$1$$

$$\frac{\sqrt{2}}{\sigma}x$$

$$\frac{2}{\sigma^2}(x^2 - \sigma^2)$$

$$\frac{2\sqrt{2}}{\sigma^3}(x^3 - 3\sigma^2 x)$$

C.3 Gram polynomials

The first four Gram polynomials $p_m(x)$, $m = 1,...4$ are:

$$1$$

$$\frac{1}{N}x$$

$$\frac{3}{N(2N-1)}(x^2 - \frac{1}{3}N(N+1))$$

$$\frac{5}{N(N-1)(2N-1)}(x^3 - \frac{1}{5}(3N^2 + 3N - 1)x)$$

C.4 *Krawtchouk polynomials*

The first four Krawtchouk polynomials k_m, $m = 1,...,4$ are:

$$1$$

$$x$$

$$\frac{1}{2}(x^2 - \frac{1}{2}N)$$

$$\frac{1}{6}(x^3 - \frac{1}{2}(3N - 1)x)$$

Appendix D

Distribution of the ratio of two normal variables

Let X and Y be independent normal random variables such that $X \sim N(0, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$. We will derive in this appendix the probability density function of the random variable $T = X/Y$.

Using the change of variables technique

$$T = X/Y \quad U = Y$$

Therefore

$$X = UT \quad Y = U$$

Then, the Jacobian of the transformation is

$$J = \begin{vmatrix} \partial x/\partial t & \partial x/\partial u \\ \partial y/\partial t & \partial y/\partial u \end{vmatrix} = \begin{vmatrix} u & t \\ 0 & 1 \end{vmatrix} = u$$

Then,

$$f_T(t) = \int_{-\infty}^{\infty} |J| f_{X,Y}(ut, u) du$$

and since X and Y are independent

$$f_T(t) = \int_{-\infty}^{\infty} |u| f_X(ut) f_Y(u) du$$

where

$$f_X(ut) = \frac{1}{\sqrt{2\pi}\,\sigma_X} e^{-t^2u^2/2\sigma_X^2}$$

$$f_Y(u) = \frac{1}{\sqrt{2\pi}\,\sigma_Y} e^{-(u-\mu_Y)^2/2\sigma_Y^2}$$

Therefore

$$f_T(t) = \frac{1}{2\pi\sigma_X\sigma_Y} \int_{-\infty}^{\infty} |u|\, e^{-\frac{1}{2}\left(\frac{u^2t^2}{\sigma_X^2} + \frac{(u-\mu_Y)^2}{\sigma_Y^2}\right)} du$$

which can be put in the following form

$$f_T(t) = \frac{1}{2\pi\sigma_X\sigma_Y} \int_{-\infty}^{\infty} |u|\, e^{-\frac{1}{2}\left[\left(\frac{u-b}{a}\right)^2 + c\right]} du$$

where

$$a = \frac{\sigma_X}{\sqrt{(\sigma_X/\sigma_Y)^2 + t^2}}$$

$$b = \frac{\mu_Y(\sigma_X/\sigma_Y)^2}{(\sigma_X/\sigma_Y)^2 + t^2}$$

$$c = \frac{(\mu_Y/\sigma_Y)^2 t^2}{(\sigma_X/\sigma_Y)^2 + t^2}$$

Therefore

$$f_T(t) = \frac{e^{-c/2}}{2\pi\sigma_X\sigma_Y} \int_{-\infty}^{\infty} |u|\ e^{-(u-b)^2/2a^2} du$$

Let $z = u - b$. Then

$$f_T(t) = \frac{e^{-c/2}}{2\pi\sigma_X\sigma_Y} \int_{-\infty}^{\infty} |z + b|\ e^{-z^2/2a^2} dz$$

$$f_T(t) = \frac{e^{-c/2}}{2\pi\sigma_X\sigma_Y} \left[- \int_{-\infty}^{-b} (z + b) e^{-z^2/2a^2} dz + \int_{-b}^{\infty} (z + b) e^{-z^2/2a^2} dz \right]$$

which is equivalent to

$$f_T(t) = \frac{e^{-c/2}}{2\pi\sigma_X\sigma_Y}\left[2\int_b^{\infty} z e^{-z^2/2a^2}dz + \int_{-b}^{b} b\, e^{-z^2/2a^2}dz\right]$$

$$= \frac{e^{-c/2}}{2\pi\sigma_X\sigma_Y}\left[2a^2 e^{-b^2/2a^2} + \sqrt{2\pi}\, ab \frac{1}{\sqrt{2\pi}}\int_{-b/a}^{b/a} e^{-x^2/2}dx\right]$$

$$= \frac{e^{-c/2}}{2\pi\sigma_X\sigma_Y}\left[2a^2 e^{-b^2/2a^2} + \sqrt{2\pi}\, ab\ erf(b/\sqrt{2}\, a)\right]$$

and replacing the values of a, b, and c

$$f_T(t) = \frac{q}{\pi(q^2+t^2)} e^{-p^2/2} + \frac{pq^2}{\sqrt{2\pi}\,(q^2+t^2)^{3/2}} e^{-p^2t^2/2(q^2+t^2)} erf\left(\frac{pq}{\sqrt{2}\sqrt{q^2+t^2}}\right)$$

where

$$p = \frac{\mu_Y}{\sigma_Y}$$

$$q = \frac{\sigma_X}{\sigma_Y}$$

Appendix E

Fitting a straight line to a set of oriented primitives

Given a set of primitives described by their position (r_i, c_i) and orientation θ_i, we are interested in fitting a straight line through them. The ideal line is described by the equation $r \sin\theta + c \cos\theta = \rho$ (see Figure 82 on page 202).

We assume the following statistical model for these primitives

$$p_{\theta,\rho}(r_i, c_i, \theta_i) = \frac{1}{2\pi\sigma_{\rho i}\sigma_{\theta i}} e^{-\frac{(r_i \sin\theta + c_i \cos\theta - \rho)^2}{2\sigma_{\rho i}^2}} e^{-\frac{\sin^2(\theta_i - \theta)}{2\sigma_{\theta i}^2}}$$

That is, the perpendicular positional error of each primitive with respect to the ideal line, and the angular error are assumed to be normally distributed. Furthermore, the positional and angular errors are assumed to be independent. Errors of different primitives are independent as well, although they do not necessarily have the same variance. The positional error variance for each primitive is $\sigma_{\rho i}^2$ and the angular error variance is $\sigma_{\theta i}^2$.

Least-squares estimates for the line parameters ρ and θ can be obtained by minimizing the weighted squared error

$$\varepsilon^2 = \varepsilon_\rho^2 + \varepsilon_\theta^2$$

where

$$\varepsilon_\rho^2 = \sum_i \frac{(r_i \sin\theta + c_i \cos\theta - \rho)^2}{\sigma_{\rho i}^2}$$

and

$$\varepsilon_\theta^2 = \sum_i \frac{\sin^2(\theta_i - \theta)}{\sigma_{\theta i}^2}$$

The values of ρ and θ that minimize ε^2 are the solutions of the system of equations

$$\frac{\partial \varepsilon^2}{\partial \rho^2} = 0$$

$$\frac{\partial \varepsilon^2}{\partial \theta^2} = 0$$

Differentiation with respect to ρ yields

$$\begin{aligned}\frac{\partial \varepsilon^2}{\partial \rho^2} &= \frac{\partial \varepsilon_\rho^2}{\partial \rho^2} \\ &= -2 \sum_i (r_i \sin\theta + c_i \cos\theta - \rho)/\sigma_{\rho i}^2 \\ &= -2\left(\sum_i r_i/\sigma_{\rho i}^2\right) \sin\theta - 2\left(\sum_i c_i/\sigma_{\rho i}^2\right) \cos\theta + 2\rho \sum_i 1/\sigma_{\rho i}^2\end{aligned}$$

Equating to zero results in

$$\rho = \bar{r} \sin\theta + \bar{c} \cos\theta$$

where

$$\bar{r} = \frac{\sum_l r_l/\sigma_{\rho l}^2}{\sum_l 1/\sigma_{\rho l}^2}$$

$$\bar{c} = \frac{\sum_l c_l/\sigma_{\rho l}^2}{\sum_l 1/\sigma_{\rho l}^2}$$

Differentiation with respect to θ yields

$$\frac{\partial \varepsilon^2}{\partial \theta^2} = \frac{\partial \varepsilon_\rho^2}{\partial \theta^2} = \frac{\partial \varepsilon_\theta^2}{\partial \theta^2}$$

and

$$\frac{\partial \varepsilon_\rho^2}{\partial \theta^2} = 2 \sum_l (r_l \sin\theta + c_l \cos\theta - \rho)(r_l \cos\theta - c_l \sin\theta)/\sigma_{\rho l}^2$$

After some algebraic manipulation, terms in $\sin 2\theta$ and $\cos 2\theta$ can be collected

$$\frac{\partial \varepsilon_\rho^2}{\partial \theta^2} = w_\rho (S_r^2 - S_c^2) \sin 2\theta + 2 w_\rho S_{rc} \cos 2\theta$$

where

$$w_\rho = \sum_l 1/\sigma_{\rho l}^2$$

$$S_r^2 = \frac{\sum_l (r_l - \bar{r})^2 / \sigma_{\rho l}^2}{\sum_l 1/\sigma_{\rho l}^2}$$

$$S_c^2 = \frac{\sum_l (c_l - \bar{c})^2 / \sigma_{\rho l}^2}{\sum_l 1/\sigma_{\rho l}^2}$$

$$S_{rc} = \frac{\sum_l (r_l - \bar{r})(c_l - \bar{c}) / \sigma_{\rho l}^2}{\sum_l 1/\sigma_{\rho l}^2}$$

Now,

$$\frac{\partial \varepsilon_\theta^2}{\partial \theta^2} = -2 \sum_i \sin(\theta_i - \theta) \cos(\theta_i - \theta)/\sigma_{\theta i}^2$$

$$= -\sum_i \sin 2(\theta_i - \theta)/\sigma_{\theta i}^2$$

$$= \sin 2\theta \sum_i \cos \theta_i/\sigma_{\theta i}^2 - \cos 2\theta \sum_i \sin 2\theta_i/\sigma_{\theta i}^2$$

$$= w_\theta(\overline{\cos 2\theta} \sin 2\theta - \overline{\sin 2\theta} \cos 2\theta)$$

where

$$w_\theta = \sum_i 1/\sigma_{\theta i}^2$$

$$\overline{\sin 2\theta} = \frac{\sum_i \sin 2\theta_i/\sigma_{\theta i}^2}{\sum_i 1/\sigma_{\theta i}^2}$$

$$\overline{\cos 2\theta} = \frac{\sum_i \cos 2\theta_i/\sigma_{\theta i}^2}{\sum_i 1/\sigma_{\theta i}^2}$$

Therefore

$$\frac{\partial \varepsilon^2}{\partial \theta^2} = (w_\rho(S_r^2 - S_c^2) + w_\theta \overline{\cos 2\theta}) \sin 2\theta + (2w_\rho S_{rc} - w_\theta \overline{\sin 2\theta}) \cos 2\theta$$

Equating to zero results in

$$\tan 2\theta = \frac{w_\theta \overline{\sin 2\theta} - 2w_\rho S_{rc}}{w_\theta \overline{\cos 2\theta} + w_\rho(S_r^2 - S_c^2)}$$

Therefore

$$\theta = \frac{1}{2} \tan^{-1} \frac{w_\theta \overline{\sin 2\theta} - 2w_\rho S_{rc}}{w_\theta \overline{\cos 2\theta} + w_\rho(S_r^2 - S_c^2)}$$

Appendix F

Computer Vision using GIPSY

GIPSY is an acronym for "General Image Processing System", a software system for the purpose of image processing and image understanding. GIPSY provides the user with a variety of commands that manipulate digital images or data derived from images.

This appendix illustrates the sequence of GIPSY commands used to perform the low-level and intermediate-level computer vision tasks developed in this dissertation. Starting with the next page, two sample GIPSY sessions are shown. These sessions illustrate the analysis of two aerial images of suburban scenes from the Sunnyvale area in California.

```
Welcome to the Ephemeral World of Gipsy (Version 4.22 05/14/87)

G:  RUN MASTER1.RUN
G:  !
G:  ! THIS RUN FILE PERFORMS LOW-LEVEL AND MID-LEVEL COMPUTER VISION
G:  ! TASKS NECESSARY FOR CULTURAL OBJECT EXTRACTION.
G:  ! THE ORIGINAL IMAGE ( RADC3.SBG ) IS A SUNNYVALE, CALIFORNIA,
G:  ! SUBURBAN SCENE.
G:  !
G:  ! GENERATE 11 BANDS OF THE CENTRAL NEIGHBORHOOD LOCAL CUBIC FACET
G:  !
G:  CNCFCT RADC3.CB5 < RADC3.SBG (E)
Odd number of rows for window ( D = 5, 1 to 21 ) --  5
Odd number of columns for window ( D = 5, 1 to 21 ) --  5
Number of bits ( D = 14, 10 to 18 ) --  18
G:  !
G:  ! RUN HARALICK'S ZERO-CROSSING OF SECOND DIRECTIONAL DERIVATIVE
G:  ! EDGE OPERATOR. USE THE INTEGRATED DIRECTIONAL DERIVATIVE
G:  ! GRADIENT DEFINITION AND COMPUTE GRADIENT THRESHOLD AUTOMATICALLY
G:  !
G:  EDGE RADC3.EI5 < RADC3.CB5 (AOTI)
Radius for 2nd derivative ( D = 0.70, 0.00 to 2.00 ) --  0.5
Confidence limit for edge direction ( D = 0.95, 0.00 to 1.00 ) -- 0.95
G:  !
G:  ! THIN EDGES BY SUPPRESSING NON-MAXIMUM EDGE MAGNITUDES IN THE
G:  ! GRADIENT DIRECTION
G:  !
G:  THNEDG RADC3.TN5 < RADC3.EI5
Enter background symbol used in input ( D = 1, 0 to 99 ) --  1
Enter band number for edge strength ( D = 1, 1 to 4 ) --  1
Enter band number for edge direction ( D = 2, 1 to 4 ) --  2
Enter band number for symbolic edge map ( D = 4, 1 to 4 ) --  4
G:  !
G:  ! COMPLEMENT BINARY EDGE IMAGE
G:  !
G:  BOOL RADC3.LN5 < RADC3.TN5, RADC3.TN5
(1) A & B      (2) A & ^B    (3) ^A & B
(4) ^A & ^B    (5) A ! B     (6)  A ! ^B
(7) ^A ! B     (8) ^A ! ^B   (9)  A XOR B
(10) ^(A XOR B)     (11) Minimum ( A, B )
(12) Maximum ( A, B )
Enter function number ( D = 1, 1 to 12 ) --  8
G:  !
G:  ! CREATE PROPERTY FILE AND CHAIN CODE FILE OF LINKED EDGES
G:  !
G:  ARCSEG RADC3.ARC, RADC3.CC < RADC3.LN5 (S)
NUMBER OF ARC LABELS TO PROCESS ( D = 1, >= 1 ) --  1
ENTER ARC LABEL ( D = 1, >= 1 ) --  1
# PROP RECS IN FRAC OF IMAGE POINTS ( D = 0.10, 0.010 to 1.00 ) -- 0.1
```

```
G:  !
G:  ! CREATE IMAGE OF LINKED EDGES
G:  !
G:  CHNSIF RADC3ARC.ID < RADC3.ARC, RADC3.CC (AX)
G:  !
G:  ! PRUNE EDGE SEGMENTS BY LENGTH
G:  !
G:  PRLINK RADC3.AP1 < RADC3.ARC, RADC3.CC
Max size which can be deleted: ( D = 10, >= 1 ) --  7
Min size which can be retained: ( D = 1, 1 to 7 ) --  1
No of pixels for computing angle: ( D = 3, 1 to 7 ) --  3
Weak (low) or strong (high)
angle dependence: ( D = 1.00, 0.10 to 10.00 ) --  2
Type 1 if decreasing weights for angle comp: ( D = 0, 0 to 1 ) -- 0
Max number of iterations: ( D = 1, 1 to 100 ) --  10
G:  !
G:  ! CREATE IMAGE OF PRUNED SEGMENTS
G:  !
G:  CHNSIF RADC3AP1.ID < RADC3.AP1, RADC3.CC (AX)
G:  !
G:  ! FORM NEAR LINEAR GROUPS USING THE ITERATIVE END POINT FIT
G:  ! PARTITIONING ALGORITHM
G:  !
G:  CHNLIN RADC3.AP2, RADC3.AC2 < RADC3.AP1, RADC3.CC (CA)
MAXIMUM DISTANCE THRESHOLD ( D = 1, 0 to 10 ) --  2
MAX # OF RECORDS IN OUTPUT PROP FILE ( D = 1, 1 to 100000 ) -- 10000
G:  !
G:  ! FIT EACH GROUP TO A LINE USING EDGE POSITION AND EDGE
G:  ! ORIENTATION INFORMATION
G:  !
G:  PRPLIN RADC3.AL1 < RADC3.AP2, RADC3.AC2, RADC3.CB5 (R)
Enter zero-crossing radius ( D = 0.70, 0.00 to 1.00 ) --  0.5
G:  !
G:  ! CREATE CHAIN CODE FILE FOR FITTED LINES
G:  !
G:  LINCHN RADC3.CC1 < RADC3.AL1 (S)
G:  !
G:  ! CREATE IMAGE OF FITTED SEGMENTS
G:  !
G:  CHNSIF RADC3AL1.ID < RADC3.AL1, RADC3.CC1 (ASX)
G:  !
G:  ! CREATE IMAGE OF LINE SEGMENT ORIENTATION SIGNIFICANCE
G:  ! ( INVERSE OF LINE ORIENTATION STANDARD DEVIATION )
G:  !
G:  CHNSIF RADC3.TSD < RADC3.AL1, RADC3.CC1 (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  56
G:  EXSIF RADC3.TSD
Examine SIF D.01 7/26/79
A: PROT OFF
A:
 MID 18 0
A: DONE
```

```
G:  EQUANT RADC3TSD.SIF < RADC3.TSD
Number of quantized levels for output image ( D = 255, >= 1 ) --  256
G:  !
G:  ! CREATE IMAGE OF LINE DISTANCE SIGNIFICANCE
G:  ! ( INVERSE OF DISTANCE STANDARD DEVIATION )
G:  !
G:  CHNSIF RADC3.RSD < RADC3.AL1, RADC3.CC1 (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  57
G:  EXSIF RADC3.RSD
Examine SIF D.01 7/26/79
A: PROT OFF
A:
 MID 18 0
A: DONE
G:  EQUANT RADC3RSD.SIF < RADC3.RSD
Number of quantized levels for output image ( D = 255, >= 1 ) --  256
G:  !
G:  ! EXTRACT LINES OF LENGTH GREATER OR EQUAL TO 5 PIXELS
G:  !
G:  XTRACT RADC3.LL1 < RADC3.AL1 (ACR)
Enter column number ( D = 1, 1 to 62 ) --  39
ENTER MINIMUM VALUE OF THE KEY  ( D = 0 ) --  50
ENTER MAXIMUM VALUE OF THE KEY ( D = 51, >= 51 ) --  999999
G:  !
G:  ! CREATE IMAGE OF LINE SEGMENTS ID
G:  !
G:  CHNSIF RADC3LL1.ID < RADC3.LL1, RADC3.CC1 (ASX)
G:  !
G:  ! EXTRACT CORNERS BETWEEN PAIR OF LINE SEGMENTS
G:  !
G:  LINJCT RADC3LL1.JL1, RADC3LL1.AJ1 < RADC3LL1.ID, RADC3.LL1 (C)
For RADC3LL1.ID
Enter row size for window ( D = 3, 1 to 256 ) --  7
Enter col size for window ( D = 3, 1 to 256 ) --  7
Enter corner angle tolerance (degrees) ( D = 15, 0 to 45 ) --  7
Multiple of st dev for corner angle ( D = 1.96,  >= 0.00 ) --  1.96
G:  !
G:  ! CREATE IMAGE OF SEGMENTS PARTICIPATING IN CORNERS
G:  !
G:  CHNSIF RADC3JL1.ID < RADC3LL1.AJ1, RADC3.CC1 (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  59
G:  !
G:  ! EXTRACT CULTURAL OBJECTS
G:  !
G:  LCHAIN RADC3LL1.AC1, RADC3LL1.AD1 < RADC3LL1.JL1, RADC3LL1.AJ1 (B)
Enter angle tolerance (degrees) ( D = 7, 0 to 45 ) --  7
Multiple of st dev for angle tolerance ( D = 1.96,  >= 0.00 ) --  1.96
Width of histogram smoothing window ( D = 5, 0 to 90 ) --  5
Max gap between collinear segments ( D = 32.00,  >= 0.00 ) --  32.
```

```
Max lat separation between collin segm ( D = 2.00,  >= 0.00 ) -- 2.
Maximum object side size ( D = 102.40,  >= 0.00 ) --  100.
Max gap at ends before completion ( D = 10.00,  >= 0.00 ) --  10.
G:  !
G:  ! CREATE IMAGE OF SEGMENTS BELONGING TO BRIGHT CULTURAL OBJECTS
G:  ! BEFORE CULTURAL OBJECT FITTING TAKES PLACE
G:  !
G:  XTRACT RADC3LL1.AE1 < RADC3LL1.AD1 (ACR)
Enter column number ( D = 1, 1 to 62 ) --  61
ENTER MINIMUM VALUE OF THE KEY  ( D = 0 ) --  1
ENTER MAXIMUM VALUE OF THE KEY ( D = 2, >= 2 ) --  999999
G:  CHNSIF RADC3AE1.ID < RADC3LL1.AE1, RADC3.CC1 (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  61
G:  !
G:  ! CREATE IMAGE OF BRIGHT CULTURAL OBJECTS
G:  !
G:  LINCHN RADC3AC1.CHN < RADC3LL1.AC1 (S)
G:  CHNSIF RADC3AC1.ID < RADC3LL1.AC1, RADC3AC1.CHN (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  61
G:  EXIT
 Elapsed CPU time:       0 00:46:01.46
```

```
Welcome to the Ephemeral World of Gipsy (Version 4.22 05/14/87)

G:  RUN MASTER2.RUN
G:  !
G:  ! THIS RUN FILE PERFORMS LOW-LEVEL AND MID-LEVEL COMPUTER VISION
G:  ! TASKS NECESSARY FOR CULTURAL OBJECT EXTRACTION.
G:  ! THE ORIGINAL IMAGE ( RADC5.SIF ) IS A SUNNYVALE, CALIFORNIA,
G:  ! SUBURBAN SCENE.
G:  !
G:  ! GENERATE 11 BANDS OF THE CENTRAL NEIGHBORHOOD LOCAL CUBIC FACET
G:  !
G:  CNCFCT RADC5.CB5 < RADC5.SIF (E)
Odd number of rows for window ( D = 5, 1 to 21 ) --  5
Odd number of columns for window ( D = 5, 1 to 21 ) --  5
Number of bits ( D = 14, 10 to 18 ) --  18
G:  !
G:  ! RUN HARALICK'S ZERO-CROSSING OF SECOND DIRECTIONAL DERIVATIVE
G:  ! EDGE OPERATOR. USE THE INTEGRATED DIRECTIONAL DERIVATIVE
G:  ! GRADIENT DEFINITION AND COMPUTE GRADIENT THRESHOLD AUTOMATICALLY
G:  !
G:  EDGE RADC5.EI5 < RADC5.CB5 (AOTI)
Radius for 2nd derivative ( D = 0.70, 0.00 to 2.00 ) --  0.5
Confidence limit for edge direction ( D = 0.95, 0.00 to 1.00 ) -- 0.95
G:  !
G:  ! THIN EDGES BY SUPPRESSING NON-MAXIMUM EDGE MAGNITUDES IN THE
G:  ! GRADIENT DIRECTION
G:  !
G:  THNEDG RADC5.TN5 < RADC5.EI5
Enter background symbol used in input ( D = 1, 0 to 99 ) --  1
Enter band number for edge strength ( D = 1, 1 to 4 ) --  1
Enter band number for edge direction ( D = 2, 1 to 4 ) --  2
Enter band number for symbolic edge map ( D = 4, 1 to 4 ) --  4
G:  !
G:  ! COMPLEMENT BINARY EDGE IMAGE
G:  !
G:  BOOL RADC5.LN5 < RADC5.TN5, RADC5.TN5
(1) A & B      (2) A & ^B    (3) ^A & B
(4) ^A & ^B    (5) A ! B     (6)  A ! ^B
(7) ^A ! B     (8) ^A ! ^B   (9)  A XOR B
(10) ^(A XOR B)     (11) Minimum ( A, B )
(12) Maximum ( A, B )
Enter function number ( D = 1, 1 to 12 ) --  8
G:  !
G:  ! CREATE PROPERTY FILE AND CHAIN CODE FILE OF LINKED EDGES
G:  !
G:  ARCSEG RADC5.ARC, RADC5.CC < RADC5.LN5 (S)
NUMBER OF ARC LABELS TO PROCESS ( D = 1, >= 1 ) --  1
ENTER ARC LABEL ( D = 1, >= 1 ) --  1
# PROP RECS IN FRAC OF IMAGE POINTS ( D = 0.10, 0.010 to 1.00 ) -- 0.1
```

```
G:  !
G:  ! CREATE IMAGE OF LINKED EDGES
G:  !
G:  CHNSIF RADC5ARC.ID < RADC5.ARC, RADC5.CC (AX)
G:  !
G:  ! PRUNE EDGE SEGMENTS BY LENGTH
G:  !
G:  PRLINK RADC5.AP1 < RADC5.ARC, RADC5.CC
Max size which can be deleted: ( D = 10, >= 1 ) --  7
Min size which can be retained: ( D = 1, 1 to 7 ) --  1
No of pixels for computing angle: ( D = 3, 1 to 7 ) --  3
Weak (low) or strong (high)
angle dependence: ( D = 1.00, 0.10 to 10.00 ) --  2
Type 1 if decreasing weights for angle comp: ( D = 0, 0 to 1 ) -- 0
Max number of iterations: ( D = 1, 1 to 100 ) --  10
G:  !
G:  ! CREATE IMAGE OF PRUNED SEGMENTS
G:  !
G:  CHNSIF RADC5AP1.ID < RADC5.AP1, RADC5.CC (AX)
G:  !
G:  ! FORM NEAR LINEAR GROUPS USING THE ITERATIVE END POINT FIT
G:  ! PARTITIONING ALGORITHM
G:  !
G:  CHNLIN RADC5.AP2, RADC5.AC2 < RADC5.AP1, RADC5.CC (CA)
MAXIMUM DISTANCE THRESHOLD ( D = 1, 0 to 10 ) --  2
MAX # OF RECORDS IN OUTPUT PROP FILE ( D = 1, 1 to 100000 ) -- 10000
G:  !
G:  ! FIT EACH GROUP TO A LINE USING EDGE POSITION AND EDGE
G:  ! ORIENTATION INFORMATION
G:  !
G:  PRPLIN RADC5.AL1 < RADC5.AP2, RADC5.AC2, RADC5.CB5 (R)
Enter zero-crossing radius ( D = 0.70, 0.00 to 1.00 ) --  0.5
G:  !
G:  ! CREATE CHAIN CODE FILE FOR FITTED LINES
G:  !
G:  LINCHN RADC5.CC1 < RADC5.AL1 (S)
G:  !
G:  ! CREATE IMAGE OF FITTED SEGMENTS
G:  !
G:  CHNSIF RADC5AL1.ID < RADC5.AL1, RADC5.CC1 (ASX)
G:  !
G:  ! CREATE IMAGE OF LINE SEGMENT ORIENTATION SIGNIFICANCE
G:  ! ( INVERSE OF LINE ORIENTATION STANDARD DEVIATION )
G:  !
G:  CHNSIF RADC5.TSD < RADC5.AL1, RADC5.CC1 (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  56
G:  EXSIF RADC5.TSD
Examine SIF D.01 7/26/79
A: PROT OFF
A:
 MID 18 0
A: DONE
```

```
G:  EQUANT RADC5TSD.SIF < RADC5.TSD
Number of quantized levels for output image ( D = 255, >= 1 ) --  256
G:  !
G:  ! CREATE IMAGE OF LINE DISTANCE SIGNIFICANCE
G:  ! ( INVERSE OF DISTANCE STANDARD DEVIATION )
G:  !
G:  CHNSIF RADC5.RSD < RADC5.AL1, RADC5.CC1 (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  57
G:  EXSIF RADC5.RSD
Examine SIF D.01 7/26/79
A: PROT OFF
A:
 MID 18 0
A: DONE
G:  EQUANT RADC5RSD.SIF < RADC5.RSD
Number of quantized levels for output image ( D = 255, >= 1 ) --  256
G:  !
G:  ! EXTRACT LINES OF LENGTH GREATER OR EQUAL TO 5 PIXELS
G:  !
G:  XTRACT RADC5.LL1 < RADC5.AL1 (ACR)
Enter column number ( D = 1, 1 to 62 ) --  39
ENTER MINIMUM VALUE OF THE KEY  ( D = 0 ) --  50
ENTER MAXIMUM VALUE OF THE KEY ( D = 51, >= 51 ) --  999999
G:  !
G:  ! CREATE IMAGE OF LINE SEGMENTS ID
G:  !
G:  CHNSIF RADC5LL1.ID < RADC5.LL1, RADC5.CC1 (ASX)
G:  !
G:  ! EXTRACT CORNERS BETWEEN PAIR OF LINE SEGMENTS
G:  !
G:  LINJCT RADC5LL1.JL1, RADC5LL1.AJ1 < RADC5LL1.ID, RADC5.LL1 (C)
For RADC5LL1.ID
Enter row size for window ( D = 3, 1 to 300 ) --  7
Enter col size for window ( D = 3, 1 to 300 ) --  7
Enter corner angle tolerance (degrees) ( D = 15, 0 to 45 ) --  10
Multiple of st dev for corner angle ( D = 1.96,  >= 0.00 ) --  1.96
G:  !
G:  ! CREATE IMAGE OF SEGMENTS PARTICIPATING IN CORNERS
G:  !
G:  CHNSIF RADC5JL1.ID < RADC5LL1.AJ1, RADC5.CC1 (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  59
G:  !
G:  ! EXTRACT CULTURAL OBJECTS
G:  !
G:  LCHAIN RADC5LL1.AC1, RADC5LL1.AD1 < RADC5LL1.JL1, RADC5LL1.AJ1 (B)
Enter angle tolerance (degrees) ( D = 7, 0 to 45 ) --  7
Multiple of st dev for angle tolerance ( D = 1.96,  >= 0.00 ) --  1.96
Width of histogram smoothing window ( D = 5, 0 to 90 ) --  5
Max gap between collinear segments ( D = 32.00,  >= 0.00 ) --  32.
```

```
Max lat separation between collin segments ( D = 2.00,  >= 0.00 )-- 2.
Maximum object side size ( D = 120.00,  >= 0.00 ) --  120.
Max gap at ends before completion ( D = 10.00,  >= 0.00 ) --  10.
G:  !
G:  ! CREATE IMAGE OF SEGMENTS BELONGING TO BRIGHT CULTURAL OBJECTS
G:  ! BEFORE CULTURAL OBJECT FITTING TAKES PLACE
G:  !
G:  XTRACT RADC5LL1.AE1 < RADC5LL1.AD1 (ACR)
Enter column number ( D = 1, 1 to 62 ) --  61
ENTER MINIMUM VALUE OF THE KEY  ( D = 0 ) --  1
ENTER MAXIMUM VALUE OF THE KEY ( D = 2, >= 2 ) --  999999
G:  CHNSIF RADC5AE1.ID < RADC5LL1.AE1, RADC5.CC1 (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  61
G:  !
G:  ! CREATE IMAGE OF BRIGHT CULTURAL OBJECTS
G:  !
G:  LINCHN RADC5AC1.CHN < RADC5LL1.AC1 (S)
G:  CHNSIF RADC5AC1.ID < RADC5LL1.AC1, RADC5AC1.CHN (ACSX)
ENTER PROPERTY # FOR OUTPUT LABELS ( D = 1, 1 to 62 ) --  61
G:  EXIT
 Elapsed CPU time:      0 01:14:55.9
```

The vita has been removed from the scanned document