AUTOMATIC DETECTION OF ROADS IN SPOT SATELLITE IMAGES

by

Sujata Das

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
Master of Science
in
Computer Science and Applications

APPROVED:

Dr. Roger W. Ehrich, Chairman	Dr. Layne T. Watson
Dr. James B. Campbell, Jr.	Dr. John W. Roach

May, 1988

Blacksburg, Virginia

AUTOMATIC DETECTION OF ROADS IN SPOT SATELLITE IMAGES

by

Sujata Das

Dr. Roger W. Ehrich, Chairman

Computer Science and Applications

(ABSTRACT)

The improved spatial resolution of the data from the SPOT satellite provides a substantially better basis for monitoring urban land use and growth with remote sensing than Landsat data. The purpose of this study is to delineate the road network in 20-m resolution SPOT-images of urban areas automatically. The roads appear as linear features. However, most edge and line detectors are not effective in detecting roads in these images because of the low signal to noise ratio, low contrast and blur in the imagery. For the automatic recognition of roads, a new line detector based on surface modelling is developed. A line can be approximated by a piecewise straight curve composed of short linear line-elements, called *linels*, each characterized by a direction, a length and a position. The approach to linel detection is to fit a directional surface that models the ideal local intensity profile of a linel in the least square sense. A Gaussian surface with a direction of invariance forms an adequate basis for modelling the ideal local intensity profile of the roads. The residual of the least squares fit as well as the parameters of the fit surface characterize the linel detected. The reliable performance of this line operator makes the problems of linking linels more manageable.

Acknowledgements

I would like to express my gratitude to Dr. Roger Ehrich for sharing with me his time and knowledge, to Dr. James Campbell for supporting a part of this work, to Dr. Layne Watson for providing valuable insights into various numerical techniques used and to Dr. John Roach for serving on my committee.

Thanks are also due to my family and friends for standing by me through difficult times and always making everything possible.

Table of Contents

List of Illustrations

1.0 INTRODUCTION

Remote sensing is broadly defined as the collection and interpretation of information about a target without being in physical contact with the object. The term remote sensing is commonly restricted to methods that employ electromagnetic energy as the basis for detecting and measuring target characteristics.

Aerial photography is the oldest form of remote sensing imagery and remains the most widely used method. Aerial photography analyses have played major roles in the discovery of many oil and mineral deposits around the world. These successes, using the visible portion of the electromagnetic spectrum, suggested that it might be possible to obtain comparable results by using other wavelength regions. In the 1960s, technologic developments enabled the image acquisition at other wavelengths, including near infrared, thermal infrared (IR) and microwave radiations. Development and deployment of manned and unmanned earth satellites in the 1960s provided an orbital vantage point for acquiring images of the earth.

The analysis of aerial photographs is one of the most active fields of application of pattern recognition and image understanding. This is due to the fact that a very large number of such images are routinely produced; these images are extremely complex, making visual interpretation by human photointerpreters very time consuming, subjective and costly. Fields of application of such pho-

tographs are varied, ranging from geography and cartography to agriculture, forestry, hydrology, geology and environmental monitoring.

Aerial images can be broadly categorized, based on the location of the imaging devices, as satellite images or aircraft images. Satellite images are mostly very small scale images and are used for global investigation of earth resources. On the other hand, aircraft images are of very high resolution and are suitable for the detailed analysis of the land use patterns in local districts. In urban areas, this includes map making, enumeration of the number of houses, delineation of vegetation, description of road networks, and so on. The results of such an analysis could be used to prepare a detailed land-use map and to grasp the situation of a district in order to facilitate future planning.

Landsat images are frequently analyzed for global investigations of the earth's resources. The improved (10-m and 20-m) resolution of SPOT images provides enough detail that analysts may interpret the data to extract useful urban thematic information. However, the resolution of these satellite images is coarser than those obtained from aircrafts. The imaging conditions for satellite imagery are, in general, harsher compared to those for aircraft imagery, yielding lower contrast ratios and more noise. This makes the analysis very difficult and challenging, and calls for a very robust analysis strategy.

The analysis procedure can be considerably eased if a set of reference locations can be established in the image. Knowledge of the reference above locations with respect to the locations of the various other features of the image can be used to analyze the rest of the image. The distinctive capillary appearance of roads makes road detection a promising starting point for the complete analysis procedure. Contextual information about location and orientation of the road network can be used to locate houses and other landmarks.

The purpose of this study is to delineate the road network in 20-m resolution SPOT images of urban areas automatically. These images comprise of three spectral bands in the visible and near infrared portions of the spectrum. The roads appear as dark lines in the images and have the highest

contrast in the near infrared band. This band of the image is thus used for the road network detection.

In this section, a brief introduction is given to the various fundamental characteristics of a digital image along with a brief review of the SPOT satellite program. The next section discusses in detail the earlier work done in the field of road detection. Section 3 presents the strategy chosen for detecting road networks along with photometric justifications as to the elegance of the approach. Section 4 presents the results obtained by using the strategy outlined in Section 3. Section 5 discusses the important observations made from this study.

1.1 Image Characteristics

Images can be described in terms of certain fundamental properties regardless of the wavelength at which the image is recorded. The common fundamental properties are brightness, contrast ratio and resolution. Tone, texture and other characteristics of images are functions of the fundamental properties.

1.1.1 Brightness

Remote sensing systems detect the intensity of electromagnetic radiation that an object reflects, emits or scatters at particular wavelength bands. Variations in intensity of electromagnetic radiation from the terrain are commonly displayed as variation in brightness of images.

Brightness is the magnitude of the response produced in the eye by light; it is a subjective sensation that can be determined only approximately. Variations in brightness can be calibrated with a gray

scale in which each distinguishable shade from black to white is a separate *tone*. In practise, most interpreters do not use an actual gray scale the way one would use a centimeter scale; they characterize areas on an image as light, intermediate, or dark in tone, using their own mental concept of a gray scale.

On aerial photographs, the tone of an object is primarily determined by the ability of the object to reflect incident sunlight, although atmospheric effects and spectral sensitivity of the detector are also factors. On images acquired in other wavelength regions, tone is determined by other physical properties of objects.

1.1.2 Contrast Ratio

Contrast ratio (CR) is the defined as

$$CR = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

where I_{max} and I_{min} denote the intensity levels at the brighest and the darkest spots in the image respectively. In addition to describing an entire scene, contrast ratio is also used to describe the ratio between the brightness of an object in an image and the brightness of the adjacent background. Contrast ratio is a vital factor in determining the ability to resolve and detect objects.

Images with a low contrast ratio are monotonous with uniform tones of gray. Low contrast may result from the following causes.

1. The objects and background of the scene may have a nearly uniform electromagnetic response at the particular wavelength band that the remote sensing system records. In other words, the scene has an inherently low contrast ratio.

2. Scattering of electromagnetic energy by the atmosphere can reduce the contrast ratio of a scene. Scatterring tends to make dark objects appear brighter and bright objects appear darker, thereby decreasing contrast recorded by an image. Scattering effect is most pronounced in the shorter wavelength portion of the photographic remote sensing band.

3. The remote sensing system may lack sufficient sensitivity to detect and record the contrast of the terrain. Incorrect recording techniques can also result in low-contrast images even though the scene has a high contrast when recorded by other means.

1.1.3 Resolution

The spatial resolution of a remotely sensed system is a measure of the smallest angular or linear separation between two objects that can be resolved by the sensor. For most sensor systems, it is simply the dimension of the ground-projected field of view of the sensor system that moves across the terrain collecting information.

The sensors on board the TIROS and NIMBUS satellites in the 1960s provided remotely sensed imagery with ground resolution of approximately 1000 X 1000 m and were the first to reveal the potential of space as a vantage point for earth resource observation. The MultiSpectral Scanner (MSS) and Thematic Mapper (TM) imaging systems of Landsat, developed in the 1970s and early 1980s provided imagery with spatial resolutions from 79 X 79 m to 30 X 30 m. The SPOT sensor system developed by the French Centre National d'Etudes Spatiales (CNES) produces images with spatial resolutions of 20 X 20 m and 10 X 10 m.

1.1.4 Other Characteristics of Images

Detectability is the ability of an imaging system to record the presence or absence of an object, although the identity of the object may be unknown. By knowing the resolution of an image, one can estimate the size of the smallest detectable object.

Recognizability is the ability to identify an object in an image. Objects may be detected and resolved and yet not be recognizable. For example, roads on an image appear as narrow lines that could also be railroads or canals. Unlike resolution, there are no quantitative measures for recognizability.

A *signature* is the expression of an object on an image that enables the object to be recognized. Characteristics of an object that control its interaction with electromagnetic energy determine its signature. For example, the spectral signature of an object is its brightness measured at specific wavelengths.

Texture is the frequency of change and arrangement of tones on an image.

An *interpretation key* is a characteristic or combination of characteristics that enables an object to be identified on an image. Typical keys are size, shape, tone and color. The association of different characteristics are valuable keys. On images of cities, one may recognize residential areas by the association of a dense street network, lawns and small buildings.

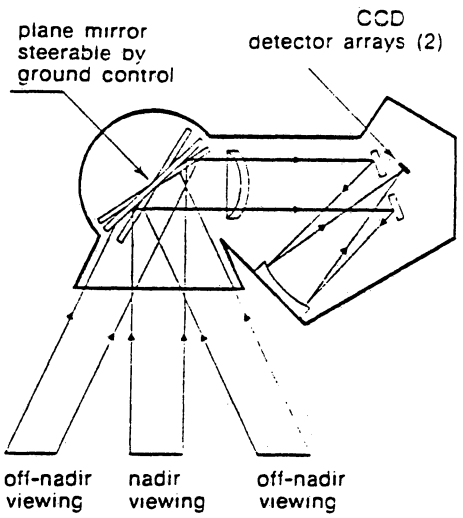
1.2 SPOT Program

In 1986, the French space agency launched a SPOT (Systeme Probatoire d'Observation de la Terre) satellite on an Arianne rocket from their launch facility in French Guiana. The orbital altitude of SPOT is 832 km, and the inclination of the orbital plane has been chosen so as to provide a sun-synchronous orbit. The south-bound orbits cross latitude 40 N at 10:00 am local sun time and ground tracks are repeated at 26 day intervals.

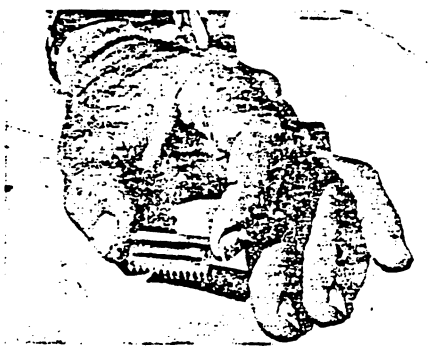
1.2.1 Imaging Systems

SPOT employs high-resolution visible (HRV) imaging systems. This imaging system is an along-track scanner. The tiltable mirror does not move during the scanning process but reflects energy into the optical system, which focuses the image onto the linear arrays of charge-coupled detectors (CCD) (Figure 1). The HRV system operates in either a multispectral mode or a high-resolution panchromatic mode. The 6000 CCDs of the high-resolution mode have a 10 by 10 m ground resolution cell and record a single panchromatic image in the 0.51 to 0.73 μm spectral band. The multispectral mode of HRV employs three arrays of 3000 CCDs with a ground resolution cell of 20 by 20 m. The multispectral mode records green, red, and reflected IR images, from which IR color images are prepared. In either mode, the ground swath is 60 km wide. SPOT employs two HRV systems that may be operated in the nadir position to acquire parallel strips of imagery with a total width of 117 km and 3 km of sidelap (Figure 2).

An innovation in SPOT is the off-nadir viewing capability (Figure 3). The mirror is tiltable up to 27 degrees on either side of vertical to obtain image strips at distances up to 475 km away from the nadir. Because of the greater viewing distance at the extreme angles, the ground resolution cells are larger and the image strips are up to 80 km wide. Off-nadir viewing provides considerable flexibility



Basic detector array (Charge Coupled Device, 1728 elements)



(2) The detectors are of the CCD (charge-coupled device) type. Each array consists of 6000 detectors arranged linearly. This forms what is known as a " push-broom " scanner since it images a complete line of the ground scene in the cross-track direction in one " look " without any mechanical scanning.

Figure 1. The SPOT Imaging System.

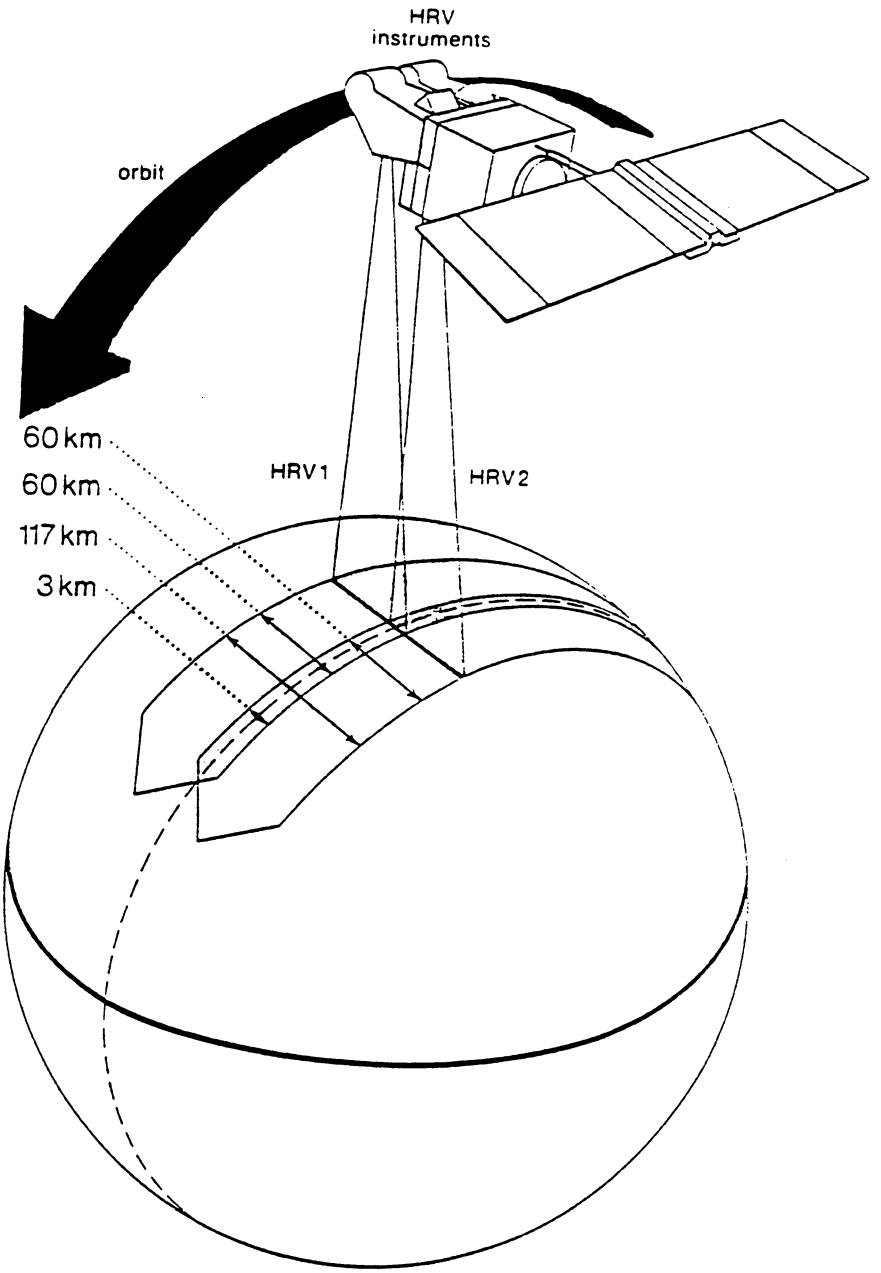


Figure 2. Nadir Viewing.

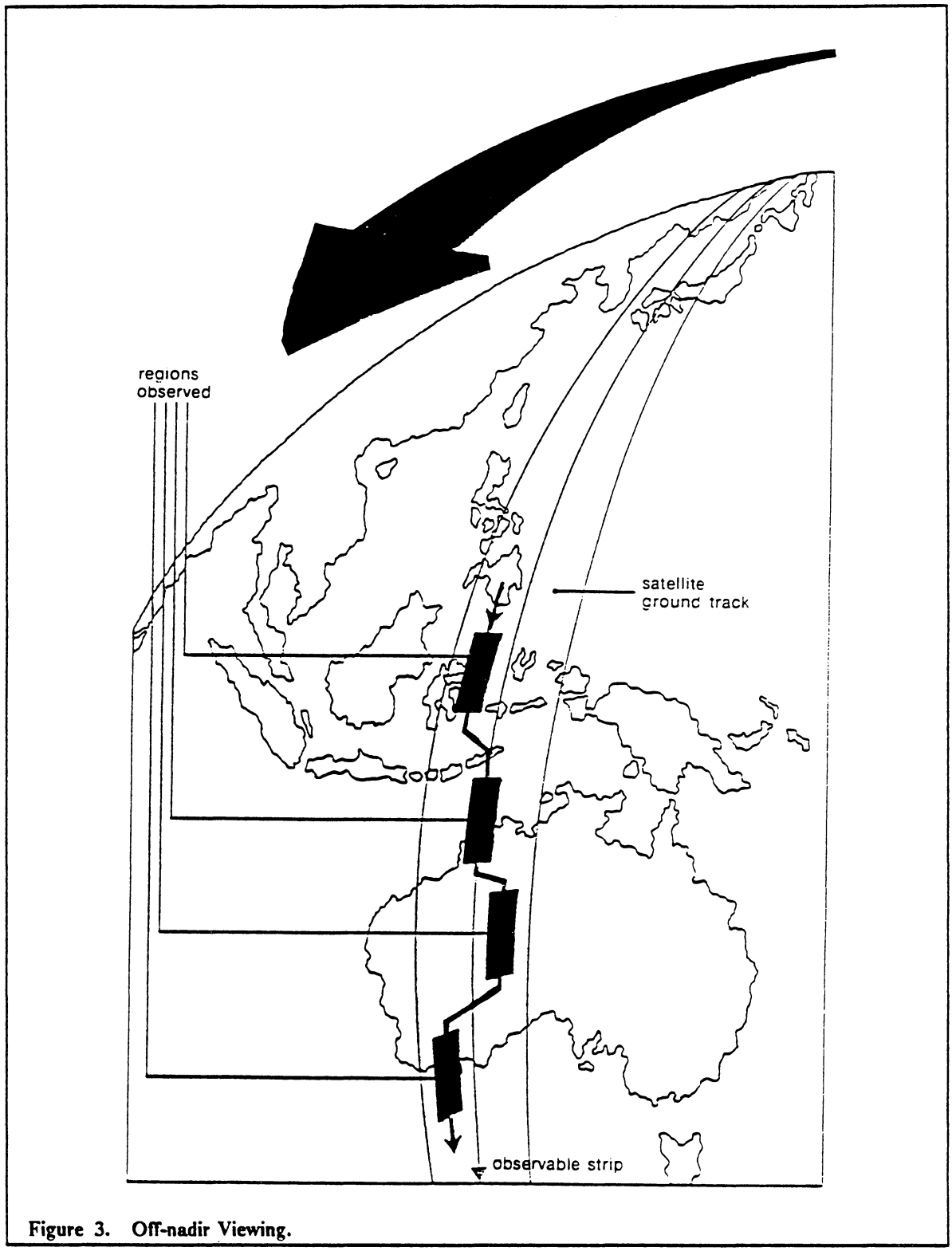


Figure 3. Off-nadir Viewing.

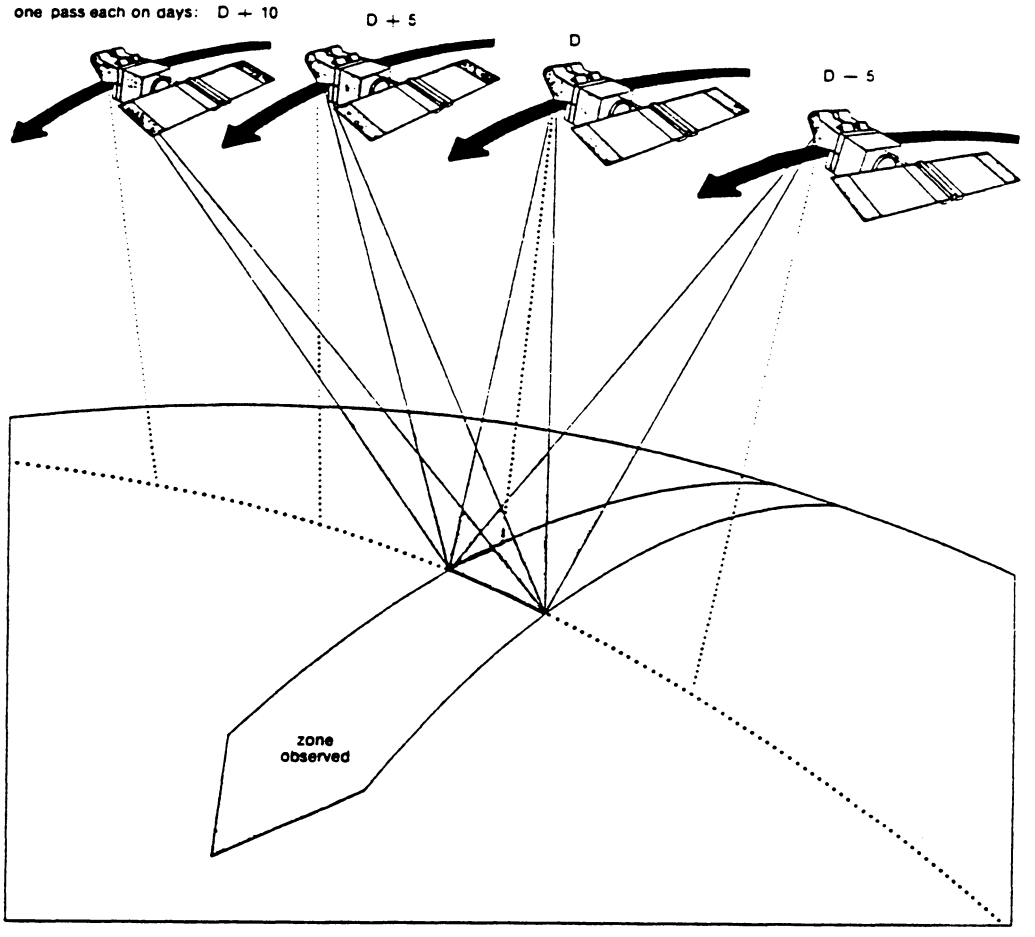


Figure 4. Revisit Capabilities.

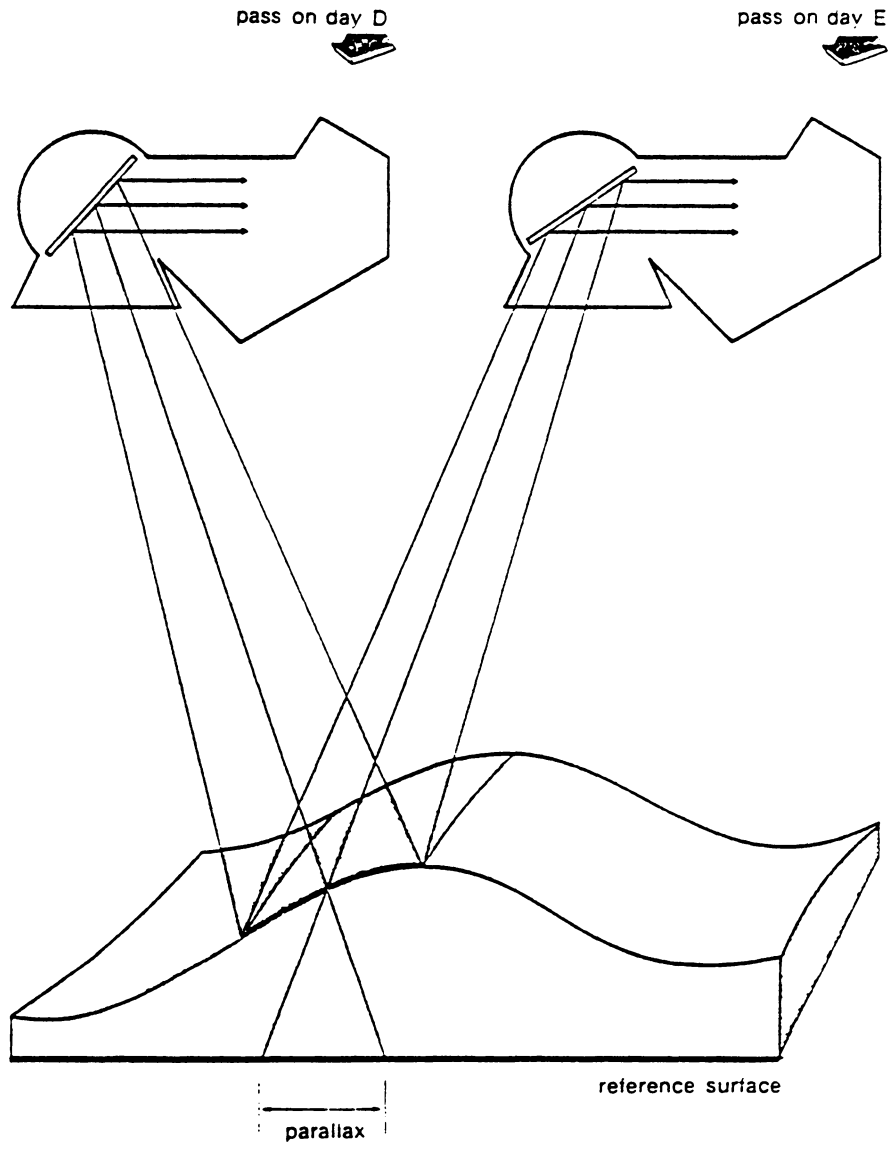


Figure 5. Stereoscopic Viewing Capabilities.

in scheduling times of image acquisition. Without this capability, an orbit path would be revisited at 26-day intervals, which is not adequate for monitoring dynamic events such as floods, volcanic eruptions, and growth cycles of crops. Localities at the equator can be imaged 7 times during a cycle; those at latitude 45 degrees can be imaged 11 times (Figure 4).

Repeated off-nadir viewing of a ground swath produces overlapping image pairs that are suitable for stereo viewing. The base-height ratio is determined by the distance between the two satellite positions divided by the altitude of 832 km. Base-height ratios range from 0.75 at the equator to 0.50 at latitude 45 degrees, which provide stereo models with strong vertical exaggeration. (Figure 5).

2.0 LITERATURE SURVEY

A great deal of work has been done towards detecting roads and linear features in aerial images. The problem of road detection reduces to the problem of edge detection and line finding in most cases. Edge detection and line finding techniques have been studied in great detail since the early days of machine perception. In spite of the large amount of previous research in this area, there is no clear basis for the choice of algorithms suitable for complex imagery. In this section, different research efforts for tackling the road detection problem are discussed.

The appearance of roads depends largely on image resolution and the spectral characteristics of the road material. Image resolution is responsible for the width of the road and the amount of internal road detail that is visible in the image. The contrast of the road with the adjacent terrain, the visible extent of the road, and any prior knowledge about the actual shape of the road and its location in the image are additional factors that have a major effect on the strategy used for locating roads in images.

2.1 *Road Detection in Aircraft Images*

The approach by Nevatia and Babu (1980) to road detection in images taken from aircraft is to detect a pair of edges running approximately parallel to each other, but with opposite orientations. The pair of edges are detected by convolving the image with a number of edge masks, thinning and thresholding these edge magnitudes, linking the edge elements based on proximity and orientation and finally, approximating the linked elements by piecewise linear segments. To find antiparallel pairs, the line segments are first sorted by their orientation. Sorting of segments is done by grouping the segments with the same orientations. A search is then made for segments that are within a certain distance of each other and whose angles are $(180 + \alpha)$ degrees apart where α is the angle of tolerance. The road is then described by an axis, a width, and its relative brightness. The proper choice of antiparallel edges that correspond to roads is a difficult problem. The antiparallel edges labelled as roads may correspond to narrow rivers or other objects that are simply formed by accidental alignment. The roads may be fragmented due to failures of the edge finding algorithms or due to shadows or obstructions. However, even these partial road descriptions have been found useful in matching these images with map-like models of the imaged area.

Similar approaches have been followed by Nagao and Matsuyama (1980) and Fua and Hanson (1987) for road detection in road images. Fua and Hanson make the road detection algorithm more robust by using model-based geometric constraints. The edges of roads are modelled as being almost straight lines in most places and can be detected locally using edge operators. They are grouped into parallel structures and linked into elongated networks that may have large scale curvature. To deal with winding curved edges, the straight edges are replaced by smoothly curved edges while the rest of the approach is retained. This approach uses a powerful technique to predict the missing components of the geometric structure e.g., if the road is fragmented or has only one side due to poor photometry or occlusions. First, the initial network of parallel edges is used to estimate

the location of the road's center and width. Next, a spline is fit at the estimated center of the road and then used to define two parallel splines that correspond to the road's edges.

A similar model is adopted by Huertas et al. (1987) to detect runways with runway markings to verify the hypothesis of the likely locations of the runways detected.

2.2 *Road Detection in Satellite Images*

In the 20-m resolution SPOT images, roads appear as linear features. Hence, the problem of road detection reduces to that of line detection in very complex imagery. Most of the earlier work on detection of roads in satellite images of lower resolution uses line detection techniques.

Bajcsy and Tavakoli (1973,1974) use crude line detection followed by iterative line following and a final thinning operation in order to yield a description of roads, with fragmentation considerably reduced. They also use the spectral properties of road materials (concrete or asphalt) to distinguish roads from rivers, streams and other such linear structures. The emphasis is more on the spectral rather than the geometric properties of the objects in the image.

Li and Fu (1974) use a syntactic approach for detecting rivers and highways in satellite images. The images are preprocessed in order to accentuate the continuity of linear features. Then, they are analyzed by tree automata. The spatial patterns of the rivers and highways are described by tree grammars. However, this method does not prove to be effective where the image is noisy, of low resolution, or when poor photometry prevails.

VanderBrug(1976) proposes the use of simple line detectors for detecting roads in satellite imagery. He experiments with the use of linear, semi-linear and nonlinear line detectors at various orientations and compares the results. It is noticed that the linear detector responds not only to lines but

also to edges. The semi-linear detector does slightly better. It does not label edges as lines but responds to background noise. The nonlinear detector, although quite insensitive to to background noise, has a tendency to produce gaps in a line. In any case, properly chosen sizes of the detectors and thresholds are essential to the performance of the detection method.

Fischler, Tenenbaum and Wolf(1981) argue that there is no currently available single coherent model suitable for reliable detection of local road presence. It is thus essential that some means for integrating information from multiple image operators and knowledge sources be devised. The scores returned from various image operators (edge and line detection operators) have little absolute meaning. They propose the use of the output from multiple edge and line operators, as well as map knowledge about the likely path of roads through an image and generic knowledge about roads (e.g., connectivity, curvature, and width constraints) for road detection. The final interpretation of the scene is achieved by using either a graph search or dynamic programming technique to optimize a global figure of merit, based on the above information sources.

The road detection problem is similar to the problem of recognizing lineaments in satellite images. Lineaments are piecewise discontinuous global lines or edges of low curvature. They generally have widely varying widths or edge characteristics and specific global and local properties that make them difficult to detect. Ehrich (1977) uses directional linear filtering applied to rectangular subregions of an image. Line segments are grown agglomeratively using a dynamic programming search technique to form complete global lineaments.

2.3 Connecting Linear Features

Linking the outputs from the early vision operators to describe a road as a single connected component is a nontrivial problem. This is more so for roads and road-like features in aerial imagery

because the actual features in the image are often broken into numerous segments or fragments due to occlusion, poor contrast with their surroundings, and changes in surface material.

Montanari (1971) deals with the problem of linking lines by constructing a line detector based on dynamic programming in which many heuristics appear embedded in the figure of merit. Martelli (1972) points out that a speed improvement is possible in the previous algorithm in the absence of serious noise if instead of doing stage by stage dynamic programming one pursues the optimal path as long as it is better than the other paths. The starting and the end points have to located for the above algorithm to work. Both the above methods have no easy way of locating the starting point. Gold and Smith (1974) employ a dynamic programming search that is constrained within an angular region called an *angular bin.* Their method also appears to be computationally very demanding, and starting points can be found only by exhaustive search.

McKeown and Pane (1985) directly deal with the problem of aligning and connecting linear features in aerial imagery. The linear features are extracted as regions. The segmentation process is designed to err towards oversegmentation of the image, since the joining of features is believed to be simpler than their decomposition. These fragments are smoothed by a Fourier approximation of the region. They are then represented by their medial axis. The physical proximity, orientation, tolerance of bumps in connection and overlap form the criteria for aligning and linking the fragments. It is important to note that it is necessary to aggregate these incomplete segmentations, filling in missing information, in order to reason about the overall scene interpretation.

3.0 A NEW OPERATOR TO DETECT ROADS

In this section, a strategy for the recognition and description of roads in 20-m resolution SPOT images is described. Although these images consist of three bands, only the infrared band is used for testing the detection method. This band has the highest contrast among all the three bands and the roads are sharper in this band. It was observed that the roads are light lines in the blue and the green bands while they are dark lines in the infrared band. Hence, the principal component image does not enhance the roads. This makes the infrared band a good data set for testing the detection process. However by using data from one band, a lot of useful spectral information is ignored. The geometric and spatial properties of roads form the basis for modelling roads. Thus any feature, e.g. river, runway, etc., that has the same distinctive geometric features as the roads is detected. Knowledge of the spectral properties of the road material can be incorporated in the detection process to distinguish roads from other linear features.

In any image understanding system, one has to compute descriptions of objects useful for their recognition. Roads have a very distinctive capillary appearance in the SPOT images. They can be categorized as being elongated line-edges with a low angle of curvature. The widths of the various roads in the SPOT image vary by a few pixels. The width as well as the brightness of each road can be taken to be approximately constant along its length.

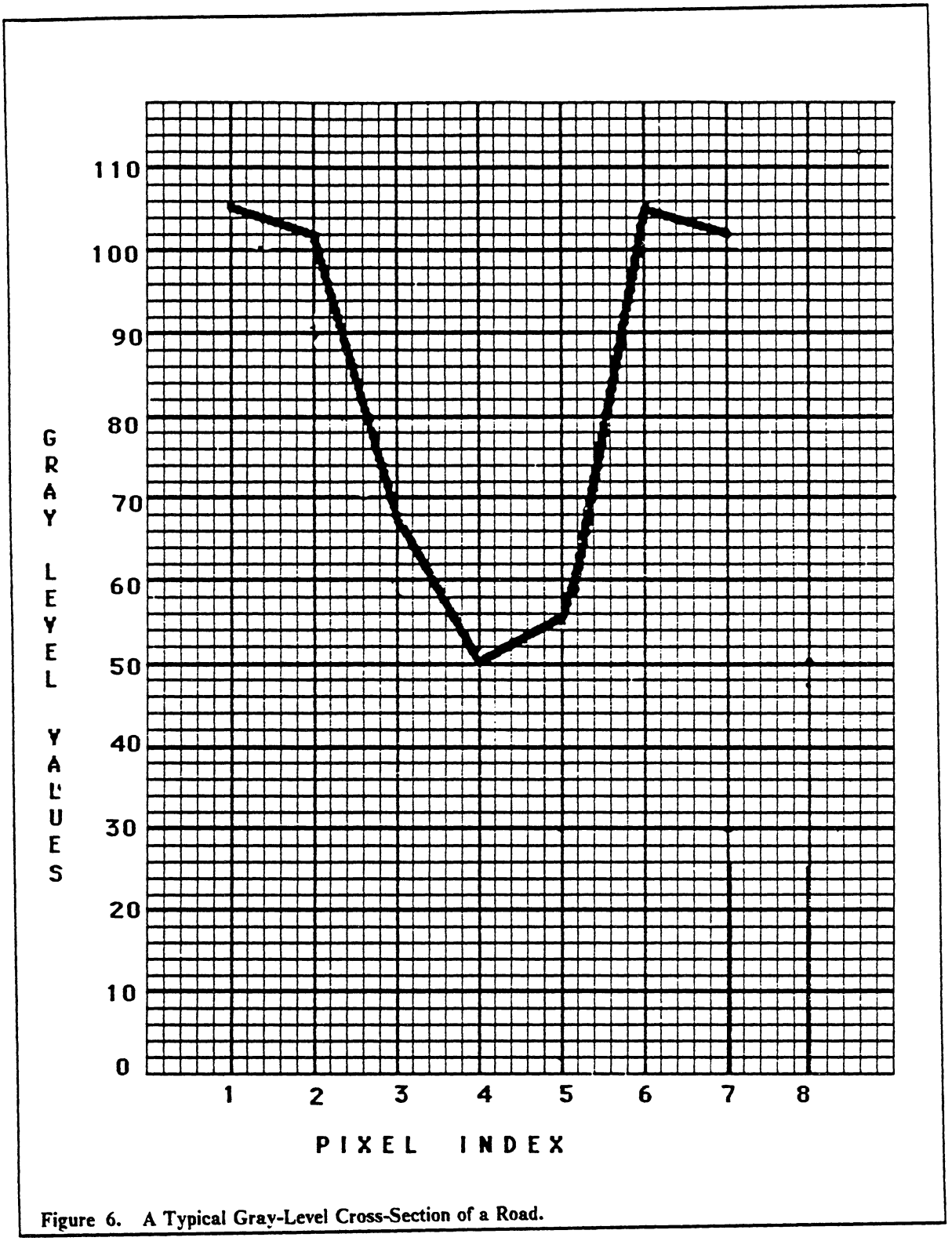


Figure 6. A Typical Gray-Level Cross-Section of a Road.

The imaging conditions, however, render these images to be of extremely low contrast, very noisy, and the line edges characterizing a road structure to be very far from an ideal linear structure. Figure 6 shows a cross-section of a typical road structure. The detection method exploits these salient features to detect the roads.

At 20-m ground resolution, the SPOT imaging device captures very little detail of the features between the two edges of the roads. An approach such as that of Nevatia and Babu can be attempted to detect a pair of **antiparallel step-edges.** However since there is not much detail visible in the roads, a single line-edge detector avoids a lot of problems encountered by Nevatia and Babu and makes the detection process more robust.

Early vision operators are very crucial to the performance of any vision system. Most modules in any conceivable vision system depend, directly or indirectly, on the performance of their edge- or region-finding operators. For example, reliable detection of edges should be expected to make the problems of linking edges more manageable. Here, the approach towards road-detection is aimed at exploiting the features of the road so as to make the output from the line-edge detection operator very reliable.

3.1 Conventional Operators

Much of the work on edge detection is based on discrete approximation to differential operators. An edge in an image corresponds to a sudden transition in the intensity surface of the underlying scene. Hence, step-edges correspond to a large first-order derivative and zero-crossings in the second-order derivative. However, the mapping from edges to edge pixels marked by derivative tests is neither one-to-one nor onto. Derivatives emphasize high frequency noise, and the higher the order of derivative that the operator relies on, the greater the sensitivity of the operator to unfiltered

noise. Thus, the operators that are based on derivatives of various orders are very ineffective in detecting roads in satellite images due to the high noise content and the low contrast of the images. Figures 9 through 14 show edges detected using different classical operators in a sample SPOT image (Scene Id: 1617275860807161250 1X; Time of Scene Center : 16:12:50; Date 08/07/86; Figures 7 and 8).

Surface fitting is among the other methods used to detect edges. Prewitt(1970) and Haralick(1984) used it as a means of estimating derivatives. Although these operators control the noise in differentiation, they have largely failed to exploit the directional character of edges. The resolution of the SPOT images is too coarse for cubic facet model [Haralick (1984)]. Further, the use of a cubic facet model may result in poor localization of the step edges in the images. According to Nalwa and Binford (1986) this is because a cubic fit forms an inadequate basis for describing an edge.

3.2 Nalwa and Binford Operator

Nalwa and Binford (1986) define a **one-dimensional** surface as one which is constant along some direction. Their operator models the intensity profile of the edges. They fit a series of **one-dimensional** surfaces to each window and accept the surface description which is adequate in the least squares sense and has the fewest number of parameters. Hyperbolic tan function was found to be an adequate model for the step-edge and its combinations adequate for roof-edges and line-edges. This tanh model has very attractive features. The tanh surface is constrained to be directional. This results in effective noise reduction without blurring the edges as severely as with other circularly symmetric smoothing operators such as the Marr-Hildreth operator. Further, it takes into account the fact that the image consists of samples of the true intensity profile "blurred" by the imaging system. For most step-edges, the tanh function with a constant is adequate:

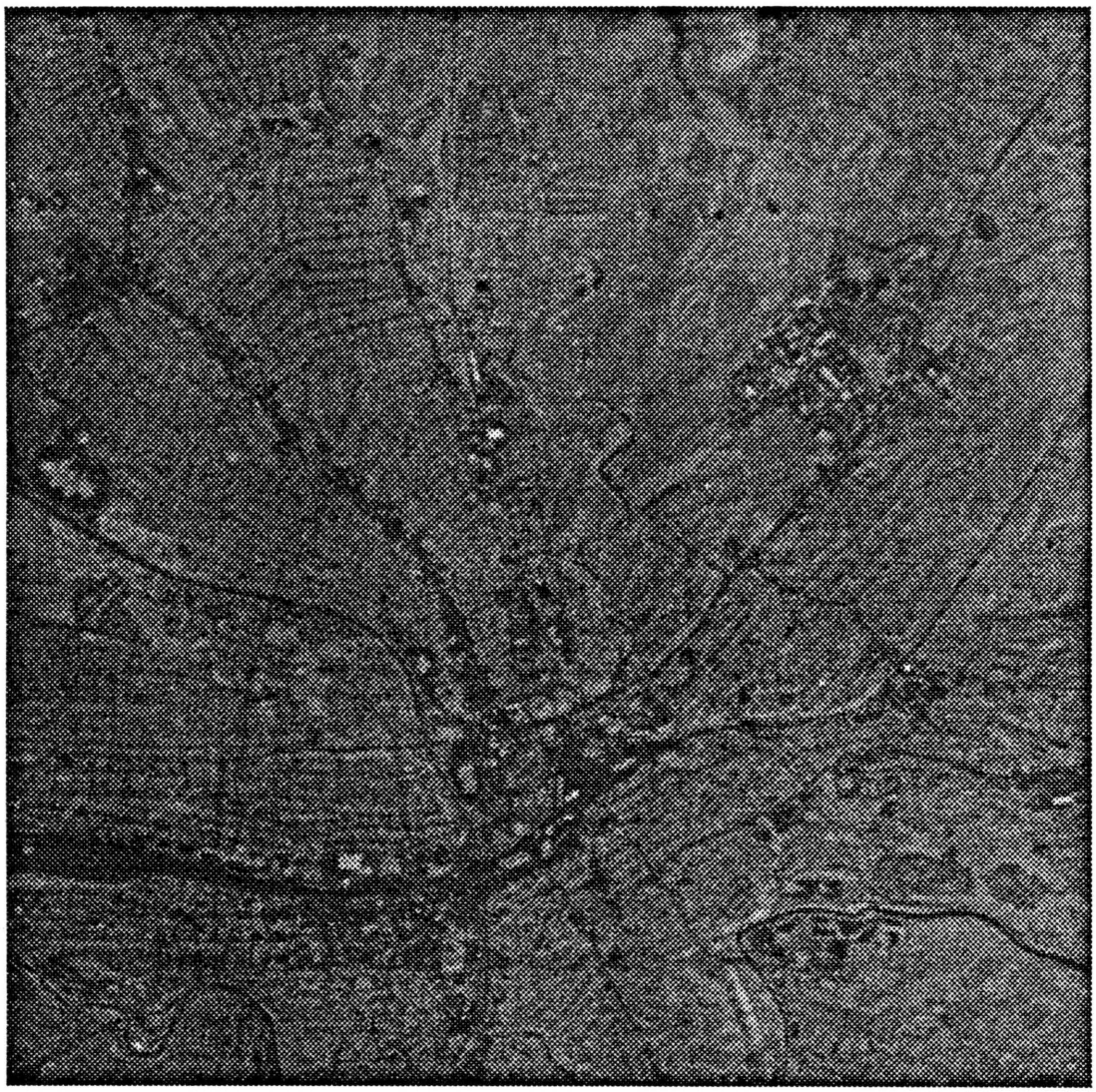


Figure 7. SPOT Image (Infrared Band) of ROANOKE.

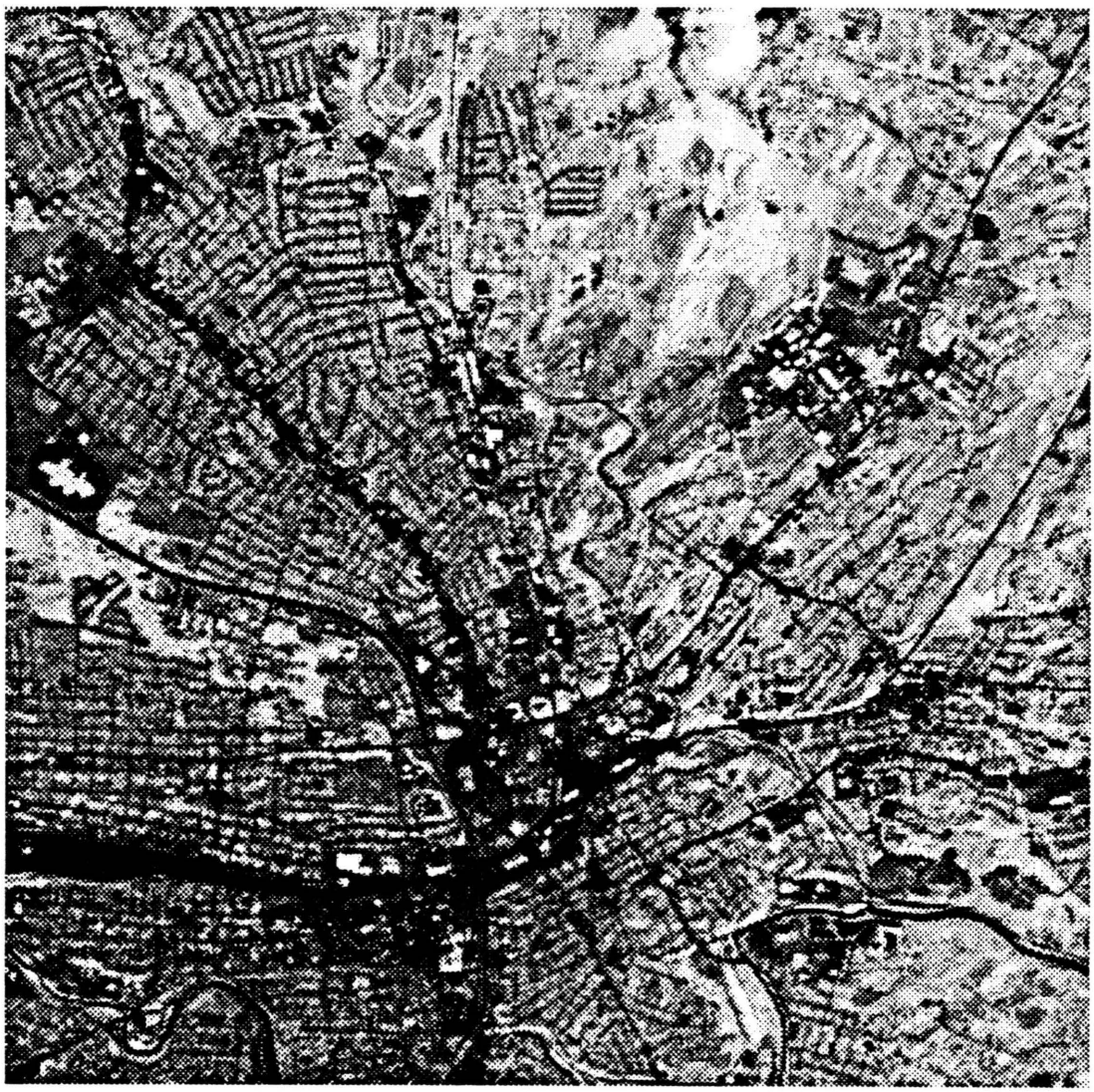


Figure 8. SPOT Image of ROANOKE After Histogram Equalization.

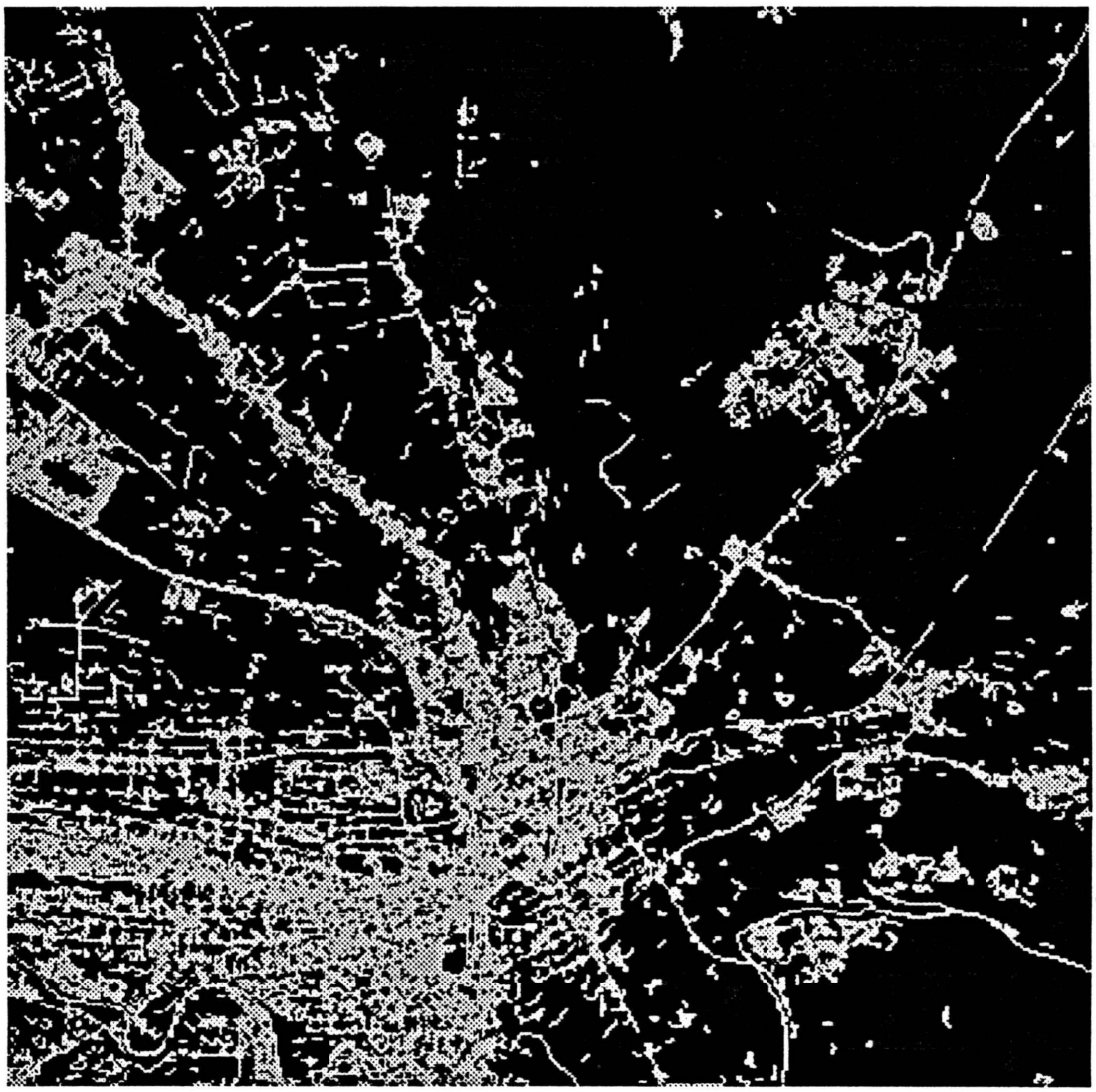


Figure 9. Edge Output from the CUBIC FACET Operator.

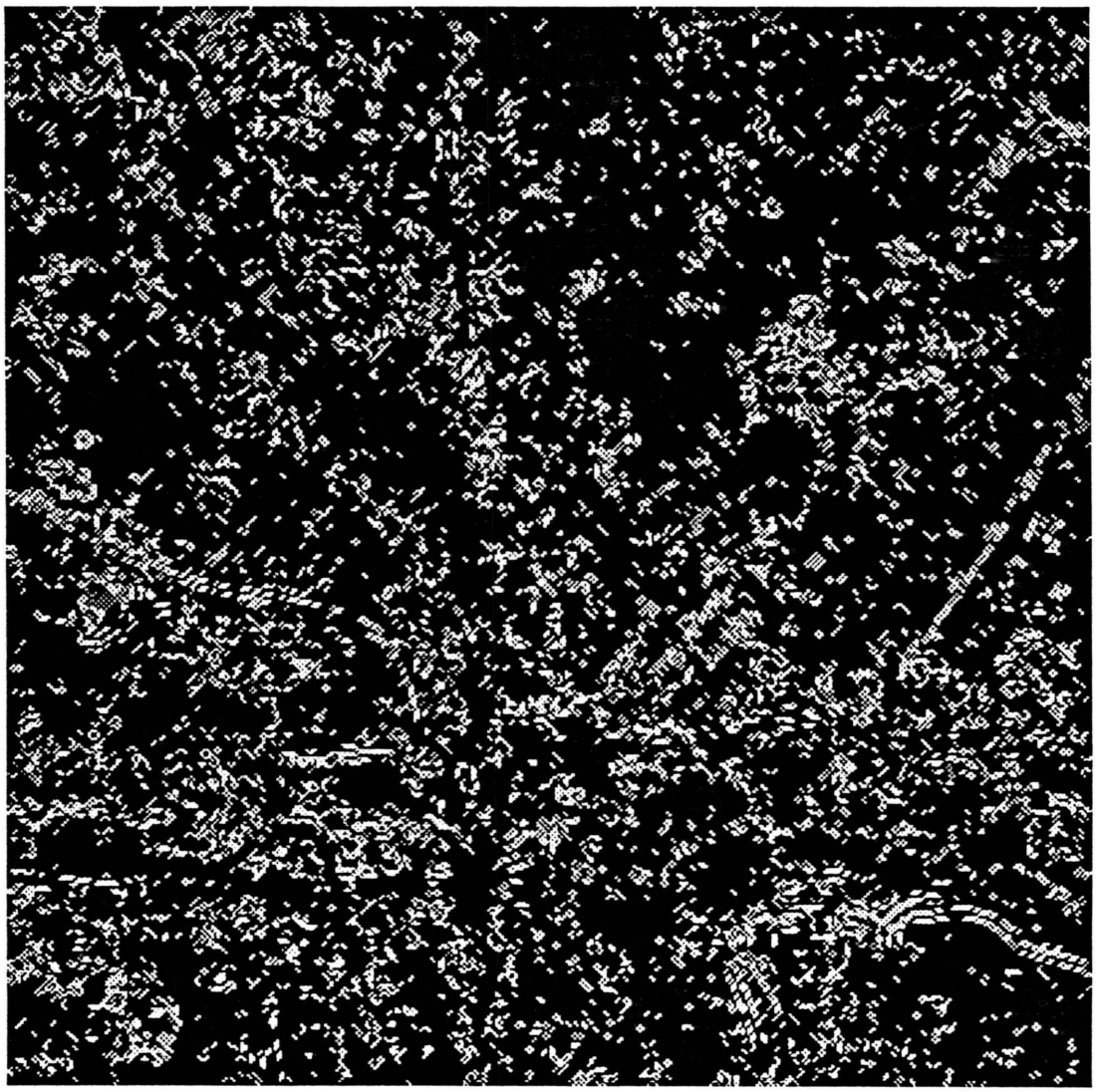


Figure 10. Edge Output from the MARR-HILDRETH Operator.

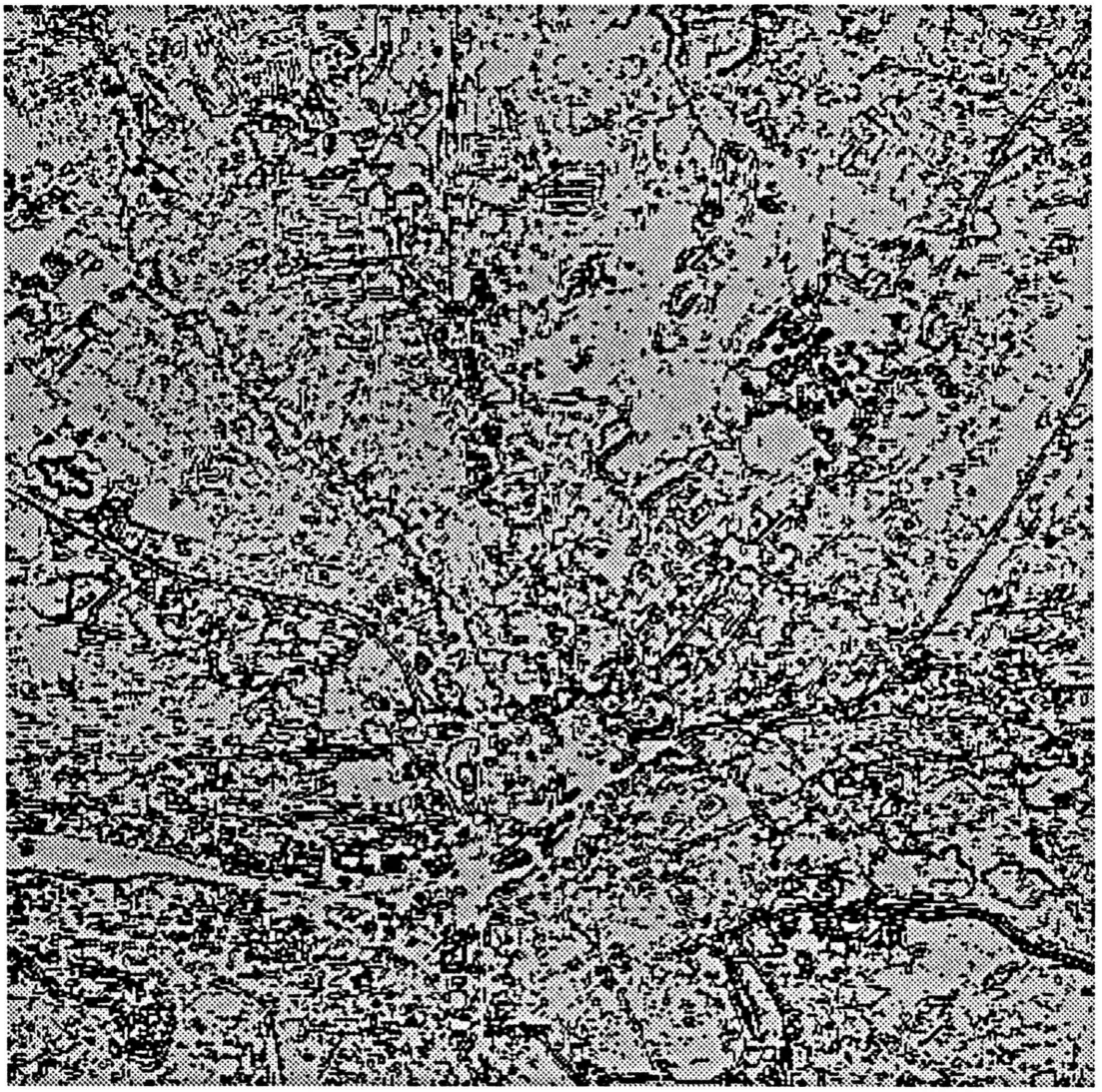


Figure 11. Edge Output from the ROBERT'S CROSS Operator.

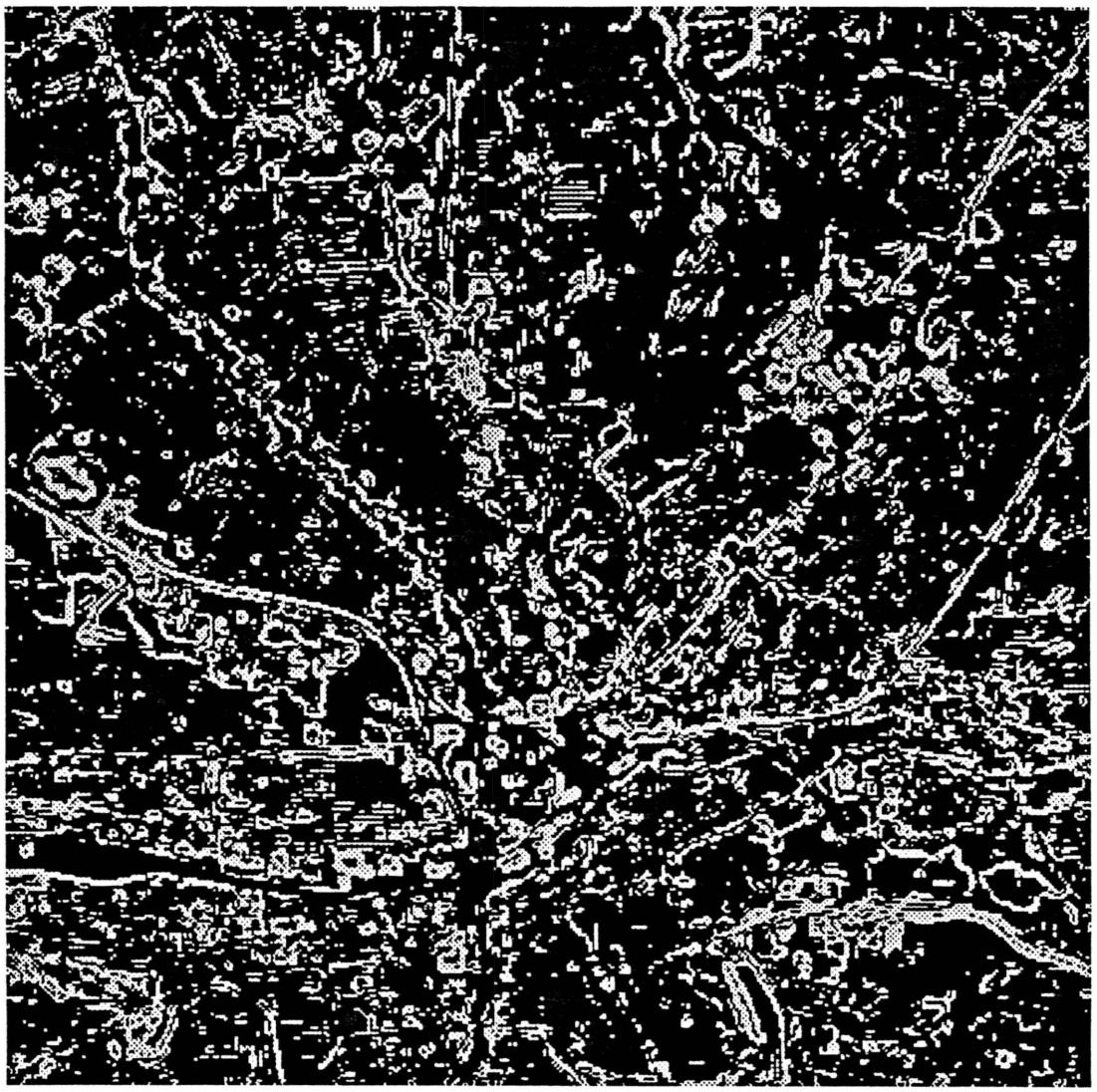


Figure 12. Edge Output from The SOBEL Operator.

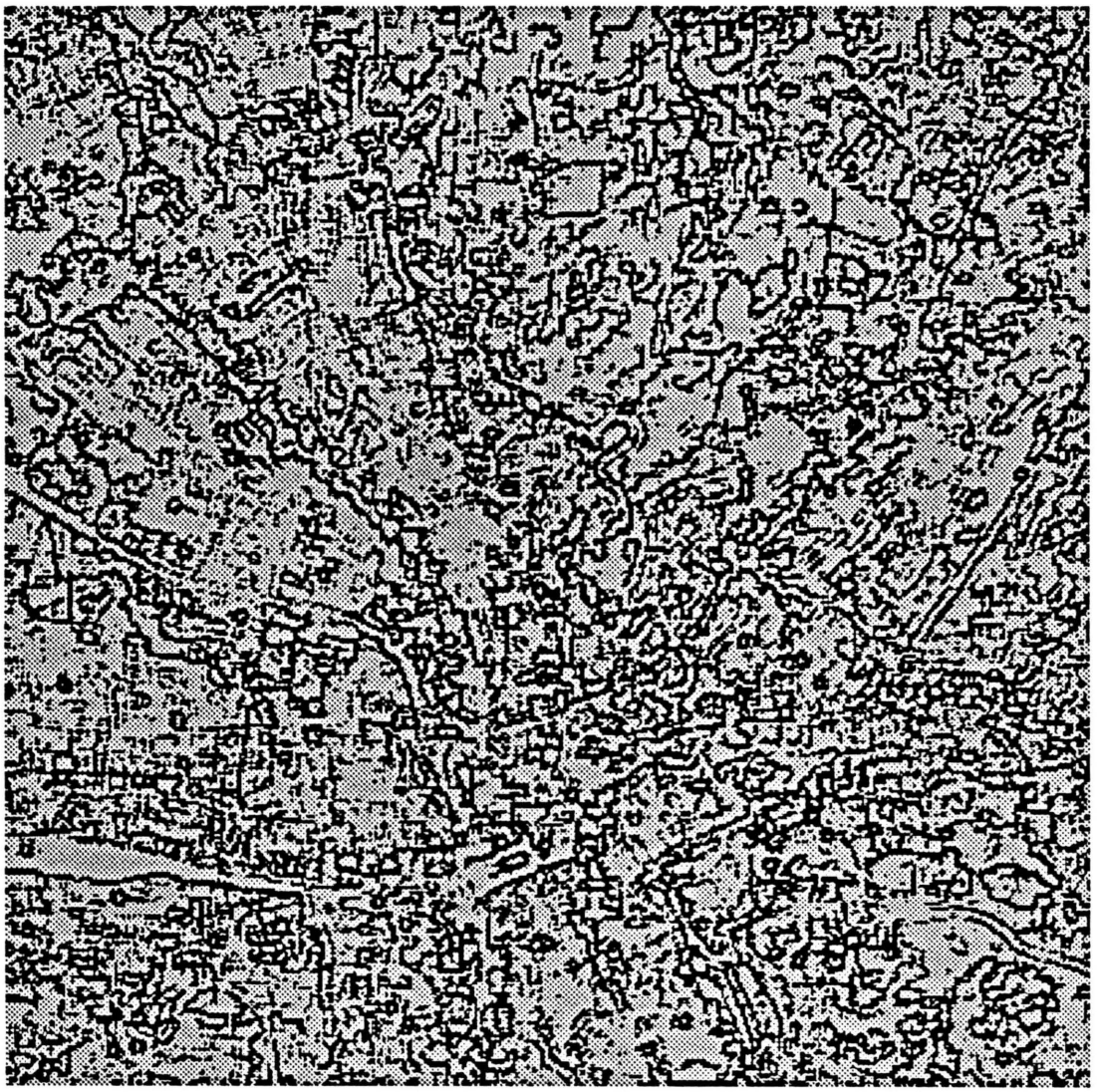


Figure 13. Edge Output from the Double Adaptive Threshold Line Operator (Watson et al. [1984]).

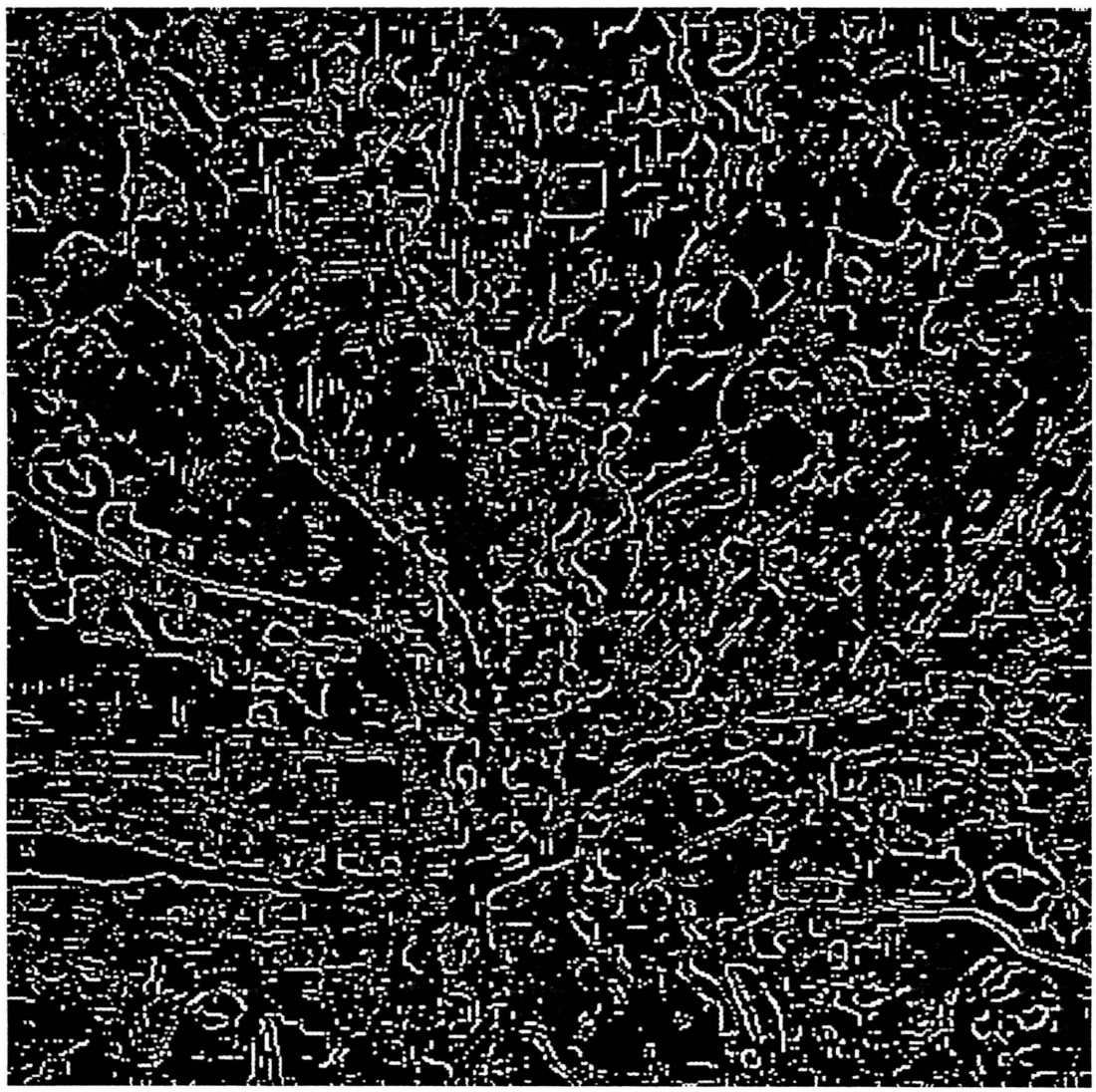


Figure 14. Edge Output from the Nevatia and Babu (1980) Operator.

$$F(z) = s \tanh(f[z + p]) + k \qquad [1a]$$

where

$$z = r \sin\theta + c \cos\theta. \qquad [1b]$$

In Eqn. [1b], (r,c) is the coordinate of the pixel in the window and θ denotes the direction perpendicular to the direction of invariance of the *one-dimensional* surface. s, p, and k in Eqn. [1a] are parameters and f is a constant determined by the blur of the imaging system. The parameters p, k, and s give useful information as to the contrast and the localization of the step-edge in the window. The angle of invariance is found first by making a rough estimate of the angle using a *one dimensional* planar surface least squares fit to the window followed by a *one dimensional* cubic fit to refine the estimate. Thresholding does not play a pivotal role in the edge detection. This operator thus gives superior performance with respect to detection, localization and resolution. However, for edges that have large deviations from the ideal step-edge, the tanh model is inadequate and it is conjectured that a cubic or a tanh with a cubic may be adequate.

Our method of detection of the line-edges is based on the approach adopted by Nalwa and Binford. Instead of using a combination of tanh and cubic functions as the model, the Gaussian function is used. The choice of such a function is based on a mathematical model of the sensor system.

3.3 A Mathematical Model of a Remote Sensing System

The sensor system measures, in a number of bands, the reflected solar radiation from or emitted thermal radiation. The remote sensing system of a MSS system is illustrated by the block diagram in Figure 15 [Kidd and Wolfe (1976)]. Light from the ground scene is attenuated and supplemented by the atmosphere before entering the sensor system. The light is reflected by the rotating mirror

through the optics and focused on the detector where it is converted to an electrical signal. The detector output is processed electronically and recorded for later image reconstruction. The block diagram for the sensor system of the SPOT satellite is similar to that of the MSS system except that mirror does not rotate.

A mathematical model, as developed by Kidd and Wolfe (1976), for the remote sensor system can be constructed by combining models of electronic and optical imaging systems with a model for radiative transfer in the Earth's atmosphere. In doing so, it is assumed that the sensor is a linear invariant system. A rather simple model of the atmosphere is seen to be adequate for system performance prediction.

Any imaging instrument with a finite sensing element or optical system will lessen the sharpness of features and have limited spatial-frequency response. Typically, the image of an impulse has a nominally Gaussian intensity distribution (Goetz et. al. (1975)). A sampled scene can be considered to be composed of an array of impulses; the image of that scene is considered to be constructed by the superposition of an array of Gaussian intensity distributions.

Further, the scene degradation caused by atmospheric turbulence is represented in the form of a transfer function by Hufnagel and Stanley (1964). The normalization transfer function is

$$T_{at}(f) = \exp[-5.82\pi^2\lambda^{-1/3}f^{5/3}S(H)/\cos\theta]$$

where λ is the spectral wavelength (of band center), f is the spatial frequency expressed in cycles/radian, θ is the nadir angle of the ground region investigated, and S(H) is an empirical function of the sensor altitude, H. This function is similar to a Gaussian curve, and the point spread function exhibits a similar shape, since Gaussian functions are Fourier transform pairs.

It is thus reasonable to expect the cross-section of a road in a SPOT image to appear like a Gaussian function because of the degraded response of the sensor system and the narrow width of the road features.

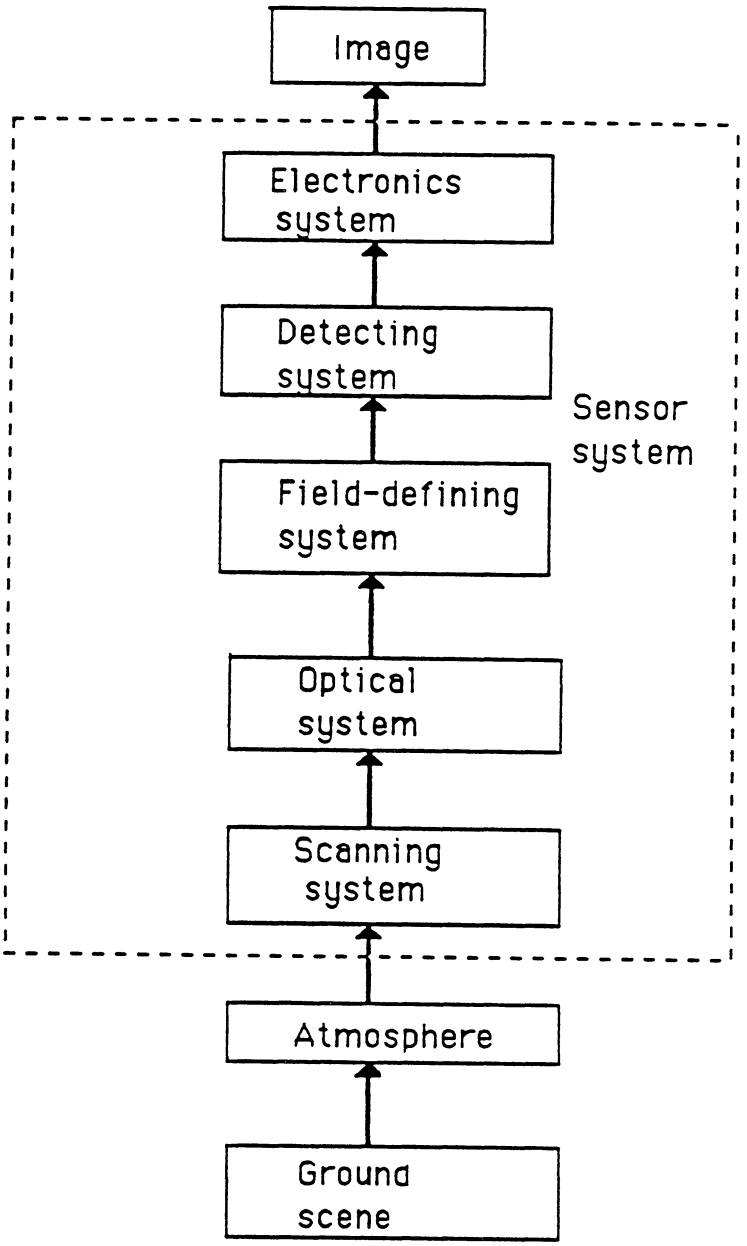


Figure 15. Block Diagram of a Sensor System.

3.4 *A Road Detector for SPOT Images*

A *one-dimensional* Gaussian surface forms an ideal model of a line-edge element in SPOT images. The strategy followed is to detect short linear line-edge elements, called *linels*, each characterized by a direction and a position, which connect together to form piecewise straight curves representing a road. In order to detect a linel, in each image window, a *one-dimensional* Gaussian surface is fit to the pixel intensities in the least squares sense. The equation of the Gaussian function is

$$G(z) = k - he^{-(w(z-p)^2)} \qquad [3a]$$

where

$$z = r\sin\theta + c\cos\theta. \qquad [3b]$$

In Eqn. [3b], (r,c) is the coordinate of a pixel in an image window and θ denotes the direction perpendicular to the direction of invariance of the *one-dimensional* surface. k, h, w, θ, and p represent parameters that have to be estimated for each pixel centred in an image window of the given image.

The least squares problem thus is:

$$\min_{k,h,p,w,\theta} \| G - Y \|_2 \qquad [4]$$

where

- $G = [g_1, g_2, ..., g_n]^T$ where $g_i = k - he^{-(w(z_i-p)^2)}$ is the one-dimensional Gaussian surface approximation pixel i with coordinates (r_i, c_i), and $z_i = r_i\sin\theta + c_i\cos\theta$.

- $Y = [y_1, ..., y_n]^T$, where y_i is the gray-level value of pixel i.

The parameters of Eqn. [3a and 3b] have a very meaningful physical interpretation. k stands for the intensity level of the background, h for the strength of the linel, w is related very definitely to the width of the linel, p stands for its position, and θ for the across direction of the road in the window. k, h, w, p and θ along with the least squares error of fit completely characterize any line-edge present in the image window.

Since roads appear as dark lines in the SPOT image, a positive value of h denotes the presence of a linel. The magnitude of h gives an indication of the strength of the linel in the window. If a linel is detected within a distance of one pixel from the central pixel, the normalized strength of the linel and the residual of the least squares fit are used to classify the linel.

Fitting the *one-dimensional* surface is equivalent to treating the data as strictly *one-dimensional* by projecting it along the direction of invariance onto a plane. By fixing the direction of invariance to the direction parallel to the direction of the road, noise in the image can be effectively reduced. The estimated width, direction and strength of the linel are very helpful in linking of the linels.

The above approach for line-edge detection is very similar to a match filtered classification technique, i.e., the closest match to noiseless representations, which in our case happens to be *one-dimensional* inverted (h > 0) Gaussian surface.

Although the modelling Gaussian function is extremely elegant, and provides us with parameters which have a very useful interpretation, it is nonlinear in its parameters. This requires solving for the parameters by an iterative process. The estimation of the solution to a nonlinear system of equations depends on the initial values assigned to the parameters. For this particular function, it was found that for certain combinations of pixel intensities in the window, the parameters converged to values that have no physical interpretation. An algorithm converges to approximately the same values if the plot of residual versus the domain of parameter estimates for the system is

bowl-shaped. A plausible explanation for the above phenomenon can be that the plot of the residual versus the domain of parameter estimates is a long trench and not a bowl. So there are many estimates of the parameters that produce the approximately the same (locally optimal) residue. This makes it impossible to obtain **the** parameter estimate that is interpretable. Moreover, a nonlinear least squares problem, such as [4], has unstable local minima because of the nonlinearity and the unpredictable combination of the intensity levels that might arise in the image window. In the next section the metamorphosis from the above model to make a model that is mathematically more tractable and maintains most of its desirable features is discussed.

4.0 IMPLEMENTATION OF THE NEW ROAD DETECTOR

The first attempt at fitting the general Gaussian surface of Eqn. [3] involved solving the nonlinear least squares problem using the LMDER subroutine of MINPACK (More' et al. [1980]). This subroutine uses a modification of the Levenberg-Marquardt method to solve the nonlinear least squares problem. But the algorithm sometimes converged to values that are not interpretable (Figure 16). It was decided to experiment with various forms of the Eqn. [3] by fixing each of the parameters in turn and observing the behaviour of LMDER.

4.1 Localization Parameter Held Constant

The localization parameter, p, in Eqn. [3], was made zero. The formulation for the function thus became

$$G_1(z) = k - he^{-(w(z)^2)} \qquad [5]$$

Note : The parameter estimates for sample problem 1 are meaningful. However, in an attempt to minimize the error of fit, LMDER gave parameter estimates for problem 2 and 3 that have no physical significance.

Sample Problem 1

DATA SET

(row,	column,	gray value)
0	-2	106
0	-1	91
0	0	58
0	1	94
0	2	109

ESTIMATES
(parameters as in Eqn. 3.)
k = 107
h = 53
p = -4.16×10^{-2}
w = 1.125
θ = 0.0

Sample Problem 2

DATA SET

(row,	column,	gray value)
0	-2	106
0	-1	67
0	0	51
0	1	55
0	2	105

ESTIMATES
(parameters as in Eqn. 3)
k = 8059.38
h = 8010.92
p = 4.9775×10^{-2}
w = 0.042
θ = 0.0

Sample Problem 3

DATA SET

(row,	column,	gray value)
0	-2	105
0	-1	67
0	0	77
0	1	74
0	2	100

ESTIMATES
(parameters as in Eqn. 3.)
k = 10636.94
h = 10568.79
p = 1.82
w = 0.028
θ = 0.0

Figure 16. Parameter Estimates to Sample Nonlinear Fitting Problems.

This implied that the algorithm tried to locate a linel positioned exactly at the center of the window. It was observed that the computed estimates were still meaningless at times.

4.2 *Width of Linel Held Constant*

Next the parameter w, which designated the width of the Gaussian, was fixed to a value w_1 that was reasonable for the test image. For the formulation

$$G_2(z) = k - he^{-(w_1(z-p)^2)} \qquad [6]$$

where w_1 is a constant and k,h, and p are parameters, LMDER mostly gave meaningful results (Figure 17). It was discovered that the parameter, w, gave a large degree of freedom to the system. In an attempt to decrement the residual and produce a globally optimal solution, a Gaussian function of large width is fitted (Figure 16). Predetermining the width of the linel in the SPOT image does not pose too many problems. The widths of the roads vary by a few pixels. In fact it was found empirically that the parameter w can be fixed to a value that represents an average of the widths of the roads, without causing too much degradation in the output of the detector.

4.3 *Background Intensity Held Constant*

It was further observed that, for the solution of eqn. [5], a very wide Gaussian function is fit and so the background value becomes very large (Figure 16). Hence, another way to tackle the problem is to make k a constant k_1. The Gaussian function then becomes

$$G_3 = k_1 - he^{-(w(z-p)^2)} \qquad [7]$$

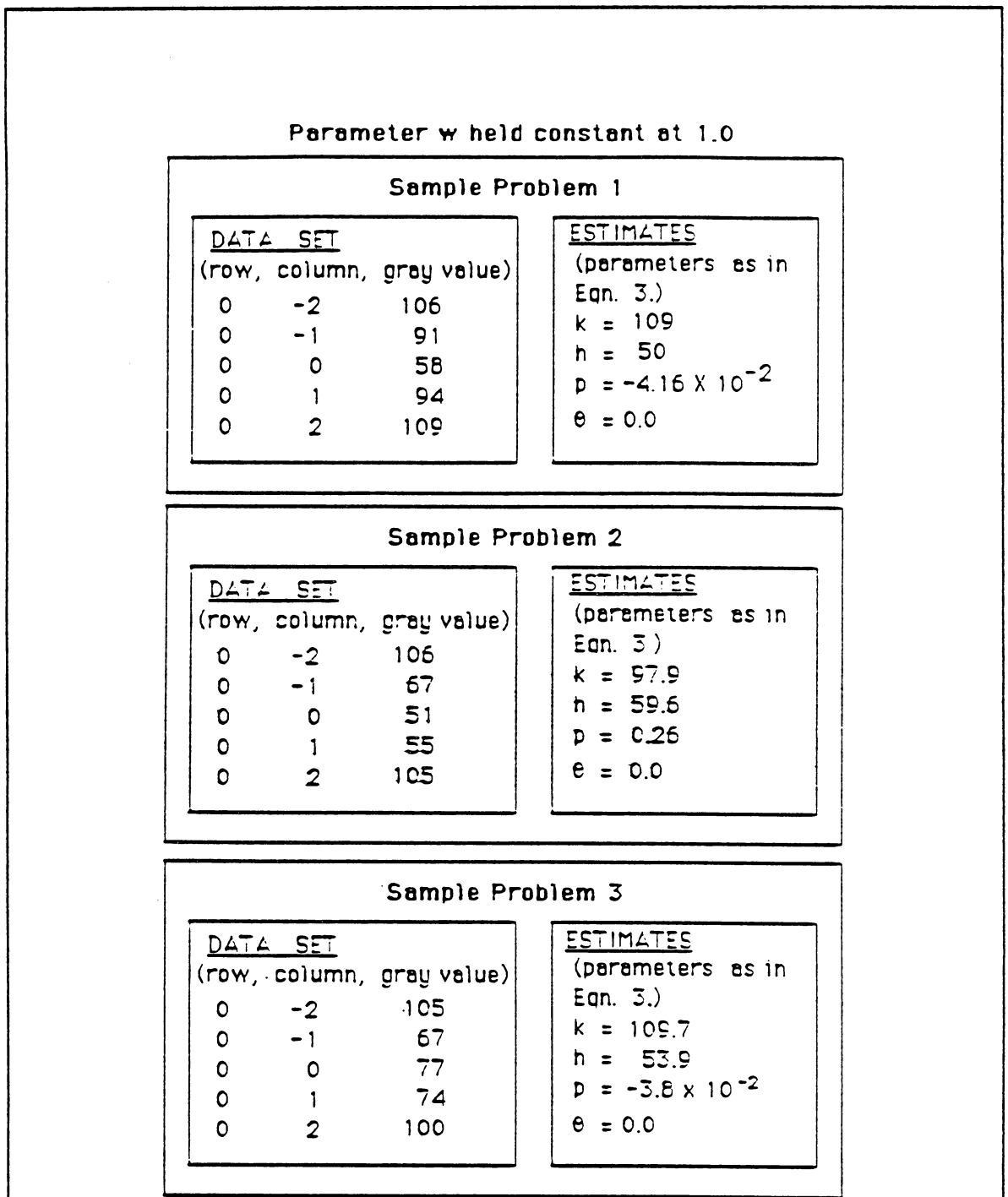


Figure 17. Parameter Estimates of Sample Nonlinear Fitting Problems with Constant Width.

where k_1 is a constant and w,p,h, and θ are parameters. It was observed that most of the results were meaningful. The value of k_1 had to be fixed to the value of the background intensity. However, the problem of estimating the background intensity beforehand is non-trivial. This is mainly due to the poor contrast of the SPOT images. It is necessary to determine k_1 accurately. This is because the parameters are highly sensitive to the values of k_1; an underestimation of k_1 always fits Gaussians of much lower strength or low magnitude of h and an overestimation of k_1 results in large positive values of h, irrespective of the intensity profile of the pixels. So, it was decided to keep k_1 a parameter of the function and hold w constant instead.

4.4 Nonlinear vs. Linear

The position and direction of invariance still make Eqn. [6] nonlinear. In an experiment to estimate the reliability and accuracy of the solutions obtained by LMDER, the position was fixed to zero. The parameters of the model then become k,h and θ. The same solutions were obtained by linearizing the model by choosing values for discrete θ and solving the equation for each distinct θ. It was observed that, sometimes, the solution using the nonlinear model LMDER converged to values that were inappropriate for our purpose. On the other hand, the linearized model gave a more accurate set of estimates (Figures 18 through 23). This finally lead us to reject the idea of using a nonlinear formulation of the problem.

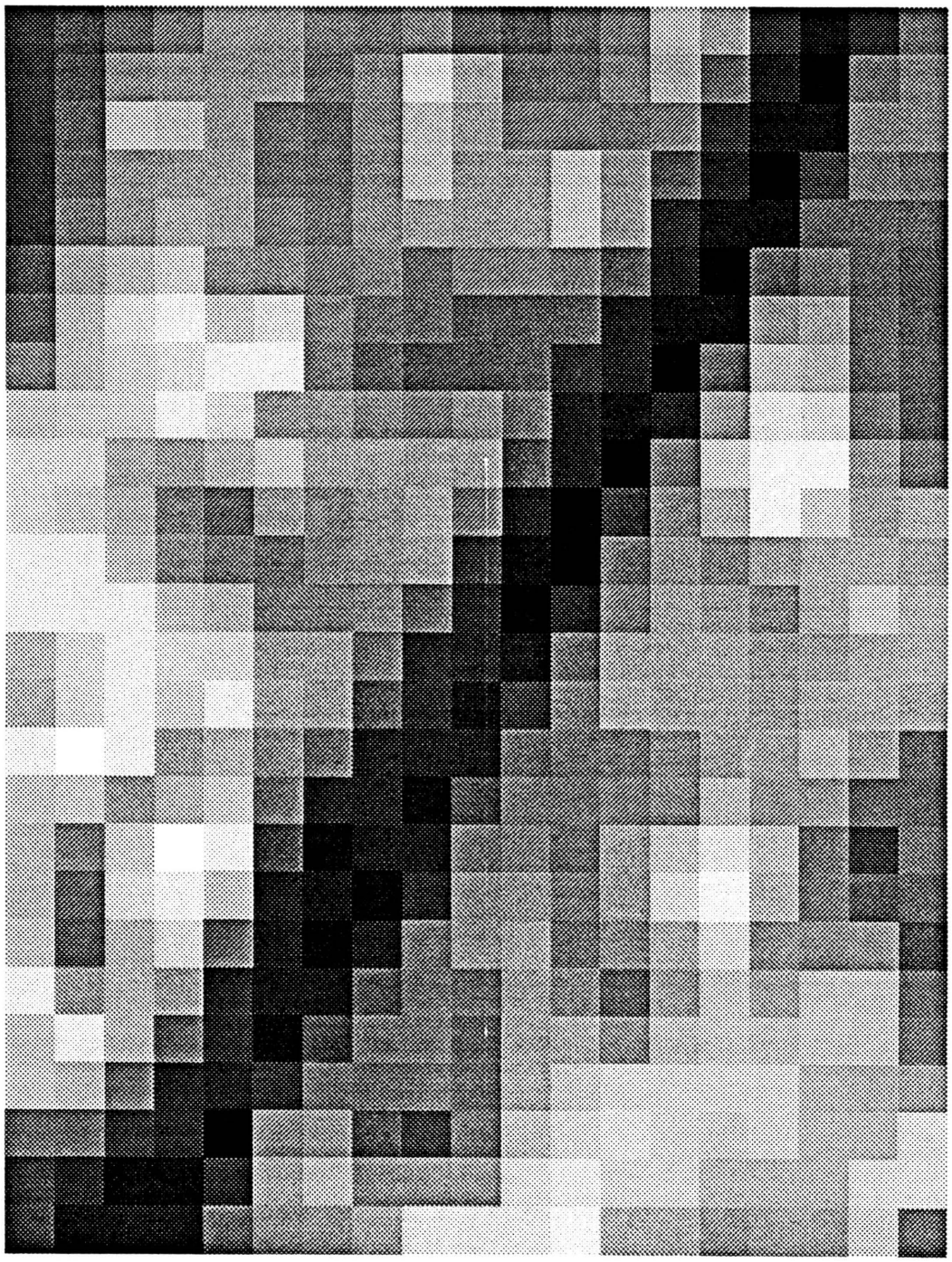


Figure 18. Sample image A for comparing solutions from Nonlinear and Linear Least Squares Methods.

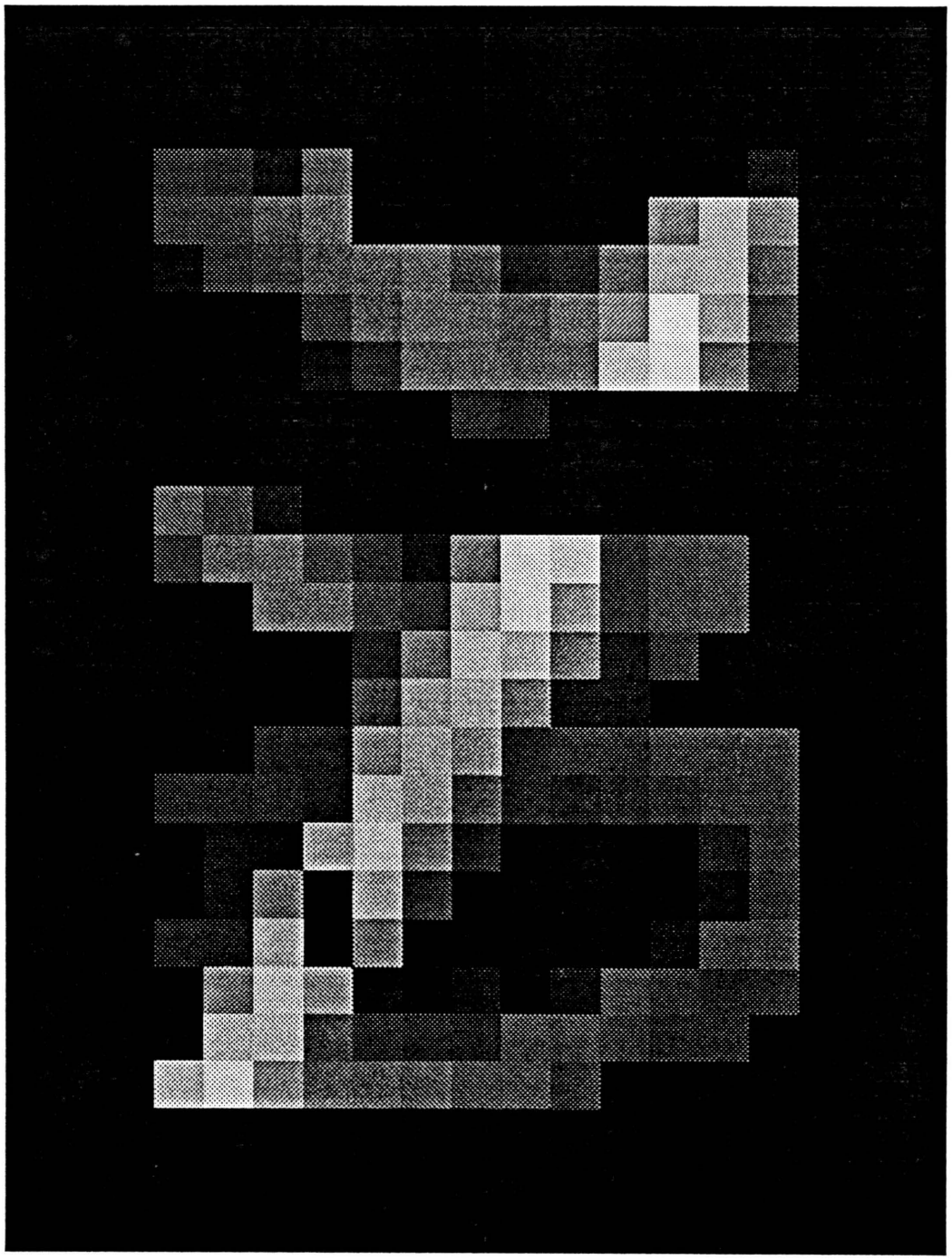


Figure 19. Heights of Gaussians from Nonlinear Least Squares fitting for image A.

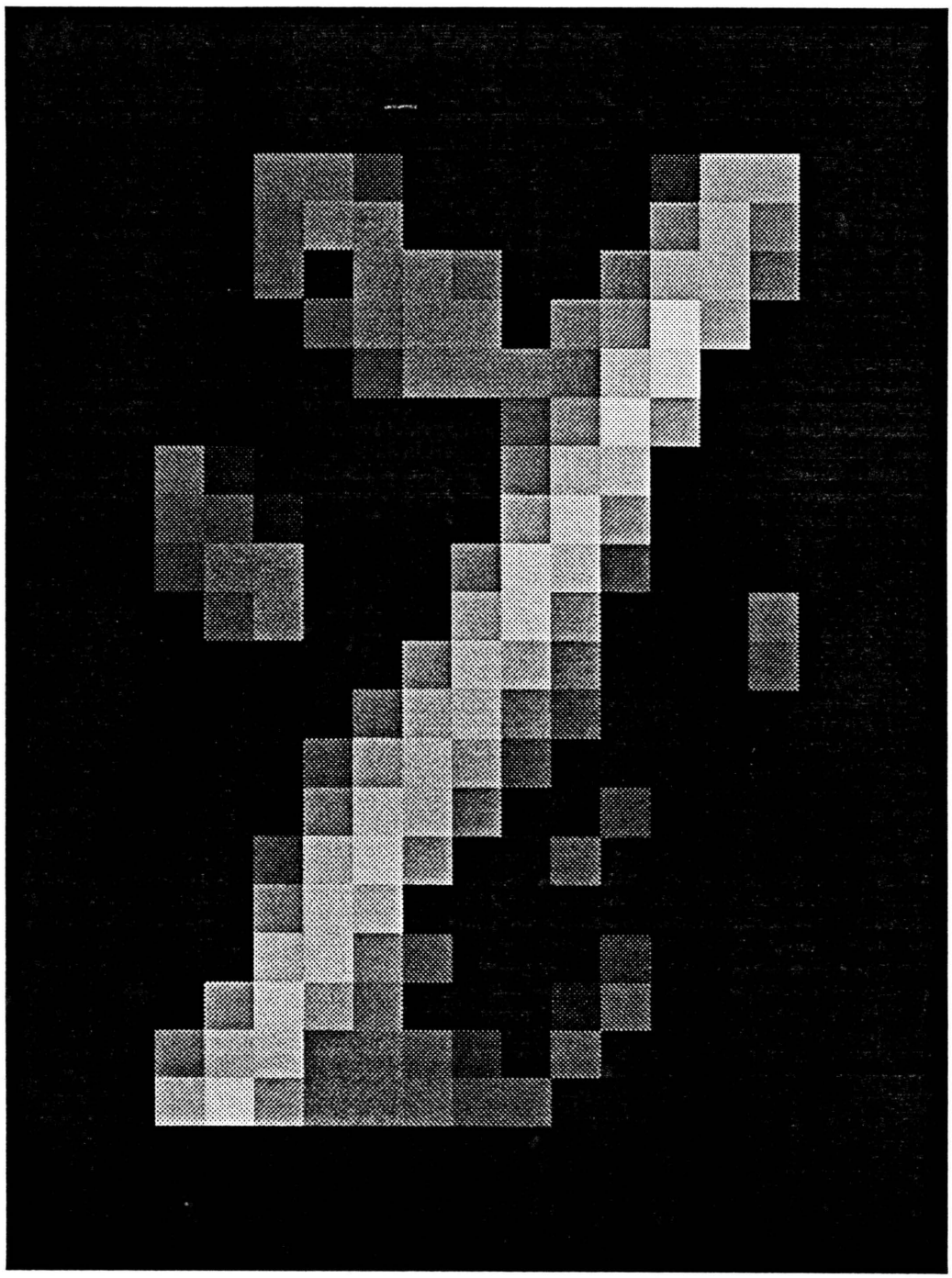


Figure 20. Heights of Gaussians from Linear Least Squares fitting for image A.

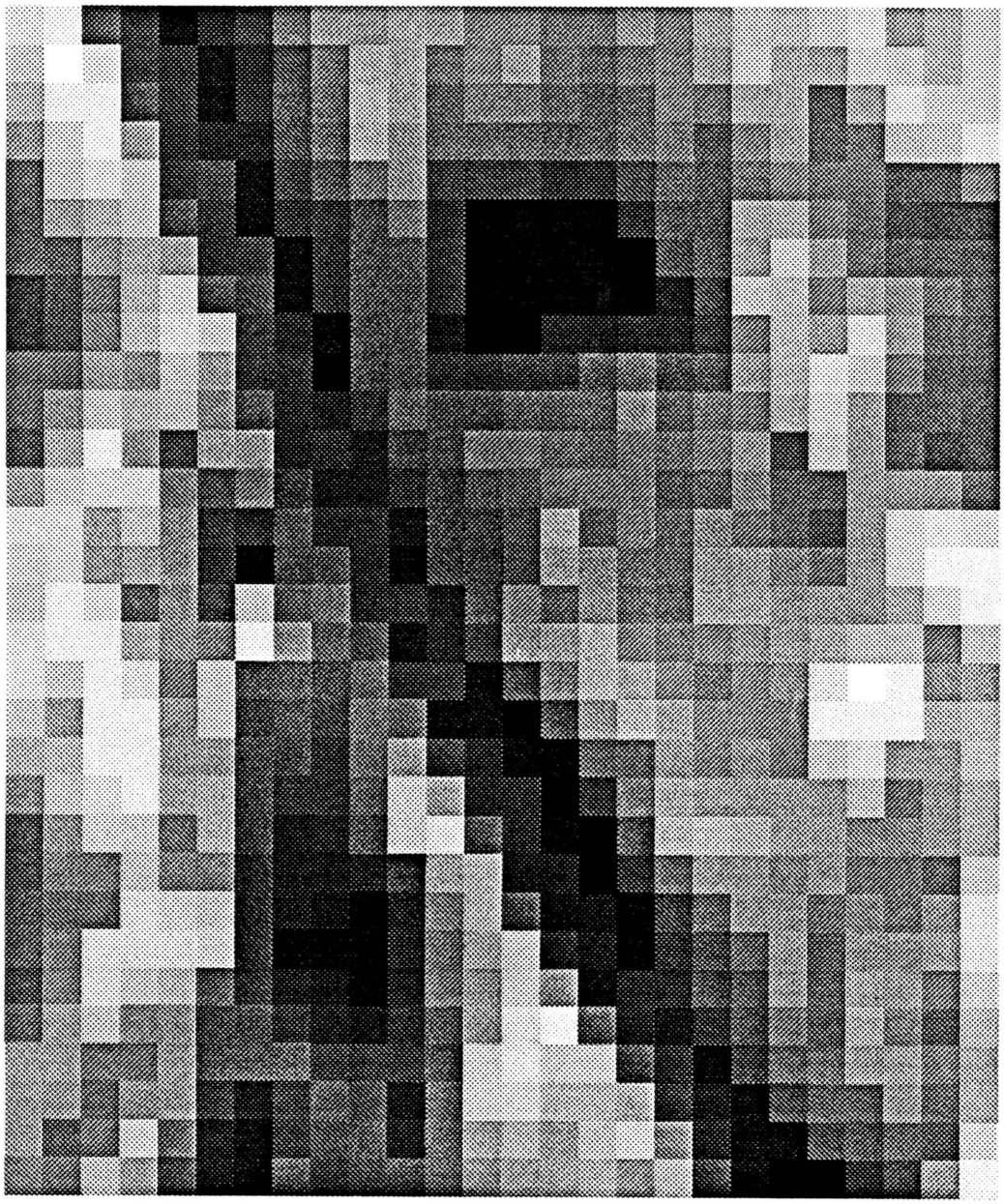


Figure 21. **Sample image B for comparing solutions from Nonlinear** and Linear Least Squares Methods.

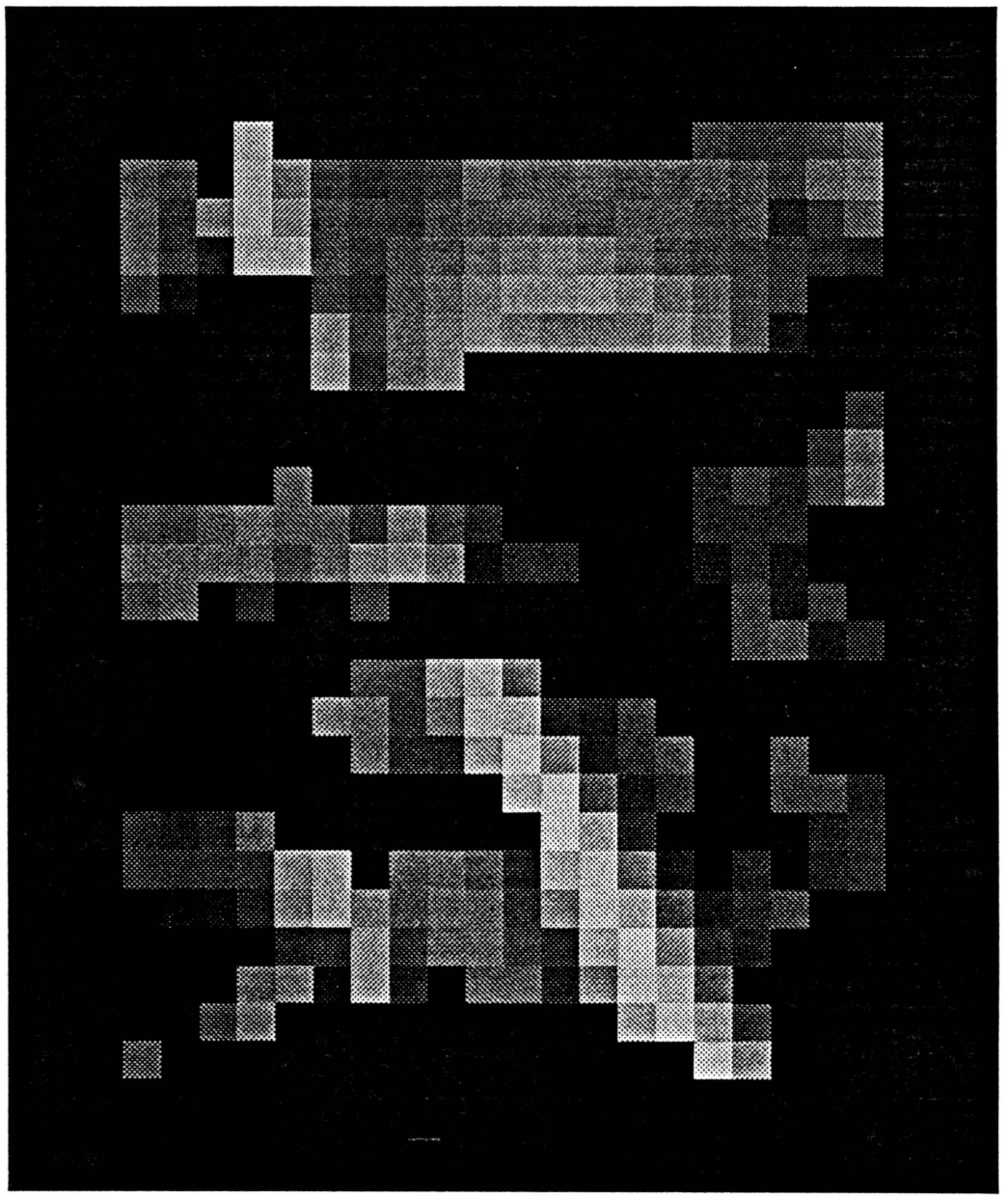


Figure 22. Heights of Gaussians from Nonlinear Least Squares fitting for image B.

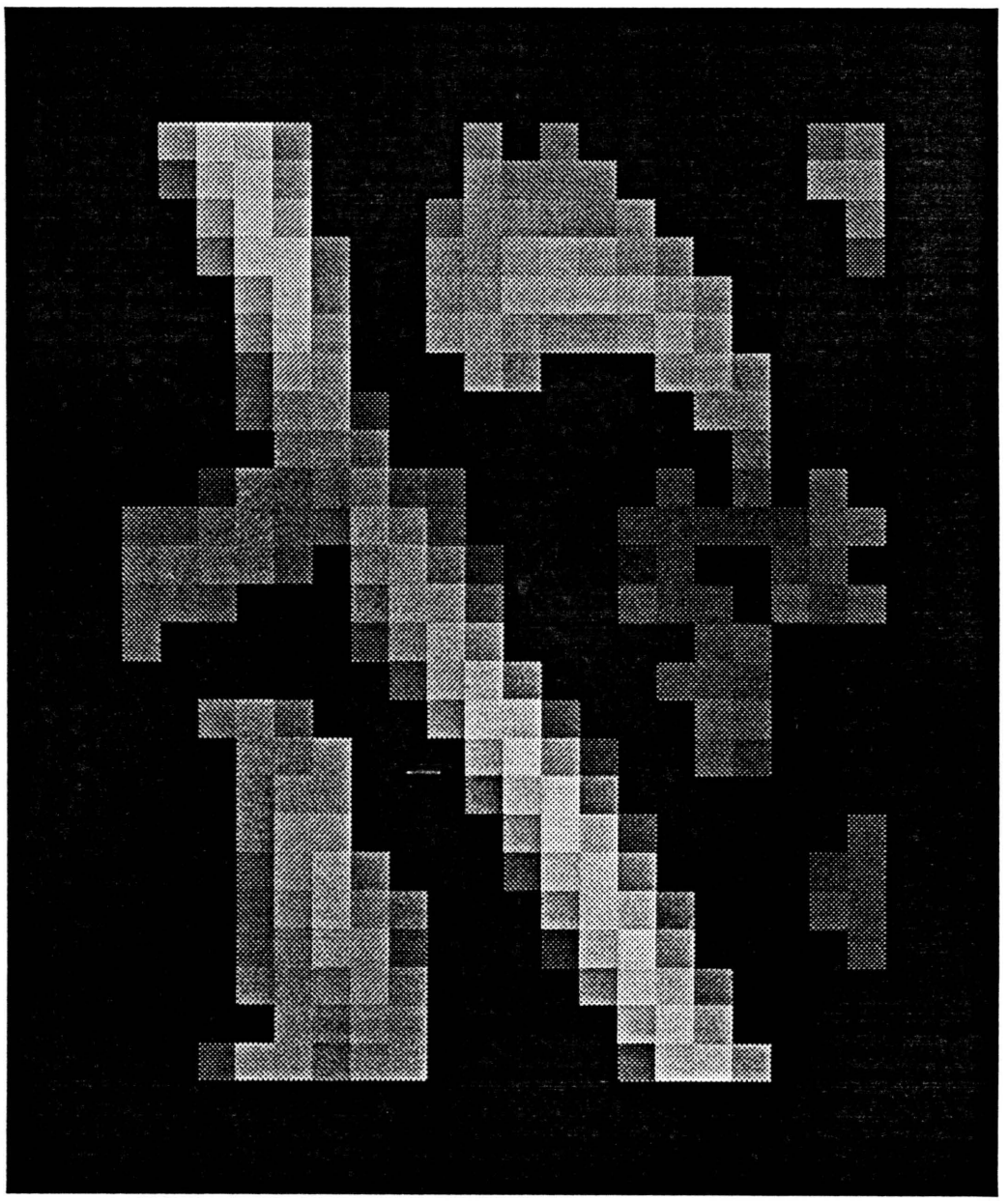


Figure 23. Heights of Gaussians from Linear Least Squares fitting for image B.

4.5 *Final Formulation of the Linel Model*

The final model then becomes

$$G_4 = k - he^{-(w(z)^2)} \qquad [8]$$

where

$$z = r\sin\theta + c\cos\theta. \qquad [9]$$

In Eqn. [9], θ takes discrete values. In Eqn. [8], w is a constant and k and h are parameters of the function. This present model loses features like sub-pixel localization and variable width of the linels. However, the model is linear and hence, computationally simple. This problem is solved by the DQRSL and DQRDC subroutines in LINPACK (Dongarra et al. [1979]).

Additional experimentation was done on Eqn. [8] by inspecting the results obtained by constraining k and h to take positive values only. The source code for this was from Lawson and Hanson[1974]. It was conjectured that the constrained least squares method will yield a better result than an unconstrained one due to the fact that this method is similar to doing a matched filtering by always forcing the function to model an ideal road element. However, it was found that although this formulation could bring out many of the roads that were occluded by shadows, etc., the operator was too sensitive to noise (Figure 24). Thus, Eqn. [8] forms the final linel detector model and the surface modelling was done using unconstrained linear least squares fitting.

The windows for the above computations were chosen such that they accentuate the directionality of the operators. Windows, elongated approximately in the direction of the invariance of the 1-D surface, were chosen. It was seen that good choice was to have the directions vary discretely by 15 degrees and the windows by 45 degrees (see Appendix A).

Figure 24. Output from Constrained Linear Least Squares Fitting for Image in Figure 7.

4.6 *Algorithm for Road Detection*

Thus, the procedure derived to detect the presence of a linel centered in an image-window can be outlined as follows. This procedure is to be repeated over the whole image by shifting the window steps of 1-pixel in the c and r directions.

1. Vary angle θ, the direction of invariance, in discrete steps.

2. For each θ, compute unconstrained linear least squares fit of the *one-dimensional* Gaussian surface and estimate the background intensity level, strength of linel and residual.

3. Among all the above fits in different directions, choose the one which has minimum residual of fit. Threshold on the residue and report the strength of the linel for the chosen fit.

This operator shall henceforth be called the *Residue Minimization Operator*. The output from this operator is a numeric image. The higher the value of a pixel, the greater the probability that the pixel is a linel.

4.7 *Matrix Formulation of the Problem*

The above procedure is mathematically equivalent to solving the following minimization problem for each pixel (Figure 25).

$$\min_{\theta} \left[\min_{\Gamma_\theta} \| A_\theta \Gamma_\theta - Y_\theta \|_2 \right]$$

$A_\theta \Gamma_\theta - Y_\theta$ can be expanded as

$$\begin{bmatrix} 1 & -e^{-(w_0 * z_1^2)} \\ 1 & -e^{-(w_0 * z_2^2)} \\ .. & ..... \\ .. & ..... \\ 1 & -e^{-(w_0 * z_n^2)} \end{bmatrix} \begin{bmatrix} k \\ h \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ .. \\ .. \\ y_n \end{bmatrix}$$

where n is the number of pixels in the image window,

Γ_θ is the vector of unknowns, i.e.,

k is the background value

h is the height of the Gaussian

$z_i, y_i, 1 <= i <= n$, are the z-value and gray value for pixel i respectively.

Figure 25. Matrix Formulation of Residue Minimization Operator.

$$\min_{\theta} \left[\min_{\Gamma_\theta} \| A_\theta \Gamma_\theta - Y_\theta \|_2 \right] \qquad [10]$$

where

- $\Gamma_\theta = [k, h]^T$

- A_θ is the n x 2 Gram matrix where n is the number of pixels in the window, and for i, 1 < = i < = n, each element of the matrix A_θ is given by

$$a_{i1} = 1 \qquad [11]$$

$$a_{i2} = -e^{-wz_i^2} \qquad [12]$$

where

$$z_i = c_i \cos\theta + r_i \sin\theta. \qquad [13]$$

(r_i , c_i) are the coordinates of pixel i, and θ is the angle perpendicular to the angle of invariance of the linel.

- $Y_\theta = [y_1, ..., y_n]^T$, where y_i is the gray-level value of pixel i.

Figure 28 shows an output for the image in Figure 27 using the Residue Minimization operator. The threshold residue is set at 100.

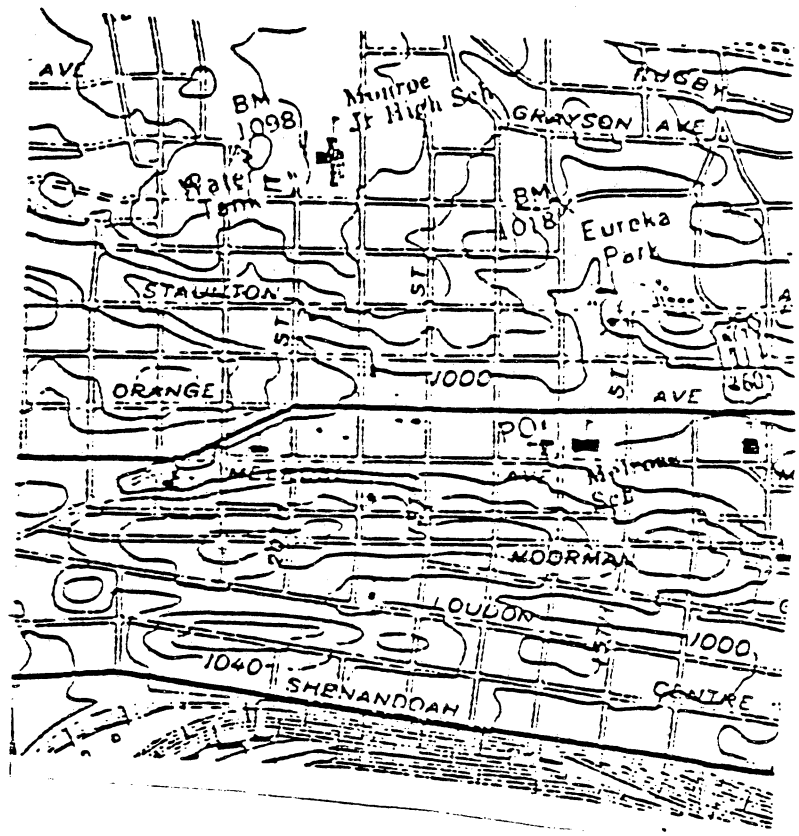


Figure 26. Ground Truth of Roads in Downtown ROANOKE.

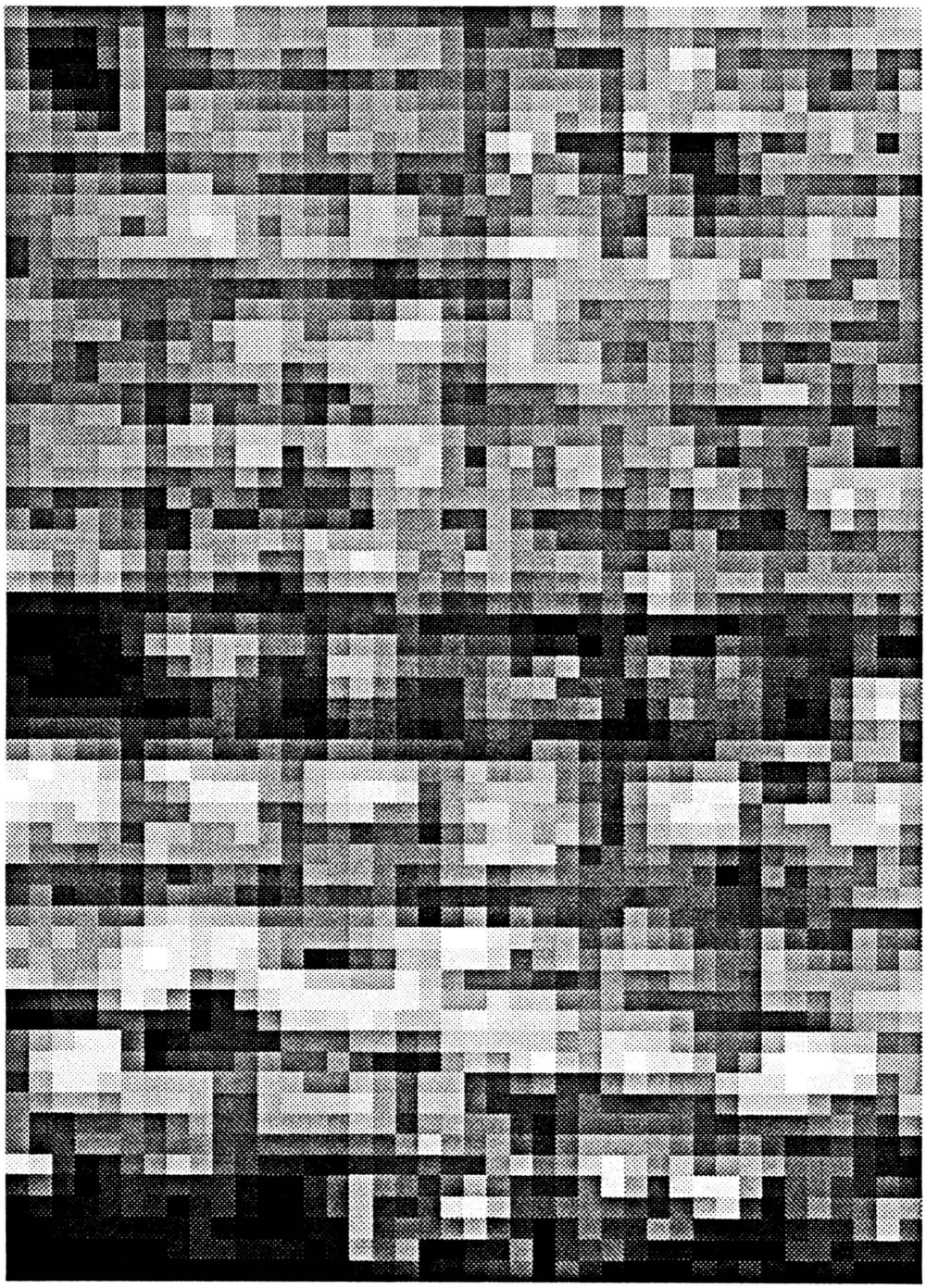


Figure 27. Part of SPOT Image Corresponding to the Ground Truth.

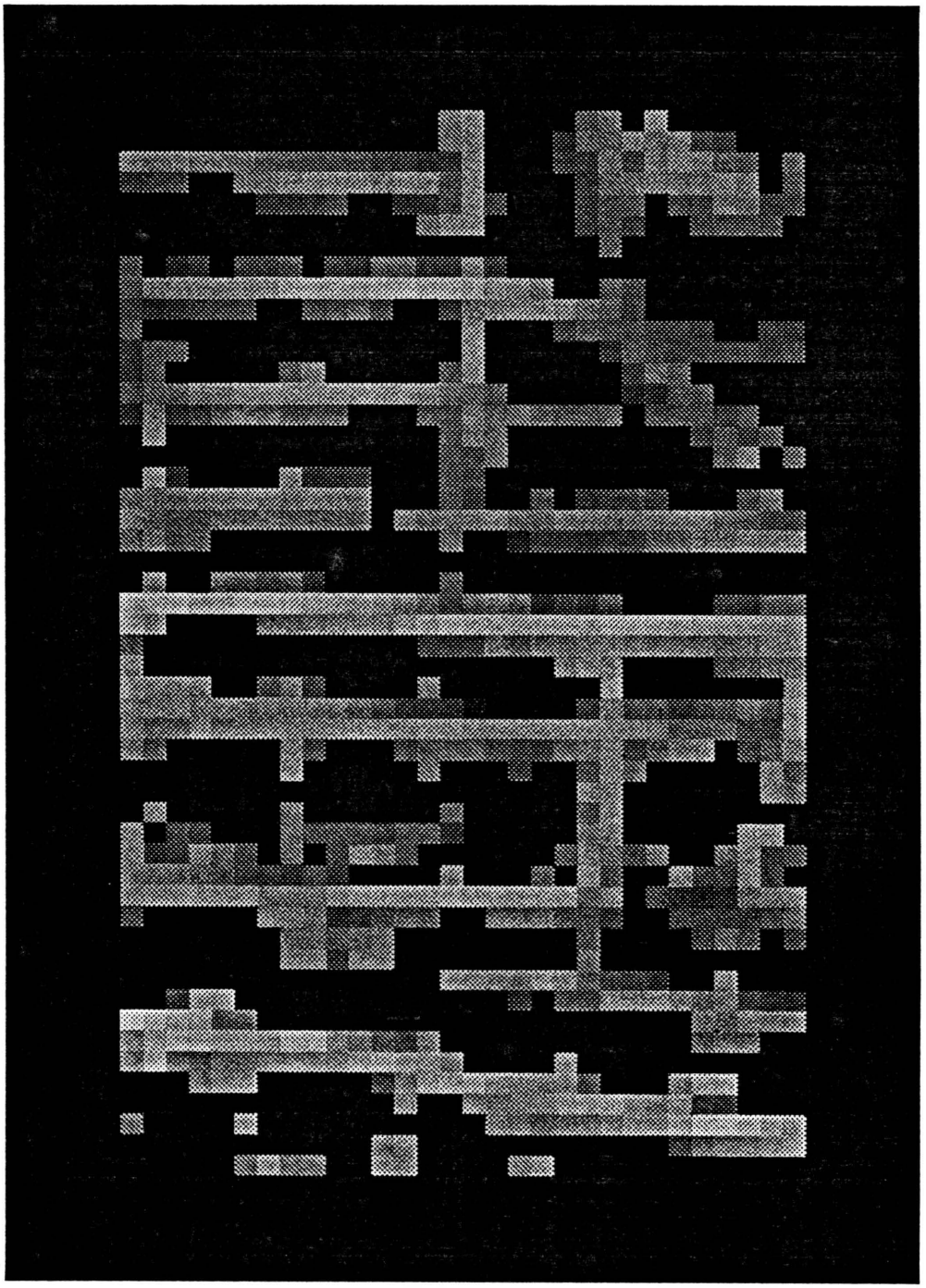


Figure 28. Output from Residue Minimization Operator for Figure 27.

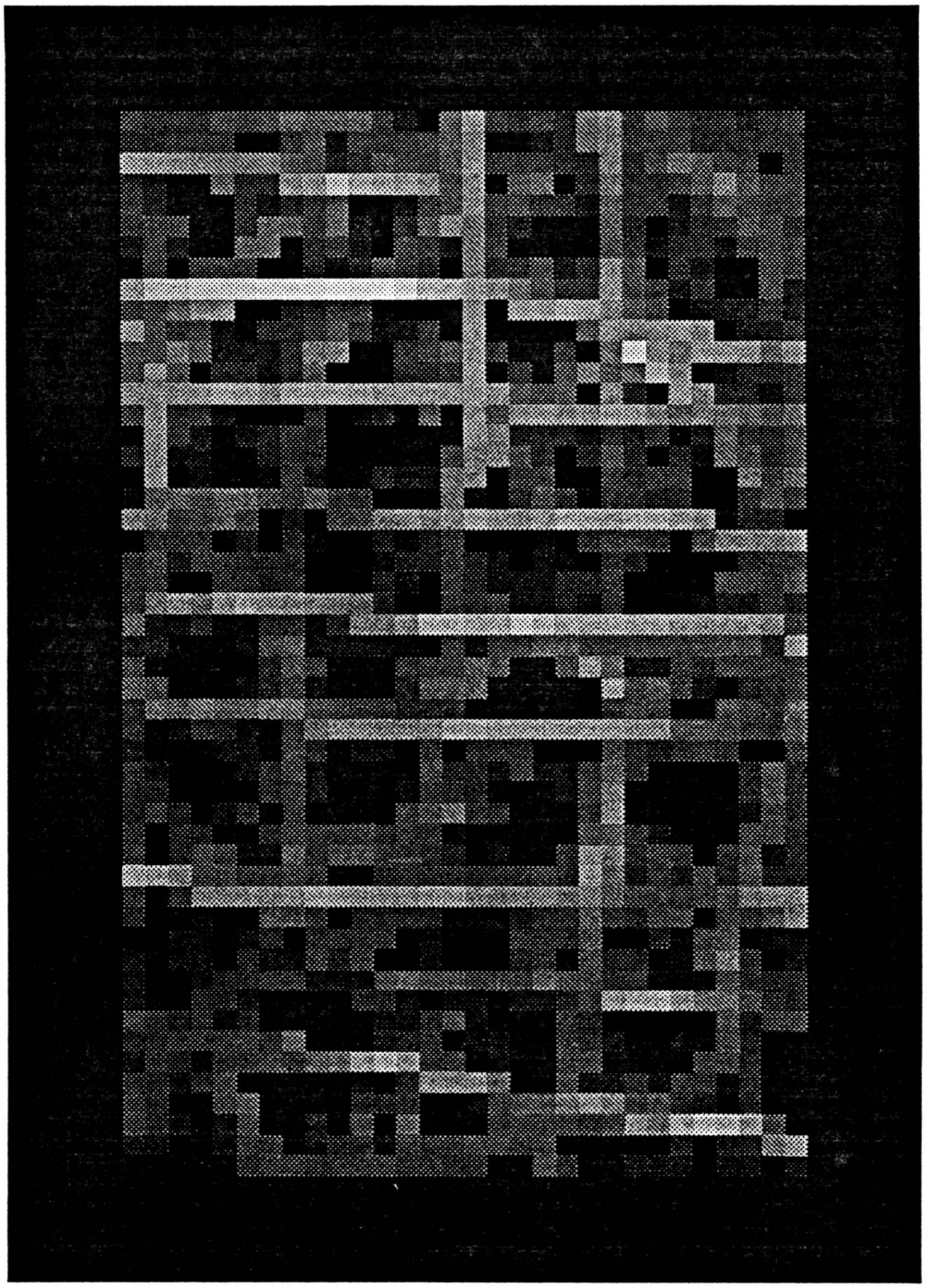


Figure 29. Output from Figure of Merit Maximization Operator for Figure 27.

4.8 *Figure of Merit Maximization Operator*

It was observed that because of the high noise content in the image, the error of fit was always high and exhibited a narrow range of variation. This introduced the need for a new measure of the probability that a given pixel lies on a road. A suggested figure of merit is

$$\frac{m h_\theta}{[r_\theta^2 + a]^l} \qquad [14]$$

where

- h_θ is the height of the Gaussian obtained from the least squares fit.
- r_θ is the residue of the least squares fit.
- a, l and m in Eqn. [14] are constants that assign relative importance to the height of the best fit Gaussian surface vs. the residue of the least squares fit.

Eqn. [14] can be thus be described as the edgeness of the pixel in that direction. Hence the problem now changes to

$$\max_\theta \left[\frac{m h_\theta}{[r_\theta^2 + a]^l} \right] \qquad [15]$$

where all the quantities in Eq. [15] are same as described as above. This operator shall be referred to as the *Figure of Merit (FoM) Maximization* operator.

Figure 29 shows the output of the Figure of Merit Maximization operator where a,m and l were set at 0, 10000 and 1 respectively. A nonmaximum suppression, in a direction θ, over a window size

of 3 was done in each of the discrete θ directions, to thin the figures of merit before the maximum was taken.

In order to compare the results obtained from unconstrained and constrained linear least squares fitting, Figures 30 and 31 were generated using constrained linear least squares fitting corresponding to Figures 28 and 29 respectively. The output from the Residue Minimization operator (Figure 30) enhances most of the roads in Figure 27. However, it can be seen to be very sensitive to noise while being very effective in extracting occluded roads. On the other hand, the output from the Figure of Merit Maximization operator of Figure 31 is very similar to that of Figure 28. This means that for the relative weights assigned to residue and heights of Gaussians, the operators are equivalent.

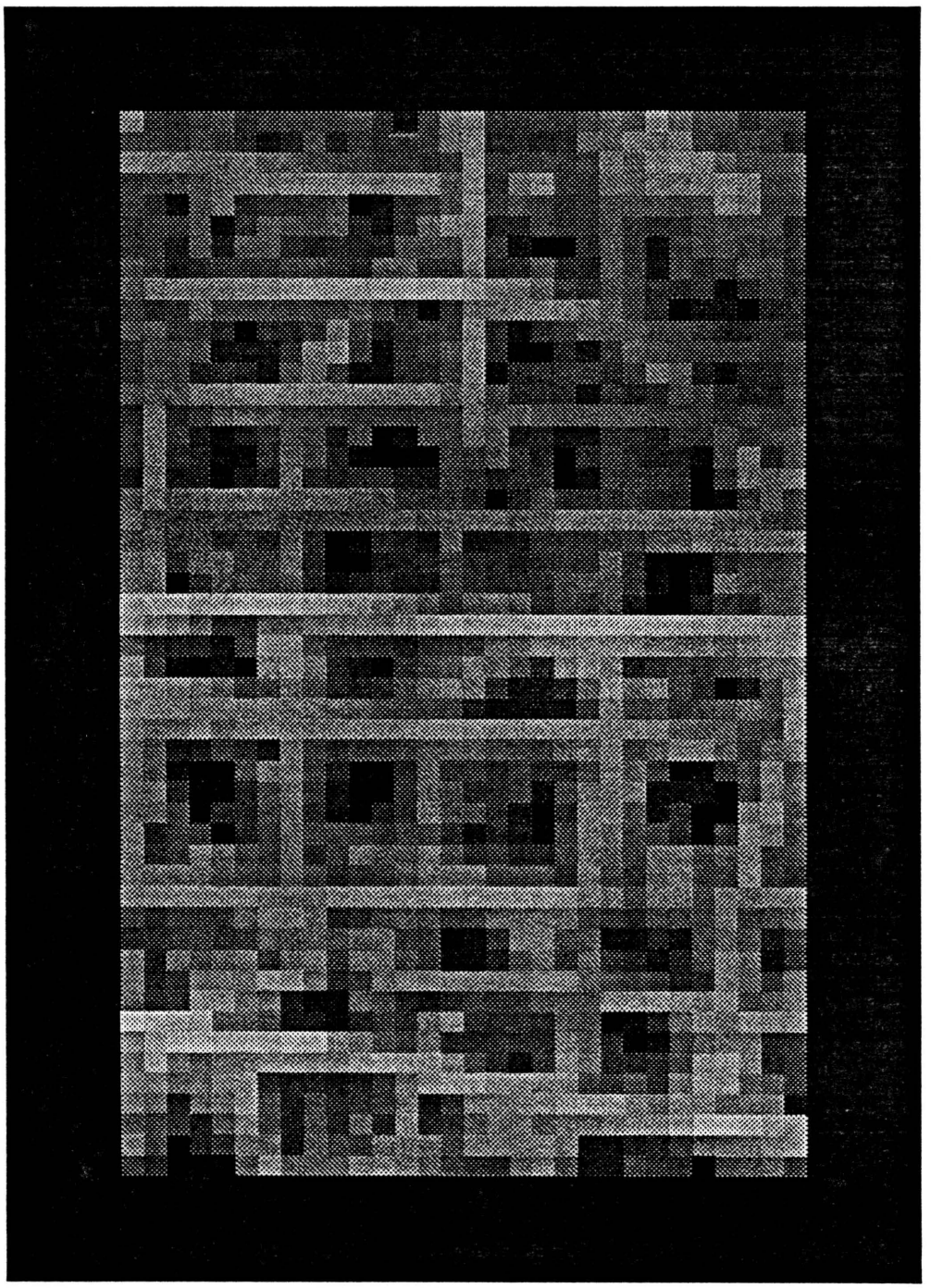


Figure 30. Output from Residue Minimization Operator for Figure 27 using Constrained Linear Least Squares Fitting.

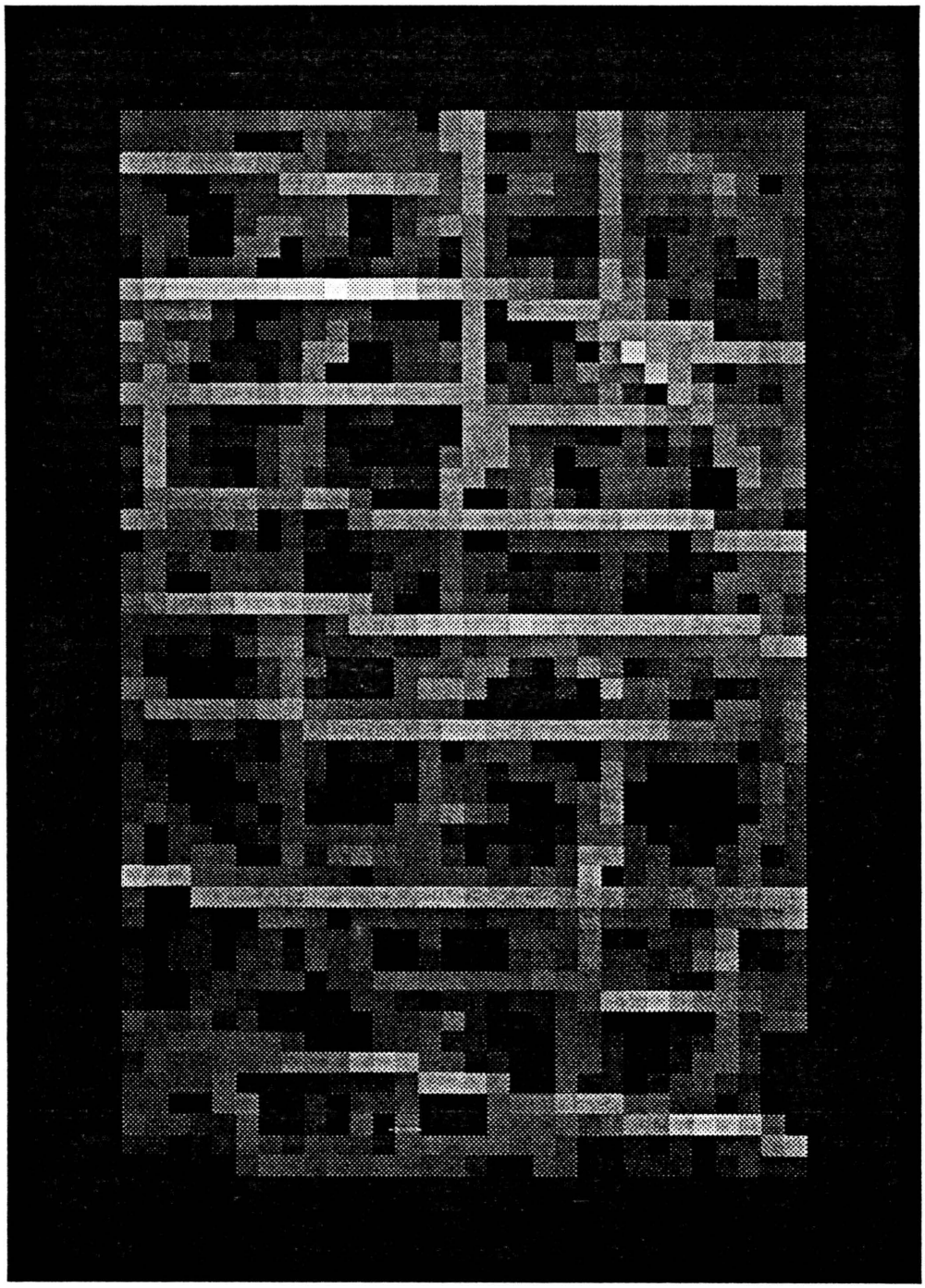


Figure 31. Output from Figure of Merit Maximization Operator for Figure 27 using Constrained Linear Least Squares Fitting.

Figure 32. Output for Image in Figure 7 using Figure of Merit Maximization Operator.

5.0 POSTPROCESSING OF THE OUTPUT FROM THE FoM MAXIMIZATION OPERATOR

A few simple techniques were used to examine their effects on the output of the Figure of Merit Maximization operator. The results from these techniques are presented below. Although the Figure of Merit Maximization operator is very effective in enhancing the roads and removing noise, from the results below, it is clear that a more sophisticated technique should be used to thin the lines.

5.1 Double Adaptive Thresholding Operator

One operator that was tested was the GIPSY LINE operator which is a line tracker that makes use of double adaptive thresholding for low contrast ridges. LINE was used sucessfully for engineering

drawings, but evidently there was much too much noise in our application for the tracker to work well.

5.2 *Semilinear Operator*

Since the lines in the figure of merit image are rather prominent, it was conjectured that Rosenfeld's semilinear operator might be useful in detecting them. This operator has a large output whenever a line is surrounded on both sides by regions of lower average intensity. The problem is that when the image noise is large and the operator is applied in all four directions, the operator becomes confused. It is likely that a more sophisticated version of this operator may still be useful.

5.3 *Iterating the FoM Maximization Operator*

Iterating the FoM Maximization operator results in very effective clearing up of noise from the image (Figure 35). This technique should be explored for thinning the roads and improving the performance of the operator.

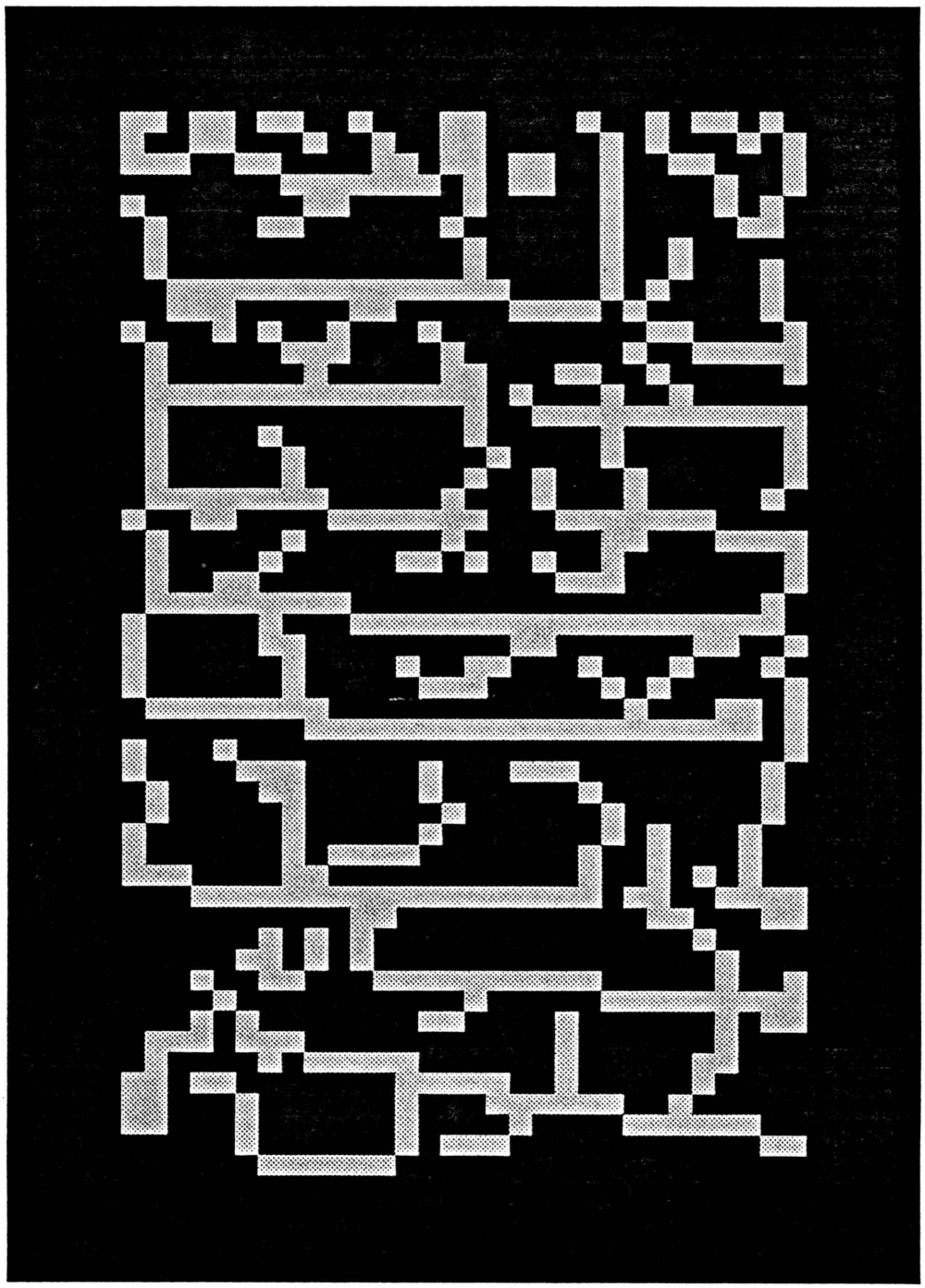


Figure 33. Output from the LINE Operator Run on Figure 29.

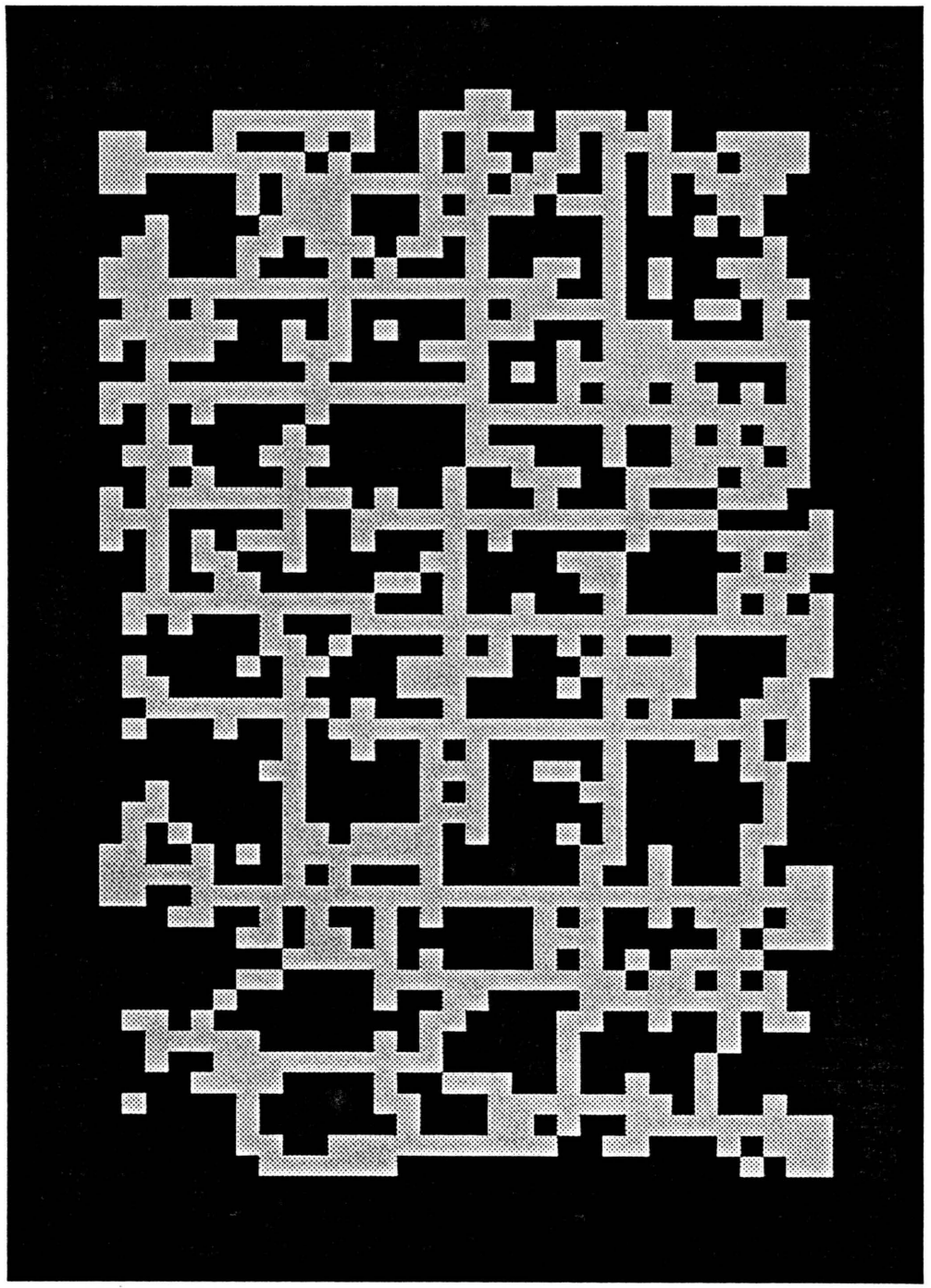


Figure 34. Output from the Semilinear Operator Run on Figure 31.

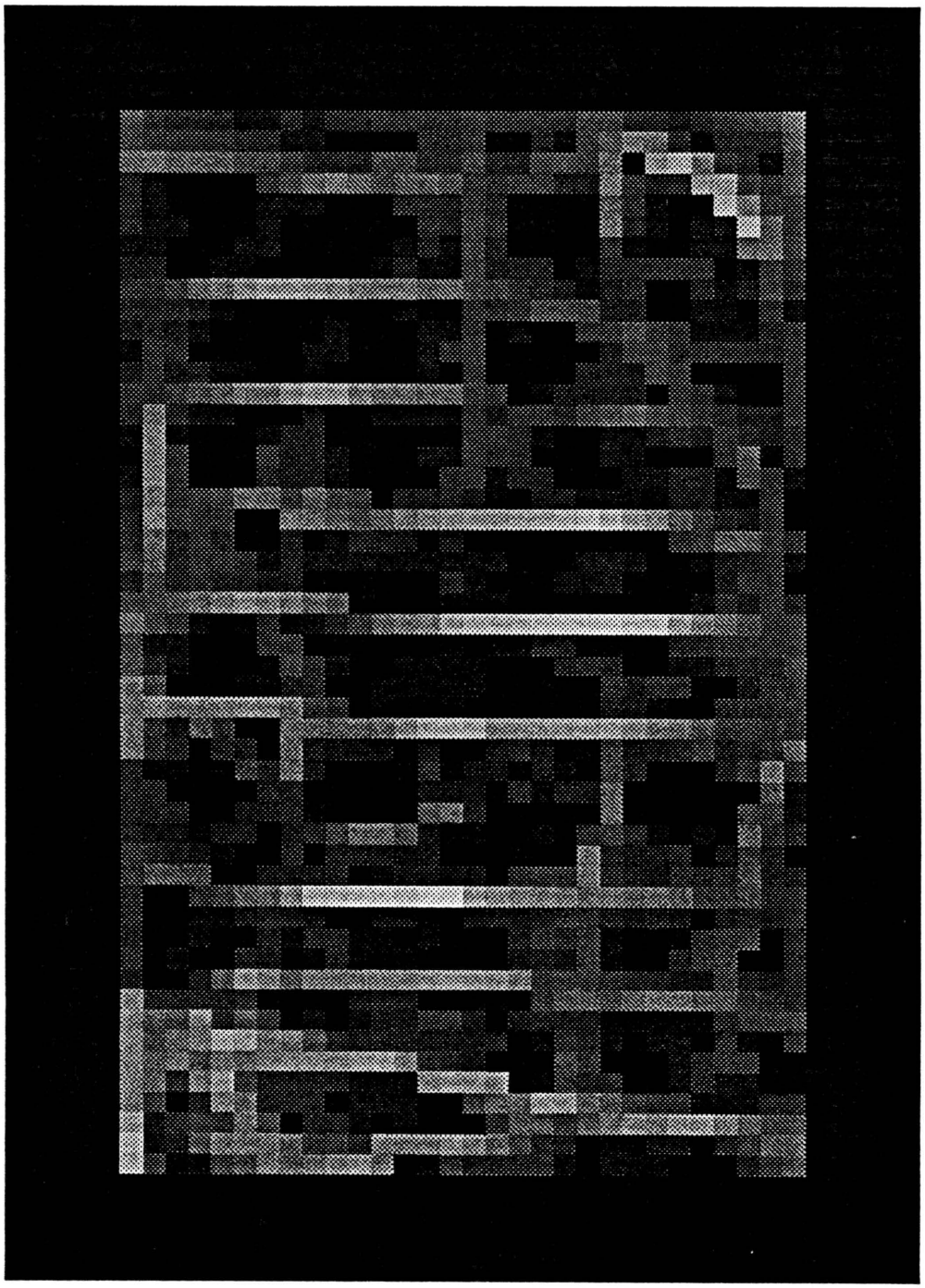


Figure 35. Output after the Second Iteration of the FoM Maximization Operator on Figure 27.

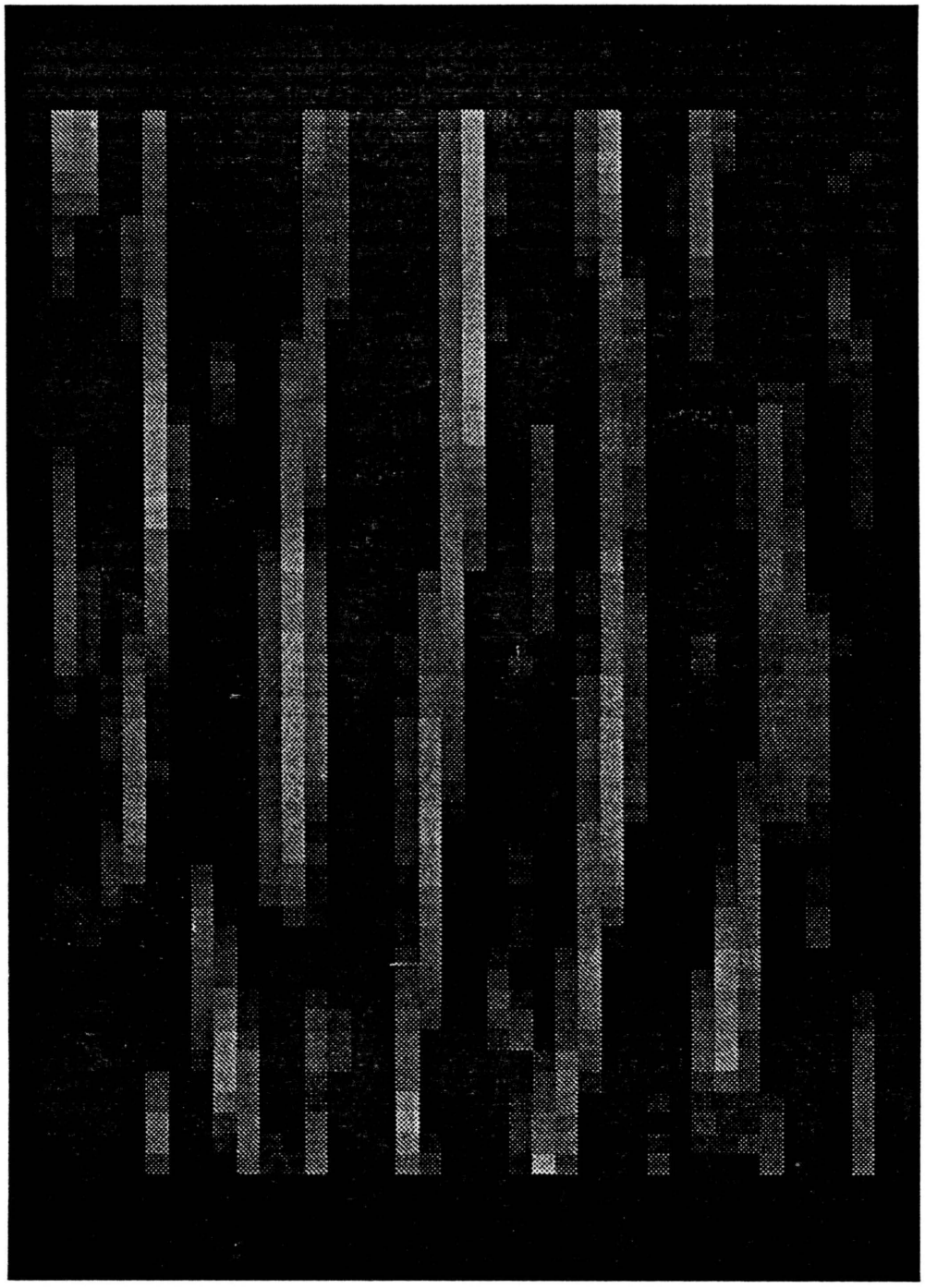


Figure 36. Road Detection in Vertical Direction Only for Figure 27.

6.0 CONCLUSION

In this thesis, the problem of and an implementation for the precise delineation of roads and other linear features appearing in 20-m SPOT satellite images was discussed. In order to detect the roads, a new line operator based on a surface modelling approach was developed. The new operator identifies the pixels that have high probability of lying on roads. These pixels have directions associated with them and each of them can be described as a *linel*. This operator was shown to be superior to the other commonly used line and edge operators.

Although this operator shows promise of being an excellent road finder for images from the SPOT satellite, better results may be obtained by further adjusting the constants a, m and l of the figure of merit (Eqn. [14]). Further new ways of combining the various steps of the algorithm should be explored in future to improve the performance of the operator. Intelligent programs incorporating prior knowledge may be built based on Gaussian surface fit technique to extract the roads under severe noise conditions. Figure 36 shows the output from a vertically constrained Figure of Merit Maximization operator when roads are known to have vertical orientations. Nonmaximum suppression is done in the horizontal direction to obtain Figure 36. The test image is that of Figure 27. The operator only looks in the vertical direction and is remarkably powerful in enhancing the vertical roads. The problem is in combining the 12 different images that result from the 5 X 11 operator.

This thesis does not deal with the problem of linking the linels to form chain-codes of the road network. Dynamic programming techniques may be used to link the linels. The heuristics for the dynamic programming may be based on the prior knowledge of the road structure of the area can be embedded in the figure of merit. Alternatively relaxation techniques may be used to link the linels. Output from the overly sensitive operator that uses constrained linear least square fitting as opposed to the unconstrained linear least square fitting, can be used to provide confidence levels to the output of the present operator and guide the linking process.

It is expected that the reliable output from the newly proposed operator will make the problems of linking the linels much simpler.

References

Bajcsy R., and M. Tavakoli, "A Computer Recognition of Bridges, Islands, Rivers, and Lakes from Satellite Pictures," Proc. of Machine Processing of Remotely Sensed Data, pp. 2A-54-2A-68, October 1973.

Bajcsy R., and M. Tavakoli, "Computer Recognition of Roads in Satellite Pictures," Proc. of Second Inter. Joint Confer. on Pattern Recognition, pp. 190-194, August 1974.

Barrow H. G., and Fishler, M. A., "An Expert System for Detecting and Interpreting Road Events Depicted in Aerial Imagery," Proc. of a Workshop on Image Understanding, Cambridge, pp. 155-156, May 1978.

Dongarra, J. J., J. R. Bunch, C. B. Moler, G. W. Stewart, "LINPACK Users' Guide", SIAM, Philadelphia, 1979.

Ehrich, R. W., "Detection of Global Edges in Textured Images," IEEE Transactions on Computers, vol. C-26, no. 6, pp. 589-603, June 1977.

Fischler M. A., J. M. Tenenbaum, and H. C. Wolf, "Detection of Roads and Linear Structures in Low Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique," CGIP 15, pp. 201-223 (1981).

Fua, P., and A. J. Hanson, "Using Generic Geometric Models for Intelligent Shape Extraction, AAAI, pp-706-7111, (1987).

Goetz, A. F. A., et al., "Application of ERTS Images and Image Processing to Regional Geologic Problems and Geologic Mapping in Northern Arizona", National Aeronautics and Space Administration, JPL, TR 32-1597, May 1975.

Gold, T. S., and H. P. Smith, "A Dynamic Programming Method for the Detection of Nearly Straight Lines in Noisy Pictures," IEEE Computing Repository, R-74-167, August 1974.

Haralick, R. M., "Digital Step Edges from Zero Crossing of Second Directional Derivatives," IEEE Transactions on Pattern Analysis and Machine and Machine Intelligence, vol. PAMI-6, pp. 58-68, January 1984.

Haralick, R. M., and L. T. Watson, "A facet model for image data," Computer Graphics and Image Processing, vol. 15, pp. 113-129, 1981.

Hueckel, M. H., "An Operator which Locates Edges in Digitized Pictures," J. ACM, vol. 18, no. 1, pp. 113-125, January 1971.

Huertas, A., W. Cole and R. Nevatia, "Detecting Runways in Aerial Images," American Association for Artificial Intelligence, pp. 712-717 (1987).

Hufnagel R. E., and N. R. Stanley, "Modulation Transfer Function Associated with Image Transmission through Turbulent Media," J. Opt. Soc. America 54, 52 (1964).

Kidd, R. H., and R. H. Wolfe, "Performance Modelling of Earth Resourses Remote Sensors," IBM Journal of Research and Development, vol. 20, no. 1, January 1976.

Lawson, C. L., Hanson R. J., "Solving Least Squares Problems," Prentice Hall, Englewood Cliffs, NJ, 1974.

Li, R. Y., and K. S. Fu, "Trees System Approach for LANDSAT Data Interpretation," Proc. IEEE Symposium on Machine Processing of Remotely Sensed Data, pp. 2A-10-2A-16, 1974.

Martelli, A., "Edge Detection using Heuristic Search Methods," Computer Graphics and Image Processing, vol. 1, pp. 169-182, August 1972.

McKeown, Jr., D. M., and J. F. Pane, "Alignment and Connection of Fragmented Linear Features in Aerial Imagery," IEEE Conference on Computer Vision Pattern Recognition, June 1985.

Montanari, U., "On the Optimal Detection of Curves in Noisy Pictures," JACM, vol. 18. pp. 335-345, May 1971.

More', J. J., B. S. Garbow, K. E. Hillstrom, "User Guide for MINPACK-1," ANL-80-74, Argome National Laboratory, Argonne, IL, 1980.

Nagao, M., and T. Matsuyama, "A Structural Analysis of Complex Aerial Photographs", New York, Plenum Press, (1980).

Nalwa, V. S., and T. O. Binford, "On Detecting Edges," IEEE Transactions of Pattern Analysis and Machine Intelligence, vol. PAMI-8, no. 6, pp. 699-714, November 1986.

Nevatia, R., and K. R. Babu, "Linear Feature Extraction and Description," Computer Graphics Image Processing, vol. 13, pp. 257-269, 1980.

Prewitt J. M. S., "Object enhancement and extraction," in Picture Processing and Psychopictories, B. S. Lipkin and A. Rosenfeld, Eds. New York: Academic, pp. 75-149, 1970.

VanderBrug G. J., "Line Detection in Satellite Imagery," IEEE Trans. Geosci. Electronics GE-14, No. 1, pp. 37-44, (1976).

Watson, L. T., K. Arvind, R. W. Ehrich, R. M. Haralick, "Extraction of Lines and Regions from Grey Tone Line Drawing Images," Pattern Recognition, vol. 17, no. 5, pp. 493-507, 1984.

Appendix A. Window indices

The following shows the indices of the rows and columns of each pixel in a centred image window of size 5 X 11. The window is directional. Hence, for different orientations of θ the pixels that are to be included in the image window change. These windows are inclined at 45 degrees to each other.

$\theta = 0$

(column row)

-2 -4

-2 -3

-2 -2

-2 -1

-2 0

-2 5

-2 4

-2 3

-2 2

-2 1

-1 -5

-1 -4

-1 -3

-1 -2

-1 -1

-1 0

-1 5

-1 4

-1 3

-1 2

-1 1

0 -5

0 -4

0 -3

0 -2

0 -1

0 0

0 5

0 4

0 3

0 2

0 1

1 -5

1 -4

1 -3

1 -2

1 -1

1 0

1 5

1 4

1 3

1 2

1 1

2 -5

2 -4

2 -3

2 -2

2 -1

2 0

2 5

2 4

2 3

2 2

2 1

θ = 45

(column row)

-1 4

0 4

-2 3

-1 3

0 3

1 3

-3 2

-2 2

-1 2

0 2

1 2

2 2

-4 1

-3 1

-2 1

-1 1

0 1

1 1

2 1

3 1

-5 0

-4 0

-3 0

-2 0

-1 0

0 0

1 0

2 0

3 0

4 0

-3 -1

-2 -1

-1 -1

0 -1

1 -1

2 -1

3 -1

4 -1

5 -1

-2 -2

-1 -2

0 -2

1 -2

2 -2

3 -2

4 -2

-1 -3

0 -3

1 -3

2 -3

3 -3

0 -4

1 -4

2 -4

1 -5

$\theta = 90$

(column row)

-5 -2

-5 -1

-5 0

-5 1

-5 2

-4 -2

-4 -1

-4 0

-4 1

-4 2

-3 -2

-3 -1

-3 0

-3 1

-3 2

-2 -2

-2 -1

-2 0

-2 1

-2 2

-1 -2

-1 -1

-1 0

-1 1

-1 2

0 -2

0 -1

0 0

0 1

0 2

5 -2

5 -1

5 0

5 1

5 2

4 -2

4 -1

4 0

4 1

4 2

3 -2

3 -1

3 0

3 1

3 2

2 -2

2 -1

2 0

2 1

2 2

1 -2

1 -1

1 0

1 1

1 2

1 5

θ = 135

(column row)

0 4

1 4

2 4

-1 3

0 3

1 3

2 3

3 3

-2 2

-1 2

0 2

1 2

2 2

3 2

4 2

-3 1

-2 1

-1 1

0 1

1 1

2 1

3 1

4 1

5 1

-4 0

-3 0

-2 0

-1 0

0 0

1 0

2 0

3 0

4 0

-4 -1

-3 -1

-2 -1

-1 -1

0 -1

1 -1

2 -1

3 -1

-3 -2

-2 -2

-1 -2

0 -2

1 -2

2 -2

-2 -3

-1 -3

0 -3

1 -3

-1 -4

0 -4

0 -5

-2 -5