Visual Sensing and Interpretation for Navigation

Takeo Kanade

Visual Sensors

Camera (B/W, Color, Stereo, Motion)

Sonar

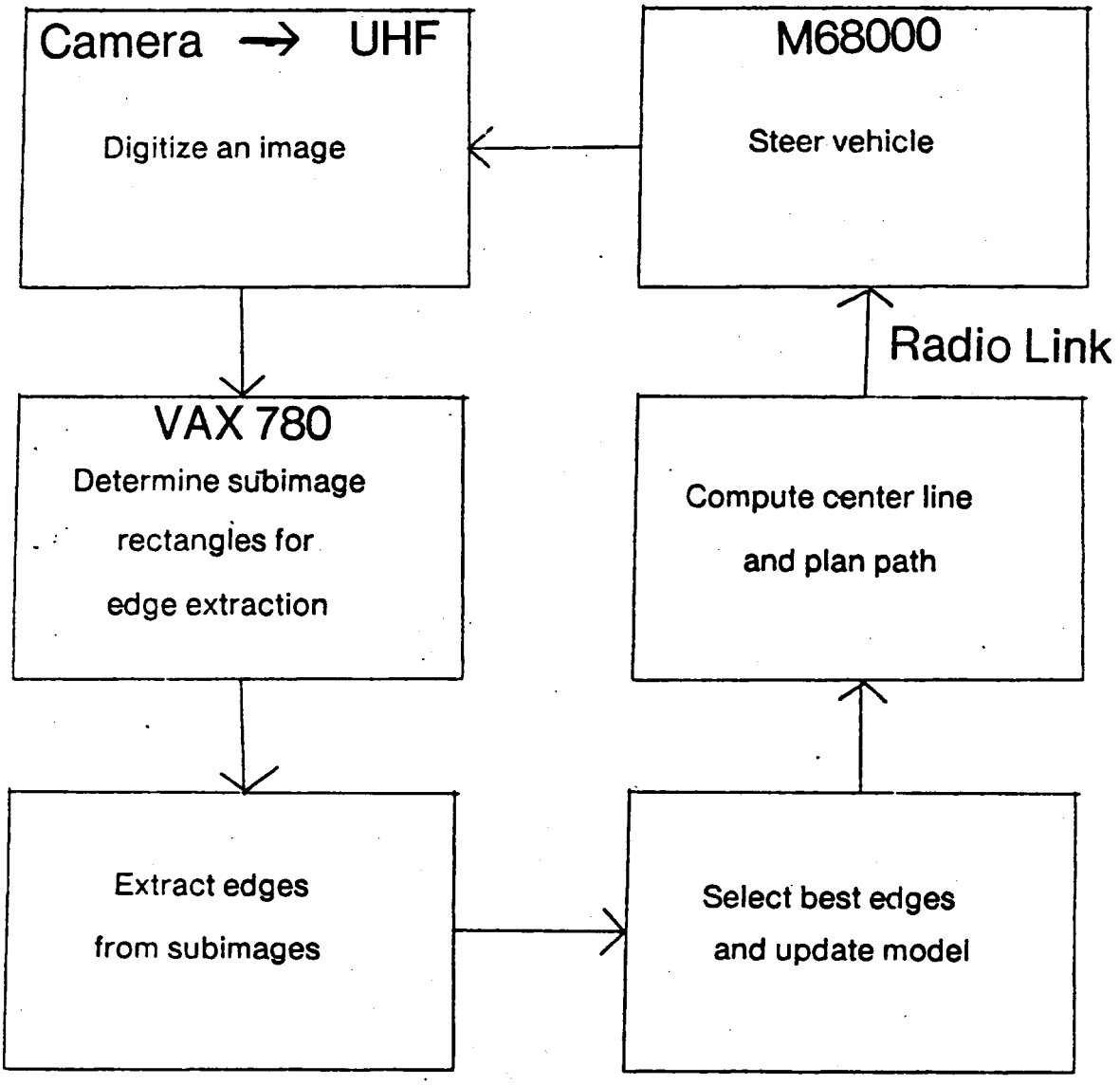
Laser Range Finder

Camera Images

- Can Recognize Color, Texture, Material, etc
- High Resolution
- High Speed Update

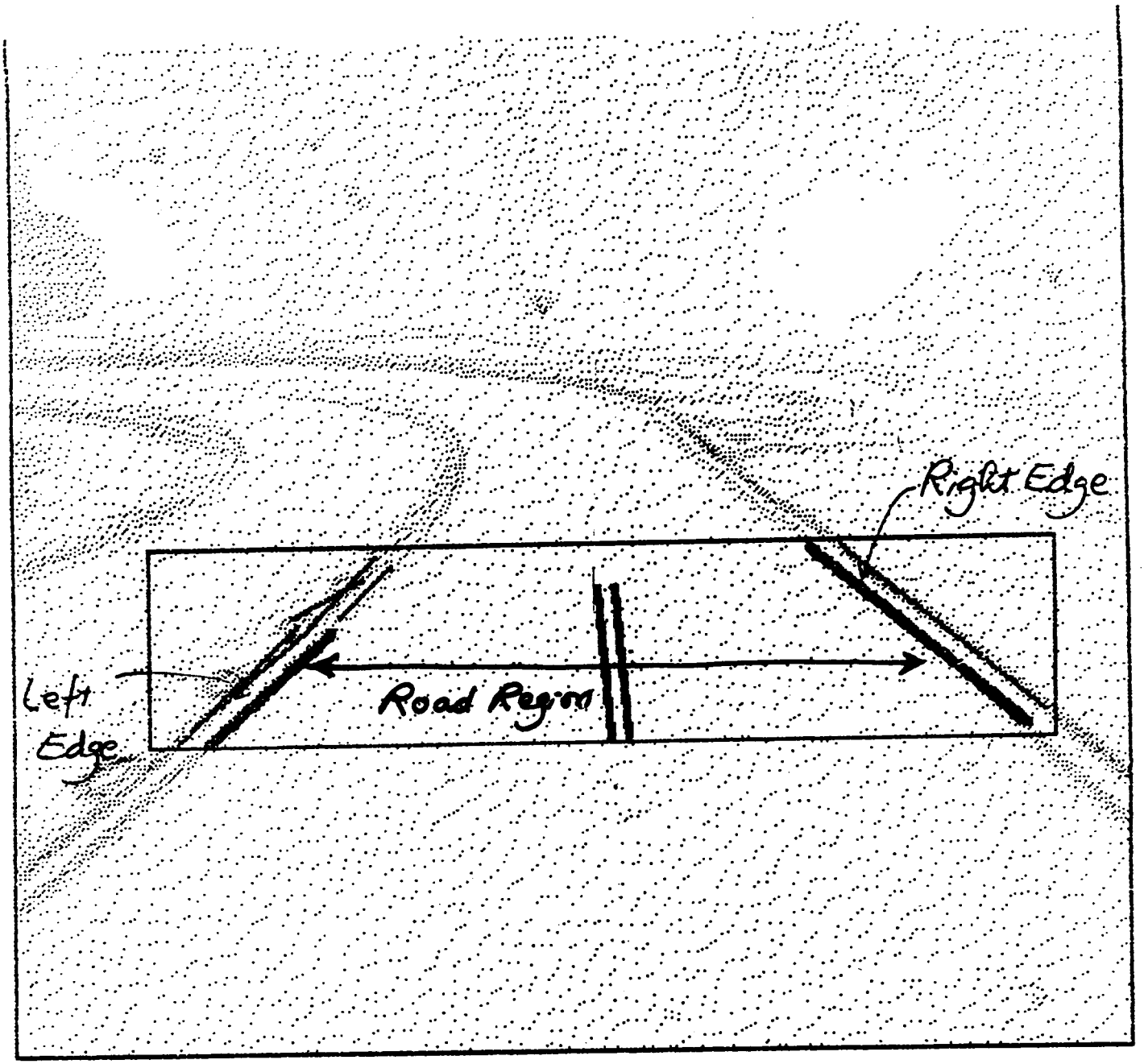
- 2D Projection
- Affected by Illumination, Shadow, View Angles, etc
- Computationally Expensive

Simple Road Following by Vision



Multiple Vision Techniques for Road Following

- Edge Detection
- Edge (Profile) Following
- Region Segmentation by Color
- Stereo for Obstacle Detection


Right Edge
Left Edge
Road Region

Stereo and Motion

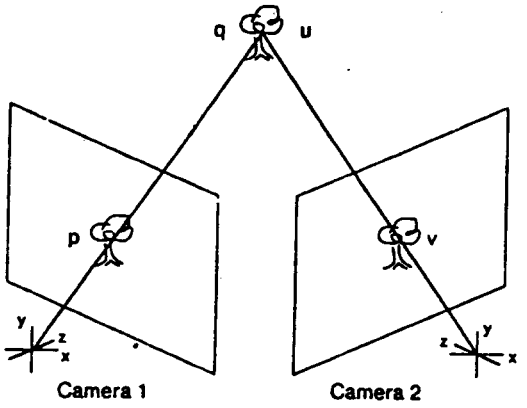


Figure 1. Camera model for optical navigation. Camera 1 defines the reference picture and coordinate system, with the origin at the "pinhole." For any point $\mathbf{p} = [\,p_x \quad p_y\,]$ in the reference image there is a point $\mathbf{q} = [\,q_x \quad q_y \quad q_z\,]$ in three-space that produced the image, at depth $z(\mathbf{p}) = q_z$. This point is expressed as $\mathbf{u} = [\,u_x \quad u_y \quad u_z\,]$ in the Camera 2, or test, coordinate system. The relationship between $\mathbf{q}$ and $\mathbf{u}$ is a function parameterized by the six camera parameters $\mathbf{c}$: three for the relative positions of the cameras and three for their relative orientation. Finally, the three-space point appears at the point $\mathbf{v} = [\,v_x \quad v_y\,]$ in the Camera 2 image plane. The points $\mathbf{p}$ and $\mathbf{v}$ are said to correspond.

Image
Image
→ Correspondences → Motion

Correspondence Problem

- Difficult
- Computationally Expensive

FIDO Vision

Computational demands of dense stereo:

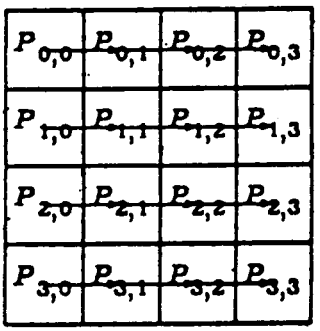
$2^3 * 2^3$ points per patch, times
$2^9 * 2^9$ candidate positions, times
$2^9 * 2^9$ points to be matched, equals
2^{42} pixel comparisons, or $> 2^{20}$ years

Need to reduce points to match and candidate positions.

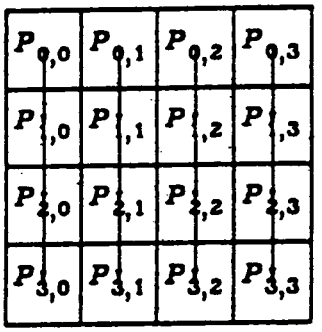
Answer:

- Select only 50 most interesting points to match
- Use hierarchical correlation
- Use stereo and motion constraints

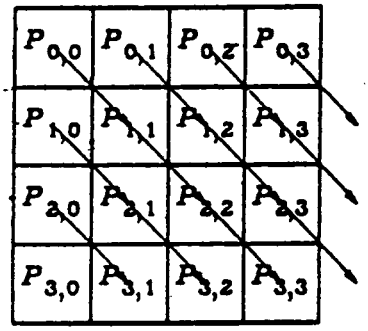
Interest Operator



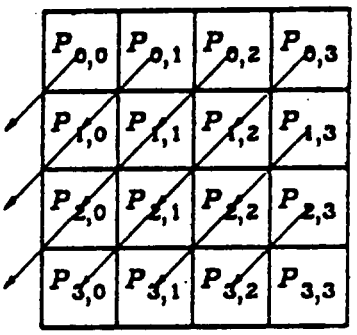
$\Sigma(P_{I,J}-P_{I,J+1})^2$



$\Sigma(P_{I,J}-P_{I+1,J})^2$



$\Sigma(P_{I,J}-P_{I+1,J+1})^2$

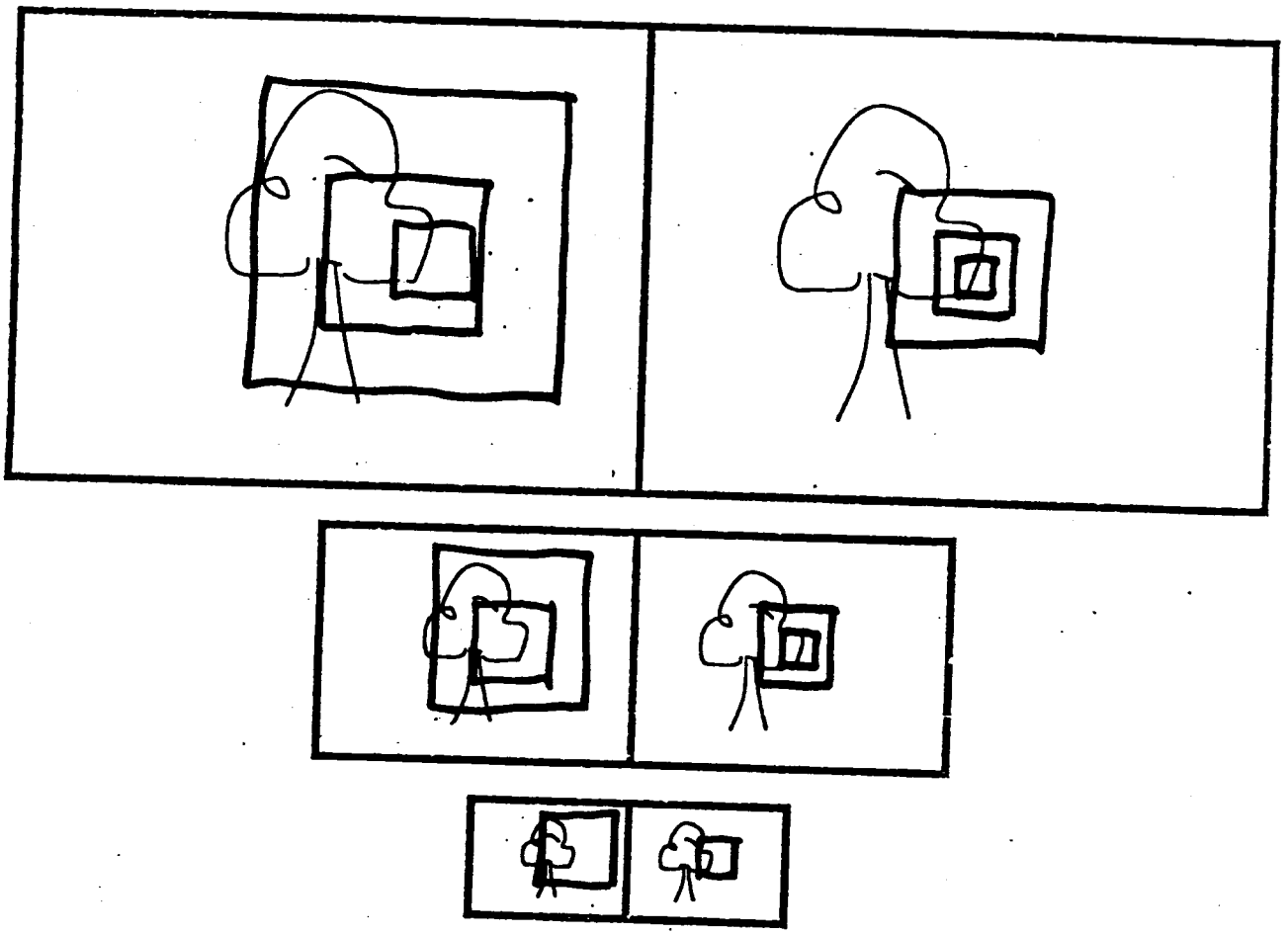


$\Sigma(P_{I,J}-P_{I+1,J-1})^2$

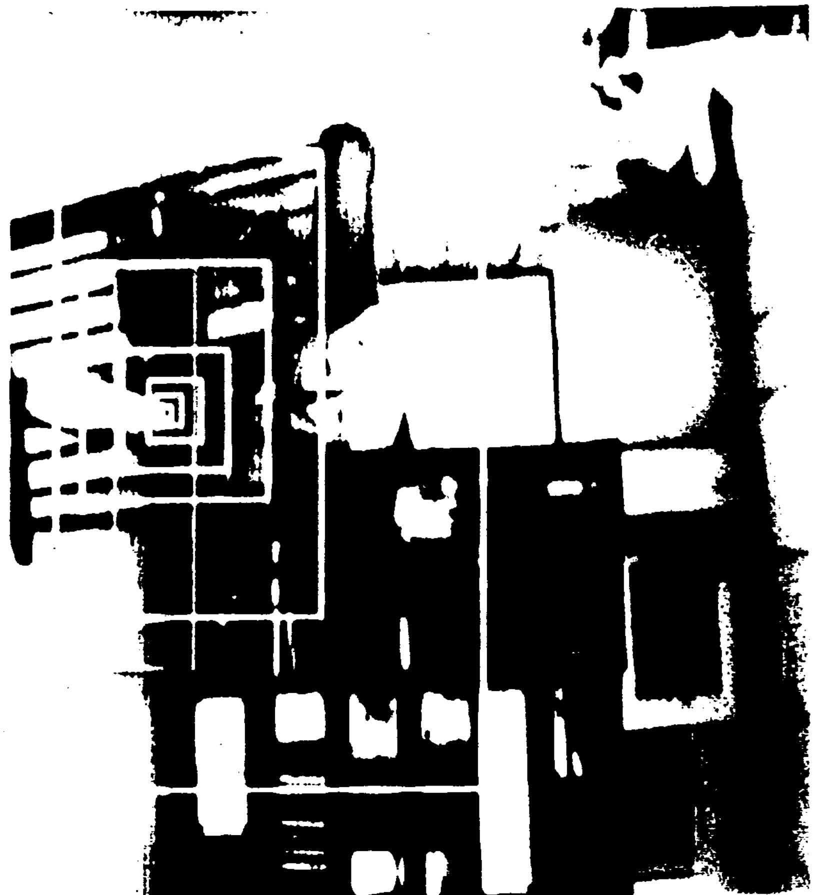
A typical interest operator window, and the four sums calculated over it ($P_{I,J}$ are the pixel brightnesses). The interest measure of the window is the minimum of the four sums.

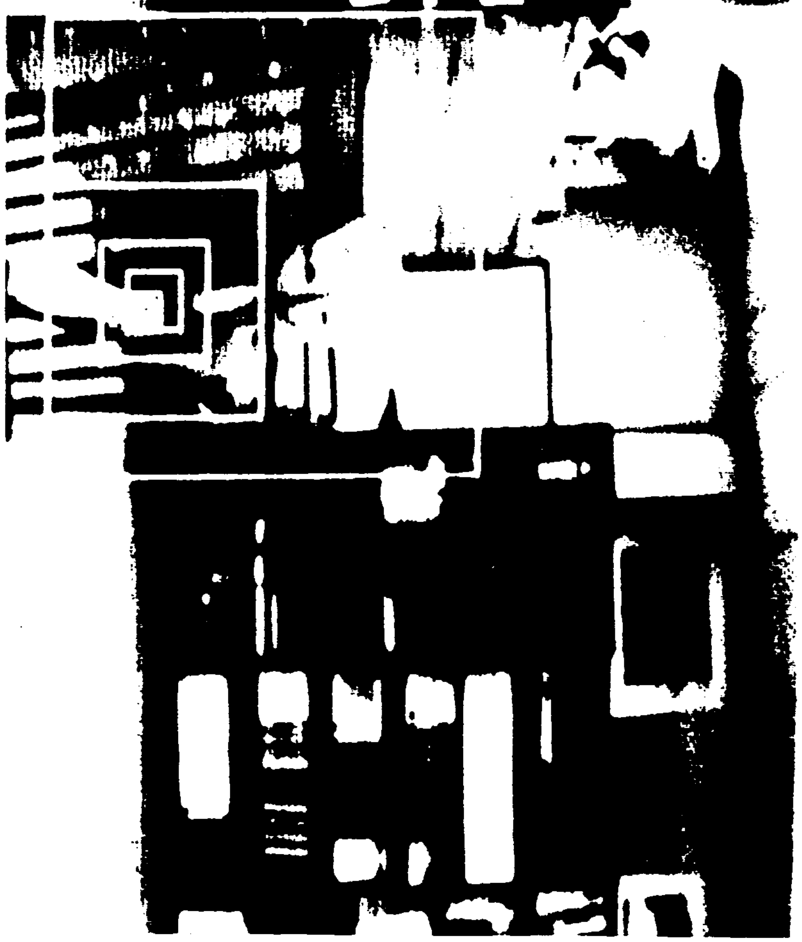
Interesting Point
= Local Maximum of Interest Measures

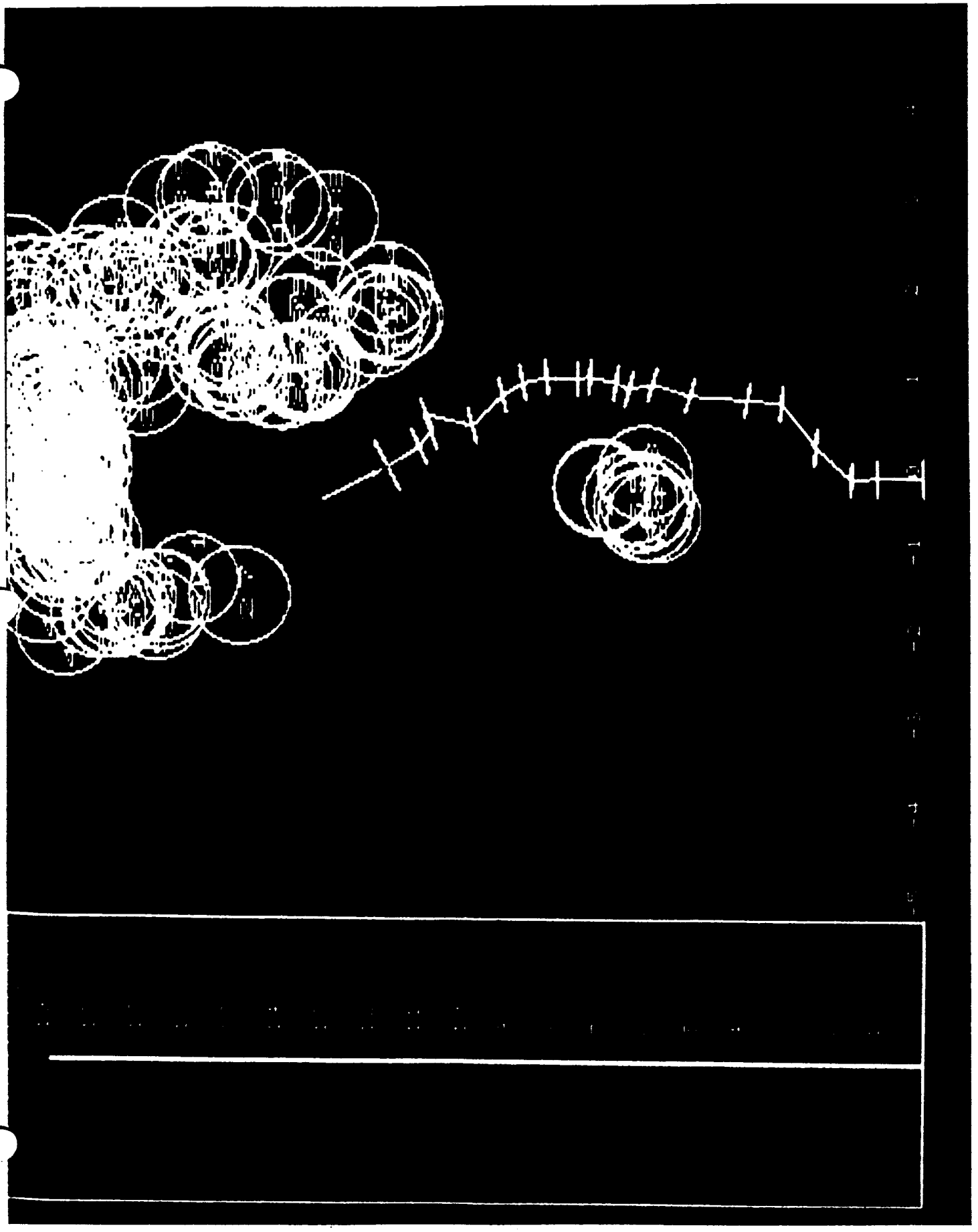
Iterative Search Using Image Pyramid



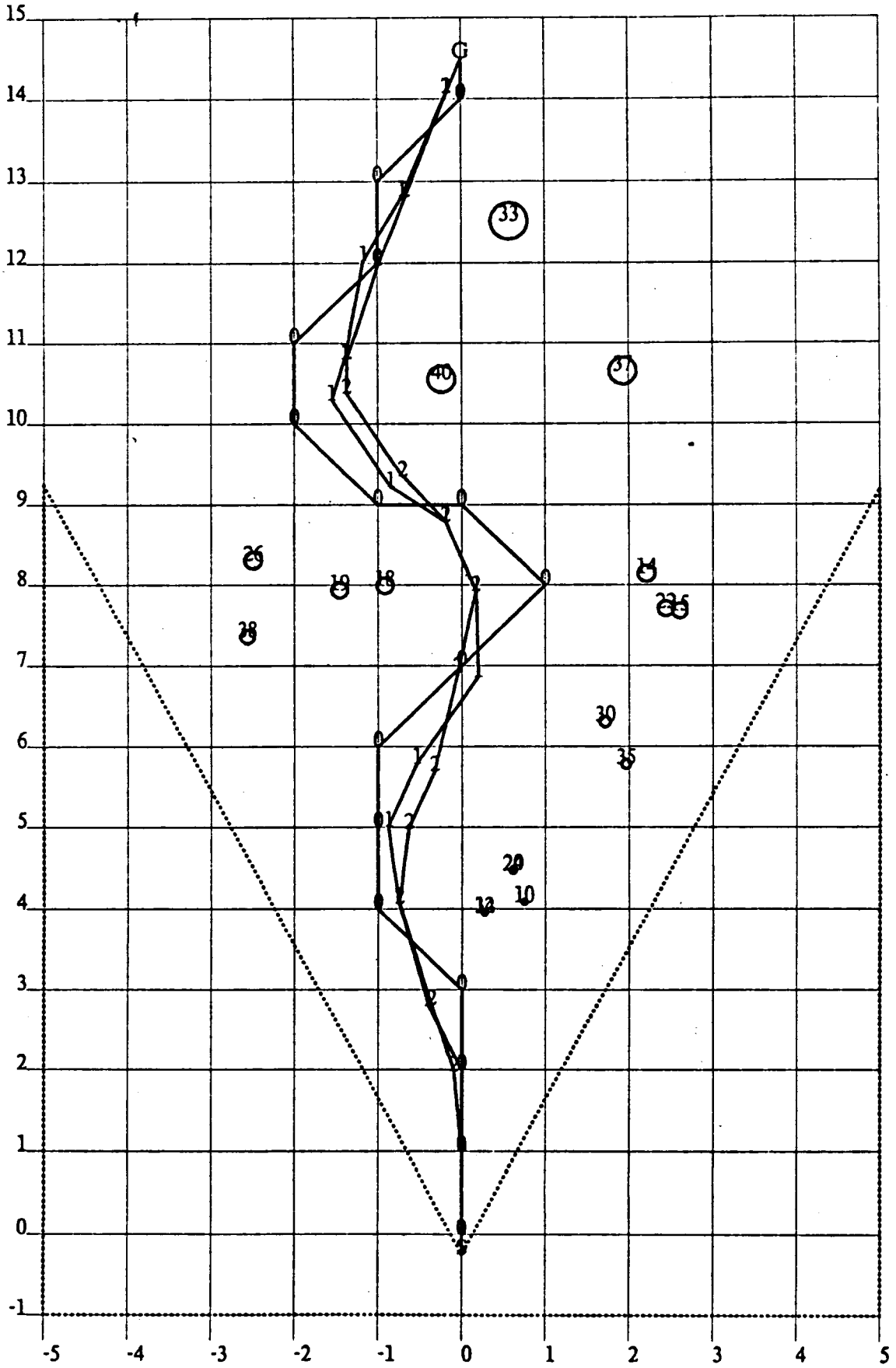
Measure of Matching $\dfrac{\sum X_i Y_i}{\sum X_i^2 \sum Y_i^2}$

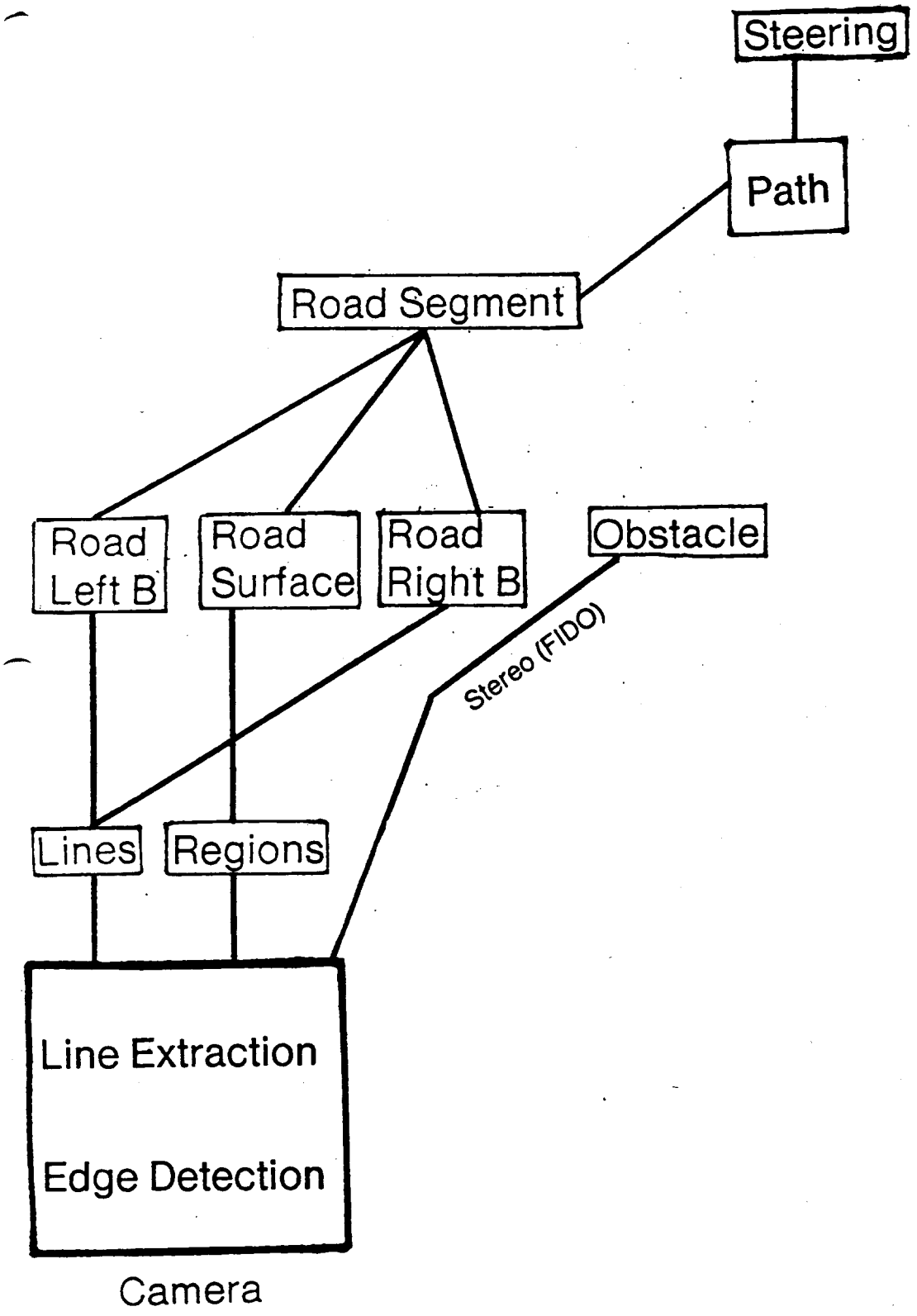






Path Planning




Steering
Path
Road Segment
Road Left B
Road Surface
Road Right B
Obstacle
Stereo (FIDO)
Lines
Regions
Line Extraction
Edge Detection
Camera

SONAR

- Simpel Range Sensing
- Inexpensive
- Relatively Fast
- Low Rate of Data

- Low Resolution
- Low Accuracy
- Noisy (Multiple/Total Reflection)

Polaroid Sonar Device

- "Instrument Grade"
- Range Limits: 0.9 to 35 feet
- Wide Beam Sonar: main lobe is 30^{o} wide
- Sonar Pulse: 1 ms chirp consisting of 56 pulses at 4 frequencies (60 kHz, 57 kHz, 53 kHz, 50 kHz)
- Elapsed Time is converted to distance: if target is 2 ft away => t = 3.55 ms

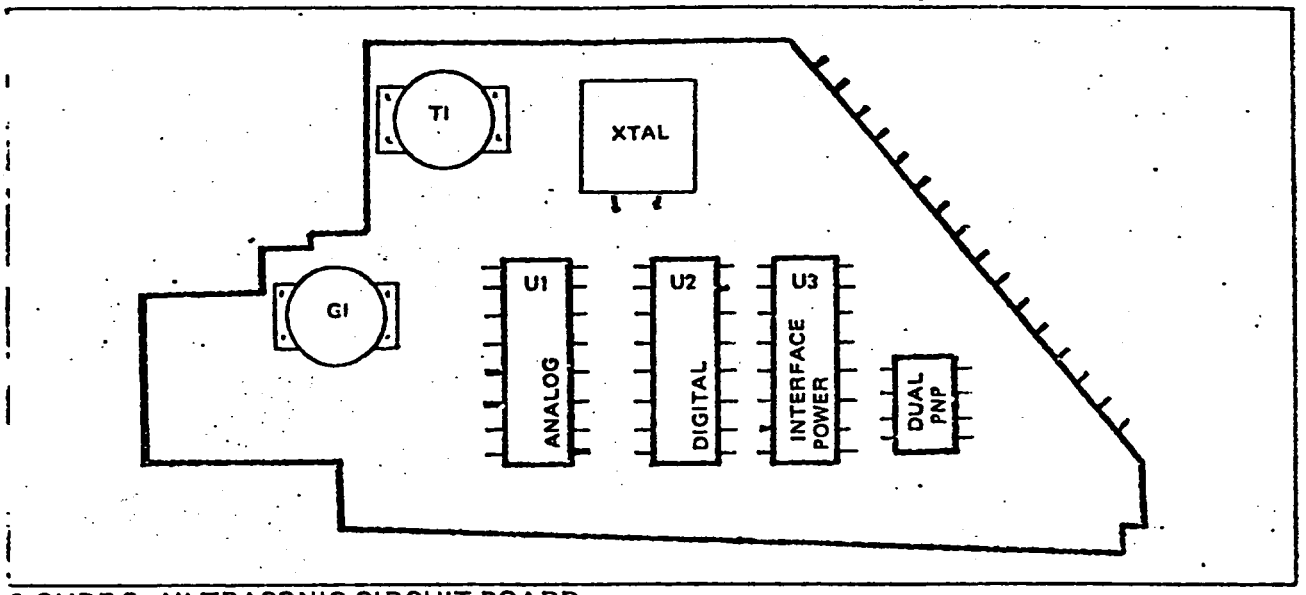


FIGURE 3. ULTRASONIC CIRCUIT BOARD

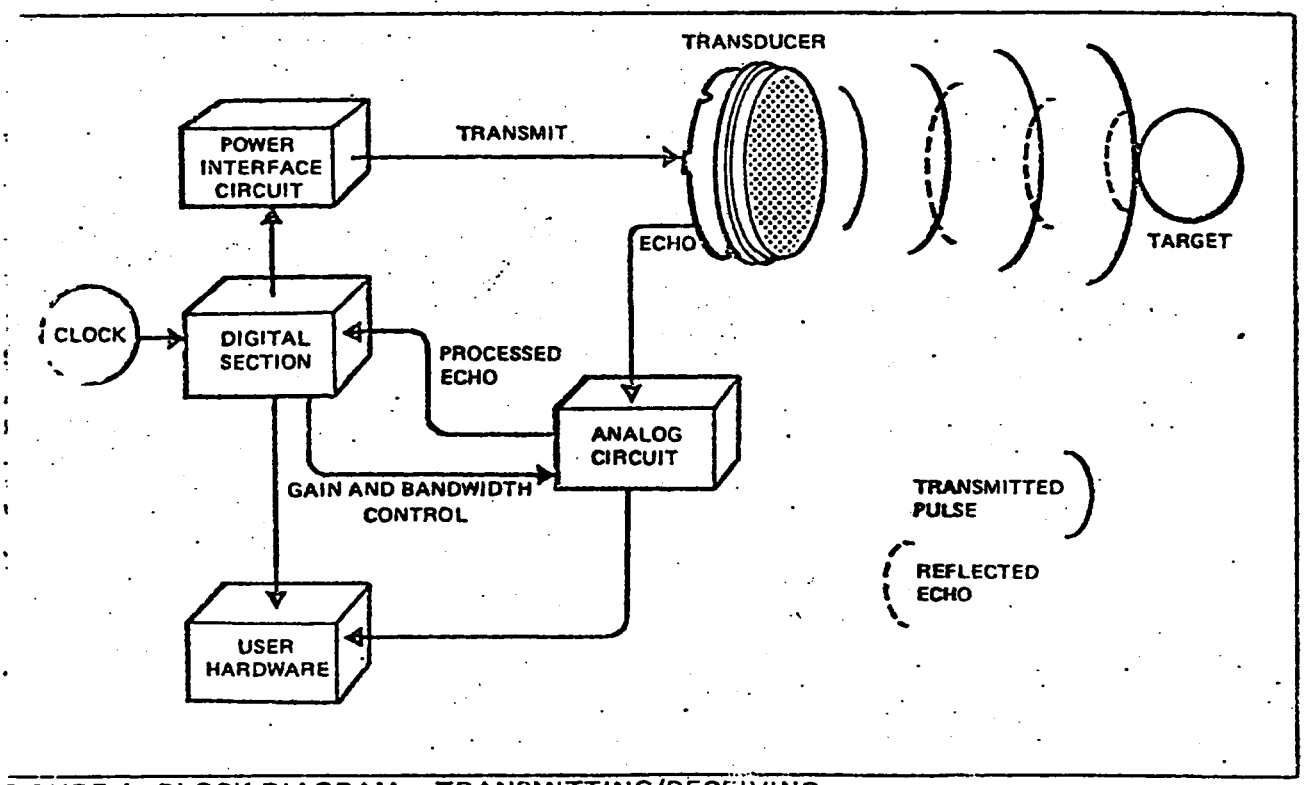


FIGURE 4. BLOCK DIAGRAM – TRANSMITTING/RECEIVING

Building Maps Using a Wide-Beam Sonar

- Sonar beam is described by a cone of probabilities.
- Map is constructed on a 2D grid.
- Mapping information is encoded in the grid using the following convention:

UNKNOWN	--->	0
OCCUPIED	--->	(0, +1]
EMPTY	--->	[-1, 0)

Empty/Occupied Probability Profile from One Sonar Reading

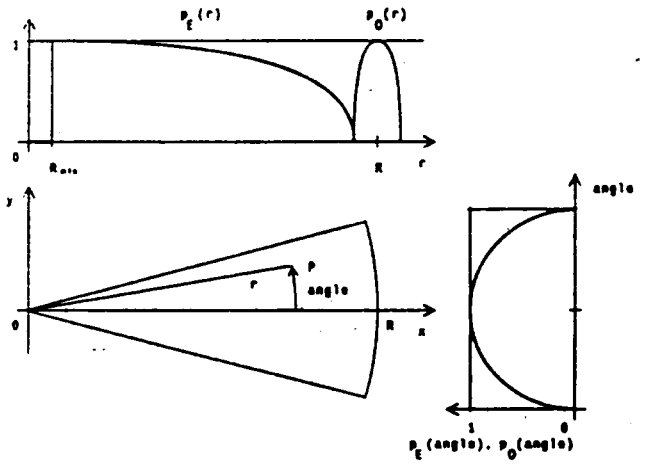
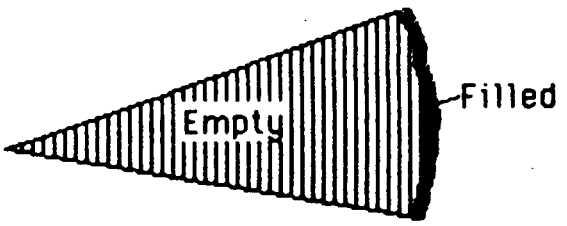
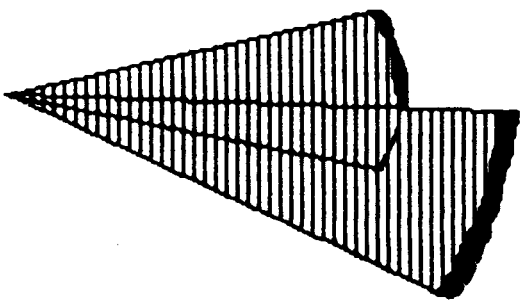


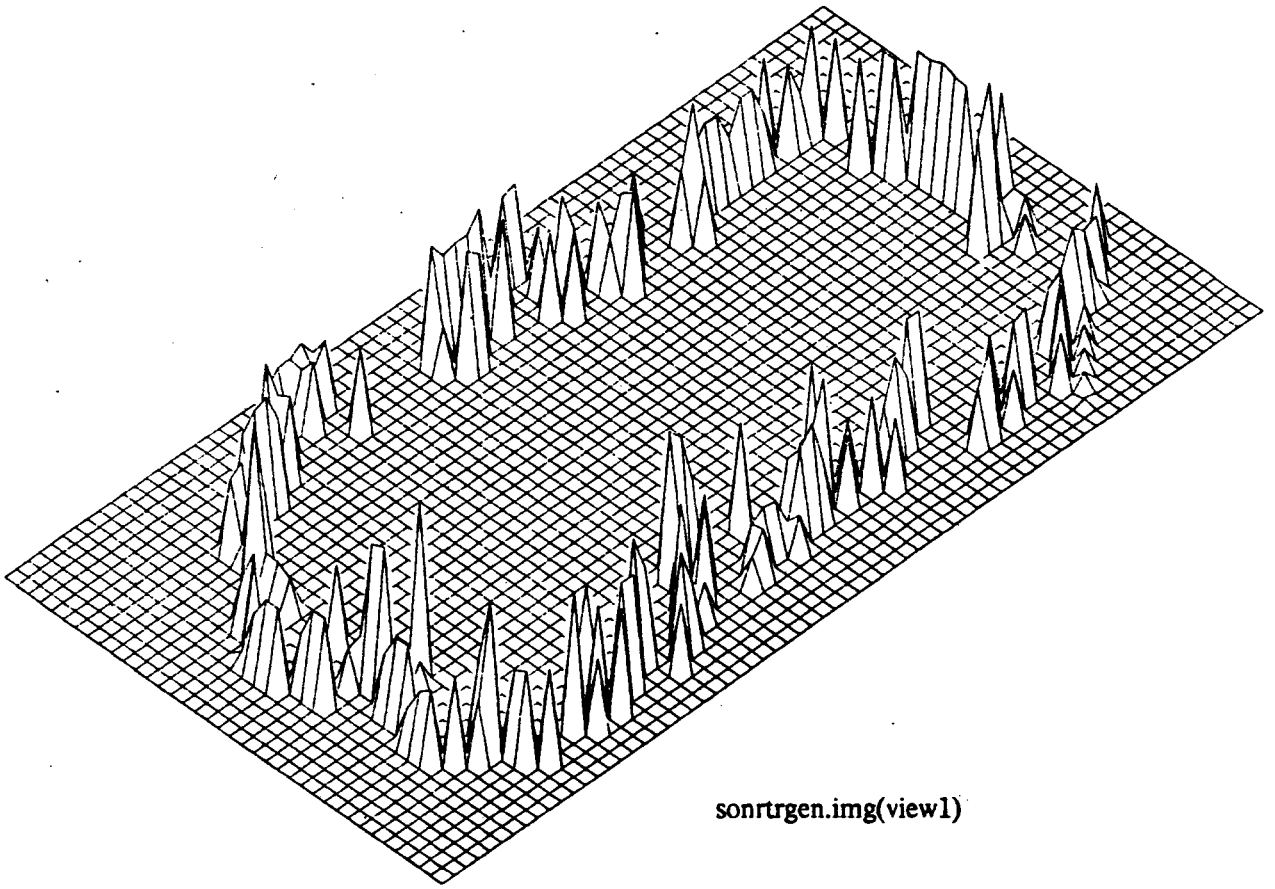
Figure 2: The Probability Profiles corresponding to the probably Empty and somewhere Occupied regions in the sonar beam. The profiles represent a horizontal cross section of the beam.

- We can combine many readings into an occupancy map.



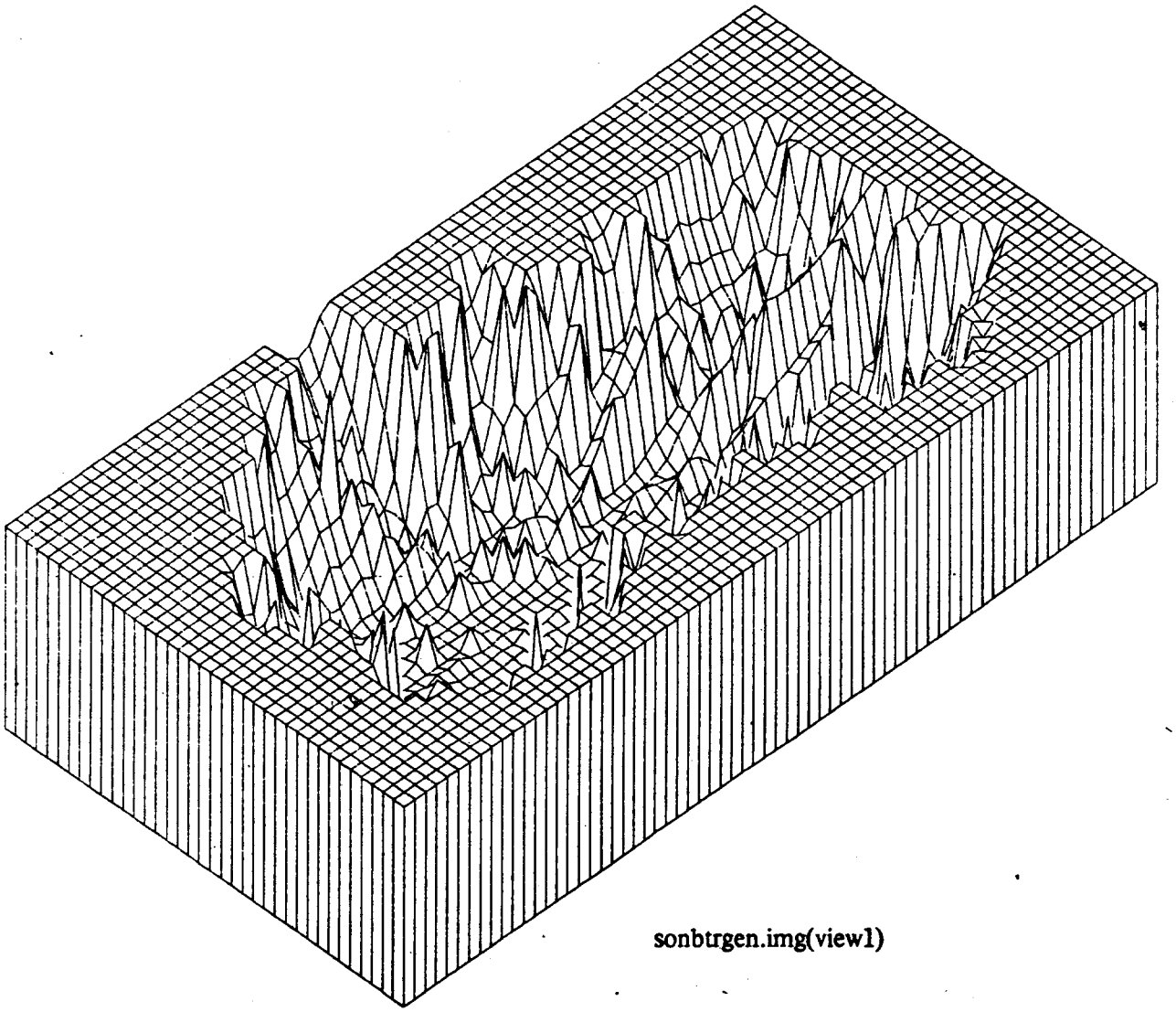


Occupiedness

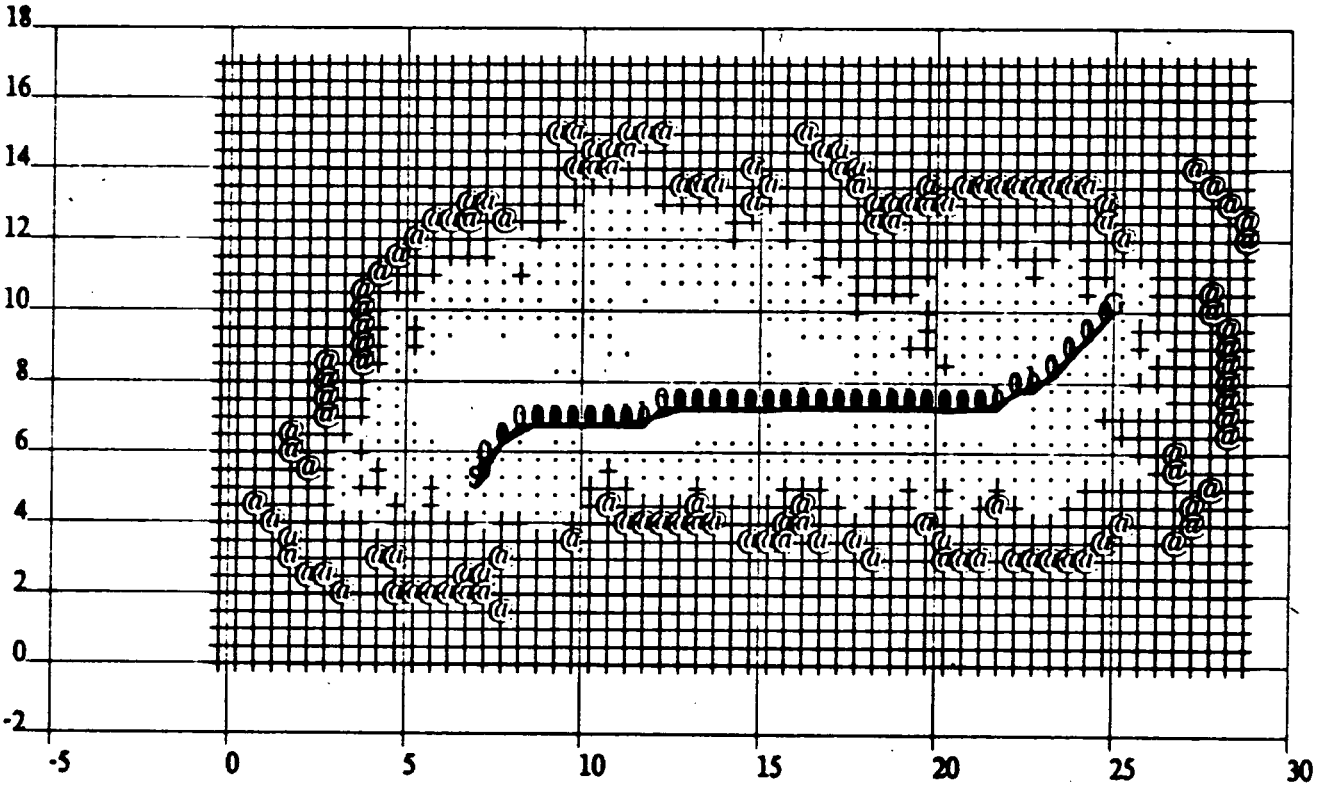


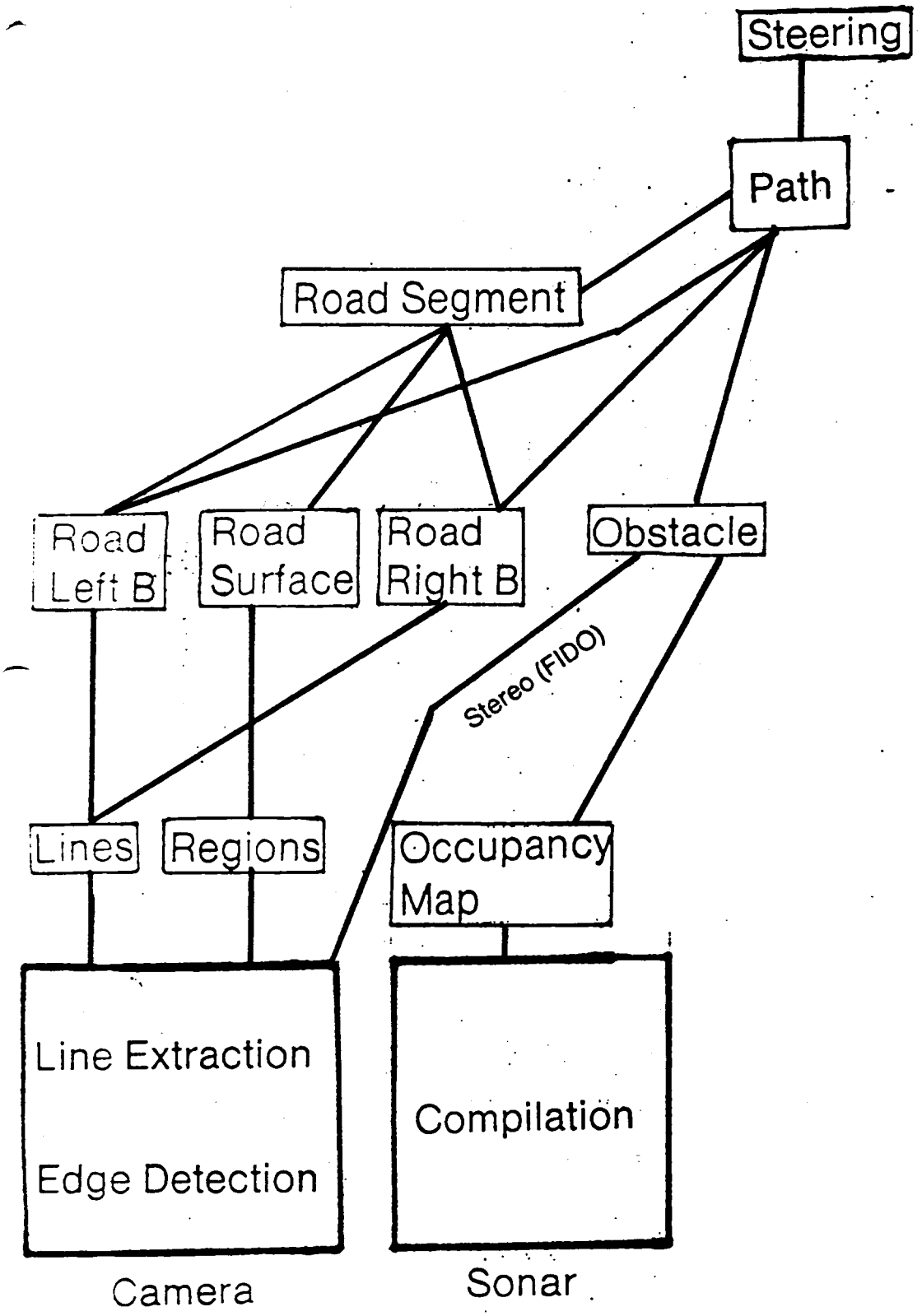
sonrtrgen.img(view1)

Emptyness



sonbtrgen.img(view1)


18
16
14
12
10
8
6
4
2
0
-2
-5
0
5
10
15
20
25
30


Steering
Path
Road Segment
Road Left B
Road Surface
Road Right B
Obstacle
Stereo (FIDO)
Lines
Regions
Occupancy Map
Line Extraction
Edge Detection
Compilation
Camera
Sonar

ERIM-ALV Range Finder

Field of View	
Horizontal	80°
Vertical	30°
IFOV	0.5°
Range Image	256 x 64 pixels
Range Noise	0.4 ft at 50 ft
Frame Rate	2 Hz
Ambiguity Interval	64 ft
Laser	100 mW, 0.82μm
Reflectance Data	Available

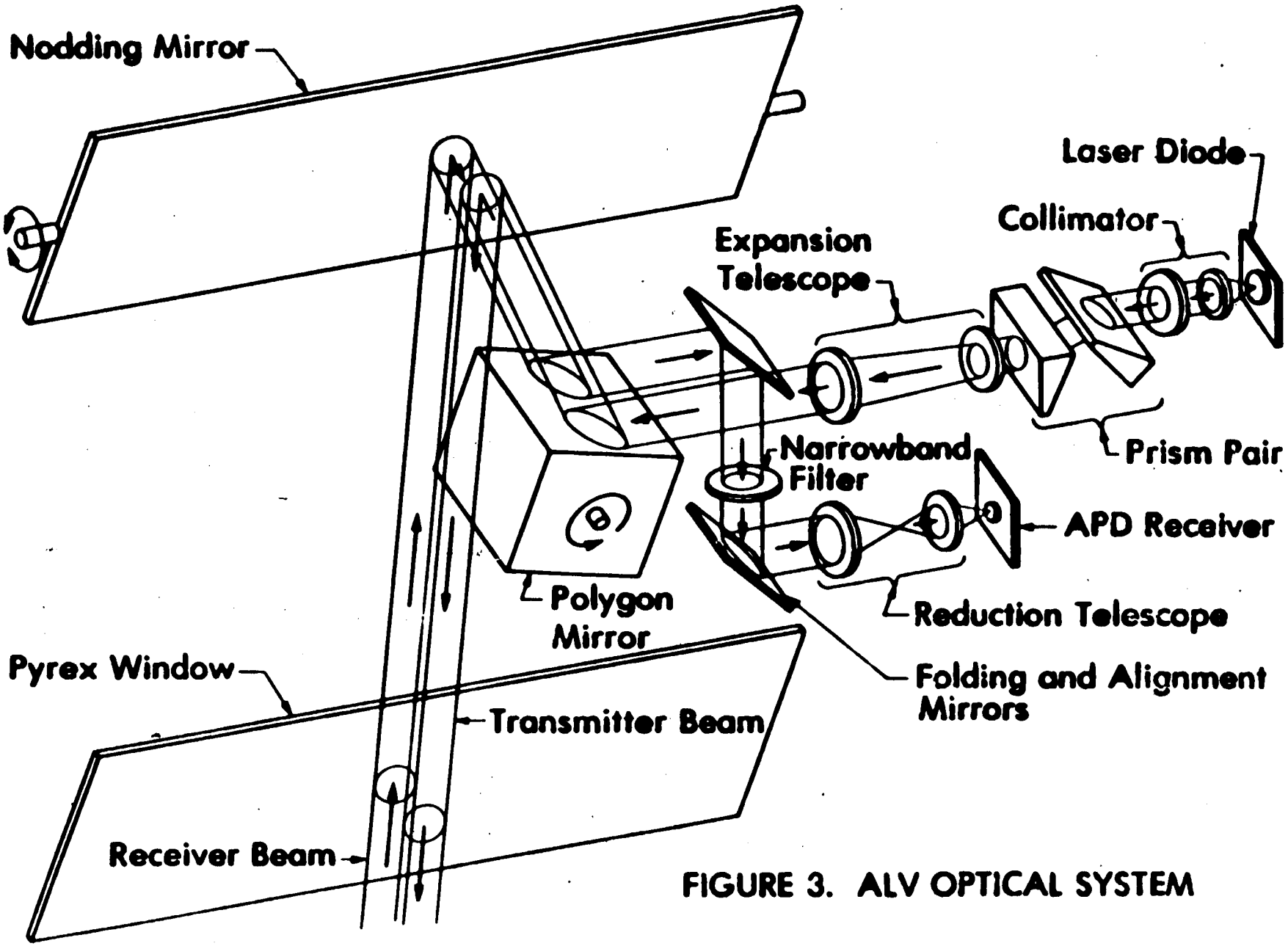
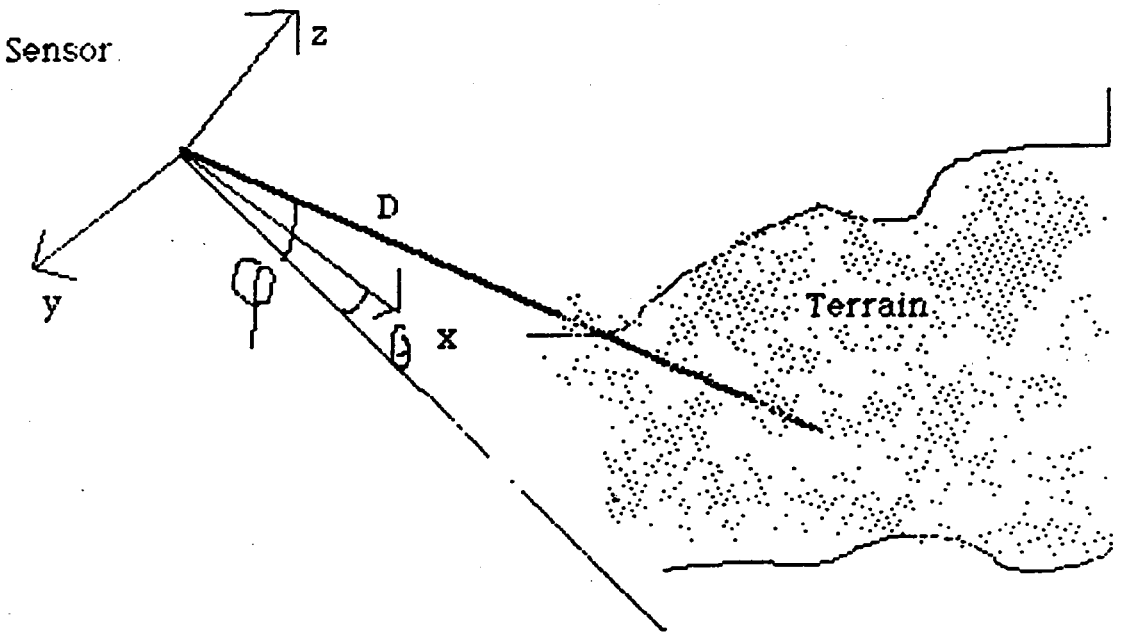
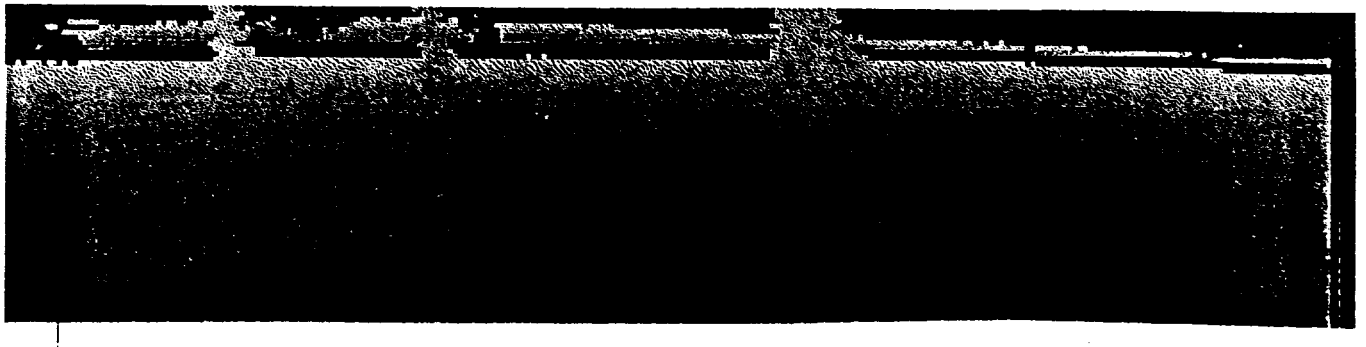


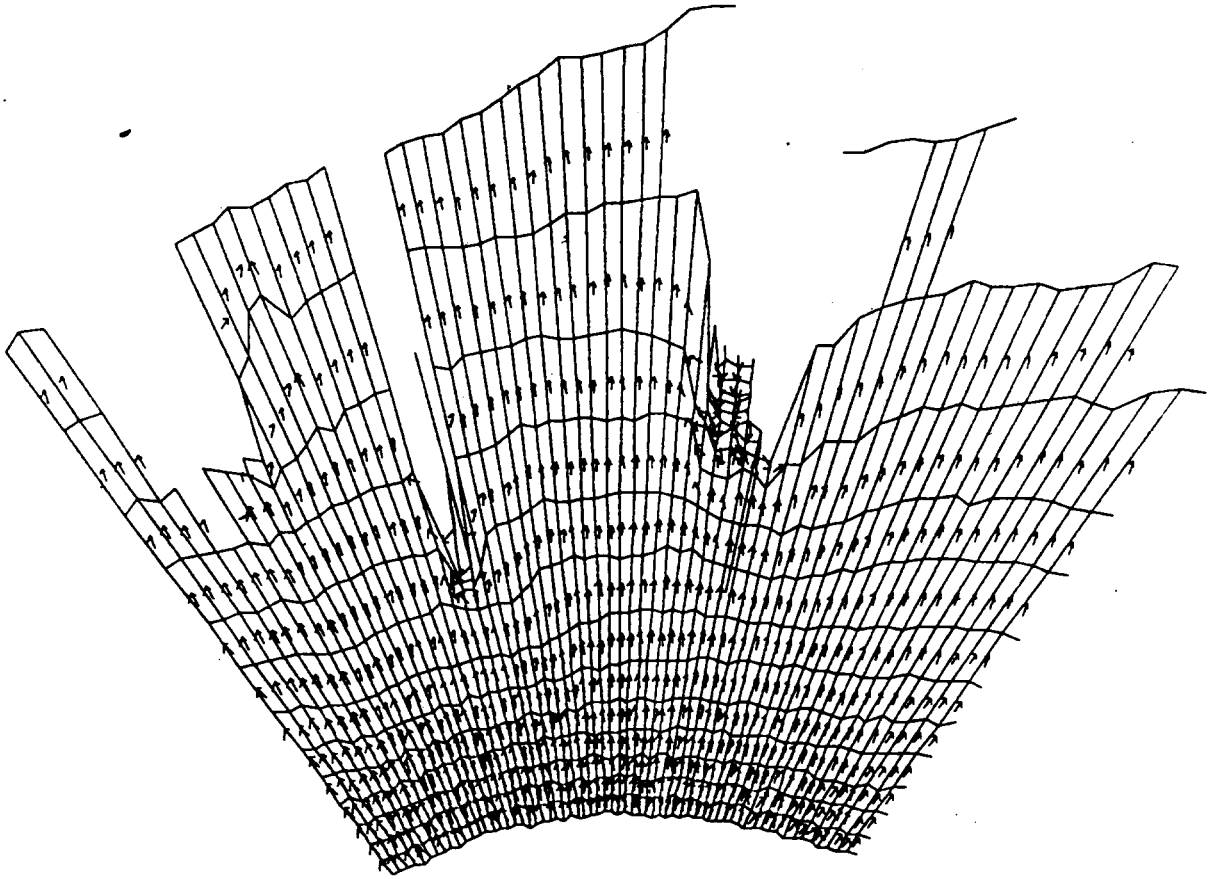
FIGURE 3. ALV OPTICAL SYSTEM


Sensor
z
y
D
x
Terrain

ERIM Raw Data

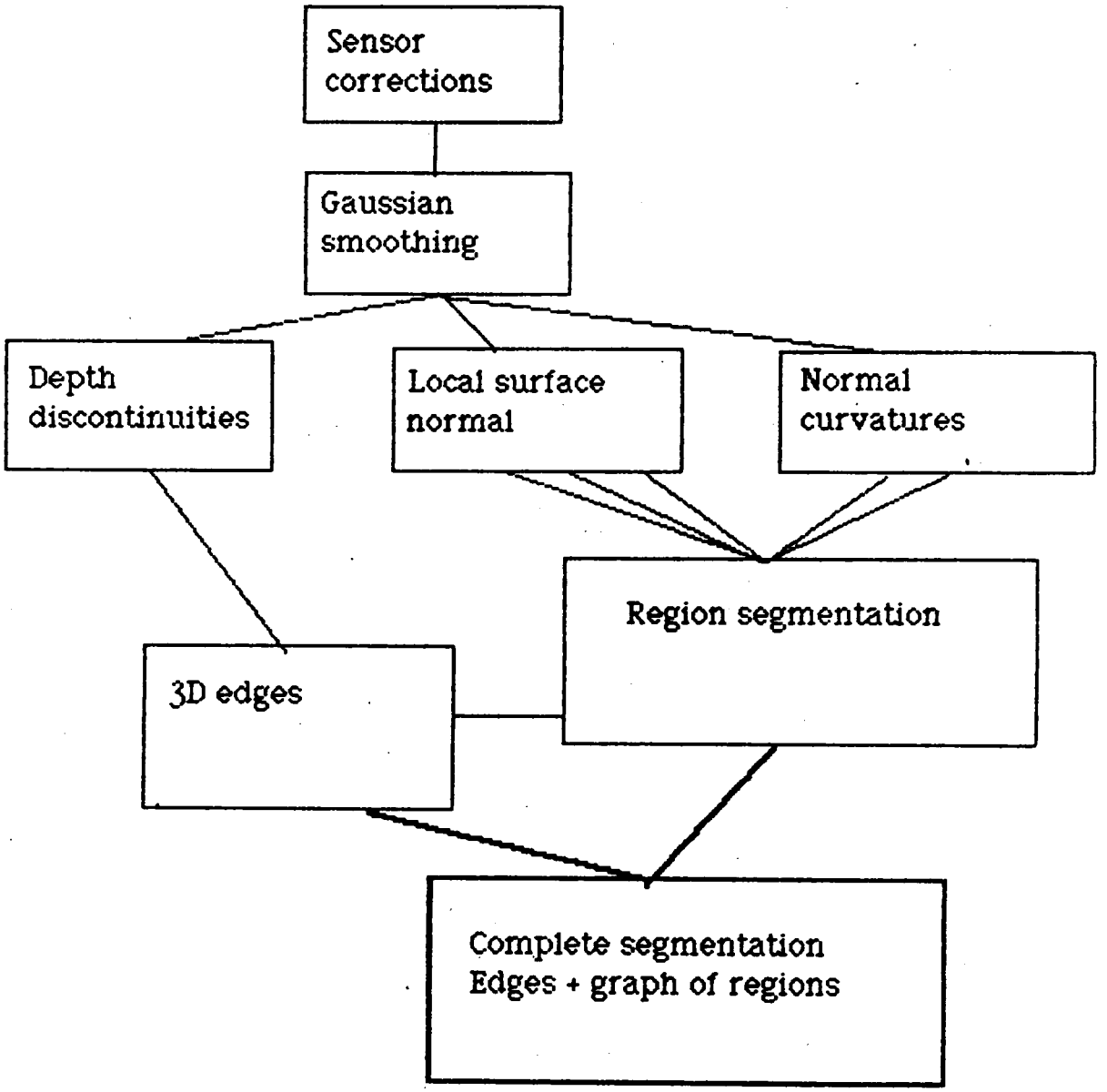


Mesh Representation with Surface Orientation

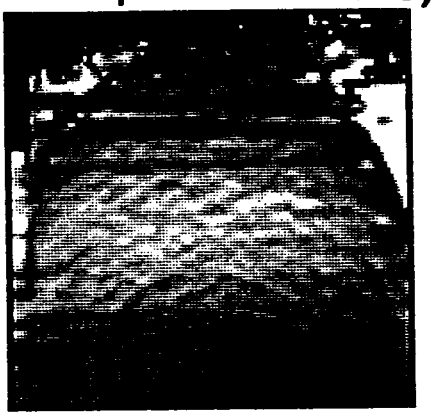


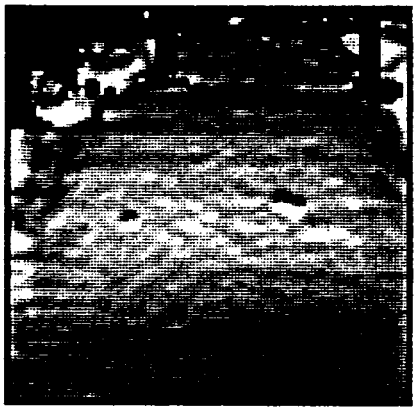
: /visk/hebert/park/erim26.img

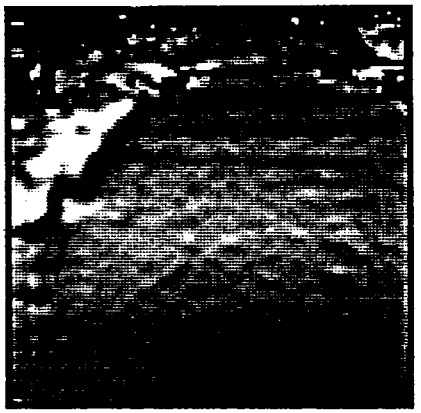
Range data processing:

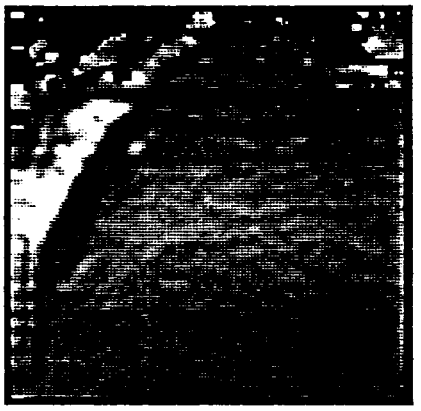


Smoothness
(Principal Curvature)

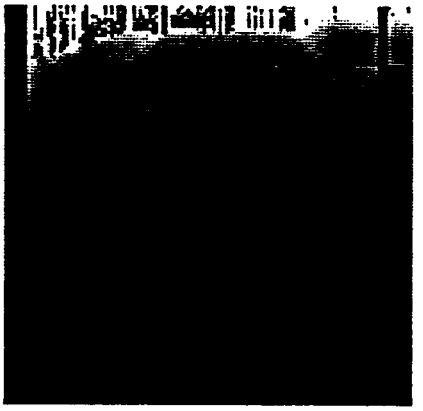


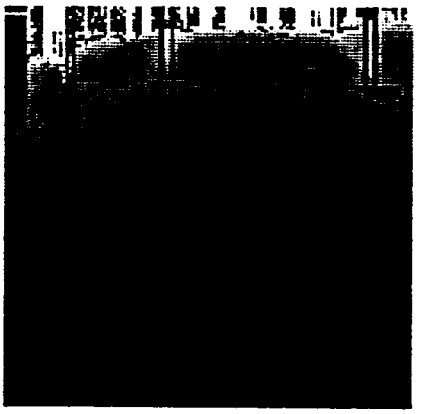


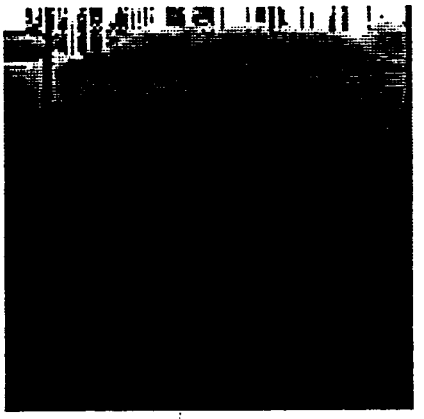


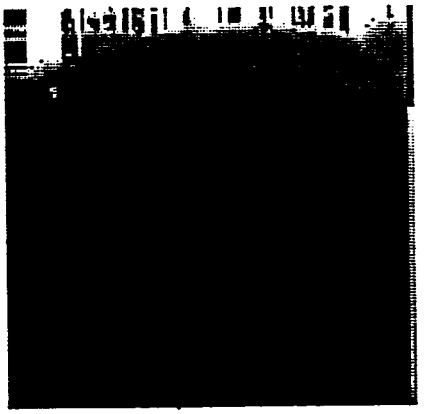


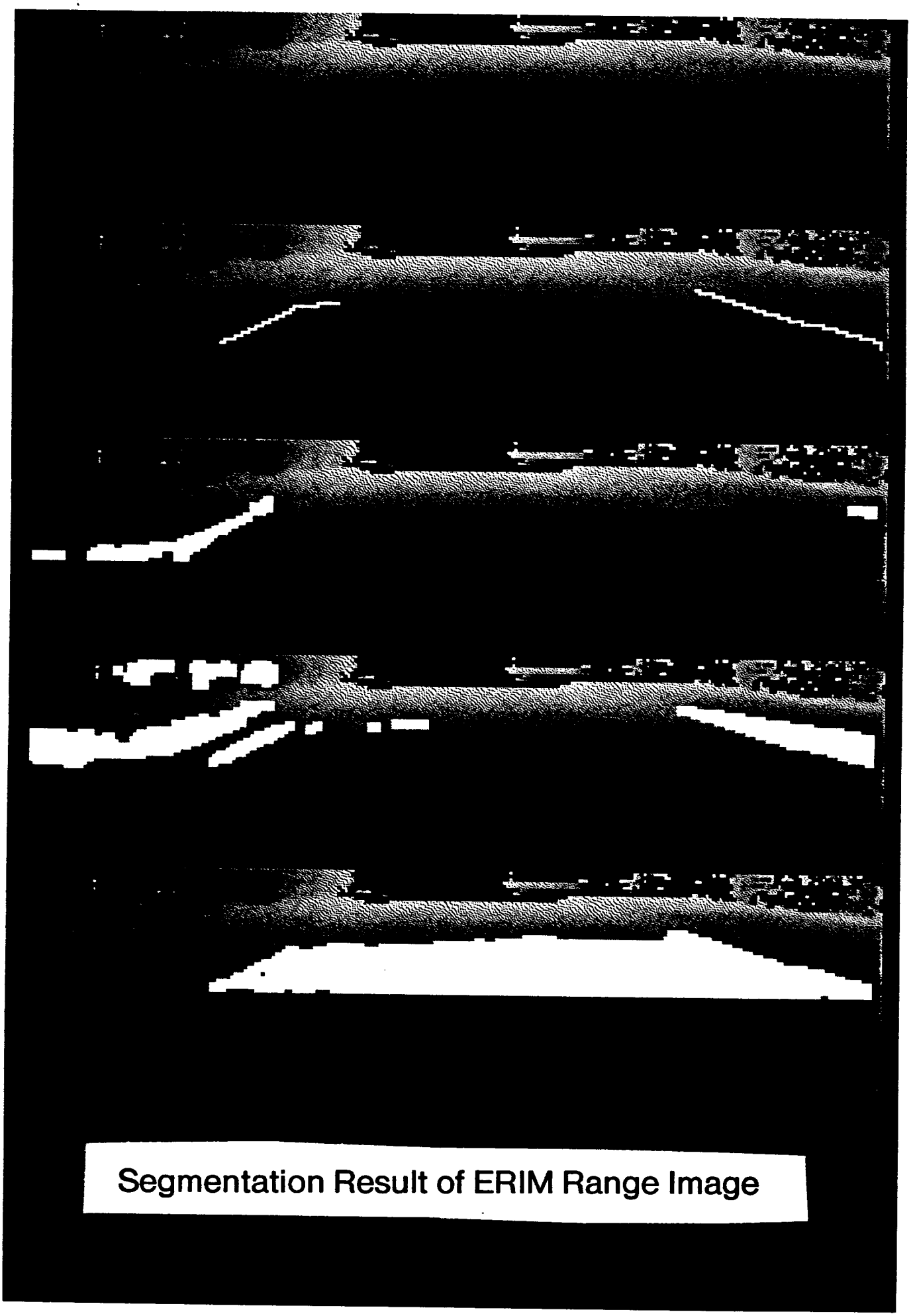
Original Data



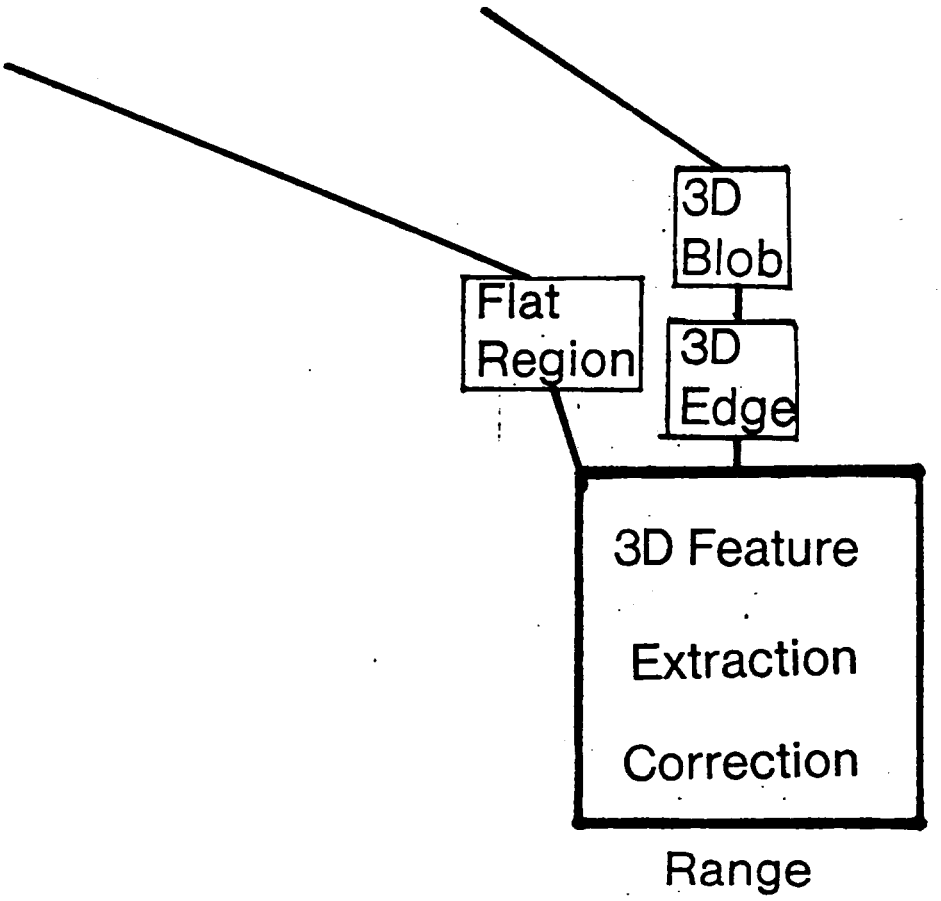


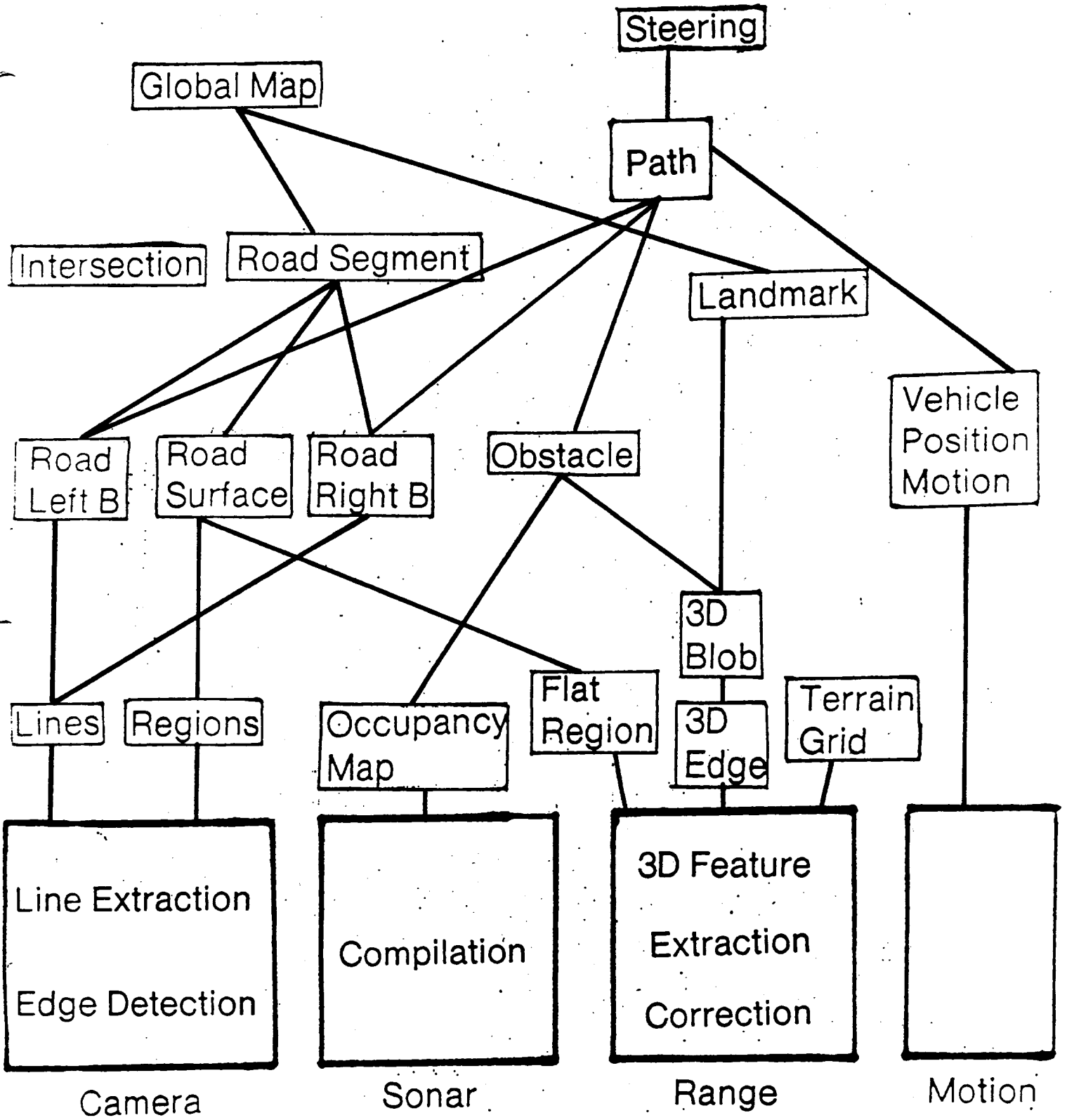






Segmentation Result of ERIM Range Image


3D
Blob
Flat
Region
3D
Edge
3D Feature
Extraction
Correction
Range


Steering
Global Map
Path
Intersection
Road Segment
Landmark
Vehicle Position Motion
Road Left B
Road Surface
Road Right B
Obstacle
3D Blob
Lines
Regions
Occupancy Map
Flat Region
3D Edge
Terrain Grid
Line Extraction
Edge Detection
Compilation
3D Feature
Extraction
Correction
Camera
Sonar
Range
Motion

Hierarchical Data Structure for Navigation