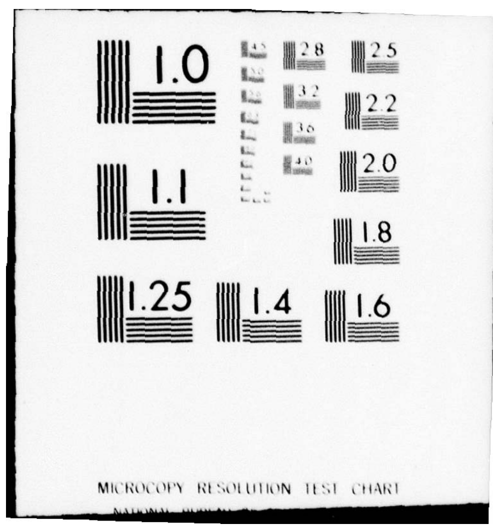
AD-A065 467 ARMY MISSILE RESEARCH AND DEVELOPMENT COMMAND REDSTO--ETC F/G 17/7
PRELIMINARY ANALYSIS OF AUTOMATIC SCENE CORRELATION BETWEEN SPE--ETC(U)
OCT 78 W W MALCOLM, J S BOLAND

UNCLASSIFIED DRDMI-T-78-85 NL

1 OF 1
AD
A 065467

END
DATE
FILMED
4-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART

LEVEL II

(12)

TECHNICAL REPORT T-78-85

U.S. ARMY
MISSILE
RESEARCH
AND
DEVELOPMENT
COMMAND

PRELIMINARY ANALYSIS OF AUTOMATIC SCENE CORRELATION BETWEEN SPECTRALLY NON-COMPATIBLE IMAGERY

William W. Malcolm
Dr. Joseph S. Boland III
Guidance and Control Directorate
Technology Laboratory

U.S. ARMY MIRADCOM

11 October 1978

Redstone Arsenal, Alabama 35809

Approved for Public Release; Distribution Unlimited.

DISPOSITION INSTRUCTIONS

DESTROY THIS REPORT WHEN IT IS NO LONGER NEEDED. DO NOT RETURN IT TO THE ORIGINATOR.

DISCLAIMER

THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION UNLESS SO DESIGNATED BY OTHER AUTHORIZED DOCUMENTS.

TRADE NAMES

USE OF TRADE NAMES OR MANUFACTURERS IN THIS REPORT DOES NOT CONSTITUTE AN OFFICIAL ENDORSEMENT OR APPROVAL OF THE USE OF SUCH COMMERCIAL HARDWARE OR SOFTWARE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
T -78-85

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
6 Preliminary Analysis of Automatic Scene Correlation Between Spectrally Non-Compatible Imagery.

5. TYPE OF REPORT & PERIOD COVERED
9 Technical Report.

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
10 William W. Malcolm
Joseph S. Boland, III

8. CONTRACT OR GRANT NUMBER(s)
12 55p.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Commander
US Army Missile Research and Development Command
ATTN: DRDMI-TG
Redstone Arsenal, Alabama 35809

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
16 1W362393A214
632303.11.21402

11. CONTROLLING OFFICE NAME AND ADDRESS
Commander
US Army Missile Research and Development Command
ATTN: DRDMI-TI
Redstone Arsenal, Alabama 35809

12. REPORT DATE
11 11 October 1978

13. NUMBER OF PAGES
52

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
14 DRDMI-T-78-85

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for Public Release;
Distribution Unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ALGORITHMS	TADS	HUGH AIRCRAFT
TEST TOWER	SPAL	NASA
SYSTEMS INPUT	LR	IRS THRESHOLD
MIRADCOM	HR	IRGTH

393 427 Jm

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A concept which might be utilized in the development of a modern attack helicopter weapon system could combine a target acquisition system and air-launched terminal homing missile to provide the capability for long range target engagement. If a laser semi-active system is employed, continuous laser designation would be required from missile launch to impact. This would increase the helicopter exposure and vulnerability to anti-aircraft weapons.

79 03 06 034

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

CONTENTS

ILLUSTRATIONS

ILLUSTRATIONS

ILLUSTRATIONS (Concluded)

INTRODUCTION

A concept which might be utilized in the development of a modern attack helicopter weapon system could combine a target acquisition system and air-launched terminal homing missiles to provide the capability for long range target engagement. If a laser semi-active system is employed, continuous laser designation would be required from missile launch to impact. This would increase the helicopter exposure and vulnerability to anti-aircraft weapons.

In order to eliminate this designation requirement, imaging missile seekers may be developed to provide the capability for automatic target tracking once acquired by the seeker, thus allowing the attack helicopter to remask after missile launch. There are two main types of imaging seekers: Those which have sensitivity in the visible (.5 to .8 μ) spectrum, and those in the infrared (3-5 or 8-14 μ).

The Army has apparently chosen to continue development of IR seekers. Size and cost constraints dictate that these seekers be low resolution units and range considerations require wide fields-of-view. These characteristics severely limit the gunner's capability to acquire and recognize the intended target by viewing the seeker imagery. Therefore the gunner must utilize some other sensor to accomplish these tasks. Assuming the attack helicopter would contain a high resolution target acquisition system through which the gunner could recognize potential targets, these targets must then be handed-off to the specific imaging missile seeker. The time required for this hand-off is of major importance in this concept.

MIRADCOM's Automatic Tracking and Integrated Fire Control A214 Missile Technology Program is investigating methods for reducing the hand-off time and thereby reducing helicopter exposure time. The initial program phases involved analysis and hardware development for providing automatic hand-off between imaging systems having the same spectral sensitivity, e.g., TV to TV, utilizing available hardware, as well as investigating problems relating to manual target hand-off. The manual hand-off mechanization requires the gunner to alternately switch the viewed video between the target acquisition system and missile seeker until the correct target has been placed within the seeker tracking gates. The results of these experiments indicate a significant amount of exposure time required to achieve this target hand-off [1,2].

As has been previously noted, the imaging seeker which has been selected for development by the Army is that with spectral sensitivity in the IR region. This decision surfaced an additional problem relating to target hand-off. The high resolution target acquisition system may have both TV and IR high resolution sensors with TV providing superior performance under specific conditions. Thus the automatic correlation system must accept targets as acquired and recognized by this TV system and automatically hand-off the selected target to the IR seeker. The technical problems related to non-compatible images are currently being investigated to determine the "best" algorithm for providing the automatic correlation.

This report presents the results of a preliminary analysis investigating automatic scene correlation between spectrally non-compatible imagery. Two edge detection algorithms were investigated and digitized video scenes from a precision target acquisition system (TV) and imaging missile seeker (IR) were utilized as correlation inputs. Two specific scenes were selected due to their different types of scene content. These were a NASA dynamic test tower and a building parking lot. Correlation and preprocessor algorithms were investigated using these inputs.

2. EDGE DETECTION ALGORITHMS

In the initial phase of this technology program, emphasis was placed on correlation of two images obtained from similar sensors, both sensitive in the .5 to .85 micron spectral range. The main considerations were scaling of the high resolution (HR) and low resolution (LR) sensor images, size of the reference array, and correlation threshold. However, for systems where the sensors have different spectral sensitivity as well as different resolution, the images differ significantly. It became obvious that additional preprocessing of the imagery prior to correlation would be required. In observing the video display of the TV and IR scenes, it appeared that if each scene could be converted to an "outline drawing" (digital array) one could preserve the important edges in the original scenes. Eventhough the modified scene would generally contain less information than the original scene, it was felt that the "outline drawing" for the two different spectral response sensors would appear similar; thus correlation could be performed. This "outline drawing" or edge map could be produced by emphasizing regions containing abrupt dark-light transitions, and de-emphasizing regions of approximately homogeneous intensity.

Two edge detection algorithms are included in this analysis (a 2×2 and a 3×3 edge detection algorithm). Each scene was evaluated using each of these algorithms. The "two by two" method is known as the Robert Cross operator [3].

Assume that the digital picture is represented by the two-dimensional function g(x,y). Then the magnitude of the gradient at pixel (i,j) can be approximated by

$$R(i,j) = \Big\{[g(i,j)-g(i+1,j+1)]^2 + [g(i,j+1)-g(i+1,j)]^2\Big\}^{1/2} . \qquad (1)$$

Equation (1) is the general form of the Roberts Cross Operator. From Equation (1) it can be seen that in picture areas of constant gray level, R(i,j) will be zero and in picture areas of high gray level change in either the x or y or both directions, R(i,j) will be large. *Figure 1* is a pixel representation of the operation computed in Equation (1).

The second edge detection algorithm operates on a 3 × 3 array of pixels centered on the pixel being investigated as shown in *Figure 2*. To determine if pixel (i,j) is an edge point in the digital picture function g(x,y), the gradient magnitudes in the x and y directions are calculated as follows:

$$
\begin{aligned}
S_X(i,j) = {} & [W_1 \cdot g(i-1,j+1) + W_2 \cdot g(i,j+1) \\
& + W_3 \cdot g(i+1,j+1)] - [W_4 \cdot g(i-1,j-1) \\
& + W_5 \cdot g(i,j-1) + W_6 \cdot g(i+1,j-1)]
\end{aligned} \tag{2}
$$

and

$$
\begin{aligned}
S_Y(i,j) = {} & [W_1 \cdot g(i+1,j-1) + W_2 \cdot g(i+1,j) \\
& + W_3 \cdot g(i+1,j+1)] - [W_4 \cdot g(i-1,j-1) \\
& + W_5 \cdot g(i-1,j) + W_6 \cdot g(i-1,j+1)]
\end{aligned} \tag{3}
$$

where, in this report $W_1=W_3=W_4=W_6=1$ and $W_2=W_5=2$ in all simulations using the 3 × 3 gradient except as noted in Section 3.B. Appendix A provides justification for selecting these values.

An estimate of the gradient at point (i,j) is given by

$$S(i,j) = [S_X(i,j)]^2 + [S_Y(i,j)]^2. \tag{4}$$

The digital picture is then reduced to binary form by comparing R(i,j) or S(i,j) to a preset threshold such that

$$
T(i,j) = \begin{cases} 1, & S(i,j) \text{ or } R(i,j) \geq GTH \\ 0, & S(i,j) \text{ or } R(i,j) < GTH \end{cases} \tag{5}
$$

where T(i,j) is the binary picture and GTH is the threshold value. If g(x,y) is of the size N × M, then T(i,j) is of the size (N-1) × (M-1) for the 2 × 2 element detector and (N-2) × (M-2) for the 3×3 detector.

The 2 × 2 edge algorithm (1) appears more sensitive to picture noise than the 3 × 3 algorithm (4). The next section will describe the results of applying these two algorithms to various digitized TV and IR scenes.

3. ANALYSIS AND SIMULATION PROGRAM

The digital simulation described in this report was performed on a Tektronix Model 4051 digital computer. The memory capability of this machine restricted the correlation surface to a 28 × 28 pixel array. To investigate the correlation surface for various low resolution scene positions required manual insertation of the corresponding 28 × 28 low resolution array.

A. SYSTEM INPUTS

The high resolution (HR) TV input imagery was obtained from MIRADCOM's Stabilized Platform

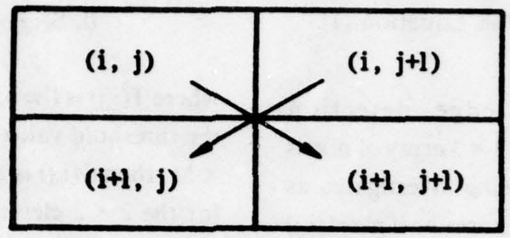


Figure 1. Pixel representation of the Roberts cross operator.

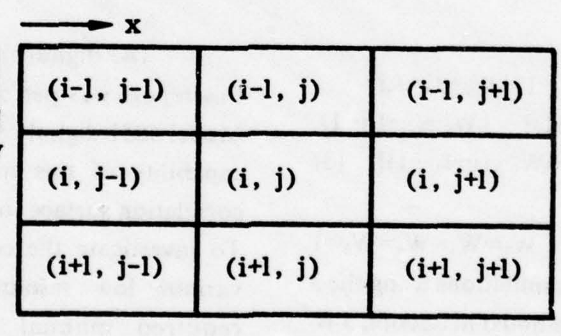


Figure 2. Pixel representation of the 3 X 3 edge operator.

Airborne Laser System (SPAL) which contains a narrow field-of-view silicon videcon. The low resolution (LR) infrared missile seeker input imagery was obtained using a Hugh Aircraft developed IRIS unit. The LR sensors field-of-view was four times larger than the HR sensor. A video field from each sensor was selected and a 240 × 256 pixel array was generated. Each pixel was quantized to eight bits or to 256 gray levels. Since the high resolution sensor's field-of-view was one-fourth that of the low resolution sensor, a single pixel was generated for each four-by-four subarray in the original field. This process was required to equalize the spatial resolution of pixels from the two images.

The two scenes used in this study, a NASA tower and a parking lot, are shown in the sequence of *Figures 3-8*. *Figures 3* and *6* are the scenes as viewed by the high resolution TV sensor. *Figures 4* and *7* represent the same TV scenes after being reduced 4:1 for use as the reference scene. *Figures 5* and *8* are the IR low resolution scenes to which the high resolution is correlated. The black square in the figures are the areas of initial correlation, while the dashed square indicates the correlation area when both the high and low resolution scenes are positioned lower to reduce gradient values. This point will be discussed later in this report.

B. ANALYSIS PROCEDURE

(1) 3 × 3 GRADIENT ALGORITHM OF NASA TOWER (*Figures 3, 4* and *5*).

The maximum size of the high resolution sensor reference array which was used in the simulation was 28 × 28. A digital overstrike plot was made of both the high and low resolution digital scenes. From these plots, a "best" guess of where the expected match point between the scenes would occur and a 28 × 28 matrix array of the low resolution sensor at this location was selected as the initial correlation analysis surface. After a complete analysis was performed the low resolution scene was moved by one or more columns and rows, equivalent to moving the sensor field-of-view, and the procedure was repeated to locate the x,y coordinates of the low resolution sensor which maximized the pixel matches between the high and low sensors. Even though this simulation required manual data insertion, a full digital simulation was performed on a CDC6600 for automatic target scan.

The first step in the simulation was to derive the gradient matrix S(i,j), Equation (4), for the 28 × 28 matrix array for the high resolution TV sensor. This 26 × 26 matrix array was converted to a binary matrix by applying Equation (5). The selecting of the proper threshold value (GTH) for the high resolution image is critical in achieving maximum correlation. This point will be discussed further in this report. It is clear that if TVGTH were set at zero, then the binary matrix would contain all ones. Similarly, if TVGTH were set above the maximum value of the gradient matrix, then the binary matrix would

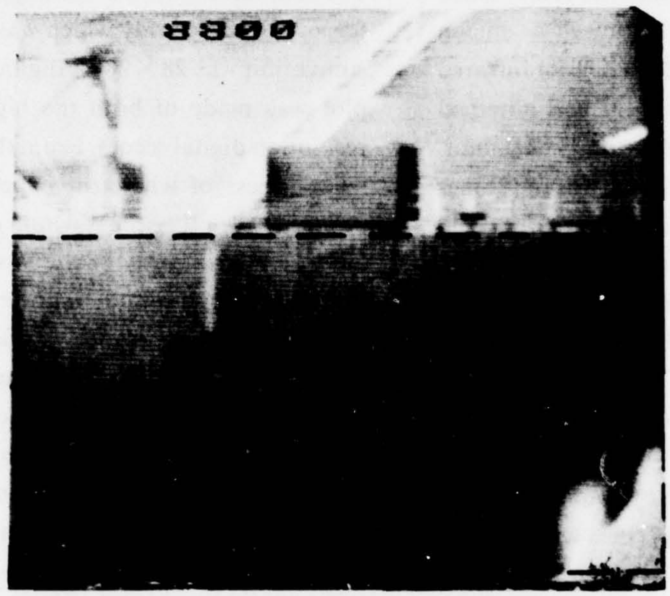


Figure 3. NASA tower high resolution TV narrow field-of-view scene input. (Solid line outlines area of initial digitized input. Dashed line outlines shifted scene input.)

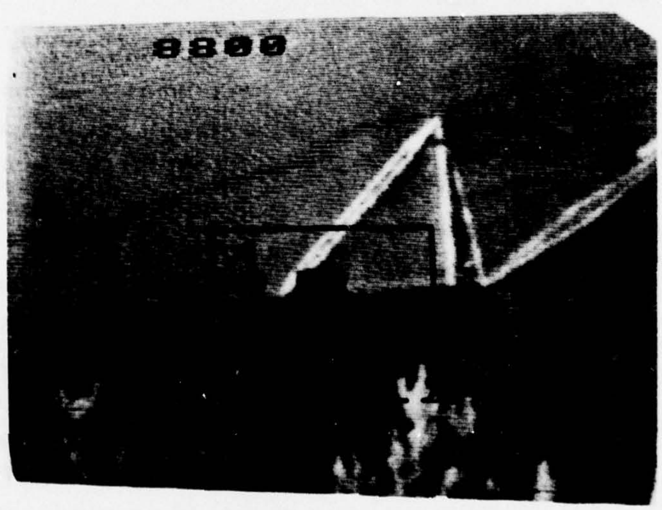


Figure 4. NASA tower high resolution TV wide field-of-view scene. Equivalent to 4:1 reduction of Figure 3. (Solid line outlines area of effective coverage. Dashed line outlines shifted scene input.)

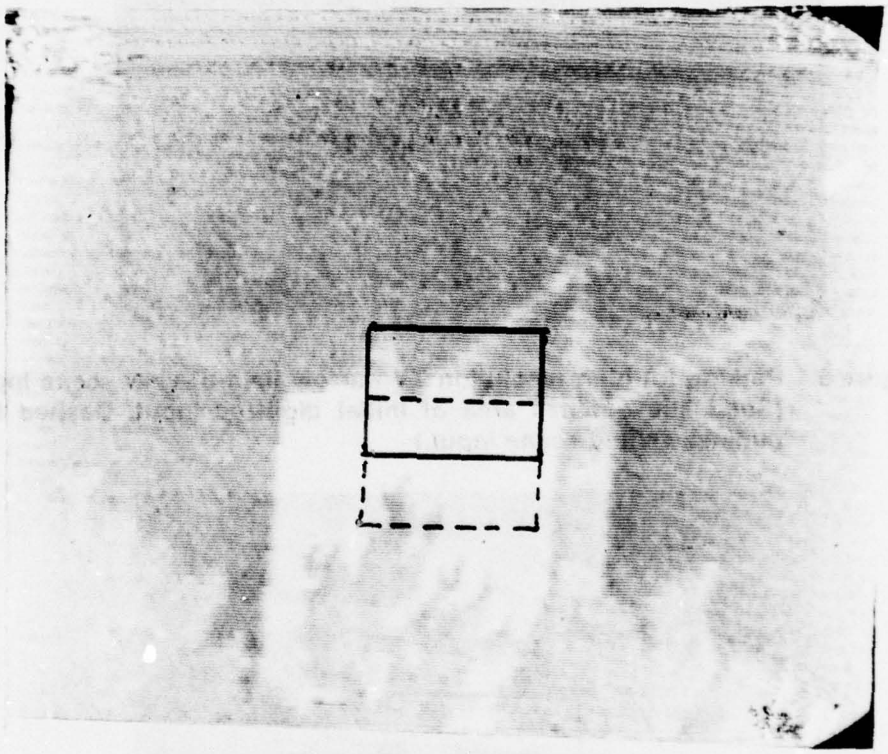


Figure 5. NASA tower low resolution IR input scene. (Solid line outlines area of initial correlation. Dashed line outlines shifted scene input.)

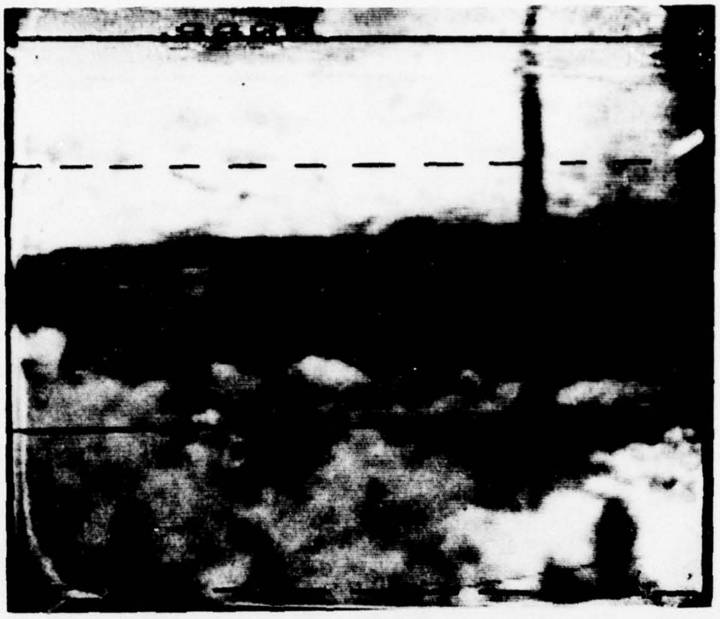


Figure 6. Parking lot high resolution TV narrow field-of-view scene input. (Solid line outlines area of initial digitized input. Dashed line outlines shifted scene input.)

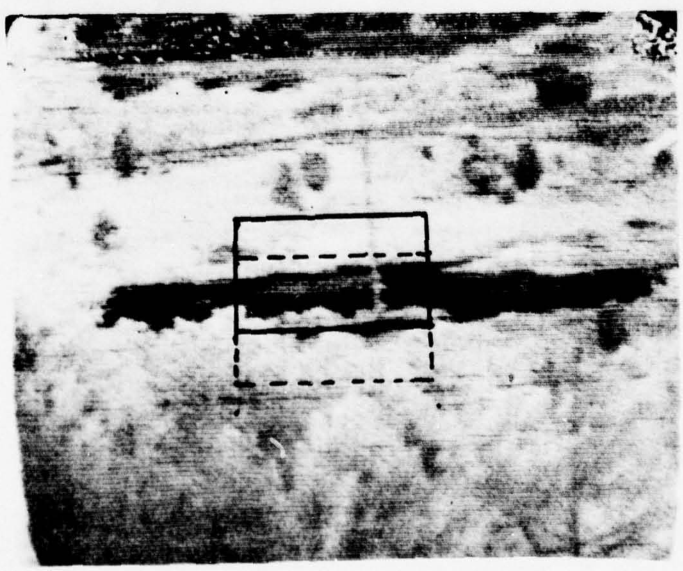


Figure 7. Parking lot high resolution TV wide field-of-view scene. Equivalent to the 4:1 reduction of Figure 6. (Solid line outlines area of effective coverage. Dashed line outlines shifted scene input.)

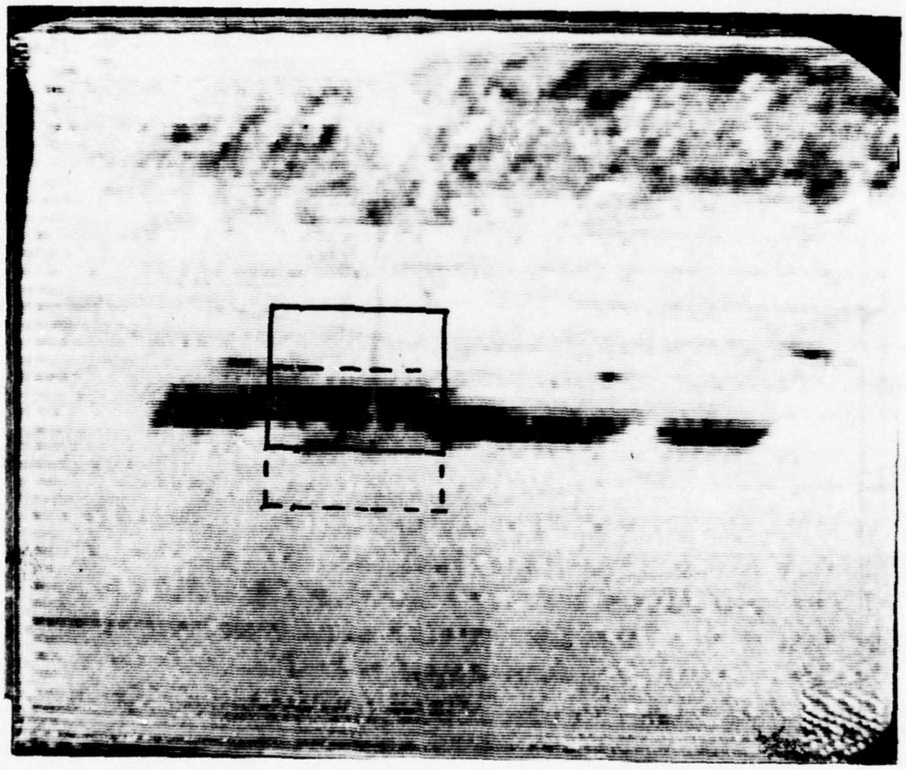


Figure 8. Parking lot low resolution IR input scene. (Solid line outlines area *of initial correlation.* Dashed line outlines shifted scene input.)

contain all zeros. Thus the proper selection of TVGTH was investigated.

Figure 9 is a plot of the number of ones (+) and zeros (-) in the binary matrix of the NASA Tower TV scene as TVGTH is varied from zero to 465, the maximum value in the gradient matrix. Results of the analysis have indicated that when the high resolution TVGTH is selected for an equal number of ones and zeros, the highest correlation peaks were achieved. In *Figure 9* this occurs with a TVGTH of 61.22. It is noted that around the zero/one crossover point significant shifts in the ratio of zeros to ones occur for small changes in threshold. It will be shown later in this report how the correlation sensitivity is influenced by variation in the high resolution sensor threshold.

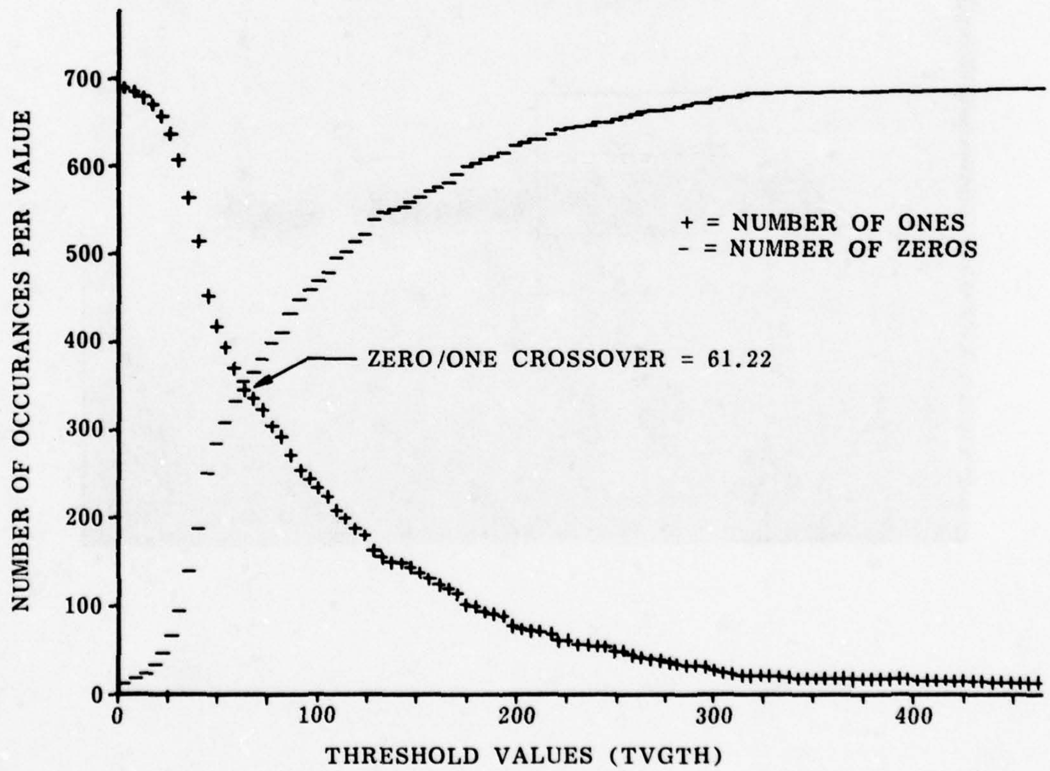


Figure 9. Plot of ones and zeros in the S(I,J) matrix for the NASA tower. (TV)

An S(i,j) gradient matrix array was generated from the low resolution IR digitized scene for the initial assumed image match point. As the analysis continued, it became evident that this initial array was not the correct match point. As with the high resolution matrix, a binary matrix must be established for the low resolution system by the selection of IRGTH. A simulation was performed by setting the high resolution TVGTH at the zero/one crossover point and varying IRGTH for the IR scene to determine the value which maximized the total number of pixel matches for the 26 × 26 array. *Figure 10* is a curve for the NASA tower for the pixel locations where the maximum number of matches occurred. The TV threshold was set at the zero-one crossover value of 61.22. The IRIS threshold at which the maximum number of matches occurred is seen to be 50.5. At this value there were 463 matches out of the possible 676 (or 68% matches). The flatness of the curve indicates the correlation is relatively insensitive to the IRIS threshold within a wide range.

In order to determine a figure of merit for correlation the following criterion was utilized

$$E_{T1}=M_{T1}-(NO_{T2}\ \mathrm{MAX}\ NZ_{T2}) \qquad (6)$$

where

E_{T1}	=	Match point magnitude at threshold IRGHT
M_{T1}	=	Total number of matches at threshold IRGTH
NO_{T2}	=	Number of ones in the high resolution matrix at TVGTH
NZ_{T2}	=	Number of zeros in the high resolution matrix at TVGTH
IRGTH	=	IRIS Threshold
TVGTH	=	TV Threshold

Figure 11 is a plot of E_{T1} for the NASA tower for various values of TVGHT. IRGHT was found to be 50.5. As will be indicated later in the report, the magnitude of E_{T1} = 120 is due to the scene content's having major changes in contrast. As the scene is changed to one where the scenes are less dynamic the value of E_{T1} decreases. However the peak location still indicates the threshold of maximum match. Also, in comparing *Figures 4* and *5* in the dashed outline, it should be noted that due to sensor location the trees have moved, reducing correlation magnitude.

As was presented previously, the relocation of the low resolution pixel array was performed manually in both x and y directions. E_{T1} for TV thresholds of 61.22 (zero/one crossover), 55, and 65 were computed as the low resolution NASA tower scene was shifted in both x and y directions.

Figures 12 and *13* indicate the results of the scene shift on E_{T1} for TV thresholds presented from the maximum match point. If the original low resolution array is used at the initial 0,0 location and the subsequent values of the maximum match are recorded as the low resolution array is displaced in

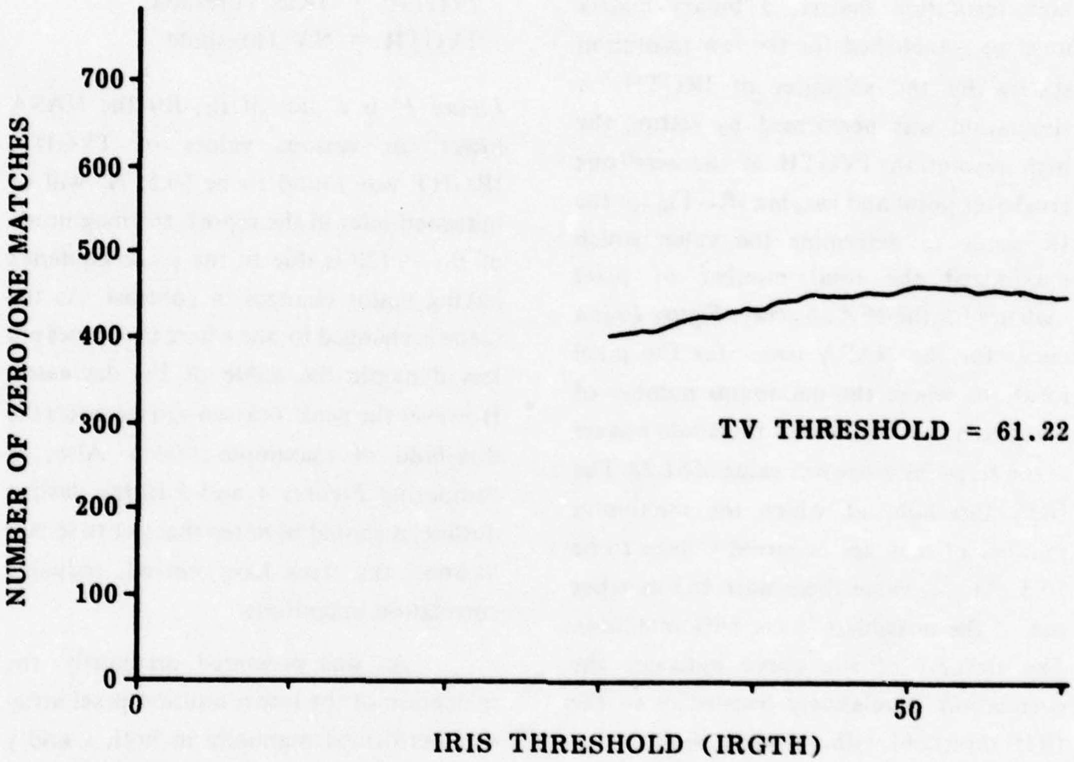


Figure 10. Number of matches of high and low resolution binary matrices for various low resolution threshold values.

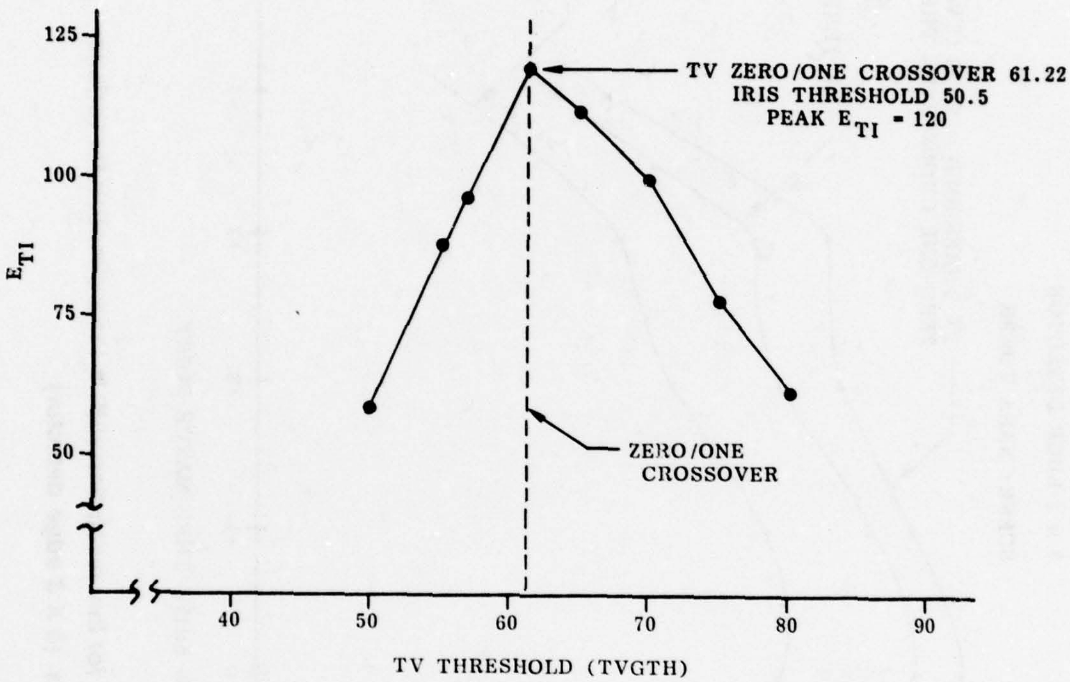


Figure 11. Plot of E_{TI} for various values of TV threshold. 3 X 3 edge detector.

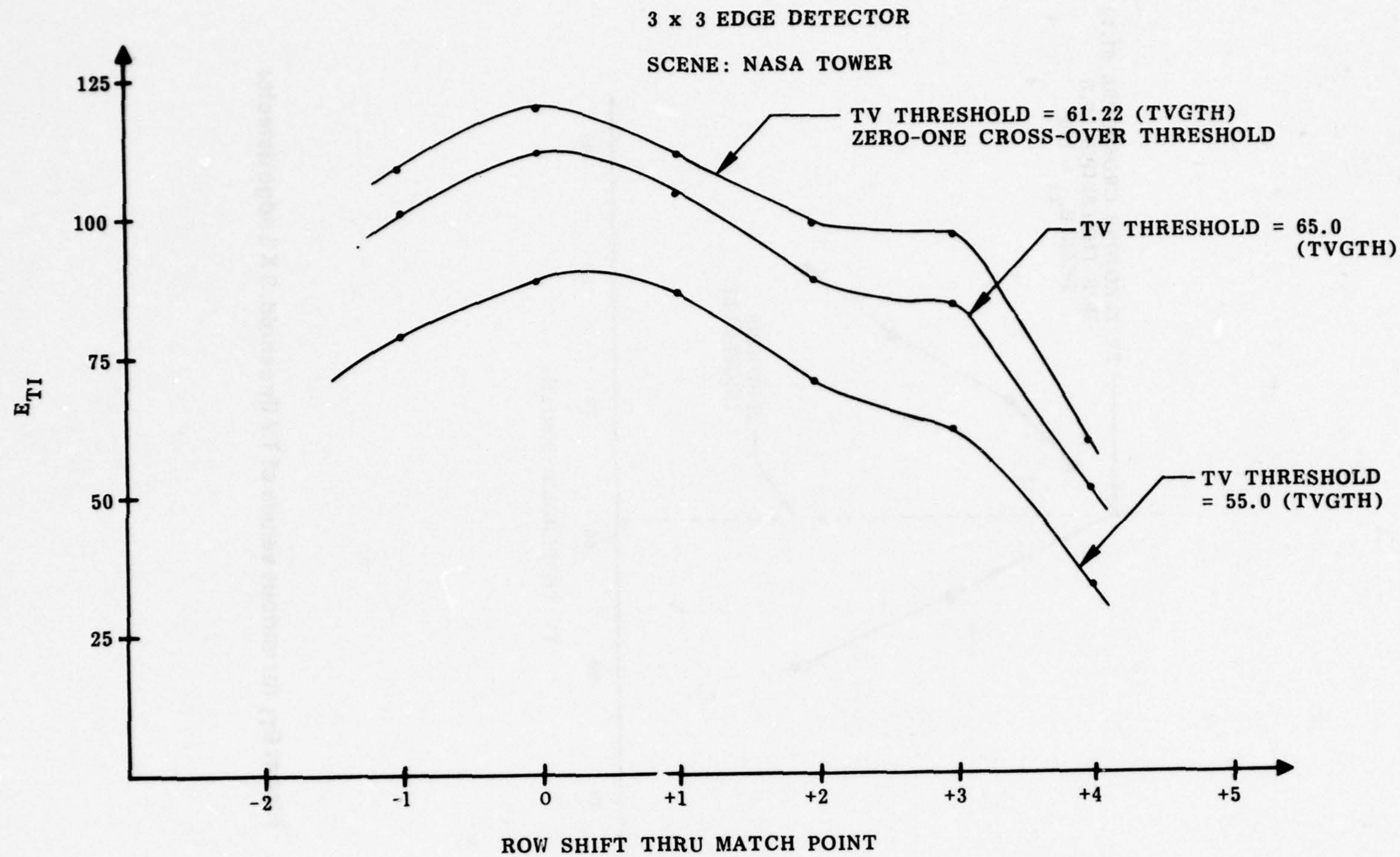


Figure 12. Values of E_{TI} for low resolution shift in the x direction for various TV thresholds. (3 X 3 edge detector)

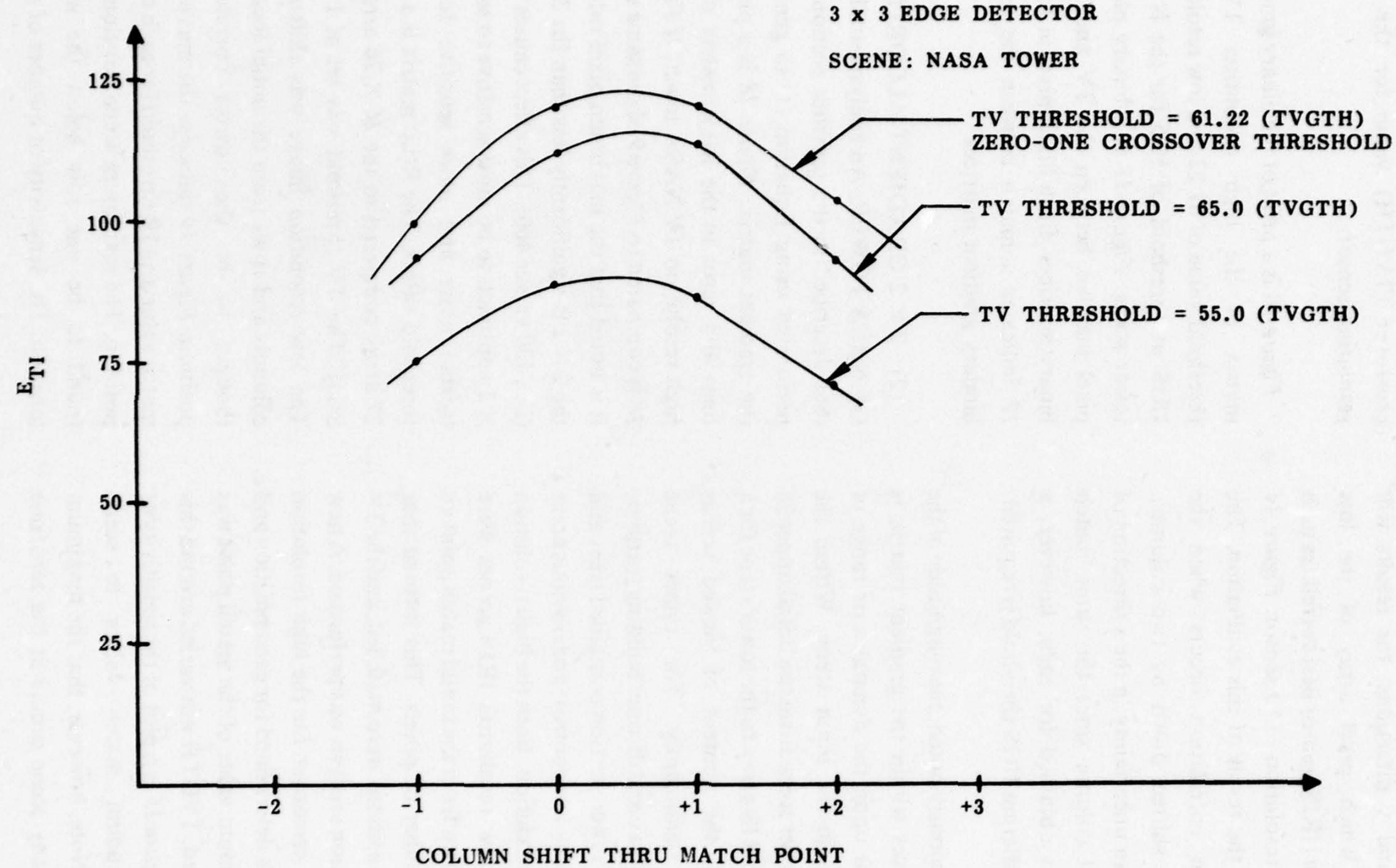


Figure 13. Values of E_{TI} for low resolution shift in the y direction for various various values of TV threshold. (3 X 3 edge detector)

both x and y directions, the results will indicate which pixel array of the low resolution (IRIS) sensor best correlates with the high resolution (TV) sensor. *Figure 14* indicates the result of this evaluation. The maximum correlation occurs when the image is shifted down by two columns. There is an uncertainty in the x direction of one pixel column since the same match value was obtained for each; however, a slightly different IRIS threshold is required.

As previously noted, the magnitude of the pixel valves within the gradient matrix is dependent upon the dynamics or range of contrasts, in the input scene. Within the NASA tower scene from the initial upper 28 × 28 TV/IRIS array to the lower scene for a 28 × 28, the content of viewed scenes differed significantly. The upper scene contained sky and distinct building features, while the lower portion contained trees and considerably less contrast and obvious areas of non-correlation. Both the high resolution TV and low resolution IRIS scenes were shifted down from the initial match point an equal number of pixels. This insured that the new positions were matched, and the 3 × 3 correlation analysis was performed. A new zero/one crossover for the high resolution image was determined for each position and the maximum value of the match point was determined. TVGTH was varied around this value. *Figure 15* is a plot of the sensitivity of the maximum match values to scene content. Note, however, that the maximum value of any scene occurs at the zero/one crossover (TVGTH) value for the high resolution sensor.

Figure 16 is a print of the binary gradient matrix of the high resolution TV at threshold value of 61.22 and low resolution IRIS at threshold of 50.5 for the NASA tower scene. *Figure 17* is a binary plot of pixel matches between the TV and IRIS binary matrices. Each black pixel in *Figure 17* indicates a match between the sensor bindary gradient matrices.

(2) 2 × 2 GRADIENT ALGORITHM OF NASA TOWER. An analysis similar to that described in the previous section was performed using Equation (1) to generate the gradient matrix. *Figure 18* is a plot of ones and zeros in the R(i,j) matrix of the high resolution TV NASA tower. If *Figure 18* is compared to *Figure 9* of the same scene it is noted that the maximum pixel value of the 2 × 2 is significantly less than the 3 × 3, (i.e., 130 versus 460). This effect causes the 2 × 2 approach to be more sensitive to sensor (scene) noise and more sensitive to the threshold values. The R(i,j) matrix is a 27 × 27 array compared to the 26 × 26 array of S(i,j). The TV threshold was set at 11.33. The low resolution image was shifted by columns and rows from the initial location thought to be the correct correlation position. *Figure 19* indicates the maximum match value and IR threshold for each scene position. The maximum scene position was found to be one row below the initial location. The sensitivity of number of pixel

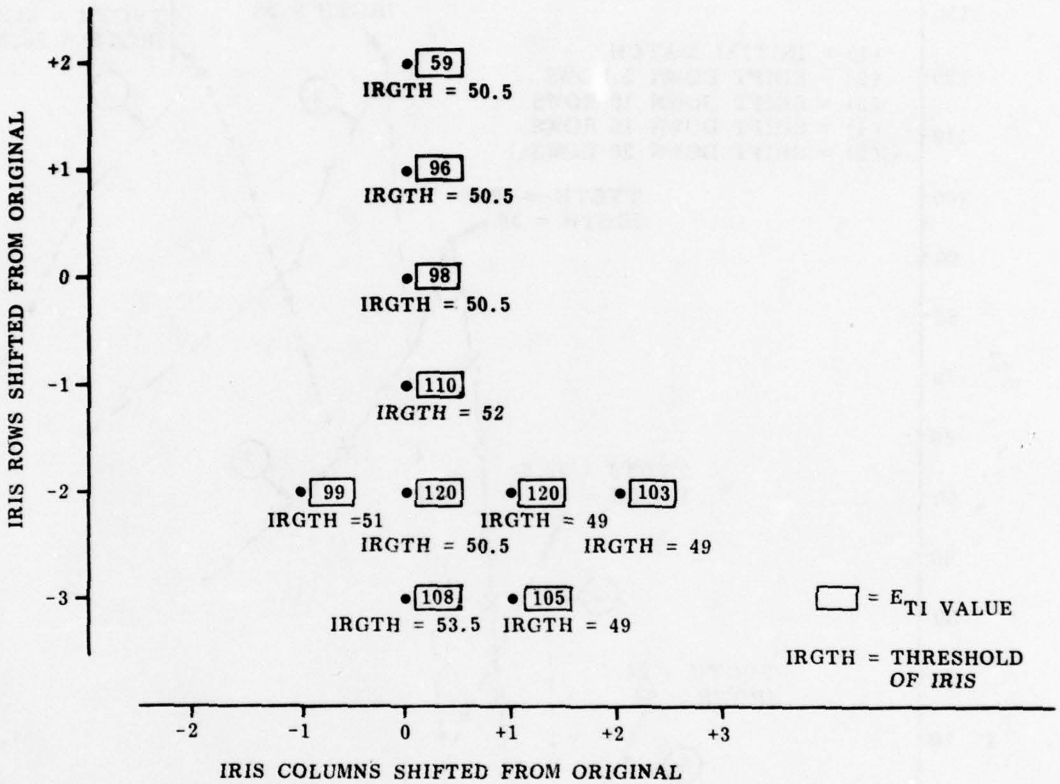


Figure 14. **Values of maximum pixel match 3 X 3 edge detector for various positions of low resolution sensors (NASA tower) (TV threshold = 61.22).**

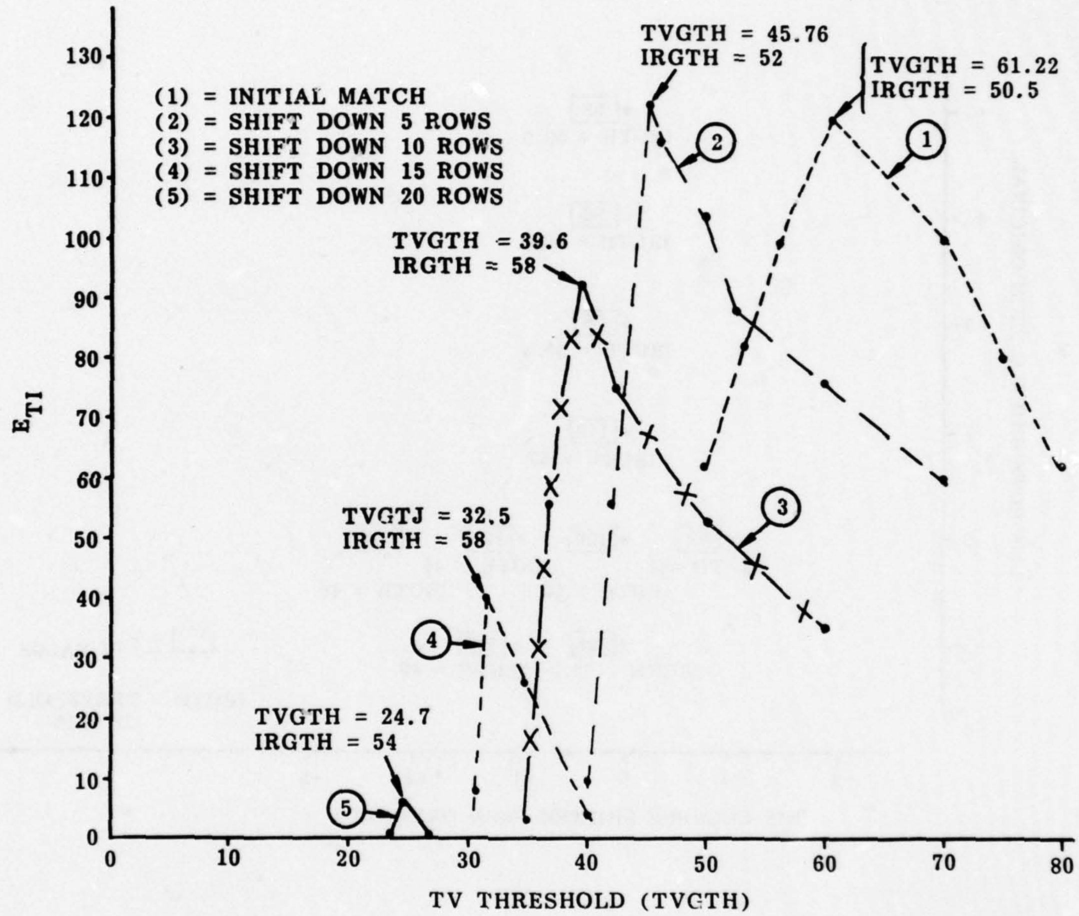
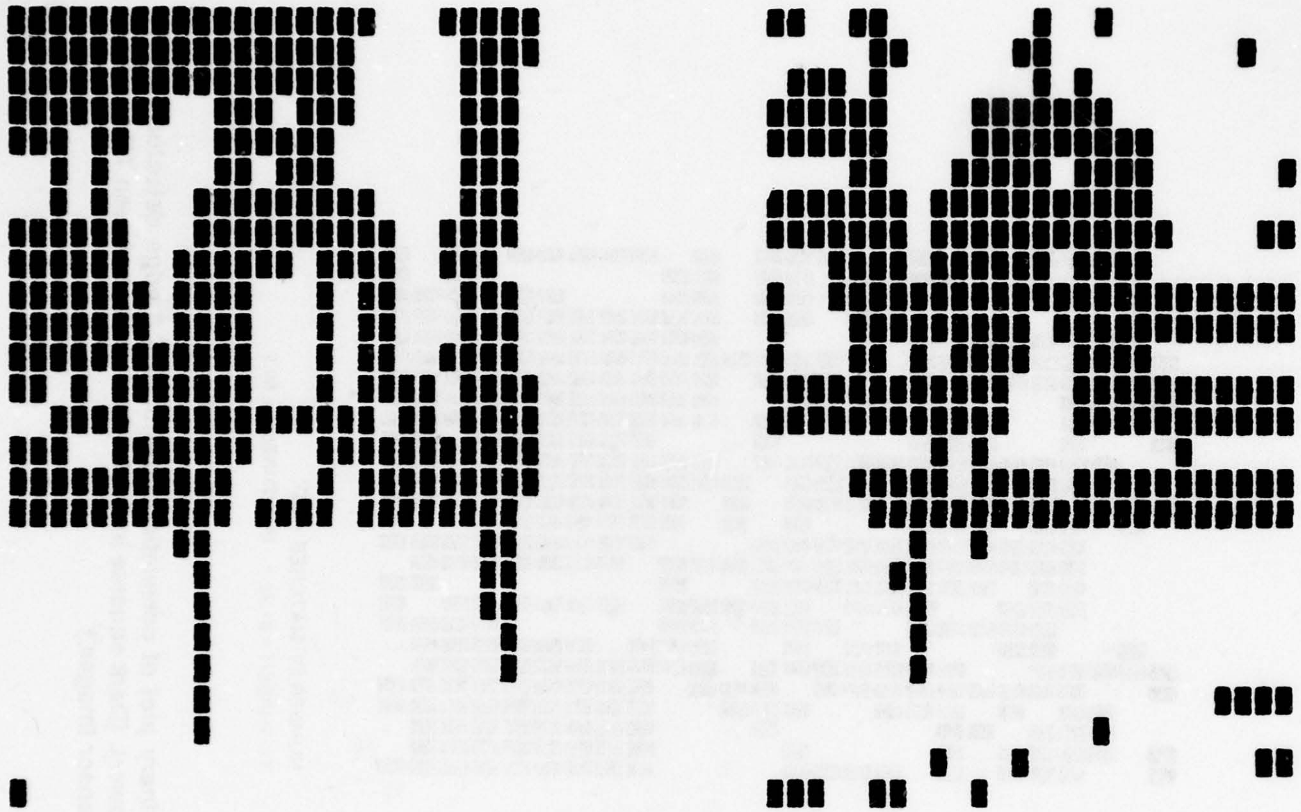


Figure 15. Sensitivity of maximum match values E_{TI} to scene content 3 X 3 edge detector. (NASA tower).

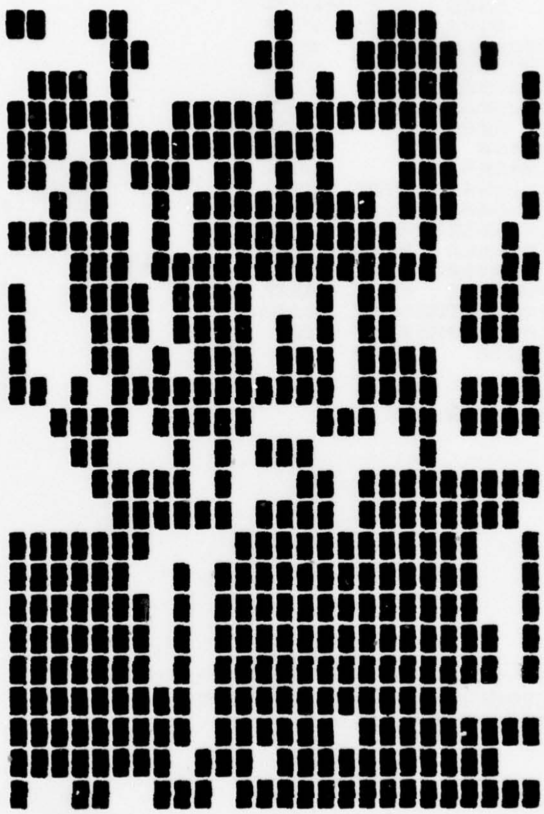


TV MATRIX
THRESHOLD USED = 61.22

■ = PIXEL ABOVE THRESHOLD

IRIS MATRIX
THRESHOLD USED = 50.5

Figure 16. NASA tower 3 X 3 binary matrix at point of scene match.



NUMBER OF MATCHES = 463

TV THRESH = 61.22 IRIS THRESH = 50.5

Figure 17. Binary plot of correlation position 3 X 3 edge detector. (NASA tower). (Dark squares indicate pixel match between TV and IRIS sensor images.)

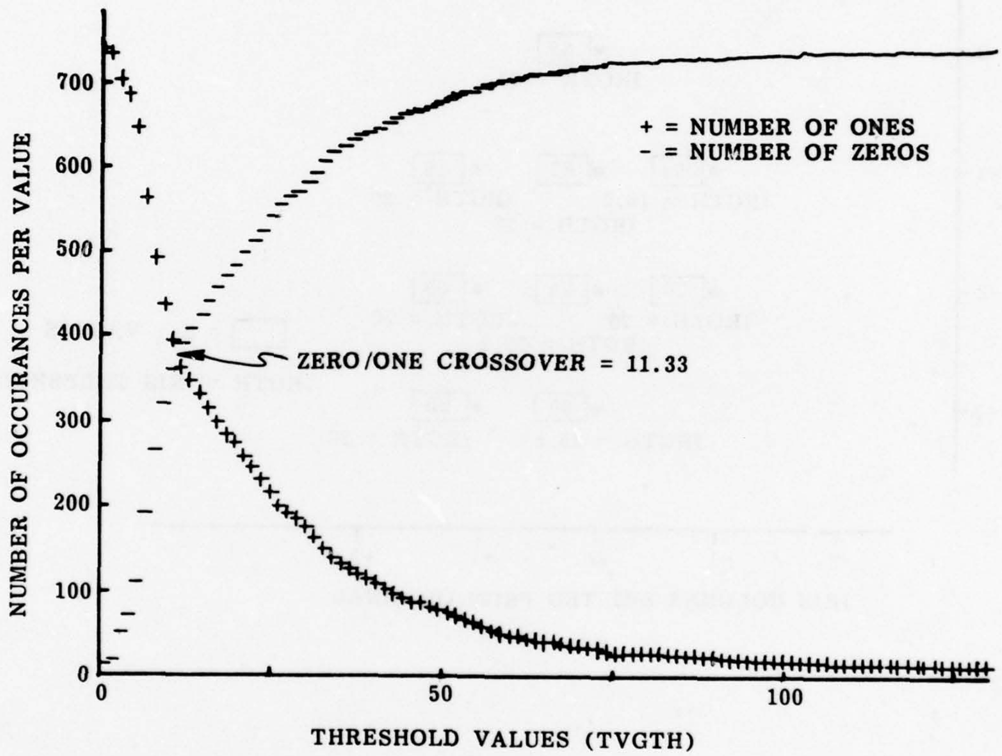


Figure 18. Plot of ones and zeros in the R(i,j) matrix for the NASA tower (TV).

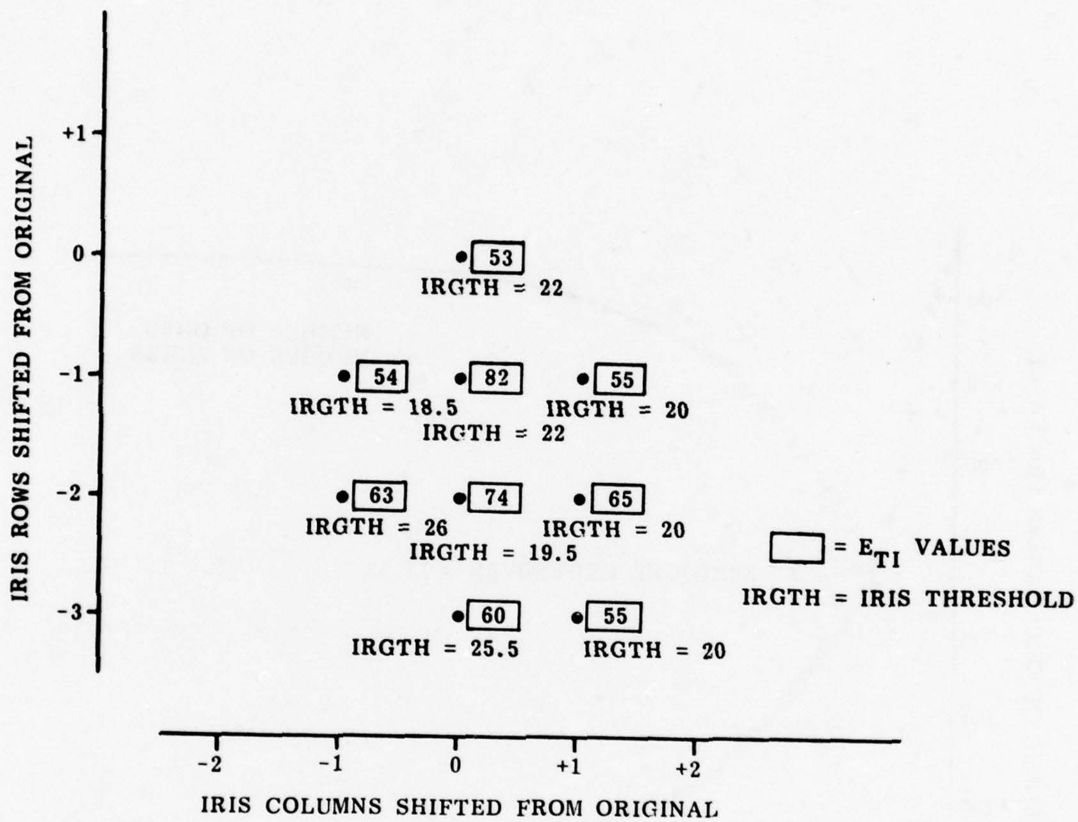


Figure 19. Values of maximum pixel match for various positions of low resolution sensor 2 X 2 edge detector. (NASA tower) (TV threshold = 11.33).

matches versus low resolution (IRIS) threshold was investigated. *Figure 20* indicates the IRIS threshold value which maximizes the number of matches to be 22.

The sensitivity of the 2 × 2 gradient matrix to scene contrast dynamics was investigated similarly to the analysis performed on the 3 × 3 matrix. Both the high (TV) and low (IRIS) resolution scenes were displaced by the same number of rows 11 and 21 from the initial match points, and the TV, IRIS thresholds versus maximum number of pixel match points were determined. *Figure 21* is a plot of the results of this investigation.

Binary matrices were generated for both the TV and IRIS images at their respective thresholds for maximum match (*Figure 22*). *Figure 23* indicates the binary plot of correlation between the images. Of the 729 total matches possible, the maximum of 461 was obtained at TVGTH = 11.33 and IRGTH = 22.

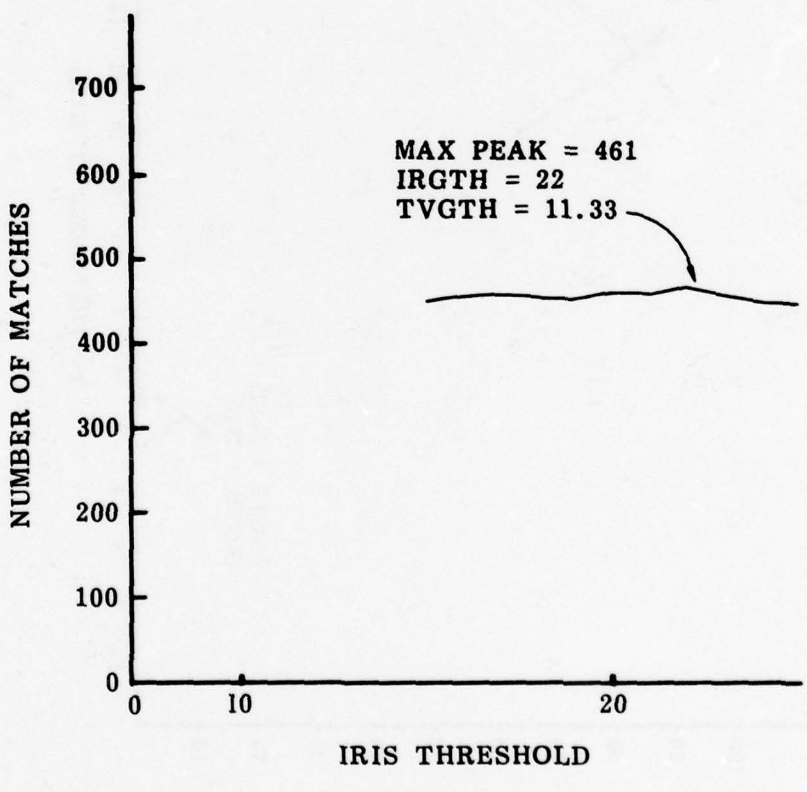


Figure 20. Number of pixel matches of high and low resolution binary matrices for various low resolution threshold values. (2 X 2 edge detector).

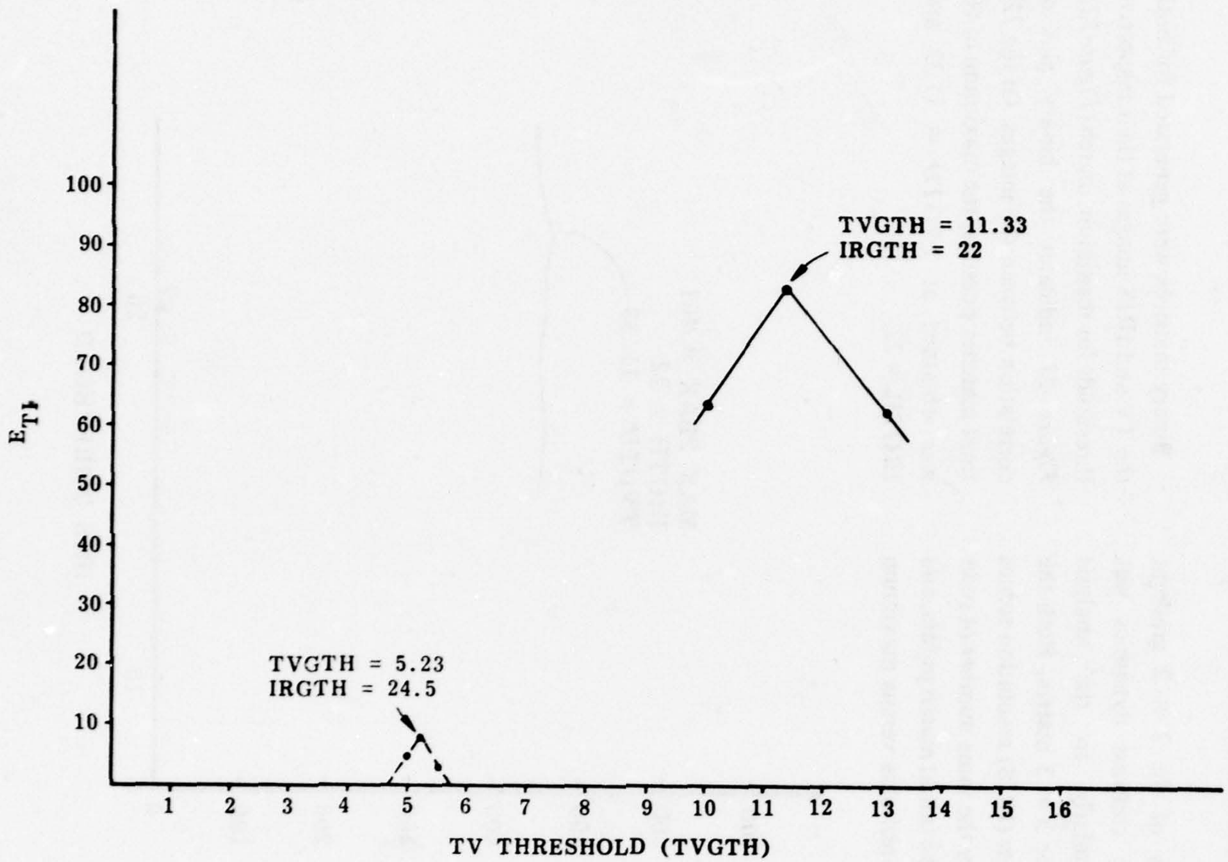
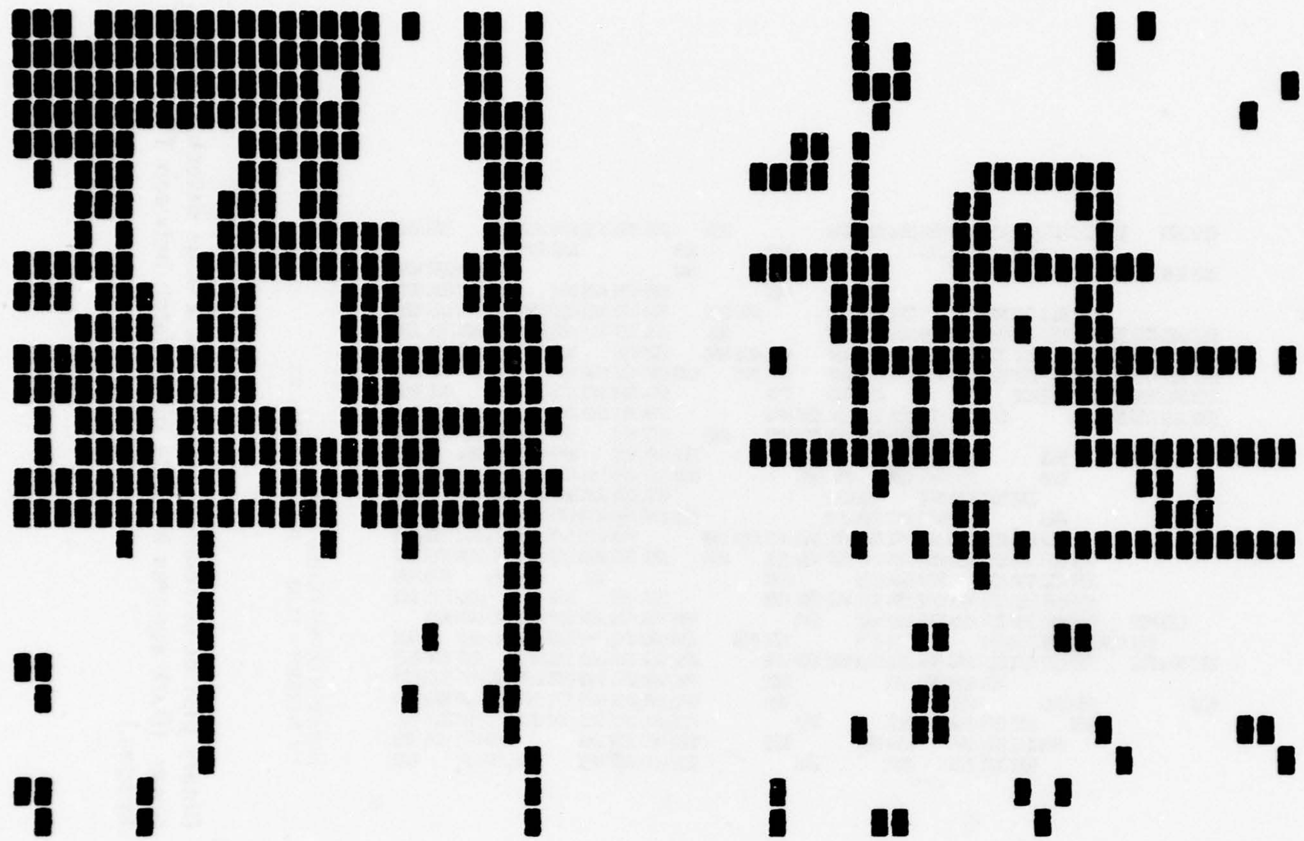


Figure 21. Sensitivity of maximum match values E_{TI} to scene content versus TVGTH 2 X 2 gradient. (NASA tower)



TV MATRIX
THRESHOLD USED = 11.33

IRIS MATRIX
THRESHOLD USED = 22

Figure 22. NASA tower 2 X 2 binary matrix at point of scene match.

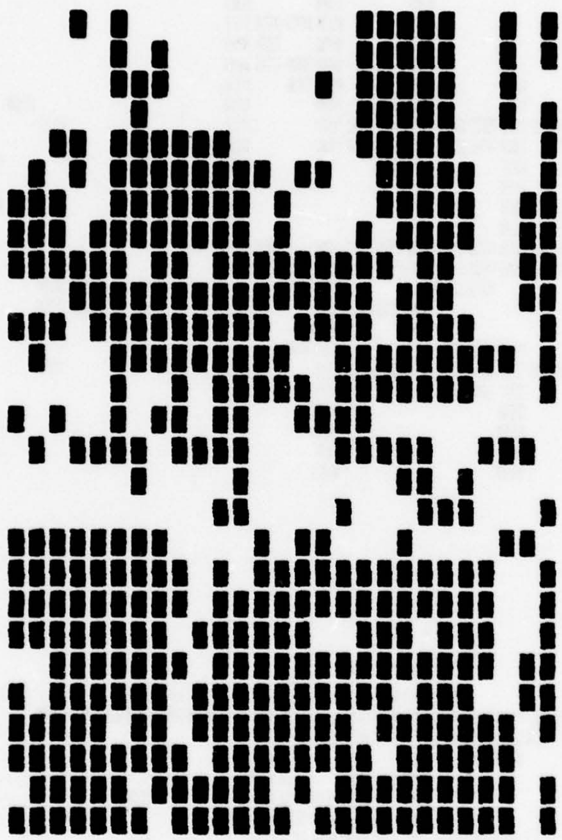


Figure 23. Binary plot of correlation position 2 X 2 edge detector. NASA tower. (Dark squares indicate pixel match between TV and IR images.)

(3) 3 × 3 GRADIENT ALGORITHM OF PARKING LOT. All the analysis results presented thus far in this report have used the NASA tower as the input scene (*Figures 3, 4* and *5*). A similar analysis was performed on a very different type of scene of black asphalt parking lot in a wooded area (*Figures 6, 7* and *8*).

A 28×28 TV high resolution input matrix was established and a plot of the zero/one crossover was established. *Figure 24* indicates the results of this simulation. It should be noted that the zero/one crossover occurs at TVGTH = 143.24 with the maximum single gradient pixel value of 630 compared to 61.22 and 460 respectively for the tower scenes.

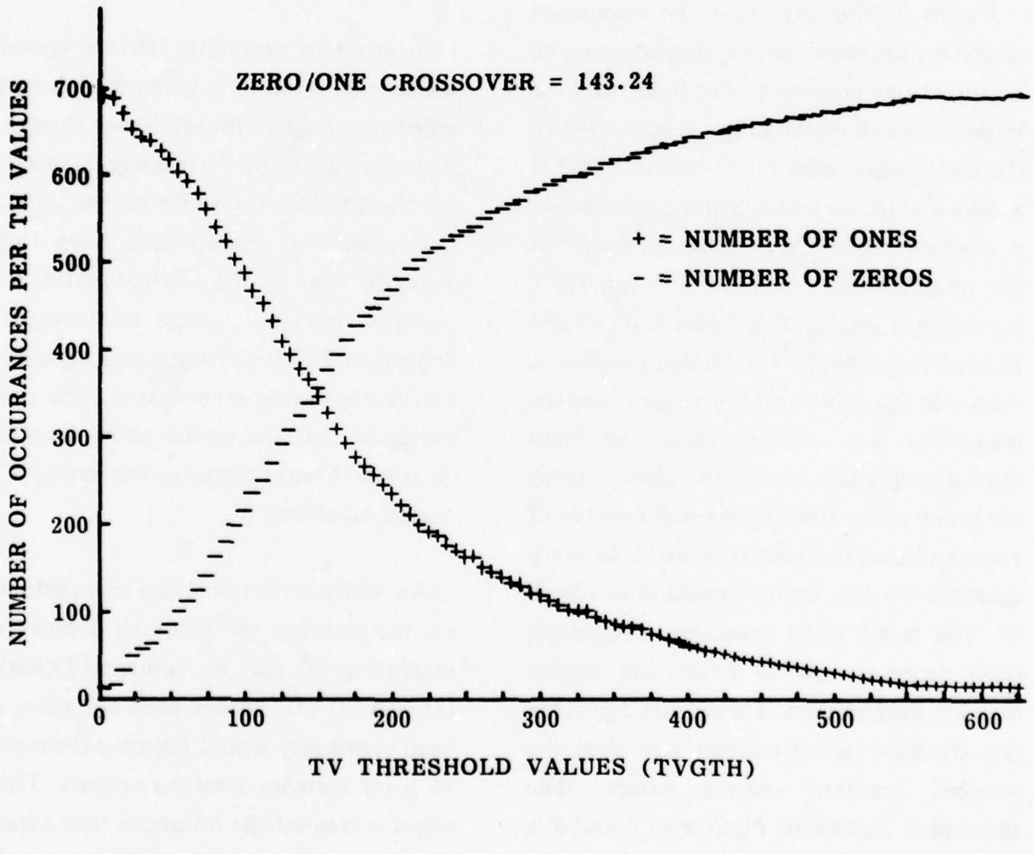


Figure 24. Plot of ones and zeros in the S(i,j) matrix for the parking lot (TV).

The low resolution (IR) scene's position was selected initially by observing the digitized pictures since the digital simulation required manual insertion of low resolution sensor movement with respect to the high resolution scene. For each chosen position, the gradient matrix of the low resolution sensor for various threshold values was correlated against the high resolution gradient matrix and the maximum match (E_{TI}, See Equation (5)) was determined.

Figure 25 indicates that the maximum match occurs when the low resolution scene is shifted one column to the right from the initial assumed match point. A sensitivity of the match point magnitude versus TVGTH at this maximum match point position was performed. *Figure 26* presents the results of the investigation. Both the TV and IRIS input scenes were shifted down 5, 10, 15 and 20 lines respectively. The 20 line position is shown in *Figures 7* and *8*. In every case the maximum E_{TI} occurs when the high resolution gradient matrix threshold is set at the point where there is an equal number of zeros and ones in its binary matrix. In every case the low resolution threshold has been 81. The match point maximum magnitude (E_{TI}) decreases as the scenes are moved down in both sensors. This occurs due to the less dynamic scene content and thus the reduced gradient matrix values. The prominent feature in *Figures 6, 7* and *8* is seen to be the power pole. As the input scenes are moved from the solid outline to the dashed outline, less of this feature exists, so the apparent decrease in E_{TI} is noted. As was noted previously, if E_{TI} is negative, the maximum match point will occur when the low resolution binary matrix is either all zeros or all ones by adjusting the IRGTH. This is clearly a non-correlation position. *Figure 27* indicates the binary matrix for both the high and low resolution sensors at the gradient matrix threshold which provided maximum match point magnitude. *Figure 28* indicates the pixel matches between the two binary matrices. The black pixels indicate agreement.

In any two randomly selected scenes in which a correlation is performed, a certain number of pixels will match even though the scenes are different. To investigate this point for the condition where both the high and low resolution images had been shifted down 20 lines for the original match point, the low resolution image was rotated 90 degrees to the high resolution image and the correlation value investigated. The results indicated that the match point magnitude (E_{TI}) was always negative indicating a "no match condition."

An additional simulation was performed on the parking lot scene to determine if increasing W_2 and W_5 values in Equations (2) and (3) to 4 rather than the value of 2 used previously would improve the number of pixel matches between sensors. This in effect increased the influence that adjacent pixel values have on the establishment of the gradient matrix as related to the diagonal elements. As was expected the values of the

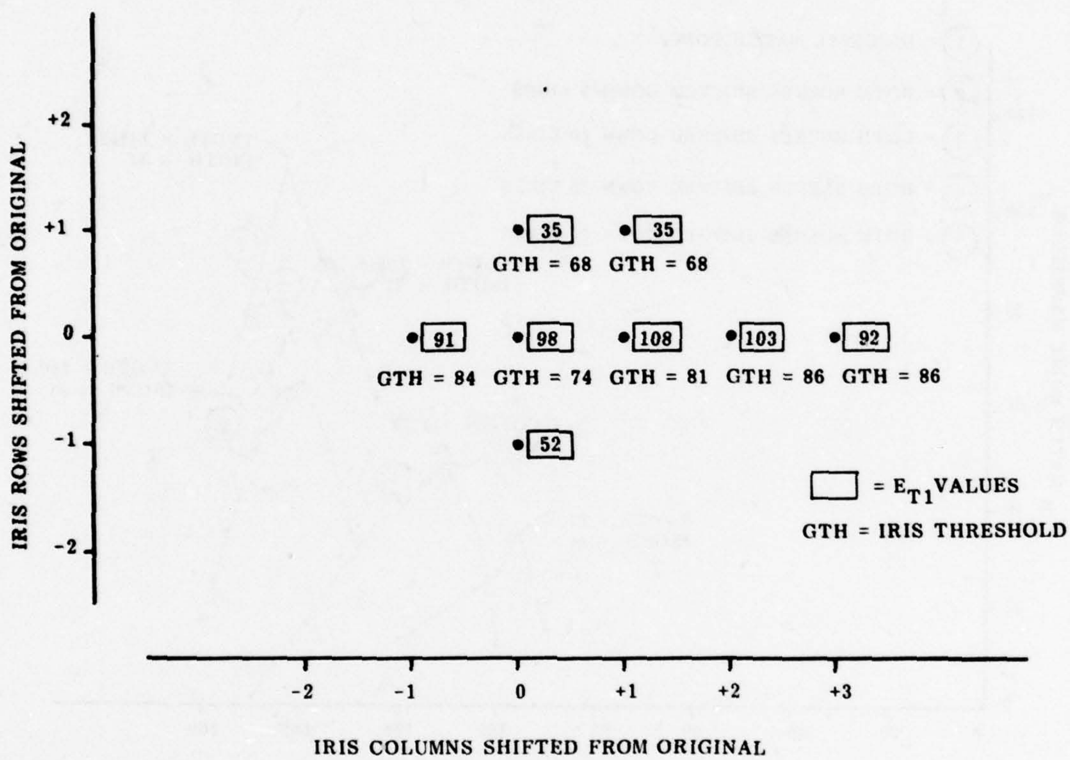


Figure 25. Values of maximum pixel match for various positions of low resolution sensor 3 X 3 edge detection. (Parking lot) (TV threshold = 143.24).

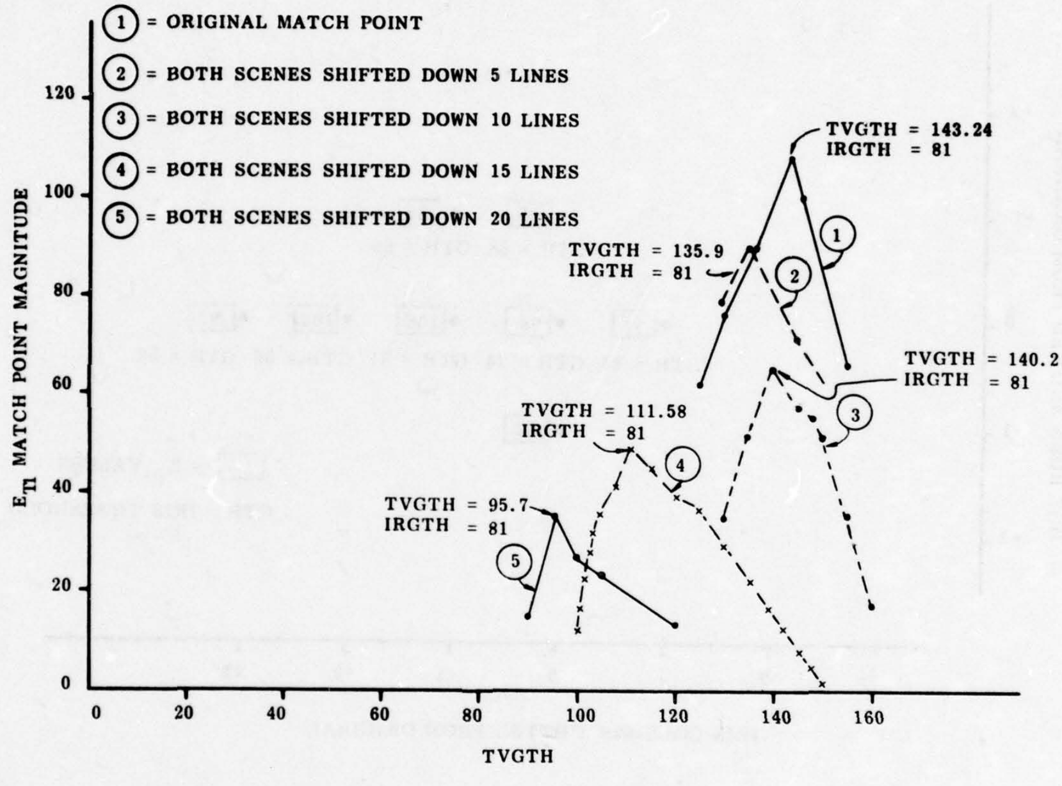
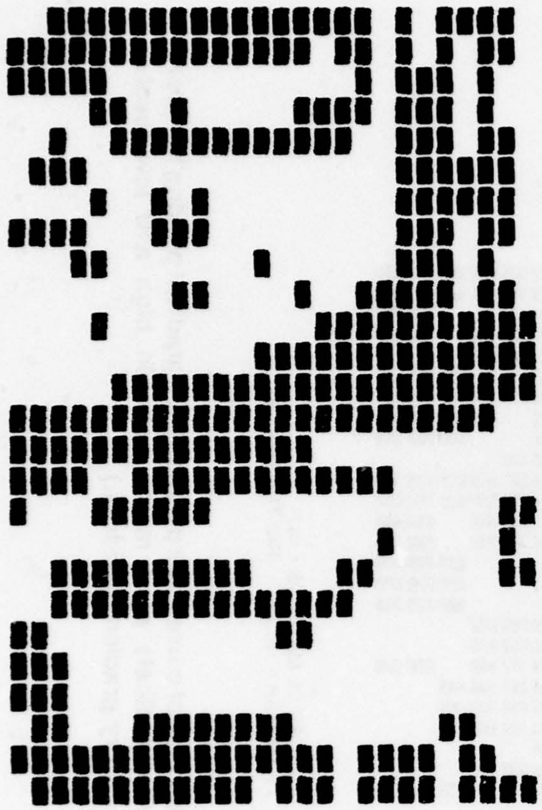
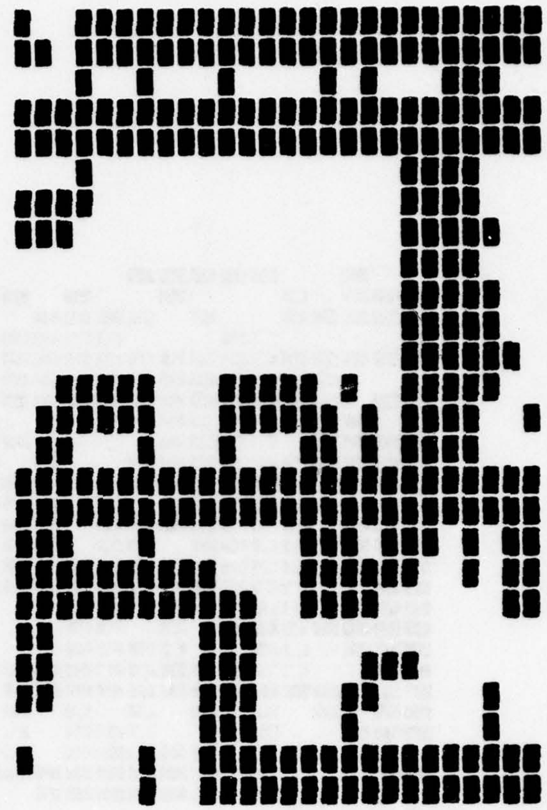


Figure 26. Sensitivity of maximum match values E_{TI} to scene content versus TVGTH 3 X 3 gradient. (Parking lot)

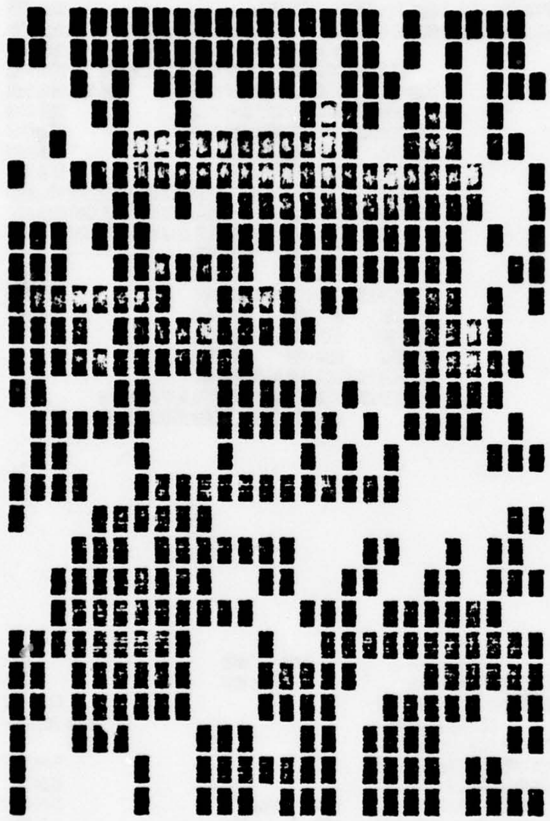


TV MATRIX
THRESHOLD USED = 143.24



IRIS MATRIX
THRESHOLD USED = 81

Figure 27. High and low resolution scenes binary gradient matrix with thresholds set at maximum match. (Parking lot 3 X 3 gradient)



NUMBER OF MATCHES = 447
TV THRESH = 143.24 IRIS THRESH = 81

Figure 28. Binary plot of correlation position 3 X 3 gradient parking lot. (Dark squares indicate pixel match between high and low resolution binary gradient matrices.)

gradient matrix increased. The high resolution gradient matrix threshold for which the zeros and ones of the binary matrix are equal increased from 143 to 220. *Figure 29* indicates the sensitivity of the match point magnitude to a high resolution sensor gradient matrix threshold. By comparing *Figure 29* to Curve 1 of *Figure 26*, it is noted that the sharpness of the peak does not change significantly. Similarly by comparing *Figure 30* to *Figures 27* and *28* to *Figure 31*, it is noted that the actual number of matches decreased by two pixels when the higher multiplier is used.

A similar simulation was performed for the case of W_2 and W_5 values of Equations (2) and (3) being set to 1. The high resolution gradient matrix threshold for which the zeros and ones of the binary matrix were equal was determined to be 104.79. The maximum match point magnitude for these conditions was for the low resolution sensor gradient matrix threshold of 62. *Figures 20f* and *20g* reflect the binary matrix and correlation pixel match for these threshold values. Comparison to *Figures 20d* and *20e* for the case where the multipliers W_2 and W_5 were set at four and *Figures 20a* and *20b* for the case of W_2 and W_5 equal two indicates the maximum number of pixel matches for this parking lot scene was achieved for the gain value of two.

(4) 2 × 2 GRADIENT ALGORITHM OF PARKING LOT. The analysis was repeated for the parking lot scenes using Equation (1) to generate the gradient matrix. The TVGTH value which made the number of zeros and ones of the binary matrix equal was found to be 31.45. As was the case with the previous analysis, an initial high/low resolution sensor scenes match area was selected and with the high resolution gradient threshold set at 31.45, the low resolution gradient threshold was varied and the maximum match value determined. The low resolution scene was then moved by rows and columns to determine which position provided the maximum. *Figure 21* indicates the results of this investigation. In this case, the initially selected positions were correct and any movement in either direction reduced the correlation peak.

The parking lot input scenes to both sensors was moved down 10 and 20 lines respectively as was done using the 3 × 3 gradient algorithm. *Figure 22* is a plot of match point magnitude versus high resolution gradient matrix threshold for the original match position and both sensor scenes moved down 10 and 20 pixel lines, respectively. The solid and dashed lines in *Figures 6* through *8* indicate the zero and 20-line positions.

4. CONCLUSIONS

This preliminary analysis of automatic scene correlation between a TV high resolution sensor (0.5 to 0.85 μ) and IR low resolution sensor (8-14 μ) for two specific scenes (NASA tower and parking lot) is best achieved if the TV gradient matrix threshold

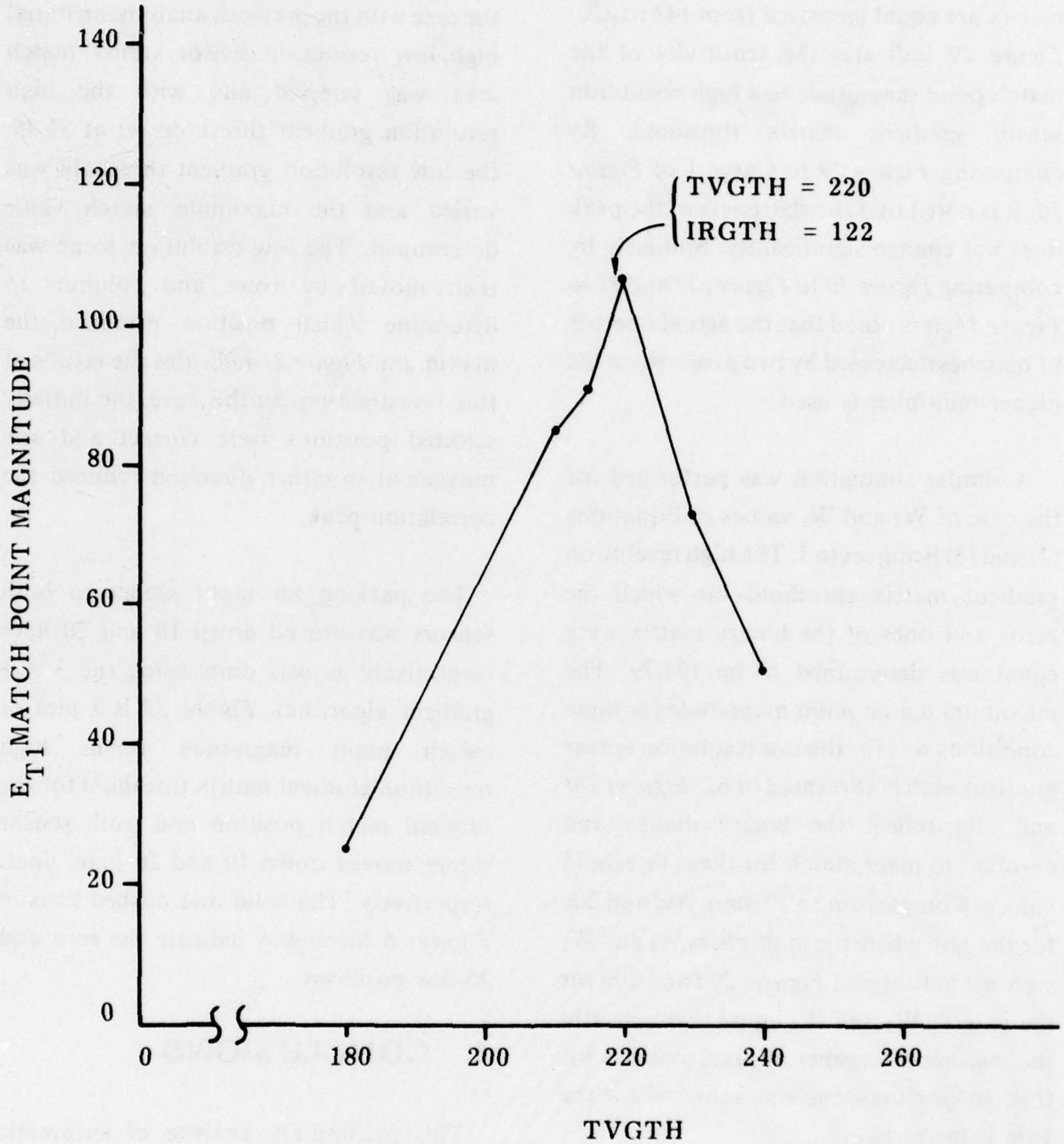
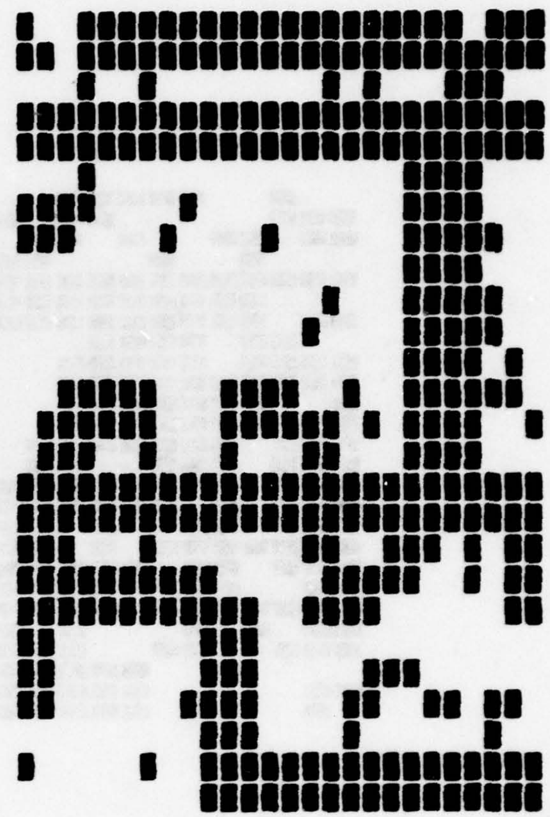


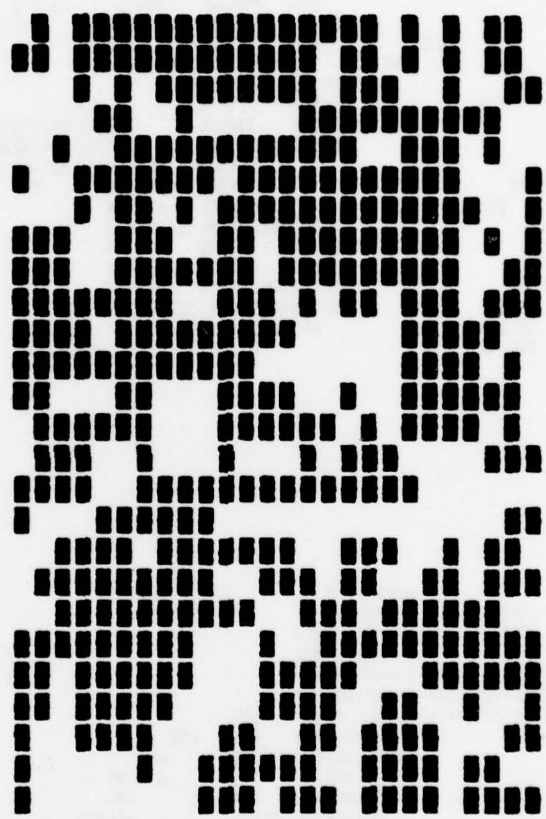
Figure 29. Sensitivity of maximum match values (E_{TI}) to TV gradient threshold for 3 X 3 gradient matrix with coefficient gain of 4. (Parking lot)

TV MATRIX
THRESHOLD USED = 220



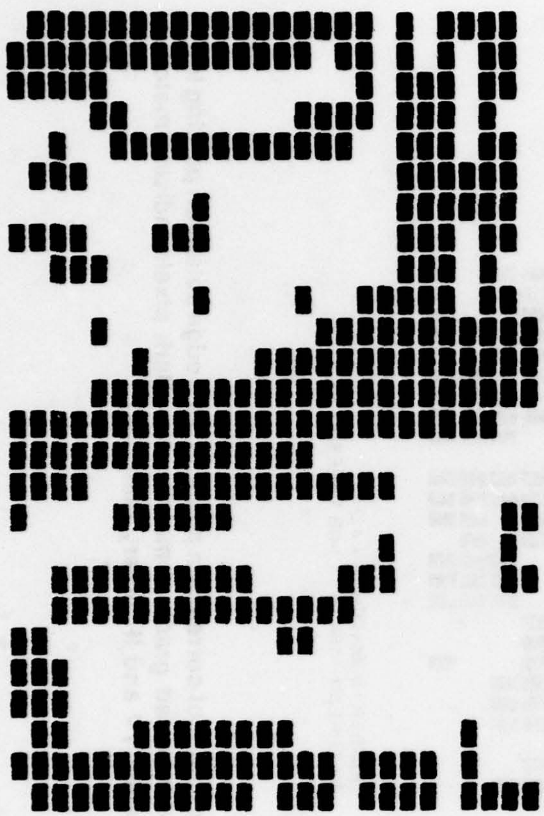
IRIS MATRIX
THRESHOLD USED = 122

Figure 30. **Parking lot 3 X 3 binary matrix at point of scene match with $W_2 = W_5 = 4$ of Equations (2) and (3).**

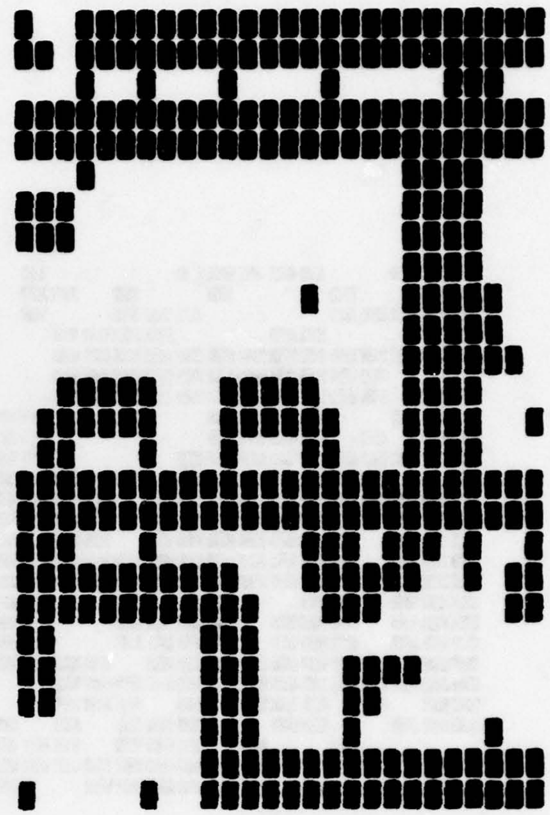


NUMBER OF MATCHES = 445
TV THRESH = 220 IRIS THRESH = 122

Figure 31. Binary plot of correlation position 3 X 3 edge detection parking lot for increased gradient matrix gain. (Dark squares indicate pixel match between TV and IR images.)

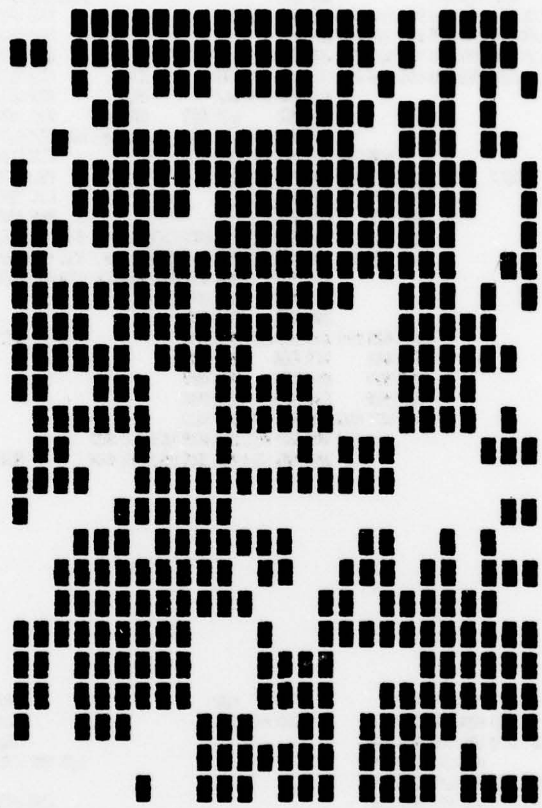


TV MATRIX
THRESHOLD USED = 104.79



IRIS MATRIX
THRESHOLD USED = 62

Figure 32. Parking lot 3 X 3 binary matrix at point of scene match with $W_2 = W_5 = 1$ of Equations (2) and (3).



NUMBER OF MATCHES = 442
TV THRESH = 104.79   IRIS THRESH = 62

Figure 33. Binary plot of correlation position 3 X 3 edge detection parking lot for decreased gradient matrix gain (Dark pixels indicate match between TV and IR binary matrices).

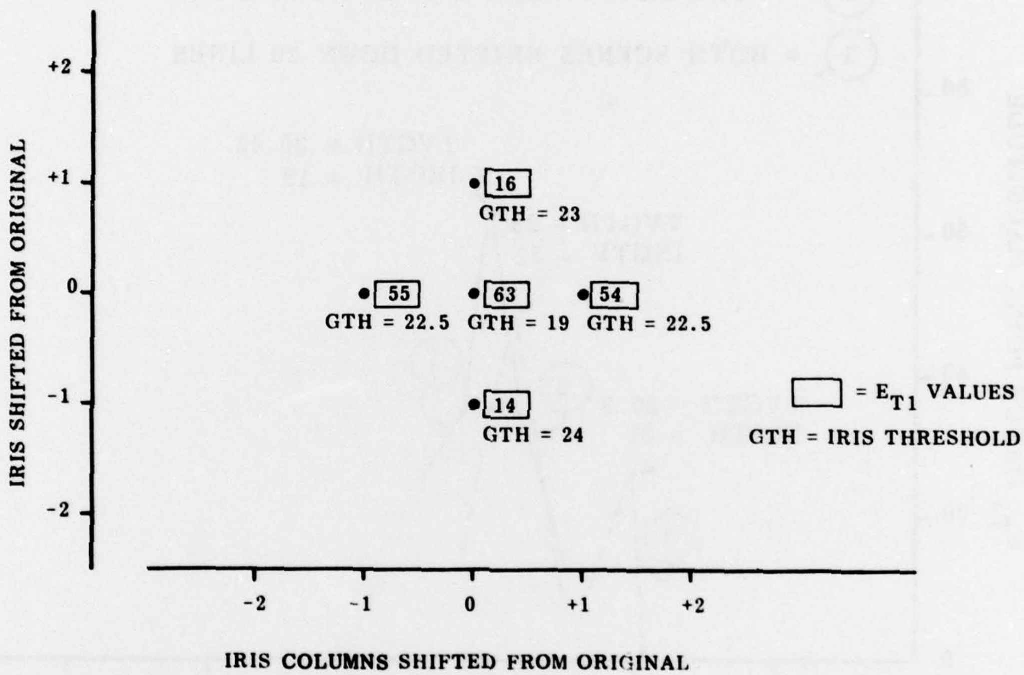


Figure 34. Values of maximum pixel match for various positions of low resolution sensor 2 X 2 edge detection (Parking lot) (TV threshold 31.45).

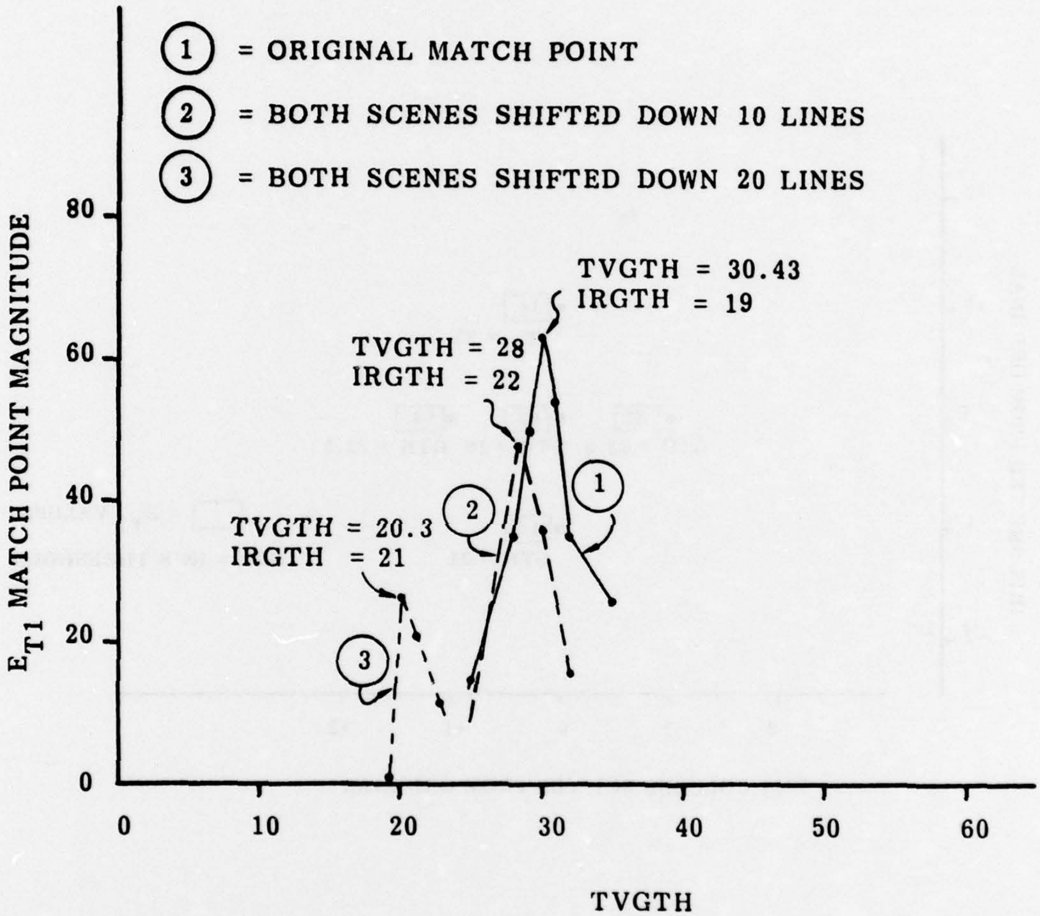


Figure 35. Sensitivity of maximum match values (E_{TI}) to scene content versus TVGTH 2 X 2 gradient (Parking lot).

(TVGTH) is set where the number of zeros and ones of the resultant binary matrix are equal. The 3×3 gradient matrix algorithm appeared less sensitive to noise and threshold values than the 2×2 algorithm. Correct correlation was achieved on both scenes using either algorithm.

The magnitude of the match point was sensitive to scene content. (The more prominent the scene features, the higher the magnitude.) Further, this limited study indicated, at least for the scenes used, that the gain coefficient values of the 3×3 gradient algorithm which produced the maximum correlation were one for the diagonal pixels and two for the adjacent pixels. These values were reflected in the appendix, although an optimal analysis was not performed.

ACKNOWLEDGMENT

The authors wish to acknowledge the technical contribution made by Dr. Jerrel R. Mitchell, in preparation of the Appendix.

APPENDIX

DERIVATION OF THE COEFFICIENTS FOR THE 3×3 GRADIENT ALGORITHM

Assume the digitized image information resides in an $N \times N$ array, g. The goal is to develop an algorithm for computing the gradient of each pixel by using the value of the pixel and its adjacent pixels, assuming a rectangular coordinate system. To be general, the i,j-th pixel of g is selected. *Figure a-1* indicates the pixel being considered, along with its adjacent pixels. The gradient of g at pixel (i,j) can be estimated by using the value of g(i,j) and two adjacent pixels. The rule for selecting the adjacent points is that both cannot lie on the same horizontal, vertical, or diagonal line through g (i,j), e.g., The pixels (i+1, j+1) and (i,j+1) are acceptable; however, (i−1, j+1) and (i+1, j−1) are not. Then, using the eight pixels surrounding (i,j), four acceptable estimates of the gradient of g at (i,j) can be computed.

$S_x^{+} \rightarrow$

$S_y^{+} \uparrow$

(i-1, j-1)	(i-1, j)	(i-1, j+1)
(i, j-1)	(i, j)	(i, j+1)
(i+1, j-1)	(i+1, j)	(i+1, j+1)

Figure A-1. A 3 X 3 typical pixel array.

As stated previously, the image gradient is a function of two variables, i.e.,

$$G = g\,(x,y) \tag{A-1}$$

from calculus

$$\Delta G \cong S_X\, \Delta X + S_Y\, \Delta Y \tag{A-2}$$

where ΔG is the change in the digitized image value for coordinate changes ΔX and ΔY; S_X and S_Y are respectively the partials of g (x,y) w.r.t. x and y (evaluated at the particular x and y coordinate).

For simplicity, $\Delta X = \Delta Y = 1$.

Using Equation (A-2) and the values corresponding to pixels (i−1, j+1), (i−1, j−1), and (i,j), the results are:

$$g(i-1, j+1) - g(i,j) \cong S_X + S_Y \tag{A-3}$$

and

$$g(i,j) - g(i-1, j-1) \cong S_X - S_Y. \tag{A-4}$$

Solving Equations (A-3) and (A-4) simultaneously gives

$$S_X \cong 1/2\,[g(i-1, j+1) - g(i-1, j-1)] \tag{A-5}$$

$$S_Y \cong 1/2\,[+g(i-1, j+1) - 2g(i,j) + g(i-1, j-1)]. \tag{A-6}$$

In a similar manner the pixels (i+1, j+1), (i+1, j−1), and (i,j) yield

$$g(i+1, j+1) - g(i,j) \cong S_X - S_Y \tag{A-7}$$

$$g(i,j) - g(i+1, j-1) \cong S_X + S_Y. \tag{A-8}$$

Solving Equations A-7 and A-8 yields

$$S_X \cong 1/2\,[g(i+1, j+1) - g(i+1, j-1)] \tag{A-9}$$

$$S_Y \cong 1/2\,[-g(i+1, j+1) + 2g(i,j) - g(i+1, j-1)]. \tag{A-10}$$

Using pixels (i,j+1), (i+1,j) and (i,j)

$$g(i, j+1) - g(i,j) \cong S_X \tag{A-11}$$

$$g(i,j) - g(i+1, j) \cong S_Y. \tag{A-12}$$

Likewise, using pixels (i−1, j) (i, j−1), and (i,j)

$$S_X \cong g\,(i,j) - g\,(i,\,j-1) \tag{A-13}$$

$$S_Y \cong g\,(i-1,\,j) - g\,(i,j) \tag{A-14}$$

The group of Equations 5A, 9A, 11A and 13A for S_X and 6A, 10A, 12A and 14A for S_Y. It is logical to average these to obtain an average estimate for the values.

$$\begin{aligned} S_X \cong 1/8 \Big|&[g\,(i-1,\,j+1) + 2g\,(i,\,j+1) + g\,(i+1,\,j+1)] \\ &- [g\,(i-1,\,j-1) + 2g\,(i,\,j-1) + g\,(i+1,\,j-1)]\Big| \end{aligned} \tag{A-15}$$

$$\begin{aligned} S_Y \cong 1/8 \Big|&[g\,(i-1,\,j-1) + 2g\,(i-1,\,j) + g\,(i-1,\,j+1)] \\ &- [g\,(i+1,\,j+1) + 2g\,(i+1,\,j) + g\,(i+1,\,j-1)]\Big| \end{aligned} \tag{A-16}$$

If Equations (2) and (3) of the main report are compared to *Equations* (A-15) *and* (A-16) then $W_1 = W_3 = W_4 = W_6 = 1$ and $W_2 = W_5 = 2$. Equations (A-15) and (A-16) have a multiplier of 1/8, which would reduce the value of S (i,j) of Equation (4) by 5.66. However, since it affects all gradient matrix values, the results will be unchanged.

The above derivation utilized four estimates of the gradient from the center pixel. There are 24 possible gradient estimates. It was found that if all were used in similar computation, the results for the 3 × 3 general array were the same as Equations (A-15) and (A-16) except that the overall multiplier changes, which does not affect the relative weight between pixels for S_X and S_Y computations.

REFERENCES

1. McIngvale, P. H., and Pitruzzello, M. C., *Preliminary Experimental Investigation of Target Hand-Off From a Target Acquisition System to an Imaging Missile Seeker*, Report RG-77-10, US MIRADCOM, July 1977.

2. McIngvale, P. H., and Pitruzzello, M. C., *Advanced Target Acquisition and Hand-Off Experiments*, Report T-78-1, US MIRADCOM, October 1977.

3. Duda, R. O., and Hart, P. E., *Pattern Classification and Scene Analysis*, New York: John Wiley and Sons, 1973, pp. 268-272.

DISTRIBUTION

	No. of Copies
Defense Documentation Center Cameron Station Alexandria, Virginia 22314	12
Project Manager Advanced Attack Helicopter	
ATTN: DRCPM-AAH, COL E. Browne	1
Mr. R. Hubbard	1
Mr. D. Reago	1
P. O. Box 209 St. Louis, Missouri 63166	
Project Manager TADS/PNVS	
ATTN: DRCPM-AAH-TP, COL C. Patnode	1
Mr. H. Khiel	1
Mr. G. Smith	1
Mr. D. Checkwick	1
P. O. Box 209 St. Louis, Missouri 63166	
Commander US Army Aviation Research and Development Command	
ATTN: DRDAV-EV, Dr. G. Marner	1
-EQ, Mr. C. Crawford	1
Mr. M. Jackson	1
P. O. Box 209 St. Louis, Missouri 63166	

DRCPM-HF,	COL R. Feist	1
	Mr. N. Mangus	1
	Mr. M. Carroll	1
	Mr. J. Liston	1
-LD,	Mr. W. Harris	1
-RK,	Mr. R. Bergman	1
-TO,	Mr. Ross Ward	1
DRXHE-MI,	Mr. D. Nichols	1
DRDMI-T,	Dr. Kobler	1
	Mr. Leonard	1
DRDMI-TG,	Mr. Huff	1
DRDMI-TGC,	Mr. Griffith	40
DRDMI-TGN,	Mr. Gambill	1
DRDMI-TGT,	Mr. Leonard	2
DRDMI-TGL,	Mr. Stripling	1
DRDMI-TGG,	Mr. Ciliax	1
DRDMI-ET,	Mr. Black	1
	Mr. Wahlheim	1
	Mr. Farris	1
	Mr. Priddy	1
DRDMI-TE,	Mr. Lindberg	1
DRDMI-TEO,	Mr. Bates	1
DRDMI-TL,	Mr. Lewis	1
DRDMI-TD,	Dr. McCorkle	1
DRDMI-TBD		3
DRDMI-TI,	(Record Set)	1
	(Reference Copy)	1

☆U.S. GOVERNMENT PRINTING OFFICE : 1979-640-122/369