IMAGE REGISTRATION USING GENERALIZED HOUGH TRANSFORMS

Simon Yam
Larry S. Davis

Computer Sciences Department
University of Texas
Austin, TX 78712

December 1980 TR 80-7

This research was supported in part by funds derived from the National Science Foundation under grant ENG-7904037.

Abstract

This paper discusses the application of Hough transform techniques to image registration. It includes a discussion of an interest operator for selecting distinctive features in any image, a generalized Hough-like transform operator for tracking arbitrary shaped/oriented objects across frames, and a fast and reliable technique - hierarchical Hough transforms (HHT) - for choosing a consistent set of tracked object locations. An example is given showing the registration of five aerial photographs.

1. Introduction

An important problem in dynamic scene analysis is registering time-varying images from frame to frame. A variety of factors contributes to making this problem difficult. They include choosing interesting features from each frame and tracking such features from frame to frame in the presence of noise and distortion.

This paper describes an image registration system which has been constructed utilizing several recently developed techniques. In this section we outline the overall organization of the system. The input to the system is a sequence of edge pictures, $f_0, f_1, \ldots f_k, \ldots$, which are produced by a simplified version of the "super slice" image segmentation algorithm [1] as described in Section 2. Section 3 of this paper gives a detailed description of the interest operator used to acquire new interesting features from new frames. Secion 4 discusses the generalized Hough transform tracking algorithm. Section 5 introduces the hierarchical Hough transform technique for determining the consistency of the tracked feature locations. Section 6 contains an experimental study which illustrates the effectiveness of the system.

It should be noted that the efficiency and the accuracy of the matching procedure can be increased significantly by exploiting additional facts about the problem. For example, if an object has been successfully tracked from one frame to the next, one can use its velocity vector to predict its location in the next frame. Efficiency can be increased by searching for matches only in the vicinity of the expected locations; accuracy can be increased by including positional errors in computation of the matching criterion function.

Although the details of the program occasionally get rather involved, the overall strategy is straightforward. The basic operations performed are summarized in figure 1. Because the program begins with no knowledge of the motions, the initial procedure for matching the first two frames is somewhat different from the strategy for matching subsequent frames. This latter strategy is the more significant of the two. On occasion, however, the matching procedure will lose track of an object, ordinarily because a major part of the object is occluded or the object has completely disappeared. When this happens, the consistency algorithm (HHT) will determine the correct location of the object or temporarily exclude that object from the set defining the registration.

The strategy employed by the system is obviously heuristic. We assume that a majority of the objects will move uniformly and maintain their relative positions across frames. This assumption is valid only if the interval between successive frames is sufficiently short; as the interval becomes longer, the objects will move more independently and rapidly so that their relative positions will not be maintained across frames. In this case, more sophisticated consistency and/or tracking algorithms must be used.

Initialization

(1) Read first two frames, f_0 and f_1.

(2) Locate control points, $C_1, C_2 .., C_n$ in f_0 by applying the interest operator.

(3) Register control points with f_1. The generalized Hough transform (GHT) operator is applied to each control point, C_i, and a set of hypothetical locations, $H_{i1}, H_{i2}, .., H_{im}$, for C_i is located in f_1.

(4) The hierarchical Hough transform (HHT) consistency algorithm is applied to determine the consistent location of each control point, C_i in f_1, by choosing one (possibly none) of C_i's hypothetical locations.

(5) The velocity vectors of the control points from f_0 to f_1 are computed.

Main loop

(6) If the number of detected control points is less than a predefined threshold, then additional control points are selected from the "interest field" of the current frame. The control points selected are predicted (by the velocity vectors) to be within the next picture frame.

(7) All the detected control points that are predicted to be outside the next picture frame are removed, and additional control points are selected. Again, they must be predicted to be within the next frame.

Figure 1. Basic system operations

(8) Read next frame.

(9) Register control points using GHT operator.

(10) Apply HHT to select the consistent set of detected control points.

(11) Compute velocity vectors.

(12) Repeat

Figure 1. Basic system operations (continued)

2. Edge Detection

The edge detection program is based on a simplified version of the "super slice" algorithm presented in [1] where a single threshold is chosen for the entire image, rather than on an object by object basis. It consists of three parts:

(1) Thresholding

A range of k thresholds, t_1 to t_k, is chosen for thresholding. The grey scale picture, P, is thresholded to produce k binary pictures. Boundaries, transitions from 0 to 1, or 1 to 0, are located in the pictures to produce k boundary pictures, b_1 to b_k. Note that a reasonable choice for thresholds is from the interval, (m-s, m+s), where m and s are the mean and standard deviation of the grey scale picture, P, respectively.

(2) Zero-crossings

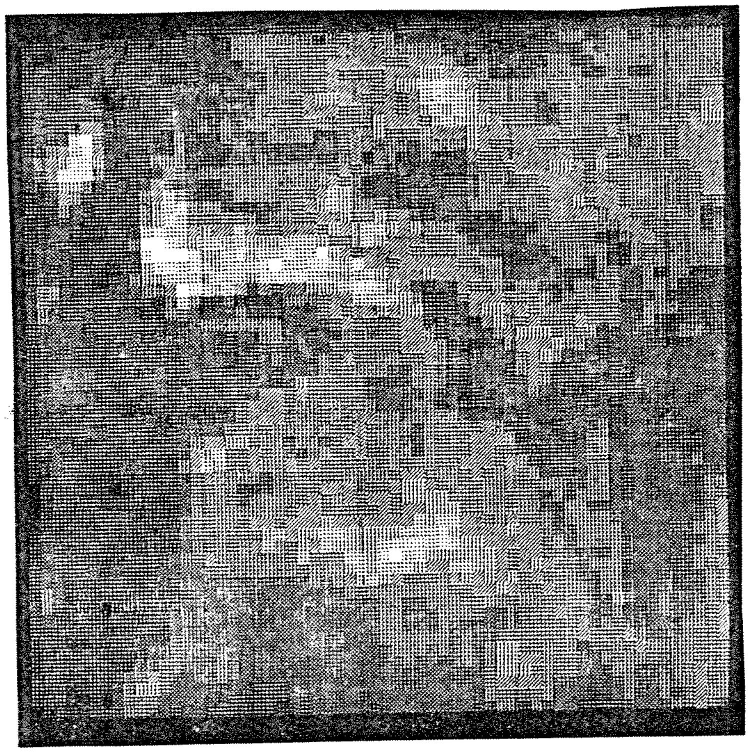
A Laplacian operator is applied to the grey scale picture, P. The zero-crossings of P, transitions from positive to negative or vice versa, are recorded to produce an edge picture, E. Marr and Hildreth [2] present a detailed discussion of edge detection based on zero-crossings.

(3) Exclusive-or

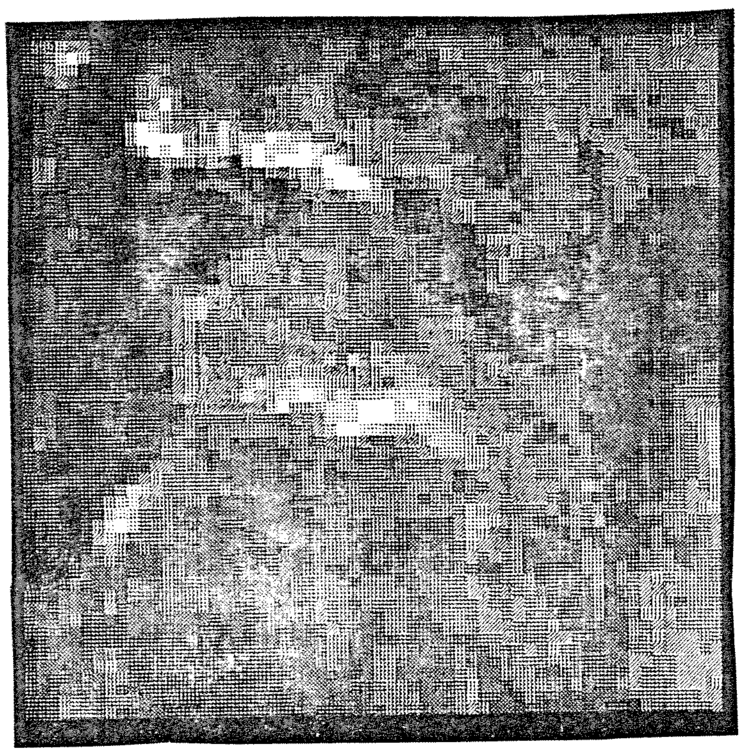
The edge picture, E, is superimposed on each of the boundary pictures, b_1 to b_k. An accumulator is set up for each boundary picture to count the number of times that there is a 1 in the edge picture but a 0 in the boundary picture, or vice versa. The boundary picture with the minimum value in its accumulator is selected as the optimal edge picture.

Thresholding by itself does not work, because it is hard to choose the right threshold. A combination of zero-crossings and thresholding allows us to obtain a good thresholded image.

Figure 2 shows five aerial photographs. Figure 3 shows the result of applying the edge detector to each of the five aerial photographs.

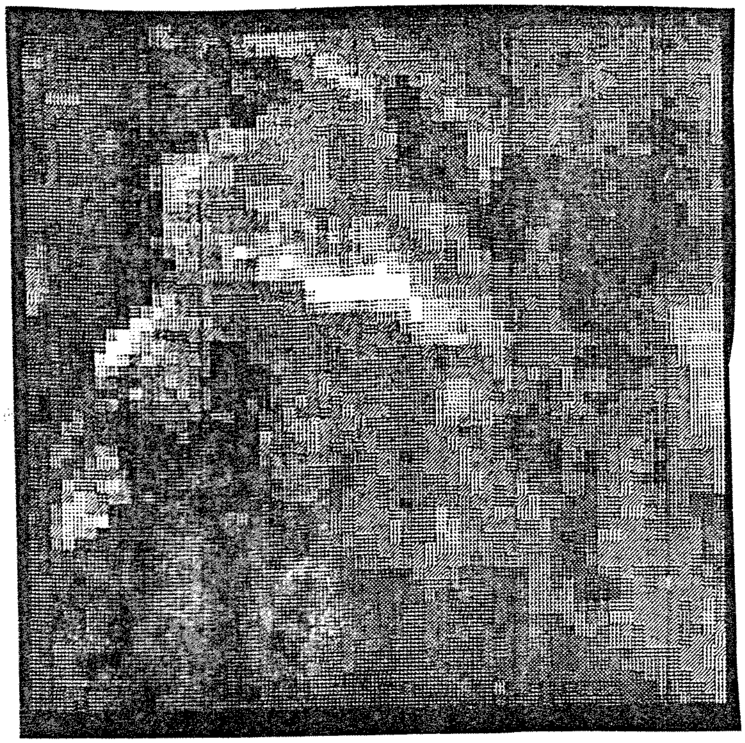


a)

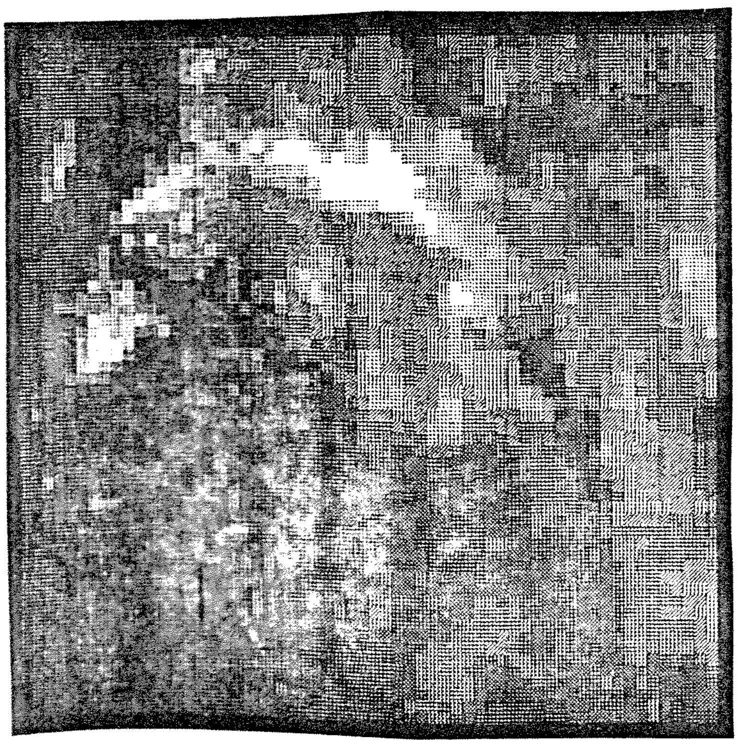


b)

Figure 2. Five original intensity images

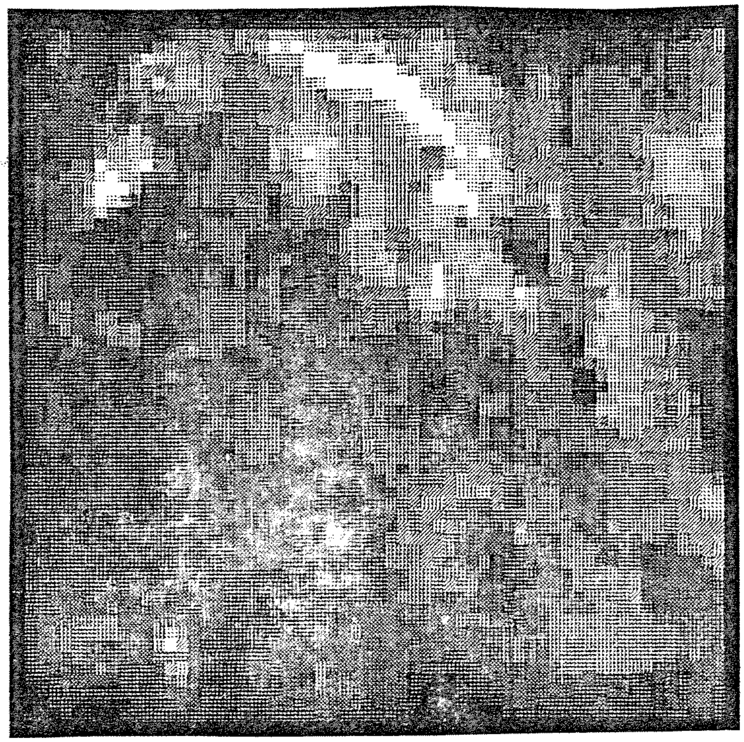


c)



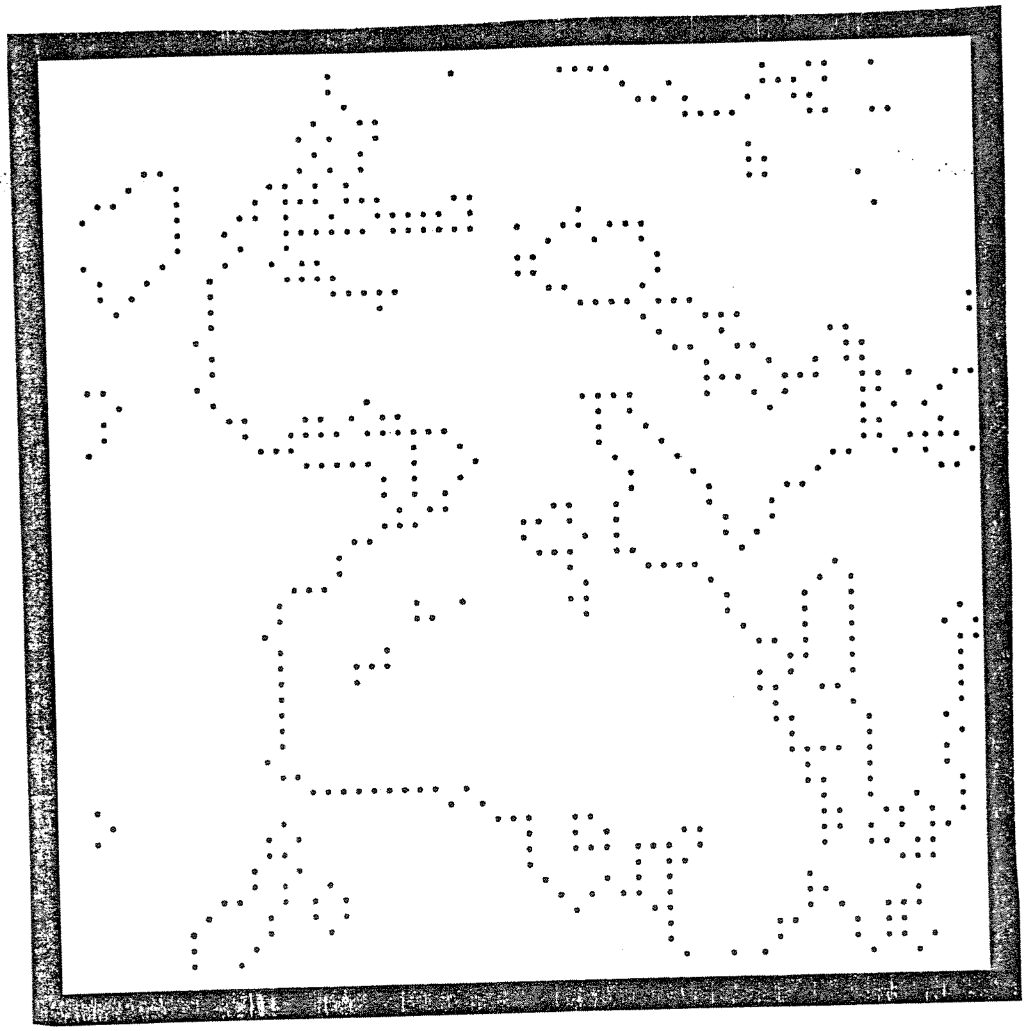
d)

Figure 2. Continued



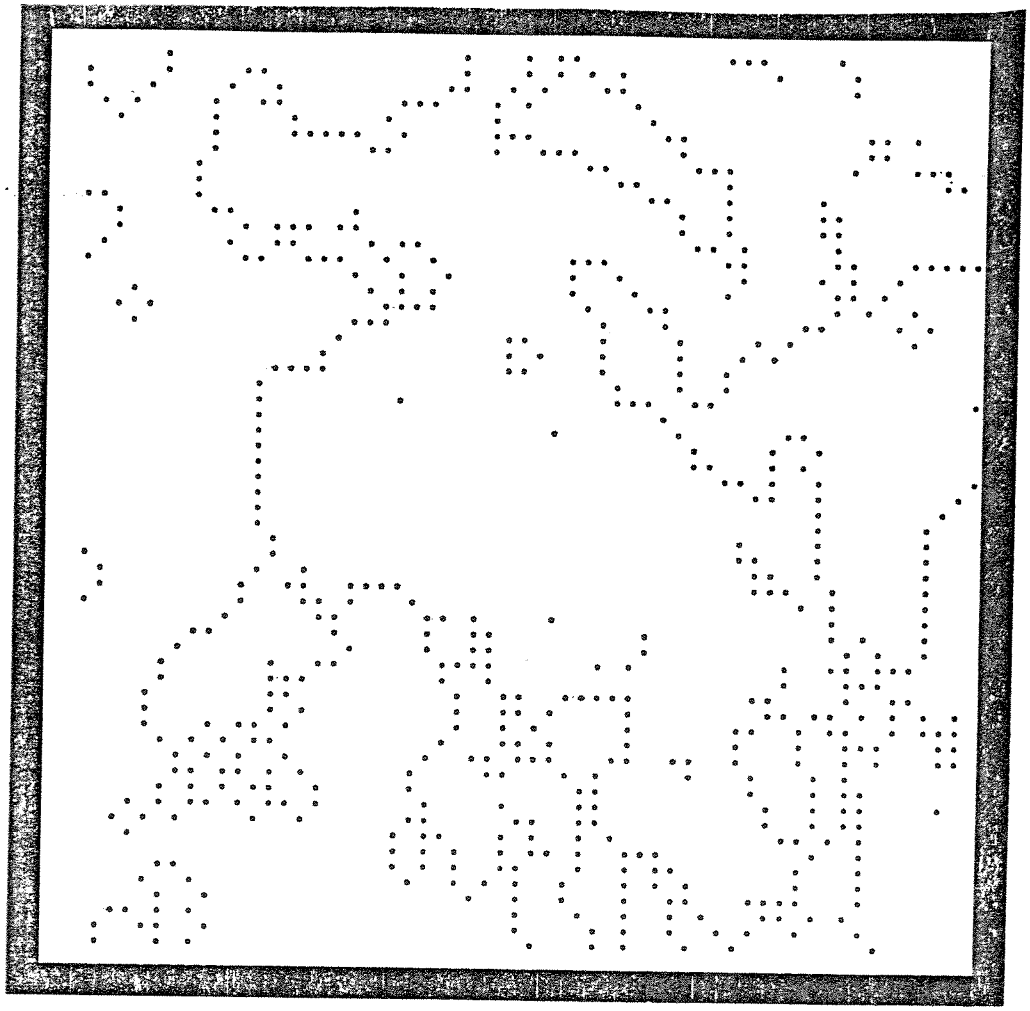
e)

Figure 2. Continued

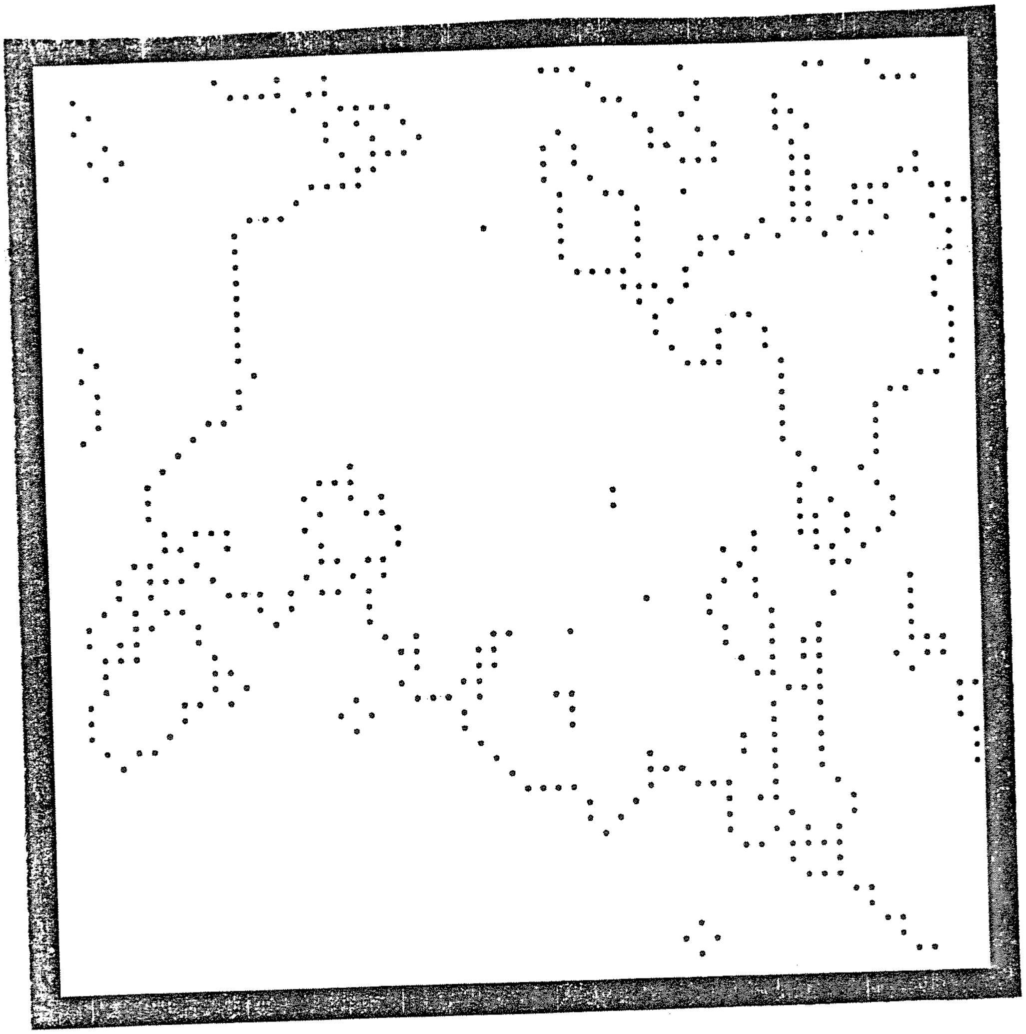


a)

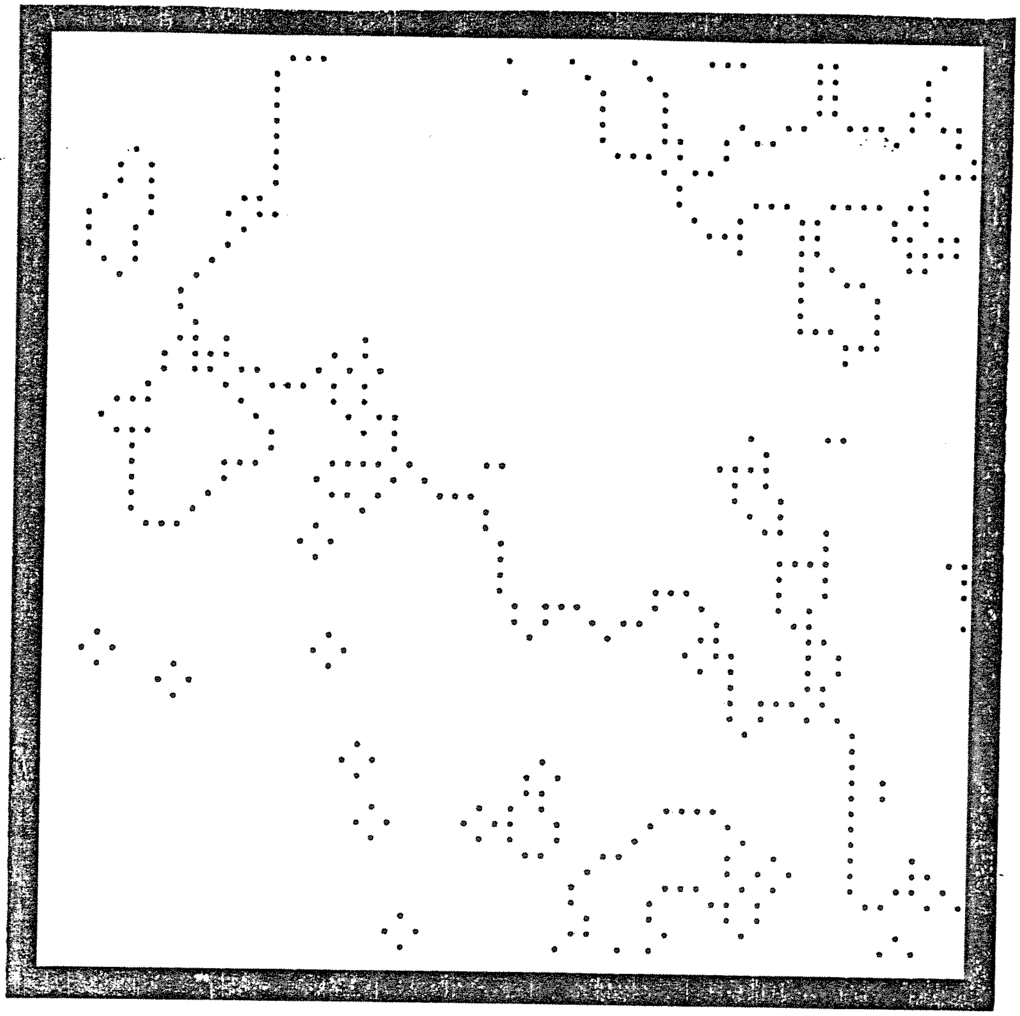
Figure 3. Binary edge arrays for Figure 1



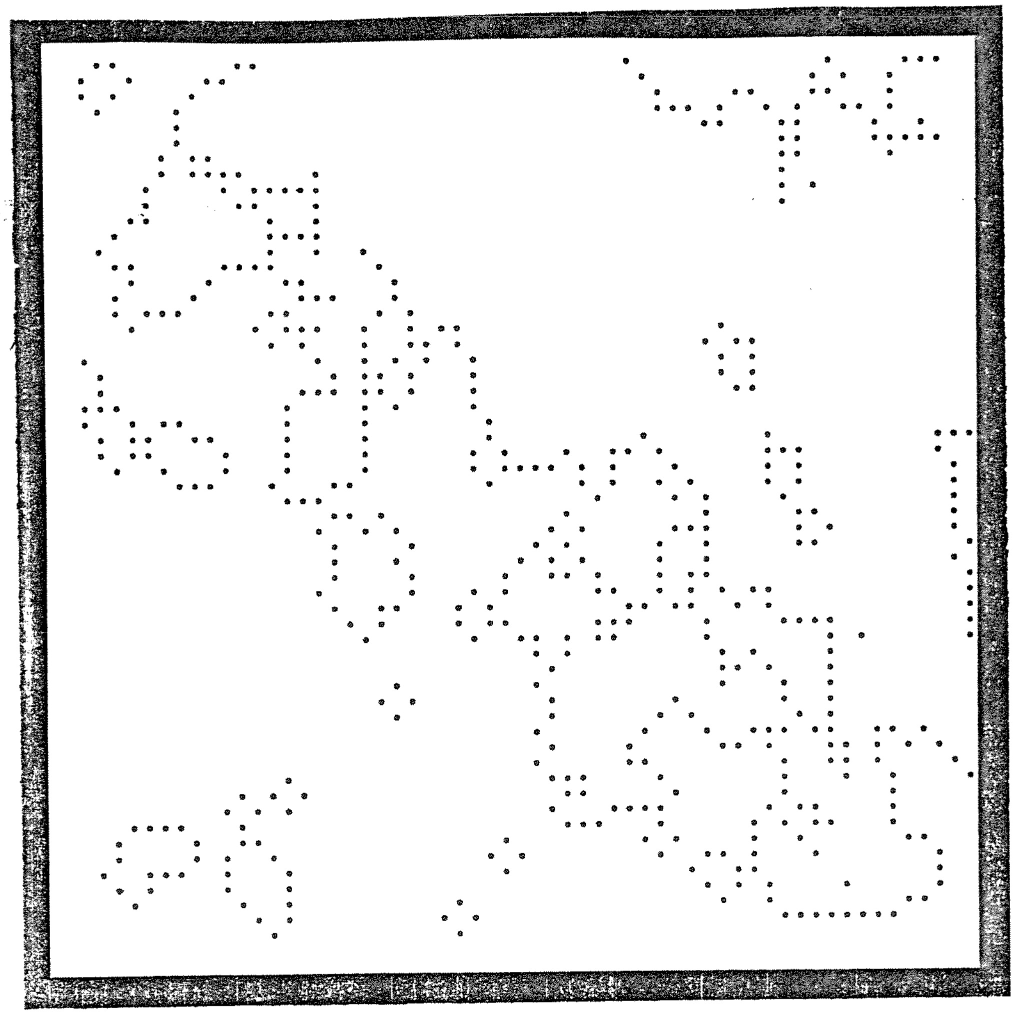
b)



c)



d)



e)

3. Interest Operator

In order to register images automatically from frame to frame, the system must first determine a set of interesting features from the initial frame so that it knows what to look for in the next frame. These interesting features are also called control points. The purpose of the interest operator is to acquire new control points. The selection of control points is very important in image registration since they must be located between two images to derive an accurate description of the image transformation. If a sufficient number of corresponding control points could be determined, then the image could be accurately registered, facilitating location of corresponding objects.

Moravec [3] presents an approach to extract control points from a scene. His interest operator is based on correlation and directional variance measures on grey scale images. In our system, however, the interest operator is applied only to binary edge pictures.

We have identified three desirable characteristics for choosing these control points when an edge array is used rather than the original intensity array.

(1) Continuity of curves

One of the desirable regions for the control point is where a long continuous curve lies; i.e., the region must contain a curve that enters from one side of the region and exits through the opposite side. The reason for this is that one would like to track the control points from frame to frame. A continuous curve reduces the chance of choosing features that are overwhelmed by noise.

(2) Local curvature high

A continuous curve in the region is not enough since there might be many other curves similar to it in the picture. We want a region distinct from other regions to generate sharp peaks in the Hough transform. Generally, curves with high curvature (many sharp corners) occur less frequently than curves with low curvature. Therefore, it is reasonable to choose the location of a control point on high local curvature regions.

(3) Uniqueness

We must stress that the control points acquired should be unique. Uniqueness may be measured by the number and relative magnitude of the peaks in the Hough transform of the original picture and the chosen region. If unique regions are chosen, ambiguous correlations can be reduced.

Hall, Davies, and Casey [4] give a detailed account of why uniqueness is important in relation to control point selection. They also present a method for selecting control points based on minimum correlation length.

In our system, continuity is measured over small square overlapping windows of specified size, typically 3*3 to 11*11. A weighted sum of the number of arc points (a), the number of end points (e), and the number of isolated points (i) is recorded as the continuity measure for the window.

The continuity measure, c, has value

$$c = \frac{a - 2e - 3i}{a}$$

Notice that we do not explicitly check for continuous curves in the window, but rather make the simplifying assumption that a large number of arc points indicates the presence of a continuous curve.

Curvature is measured using the algorithm CURVATURE which is defined over 3*3 windows. This local curvature is integrated to obtain a curvature value for the entire window. The algorithm is described with reference to figure 4.

```
i-----i-----i-----i
i P8  i P1  i P2  i
i-----i-----i-----i
i P7  i  X  i P3  i
i-----i-----i-----i
i P6  i P5  i P4  i
i-----i-----i-----i
```

Figure 4

(1) If X is an edge point then perform (2) to (3), otherwise return with curvature equal to zero.

(2) If X is an isolated point or an end point, then exit with curvature equal to zero, otherwise do (3).

(3) For each pair of edge points P_i, P_j, compute the dot product of vectors $\overrightarrow{XP}_1$ and $\overrightarrow{XP}_j$, and return with curvature equal to the maximum of the dot products.

Unique control points are chosen by computing local maximum of the sum of the continuity and curvature measures. Figure 5 shows the "interest field" (i.e., the sum of continuity and curvature), for frame 1.

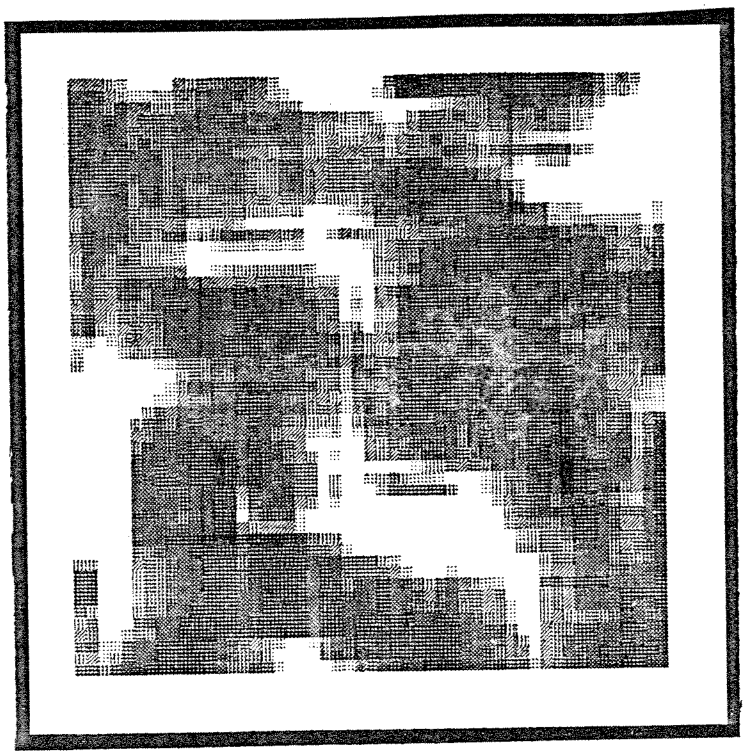


Figure 5. Interest field for frame 1

Local maxima of the field are chosen as the control points for the frame. Figure 6 shows the control points chosen from frame 1 on the basis of the values in the "interest field".

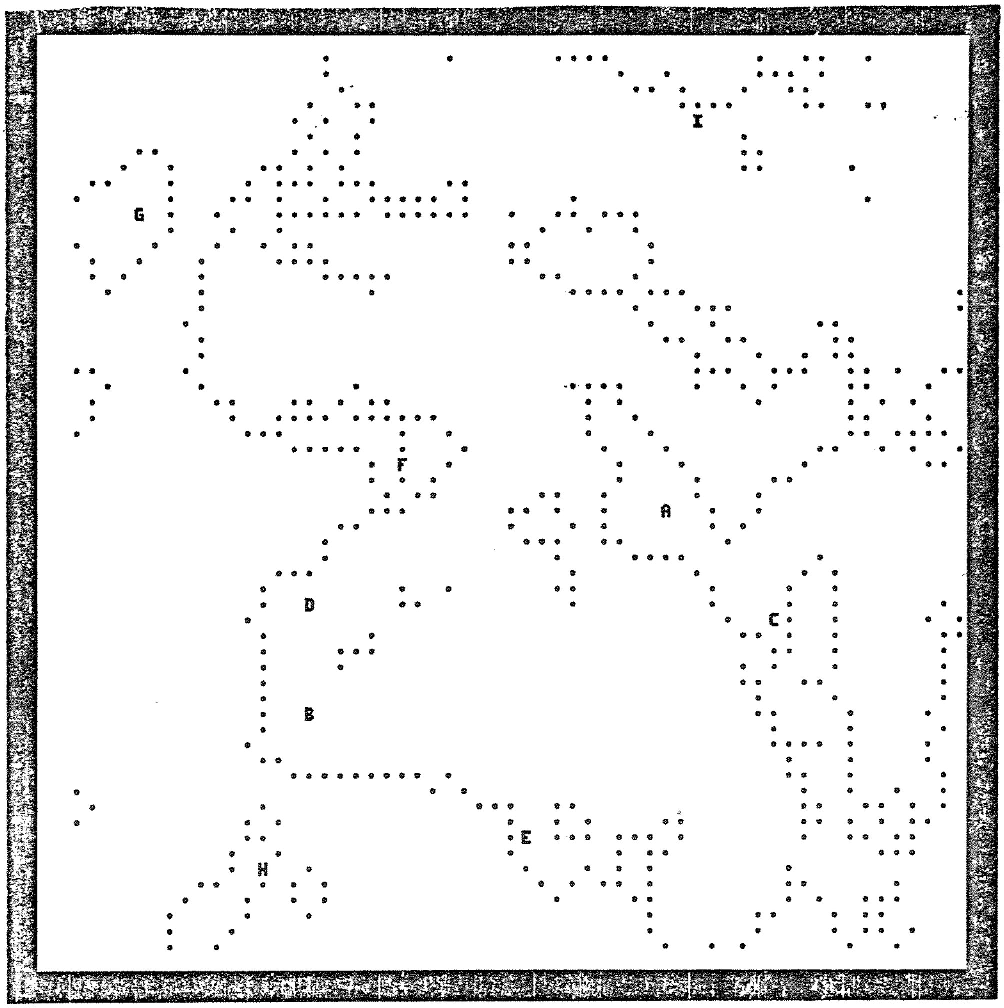


Figure 6. Control points for frame 1

Algorithm Match

For each $E_i = (X_i, Y_i)$ in E do

For each $D_j = (D_{X_j}, D_{Y_j})$ in H(B,P) do

$$h(X_i + D_{X_j},\ Y_i + D_{Y_j}) := h(X_i + D_{X_j},\ Y_i + D_{Y_j}) + 1\ ;$$

Note that algorithm MATCH only performs position invariant matching which is adequate for the present application of the system. However, algorithm MATCH can be easily extended to handle orientation and scale invariant matching; Ballard [5] and Davis and Yam [6] present descriptions of such extensions.

5. Hierarchical Hough Transform (HHT) for Maximal Set Selection

If the GHT operator were used to detect a single match position for each control point, then it would often fail to detect the correct position. Therefore, instead of choosing one maximum from the transform, a list of local maxima, H_{i_j}, j=1 to m, are chosen from the transform corresponding to the control point, C_i. The consistency algorithm then determines which of the maxima, if any, from H_{i_j} corresponds to the "correct" location of the control point, C_i.

A set of control points is said to be consistent if they maintain all their pairwise relative positions (to within some tolerance) across frames. A consistent set is maximal if it contains the largest number of control points amongst all consistent sets.

There are many consistency algorithms that can be applied to this problem. For example, see Bolles [7], Fischler and Bolles [8], Rosenfeld and Ranade [9]. The consistency algorithm chosen is based on computing a second Hough transform. We call this algorithm the hierarchical Hough transform (HHT). (See Davis [9] for a detailed discussion.) The HHT algorithm used in the system contains two levels. The first level is performed by the previously described GHT operation. Its function is to determine all the hypothetical control point locations, $H = \{H_{i_j}\}$, i=1 to n, j=1 to m, in the next frame, given a set of control points, $C = \{C_i=(X_i,Y_i)\}$, i=1 to n, in the original frame. The input to the second level is H(C,P), the Hough representation of the control points with respect to an arbitrary point, P. The following procedure is then applied to the set of hypothetical control point locations.

6. Algorithm Match1

For each $D_i = (DX_i, DY_i)$ in $H(C,P)$ do

For each $H_{i_j} = (X_{i_j}, Y_{i_j}) \in H$

$$h(X_{i_j} + DX_i, Y_{i_j} + DY_i) = h(X_{i_j} + DX_i, Y_{i_j} + DY_i) + 1$$

High values of the transform array, h, will correspond to the possible locations of P. Once the location of P is determined, the control points can be located. In practice, smoothing and/or sharpening procedures must be added to algorithm Match1 to account for the positional errors produced by the matching algorithms. In the system, the control point locations are allowed to deviate within the specified window size.

7. An Experiment

Five aerial photographs are given as shown in figure 2. The "super slice" edge detection program is applied to the grey scale image of the photographs, and five edge pictures are obtained as shown in figure 3. The interest operator is applied to the first edge picture and ten control points, labeled A to J, are selected as shown in figure 6. The control points are then registered with the second edge picture using the GHT operator. For each of the ten transforms (corresponding to the ten control points), five maximums are selected as the hypothetical locations of the control points, also labeled A to J, as shown in figure 7. (There are a total of 10*5 = 50 hypothetical locations for the 10 control points.) The HHT consistency algorithm is then applied to the set of hypothetical control point locations, and the maximal consistent set of hypothetical control point locations is obtained. Figure 8 shows the maximal consistent set for frame 2. If the number of elements in the maximal set is less than ten, then additional control points are selected from frame 2. The elements in the maximal set together with the new control points are then used to register with frame 3. Note that if one of the control points or maximal set elements is predicted to be outside the next picture frame, then that control point or maximal set element is removed and a new control point is selected. The same procedure is repeated for frames 4 and 5. Figure 9 shows the results of the registration for the remaining frames.

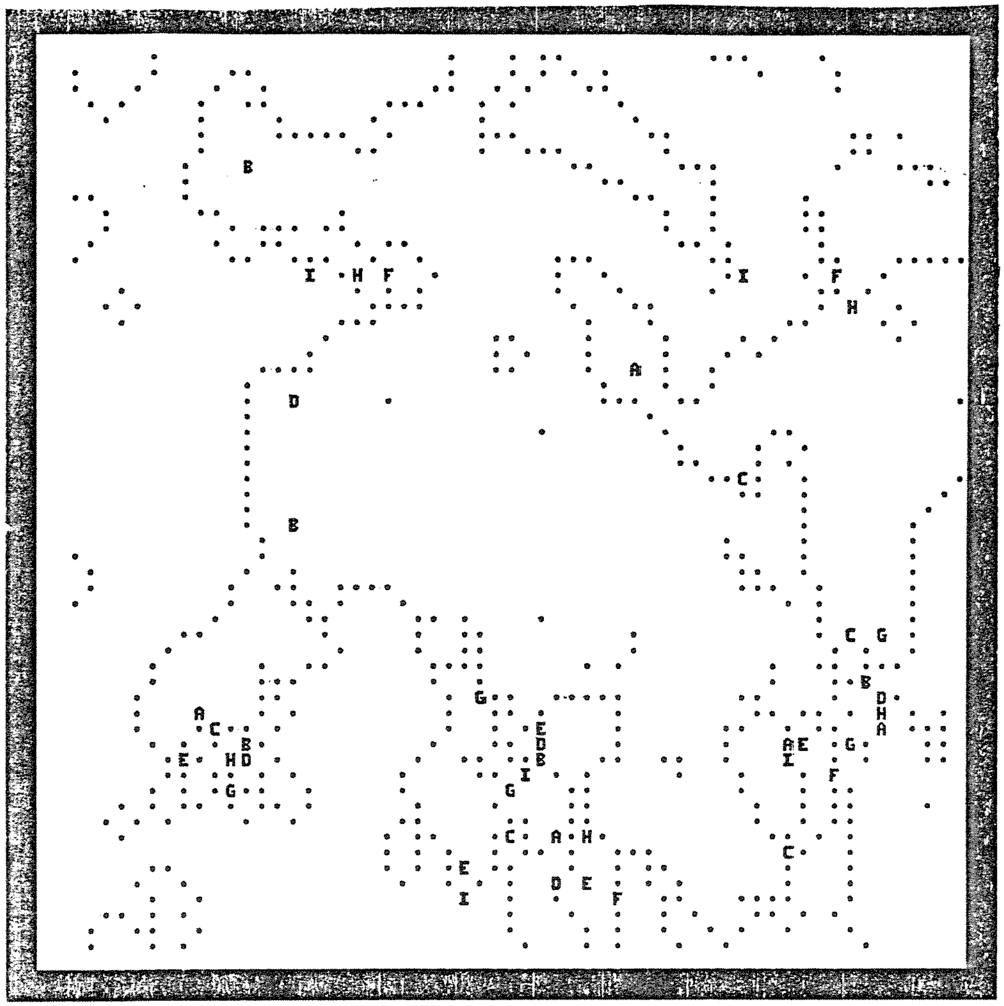


Figure 7. Control point matches in Frame 2

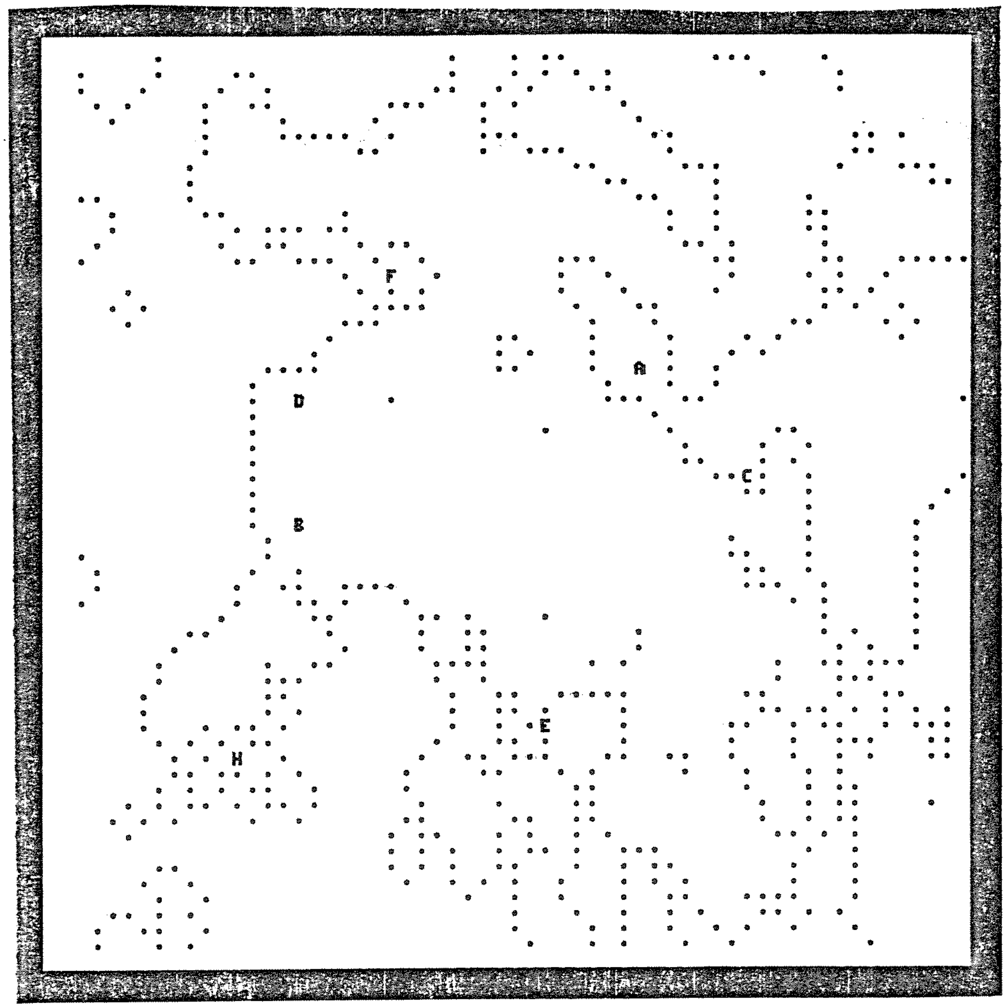


Figure 8. Maximal consistent set for frame 2

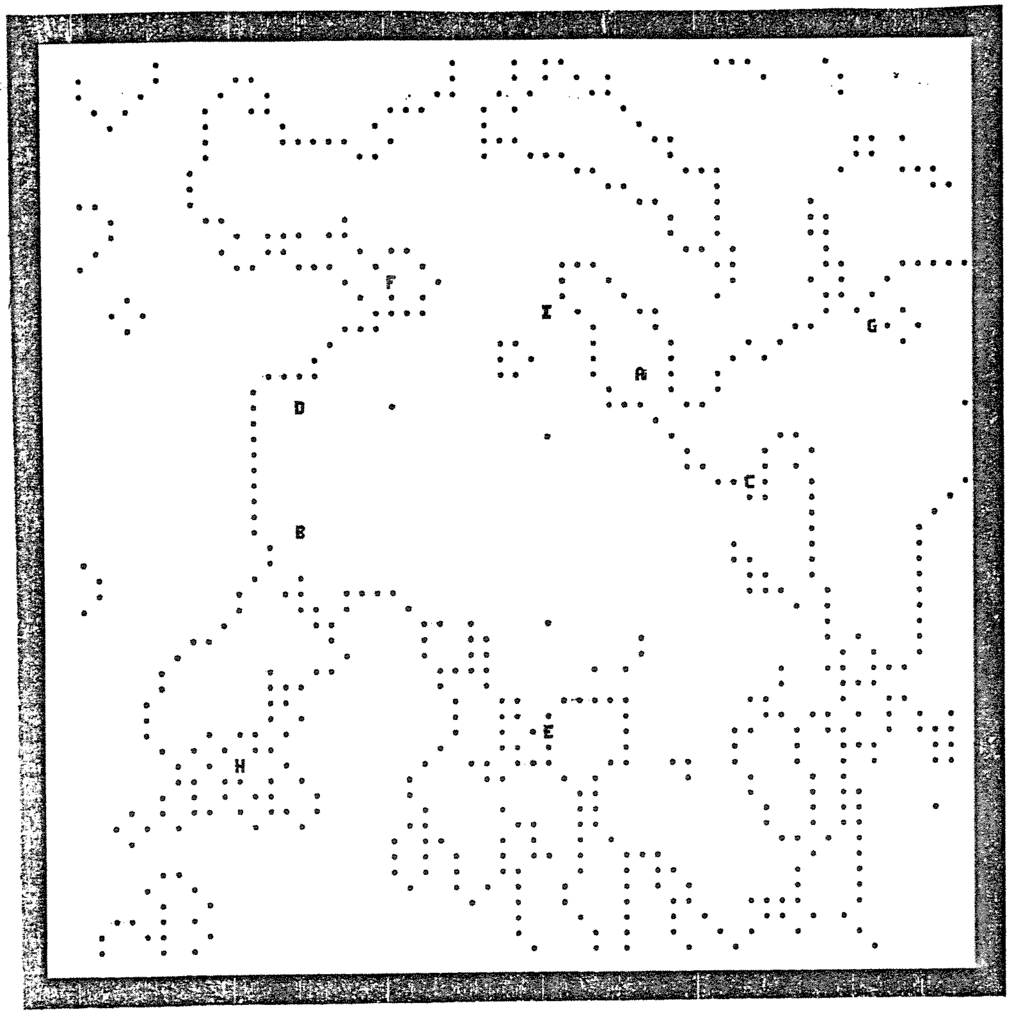


Figure 9a. Control points for frame 2

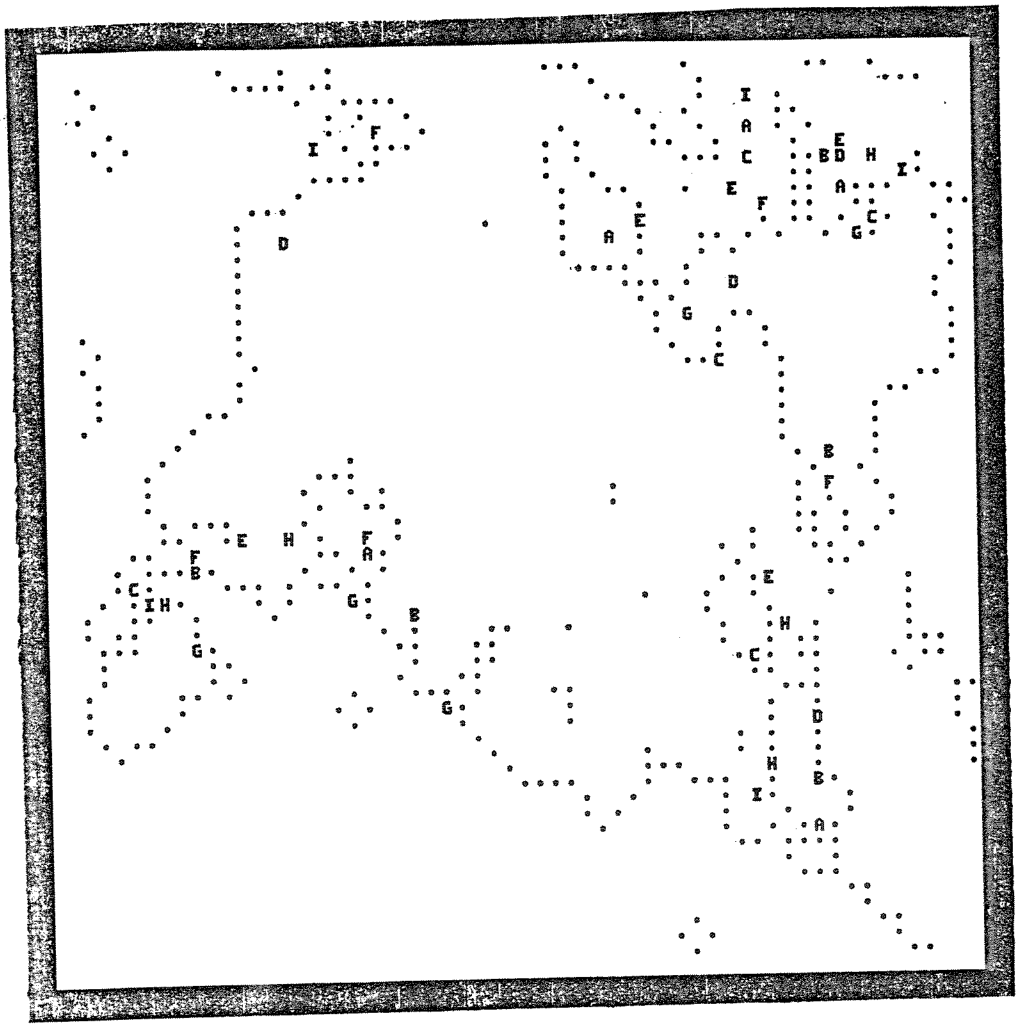


Figure 9b. Control point matches in frame 3

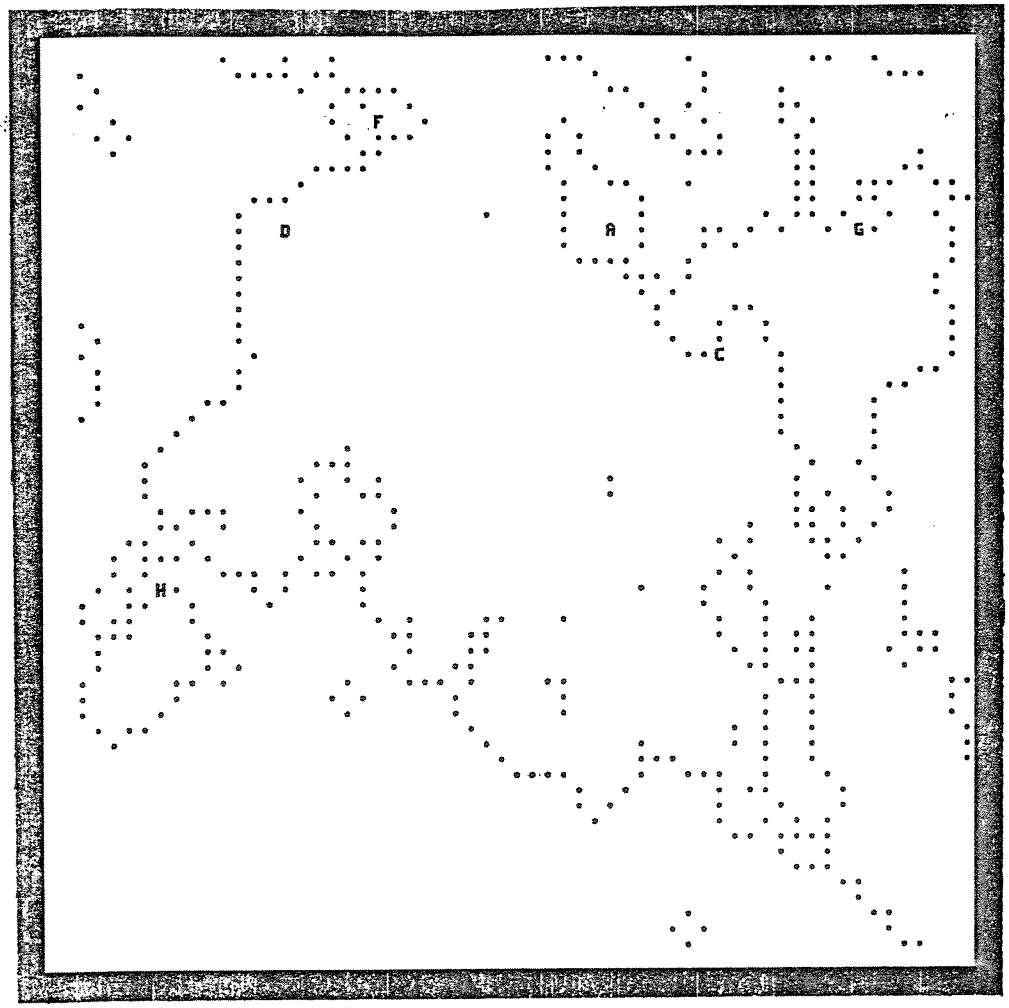


Figure 9c. Maximal consistent set in frame 3

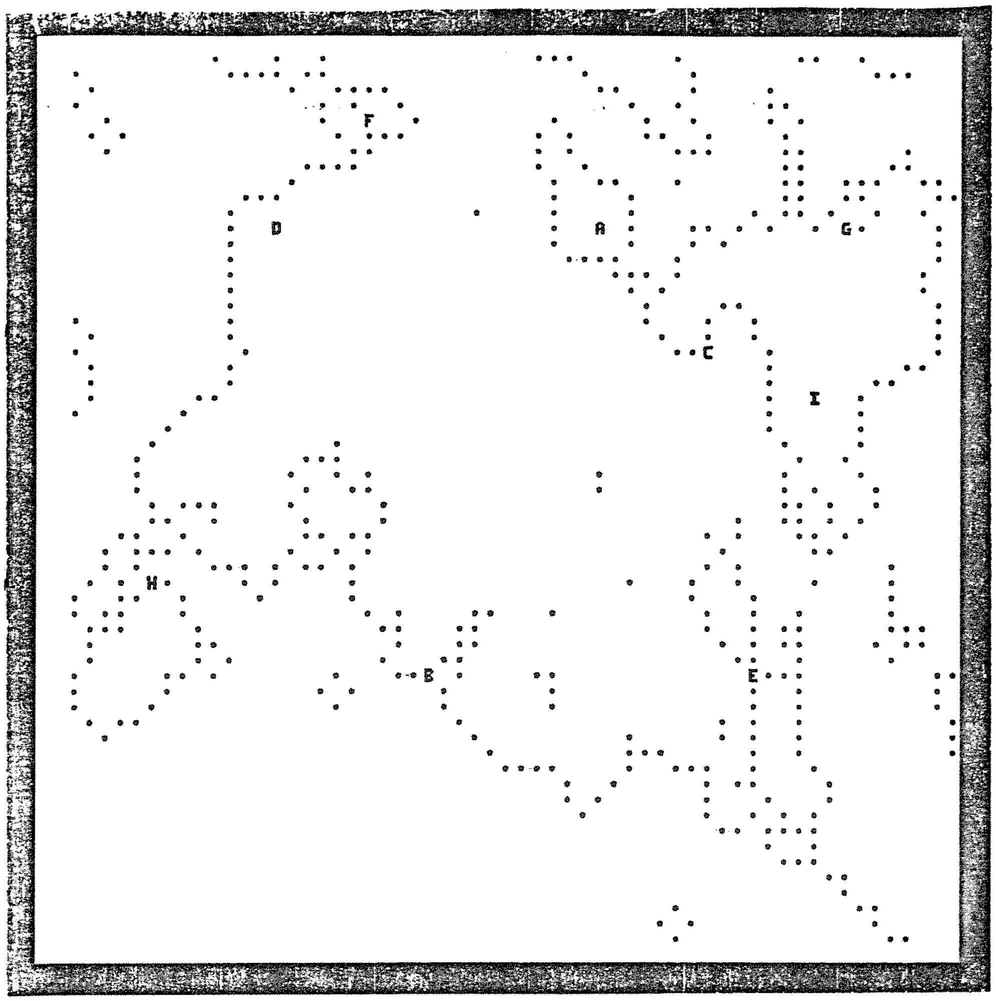


Figure 9d. Control points for frame 4

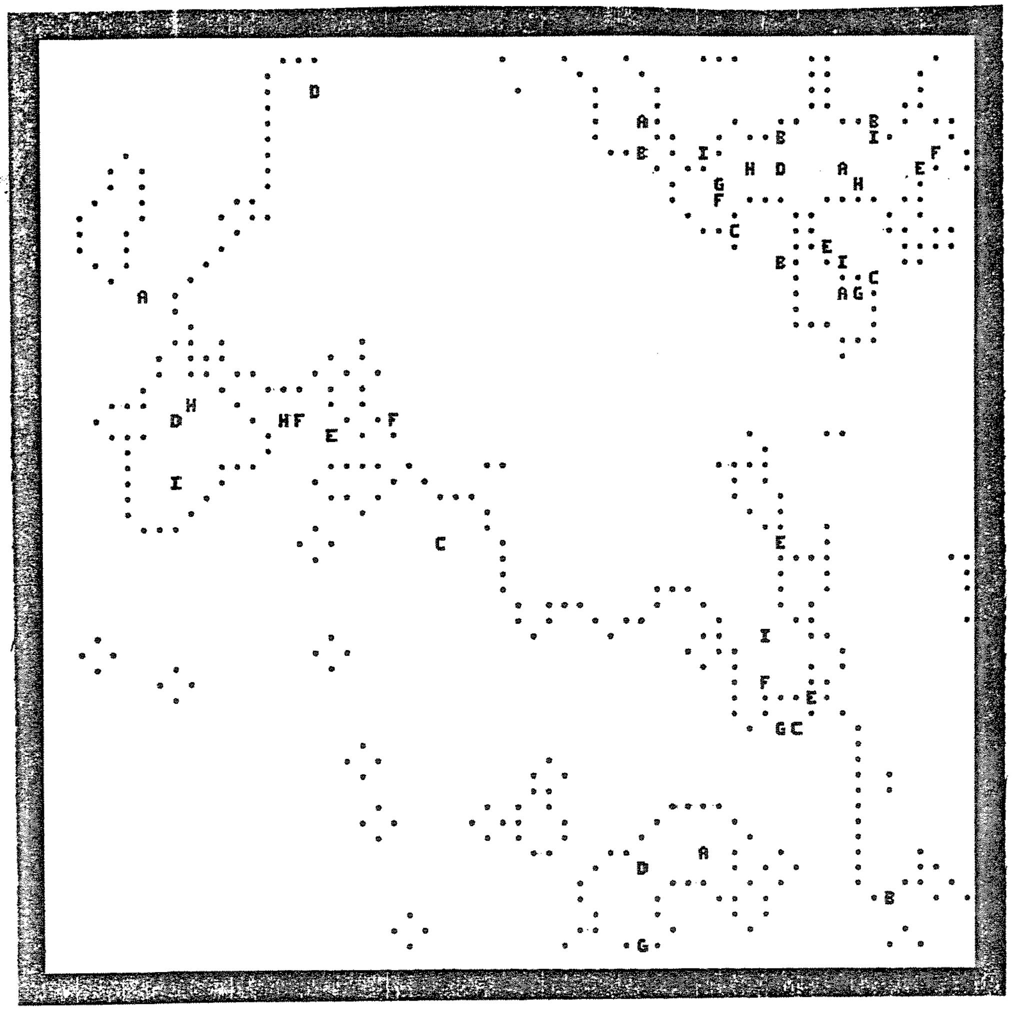


Figure 9e. Control point matches in frame 4

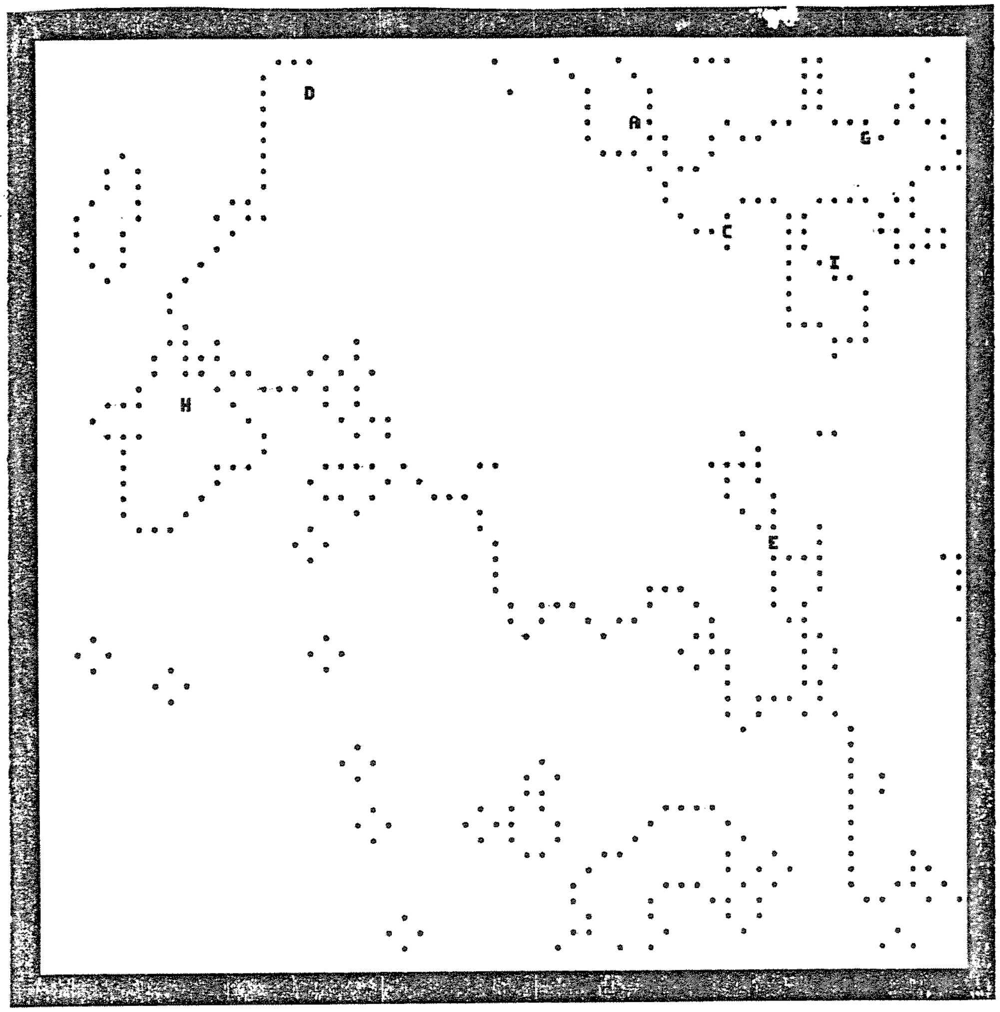


Figure 9f. Maximal consistent set in frame 4

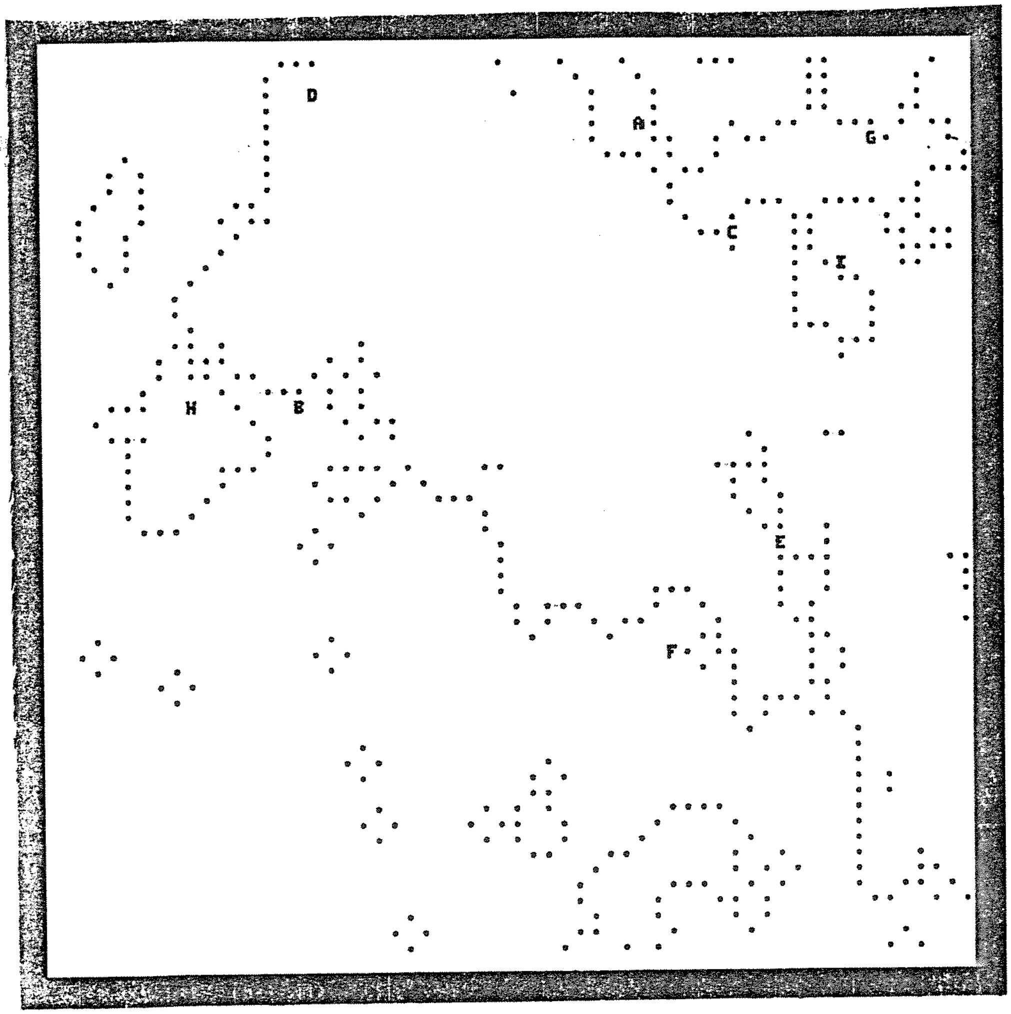


Figure 9g. Control points in frame 4

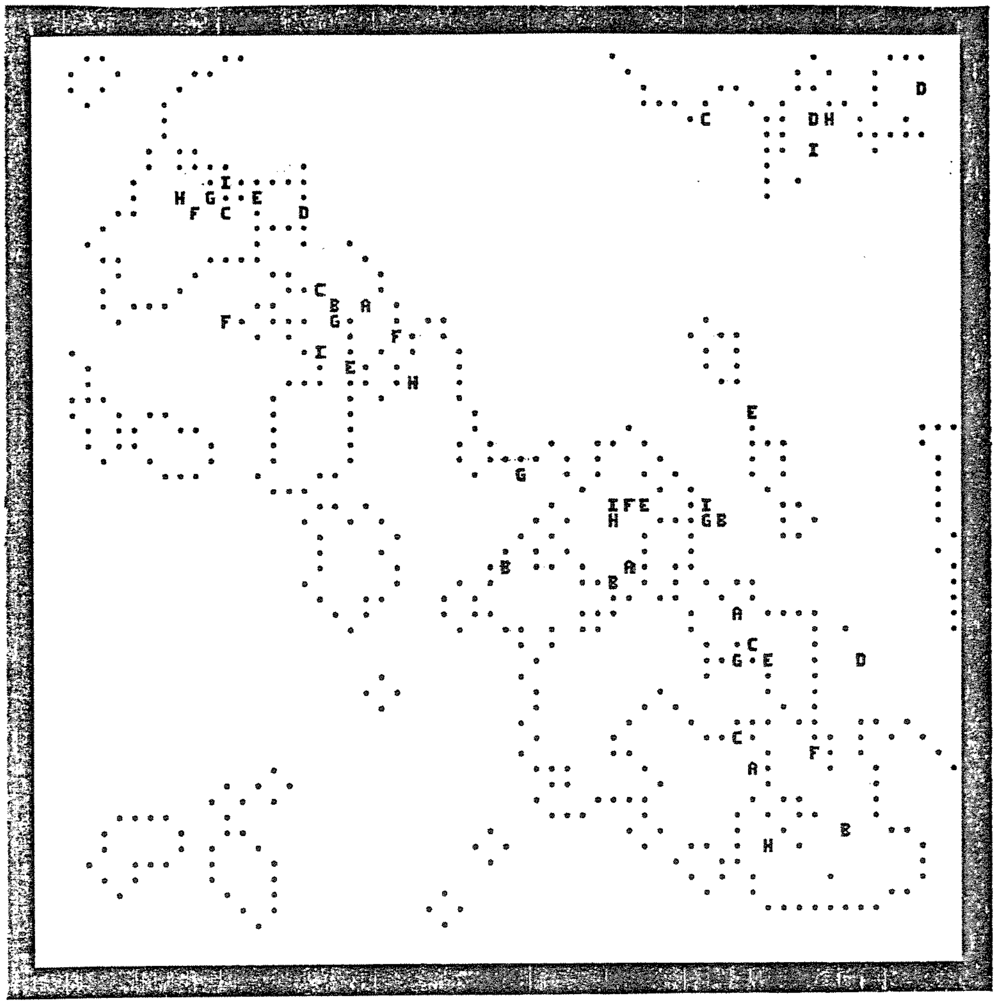


Figure 9h. Control point matches in frame 5

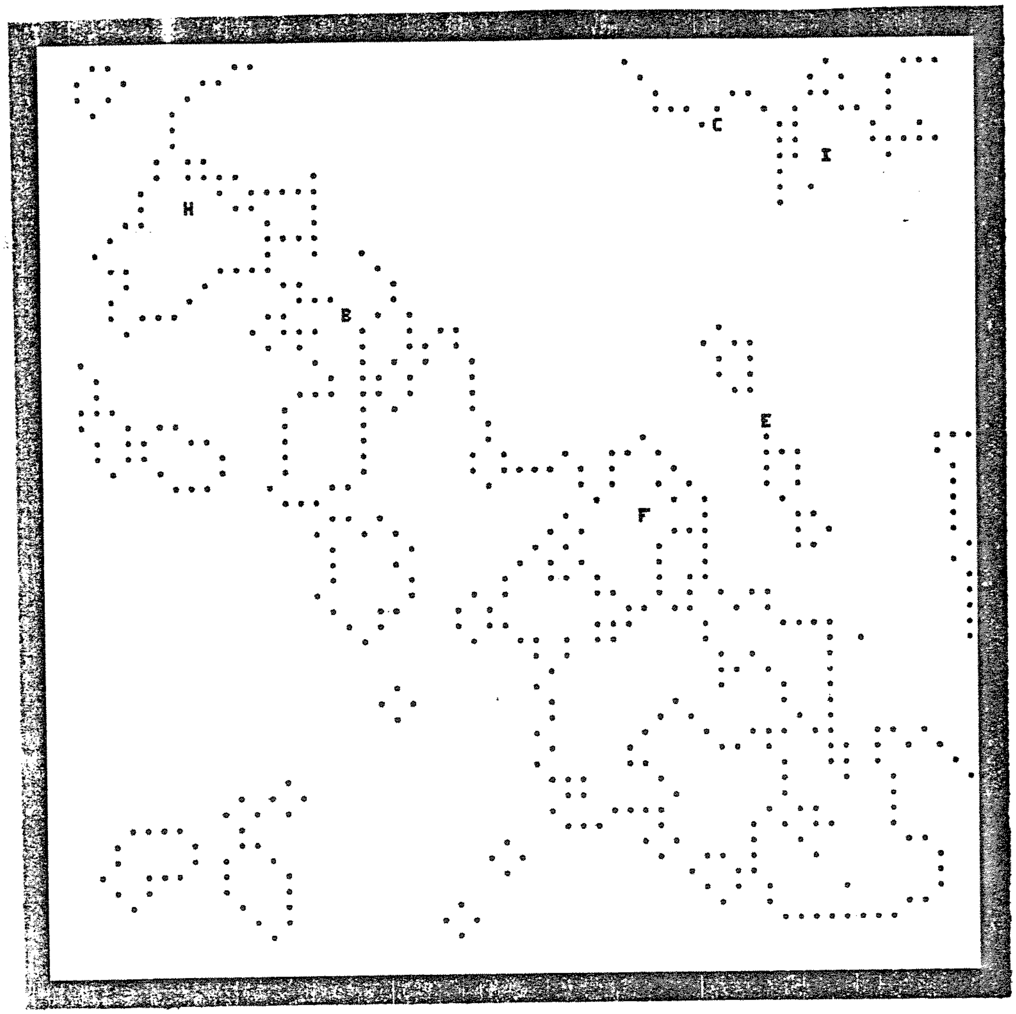


Figure 9i. Maximal consistent set in frame 5

8. Conclusion

We have presented an automatic image registration system based on the generalized Hough transform and the hierarchical Hough transform techniques. A new way of selecting interesting features from edge pictures is also presented. The system was successfully applied to the registration of a sequence of five aerial photographs.

Acknowledgement: We are grateful to Dr. David Milgram of the Lockheed Corporation, Palo Alto, CA, for providing us with the image sequence used in this paper.

9. References

1. D. Milgram, "Region extraction using convergent evidence," Comp. Graphics and Image Processing, 11, 1979, 1-12.

2. D. Marr and E. Hildrith, "The theory of edge detection," MIT A.I. memo, April, 1979.

3. H. P. Moravec, "Towards automatic visual obstacle avoidance," Fifth International Joint Conference on Artificial Intelligence, M.I.T., August, 1977.

4. E. L. Hall, D. L. Davies, and M. E. Casey, "The selection of critical subsets for signal, image, and scene matching," IEEET-PAMI, 2, No. 4, July, 1980.

5. P. Merlin and D. Farber, "A parallel mechanism for detecting curves in pictures," IEEET-Computers, 24, pp. 94-95, 1975.

6. J. Sklansky, "On the Hough technique for curve detection," IEEET-Computers, 1977.

7. D. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," Univ. of Rochester Comp. Science, TR-55, Oct., 1979.

8. L. Davis and S. Yam, "A generalized Hough-like transformation for shape recognition," Univ. of Texas Comp. Science, TR-134, Feb., 1980.

9. R. C. Bolles, "Robust feature matching through maximal cliques," Proc. SPIE Technical Symp. on Imaging and Assembly, Washington, D.C., April, 1979.

10. S. Ranade and A. Rosenfeld, "Point pattern matching by relaxation," PR-12, pp. 269-275, 1980.

11. L. Davis, "Hierarchical generalized Hough transforms and line-segment based generalized Hough transforms," in preparation.