Seventh International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Range, Forest and Wetlands Assessment

June 23 - 26, 1981

Proceedings

Purdue University
The Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

IMAGE PROCESSING FOR CARTOGRAPHIC APPLICATIONS

M.H. LOEW, R.L. PICKHOLTZ, L. GOLDMAN

George Washington University
Washington, D.C.

I. ABSTRACT

The goal of classifying objects of cartographic interest in aerial photographs was approached using techniques from pattern recognition and image processing. Bridge and airport images were chosen as the initial objects of interest and segments of photographs containing them were digitized for the data base. Edge-detection and Hough transform algorithms identified structures as candidate bridges; additional decision logic (using global contrast and other attributes) further reduced the set. Results indicate the feasibility and low computational cost of the approach.

II. INTRODUCTION AND BACKGROUND

In aerial cartographic analysis using optical images, it is unfortunate that at present few of the classical pattern recognition theories are immediately applicable. If we are to be successful with developing algorithms which actually work on real images, some ad hoc methods must be used. The reason for this is that it is rare that cartographically interesting objects can be assigned a priori statistical distributions such as are needed in a decision theoretic approach. Templates are not likely to be successful because even approximate shape for the same kind of object is too variable, let alone orientation variability. We therefore wish to examine the basic nature of the objects of interest. In aerial cartography, the basic nature can be dichotomized:

(1) Natural and gross man-made objects such as forests, fields, water, city streets, etc. These objects are properly characterized by the texture, reflectivity and fine structure.

(2) Discrete man-made objects such as bridges, roads, canals, airports, storage containers, specific industrial sites. These man-made objects have the distinguishing feature of having unique geometries, shape, boundaries, and fairly high contrast.

The natural and gross man-made objects probably will allow some degree of computerized recognition by taking advantage of the texture and fine structure. Algorithms exist that discriminate forest, field, urban area, water, or none of the above using sample statistics of sections of images.[1] The sample statistics include averages, correlation, and absolute value, all of which are very easy to compute; the results are compared against an empirically-determined threshold. Geometry is not examined and the algorithm works very well when only one of the four objects is in the scanning scene. It remains to be determined how such algorithms, which depend on averages of the gross image, work when there is an overlap with other boundary objects.

In this paper we are concerned principally with the second of the categories of cartographically important objects, namely the discrete man-made ones. Objects in this class lend themselves to enhancement; and the recognition of them is enormously benefited by various types of digital image preprocessing such as linear transforms, edge detectors, local contrast changes and the use of a priori knowledge of the geometrical characteristics of objects in question. The main emphasis in this work was the development of a set of working tools which could thus be incorporated into a transportable algorithm that would work on actual images. To illustrate the principles and to develop a

Supported by contract DAAK70-79-C-0147 with the U. S. Army Engineer Topographic Laboratories.

concrete example, we directed our efforts principally to bridge-over-water recognition. The objective was to have the algorithm be effective regardless of the other objects in the field, even those that superficially might resemble a bridge. To do this we developed and refined set of digital preprocessing techniques such as edge detectors, image smoothers, straight-line transforms (quantized Hough and other variants), thresholding techniques, and variations on the medial axis transform. In addition, we incorporated the *a priori* knowledge of bridge characteristics including the global environment. The actual process of bridge identification then consists of applying the preprocessing operations in sequence and makes a "bridge" decision if all tests are met. In a sense, this procedure can be viewed as a syntactical approach when the pattern primitives are the individual operations (edge detection, Hough transform, thresholding, etc.)

III. APPROACHES TO AUTOMATIC CARTOGRAPHIC IMAGE CLASSIFICATION

A. PRELIMINARIES

Classification of a cartographic image is used here to mean the assignment of a class label to the image or a subimage. The label would denote a standard cartographic category--a bridge, road, river, etc.--and in general would identify a set disjoint from those corresponding to other labels. The simplest approach is template matching, which compares templates of prototypical cartographic objects to the image being examined. The template that is closest (as determined with a suitable metric) is declared the class of the image. Template matching usually is computationally easy and fast; it is, however, strongly dependent on the orientation and scale of the image; the quality of the image can also affect the match.

That sensitivity gave rise to the use of features that jointly characterize the pattern and that exhibit less variation as the image departs from the prototype or is corrupted by noise. Features are measurements made on the pattern and are chosen to be both easy to extract and effective in separating the classes.

B. PREPROCESSING

An image in analog form (e.g., a continuous-tone photograph) must undergo several preprocessing operations before it is in a form suitable for feature extraction. The first process usually applied to an input picture is that of digitization, in which this original picture is converted into another two-dimensional representation that is suitable for digital computer processing. The second process often is that of edge detection and thresholding, whereby the digitized image is examined for large changes in intensity, followed by a thresholding operation that retains only those edges that are likely to be significant in subsequent processing. Next is usually an image segmentation process, in which certain characteristics (e.g., texture) are utilized to subdivide the image into distinct nearly-homogeneous regions. Feature extraction follows; typical features in cartographic images include orientations and lengths of straight lines, regularity or periodicity, aspect ratios, and relative sizes. Finally, the features are used in a classifier that may evaluate a linear or nonlinear function of the features and compare it to a threshold or use the features sequentially, stopping when a decision can be made. Many classifier designs exist and tradeoffs between feature complexity and classifier complexity are inevitable.

C. INITIAL CLASSIFICATION

The major categories of cartographic objects and some examples are:

1. Road and Road-Like (Ribbon)
   - Dual Highways
   - Other Roads
   - Railroad Tracks
   - Pipelines
2. Simple Large Scale "Area" Target (Irregular)
   - Cemetery
   - Quarry or Borrow Pit
   - Marsh and Swamp
   - Forests
3. Small "Area" Targets (Regular)
   - Isolated Buildings
   - Storage Tanks
4. Complex (Aggregate) Targets--Large Scale
   - Airport
   - Urban Area
   - Industrial Area
   - Railroad (Switchyard) area
5. Water-Bearing
   - Rivers
   - Streams
   - Lakes
   - Canals

6. Appurtenances of Water-Bearing Bodies (cued by bodies of water)

   - Dams
   - Bridges (road or railroad may also cue)
   - Shoreline
   - Rapids and falls

7. Miscellaneous Category

   - Tunnel Entrances

The indicated categories are appropriate for the just level of a hierarchical procedure in which only a few features are used to make the initial separation. The examples of subcategories typify the second-level classification problem. It is likely that specialized techniques will be necessary to allow subdivision of the major categories. One approach is presented in the following section.

IV. ROUTINES FOR BRIDGE PATTERN DETECTION

The segment of the project discussed in the following has as its goal the generation of a small set of structures from an image that are potential bridges. These bridge candidates are sets of parallel lines that are long and relatively close together. Each possible bridge is then subject to a set of tests that result in the structure being labelled "bridge" or "not a bridge". The process involves four primary steps:

(1) Pre-processing (if desired or necessary to down-sample, reduce noise or both),

(2) Edge detection,

(3) Recognition of long, parallel lines from the edge-detected field,

(4) Testing the resulting lines for "bridge" or "non-bridge" conditions.

Each of these steps is discussed in turn.

A. PREPROCESSING TECHNIQUES

Two preprocessing routines have been developed: image down-sampling and image noise reduction. The potential desirability of these two procedures are clear. For example, the original images being processed have been digitized to 256 pixels/inch. The image scale, however, may be such that the structures to be detected are on the order of millimeters or more. Thus, a large computational advantage may be gained with minimal relevant information loss by down-sampling the image to, say, 64 pixels/inch. The benefit is a factor of 16 reduction in the number of pixels to be processed while retaining essentially all relevant detail.

Two methods for down-sampling are: first, by simply extracting every n-th pixel in every n-th row (giving an n-to-1 down-sampling); and second, by collapsing every n-by-n square of pixels into a single pixel with a value equal to the average of the n^2 pixels. The first method is computationally much faster and appears to be equally effective in the images being used. Almost all of the examples shown in the following are the result of processing sections of a 4-to-1 down-sampled image resulting from method one.

Two types of noise reduction may be performed: (a) eliminating isolated noise "spikes" via thresholding; or (b) image smoothing using one of three convolutional masks.

Noise spike elimination is performed as follows: the value of each pixel is compared to the average of its eight nearest neighbors. If the absolute value of the difference between the pixel value and the average is greater than a threshold, the pixel is set equal to the average. The threshold level may be dynamically varied over the image by setting it equal to a specified number of standard deviations above the 8-neighbor average. Thus, regions of "smoothness" are compared to a low threshold while rapidly varying regions are compared to a high threshold.

Image smoothing is achieved by specifying one of three 3 x 3 convolutional masks, providing a range of smoothing. The three masks are:

Mask 1: $$\frac{1}{9}\cdot\begin{array}{|c|c|c|}\hline 1 & 1 & 1\\\hline 1 & 1 & 1\\\hline 1 & 1 & 1\\\hline\end{array}$$

Mask 2: $$\frac{1}{10}\cdot\begin{array}{|c|c|c|}\hline 1 & 1 & 1\\\hline 1 & 2 & 1\\\hline 1 & 1 & 1\\\hline\end{array}$$

Mask 3: $$\frac{1}{16}\cdot\begin{array}{|c|c|c|}\hline 1 & 2 & 1\\\hline 2 & 4 & 2\\\hline 1 & 2 & 1\\\hline\end{array}$$

As can be seen, the masks differ in the weighting given to pixels nearer the central pixel. The convolutions are all normalized so that the average image brightness is unchanged.

B. EDGE DETECTION

A review of the literature reveals a large number of edge detection algorithms for digital images, most falling into one of two categories: template-matching operators[2,3] and differential operators[4,5,6]. The latter group includes both 2 x 2 and 3 x 3-pixel operators. Comparisons of these techniques[7,8,9] in terms of performance and computational simplicity suggest that the group of 3 x 3 differential operators may be preferred for the current task. Thus, it is to this group that attention was directed.

In general, the 3 x 3 differential operators are of the form:

$$\underline{F} = \begin{array}{|c|c|c|} \hline 1 & W & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -W & -1 \\ \hline \end{array} \qquad \underline{G} = \begin{array}{|c|c|c|} \hline 1 & 0 & -1 \\ \hline W & 0 & -W \\ \hline 1 & 0 & -1 \\ \hline \end{array}$$

With W = 1, the operator is known as the "Prewitt" or "smoothed gradient" operator. With W = 2, it is the "Sobel" operator. $\underline{F}$ and $\underline{G}$ are simultaneously convolved with the digitized image and an edge is judged to be present if:

$$(\underline{F} \cdot \underline{B})^2 + (\underline{G} \cdot \underline{B})^2 \geq A$$

where A is a pre-set threshold value, and B is the 3 x 3 subimage currently being tested. An alternative threshold expression compares the sum of the absolute values rather than the sum of the squares, for computational efficiency.

Numerous other techniques for edge detection have been tabulated in the literature and are not discussed further here.

C. LINE RECOGNITION

The output of an edge-detection algorithm such as described above is, in general, a set of discrete, disconnected edge segments, often appearing to be randomly placed. The next test is usually to use a "linking" or "line building" algorithm. These algorithms generally look for line segments that fall within a given tolerance of distance and orientation of each other. When such segments are found, they are linked, forming a single longer line segment. This process continues until all long lines (if present) are built up. All segments which are below some threshold value in length are dropped.

The above process, however, can be time consuming and, for the task of (for example) detecting bridges may not even be necessary. In the present task, we are not so much interested in line-building as in the specific question "are there long lines present in the image?" and, if so, "are there parallel lines whose lengths are much greater than their separation?". These questions arise since bridges are, of course, composed of long, close parallel lines. These types of questions may be answered using a Hough transform.

Basically, a Hough transform operates on a set of predetermined feature points in the image (or X-Y) space. The set of edge points resulting from an edge detection operation is such a set of feature points. The Hough transform uses these points to generate a set of points in rho-theta space, where the rho and theta values are coordinates of a line in X-Y space. That is, it is a line-to-point transformation.

To explain this transformation, assume that some line L exists in the image (X-Y) space). This line can be described by two coordinates: rho(R), the perpendicular distance of the line to the origin (which may be selected, for example, to be the lower left corner of the image), and theta (Θ), the angle that the perpendicular makes with the X-axis (Figure 1). Now, let B be an edge point of L detected by an edge operation. B is thus a "feature point". An infinite number of lines may pass through this point. Suppose, however, that we quantize the angles of the candidate lines into, say, eight values between 0 and 180° (22.5° increments). Thus, we allow one of eight possible lines to pass through B (Figure 2). We now plot the coordinates of each of these eight lines in (R-Θ) space (Figure 2). In general, the result is a curve as in Figure 2.

Now, suppose we have quantized rho into, say, n levels so that (R-Θ) space is represented by a 2-dimensional (R-Θ) matrix Figure 3). We now perform the plotting procedure for each of the 8 possible lines by incrementing (by 1) the appropriate cell in the (R-Θ) matrix. An identical process is performed for every feature point in X-Y space (i.e., every detected edge point). The final result is a matrix (in which each element defines a particular (R-Θ) pair) whose entries are equal to the number of times each cell was incremented. That is, a cell with a value of 20 implies that 20 edge points were detected each of which has one of its eight possible lines possessing those (R-Θ) coordinates. Since, of course, all lines with the same coordinates are collinear, a (R-Θ) matrix element with a high value is the same as saying that a large number of feature (i.e., edge) points lie along the same line. Thus (R-Θ) elements with values above some threshold are judged to be lines (or edges) that exist in X-Y space.

Cells with the same value of Θ but differing R's represent parallel lines in the image. Cells exceeding the threshold that have the same Θ with R's that are close together define image lines that are long, parallel, and close together. These are the potential bridges.

Although more to the point, the above process is still calculation-intensive with fine quantization, even if the number of feature points is relatively small. However, if the feature points are generated by the edge detection mechanism previously described, then we already have an estimate of the orientation of the line passing through each edge point: namely,

$$\gamma = \arctan[(\underline{F}\cdot\underline{B})]/(\underline{G}\cdot\underline{B})]$$

where $\underline{F}$,$\underline{G}$, and $\underline{B}$ are defined before. Thus, rather than calculating rho and theta for a number of possible lines (the number depending on the quantization of theta) and incrementing each of the appropriate (R-Θ) cells, we do it only for that line most likely to pass through each feature point (i.e., that with coordinates R,Θ)[10]. Since we already know Θ, we calculate rho:

$$R = X\cdot\cos\Theta + Y\cdot\sin\Theta$$

where X and Y are the image space coordinates of the feature point. The net result is a single addition calculation (of R) for each feature point to obtain the Hough transform.

Two weaknesses of the Hough transform become evident when it is applied to a large input image: first, that a large amount of memory is required to store the (R-Θ) matrix if fine quantization of rho is desired; and second, that no information is provided about where in the image the lines exist (i.e., no end points are provided). These difficulties may be avoided by using the following scheme: divide the image into a number of small (for example 32 x 32 pixel) sub-images. Then operate on these subimages individually. As will be described next, the final result of this scheme is a relatively small set of "bridge candidates" (long, close parallel lines). These may then be conveniently subjected to a series of tests or further analyses to determine its "bridge" or "non-bridge" status.

D. ADDITIONAL ANALYSIS

Although work on this step has just been started, the procedure is relatively straightforward once the bridge candidates have been isolated. In particular, one could do the following:

(1) decompose the image into textured segments (corresponding to, for example, water, land-urban, and land-rural via one of several proposed algorithms[11,12,13]. Then, long parallel lines over, say, water, are very probably bridges.

(2) perform a series of simple "environmental" tests on each candidate to determine its context. For example, one could determine a global contrast, as a measure of the mean and variance of a block of pixels on either side of the lines. Those over water would tend to uniformity and equality (i.e., similar means and small variances) for blocks or pixels on either side, since these would correspond to water. Another simpler procedure is to calculate the "edginess" of the region around a bridge candidate. Edginess is taken to be the number of detected edge points in the region divided by the total number of points in the region. Edginess is a common textural statistic and, again, we would expect the water regions (i.e., subimages containing bridges over water) would have a much lower edginess than other regions, due to its high uniformity.

As an example, we shall use the edginess value defined above to attempt a classification of four potential bridges. As before, these candidates are taken to be sets of long, close parallel lines extracted from image segments.

Figure shows a raw image of candidate number 4 and its edge-detected version, with Figure the Hough transform matrix. The parallel structure of the image is evident in the matrix. From experience, we may set the edginess threshold at, say, 0.20. Those candidates located in images with edginess values less than 0.20 will be judged as being over water, and therefore, bridges. Results are shown in Table 1.

Table 1
Edginess of Bridge Candidates

Candidate No.	# edge points (excluding bridge pts)	Edginess	Thres-hold	Classifi-cation	Actual
1	406	0.41	0.20	Non-bridge	Highway
2	84	0.09	0.20	Bridge	Bridge
3	313	0.31	0.20	Non-bridge	Highway/ Overpass
4	131	0.15	0.20	Bridge	Bridge

E. SUMMARY

Bridge candidates are extracted from raw digitized images by performing a Hough transform on an edge-detected image and keep-ing only those entries corresponding to long parallel lines that are close together. These are represented by matrix entries in the same column (same angle) and in rows that are adjacent or nearly adjacent. Further testing is then done on this small set of potential bridges to determine whether or not a bridge is present. An example using image edginess in the neighborhood of the candidate is given.

V. CONCLUSIONS

Progress has been made during the past year in the development of computer techniques to extract items of cartographic interest from aerial photographs. The problem as originally posed was the construction of a system that can perform automatic or semi-automated cartographic analysis; a number of subsidiary problems were identified that provided the pattern-recognition framework for the overall problem. To date we have examined two types of man-made cartographic objects--bridges and runway configurations--in a variety of surroundings. Results were good, indicating that it is possible to locate accurately those objects in the image, using a combination of edge-detection and transform techniques and certain a priori information about the two types of objects.

The success to date provides support for the use and enhancement of the Hough transform as a locator of lines and line segments; since lines often characterize man-made cartographic objects, there is value in further development of that tool. The use of a large number of picture segments, coupled with a variety of object/area types is necessary for completion evaluation of that approach to feature extraction.

The success of various features used in recognition of cartographic objects depends in part on the spatial resolution and intensity quantization of the digitized image. Although the present values of resolution and quantization are adequate, it may be that less of either or both would also be adequate; in cases where memory is limited one may have to make a trade between number of pixels and number of gray levels per pixel. The effect of such a trade on the effectiveness of various features is unknown.

Acknowledgment

The authors appreciate the contributions of Dr. F. Rohde and Dr. P-F. Chen, U. S. Army Engineer Topographic Laboratories.

VI. REFERENCES

1. Chen, P.F., Unpublished notes on experimental cartographic system, U.S. Army Engineer Topographic Laboratories, 1979.

2. Kirsch, R., "Computer Determination of the Constituent Structures of Biological Images, " Comput. Biomed. Res., Vol. 4, No. 4, No. 3, 315-328, 1971.

3. Prewitt, J.M.S., "Object Enhancement and Extraction," in Picture Processing and Psychopictorics, Lipkin and Rosenfeld, Eds., Academic Press, NY, 1970.

4. Duda, R.O., and P.E. Hart, Pattern Classification and Scene Analysis, New York, NY, J. Wiley, 1973.

5. Frei, W., and C.C. Chen, "Fast Boundary Detection: A Generalized and New Algorithm," IEEE Trans. Comput., Oct. 1977.

6. Roberts, L.G., "Machine Perception of 3-dimensional Solids," in Optical and Electro-optical Information Processing, J.T. Tippet, et. al., Eds., MIT Press, Cambridge, MA, 159-197, 1965.

7. Abdou, I., and W.K. Pratt, "Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors," Proc. IEEE, Vol. 67, No. 5, 753-763, 1979.

8. Deutsch, E.S., and J.R. Fram, "A Quantitative Study of the Orientation Bias of Some Edge Detector Schemes," IEEE Trans. Comput., Vol. C-27, No. 3, 205-213, March 1978.

9. Fram, J.R., and E.S. Deutsch, "On the Evaluation of Edge Detector Schemes and Their Comparison with Human Performance," IEEE Trans. Comput., Vol. C-24, No.6, 616-628, June 1975.

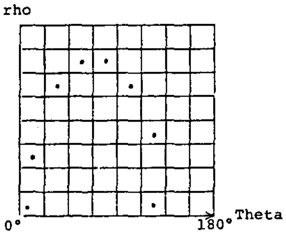
10. Gorman, F., and M.B. Clowes, "Finding Picture Edges Through Collinearity of Feature Points," IEEE Trans. Comput., Vol. C-25, No. 4, 449-456, April 1976.

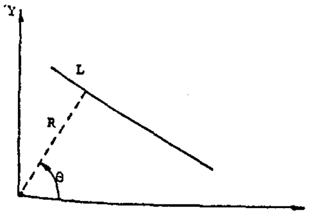
11. Haralick, "Statistical and Structural Approaches to Texture," Proc. IEEE, Vol. 67, No. 5, 786-804, 1979.

12. Hawkins, J.K., "Textural Properties for Pattern Recognition," in Picture Processing and Psychopictorics, Lipkin and Rosenfeld, Eds., Academic Press, Inc., 1969.

13. Thompson, W.B., "Textural Boundary Analysis," IEEE Trans. Comput, March, 1977, 574-577.

Fig. 3. The R-θ Matrix →





← Fig. 1. Hough transform definition

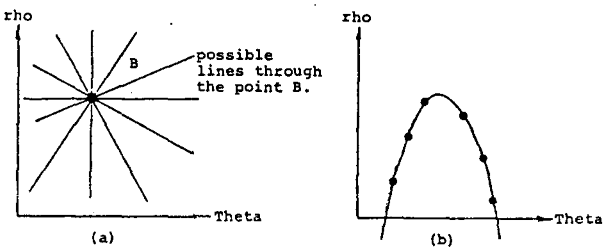


Fig. 2. Hough Transform space

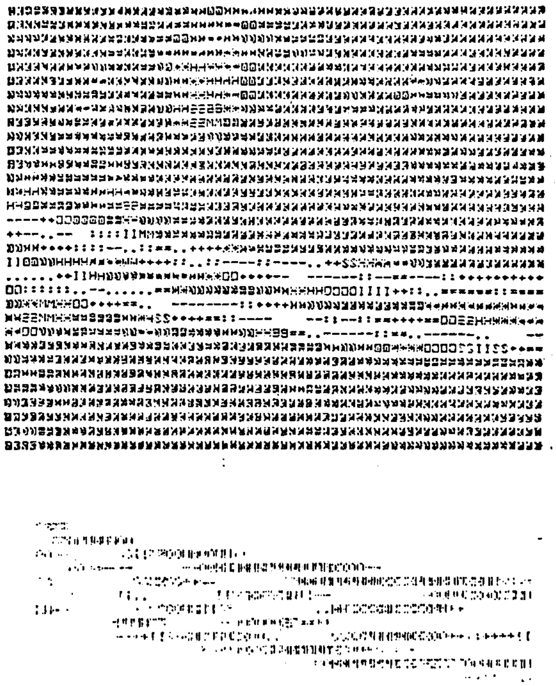


Fig. 4. A raw image (upper) and the non-blank portion of the image after edge detection (lower).

Rho						Theta							
1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	1	0	1	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0	0	0	[illegible]	[illegible]
4	0	0	0	0	0	1	0	0	1	0	1		
5	0	0	2	2	2	0	0	0	0	1	0	1	
6	0	0	1	0	3	2	1	0	0	1	7	3	
7	0	1	0	1	1	0	0	1	0	0	4	3	
8	1	0	1	0	1	4	2	1	0	0	7	3	
9	2	1	1	2	2	0	2	3	1	1	7	2	
10	7	1	1	0	0	0	1	0	2	1	3		
11	15	1	4	0	0	0	0	0	1	0	3		
12	14	4	0	1	0	0	0	0	0	0	0		
13	12	1	0	0	0	0	0	0	0	1	1		
14	4	1	0	1	0	0	0	0	0	0	1		
15	0	1	0	0	0	0	0	0	0	0	0		
16	0	0	0	0	0	0	0	0	0	0	0		
17	0	0	0	0	0	0	0	1	0	0	0		
18	0	0	0	0	0	0	0	0	0	0	0		
19	0	0	0	0	0	0	0	1	0	0	0		
20	0	0	0	0	0	0	0	0	0	0	0		
21	0	0	0	0	0	0	0	0	0	0	0		
22	0	0	0	0	0	0	0	0	0	0	0	[illegible]	
23	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1	[illegible]	
24	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0	0

Fig. 5. Hough Matrix for Fig. 4.