SIZE AND POSITION INDEPENDENT PATTERN RECOGNITION

By

DENNIS LEROY CRANSTON

Bachelor of Science

Oklahoma State University

1966

Submitted to the faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF BUSINESS ADMINISTRATION
July, 1967

SIZE AND POSITION INDEPENDENT PATTERN RECOGNITION

Report Adviser:

Wm. J. Meinhardt

Report Adviser

Richard R. Weeks

Norman N. Durham

Dean of the Graduate College

PREFACE

Part of the study of Organizational Theory is overlapped with Behavioral Science which is the science of man's interaction with his environment. One of the end results from the study, in my opinion, is to develop processes that will allow man or more specifically the administrator to dwell on conceptual problems rather then problems which might be best done automatically.

One area that research is progressing in for this end result is pattern recognition. While the immediate applications of pattern recognition seem far removed from management, yet so was the work being done with electricity in the 18th Century which has progressed to the modern day computers.

I wish to express my appreciation to Dr. W. Meinhart for his encouragement preparatory to and during the course of study. I also wish to thank my wife, Sharon, for her patience and perseverance in preparing this report.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Definition of Pattern Recognition

A pattern for pattern recognition purposes is a set of nonrandomly arranged, intersecting lines. Usually these lines form the outline of a recognizable object.

Recognition concerns the ability of a system to acknowledge the presentation of a pattern to its input and the ability to provide a particular output for a given pattern. The recognition system is analogous to a black box with an input terminal and a single or multiple output terminals. Each output terminal can give a binary answer, i.e., one or zero, yes or no.

The pattern recognition system operates by presenting a series of patterns called the pattern set to the input terminals, one pattern at a time. The pattern recognition system in the case of the single output gives a yes answer only if the pattern presented to the system satisfies the requirements of the system, otherwise, it gives a no answer. For the multiple outputs, the system gives a yes answer only on the output which corresponds to the pattern presented to it. If the pattern does not satisfy the requirements for any of the outputs, all the output terminals give a no answer.

Current Areas of Application

Neurological researchers use pattern recognition methods to investigate theories about the internal operation of the human brain. The scientists in this area are concerned with the development of basic elements and their interconnection to duplicate on an element by element level the response of the brain to stimulation by aural, visual, internal and other patterns.

Uses in the data processing field have thus far provided the greatest practical application of the techniques of pattern recognition. Magnetic Ink Character Recognition (MICR) and optical character recognition equipment have become common items for use by banks and other companies for input to computers. While the techniques in use are rather crude in comparison with theoretical techniques currently being worked on, they provide a standard by which experimental methods can be measured against.

Most of the theoretical work in pattern recognition being done today can be classed, for lack of a better title, as the simulation of the human brain. The work in this area views the human pattern recognition system as a black box and the researchers are interested in obtaining the same responses as those obtained from the brain without regard to how the brain accomplishes this task. The pattern recognition theories in this area can be divided into two approaches: those that are derived from the Perceptron theory originated by Rosenblatt (1) and other methods. One of the other basic approaches is in the n-truple scheme developed by Bledsoe and Browning (2).

The Bledsoe-Browning Method

The Bledsoe-Browning Method basically is a technique that uses submatrices to scan an input matrix. The input matrix is an N x N grid which has on it a pattern comprised of 1's on the outline of the pattern and zeros elsewhere. (See Figure 1)

X-Axis

1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0
9	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
10	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
11	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
12	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
13	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
14	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Y-Axis

Figure 1. Input Matrix

A set of submatrices each containing a particular feature is provided. An example set is shown in Figure 2.

Type-A				Type-B				Type-C				Type-D			
0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
0	1	1	1	0	1	0	0	1	1	1	0	0	0	1	0
0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0
0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0

Type-A Type-B Type-C Type-D

Figure 2. Set of Submatrices

Each submatrix is to be moved one position at a time over the input matrix until all possible positions have been covered. When the submatrix matches cell by cell with the corresponding squares in the input matrix the X and Y position and the type of submatrix is stored. For the example shown, the list in Table I would be developed.

TABLE I

TABLE OF MATCHES

Type	X-Pos.	Y-Pos.
A	7	4
B	1	4
C	7	8
D	12	3
D	12	8

When all of the submatrices have been used, the list of positions and types is compared with a master list, containing a list of factors for each of the patterns in the set of patterns that the system has been given. When the list developed by the scanning process matches a particular list of factors in the master list the pattern is said to be recognized.

Limitations of the Bledsoe-Browning Method

There are two limitations to the Bledsoe-Browning Method which restricts its use to adjusted patterns. This is not a criticism of their method but is an emphasis of the points where further development is needed.

For the Bledsoe-Browning Method and for most all other methods as well, the size of the pattern has to be fixed from one trial to the next. Also the position of each pattern in the input matrix is fixed for all trials. Once the system has been given the factors to recognize a pattern, that pattern must remain in the same grid positions as originally set up.

The purpose of this paper is to propose a technique by which these limitations can be removed. The standard by which the results of this technique will be measured will be the ability of the pattern recognition system to give the correct response to a pattern out of a set of similar patterns without regard to the size, position, and rotation of the pattern.

CHAPTER II

THEORETICAL DESIGN OF THE PATTERN RECOGNITION PROCEDURE

General Approach

The Size and Position Independent Pattern Recognition (SPIPR) System is composed of four phases. Upon presentation of the pattern to the input matrix, the phases are taken sequentially until the pattern is recognized. The system is assumed to have previously been given the parameters of the patterns for which the system is to recognize.

Phase One

The first phase is a semi-mechanical process of scanning the input matrix with the submatrices and determining the points at which they coincide.

If the subset of features particular to a pattern P_n is denoted S_n where n is the pattern number, then the set of features necessary to recognize patterns $P_1 . . . . P_m$ will be the sum of the subsets $S_1 . . . . S_m$, i.e., $S_{sum} = \sum_{i=1}^{m} S_i$. However, it can be shown that S will be considerably smaller than the S_{sum} indicated above.

If the elements in the input matrix are either 1's or 0's and the size of the input matrix is k x k then the number of possible patterns is $2^{k \times k + 1} - 1$. The submatrix is h x h giving $2^{h \times h + 1} - 1$ possible submatrices or features. Since h is considerably less than k then $2^{h \times h + 1} - 1 << 2^{k \times k + 1} - 1$. In practice the number of feasible submatrices will be a fraction of those theoretically possible because some of them can be derived from rotating others.

The scanning process involves taking each submatrix in turn, putting it over each possible location on the input matrix and comparing the corresponding elements. If all elements match, then the x and y position of the submatrix on the input matrix is recorded along with the submatrix type or number and its rotation. The submatrix is then rotated an increment and again passed over the input matrix as above. When the submatrix has been completely rotated, the next is used in the same fashion. The table developed in the above procedure contains, for each match, the submatrix type, the rotation, the x position and the y position. This table is then used as input to the next phase.

Phase Two

The information in the table generated in the scanning phase is tabulated on the basis of feature type. The generated table then shows the number of features found of each type which can then be compared with the stored tables. When a table in the library of stored pattern parameters matches the generated table, this set of parameters for the corresponding pattern are extracted. When the entire library has been compared, the result is a new library containing just the patterns that have the same submatrix type distribution as the pattern that had been scanned in the input matrix.

Phase Three

The third phase is similar to the second except that the generated table is tabulated on the basis of rotation index within feature type.

This is then compared with the library resulting from the previous phase with the matching patterns again being extracted.

Phase Four

The x and y position of each submatrix is used to prepare an array showing the distance from every feature on the input matrix to every other feature, i.e., the distance d from the i-th feature to the i-th feature is $d = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$. After the distance matrix has been calculated it is standardized with the largest value being set equal to a constant and the others adjusted proportionately. This is done to convert all patterns to a standard size so that regardless of the size of pattern presented, it can be recognized.

The determinate of the standardized distance matrix is then calculated to provide a distance matrix index. While the determinate of the distance matrix is not necessarily unique for each such matrix, the probability of two different distance matrices from two different patterns being the same, particularly after the above selection phases, is so low that difficulties with the use of it are not foreseen.

The determinate is then compared with the determinates in the list that resulted from the previous phase. When a determinate in the list is found that is equal to the calculated determinate, the corresponding pattern is the pattern that matches the input given to the system.

General Flow

It has been assumed in the above discussion that the number of possible matching patterns is sufficiently large so that it was necessary to carry the process through to the last phase to obtain an unambiguous response. As an examination of the Generalized Flow Model Figure 3 will

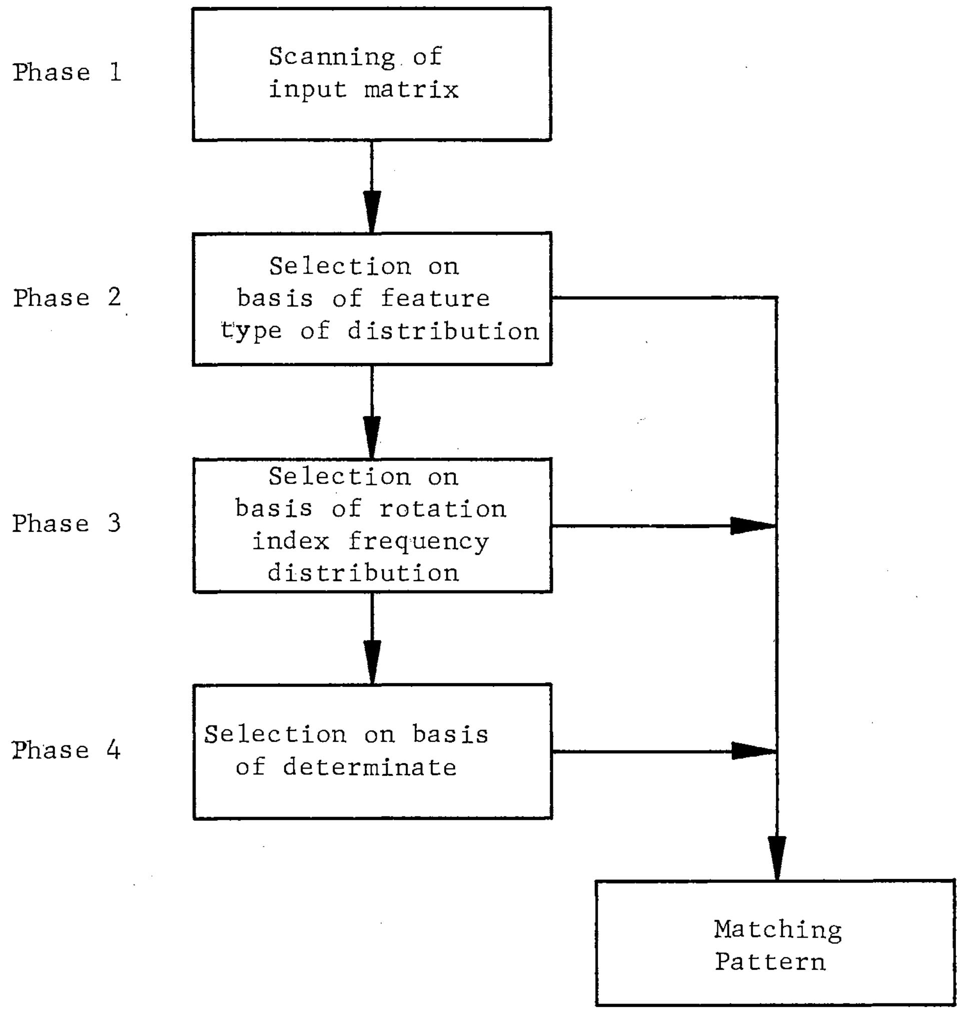

Figure 3. Generalized Flow Model

show, this assumption is not necessary in a normal situation. Whenever the number of possibilities left at the end of a phase is equal to one, it is unnecessary to continue the process.

While Figure 3 shows the flow to go step by step with each phase completed before the next one starts, the optimum flow would actually be in a concentric form. Using this form, a set of parameters for a pattern would be taken out of the library and compared on the basis of the feature type distribution. If it matched it would then be compared on the basis of the feature type rotation distribution rather then extracted as it would be in the case of the sequential flow. This process of comparisons is continued until a mismatch occurs at which point the pattern is discarded and another one is pulled from the library.

Rotation Index

The rotation index referred to in the previous discussion is a dimensionless integer indicating the number of increments a submatrix has been revolved about its midpoint. Because of the necessity for keeping the sides of the submatrix parallel to the axis of the input matrix, the rotation is not a true rotation. The same effect is arrived at by shifting the contents of each cell across and down keeping within the original layer as counted from the midpoint. This is illustrated in Figure 4.

$$\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ & 1 & \end{array} \qquad \begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ & 2 & \end{array} \qquad \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ & 3 & \end{array} \qquad \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ & 4 & \end{array} \qquad \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ & 5 & \end{array}$$

Rotation Index

Figure 4. Submatrix Rotation

Up to a certain transition point the maximum number of increments that can be taken is limited to 8. The use of a matrix with an odd number of elements on a side allows the use of the center cell to rotate around, simplifying the calculation of the next position for each cell. For a matrix of size k x k with k being odd the number of positions a cell will move for each increment is determined by the distance it is from the center cell $(\frac{k+1}{2}, \frac{k+1}{2})$; i.e., a cell one cell away from the center would move one position each increment, a cell two cells away would move two positions, etc.

Recognition Time

If this system were to be implemented, it would seem logical to eliminate the second phase and to go directly to the third and fourth phases. While this would give the same results, it can be shown that this is not the quickest method. Using the number of compares necessary to respond with the correct pattern as the criteria, the number needed for the four-phase alternative is:

$$N \cdot F + C_1 \cdot N \cdot F \cdot R + C_1 \cdot C_2 \cdot N/2$$

and for the three-phase alternative

$$(N \cdot F \cdot R + N)/2$$

where N = Number of patterns in library

F = Number of features used

R = Maximum number of rotation increments

C_i = %/100 extracted as a result of the previous i-th phase.

Therefore, if the four-phase alternative to use less time then the three-phase, the following must be true

$$N \cdot F + C_1 \cdot N \cdot F \cdot R + C_1 \cdot C_2 \cdot N/2 < (N \cdot F \cdot R + N)/2$$

if $C_1 = C_2 = .1$ (10%) and $R = 8$, then:

$$N \cdot F + .8 \cdot N \cdot F + .005 \cdot N < (8 \cdot N \cdot F + N)/2$$

$$N(1.8 \cdot F + .005) < N(8 \cdot F + 1)/2$$

$$1.8 \cdot F + .005 < 4 \cdot F + .5$$

Since $1.8 \cdot F$ is always less than $4 \cdot F$ the four-phase alternative will respond in less time then the three-phase. At small values of N and/or large values of C_1 and C_2 it is possible for the inequality sign to reverse; however, this is not usually the case. The equation for the four-phase alternative assumes a sequential flow; however, if a concentric form were used as discussed previously, the left-hand side of the inequality would be divided by two. This would allow a much greater reduction in response time for the four-phase alternative over the three phase.

An Example of the Basic Process

Referring to Figure 2 in Chapter I, it will be noted that submatrix Type-D is a rotation of Type-B. Therefore, by the use of rotation only three different types of features are needed for this example.

Scanning the submatrices A, B, and C shown in Figure 2 across the input matrix shown in Figure 1 gives the set of matches shown in Table II.

TABLE II

TABLE OF MATCHES

Match No.	Type	X-Pos.	Y-Pos.	Rotation
1	A	7	4	1
2	B	1	4	1
3	B	12	3	3
4	B	12	8	3
5	C	7	8	1

From this the distance matrix shown in Table III would be calculated. (Decimals are omitted)

TABLE III

DISTANCE MATRIX

	1	2	3	4	5
1	0	6	4	5	4
2	6	0	9	13	5
3	4	9	0	5	5
4	5	13	5	0	5
5	4	5	5	5	0

The frequency distribution of feature types would be: A-1, B-3, C-1. If a set of recognizable patterns have previously been given such as in Table IV the subset based on matches of the feature type frequency distribution (column I) can easily be selected.

From Table IV column I only pattern numbers 4, 5, and 7 meet the criteria. To decide among them it is necessary to look at the frequency distribution of rotations within types. From Table II this would be A,1 -1; B,1 -1; B,3 -2; and C,1 -1. Comparing this with Table IV column II shows that pattern numbers 4 and 7 match this breakdown.

The next step is to calculate the determinate of the distance matrix. For the values in the distance matrix shown in Table III the determinate is 27,400. Comparing this with the determinate values for Pattern Nos. 4 and 7 in Table IV (column III) only Pattern No. 4 matches and therefore is the correct response of the system to the original pattern presented to the input matrix.

TABLE IV

PARAMETERS OF RECOGNIZABLE PATTERNS

	I	II			III
Pattern No.	Feature Type A B C	Feature Type	Rotation	Frequency	Determinate
1	1 1 1	A	1	1	9,470
		B	1	1	
		C	1	1	
2	1 2 1	A	1	1	10,100
		B	1	1	
		C	1	1	
3	1 2 1	A	1	1	33,200
		B	1	1	
		B	4	1	
		C	1	1	
4	1 3 1	A	1	1	27,400
		B	1	1	
		B	3	2	
		C	1	1	
5	1 3 1	A	1	1	19,430
		B	1	3	
		C	1	1	
6	1 3 2	A	2	1	42,650
		B	1	1	
		B	3	2	
		C	1	1	
		C	2	1	
7	1 3 1	A	1	1	33,200
		B	1	1	
		B	3	2	
		C	1	1	

CHAPTER III

COMPUTER IMPLEMENTATION

General Approach

A computer program was written to investigate the characteristics and problems resulting from implementing the Size and Position Independent Pattern Recognition System. For ease in programming and debugging the program was written in two parts. The first part corresponds to the phase one description in Chapter II and the second follows the concentric alternative for phases, two, three and four. The two parts were written in FORTRAN IV for the IBM 7040.

The size of the input matrix was set at 35 x 50 and the size of the submatrices at 3 x 3. The 3 x 3 submatrix gives the possibility of 1023 ($2^{3 \cdot 3 + 1} - 1$) features. To reduce this number of possibilities to a manageable size it is necessary to define a feature for this application.

A feature for this use is defined as a unique (i.e., cannot be obtained by rotating or shifting another feature) irregular structure of 1's and 0's. The use of the word 'irregular' means that the feature is not a reflection about both its axis and 'structure' means that all cells containing a 1 are adjacent to at least one other cell containing a 1.

Using this definition the number of features that will be used is four. These are shown in Figure 5.

0	1	0	0	1	0	0	1	0	0	1	0
0	1	0	0	1	1	0	1	1	1	1	1
0	0	0	0	0	0	0	1	0	0	1	0
	1			2			3			4	

Feature Number

Figure 5. Features

Scanning Program

The scanning program listed in Appendix A was written closely following the theoretical description in Chapter II. Assuming a knowledge of FORTRAN IV a brief description is as follows:

Lines 12-15. Read in the four submatrices used in scanning.

Lines 16-23. Read in only the lines of the input matrix that contain the pattern. This is done rather than reading in the entire matrix to conserve time.

Lines 24-35. These steps set up the indexing required for scanning.

Lines 36-41. At this point the submatrix T(JMS, IMS) is compared with the input matrix ZMAT(L,K). If all the elements match the program continued to Line 42.

Lines 42-49. If the submatrix matched, the parameters are stored in the table TAB.

Lines 50-58. The submatrix is rotated one position each time.

Lines 64-68. The table generated by scanning is tabulated into the frequency type distribution table and the feature type rotation distribution table.

Lines 69-78. The x and y position is used to calculate the distance matrix from which the function (listed in Appendix B) in Line 78 calculates the determinate.

Lines 79-83. The various tables developed in this program are punched out for use by the selection program.

Selection Program

The selection program using the concentric form of search hunts through the library stored on magnetic tape until it finds the matching pattern. The listing is in Appendix C.

Lines 12-19. To eliminate the need for maintaining a magnetic tape the library was stored on cards and put on tape each time the system was used.

Lines 20-29. The parameters for the pattern to be recognized are read in.

Line 33. The pattern parameters to be compared are read in from the library on tape unit 0.

Lines 38-40. The feature type distribution is compared and if it matches the program continues to the next step.

Lines 41-56. The feature type rotation distribution is compared.

Lines 57-59. The determinate is compared.

Lines 60-63. The results of the search along with various statistics are printed out.

CHAPTER IV

EXPERIMENTAL RESULTS

Operation

The pattern parameter library consisted of different patterns ranging from simple patterns with eight or less features to the more complex with 15 or more. Using the four submatrices the scanning program took an average of .02 hour or 1.2 minutes for each pattern. A slight variation in run time was due to the calculation of the determinate which due to the algorithm used varied proportionately with F(F-1) where F is the number of features found.

Although a limited pattern set was used, the time for each pattern searched was estimated to be less than .002 hour (about one sec.) for each pattern in the library.

Ambiguities in Selection

A problem that occurred in the Selection Program was in the comparison of the determinates. In the process of calculating the determinate, (i.e., distance calculation, standardization, determinate calculation) the large number of multiplications using finite arithmetic causes a round off error in the value of the determinate. The solution to this, although not implemented, could have been to use an interval of ± 1% around the value.

CHAPTER V

CONCLUSION

The system developed above was viewed as a theoretical way of removing the limitations inherent in most pattern recognition systems developed to the present time; and as a theory, the application of it would require further refinement. Some of the refinements necessary to apply the theory are discussed below.

Increment Size Reduction

The system up to this point has been discussed from an entirely deterministic viewpoint, i.e., all of the results of the different distributions and comparisons have been assumed to have exact values. This deterministic method holds true up to a certain transition point. The change in the size of a pattern has been taken as a multiple of a distance between adjacent cells in the input matrix. At the transition point the value of the increment of change in size of the pattern approaches zero. In like fashion, the possible rotations of a pattern go from 8 to infinity.

An example of the complexities that arise from the reduction in increment size can be ascertained from a comparison of Figure 4 in Chapter II where the rotation increment as 1 with Figure 6 below with an increment of .5.

$$\underset{1}{\begin{array}{cccc}0&0&1&0\\0&0&1&0\\0&0&1&0\\0&0&0&0\end{array}} \qquad \underset{2}{\begin{array}{cccc}0&0&1&0\\0&0&1&0\\0&1&0&0\\0&0&0&0\end{array}} \quad \text{or} \quad \underset{2}{\begin{array}{cccc}0&0&0&1\\0&0&1&0\\0&0&1&0\\0&0&0&0\end{array}} \qquad \underset{3}{\begin{array}{cccc}0&0&0&1\\0&0&1&0\\0&1&0&0\\0&0&0&0\end{array}}$$

Rotation Index

Figure 6. .5 Increment Rotation

As can be seen, this leads to an ambiguity as exemplified by the middle two features.

The solution to this problem is to increase the number of elements in the input matrix while keeping the overall size in the same proportion to the average size of the patterns so that the lines in the pattern overlap several cells. This means that when the submatrices are scanned across the input matrix, the occurrence of a feature cannot be determined exactly as it was in the deterministic method, but must be determined on the basis of the percentage of matching cells. This percentage is then examined by a probabilistic analysis. If the percentage of matching cells is high; then there is a high probability that the area of the input matrix, being compared to the submatrix, contains the same feature as is in the submatrix. The implementation of this would require a confidence level below which the result of the comparison is no and above which the result is yes.

The comparison of the derived feature type distributions and the feature type rotation distributions with the corresponding distributions in the library would also have to be based on probabilities, possible using a correlation approach.

The determinate of the distance matrix would have to be compared to the determinates in the library on a probabilistic method also.

Resolution

The resolution of a pattern recognition system can be defined as the ability of the system to respond with the correct result to a small increment of change in the size or rotation of a pattern. As an example if the input area were 10 units of measure by 10 units and this area contained a 10 x 10 matrix then the smallest increment of change in size would be one unit. If, however, the input area contained a 1000 x 1000 matrix, then the increment would be .01 units. Therefore the resolution of the 1000 x 1000 matrix is said to be one hundred times greater than the 10 x 10 matrix of the same area.

The resolving ability of this system is important when the system is used in a practical situation because the smallest incremental change recognizable by the system must be less than or equal to the smallest incremental change expected in the size of the patterns to be used.

Practical Applications

While the Size and Position Independent Pattern Recognition System is unsuitable for immediate real time operations such as character recognition equipment due to its multiple steps, it does have a wide range of feasible applications. Some of the possible applications are as follows:

Weather Forecasting. Using a weather map as the input pattern, the system could provide weather forecasts automatically.

Location Finding. With the ship or plane mounted radar set as the input, the system could provide a continuous indication of the location of the vehicle.

In general, any time the situation can be displayed in two dimensions and consists of irregular repeated parts, the Size and Position Independent Pattern Recognition System can be used to analyze and reach conclusions about the data.

BIBLIOGRAPHY

(1) Rosenblatt, Frank, Principles of Neurodynamics, Spartan Books, Washington, D.C. (1962).

(2) Highleyman, W. H. "Further Results on the N-tuple Character Recognition Method." IRE Transactions on Electronic Computers. Vol. EC-10. (March, 1961) p. 97.

APPENDIX A

```
C        SCANNING PROGRAM
C
         DIMENSION ZMAT(35,50),S(35),SUB(4,3,3),T(3,3),TAB(25,4)
         DIMENSION FD(4),FDR(4,8)
         COMMON ZMAT
     500 FORMAT(911)
     501 FORMAT(13,3511)
     505 FORMAT(A6)
     700 FORMAT(A6,F5.0)
     701 FORMAT(A6,8F5.0)
     702 FORMAT(A6,E16.8)
C
C        READ IN FEATURES
         DO 5 K=1,4
       5 READ(5,500)((SUB(K,I,J),I=1,3),J=1,3)
C
C        READ IN PATTERN
         READ(5,505)IDEN
      10 READ(5,501)J,(S(I),I=1,35)
         IF(J.EQ.99) GO TO 20
         DO 15 I=1,35
      15 ZMAT(I,J)=S(I)
         GO TO 10
C
C        PHASE 1 SCANNING
      20 NOF=1
         DO 36 NF=1,4
         DO 22 I=1,3
         DO 22 J=1,3
      22 T(I,J)=SUB(NF,I,J)
         DO 35 NR=1,8
         DO 30 I=1,50
         DO 30 J=1,35
         IM=I+2
         JM=J+2
         DO 25 K=I,IM
         IMS=K-I+1
         DO 25 L=J,JM
         JMS=L-J+1
         IF(ZMAT(L,K).NE.T(JMS,IMS)) GO TO 30
      25 CONTINUE
C        TO HERE IF FEATURE MATCHES
         TAB(NOF,1)=NF
         TAB(NOF,2)=NR
         TAB(NOF,3)=J+1
```

```
      TAB(NOF,4)=1+1
      NOF=NOF+1
      IF(NOF.GT.25) CALL EXIT
   30 CONTINUE
      SAV=T(1,1)
      T(1,1)=T(2,1)
      T(2,1)=T(3,1)
      T(3,1)=T(3,2)
      T(3,2)=T(3,3)
      T(3,3)=T(2,3)
      T(2,3)=T(1,3)
      T(1,3)=T(1,2)
      T(1,2)=SAV
   35 CONTINUE
   36 CONTINUE
      NOF=NOF-1
C
C     BUILD FEATURE DISTRIBUTION TABLES
      DO 60 I=1,NOF
      J=TAB(I,1)
      FD(J)=FD(J)+1.
      K=TAB(I,2)
   60 FDR(J,K)=FDR(J,K)+1.
      DO 70 I=1,NOF
      DO 70 J=I,NOF
      ZMAT(I,J)=0.
      ZMAT(J,I)=0.
      IF(I.EQ.J) GO TO 70
      D=SQRT(((TAB(J,3)-TAB(I,3))**2)+((TAB(J,4)-TAB(I,4))**2))
      ZMAT(I,J)=D
      ZMAT(J,I)=D
   70 CONTINUE
      DET=DETER(NOF)
      DO 75 I=1,4
   75 WRITE(7,700)IDEN,FD(I)
      DO 77 I=1,4
   77 WRITE(7,701)IDEN,(FDR(I,J),J=1,8)
      WRITE(7,702)IDEN,DET
      CALL EXIT
      END
```

APPENDIX B

```
      FUNCTION DETER(NOF)
      COMMON ZM(35,50)
C
C     STANDARDIZE THE MATRIX
      ZMAX=0.0
      DO 5 I=1,NOF
      DO 5 J=I,NOF
    5 IF(ZM(I,J).GT.ZMAX)ZMAX=ZM(I,J)
      DO 8 I=1,NOF
      DO 8 J=1,NOF
    8 ZM(I,J)=(25./ZMAX)*ZM(I,J)
      SUM=0.
      DO 15 I=2,NOF
      LIM=NOF + I-1
      PROD=1.
      DO 10 J=I,LIM
      K=J-I+1
      L=J
      IF(J.GT.NOF)L=J-NOF
      PROD=PROD*ZM(L,K)
   10 CONTINUE
      SUM=SUM+PROD
   15 CONTINUE
      DETER=SUM
      RETURN
      END
```

APPENDIX C

```
C       SELECTION PROGRAM
C
        DIMENSION FD(4),FDR(4,8),FDA(4),FDRA(4,8)
    500 FORMAT(A6,F5.0)
    501 FORMAT(A6,8F5.0)
    502 FORMAT(A6,E16.8)
    600 FORMAT(1H1)
    601 FORMAT(1X,14HINPUT PATTERN ,A6,4H IS ,A6,4X,3I6)
    602 FORMAT(1X,14HINPUT PATTERN ,A6,10H NOT FOUND,4X,3I6)
        DATA EN/3HEND/
        DATA IEN/3HEND/
        WRITE(6,600)
        REWIND 0
      1 DO 5 I=1,4
        READ(5,500)IDEN,FD(I)
      5 IF(IDEN.EQ.IEN) GO TO 10
        DO 6 I=1,4
      6 READ(5,501)IDEN,(FDR(I,J),J=1,8)
        READ(5,502)IDEN,DET
        WRITE(0)IDEN,FD,FDR,DET
        GO TO 1
C
C       READ IN PATTERN TO BE RECOGNIZED
     10 WRITE(0)EN,FD,FDR,DET
        ENDFILE 0
     12 REWIND 0
        DO 14 I=1,4
     14 READ(5,500)IDENA,FDA(I)
        DO 15 I=1,4
     15 READ(5,501)INDEA,(FDRA(I,J),J=1,8)
        READ(5,502)IDENA,DETA
        IFR=0
        IF=0
        N=0
     20 READ(0)IDEN,FD,FDR,DET
        IF(IDEN.EQ.IEN) GO TO 50
        N=N+1
C
C       COMPARE FEATURE TYPE DISTRIBUTION
        DO 25 I=1,4
        IF=IF+1
     25 IF(FD(I).NE.FDA(I)) GO TO 20
C
C       COMPARE FEATURE TYPE ROTATION DISTRIBUTION
        DO 40 I=1,8
```

APPENDIX C (CONTINUED)

```
      K=I-1
      DO 35 M=1,4
      DO 30 J=1,8
      L=J+K
      IF(L.GT.8)L=L-8
      IFR=IFR+1
      IF(FDR(M,J).NE.FDRA(M,L)) GO TO 40
   30 CONTINUE
   35 CONTINUE
C     HERE IF MATCH 4 X 8
      GO TO 45
   40 CONTINUE
      GO TO 20
C
C     COMPARE ON DETERMINATE
   45 IF(DET.NE.DETA) GO TO 20
C     PATTERN FOUND
      WRITE(6,601)IDENA,IDEN,N,IF,IFR
      GO TO 12
   50 WRITE(6,602)IDENA,N,IF,IFR
      GO TO 12
      END
```

VITA

Dennis Leroy Cranston

Candidate for the Degree of

Master of Business Administration

Report: SIZE AND POSITION INDEPENDENT PATTERN RECOGNITION

Major Field: Management

Biographical:

Personal Data: The author was born in DesMoines, Iowa on August 8, 1941, the son of Merle L. and Florence L. Cranston. On May 26, 1963, he married Sharon Jean Hanson.

Education: The author graduated from Cushing High School in Cushing, Oklahoma, in 1959. In September, 1959, he entered Oklahoma State University, Stillwater, Oklahoma, where he received a Bachelor of Science Degree in Management in May, 1966. In June, 1966, he was admitted to the Graduate College of Oklahoma State University, Stillwater, Oklahoma, where he completed the requirements for the Master of Business Administration Degree in July, 1967.

Professional Experience: The author was employed by International Business Machines as a Customer Engineer from September, 1961 to February, 1963. From March to December, 1963, he was employed by the Service Bureau Corporation. He was a Systems Analyst at the Oklahoma State University Computing Center from January, 1963 to May, 1967 at which time he served as a Graduate Assistant until May, 1967.

Membership in professional societies: The author is a member of Association for Computing Machinery and Beta Gamma Sigma.

Name: Dennis Leroy Cranston   Date of Degree: July, 1967

Institution: Oklahoma State University   Location: Stillwater, Oklahoma

Title of Study: SIZE AND POSITION INDEPENDENT PATTERN RECOGNITION

Pages in Study: 28   Candidate for Degree of Master of Business Administration

Major Field: Management

Scope and Method of Study: An expansion and modification of the Bledsoe-Browning n-truple technique was developed to remove the size, position and rotation limitations on the presentation of a pattern to the input matrix. The pattern recognition system developed, utilized a multi-step process involving the distributions of the features and the conversion of the absolute locations of the features on the input matrix to relative locations.

Findings and Conclusions: The Size and Position Independent Pattern Recognition System was found to distinguish between similar patterns and to recognize patterns varied in size, position and rotation.

ADVISER'S APPROVAL ______________________