SPACE RESEARCH COORDINATION CENTER

UNIVERSITY OF PITTSBURGH
1787
VERITAS ET VIRTUS

A METHOD OF RECOGNITION OF HAND DRAWN LINE PATTERNS

BY

M. J. ZOBRAK AND T. W. SZE

DEPARTMENT OF ELECTRICAL ENGINEERING

SRCC REPORT NO. 59

UNIVERSITY OF PITTSBURGH
PITTSBURGH, PENNSYLVANIA
22 NOVEMBER 1967

The Space Research Coordination Center, established in May, 1963, has the following functions: (1) it administers predoctoral and postdoctoral fellowships in space-related science and engineering programs; (2) it makes available, on application and after review, allocations to assist new faculty members in the Division of the Natural Sciences or the School of Engineering to initiate research programs or to permit established faculty members to do preliminary work on research ideas of a novel character; (3) in the Division of the Natural Sciences it makes an annual allocation of funds to the interdisciplinary Laboratory for Atmospheric and Space Sciences; (4) in the School of Engineering it makes a similar allocation of funds to the Department of Metallorgical and Materials Engineering; and (5) in concert with the University's Knowledge Availability Systems Center, it seeks to assist in the orderly transfer of new space-generated knowledge in industrial application. The Center also issues periodic reports of space-oriented research and a comprehensive annual report.

The Center is supported by an Institutional Grant (NsG-416) from the National Aeronautics and Space Administration, strongly supplemented by grants from The A. W. Mellon Educational and Charitable Trust, the Maurice Falk Medical Fund, the Richard King Mellon Foundation and the Sarah Mellon Scaife Foundation. Much of the work described in SRCC reports is financed by other grants, made to individual faculty members.

A METHOD OF RECOGNITION OF HAND DRAWN LINE PATTERNS

M. J. Zobrak and T. W. Sze

Department of Electrical Engineering
University of Pittsburgh
Pittsburgh, Pennsylvania

Proceedings of The First Annual Princeton Conference on Information Sciences and Systems, 1967.

A METHOD OF RECOGNITION OF HAND DRAWN LINE PATTERNS

M. J. Zobrak
Department of Electrical Engineering
University of Pittsburgh
Pittsburgh, Pennsylvania

T. W. Sze
Department of Electrical Engineering
University of Pittsburgh
Pittsburgh, Pennsylvania

SUMMARY

A method utilizing the direction of pen motion during sequential increments of time is presented as a means of recognizing hand drawn line patterns. During the time a pattern is drawn, d+1 discrete points are measured. From these, d quantized directions are obtained and used as parameters in a discriminant function. The coefficients, or weights, of the function are determined through a learning algorithm. Experiments were performed using ten different sets of children's writing. The complete English Alphabet, in capital letters, and the numerals comprised each set. These patterns had various degrees of distortion, rotation, and continuity. After a training phase in which all ten sets were used, the error during recognition was 6.0%.

INTRODUCTION

One of the major problems in any pattern recognition system is that of feature extraction. In most line pattern recognition systems, feature extraction involves data processing of input coordinate points to produce features such as length of line pattern, center of gravity, curvature, etc. In this paper feature extraction is simplified since only one type of feature, based on approximated spatial derivative information, is extracted. The spatial derivatives are resolved into directions, coded with "Direction Numbers", and assembled into a feature vector. The element positions of particular features in the vector coincide with the time sequence in which the features were extracted.

DIRECTION TIME FEATURES

In order to understand how the features are extracted, it is helpful to briefly consider the construction of the input device.[1] A pad simulates a portion of the first quadrant of a rectangular coordinate system. As a pen moves along the pad, the x,y coordinates of the pen are stored into successive locations of a memory system each time the coordinates change. The points in Figure 1(a) can be considered as being on the pad. If a pen is placed at point 1 and moved through points 3 and 5 to point 7, as in Figure 1(b), a direction sequence is obtained with elements "down, right, up". If the points 7, 5, 1, 3, 7 are traversed in that order, as in Figure 1(c), the sequence obtained is

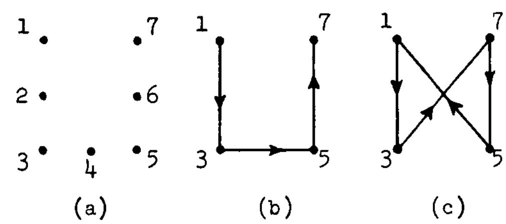

Figure 1. Direction sequence illustration

"down, up left, down, up right". The directions down, right, etc. are quantized into sectors and each sector is represented by a number. Figure 2 illustrates this idea with eight quantized sectors. Each sector is 45 degrees wide and sectors 0 and 4 are centered about the

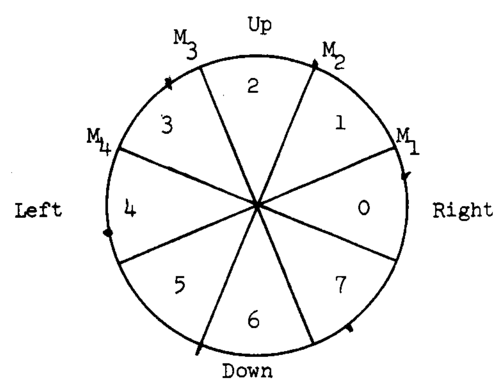

Figure 2. Quantization of directions

horizontal. Applying this numerical representation of directions to the direction sequence "down, right, up" results in the sequence "6, 0, 2", where each number will be called a direction number. Similarly, the sequence "down, up left, down, up right" becomes the sequence of direction numbers "6, 3, 6, 1". This last sequence can be used to represent the pattern of Figure 1(c) when drawn as indicated. It cannot be used to represent a pattern whose final appearance is the same as that in Figure 1(c) but whose stroke sequence is different. This is a seeming disadvantage of the method. It has been verified, however, that there is a considerable amount of invariance in the order of strokes used in constructing a hand-drawn line pattern[2]. Furthermore, in some applications the sequence of strokes in drawing the patterns is intended to be invariant[3].

Since, in fact, a large number of coordinate points are stored for any typical pattern (the coordinate grid is structured with 100 lines per inch in the device under consideration), it is evident that any pattern drawn on the pad can be represented as precisely as desired by using a sequence of direction numbers (within the resolving ability of the coordinate grid and quantization shown in Figure 2)--the precision increasing as the number of direction numbers increases. For example, using the additional points 2, 4, and 6 in Figure 1(a). It is also evident that such a sequence represents the direction in which the pen was moving during successive increments of time. It therefore can be considered a feature vector with direction numbers as elements--a direction-time feature vector. Hereafter, such vectors will be denoted as

$$\underline{X} = [x_1, x_2, \ldots, x_j, \ldots, x_d]^t \qquad (1)$$

where

x_j is a direction number indicating the direction of motion of the pen in the jth increment of time.

THE RECOGNITION METHOD

The development which follows assumes the existence of a prototype vector for each of the recognizable categories. Each prototype is the standard for its category. The simplest way of obtaining such vectors is to construct well drawn patterns and then store their feature vectors as the prototype vectors. The ith recognizable pattern category is denoted by χ_i while its prototype vector is denoted by $\underline{P}_i$.

Normalization

As implied previously, a prime consideration in developing this method was simplicity. Accordingly, only size normalization is done. A simple scheme, based on the fact that the number of data points is indicative of pattern size, is used. For example, drawing the letter "L" one inch high might result in approximately 180 coordinate points stored in the memory system whereas an "L" two inches high would have approximately 360 coordinate points stored.

Normalization could be achieved if some number of points, less than the total number stored, were used to represent the pattern. The number chosen must be one greater than d, the number of elements desired in the feature vector. Furthermore, the points must be chosen such that they are in fact representative of the pattern. A promising, convenient, and simple choice is to include the first and last coordinate points and d-1 approximately equally spaced points in between.

Extracting Features

If the d+1 points representing a size normalized pattern are considered in the same time sequence in which they were traversed during the drawing of the pattern, a d dimensional direction-time feature vector can be developed. Let the sequence $s_1, s_2, \ldots, s_d, s_{d+1}$ represent the points. This implies that the pen first contacted the pad at s_1. Then possibly after moving some distance, it passed over s_2 and similarly for $s_3, \ldots, s_d$. Finally, the pattern was completed with the pen stopping at s_{d+1}. No restriction is made on the number of times the pen may have left the pad and returned in between s_1 and s_{d+1}. Lack of this restriction suggests a need for a signal from the user to the system to indicate that a pattern has been completed and recognition should commence. In order to extract feature x_j, the slope,

$$m_j = (a_j - a_{j+1})/(b_j - b_{j+1}) = N_j/D_j$$

of an imaginary line drawn from s_j to s_{j+1} is obtained and quantized utilizing the slopes of the four lines in Figure 2. The quantization can be expressed as

$$x_j = 0 \text{ or } 4 \text{ if } M_4 < m_j \leq M_1$$

$$x_j = 1 \text{ or } 5 \text{ if } M_1 < m_j \leq M_2$$

$$x_j = 2 \text{ or } 6 \text{ if} \begin{Bmatrix} M_2 < m_j < \infty \\ \infty < m_j \leq M_3 \end{Bmatrix}$$

$$x_j = 3 \text{ or } 7 \text{ if } M_3 < m_j \leq M_4$$

Discrimination between the two possible choices for x_j is accomplished using the following relationships:

$$x_j = 0, 1, 2, 6, \text{ or } 7 \quad \text{if } D_j < 0$$

$$x_j = 0, 1, 2, 3, \text{ or } 4 \quad \text{if } N_j < 0$$

After d passes through this decision structure, the direction-time feature vector $\underline{X}$ is obtained.

Recognition

Frequently recognition is based on a set of linear discriminant functions of the form $\{g_i(\underline{X}): \; g_i(\underline{X}) = \underline{W}_i \cdot \underline{X}, \; i = 1, 2, \ldots, R\}$. To be effective such functions require the elements of $\underline{X}$ to possess an ordering relationship similar to the real numbers. Since in the method presented here the elements of $\underline{X}$ are in fact elements of Z_8, the set of integers mod 8, such functions do not suffice (i.e. the ordering has zero in between seven and one). Instead, a set of functions of the form

$$\left\{g_i(\underline{P}_i, \underline{X}): \;\; g_i(\underline{P}_i,\underline{X}) = \underline{W}_i \cdot \underline{\phi}(\underline{P}_i,\underline{X}), \; i=1,2,\ldots,R\right\}$$

is used where g_i is a metric in the d dimensional space of the feature vectors. The output is determined by $\min_i \left\{g_i(\underline{P}_i,\underline{X})\right\}$. Clearly, g_i is a metric if the individual components in the linear combination which comprises g_i are metrics. Accordingly, if the jth elements of $\underline{W}_i$, $\underline{P}_i$, and $\underline{X}$ are denoted by w_{ij}, p_{ij}, and x_j. Then one can write

$$g_i(\underline{P}_i,\underline{X}) = \sum_{j=1}^{d} w_{ij} \; \phi_j(p_{ij},x_j).$$

If the restriction that at least one of the $w_{ij} > 0$ is imposed, then g_i is a metric if ϕ_j is a metric. When ϕ_j is defined as

$$\phi_j(p_{ij},x_j) = \begin{cases} |p_{ij} - x_j| & \text{if } |p_{ij} - x_j| \leq 4 \\ |8 - p_{ij} - x_j| & \text{if } |p_{ij} - x_j| \quad 4 \end{cases}$$

its metric properties are easily obtained.

Training

While the function $\underline{\phi}(\underline{P}_i,\underline{X})$ is itself worthy of some note as a discriminant function, the errors it produces can be decreased by a suitable selection of weight vectors.* Some strides have been taken toward determining these weight vectors when assumptions can be made about the parameters which characterize the probability distributions of the patterns.[4] When no such assumptions can be made non-parametric methods must be employed. Non-parametric methods involve attempts to recognize known patterns presented sequentially to the system. After each attempt the weight vectors are either modified if the response was incorrect or not modified if the response was correct. In the method employed here for example, if $\underline{X}$ is the kth pattern in the sequence and it is known to be in category i and $\underline{W}_i \cdot \underline{\phi}(\underline{P}_i,\underline{X}) < \underline{W}_j \cdot \underline{\phi}(\underline{P}_j,\underline{X})$, $j=1,2,\ldots,R$: $j \neq 1$, then $\underline{X}$ is correctly recognized as being in $\mathcal{X}_i$ and no modification takes place. If, however, $\underline{X}$ is known to be in category i and $\underline{W}_j \cdot \underline{\phi}(\underline{P}_j,\underline{X}) \leq \underline{W}_i \cdot \underline{\phi}(\underline{P}_i,\underline{X})$ for j in some non-empty set of integers J, then the set of weight vectors $\left\{\underline{W}_h: \; h=1,2,\ldots,R\right\}$ is modified according to the following training rule:

$$\underline{W}_h(k+1) = \begin{cases} \underline{W}_i(k) - c(k)\underline{\phi}(\underline{P}_i,\underline{X}(k)) & h=i \\ \underline{W}_j(k) + \alpha_j(k)c(k)\underline{\phi}(\underline{P}_j,\underline{X}(k)) & h=j, \; j \in J \\ \underline{W}_h(k) & \text{otherwise} \end{cases}$$

where k denotes the number of modifications, $c(k) > 0$, $\alpha_j(k) > 0$, $\sum \alpha_j(k) = 1$, and at least one $w_{hm}(k) > 0$. $w_{hm}(k)$ is the mth element of $\underline{W}_h(k)$.**

The proof that this training rule results in a set of weight vectors which can correctly recognize all the patterns used in training, if such a set exists, is given in reference 8.

EXPERIMENTS

The method was simulated on an IBM 7090 computer using the MAD language. Lack of the aforementioned input device necessitated drawing the patterns on graph paper and storing the X,Y coordinate data on punched cards.

An arbitrary, interesting, and convenient choice for the patterns to be recognized was the set of alphanumeric characters with no distinction between "O" and zero. This placed R, the number of patterns recognizable by the system, at 35. More intuitive than arbitrary was the selection of d, the dimension of the feature vector, as 10 and the direction quantization as shown in Figure 2. Also, for simplicity $c(k) = 1/8$, $\alpha_j(k) = 1$ for the j which denotes the minimum of $\left\{\underline{W}_j \cdot \underline{\phi}_j \; : \; j \in J\right\}$, and all other $\alpha_j(k) = 0$, for all k.

Data

Data for the program consisted of 10 sets of alphanumeric characters. Representative samples are shown in Figure 3. Set 1 was reproduced from a leading text on handwritting[6], and was used as the prototype set. The sequence of strokes used in drawing each of the other sets was the same as that shown on the prototype set.

Results

In one experiment the program trained on two sets of characters (Sets 2 and 3), and then it attempted to recognize all 10 sets. Initially all weight factors were set equal to unity. The results are viewed with mixed emotions. For although the weights were modified such that Sets 2 and 3 were recognizable with no errors

---

* With all elements of all weight vectors set equal to one the error rate is 22%.

** Duda and Fossum[5] describe a similar non-parametric method. Their discriminant functions are of the form $g_i(\underline{X}) = \underline{W}_i \cdot \underline{X}$ and their training rule is applied when, for some $j \in J$, $\underline{W}_j \cdot \underline{X} \geq \underline{W}_i \cdot \underline{X}$. The training rule is

$$\underline{W}_h(k+1) = \begin{cases} \underline{W}_i(k) + c(k) \, \underline{X}(k) & h=i \\ \underline{W}_j(k) - \alpha_j(k)c(k)\underline{X}(k) & h=j, \; j \in J \\ \underline{W}_h(k) & \text{otherwise} \end{cases}$$

Set 1

Set 2

Set 3

Set 4

Set 5

Set 6

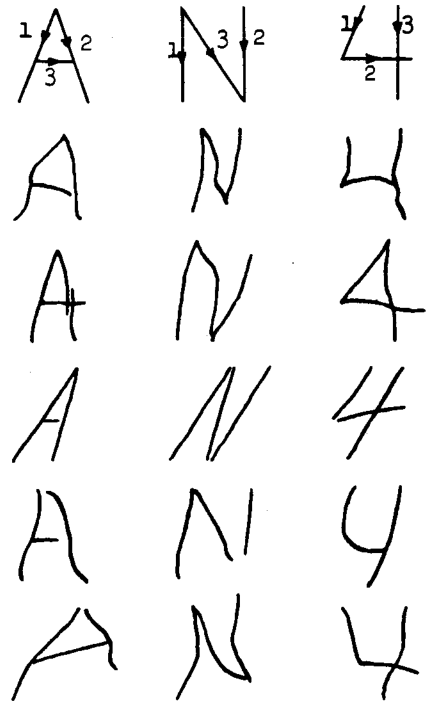

Figure 3. Samples of patterns used in experiments

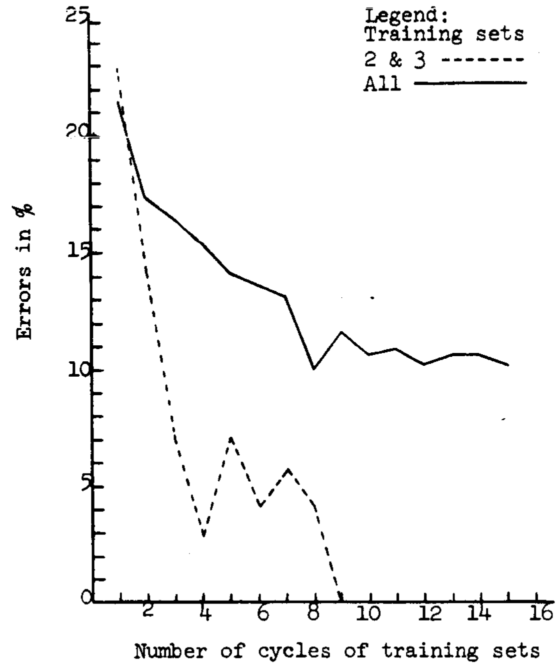

Figure 4. % error during training phase

after the system had observed them nine times, the error rate after attempting to recognize the 350 patterns in the 10 sets (using the modified weights) was considerably higher at 16%***. The convergence of the error rate during the training phase is shown in Figure 4.

More significant and gratifying results were obtained when the program was allowed to train on all 10 pattern sets. Again all weight factors were initially set at unity. Then after having observed each of the 350 patterns 15 times, an attempt was made to recognize them. The error rate for this attempt was 6.0%. Figure 4 depicts the convergence of the error rate during the learning phase.

The execution time of the IBM 7090 in recognizing the 350 patterns was 87 seconds. This implied a recognition rate of one pattern per 0.25 seconds.

*** A correct response is considered as any response which correctly identifies the unknown pattern and only that pattern. Any other response is an error.

CONCLUSIONS

Aside from the basic conclusions that pattern recognition is feasible using direction-time feature vectors, it is evident that use of a more efficient machine language program, as opposed to the compiled version used here, and the faster computers available today should pr duce a significant increase in speed. Accordingly, it can also be concluded that use of direction-time feature vectors permits rapid recognition of patterns.

The fact that the method is fast, simple, and able to recognize reasonably well drawn alphanumeric characters with a low error rate, suggests its use as a man-machine communicating device. The particular application in mind is that in which a user takes a few moments to teach the system his style of drawing the alphanumerics, mathematical symbols, etc., then proceeds to write, edit, or insert data into a computer program. Such an application suggests several areas for extending this research.

First, it could not have gone unnoticed that a great many parameters were set arbitrarily or intuitively, i.e. the number of sectors, the slopes of the lines for extracting features and the number of elements in the feature vector. A method of optimally determining some or all of these parameters based on the user's style of drawing seems both beneficial and interesting to pursue. Second, the development of functional units capable of performing the functions implied in the method also appears to be a challenging and worthwhile task. Finally, the learning scheme appears to converge slowly. This may or may not be a function of the patterns used. Nevertheless, in view of the fact that learning methods with varying rates of convergence exist for other discriminant functions[7], it seems reasonable to conduct further investigation in this area.

REFERENCES

1. M. R. Davis and T. O. Ellis, "The RAND tablet: A man-machine graphical communication device", AFIPS Joint Computer Conference Proceedings, Vol. 26, pp. 325-331, December, 1964.

2. W. Teitelman, "New methods for real time recognition of hand drawn characters", Bolt, Beranek, and Newman, Inc., Report No. 10105 June, 1963.

3. M. J. Zobrak, "A method for rapid recognition of hand drawn line patterns", M.S. Thesis University of Pittsburgh, Pittsburgh, Pennsylvania, April 1966.

4. G. S. Sebestyen, Decision Making Processes in Pattern Recognition. New York: MacMillan 1962.

5. R. O. Duda and H. Fossum, "Pattern Classification by interatively determined linear and piecewise linear discriminant functions," IEEE Trans. on Electronic Computers, Vol. EC-15, pp 220-232, April 1966.

6. Peterson Handwriting System, Greensburg Pennsylvania: Peterson Handwriting Company, 1960.

7. N. J. Nilsson, Learning Machines: Foundations of Trainable Pattern Classifying Systems. New York: McGraw-Hill, 1965.

8. C. C. Li, T. L. Teng, M. J. Zobrak, and T. W. Sze, "On Recognition of Handwriting Chinese Characters", First Annual Princeton Conference on Information Sciences and Systems, 1967.