Research and Development Technical Report
ECOM-01901 29

GRAPHICAL-DATA-PROCESSING RESEARCH STUDY AND EXPERIMENTAL INVESTIGATION

SEVENTH QUARTERLY REPORT

By

R. O. Duda P. E. Hart J. H. Munson

FEBRUARY 1968

ECOM
UNITED STATES ARMY ELECTRONICS COMMAND · FORT MONMOUTH, N.J. 07703
CONTRACT DA 28-043 AMC-01901(E)
STANFORD RESEARCH INSTITUTE
Menlo Park, California 94025

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

The citation of trade names and names of manufacturers in this report is not to be construed as official Government endorsement or approval of commercial products or services referenced herein.

Disposition

Destroy this report when it is no longer needed. Do not return it to the originator.

TECHNICAL REPORT ECOM-01901-29 FEBRUARY 1968

GRAPHICAL-DATA-PROCESSING RESEARCH STUDY AND EXPERIMENTAL INVESTIGATION

REPORT NO. 29, SEVENTH QUARTERLY REPORT

1 SEPTEMBER TO 30 NOVEMBER 1967

SRI Project 5864

CONTRACT DA 28-043 AMC-01901(E)

Continuation of Contract DA 36-039 AMC-03247(E)

Prepared by

R. O. DUDA P. E. HART J. H. MUNSON

STANFORD RESEARCH INSTITUTE
MENLO PARK, CALIFORNIA 94025

For

U.S. ARMY ELECTRONICS COMMAND, FORT MONMOUTH, N.J. 07703

ABSTRACT

This report describes the continuing development of scanning, preprocessing, character-classification, and context-analysis techniques for hand-printed text, such as computer coding sheets in the FORTRAN language.

Both edge-detection and topological preprocessing are coupled with classification by a learning machine and used to process a large file of characters printed by a single author. The two systems are combined to achieve a recognition rate considerably better than our previous results. No other comparable results on unconstrained hand printing with a full alphabet are known to us.

The same methods are also applied to a well-known file of hand-printed characters collected by Highleyman. The combination of preprocessing and classification methods has achieved performance better than that reported for any other recognition system.

CONTENTS

ABSTRACT ii

LIST OF ILLUSTRATIONS iv

I INTRODUCTION 1

II INTRA-AUTHOR EXPERIMENTS ON THE JM DATA FILE 2

A. Introduction 2

B. The JM Data File 2

C. Legibility of the JM Text 3

D. TOPO 3-CALM Experiment 1 4

E. PREP-CALM Experiment 11 6

F. PREP-CALM Experiment 12 6

G. PREP-CALM Experiment 12A 8

H. PREP-CALM Experiment 13 8

III EXPERIMENTS ON THE JM DATA FILE WITH COMBINED CLASSIFIERS . 10

A. Introduction 10

B. TOPO 3-CALM Experiment 1 and PREP-CALM Experiment 11 Combined 10

C. TOPO 3-CALM Experiment 1 and PREP-CALM Experiment 12A Combined 11

D. Summary 12

IV EXPERIMENTS WITH HIGHLEYMAN'S DATA 13

A. Introduction 13

B. Nearest-Neighbor Classification 14

C. Edge-Detecting Preprocessing and Piecewise-Linear Classification 15

D. Human Performance 16

E. Conclusions 17

REFERENCES 19

DD FORM 1473

ILLUSTRATIONS

Fig. 1 A Fragment of Hand-Printed Text from a Single Author . . . 4

Fig. 2 Learning Curves for TOPO 3-CAIM Experiment 1 5

Fig. 3 Learning Curves for PREP-CAIM Experiment 11 7

Fig. 4 Tradeoff Curve for Combined Systems on Single-Author Date 11

I INTRODUCTION

This report describes the continuing development of scanning, preprocessing, character-classification, and context-analysis techniques for hand-printed text. The particular subject matter of our investigation is hand-printed FORTRAN text on standard computer coding sheets, with a 46-character alphabet. The reader is referred to the previous reports of this project for background and supplementary material.

In Sec. II, we describe a single author's file of 2,999 hand-printed characters, used to continue the intra-author recognition experiments beyond the preliminary experiments described in the last Quarterly Report. The TOPO 3-CALM and PREP-CALM preprocessor-classifier systems were applied to this file, and performance was observed far exceeding any previously seen in multi-author experiments.

In Sec. III, we show the results of combining the action of the systems treated in Sec. II. The combined system recognized independent test data with 97-percent accuracy and no rejects. This is our best recognition score to date, and we know of no comparable results reported for the recognition of unconstrained hand printing with a full alphabet.

A collection of experiments on a well-known set of hand-printed data collected by Highleyman is described in Sec. IV. The PREP-CALM system performed considerably better than any of several previously reported methods, none of which involved extensive preprocessing of the data.

II INTRA-AUTHOR EXPERIMENTS ON THE JM DATA FILE

A. Introduction

In Sec. III of the preceding Quarterly Report, we described several limited experiments on hand printing from a single author. These experiments indicated a great reduction in error rate, compared to the rates obtained in multiple-author experiments to date. We concluded that "The results of these experiments should be considered somewhat tentative ... the test samples were statistically small ... the data were taken from coding sheets in which 20 alphabets were written on successive lines at one sitting."

We have now performed the follow-on experiments pointed to in the preceding report, using a large file of data including training and test data from actual coding sheets. These experiments have borne out the dramatic improvement in recognition-test error rate suggested by the earlier experiments.

B. The JM Data File

The JM data file consisted of 2,999 characters in the 46-category FORTRAN alphabet, hand-printed by John Munson. This author was chosen as the source of the file because of the existence of a number of actual coding sheets prepared by him on the proper forms during the development of SDS 910 FORTRAN computer programs.

The first 920 character patterns in the file were the 20 alphabets (Sequence Nos. 50-69) used for the previously reported intra-author experiments. Added to these were 2,079 characters gathered from four separate coding sheets, written at different times over a period of a few months. Each line on a coding sheet was given a unique sequence number, ranging from 1,000 to 1,111.

The first five alphabets in the file (Sequence Nos. 50-54, patterns 1-230) were reserved for possible testing but were not used. The training set contained 1,727 patterns. It consisted of the remaining 15 alphabets

(Sequence Nos. 55-69, patterns 231-920) and 1,037 characters of text (Sequence Nos. 1,000-1,056, patterns 921-1,957). The test set contained 1,042 characters taken from two coding sheets (Sequence Nos. 1,057-1,111, patterns 1,958-2,999). About one-third of the test data came from the same sheet as some of the training data; the remainder came from a separate sheet that was written separately from any of the training data.

The inclusion of the hand-printed alphabets in the training data ensured that each of the 46 character types would be represented. The character types were not evenly represented in the text material. Their appearance was determined fortuitously by the text that happened to be chosen.

The same training and test sets were employed throughout the several experiments to be described.

C. Legibility of the JM Text

A fragment of the actual test data is shown in Fig. 1. It may be seen that the printing is fairly legible; it is by no means highly regular. The printing was done with a little care, but with no labored attention to the quality of individual characters. The coder was actually preparing a program text for keypunching, although aware that the sheet might some day be used in recognition experiments. Thus, although the test data were not completely "candid" data, they were generated under conditions that closely model a system in which workers were preparing material for machine input.

Ten human subjects were asked to classify the test set characters, which were presented (in random order) in quantized form on the cathode ray tube attached to the SDS 910 computer. The average error rate was 0.72 percent; assuming a normal distribution of scores, the "true" error rate was 0.72 ± 0.17 percent with 95-percent confidence. (If the 10 responses for each character were used to reach a group decision, only 2 errors [0.2 percent] were made. This would indicate that the individual errors were largely uncorrelated.) These rates do not include the few typographical errors made by the subjects in typing their responses. The rates also do not include six patterns found to be mislabeled in the test data file.

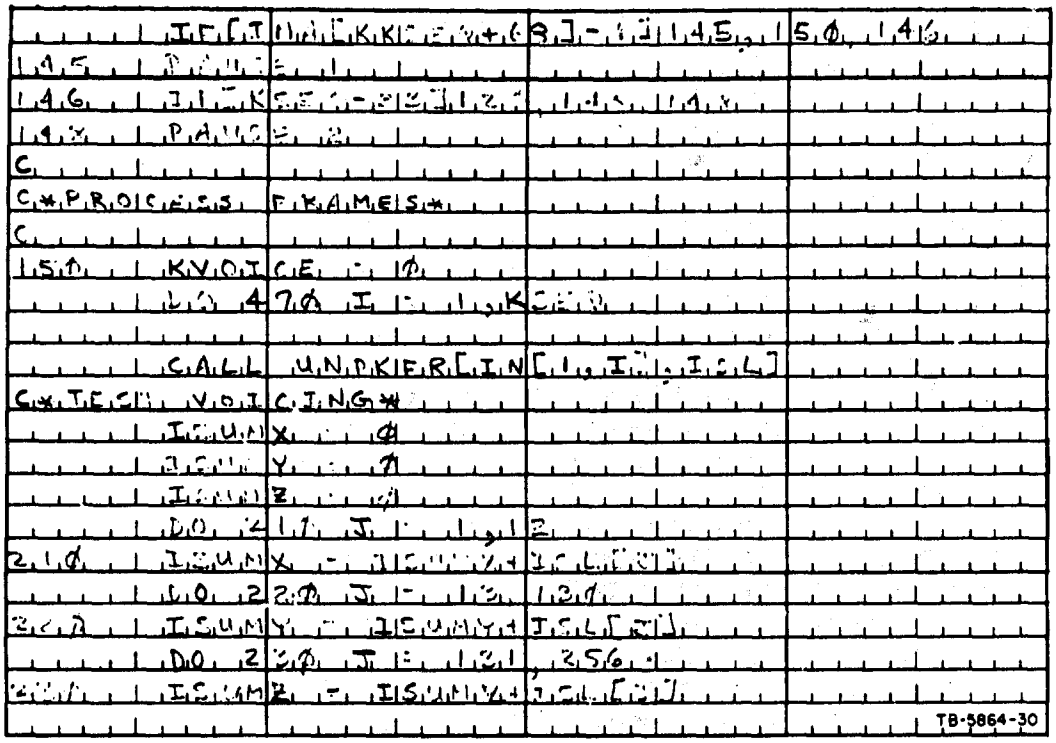

FIG. 1 A FRAGMENT OF HAND-PRINTED TEXT FROM A SINGLE AUTHOR

D. TOPO 3-CALM Experiment 1

The complete JM file was preprocessed by the SDS 910 computer program TOPO 3. TOPO 3 was a minor revision of the TOPO 2 program that has been described in earlier reports. TOPO 3 was rearranged to make it run considerably faster than TOPO 2, and the set of features in the output feature vector was slightly different. For all practical purposes, however, the topological features produced by TOPO 3 (describing character enclosures, concavities, stroke tips, profiles, size, and so on) were the same as those from TOPO 2.

The output feature vectors from TOPO 3 were processed by the CALM learning-machine simulation, which implemented a 16-category linear machine. The training and test sets defined above were used.

The results of the experiment are shown in Fig. 2. For the first 5 iterations, training was performed only on the 690 characters from the 15 alphabets within the training set. Thereafter, the full training set was

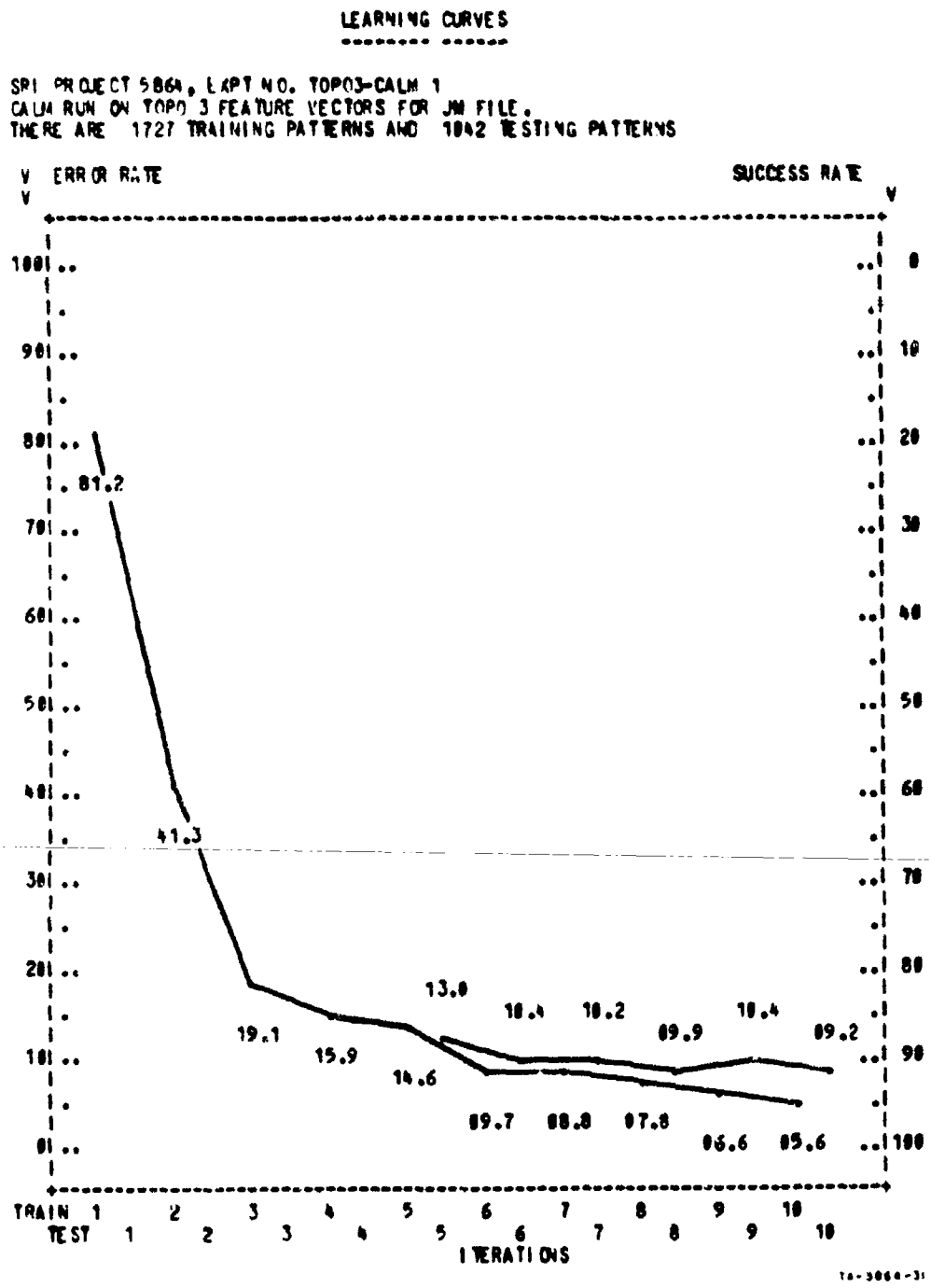

FIG. 2 LEARNING CURVES FOR TOP03-CALM EXPERIMENT 1

used. Test readings were not taken until after the fifth training iteration. At the time of the first test reading (Iteration 5), the machine had only been trained on characters from the alphabets. The test error rate dropped from 13 percent to 10 percent between the fifth and sixth iterations, owing to the expansion of the training set to include the text characters.

The training error rate reached 6 percent in 10 iterations, and the test error rate reached 9 percent. The test error rate was approximately the same as that of TOPO 2-CALM Experiment 4, described in the preceding Quarterly Report, in which much smaller training and test sets were used. The larger amount of training data compensated for the increase in difficulty of recognizing characters from text on actual coding sheets, compared with the characters in alphabets.

E. PREP-CALM Experiment 11

The JM file was preprocessed by the computer simulation of the edge-detecting preprocessor, PREP 24A. In this run, the patterns were only preprocessed in one view. The resulting feature vectors were presented to CALM for processing in PREP-CALM Experiment 11.

The results are shown in Fig. 3. As in Experiment TOPO 3-CALM 1, only the 690 alphabet patterns were used for training in the first 5 iterations, and the full training and test sets were used thereafter. The training error rate reached 1 percent; the test error rate reached 12 percent.

F. PREP-CALM Experiment 12

In PREP-CALM Experiment 12, PREP 24A was used to preprocess each hand-printed pattern in nine different views. The advantage of nine-view over one-view preprocessing with the edge-detecting masks has been shown in other experiments previously reported during this project.

In running CALM on the nine-view preprocessed feature vectors, nine training iterations were first performed over the entire training set. During this sequence of iterations, each view of each training pattern was presented once for training. The test patterns were then presented for nine-view testing. In this case, the classification was done "by

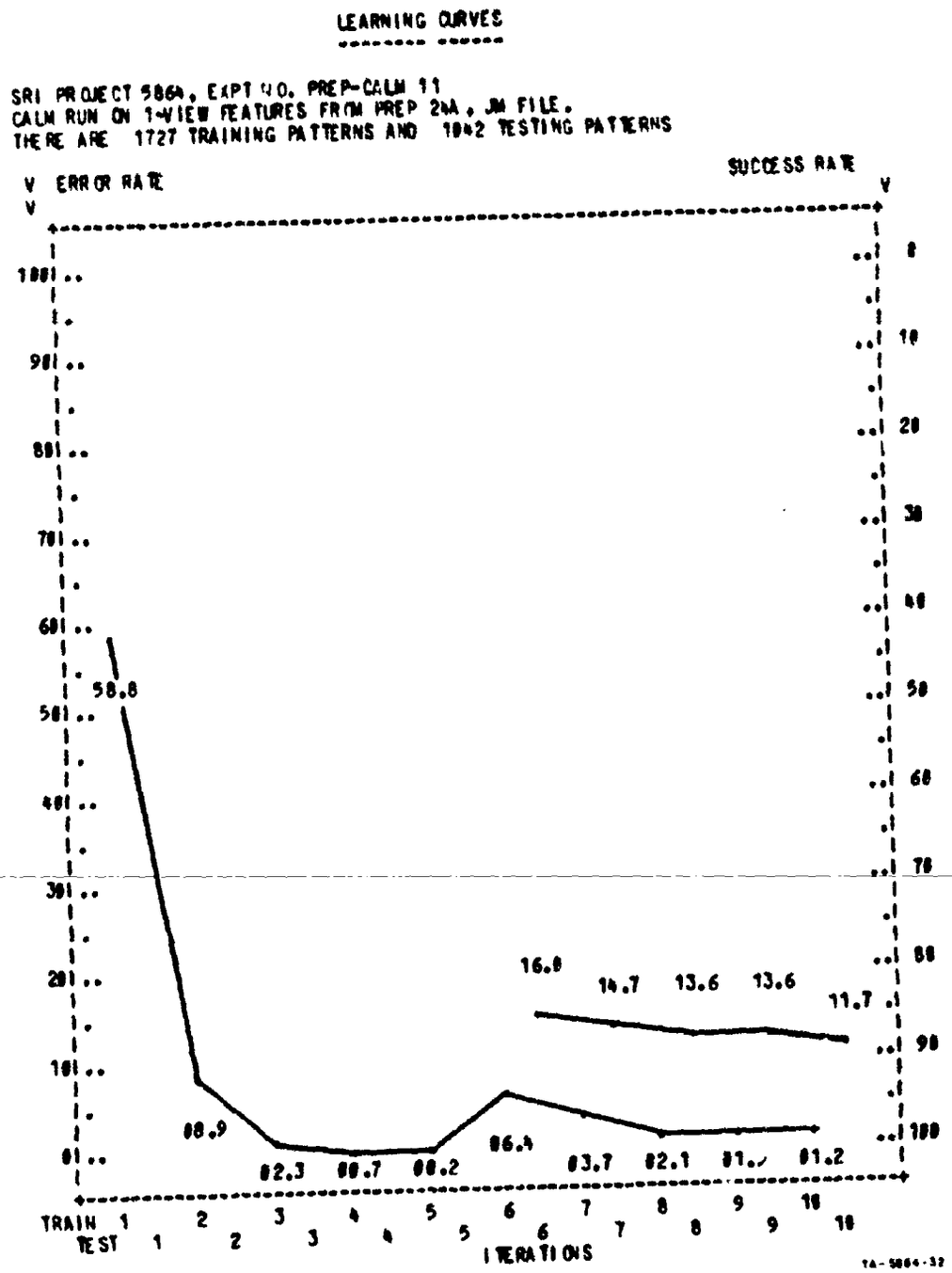

FIG. 3 LEARNING CURVES FOR PREP-CALM EXPERIMENT 11

views." As each view was presented, the learning machine was forced to make a category decision. A vote was taken among the nine single-view decisions to produce the final decision.

The sequence of nine training iterations followed by a nine-view test iteration was repeated three times. The following results were obtained:

	Iteration 9	Iteration 18	Iteration 27
Training error rate	15%	11%	10%
Nine-view test error rate	6%	6%	5%

G. PREP-CALM Experiment 12A

Nine-view classification "by categories" is an alternative to classification "by views." In classification by categories, an accumulator register is employed for each category. The registers are initially zeroed and, as each view is presented, the Dot Product Unit sums are added into the registers for the corresponding categories. After all views have been presented, the character is assigned to the category with the largest accumulated total. We have noted previously that the two methods of multi-view classification yield comparable results (Report No. 22, the final report for Contract DA 36-039 AMC-03247(E), page 46).

In PREP-CALM Experiment 12A, the weights of the trained learning machine from Experiment 12 (at the 27th iteration) were reused. The CALM program was slightly modified to classify the test patterns by categories, instead of by views. The resultant test error rate was 4 percent, versus 5 percent for the former experiment.

H. PREP-CALM Experiment 13

PREP-CALM Experiment 13 was motivated by the following observation concerning nine-view testing by categories: the result obtained by presenting nine different feature vectors (views) and accumulating the DPU sums can also be obtained by adding together the nine feature vectors, component by component, and presenting the result as a single feature vector. In other words, it makes no difference whether the data representing the nine views are added together at the feature-vector level or at the DPU sum level. This is a consequence of the linear nature of the DPU read operation.

The question arises: What would be the effect of applying this change in policy to the training patterns as well as the test patterns? In order to answer this question, we accumulated the nine feature vectors for each pattern in the .IM file into a single feature vector. (Because the original feature vectors has binary components of +1 and -1, the new vector had components ranging from -9 through +9.)

The accumulated feature vectors, arranged into the usual training and test sets, were used as input to CAIM. In 10 iterations, the training error rate reached 0.8 percent. The test error rate was 7 percent after 5 iterations and 9 percent after 10 iterations.

In view of the equivalence we have just described, the test patterns for Experiment 12A and for Experiment 13 are effectively identical. The poorer performance in the latter experiment must be a result of the different training histories. We hypothesize that the separate presentation of each view forces the learning machine to "train harder," intuitively speaking--that more mileage is obtained from the data because each view represents a separate pattern to challenge the machine.

Thus we have observed that the best performance is obtained by grouping all the views of the test pattern together and testing by categories, while using the views of the training patterns seperatly.

III EXPERIMENTS ON THE JM DATA FILE WITH COMBINED CLASSIFIERS

A. Introduction

The technique of combining the TOPO-CALM preprocessor-classifier system with the PREP-CALM system, in order to reduce the classification error rate, was anticipated in the Second Quarterly Report of this project (pages 9-10). The technique was first tested on the limited sample of single-author data used for our first intra-author experiments, and it gave a definite improvement in performance (Sixth Quarterly Report, pages 15-16). We have now applied the technique to a more adequate set of test data, namely, the test data from the JM file described in the preceding section of this report. The TOPO-CALM system was combined with both the one-view and nine-view versions of the PREP-CALM system.

B. TOPO 3-CALM Experiment 1 and PREP-CALM Experiment 11 Combined

The first combined experiment was performed by adding together the learning-machine responses for the test patterns from TOPO 3-CALM Experiment 1 and those from PREP-CALM Experiment 11. For each test pattern, the two Dot-Product-Unit sums in each of the 46 categories were added to form a new set of 46 sums on which the classification decision was to be based. Prior to the addition, the sets of sums from the two experiments were scaled by an empirically determined scale factor so that they would have approximately the same overall range of values and neither set would overwhelm the other in the addition.

The test error rate using the combined sums was 4 percent. This value may be compared with those from the two experiments using the individual machine combinations, namely, 9 percent (TOPO 3-CALM Experiment 1) and 12 percent (PREP-CALM Experiment 11)

Combining the two preprocessor-classifier systems in parallel is evidently a powerful method for improving performance. The improvement implies that the particular errors made by one system are to a considerable degree independent of the errors made by the other--otherwise, the combined system would behave much like either of the individual ones.

C. TOPO 3-CALM Experiment 1 and PREP-CALM Experiment 12A Combined

In order to combine the learning-machine responses to the test data of TOPO 3-CALM Experiment 1 with those of PREP-CALM Experiment 12A, it was necessary to condense the nine-view responses of the latter to a single response. This was done by using the accumulated Dot-Product-Unit sums (formed during the classification-by-categories process) to represent the response of the nine-view PREP-CALM system to the pattern as a whole. Obtaining the accumulated sums for this purpose was, in fact, the prime motivation for performing Experiment 12A.

The accumulated sums from the PREP-CALM system were scaled and added to the sums from the TOPO-CALM system, just as in the other combined experiment described above. Using the combined sums as the basis of classification, we observed an error rate of 3 percent. This compares with test error rates of 9 percent for the TOPO 3-CALM system alone and 4 percent for the nine-view PREP-CALM system alone.

By examining the distribution of the difference between the largest and the second largest combined sums, we obtained a tradeoff curve of errors vs rejects for the combined system. This curve is presented in Figure 4. If the reject margin of the combined machine were set, for example, to reject 3 percent of the test patterns, the error rate would be reduced to 1-1.2 percent. Beyond this point, the rate of return (in terms of error reduction) diminishes.

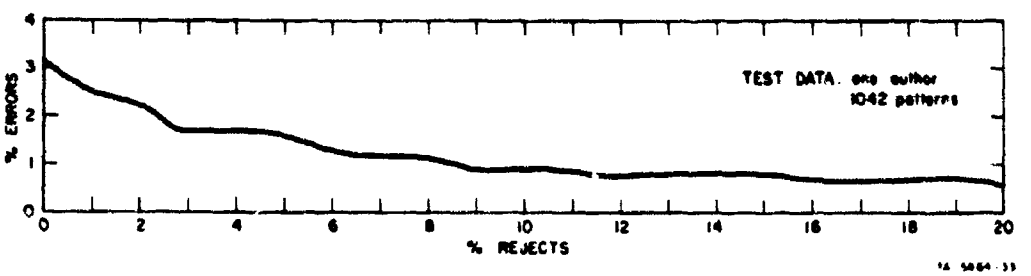

FIG. 4 TRADEOFF CURVE FOR COMBINED SYSTEMS ON SINGLE-AUTHOR DATA

D. Summary

The performance just described is by far the best performance that we have achieved to date on a significantly large body of hand-printed data. To our knowledge, no reported experiments or operational systems have achieved comparable performance on relatively unconstrained hand printing with a full alphabet. Let us summarize briefly the factors pertinent to this result.

We have attempted elsewhere in this report to indicate the quality of the hand-printed test data. We would suggest that the quality is comparable to that expected of data prepared by workers for machine input if the workers were reasonably motivated but had no particular training in forming characters and observed no detailed constraints. The addition of such training and constraints should reduce the variability of the printing to a level so low that the same recognition system would experience an error rate of much less than 3 percent. This approach may be necessary for systems in which text recognition with good accuracy is to be performed without the aid of sophisticated context analysis.

Looking the other way, the single-author result is far better than the multi-author results, which give an indication of the system's performance with the unconstrained printing of an untutored population. Considerable education and constraint would evidently have to be applied to a population in order to achieve high recognition rates.

The recognition system has arrived at its present level of performance through the successive incorporation of several new features, whose progress has been detailed in many of the previous Quarterly Reports. Starting with the original PREP and CALM structures, which were implemented both in hardware and in computer simulations, major additions have been the nine-view preprocessing, the TOPO preprocessors, and the parallel combining of preprocessor-classifier systems. Each of these building blocks plays an important role in the final result.

IV EXPERIMENTS WITH HIGHLEYMAN'S DATA

A. Introduction

One of the recurring problems in evaluating pattern-recognition results reported in the literature is that few authors give sufficiently detailed descriptions of the data they use. This makes it very difficult to make fair comparisons of different pattern-recognition procedures. One set of data, however, has been used as a standard of comparison by several researchers: the set of hand-printed characters collected, quantized, and encoded by Highleyman.[1-4] Since these data were readily convertible to our standard 24 × 24 format, we decided to apply our techniques to them.

Highleyman's data set consists of 50 alphabets of hand-printed characters. Each alphabet was printed by a different individual, and each contains 36 characters (the 10 numerals and 26 upper-case letters) quantized and represented as 12 × 12 binary (black-white) array. The great amount of variability encountered in the data has tended to rule out the simpler approaches, such as the use of decision trees, and the methods used have been more or less statistical in spirit.

One common characteristic of these methods has been the use of some or all of the patterns to fix the values of free parameters in the classifier. In those cases where the first 40 alphabets (the training data) were used to determine parameters and the last 10 alphabets (the testing data) were used to provide an independent test, the performance on the test data was always much worse than the performance on the training data. For example, Chow[4] obtained a 2.1-percent error rate on the training data, but a 41.7-percent error rate on independent test, and this represents the best performance reported to date.

Similar discrepancies have been noted by other investigators[2,4-5] and have usually been attributed to the small number of samples available for characters having so much variability. There is no doubt that a

---

* References are listed at the end of this report.

larger number of samples would reduce the size of this discrepancy, for in the case of infinate training and testing sets, the error rates should be the same. It is not clear, however, how much the test error rate would be reduced, or how many samples would be needed to estimate the best achievable performance.

In this section we shall describe the results of three different experiments with Highleyman's data. The first used a nonparametric classification procedure that exchanged the need for assumptions about the pattern distributions for the need for a large number of patterns. The second used edge-detecting preprocessing prior to classification to remove some of the variability in the characters and to exploit simple a priori knowledge about the data. In the third experiment, the ability of people to recognize the test data was measured to provide an objective performance standard.

B. Nearest-Neighbor Classification

The use of a nearest-neighbor (NN) machine to classify patterns was described in the Sixth Quarterly Report. From a statistical standpoint, the NN rule is a nonparametric decision rule that assigns an unclassified pattern to the class of the nearest of a set of correctly classified reference patterns. When the set of reference patterns is large, the error rate of the NN rule is less than twice the minimum possible error rate. Specifically, if

P_o = Bayes probability of error

P = large-sample NN probability of error

N = Number of classes,

then, under very weak regularity conditions,

$$P_o \leq P \leq 2P_o - \frac{N}{N-1} P_o^2 ,$$

and these bounds can be shown to be the tightest possible.

When the NN rule was applied to Highleyman's data, the training patterns were used as the reference patterns for the classification of the testing data. No preprocessing of the data was performed, each pattern

being viewed as a 144-component binary vector. A test pattern was classified by measuring the Hamming distance between the test pattern and each of the 1440 training patterns, and by assigning the test pattern to the class of the nearest training pattern; ties with patterns in different classes were broken arbitrarily.

The error rate resulting from applying this procedure to the testing data was 47.5 percent. If the training set were large enough for the large sample results to hold, this would mean that the minimum error rate would lie somewhere between 27.6 percent and 47.5 percent. We shall see that the minimum error rate is probably less than 11.4 percent, and, hence, that the training data is not a sufficiently large sample in the nearest-neighbor sense.

C. Edge-Detecting Preprocessing and Piecewise-Linear Classification

One of the big differences between Highleyman's data and the data we have been using in our experiments is that broken and fragmented characters appear frequently in Highleyman's data. This ruled out the use of the TOPO programs to extract features. However, all that was needed to use the PREP 24A simulation of the 1024-image optical preprocessor was to expand the 12 × 12 figures to 24 × 24 figures. This was done merely by copying each row and column twice.

A PREP-CALM experiment was run using the expanded patterns just as we used our own data in the experiments described in the Second and Third Quarterly Reports. The 84-bit feature vectors were obtained for 9 views of each character. These formed the input for the CALM simulation of a 36-category Piecewise-Linear Learning Machine having two Dot Product Units per category and a training margin of 85.

After 18 iterations of the training data (by which time all views of all of the training patterns had been encountered twice), testing was performed. All nine views of each test pattern were presented, and the class appearing most often among the nine individual responses was selected for the pattern. The resulting error rate for all 36 classes was 31.7 percent. Repetition of this experiment using the 10 numerals alone yielded an error rate of 12.0 percent. Both of these results are

significantly better than previously reported results, but this performance still falls short of human performance.

D. Human Performance

In 1960, Neisser and Weene reported an average error rate of 4.1 percent made by a group of nine people in recognizing hand-printed upper-case letters and numerals, and they indicated that 3.2 percent was probably a good estimate of the minimum possible error rate for their data.[10] These results apply to a 34-category alphabet, since confusions between I and l or between O and Ø were not counted as errors. Most importantly, the characters used were reproduced photographically with high resolution and apparently with good gray scale, whereas Highleyman's data are low-resolution two-level gray-scale figures; thus, these rates do not apply to Highleyman's data.

To estimate human error rates on Highleyman's data, we performed a simple, computer-controlled experiment involving 10 people who, though aware of the existence of Highleyman's data, had not seen the test data before. The experimental procedure had two phases: a training phase in which the subjects familiarized themselves with both the equipment and the data by viewing the training data under test conditions, and a testing phase in which performance was recorded. In both phases, the characters were selected randomly without replacement from 10 alphabets printed by 10 different writers; the training phase used the first 10 alphabets, while the testing phase used the last 10.

The characters were displayed as a 12 · 12 array of points (bright points for the figure) occupying a 0.3-inch square centered in a 3 · 4.5-inch oscilloscope screen. Each subject was free to take as long as he wished in making up his mind, and when a decision was reached he reported it by striking the corresponding typewriter key. This caused the subject's decision to be recorded, the correct character to be typed out if a mistake had been made, and the next character to be displayed. We chose to maintain the error response during the testing phase because it noticeably sustained the subject's attention and induced him to perform well.

Most subjects were satisfied with the training phase after they had seen 75 to 100 characters, and volunteered to move on to the testing phase. On the test data, their error rates ranged from 13.6 percent to 18.3 percent, with an average error rate of 15.7 percent. Assuming a normal distribution of scores, this indicates that with 95-percent confidence the true mean error rate is 15.7 percent ± .9 percent.

These numbers include a fair proportion of errors due to confusions between I and 1 and O and 0. If these errors are not counted, the mean error rate drops to 11.5 percent, which is still considerably greater than the 4.1 percent reported by Neisser and Weene for their unquantized characters. If the I-1 and O-Ø distinctions are retained, but if a plurality vote of the 10 separate responses is used to classify the characters (ties being broken arbitrarily), then an error rate of 11.4 percent results. We believe that this value is close to the minimum error rate achievable with Highleyman's data and that the performance of other methods on the 36-character test data should be viewed relative to this standard.

E. Conclusions

The 47.5-percent error rate obtained by nearest-neighbor classification is typical of the error rates achieved by other general classification techniques.[3,4,5] If 11.4 percent is the minimum achievable error rate, then the 47.5-percent result indicates that the amount of training data is much too small for NN classification, and this is probably true for the other general methods as well.

By employing edge-detecting preprocessing followed by 9-view classification by a piecewise-linear machine, we obtained an error rate of 31.7 percent. While this represents a significant improvement over previously reported results, it is still far too high to be practical. However, the best performance we can ever expect on Highleyman's data is approximately 11 percent, which in turn seems to be much too high.

The reason for most of these errors is clear to anyone who has ever looked at Highleyman's data. Aside from the basic indistinguishability of O's from 0's and many I's from 1's, most of the difficulty is due to either inadequate resolution or breaks in the characters. It is extremely doubtful

that more sophisticated preprocessing and classification could ever overcome these fundamental difficulties. Thus, while Highleyman's data has served as an interesting vehicle for comparing our classification methods with others, its basic characteristics severely limit its usefulness for hand-printed character-recognition research.

REFERENCES

1. W. H. Highleyman and L. A. Kamentsky, "Comments on a Character Recognition Method of Bledsoe and Browning," IRE Trans. on Electronic Computers, Vol. EC-9, p. 263 (June 1960).

2. W. W. Bledsoe, "Further Results on the N-tuple Pattern Recognition Method," IRE Trans. on Electronic Computers, Vol. EC-10, p. 96 (March 1961).

3. W. H. Highleyman, "An Analog Method for Character Recognition," IRE Trans on Electronic Computers, Vol. EC-10, pp. 502-512 (September 1961).

4. C. K. Chow, "A Recognition Method Using Neighbor Dependence," IRE Trans. on Electronic Computers, Vol. EC-11, pp. 683-690 (October 1962).

5. W. H. Highleyman, "Linear Decision Functions, with Application to Pattern Recognition," Proc. IRE, Vol. 50, pp. 1501-1514 (June 1962).

6. R. O. Duda and H. Fossum, "Pattern Classification by Iteratively Determined Linear and Piecewise Linear Discriminant Functions," IEEE Trans. on Electronic Computers, Vol. EC-15, pp. 220-232 (April 1966).

7 W. H. Highleyman, "Data for Character Recognition Studies," IEEE Trans. on Electronic Computers, Vol. EC-12, pp. 135-136 (April 1963).

8. T. M. Cover and P. E. Hart, "Nearest Neighbor Pattern Classification," IEEE Trans. on Information Theory, Vol. IT-13, pp. 21-27 (January 1967).

9. P. E. Hart, "An Asymptotic Analysis of the Nearest-Neighbor Decision Rule," Stanford Electronics Laboratories Technical Report No. 1828-2 (SEL-66-016) (May 1966).

10. U. Neisser and P. Weene, "A Note on Human Recognition of Hand-Printed Characters," Information and Control, Vol. 3, pp. 191-196 (1960).

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified

Stanford Research Institute 333 Ravenswood Avenue Menlo Park, California 94025	UNCLASSIFIED N/A

REPORT TITLE

GRAPHICAL-DATA-PROCESSING RESEARCH STUDY AND EXPERIMENTAL INVESTIGATION

DESCRIPTIVE NOTES (Type of report and inclusive dates)

Report No. 29, Seventh Quarterly Report - 1 September to 30 November 1967

AUTHOR(S) (First name, middle initial, last name)

Richard O. Duda Peter E. Hart John H. Munson

REPORT DATE	TOTAL NO. OF PAGES	NO. OF REFS
February 1968	29	10

CONTRACT OR GRANT NO.	ORIGINATOR'S REPORT NUMBER(S)
DA 28-043 AMC-01901(E) PROJECT NO. IP6-20501-A-448-02-45	SRI Project 5864 7th Quarterly Report
c. d.	OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ECOM-01901-29

DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution in unlimited.

SUPPLEMENTARY NOTES	SPONSORING MILITARY ACTIVITY
	U.S. Army Electronics Command Fort Monmouth, New Jersey 07703 Attn: AMSEL-NL-A-1

ABSTRACT

This report describes the continuing development of scanning, preprocessing, character-classification, and context-analysis techniques for hand-printed text, such as computer coding sheets in the FORTRAN language.

Both edge-detection and topological preprocessing are coupled with classification by a learning machine and used to process a large file of characters printed by a single author. The two systems are combined to achieve a recognition rate considerably better than our previous results. No other comparable results on unconstrained hand printed with a full alphabet are known to us.

The same methods are also applied to a well-known file of hand printed characters collected by Highleyman. The combination of processing and classification methods has achieved performance better than that reported for any other recognition system.

DD FORM 1 NOV 68 1473 (PAGE 1)

S/N 0101-807-6801

UNCLASSIFIED

Security Classification

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Pattern Recognition						
Adaptive Systems						
Learning Machines						
Preprocessing						
Classification						
Context Analysis						
FORTRAN Syntax Analysis						

DD FORM 1 NOV 68 1473 (BACK)
(PAGE 2)

UNCLASSIFIED

Security Classification