NAVSWC TR 91-74

AD-A238 294

HANDWRITTEN CHARACTER RECOGNITION USING FEATURE EXTRACTION AND NEURAL NETWORKS

BY J. J. FULLER A. FARSAIE T. DUMOULIN

WEAPONS SYSTEMS DEPARTMENT

12 FEBRUARY 1991

Approved for public release; distribution is unlimited.

NAVAL SURFACE WARFARE CENTER

Dahlgren, Virginia 22448-5000 • Silver Spring, Maryland 20903-5000

HANDWRITTEN CHARACTER RECOGNITION USING FEATURE EXTRACTION AND NEURAL NETWORKS

BY J. J. FULLER A. FARSAIE T. DUMOULIN
WEAPONS SYSTEMS DEPARTMENT

12 FEBRUARY 1991

NAVAL SURFACE WARFARE CENTER
Dahlgren, Virginia 22448-5000 • Silver Spring, Maryland 20903-5000

FOREWORD

In this report, a novel set of affine invariant features and an accompanying neural network are described for performing two-dimensional pattern recognition, specifically, handwritten character recognition. The results achieved by the network on single characters, distinct words, and similar words are presented. The Concepts and Technologies Branch (G42) is applying artificial neural systems technology to areas which require fast, accurate, and robust recognition and classification.

This study was partially funded by the Office of Naval Research, Summer Faculty Program.

This technical report was reviewed by Dr. Kenneth F. Caudle, Head of the Advanced Weapons Division.

Approved by:

Kurt F. Mueller

KURT F. MUELLER
Deputy Department Head
Weapon Systems Department

CONTENTS

INTRODUCTION

The problem of recognizing handwritten characters and/or words is one that has delighted, intrigued, and puzzled researchers in Artificial Intelligence for many years. The solution to the problem could result in untold savings in time and money to business, industry, and government.

Many significant advances have been made in this area including the development of the Neocognitron-a hierarchichal neural network organized like the visual cortex and consistent with the visual system described by Hubel and Wiesel.[1] The network has the advantages of being affine invariant and of being relatively insensitive to noise and distortion. It has the disadvantage of being one of the most complicated neural networks ever devised.

Other approaches have been taken and even some commercially available software products have been developed. It is probably still correct to say, however, that the final solution has not been achieved. The problem of segmentation looms large over any attempt at solution. A new approach--combining features initially developed for target recognition with a simple neural network—will be described. While the approach taken does not purport to solve the problems of segmentation and proliferation of correct forms described above, it will be shown to have the advantages of simplicity and accuracy.

APPROACH

Fuller and Farsaie[2] were originally described and used a set of new features (Theta Neighbors) for identification of and differentiation among targets. We formalize the notion of Theta Neighbors as follows:

Let I represent a digitized image contained in an n x m pixel grid. Let (X,Y) denote the center of mass of I. Without loss of generality, we may assume that (X,Y) = (0,0) since a simple translation would accomplish the desired result. Let theta be a fixed angle - 0 < theta < 360 degrees- and let (i,j) and (l,m) be two pixels in the n x m pixel grid. We say that (i,j) and (l,m) are Theta Neighbors if and only if for some angles phi[1] and phi[2]

1. (i,j) and (l,m) are both pixels in I
2. (i,j) = (R*cos(phi[1]),R*sin(phi[1])
   (l,m) = (R*cos(phi[2]),R*sin(phi[2])

Thus (i,j) and (l,m) lie on the same circle of radius R about the origin

3. ABS(phi[1]-phi[2]) = theta where ABS denotes absolute value.

A set of features may be constructed as follows:
Let {theta[1],theta[2],...,theta[n]} be a set of distinct angles with 0 < theta[i] < 360 for i = 1,2,...,n. Let N(theta[i]) denote the number of pixels in I which have let theta[i] neighbors and let F[i] = N(theta[i])/A where A is the area of the circle of minimum radius needed to enclose.

The strength of Theta Neighbors in identifying patterns is summarized in the following propositions:

Proposition 1: Let I be as above and let I(theta[i]) be the image obtained when I is rotated by theta[i] about its center of mass. Then N(theta[i]) may be obtained by simply counting the pixels in the intersection of I and I(theta[i]).

Proposition 2: The set of features {F[1],F[2],...,F[n]} described above is affine invariant.

We omit formal proofs of propositions 1 and 2 and we simply state that rotation and translation invariance of the features is obvious since Theta Neighbors are determined with respect to the center of mass of I. Scale invariance is apparent if it is recognized that N(theta[i]) is the area of a particular subset of a circle. As the radius of this circle increases or decreases, N[theta[i]) will increase or decrease proportionately with the square of the radius, or equivalently with the area.

Perhaps more important than the calculation of the features just described is the motivation for choosing them in the first place. This motivation comes from the political cartoonist who--when constructing a caricature of a famous person--chooses and then emphasizes a prominent feature of that person. An obvious question is "How are the prominent features identified?"

While this question is difficult to answer quantatatively, one can say prominent features are those which humans identify as being substantially different from some internalized norm. By performing a series of mental subtractions from a norm, humans accomplish recognition.

This reasoning is the basis for choosing the features {F[1],F[2],...,F[n]} described above. If the image I is a solid circle, it is seen that F[i] = 1 for i = 1,2,...,n. Any point in the circle would have a Theta Neighbor for any value of theta and N(theta[i]) would simply be the area of the circle. As I departs from being a circle--the network's internalized norm--the values of F[i] measure how big/small this departure has become. Values of F[i] substantially different from 1 or very close to 1 quantatatively identify "prominent" features of I. The need for a neural network may be questioned at this point since the features we have described are invariant. In Reference 2 the problems with aliasing and jaggies are documented when each target has only one correct form. Handwritten character recognition has not only the problems of aliasing and jaggies, but also the problem of a multitude of correct forms for characters or words. Neural networks were chosen as the method to deal with these problems.

Previously, Kohonen networks were used to learn the extracted features. While the results were good, these networks suffer from long training times and problems when new examples of old classes are added late in training. For these reasons, the Cluster Euclidean network was chosen for use in testing and training. This network is descibed by Lippmann in Reference 3.

TRAINING AND TESTING

Due to the many correct forms of handwritten characters, a definitive test is difficult to devise. To establish the validity of the approach taken in this paper, it was decided to test the network in the following three ways:

1. Single characters--to establish the capability of the network to distinguish among characters having little structure
2. Distinct Words--to establish the capability of the network to differentiate among nonsimilar words
3. Similar words--to establish the capability of the network to differentiate between very similar words

To accomplish training and testing, a program was written which allowed the user to use a mouse to draw characters and write words in a 100 x 100 pixel grid. Two features were chosen corresponding to theta = 45 degrees and theta = 90 degrees.

SINGLE CHARACTERS

In the single characters training, the user was told to

1. write legibly and
2. give the network 5 examples each of the characters 2,4,6, and 8.

For testing, the same tester was told to

1. write legibly
2. show the network the sequence 2,4,6,8 five times and to record the responses.

The results are shown in the following table:

	Character			
	2	4	6	8
Number Correct	4	5	5	5

Total correct 19/20 = 95%

The only incorrect response occurred when the network classified a 2 as a 6. This is understandable when one considers the invariance of the features and the fact that a handwritten 6 often looks like a mirror image of a handwritten 2.

DISTINCT WORDS

To train the network to differentiate between distinct words, a tester was instructed to

1. Write legibly
2. Show the network 2 examples each of the "words" KOHO, ART I, and BPROP

For testing, the tester was instructed to

1. Write legibly
2. Repeat the sequence KOHO, ART I, and BPROP 4 times and record the results

The network correctly classified all 16 words it was presented during the test.

SIMILAR WORDS

To train the network to differentiate between similar words, the tester was instructed to

1. write legibly
2. Show the network examples of the words "BELL" and "BiLL" until it appeared the network was capable of making a distinction between them. This required approximately 20 examples of each word.

For testing, the tester was instructed to

1. Write legibly

2. Repeat the sequence "BELL" and "BiLL" 6 times and record the results.

The results showed that a correct identification was made in 11 of the 12 test cases. With no additional training, the network was asked to identify the words "BELLs" and "BiLLs." The network correctly identified "BELLs" as being most similar to "BELL." With one more training example, the network was also capable of identifying "BiLLs" as being most similar to "BiLL."

The overall performance of the network is summarized in the following table.

	Single Character	Distinct Words	Similar Words
Training Examples	5	2	20
Percent Correct	95	100	92

SUMMARY AND CONCLUSIONS

The results presented in this paper show great promise for the use of Theta Neighbors and neural networks for handwritten character recognition. In addition, nothing in this paper was peculiar to handwritten character recognition. Hence the approach may be applied to the problem of pattern recognition in general. The features described are easily calculated, and the accompanying neural network is easy to program and quick to train.

By using the features described in this paper and a more complicated neural network, the authors are achieving good results on the problem of three-dimensional target recognition. The results of this work should appear soon.

The authors expect to continue to investigate the problem of target recognition using neural networks and features such as those described here. Emphasis will be placed on partially obscured targets in three dimensions.

REFERENCES

1. Hubel, D. H. and Wiesel, T. N., "Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex," Journal of Physiology, Vol. 160, 1962, pp. 106-154.

2. Fuller, J. and Farsaie, A., "Invariant Target Recognition Using Feature Extraction," in Proceedings of the IJCNN, Washington, DC, Jan 1990, Vol. II, pp. 595-598.

3. Lippmann, R. P., "An Introduction to Computer with Neural Nets," IEEE ASSP Magazine, Apr 1987, pp. 11-13.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	12 February 1991	Interim Report

4. TITLE AND SUBTITLE
Handwritten Character Recognition Using Feature Extraction and Neural Networks

5. FUNDING NUMBERS

6. AUTHOR(S)
J. J. Fuller, A. Farsaie, and T. Dumoulin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Surface Warfare Center
White Oak Laboratory (Code G42)
10901 New Hampshire Avenue
Silver Spring, MD 20903-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
NAVSWC TR 91-74

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A novel set of affine invariant features and an accompanying neural network are described for performing two-dimensional pattern recognition. The features are described and shown to be affine invariant. The validity of the approach and the need for a neural network are demonstrated by applying the features and the network model to handwritten character recognition.

14. SUBJECT TERMS
neural networks character recognition feature extraction

15. NUMBER OF PAGES
18

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and its title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

BLOCK 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If Known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- **DOD** - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- **DOE** - See authorities.
- **NASA** - See Handbook NHB 2200.2
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - Leave blank.
- **DOE** - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- **NASA** - Leave blank.
- **NTIS** - Leave blank.

Block 13. Abstract. Include a brief ***(Maximum 200 words)*** factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code ***(NTIS only)***

Blocks 17.-19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.