The South African Institute of Computer Scientists
and
Information Technologists

Proceedings

of the

1996 National Research and Development Conference

Industry meets Academia

Interaction Conference Centre, University of Natal,
Durban.
26 & 27 September

Edited by
Vevek Ram

ISBN 0-620-20568-7

Cover printed by Natal Printers (Pty) Ltd, Pietermaritzburg
Copying by the Multicopy Centre, University of Natal, Pietermaritzburg
Binding by Library Technical Services, University of Natal, Pietermaritzburg

FOREWORD

This book is a collection of papers presented at the National Research and Development Conference of the Institute of Computer Scientists and Information Technologists, held on 26 & 27 September, at the Interaction Conference Centre, University of Natal, Durban. The Conference was organised by the Department of Computer Science and Information Systems of The University of Natal, Pietermaritzburg.

The papers contained herein range from serious technical research to work-in-progress reports of current research to industry and commercial practice and experience. It has been a difficult task maintaining an adequate and representative spread of interests and a high standard of scholarship at the same time. Nevertheless, the conference boasts a wide range of high quality papers. The program committee decided not only to accept papers that are publishable in their present form, but also papers which reflect this potential in order to encourage young researchers and to involve practitioners from commerce and industry.

The organisers would like to thank IBM South Africa for their generous sponsorship and all the members of the organising and program committees, and the referees for making the conference a success. The organisers are indebted to the Computer Society of South Africa (Natal Chapter) for promoting the conference among its members and also to the staff and management of the Interaction Conference Centre for their contribution to the success of the conference.

On behalf of the Organising Committee
Vevek Ram
Editor and Program Chair
Pietermaritzburg, September 1996

Organising Committee

Conference General Chairs
Mr Rob Dempster and Prof Peter Warren (UNP)

Organising Chair
Dr Don Petkov (UNP)

Secretariat
Mrs Jenny Wilson

Program Chair
Prof Vevek Ram (UNP)

Program Committee

Prof Peter Wentworth, Rhodes
Dr Milan Hajek, UDW
Prof Derek Smith, UCT
Prof Anthony Krzesinski, Stellenbosch
Dr Don Petkov, UNP
Mr Rob Dempster, UNP
Prof Peter Warren, UNP

Table of Contents

List of Contributors

Vladimir B Bajic
Centre for Engineering Research,
Technikon Natal,
P O Box 953
Durban 4000

C N Blewett
Department of Accounting
University of Natal
King George V Avenue
Durban 4001

Justin Cansfield
Department of Accounting
University of Natal
King George V Avenue
Durban 4001

Tom Considine
Apron Services (Pty) Ltd
P O Johannesburg
International Airport
1600

Eric Cloete
School of Electrical Engineering
Cape Technikon
Box 652
Cape Town

I Cloete
Computer Science Department
University of Stellenbosch
Stellenbosch
7600

O A Daini
Department of Computer Science
University of Botswana
Gaborone
Botswana

Nirvani Devcharan
Umgeni Water
Box 9
Pietermaritzburg
3200

P J A de Villiers
Department of Computer Science
University of Stellenbosch
Stellenbosch
7700

Ruth de Villiers
Department of Computer Science and
Information Systems
UNISA
Box 392, Pretoria, 0001

G J Erwin
Business Information Systems
University of Durban-Westville
Private Bag X54001
Durban 4000

G Ganchev
Computer Science Department
University of Botswana
PBag 0022
Gaberone, Botswana

J Geldenhuys
Department of Computer Science
University of Stellenbosch
Stellenbosch
7700

Louise Gibson
BIS, Dept Accounting & Finance
University of Durban
Pvt Bag X10
Dalbridge 4014

Mike Hart
Department of Information Systems
University of Cape Town
Rondebosch
7700

M. Hajek
Department of Computer Science
University of Durban-Westville
Pvt Bag X54001
Durban 4000

A C Hansen
Dept of Agricultural Engineering
University of Natal
Private Bag X01
Scottsville 3209

J M Hattingh
Department of Computer Science
Potchefstroom University for CHE
Potchefstroom 2520

Boris Jankovic
Centre for Engineering Research
Technikon Natal
P O Box 953,
Durban 4000

Meredith McLeod
Centre for Engineering Research,
Technikon Natal,
P O Box 953
Durban 4000

Paula Kotze
Department of Computer Science and
Information Systems
UNISA
Box 392
Pretoria, 0001

D Moodley
Computer Management Systems
Box 451
Umhlanga Rocks
4320

J W Kruger
Vista University
Soweto Campus
Box 359
Westhoven 2124

Andrew Morris
P O Box 34200
Rhodes Gift
7707

A C Leonard
Dept of Informatics
University of Pretoria
Pretoria
2000

R C Nienaber
Technikon Pretoria
Dept of Information Technology
Private Bag X680
Pretoria 0001

Ben Laauwen
Laauwen and Associates
P O Box 13773
Sinoville
0129

E Parkinson
Department of Computer Science
University of Port Elizabeth
Box 1600
Port Elizabeth 6000

Mari Le Roux
Information technology, development: project
leader
Telkom IT 1015
Box 2753
Pretoria 0001

Don Petkov
Department of Computer Science and
Information Systems
University of Natal
PBag x01
Scottsville 3209

P W L Lyne
Dept of Agricultural Engineering
University of Natal
Private Bag X01
Scottsville 3209

Olga Petkov
Technikon Natal
Box 11078
Dorpspruit 3206
Pietermaritzburg

Rose Mazhindu-Shumba
Computer Science Department
University of Zimbabwe
Box MP167
Harare, Zimbabwe

N Pillay
Technikon Natal
Box 11078
Dorpspruit 3206
Pietermaritzburg

V Ram
Department of Computer Science and
Information Systems
University of Natal
PBag x01
Scottsville 3209

Melinda Redelinghuys
Department of Computer Science and
Information Systems
UNISA
Box 392
Pretoria, 0001

Karen Renaud
Computer Science and Information Systems
UNISA
Box 392
Pretoria, 0001

H N Roux
Department of Computer Science
University of Stellenbosch
Stellenbosch
7700

T G Scott
Department of Computer Science
Potchefstroom University for CHE
Potchefstroom
2520

T Seipone
Department of Computer Science
University of Botswana
Gaborone
Botswana

Derek Smith
Department of Information Systems
University of Cape Town
Rondebosch
7700

Anette L Steenkamp
Department of Computer Science and
Information Systems
UNISA
Box 392
Pretoria, 0001

T Steyn
Department of Computer Science
Potchefstroom University for CHE
Potchefstroom 2520

H. Suleman
Department of Computer Science
University of Durban-Westville
Pvt Bag X54001
Durban 4000

A Vahed
Department of Computer Science
University of Western Cape
Private Bag X17
Bellville 7530

A Van der Merwe
Computer science and Informations Systems
UNISA
P O Box 392
Pretoria,0001

Tjaart J Van Der Walt
Foundation for Research and Development
Box 2600
Pretoria, 0001

K Vavatzandis
Department of Information Systems
University of Cape Town
Rondebosch
7700

Y Velinov
Dept Computer Science
University of Natal
Private Bag X01
Scottsville 3209

H Venter
Department of Computer Science
University of Port Elizabeth
Box 1600
Port Elizabeth 6000

H L Viktor
Computer Science Department
University of Stellenbosch
Stellenbosch
7600

R Von Solms
Department of Information Technology
Port Elizabeth Technikon
Private Bag X6011
Port Elizabeth 6000

A J Walker
Software Engineering Applications
Laboratory
Electrical Engineering
University of Witwatersrand
Johannesburg

P Warren
Computer Science Department
University of Natal
P/Bag X01
Scottsville 3209

Max Watzenboeck
University of Botswana
Private Bag 0022
Gaberone
Botswana

FEATURE EXTRACTION IN NEURAL NETWORKS: AN APPLICATION IN HANDWRITTEN CHARACTER RECOGNITION

Deshendran Moodley
Computer Management (Pty) Ltd
P O Box 451, Umhlanga Rocks, 4320

Vevek Ram
Department of Computer Science and
Information Systems, University of Natal,
Private Bag X01, Scottsville, 3209

Abstract

This paper examines and provides an overview of feature extraction techniques that are currently used in neural networks for image recognition. These include moment invariants, Zernike moments, Fourier descriptor techniques, Gabor and wavelet filters and the Neocognitron. An implementation of an handwritten character recognition system is discussed to illustrate the practical significance of feature extraction. The methods of Zernike moments and the Neocognitron are used for feature extraction and the multilayered perceptron is used as the classifier. The results are also compared to the same application without feature extraction.

Introduction

Shape and pattern recognition is an essential function of computer-based vision systems used in industrial automation. Artificial Neural Networks have played a prominent role in this area and many diverse manufacturing applications such as product inspection and packaging, robotics and remote sensing have been successfully developed. In a typical application, a feature vector which is a numerical representation of an image, is presented to an ANN for classification. In most cases, the overwhelming size of the feature vectors require networks which are computationally impractical. Feature extraction is a process which drastically reduces the size of the feature vector when compared to the original image, by reducing redundancy and retaining only the information necessary for discrimination. Many techniques for achieving feature extraction exist and this paper examines and provides an overview of those that are currently used.

The multilayered perceptron (MLP) (Rumelhart, Hinton & Williams, 1986), a popular neural network architecture, consists of three layers, the input layer, the hidden layer and the output layer with each layer containing processing elements, called neurons, which are connected to the neurons in the previous layer. The classification process involves the presentation of a feature vector to the input layer and the producing of a response from a specific neuron in the output layer, corresponding to an output class. Initially a set of prototype pairs of feature vectors and desired responses are presented repetitively to the network. After each presentation, differences between the actual output and the desired output of the network are used to modify the strength of connections within the network. Eventually the network reaches a state where these differences are negligible and the network is ready to perform classification. For digital image recognition the digital image forms the feature vector and the output classes are all the possible classification categories for the image. The use of the multilayered perceptron in digital image recognition has two limitations. Firstly, the unnecessary processing of redundant information not needed for discrimination between output classes. For example, if the task was to classify fruit the information about the size, position or how the fruit is rotated in the image is of no significance. Secondly, classification of large digital images necessitates a large number of neurons in the input layer, which leads to an increase in the computational resources needed for the operation of the network and the time needed for training the network.

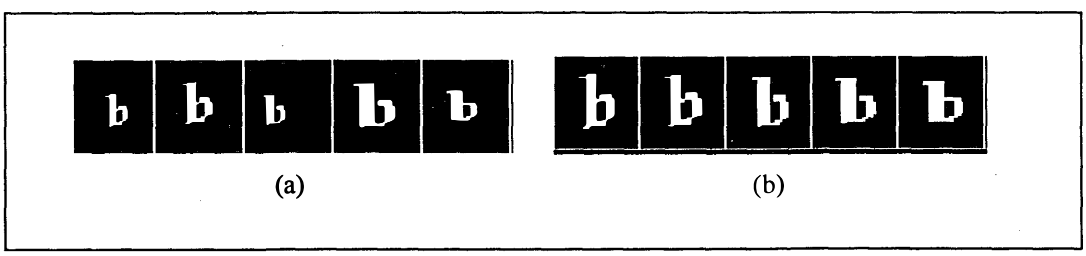
The above limitations of the MLP can be overcome by the introduction of a feature extraction stage in which the digital image is processed so that only relevant properties of the image required for discrimination between the output classes are extracted. This new feature set will then form the input to the neural network. In this way redundant information will be removed and the size of the feature set reduced. Generally feature extraction methods must have the following properties (Khotanzad & Lu, 1991). Extracted features must retain much of the discriminant information present in the original data. Features should have small intra-class variance, that is, slightly different shapes with similar general characteristics should have similar numerical values. Features must also have large inter-class differences, that is, features from different classes should be quite different numerically.

Feature extraction methods

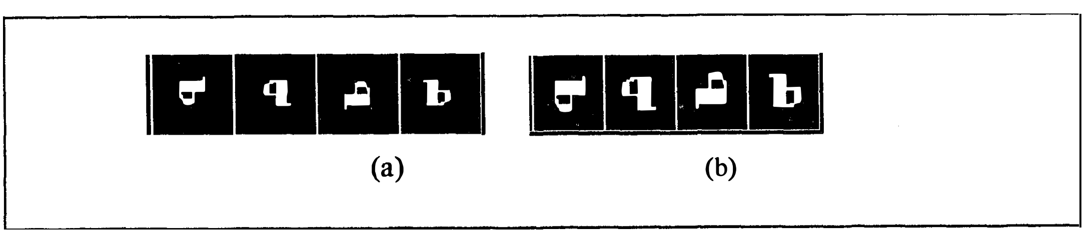
The shape of objects in the image and variations in texture within the image are especially important in digital image recognition. Different textures in images can be distinguished by their preferred direction or orientation in the image and by their spatial frequency, that is, how fine (more detail) or coarse (less detail) the pattern is. The detection of variations in texture helps in the segmentation of an object and its background in the image and the identification of the shape of the object. The feature extraction methods discussed below can be classified by these properties of the image which they capture. Moment invariants (Jain, 1989), Zernike moments (Khotanzad and Hong, 1990), Fourier descriptors (Kulkarni, 1994) and the Neocognitron (Fukushima, 1988) capture the shape of an object in the image while Gabor filters (Daugman, 1988) and wavelets (Mallat, 1989) help to distinguish between textures within an image. Since Zernike moments and the Neocognitron are selected for the experiments, they are discussed in detail. The other methods are included for the sake of completeness.

Zernike moments

Zernike moments (Khotanzad and Hong, 1990) have been used previously for shape recognition tasks. The representation consists of 47 positional, size and orientation invariant features which are the magnitudes of the complex valued Zernike moments of a digital image. The original image can be reconstructed from these moments, thus giving an indication of the quality of the representation. Firstly, the image is standardised with respect to size and position. Thereafter the rotational invariant Zernike moments of these standardised images are calculated. Figure 1 shows five instances of the letter 'b' with variations in size and position before and after standardisation and figure 2 shows four instances of the letter 'b' with variations in orientation before and after standardisation. The first four Zernike moments of the images in figure 2b are shown in table 1.



1. Five instances of the character b (a) before and (b) after standardization



2. The letter b rotated through angles of 90°, 180°, 270°

Table 1 The magnitudes of 4 Zernike moments for the rotated images in figure 2

	90°	180°	270°	0°
|A[2,0]|	207.935751	207.935751	207.935751	207.935751
|A[2,2]|	3.246915	3.246915	3.246915	3.246915
|A[3,1]|	52.198420	52.198420	52.198419	52.198419
|A[3,3]|	2.201566	2.201566	2.201566	2.201566

The Neocognitron

The Neocognitron (Fukushima, 1988), is a multilayered neural network architecture, based on the feed forward architectures in biological systems, which combines the feature extraction and classification stage. The Neocognitron is insensitive to variations in position and distortions in the input image. This architecture has been used predominantly for shape recognition tasks. The Neocognitron incorporates the feature extraction stage within the network and this in itself is both convenient and important. The introduction of a separate feature extraction stage, for example Zernike moments results in uncertainty about the degree of processing carried out by the neural network as the feature extraction stage may tend to oversimplify the classification process. The Neocognitron is insensitive to scale, changes in position and even distortions of objects in the input pattern. The unsupervised training used for the Neocognitron resembles closely the processes in biological visual systems and the system determines the features to be detected. Even though this training method does not produce good results in practice, it does give insight into the operation of biological visual systems. To produce better results, but moving away slightly from the biological paradigm, supervised training methods have been developed. These are more suited to pattern recognition and allow the teacher to control which features are to be detected.

Moments invariants

Moment invariants are derived from the central moments, m_{pq}, of an image (Jain, 1989)

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y)$$

These moments provide a positional, size and orientation invariant representation of an image. They have been used for shape recognition and provide a representation consisting of seven features. A significant disadvantage is that there is no measure of how good a representation of the original image these moments provide.

Fourier descriptors

Fourier descriptors provide a positional, size and orientation invariant representation of a digital image. This method has been investigated for 2-D aircraft recognition (Kulkarni, 1994). New ways in which to use Fourier descriptors for image representation are currently being investigated (Kauppinen & Seppanen,1995).

Gabor filters

Gabor filters (Daugman, 1988) operate similarly to structures in biological systems, that is, they categorise areas of an image into different ranges of orientations and spatial frequencies. It is one of the few representations that can simultaneously distinguish between orientations and spatial frequencies which is important for texture classification. However the calculation of the Gabor filters for a digital image is difficult.

Wavelets

The wavelet representation (Mallat, 1989) uses the differences between images at different resolutions to represent a digital image. This representation also, but to a lesser extent, is able to simultaneously distinguish areas of the image according to different orientations and spatial frequencies. Calculation of this representation is less difficult than for Gabor filters.

The application of handwritten digit recognition was chosen to practically illustrate the role of feature extraction in image recognition. Two feature extraction models, Zernike moments and the Neocognitron using supervised training, were chosen as suitable for two dimensional shape recognition. Three experimental applications were designed and implemented to demonstrate the significance of introducing a feature extraction stage into the handwritten digit recognition system:

- Experiment I consisted of a system with no feature extraction.
- Experiment II consisted of a system with a feature extraction separate feature extraction stage using Zernike moments.
- Experiment III consisted of a system which combined the feature extraction and the classification stages in the form of the Neocognitron.

These experiments were tested on data obtained from the National Institute of Standards and Technology (NIST) in the USA. The data consisted of isolated handwritten digits. Even though in some instances the characters were cut off or parts of two characters were found in some images (segmentation error), the segmentation process used to produce the isolated characters was considered overall to be quite good since:

- the characters had minimal variations in size.
- the characters had little or no variations in position as all the characters filled the entire image.

A selection of these characters were rotated and a new data set, data set II, was constructed which contained a combination of the rotated and the normal characters. The experiments were also tested on data set II. The results of these experiments are displayed in table 1 and table 2.

Table 2 Performance of the three experiments on data set I.

Experimental application	Accuracy rate on training set	Accuracy rate on test set
I. MLP	100%	83.8%
II. Zernike moments	55-65%	50 - 60%
III. Neocognitron	-	41%

Table 3 Performance of the three experiments on data set II.

Experimental application	Accuracy rate on training set	Accuracy rate on test set
I. MLP	100%	47%
II. Zernike moments	100%	67%
III. Neocognitron	-	28%

The results of these experiments can be summarised as follows:

- The introduction of a feature extraction method in the image recognition system, developed in experiment II, reduces the size of the feature vector and thus the amount of resources needed for implementing, training and testing of the artificial neural network. This is especially noticeable in the number of neurons needed for the input layer.
- The feature extraction process allows the system to be invariant to certain irrelevant properties of the input data.
- The disadvantages of introducing a feature extraction stage into the system, is the additional processing that is required and the loss of information that occurs when reducing the input data. This must be weighed against the advantages of invariancy and reduction in size of the feature vector which feature extraction provides. It can thus be concluded that if there are negligible variances in parameters such as size and position, then the only advantage of the feature extraction stage is reduction in size of the feature vector.

Conclusions

An artificial neural network, the multilayered perceptron, has been investigated for digital image recognition applications. The limitations of the multilayered perceptron have been addressed by the introduction of a feature extraction stage. An experiment in the recognition of handwritten characters demonstrated the practical effectiveness of two particular methods namely Zernike moments and the Neocognitron. There are many other applications where feature extraction can certainly play a major role in contributing to the effectiveness of recognition systems, and it is hoped that this research has provided some insight in order to make better decisions regarding their design and development.

References

Daugman, J. G. (1988), Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol 36 no 7, pp. 1169- 1179.

Fukushima, K. (1988), A Neural Network for Visual Pattern Recognition, *IEEE Computer*, Vol 21, no. 3, pp. 65- 74.

Jain, A. K. (1989), *Fundamentals of digital Image Proessing*, Prentice-Hall, United States of America.

Kauppinen, H., Seppanen, T. Pietikainen, (1995), An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D shape Classification, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 12, no 2, pp. 201- 206.

Khotanzad A., and Lu J. -H. (1991), Shape and Texture Recognition by a Neural Network, in *Artificial Neural Networks and Statistical Pattern Recognition*, edited by Sethi I. K. & Jain A.K. , Vol. II, pp. 109-131, Elsevier Publishers B. V, Amsterdam.

Khotanzad, A., Hong, Y. H. (1990), Invariant Image Recognition by Zernike Moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, no. 5, pp. 489-498.

Kulkarni, A. D. (1994), *Artificial Neural Networks for Image Understanding*, Van Nostrand Reinhold, New York.

Rumelhart, D. E., Hinton, G. E. , Williams, R. J. (1986), Learning Intern Representations by Error Propagation, In *Parallel Distributed Processing - Explorations in the Microstructure of Cognition, Volume 1: Foundations*, edited by Rumelhart, D. E., McClelland, J. L.and the PDP Research Group, pp. 318-362, MIT Press, United States of America.