



City Research Online

City, University of London Institutional Repository

Citation: Roy, D. M. (1996). Gestural human-machine interaction using neural networks for people with severe speech and motor impairment due to cerebral palsy. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/31058/>

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

**Gestural Human-Machine Interaction using Neural
Networks for People with Severe Speech and Motor
Impairment due to Cerebral Palsy**

David Michael Roy

Submitted for the degree of Doctor of Philosophy
to
the City University

Department of Systems Science
Department of Clinical Communication Studies

January 1996

Contents

1	Introduction	1
1.1	Background and Rationale	1
1.2	Aims and Objectives	4
1.3	Previous and Related Work	4
1.3.1	Use of Gesture and Sign in severely speech and motor impaired Populations.....	4
1.3.2	Gestural Human Machine Interaction for People with Motor Impairment.....	7
1.3.3	Gestural Human-Machine Interaction for Other Populations	9
1.4	Overview of Thesis	12
2	Systems Approach to Human Machine Interaction	15
2.1	Introduction	15
2.2	Human-Machine Interaction as Emergent Behaviour	15
2.3	Eliciting and Identifying Candidate Behaviours for HMI.....	19
2.4	Machine Perception of Human Behaviour	21
2.4.1	High Bandwidth Efferent Human Machine Interaction	21
2.4.2	System Architecture for Gestural HMI.....	21
2.4.3	Automatic Gesture Recognition Study Objectives.....	22
3	Human Factors	24
3.1	Introduction	24
3.2	User Profile	24
3.2.1	Definition of Target Population of Primary Users	24
3.2.2	Specifying the Secondary Users of the HMI	25
3.2.3	Cerebral Palsy: Nature of the Condition	25
3.2.4	Communicative Interaction of People with SSMICP	29
3.3	Human-Machine Interfaces used by People with SSMICP	29
3.3.1	Electronic Assistive Technology	29
3.3.2	Assistive Technology for Communication	31
3.4	Summary and Implications for Gestural HMI Design	32
4	Gestural Ability Pilot Study	34
4.1	Introduction	34
4.2	Subject Selection.....	34
4.3	Subject Profile	35
4.4	Experimental Design.....	38
4.5	Results	40
4.5.1	Previously Documented Gestural Ability	40

4.5.2	Elicited Gestural Ability	41
4.6	Discussion	45
5	Gestural Data Collection for Pattern Analysis	46
5.1	Introduction	46
5.2	Subject Selection	46
5.3	Subject Description	47
5.4	Gestural Subset.....	48
5.5	Experimental Rig	49
5.6	Gesture Tracking System	51
5.7	Gestural Data Collection Sessions	53
5.8	Results and Discussion.....	54
5.8.1	Gesture co-articulation and Timing.....	54
5.8.2	Rate of Production of Gestures	54
6	Examination and Processing of Gestural Data Stream	56
6.1	Introduction	56
6.2	Gesture Animation	56
6.3	Body Model.....	57
6.4	Qualitative Examination of Gesture Set using Animation	59
6.5	Developing a Strategy for Gesture Recognition.....	64
6.5.1	Key Factors Affecting Gesture Recognition	64
6.5.2	Fixed Time Window Recognition Algorithms	65
6.5.3	Advantages and Disadvantages of Feed Forward Neural Networks..	66
6.6	Segmentation	67
6.7	Signal Conditioning/Data Reduction	70
7	Gesture Classification using Neural Networks	76
7.1	Introduction	76
7.2	Neural Network Description	77
7.2.1	Elementary Neuron Model.....	77
7.2.2	Activation Function.....	77
7.2.3	Network Architecture.....	79
7.2.4	Network Learning using Back-propagation of Errors.....	81
7.2.5	Weight and Bias Initialization	82
7.3	Determining a Performance Measure for Gesture Recognition ...	82
7.4	Preparation of Gesture Sets for Training and Validation	83
7.5	Network Complexity	83
7.5.1	Method	84
7.5.2	Results and Discussion.....	84
7.6	Training Behaviour of Neural Networks using Gestural Data	88
7.7	Comparison of Activation Functions	91

7.7.1	Method	91
7.7.2	Results and Discussion.....	92
7.8	Comparison with k-Nearest Neighbour Method	95
7.8.1	Method	95
7.8.2	Results and Discussion.....	95
7.9	Summary and Discussion	99
8	Feature Set Comparison using Neural Networks	101
8.1	Introduction	101
8.2	Gesture Segment Length	101
8.2.1	Method	101
8.2.2	Results and Discussion.....	102
8.3	Forearm Orientation	107
8.3.1	Method	107
8.3.2	Results	110
8.4	Scalar and Vector Velocity	115
8.4.1	Method	115
8.4.2	Results and Discussion.....	116
8.5	Curvature and Plane of Motion	119
8.5.1	Method	119
8.5.2	Results	120
8.6	Summary and Discussion	123
9	Conclusions	125
9.1	Summary	125
9.2	Interpretation and Implications of Research	128
9.3	Future Work.....	129
	Appendix A: Gesture Elicitation Session Transcripts	132
	Appendix B: Neural Network Training Algorithm	155
	Appendix C: Confusion Matrices	158
	Appendix D: Hinton Diagrams of W1 and W2	217
	References	233
	Glossary	243

List of Figures

Figure 2.1	Proposed conceptualisation of human-machine interaction	16
Figure 2.2	Venn Diagram of Human Behaviour	18
Figure 2.3	Systems model of human-human interaction used to elicit candidate behaviour for HMI	20
Figure 2.4	Example architecture for gesture recognition system	22
Figure 2.5	Intellikeys® expanded membrane keyboard	30
Figure 2.6	Two examples of proprietary switches used to harness movement from people with SSMICP	30
Figure 2.7	An example of a VOCA using indirect selection using single switch scanning of up to 128 items	31
Figure 4.1	Histograms indicating number of instances of use of body parts involved in gestural repertoires for subjects S1 to S6	43
Figure 4.2	Histograms indicating number of instances of use of body parts involved in gestural repertoires for subjects S7, S9, S11, S12	44
Figure 5.1	Subject performing dynamic arm gesture with magnetic tracker receiver attached distally to the right forearm	47
Figure 5.2	The “Flock of Birds” magnetic tracker receiver was attached distally using a velcro and elastic wristband developed for the study	50
Figure 5.3	Plan view of experimental rig	51
Figure 5.4	Schematic of “The Flock of Birds” magnetic tracker	52
Figure 5.5	Diagram showing the transmitter and receiver modules and the co-ordinate system used by the “Flock of Birds” magnetic tracker	53
Figure 6.1	Dynamic CP arm gestures animated using body model	57
Figure 6.2	Body model comprising 32 polygons used to animate the “Flock of Birds” magnetic tracker arm movement data	58
Figure 6.3	Corresponding video frames (left) and computer animated frames of magnetic tracker data (right) for the gesture “rainbow” exemplar 1	61
Figure 6.4	Successive corresponding video frames (left) and computer animated frames of magnetic tracker data (right) for the gesture “rainbow” exemplar 2	62
Figure 6.5	Successive corresponding video frames (left) and computer animated frames of magnetic tracker data (right) for the gesture “surrender”	63
Figure 6.6	Successive corresponding video frames (left) and computer animated frames of magnetic tracker data (right) for the gesture “stroke the cat”	64
Figure 6.7	Histogram showing the number of gestures that were unsegmentable due to an inconsistent gestural form or severe “glitching”	69
Figure 6.8	Frequency response of low-pass chebychev IIR filter type I used to filter the gestural data stream	71
Figure 6.9	Gestural stream position data before ((a), (c)) and after ((b), (d)) filtering	72
Figure 6.10	Plots of position data for one exemplar of “rock a baby” showing manual segmentation	73
Figure 6.11	“Plots of position data for one exemplar of “hot”	74
Figure 6.12	Power spectral density of the gestural data stream	75
Figure 7.1	Elementary backpropagation neuron	77
Figure 7.2	Graph of tan-sigmoid activation function	78

Figure 7.3	Graph of log-sigmoid activation function	79
Figure 7.4	Architecture of a fully connected feedforward neural network	79
Figure 7.5	Schematic representation of two layer feedforward neural network architecture	80
Figure 7.6	Average recognition rate for 26 gestures comparing FFNNs with from 1 to 20 nodes in the hidden layer	85
Figure 7.7	Results for gestures recognised at or above 80% for one to 20 nodes in the hidden layer	86
Figure 7.8	Plot of neural network backpropagation training parameters for 12-16t-26 network	90
Figure 7.9	Gesture recognition rate against number of training epochs for all 26 gestures and for the best 12 gestures	91
Figure 7.10	Gesture recognition results for six different combinations of activation functions in the hidden and output layers	94
Figure 7.11	Comparison of Activation Functions: Average recognition rate for 26 gestures for experiments 60 to 65	95
Figure 7.12	Gesture recognition results for Euclidean distance k-nearest neighbour classifier with k=1 to 9 using trte2i14p feature set	97
Figure 7.13	Gesture recognition results for Euclidean distance k-nearest neighbour classifier with standard normalisation with k=1 to 9 using trte2i14p feature set	98
Figure 8.1	Average recognition rate for 26 gestures for feature sets involving gesture segment length	104
Figure 8.2	Recognition results of gesture segments from 160ms to 1120ms for gestures recognized at or above 80%	105
Figure 8.3	Recognition rate for 26 gestures for feature sets involving forearm orientation	111
Figure 8.4	Recognition results feature sets containing forearm orientation information for gestures recognized at or above 80%	112
Figure 8.5	Average recognition rate for 26 gestures. Comparison of results for networks of 16 and 20 nodes in the hidden layer for feature sets containing forearm orientation	113
Figure 8.6	Comparison between 16 and 20 hidden nodes ANN architectures for feature sets containing forearm orientation information	114
Figure 8.7	Recognition rate for 26 gestures for feature sets involving scalar and vector velocity	117
Figure 8.8	Recognition results feature sets containing scalar and vector velocity information for gestures recognized at or above 80%	118
Figure 8.9	Recognition rate for 26 gestures using curvature and plane of motion features	121
Figure 8.10	Recognition results feature sets containing curvature and plane of motion information for gestures recognized at or above 80%	122

List of Tables

Table 4.1	Details of subjects chosen for gestural ability pilot study	36
Table 4.2	Existing Methods of Expressive Communication	37
Table 4.3	Mobility and Powered Mobility Access Method	38
Table 4.4	Concept Categories used to Elicit Gestures	39
Table 4.5	Unaided Modalities of Expression	41
Table 5.1	Gesture Sub-set	49
Table 5.2	“Flock of Birds” output record per sample	52
Table 5.3	Summary statistics for instrumented gestural data collection sessions	55
Table 6.1	Gestures categorised in terms of movement characteristics	60
Table 6.2	The number and proportion of gestures of each class that were not manually segmentable.	69
Table 7.1	Seven gestures recognised at or above 80% with only 4 hidden nodes	87
Table 7.2	Additional 6 gestures recognised at or above 80% with 16 hidden nodes ..	88
Table 7.3	Experiments with varying combinations of activation function	92
Table 7.4	Summary results showing best recognition rates based on for k-nearest neighbours and comparison	99
Table 8.1	Feature sets of increasing sample size representing increasing gesture segment length	102
Table 8.2	Gestures grouped to show the minimum gesture segment length (GSL) needed to recognise each gesture at or above 80%.	106
Table 8.3	Gestures misrecognised at or greater than 60% for GSL of 640ms	107
Table 8.4	Description of feature vectors p involving forearm orientation and the associated feedforward neural network	108
Table 8.5	Description of feature vectors p involving scalar and vector velocity and the associated feedforward neural network	115
Table 8.6	Description of feature vectors p involving curvature and plane of motion and the associated feedforward neural network	119
Table 8.7	Confusion matrix of three visually similar arm gestures	123
Table 8.8	Summary results for feature set study	124

Acknowledgements

I am deeply grateful to all my students in the U.K., and more recently my student collaborators in the USA. I would like to express my thanks to Russell Batchelor and Andy Hardy who originally provided me with the 'at the chalk-face' experience of the need for technology to enhance communication for learning, leisure, work and as a tool for the creative spirit.

I would like to especially acknowledge the student collaborators of John G. Leach School, Delaware, HMS School for Cerebral Palsy, Widener Memorial School, Philadelphia, their therapists, teachers, care-givers and parents for sharing our vision and generously offering their enthusiasm, time and effort in support of the work. Particular mention should be made of the support of: Beth Latcham, Pam Hyde, Dr. Connie Ames of John.G. Leach School; Dr. Tom Sicoli of the A.I. duPont Children's Hospital School; Dr. Diane Gallagher, Ed Contaldi, Christine Coia, and Meredith Cracknell of HMS School and Michelle Levin of Widener Memorial School.

Thanks to my supervisors Professor Bob Fawcus of the Department of Clinical Communications Studies, Dr. Mike Floyd, Director of the Rehabilitation Resource Centre for their guidance, advice and encouragement. I wish to state my gratitude to Professor Ewart Carson, System Science Department and Director of the Institute of Health Sciences for generously offering me a SERC studentship at the City University. Special thanks to Dr. Roman Hovorka, Dr. Ron Summers, Dr. Tim Ellis for their guidance and reading of my draft manuscript.

Particular acknowledgement is due to Dr. Rick Foulds and Dr. William Harwin of the Center for Applied Science and Engineering in Rehabilitation, University of Delaware for facilitating my stay in the USA as a visiting scholar to the University of Delaware. Thanks to Dr. Roman Erenshteyn for his mentorship in the area of pattern recognition. Thanks to Dr. Michael Alexander and Dr. Freeman Miller of Alfred. I. duPont Institute Children's Hospital of the Nemours Foundation for providing advice relating to clinical issues and in recommending subjects.

My thanks to colleagues who have contributed to my thinking in gesture, psycholinguistics, language acquisition, in particular: to Professor Bencie Woll for our creative discussion on the nature of sign. In the Netherlands: Dr. Hans van Balkom; Margriet Heim for sharing her recent research work with infants with cerebral palsy; Dr. Ann Mills for inviting me to participate at the Morphology and Phonology of Sign workshop. Thanks to: Bruce Baker for our discussion on the polysemous nature of symbols; Dr. Charlie Robinson, for the opportunity to see Pittsburgh and discuss computational intelligence; Dr. Catherine Pelachau and Dr. Norman Badler of University of Pennsylvania for the invitation to learn more about facial animation and recognition; Sister Jean Antony and Rebecca Mercuri of Immaculata College; John Eulenberg, University of East Lansing. Thanks to the ACM CHI'94 Doctoral Consortium Committee for inviting me to Boston, in particular Dr. Marilyn Mantei and Dr. Andrew Monk.

I would like to thank the numerous individuals, colleagues, and friends in the UK, Europe and the USA who have added to my understanding of the problem and inspired me with renewed enthusiasm.

Lastly, but by no means least, I wish to acknowledge Marilyn Panayi for her collaboration, inspiration and enthusiasm. In particular, acknowledgement is due for her assistance in working with the student collaborators and acting as facilitator during data collection.

This thesis reports work that was conducted at the City University, London, UK and as a visiting scholar to the University of Delaware at the Centre for Applied Science and Engineering in Rehabilitation, University of Delaware, USA. Gestural data collection was performed at John G. Leach School, Delaware, HMS School for Cerebral Palsy and Widener Memorial School, Philadelphia, Pennsylvania, USA. Financial support for the visiting scholarship was provided by the Nemours Foundation, USA. This Ph.D. research was supported by a studentship from the Science and Engineering Research Council, U.K.

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

The long-term aim of this research is the development of a robust and appropriate method of high efferent bandwidth gestural human-machine interaction (HMI) that enhances and extends the multimodal expressive abilities of people with severe speech and motor impairment due to cerebral palsy (SSMICP). A human-factors driven approach was adopted to generate and identify candidate behaviour for gestural HMI. Neural methods were applied to investigate the automatic recognition of human movement with a high noise component using spastic-athetoid cerebral palsy arm movement data.

Human-machine interaction was considered as an emergent property leading to the development of a methodology based on human-human interaction to elicit a wide range of spontaneous or near spontaneous gestures. Twelve subjects with SSMICP aged five to 18 years took part in a gestural ability pilot study. From 30 to 141 concepts presented verbally were used to elicit a wide range of spontaneous or near spontaneous gestural responses. Subjects were encouraged to express each concept in any way they wished. Frequently gestural ability was beyond that anticipated by therapists, educators, parents and physicians. Therapeutic, educational, and medical records did not predict gestural ability observed in the study. Analysis of video-taped sessions indicated that gestures were frequently articulated using multiple parts of the body. Nine out of ten subjects used either the right or left arm more frequently than any other body part.

Instrumented gestural data comprising a subset of 27 gestures from a 17 years old subject with spastic-athetoid quadriplegia was used to investigate automatic gesture recognition. Co-articulated dynamic arm gestures were elicited in random order and gestural data recorded at 100 samples/second using a six-degree-of-freedom magnetic tracker attached distally to one forearm. The gestural data stream was examined using a simple body model developed using MATLAB¹ and animated on a Silicon Graphics Workstation. In the absence of suitable features to automatically segment the gestural data stream, gestures were manually segmented.

Low-pass filtering was used to remove "jerkiness" and data reduction was achieved through re-sampling. The use of time-delay feedforward neural networks was investigated using features extracted over a fixed time interval as input. Neural network classifiers outperformed two k-nearest neighbour methods. Time windows of 160ms to 1120 ms were compared. A span of 640ms comprising four time samples yielded the optimum rate of recognition. Feature sets containing measures of position, forearm orientation, scalar and vector velocity, curvature and plane of motion were compared. A feature set comprising four time intervals of x,y,z position gave highest recognition rate. 12 gestures were recognised at or above 80% with an average recognition rate of 90%. Maximum results for all 26 gestures was 55%. Results suggest that the fixed time window approach coupled with low pass filtering may be a feasible method for the computer recognition of noisy gestural movement. Conversely, the results show that it is possible for people classed as having no functional use of upper extremities by traditional assessment techniques to produce a repertoire of dynamic arm gestures with sufficient consistency to be recognised by machine.

¹ MATLAB is a trademark of The Math Works Inc.

List of Abbreviations

AAC	Alternative and augmentative communication
ADL	Activities of daily living
ANN	Artificial neural network
CFM	Confusion matrix
CP	Cerebral palsy
FFNN	Feedforward neural network
GSL	Gesture segment length
HCI	Human computer interaction
HHI	Human-human interaction
HMI	Human-machine interaction or human-machine interface
kNN	k-nearest neighbour
SGI	Silicon Graphics Incorporated
SSE	Sum-squared error
SSMICP ...	Severe speech and motor impairment due to cerebral palsy
VOCA	Voice output communication aid

Chapter 1

Introduction

1.1 Background and Rationale

“We are now witnessing the appearance and rapid development of portable, wearable, and environmental computer technologies. If current developments continue, we may ultimately see a computer and network interface in every effective object in our environment. We will transact more and more of our economic and emotional business through technology. The obvious question is what the effect and affect of such technology be?”

Myron Kruger, 1993 “The Experience Society”, Presence vol. 2, no. 2, pp.162-168.

The range of activities that are computer mediated is rapidly expanding. Increasingly our everyday interactions are computer mediated. This phenomenon has been described as “ubiquitous computing” (Buxton & Card, 1994). However, existing human-machine interfaces¹ only harness a fraction of human expressive and manipulative ability. As the computing power available for purely interface related activities has grown, the *afferent* bandwidth of human-machine interaction (HMI) has increased substantially e.g. multimedia computing, graphical user interfaces (Shneiderman, 1992, Maybury, 1993). There has not yet been a corresponding increase in the *efferent* bandwidth. Even in the emerging field of virtual reality, the vast majority of research papers relate to physical affect more than effect at the interface.

The work presented in this thesis is concerned with addressing this issue, particularly in

¹ The terms “human-machine interface” and “human-machine interaction” are used in preference to “human-computer interface” and “human-computer interaction” in anticipation of all machines incorporating computer technology. In addition, the envisaged applications of gestural human-machine interaction extend far beyond the traditional concept of a computer system.

relation to the needs of disabled people. It is proposed that as computer technology advances *any human behaviour will become a candidate for human-machine interaction*. This poses the questions: *which behaviours should be used and how?* The investigation of this research question involves both human factors and technological issues equally. Central to the approach adopted in this thesis is the notion that machines are more likely to closely fulfil the needs of people if, on balance, human factors issues drive and shape technological development (Shneiderman, 1992, Ehn et al., 1993, Newell & Cairns, 1993, Kyng et al., 1995, Greenbaum & Kyng, 1995).

A number of candidate HMI behaviours are readily identifiable e.g. speech, handwriting, sign-language, manipulation of objects in virtual environments, speech with gestures. Each of these examples is derived from behaviour that developed independently of the computer. Each exists independently of the computer.

However, before music or typewriter keyboards existed it would have been hard to anticipate that it was possible to play Bach or that the average person could master touch-typing or that this type of behaviour could be useful and as prevalent as it is today. This is an example of how technology can facilitate the emergence of new behaviours that are not readily observed in everyday life.

Thus, *the set of behaviours useful for human-machine interaction is greater than the set of behaviours readily observed in everyday life*.

It is proposed that humans possess many latent abilities that may emerge through interaction with machines. In other words, *human-machine interaction can be considered to be emergent behaviour*. Thus, human factors research into gestural human-machine interfaces does not need to restrict itself to harnessing "natural" modalities of expression. In the case of the people with severe speech and motor impairment due to cerebral palsy (SSMICP), "natural" expressive ability is constrained by neurological impairment although cognitive function may be intact. Interaction with other people and the environment can be limited and frustrating with a high rate of failure. In this case, computer recognition of such behaviour would be of limited benefit.

The field of augmentative and alternative communication (AAC) has developed to supplement "natural" modalities and to offer more effective "alternate and augmented" modalities of communicative expression. An early example of AAC would be a simple board containing words or picture symbols that were selected through pointing (Feallock, 1958, Goldberg & Fenton, 1960, Sayre, 1963). The advent of computer technology and affordable speech synthesis led to the development of the voice output communication aid (VOCA). A typical configuration would require the targeting of one or more electromechanical switches to select graphic symbols in combinations that are translated to words and phrases output using a speech synthesizer. An historical perspective is given by Zangari et al. (1994).

However, as with non-physically impaired populations, AAC users frequently communicate multimodally. Typically, they use combinations of dysarthric speech, vocalisation, eyegaze, facial expression and gesture. Unlike unimpaired populations, this population finds precise targeting particularly challenging, the very behaviour afforded by the low-technology pointing board or high-technology switches of AAC devices. For many people with severe speech and motor impairment due to cerebral palsy targeting switches is difficult or impossible.

It is hypothesised that:

there are other behaviours that can be recognized by technology, in particular gesture; these can be developed into new methods of HMI in general and AAC in particular; these methods will offer advantages over conventional switch based human machine interaction.

For the purposes of this thesis, *gesture* is liberally defined as:

any movement of the body that can be used for the purposes of communication.

The gestures performed by subjects in this research project were not prescribed, instead they were generated from the subjects' knowledge of the world. It is recognised that gesture can be classified using a number of taxonomies (Efron, 1941, Birdwhistle, 1970, Kendon, 1980, McNeill, 1992) however, detailed categorisation of the subjects' gestures lies outside the scope of this thesis and is left as an area for future work.

Gestural human-machine interaction involves the transduction of human movement into streams of data orders of magnitude larger than conventional switch-based HMI. The transduction of movement can be provided by body instrumentation or video camera. The body instrumentation approach was adopted for this study as there are many fundamental problems of computer vision (e.g. occlusion, object invariance) that have yet to be adequately solved before reliable body position data can be extracted and processed (Bichel, 1995). The use of body instrumentation allows the investigation of the computer recognition of human movement based on the assumption that high fidelity body position data is available.

However, even when accurate body position data is available, our knowledge of how to automatically recognise communicative and expressive movement such as gesture or sign language is still in its infancy. As reflected in the literature review presented in this chapter, a number of techniques have been applied to this problem, ranging from hidden dynamic time warping, hidden Markov models, dynamic programming, feedforward and recurrent neural networks, knowledge based methods. Various means have been devised to extract features from the gestural data stream; however, there is still an insufficient body of evidence to offer clear guidance as to which scheme to adopt for a particular set of gestures. One of the key issues is that of automatic segmentation. The most common approach is to look for the boundaries between a) one gesture and another and b) gestural and non-gestural movements. However, the literature on transcription, manual coding, and segmentation of communicative and expressive movement (i.e. verbal and non-verbal elements of language) also reveals the difficulties associated with this process (Bloom & Lahey, 1978, van Balkom & Heim, 1991). It is proposed that the task of segmentation is as complex as the task of recognition. This implies that approaches that involve determining the beginning and end of gestures based on simple features (e.g. movement thresholds) are likely to be useful only in very limited situations. The approach favoured in this thesis is one of segmentation by recognition, a strategy that intuitively seems close to the method used in manual segmentation.

The movement of people with SSMICP is considerably different in nature to that of people without motor impairment in many respects e.g. range, quality, timing, effort, and controllability. Compared to neurologically unimpaired movement, the ratio of "signal" to "noise" in cerebral palsy movement is low. However, the nature of the noise and the relationship to the signal is not well understood. As movement characteristics and abilities vary considerably between individuals with SSMICP, the nature of the movement signal is likely to be idiosyncratic. Thus, from a technological perspective, the

problem can be described as the computer recognition of gestures with a *high noise component* where the movement signal is likely to *vary considerably between individuals*.

Thus, the problem of automatic recognition of gestures of people with SSMICP is different from the recognition of gestures from people without impairment. Therefore, it is a large assumption that a system that is designed to perform optimally on the gestures of people without motor impairment will perform optimally with people with SSMICP. The validity of this argument is supported by the results of Cairns (1993). He found that although recognition rates using two classification methods were similar for people without motor impairment, they differed significantly when the recognition algorithms were presented with data from disabled people. Consistent with this perspective, research reported in this thesis was concerned only with the automatic recognition of the gestures from people with SSMICP.

1.2 Aims and Objectives

The ultimate purpose of the research effort is to acquire the knowledge to develop a viable and robust method of gestural human-machine interaction for people with SSMICP that maximally enhances and extends the user's expressive abilities. Potential applications of this technology include: AAC device, computer/internet access, interaction in virtual environments, wheelchair/robotic control, smart house control.

As previously stated, a human factors led approach has been adopted. It is recognised that human-factors and technological research and development are closely linked. In some respects the two areas can be treated independently, in other respects they are highly dependent on each other.

The programme of research focuses on both areas:

- The gestural abilities of people with severe speech and motor impairment due to cerebral palsy.
- The automatic recognition of the gestures of people with cerebral palsy.

More specifically the research attempted to:

- Develop appropriate cognitive frameworks suitable for exploring and developing gestural repertoires that are candidate behaviours for human-machine interaction.
- Develop and document the gestural repertoires of people with SSMICP.
- Collect gestural data using body instrumentation from people with SSMICP.
- Develop and compare gesture recognition algorithms based on neural networks using data from people with SSMICP.
- Make recommendations for further research in this area.

1.3 Previous and Related Work

1.3.1 Use of Gesture and Sign in severely speech and motor impaired Populations

Gesture is recognised as an integral part of human expression and communication (Wundt, 1921, Efron, 1941, Cherry, 1957, Leroi-Gourhan, 1964, Ekman and Friesen, 1969, Birdwhistle, 1970). Like language, gesture conveys information, often adding to the information content of speech. It develops alongside language, and like speech, it

increases in sophistication as a child develops (Vygotsky, 1937, Werner & Kaplan, 1963, Riseborough, 1982, McNeill, 1992). McNeill asserts that a linkage between gesture and speech exists from an early stage forming part of a developing language-gesture system. This system takes ten or so years to mature into a system that approaches that of an adult. A current review of research into language and gesture can be found in Emmorey and Reilly (1995).

In the case of children with severe speech and motor impairment, this developmental process is disrupted. There have been various attempts to understand the communicative processes of such children (Lloyd et al., 1990, van Balkom, 1991, Kraat, 1991, Gerber & Kraat, 1992, Soto & Olmstead, 1993, Heim, 1994, Helmquist, 1994, Letto et al., 1994). Although adequate developmental models have yet to be devised (Levelt, 1994, von Tetzchner et al., in press), experience shows these populations can acquire language and frequently use gesture and other non-verbal methods of communication.

Light et al. (1985) analysed the communicative interaction patterns of eight congenitally physically disabled children (between the ages of four and six) and their primary caregivers. They found that the children used multiple modes to communicate. 81.8% of their communicative turns were conveyed by means other than their AAC language board, e.g. vocalization, gesture, eye-gaze used alone or in combination. Only 18.2% of turns involved use of their communication board. Another important observation was that their preferred modalities of expression showed marked variations between subjects. Across subjects, no overall modality preference could be determined.

Idiosyncratic systems of non-spoken communication often develop without formal intervention between children with SSMICP and their familiar communication partners, particularly their parents. However, Houghton et al. (1987) and Rowland (1990) showed that spontaneous attempts at communication that are subtle and idiosyncratic may be ignored or misinterpreted. A recent AAC user survey conducted by Murphy et al., Stirling University (1995) found that 22.2% of users did not use their AAC systems at all for "informal" use e.g. at home or with friends.

The formalisation of development of an idiosyncratic gestural system (e.g. the creation of a gesture dictionary for each individual) has been suggested as good practice in a number of texts on AAC (Musselwhite & St. Louis, 1982, Siegel-Causey & Guess, 1989), but in practice this seldom happens. The negative aspect of individualised gestural repertoires is that typically they exhibit limited transparency and have to be learnt before they can be used by unfamiliar people (although documentation of their gesture dictionary carried by the user can help in this respect).

Instances of documented development of formal systems include Hamre-Nietupski et al. (1977). They identified around 160 "natural" gestures. They remark that the advantages of using their system included ease of production through "gross motor" movements, and a high level of transparency through iconicity. They provide a detailed guide for use within a functional curriculum.

Many communicative gestures and sign language lexemes involve substantial use of hand shape and hand movement. People with SSMICP often have limited hand control and a high degree of spasticity potentially severely limiting their expression in these modalities.

A comprehensive review of gestural and sign based AAC is presented in Musselwhite and St. Louis (1982). Duffy (1977 reported in Musselwhite & St. Louis, 1982) developed a system of 471 signs formed by combinations of "gross" gestures, some accompanied

by vocalisation. The system was initially developed for four subjects with quadriplegic athetoid cerebral palsy aged seven to 15. Wherever possible signs were iconic.e.g. crossing the legs for PANTS (trousers) or pantomiming pulling up trousers. A general sign for time would precede time category signs e.g. one o'clock = TIME + 1, days of the week (Sunday = DAY +1), months were indicated by the sign for time plus the first letter of the month and another letter (December = TIME + D + "eeee"). Questions were indicated by making the sign QUESTION followed by the appropriate sign (WHO, WHAT, WHERE). Facial movements and whole body movements were used in expressing feelings. Other categories included people, places, adjectives, prepositions and pronouns.

Similarly Skelly et al. (1979 reported in Beukelman & Mirenda, 1992) developed Amer-Ind Gestural Code based on American Indian Hand Talk. The repertoire includes 250 concepts labels e.g.: QUIET, made by holding the index finger of the hand to the lips. This sign may have many referents e.g. silent, calm, dormant hush, low (noise), mute, noiseless, serene, silence, silent, still, tranquil. The intended meaning is determined through context. This feature enables the expansion of the repertoire to a vocabulary equivalent to 2500 words. The initial limited repertoire is expanded through "agglutination", the principle that allows for the invention of new ways to express concepts e.g. insane = BRAIN + FLY + DISTANT. 80% of the repertoire can be executed using one hand and requires moderate motor control. They report that 80%-88% of the hand signals could be recognised by untrained observers. However more recent work by Doherty et al. (1985) suggested only 50-60% of the hand signals are recognisable by non-disabled adults when presented without reference to their conceptual categories.

An example of a system developed by an AAC user taking into account physical abilities and limitations is the White's Gestural System for the Lower Extremities created by Cathy White who has a severe hearing loss and cerebral palsy with severe upper extremity involvement. The system, developed together with her mother, comprised 125 "leg signs", using foot, toe, heel, knee, ankle, calf and thigh touch points to convey messages in a variety of linguistic categories such as people, actions, or objects. (Huer, 1987 reported in Beukelman and Mirenda, 1992).

Musselwhite and St. Louis (1982:124) discuss and review pantomime used in language therapy for severely communicatively impaired populations. They suggest that pantomime is accessible as both an input and output system due to its high iconicity, although they remark that "pantomime would probably not be selected as the primary long-term approach due to its relative inefficiency and limited scope".

The "movement based" approach to language development originally developed for deaf-blind children by Van Dijk (1966) has been adapted for use by people with severe speech and motor impairment by Siegel-Causey & Guess (1989). Based on the principle of learning through doing, the adapted technique comprised six levels: *nurturance*, development of a warm positive relationship conducive to communicative interaction; *resonance*, rhythmic movements involving direct physical contact with the aim of shifting attention from self to the external world; *co-active movement*, extension of resonance to develop sequence and anticipation; *non-representational*, teaching relationship between world and graphic representation; *deferred imitation*, teaching imitation of facilitator movement; *natural gestures*, facilitating communicative gestures that are "self-developed".

The advantage of prescribed gesture and sign systems is that once learnt they form a

common system of communication. However, given the significant variation in physical ability found in cerebral palsy (CP) populations, the disadvantage is that many individuals will only be able to access a subset of the gesture or sign lexicon. On the other hand, a system developed to match the idiosyncratic range of abilities of an individual, e.g. Whites Gestural System, has the advantage that it can be tailored to utilise the unique range of abilities of an individual. This is likely to result in a substantially larger lexicon. The disadvantage is that communication partners have to learn a system that applies to only one individual.

Computer recognition offers the possibility of the gestural HMI being used as a translation system. In this case, the gestural HMI could be trained to recognise an individual's idiosyncratic gestural repertoire. The machine could perform a translation into a commonly understood form e.g. gesture to speech. In this way, the lexicon size could be maximised while still being understood by all communication partners.

Gesture and sign based AAC systems have probably become less frequently promoted by speech and language therapists with the advent of switch-input electronic communication devices. However, the gesture and sign systems reported in this section, suggest that gestural HMI comprising small lexicons can be developed into a viable communication system.

1.3.2 Gestural Human Machine Interaction for People with Motor Impairment

Harwin (1990) at Cambridge University, pioneered research into the computer recognition of head gestures of people with severe speech and motor impairment due to cerebral palsy. He worked with a 23 year old non-vocal cerebral palsy quadriplegic subject who had developed a relatively unusual mode of communication involving tracing the shapes of letters of the alphabet with head and eye movements. A "Polhemus Isotrak" magnetic tracker was attached to the head to capture head movement. A simple head model, together with the concept of an imaginary head-stick cutting a plane was used to transform the six-degree of freedom magnetic tracker data to two-dimensions. This reduced the recognition task to one of recognising two dimensional movement. The recognition algorithms employed hidden Markov models, a syntactic-statistical pattern recognition method to classify head gestures.

The gestural data stream was automatically segmented by thresholding movement variance in the x and y direction over 10 samples. Six planar features were extracted from each sample and assigned a symbol. The symbol was determined by finding the nearest cluster centre. Cluster centres were determined using a k-means clustering algorithm using a set of training data. After feature extraction and clustering, each gesture was represented by a sequence of symbols. These data were input to a set of hidden Markov models previously derived from training data to classify the gestures. Recognition rate of 83% for two head gestures (yes, nodding) and (no, shaking). In the second application using five head gestures ("yes", "no", "C", "L", "W") rates of 51% were achieved. The results showed that recognition of head gestures from people with SSMICP is feasible, but the recognition rate needs to be increased. A number of improvements were suggested for future work aimed at increasing the recognition rate. These included: improved feature extraction/gesture coding, and employing larger models trained using a larger set of training data. This work highlighted the difficulty encountered in attempting to automatically segment head gestures from movement data

containing a significant athetoid component.

Perricos (1993, 1994) continued the work on head gestures recognition at Cambridge University. He constructed a real-time recognition system based on principle coefficient analysis and dynamic time warping. As with Harwin's system, the gestures have to be segmented before they can be classified. This was achieved using a "tremor filter" which determined a "still threshold" and a "movement threshold". The gestures used were six directional gestures: "yes", "no", "up", "down", "no", "left", "right", and up to nine "complex" gestures e.g. "T", "O". In an early report (1993) he outlines preliminary results recognition rates of 54.6% for a subject with quadriplegic cerebral palsy and limited head control, and 92.5% for a subject with paraplegic athetoid cerebral palsy and relatively good head control. In a more recent publication he reports results for six subjects: "five subjects with varying levels of athetoid cerebral palsy" and one with Friedrichs ataxia. Real-time recognition results ranged between 44.2% and 98.8%. This was an average of 6.3% below human recognition rates.

Cairns (1993) at the MicroCentre, University of Dundee looked at the computer recognition of dynamic arm gestures from people with and without motor impairment. Disabled subjects had motor impairment due to cerebral palsy and were between the ages of 9 and 55. The subjects were "asked to provide examples of gestures they found easy and natural to make. No definition of the form the gesture should comprise was given. No meaning was attached to what the gesture should mean". The size of each subject's gestural repertoire was from two to five gestures. The gestural repertoires of only two of the nine disabled subjects appear to have had any symbolic referent either iconic or linguistic. The other gestures were described in terms of their component movements e.g. "raise arm from rest position to almost horizontal", "small oscillation almost at rest position", "up and down in the z-plane, arm from rest to head", "movement in the z-plane in front of the body", "hand moved up to mouth (food sign)", "hand moved up to hair (brush) sign". Cairns reports substantial difficulty in automatically segmenting the gestures despite asking the subjects to pause between gestures. The gestural data were semi-automatically segmented and 15 feature sets derived from this data. Feature sets contained three-dimensional movement information including measures of position, velocity, acceleration, and power spectral density. Early pilot work examined the use of linear discriminant analysis (LDA), feedforward neural networks (FFNN) trained using backpropagation, and recurrent neural networks (RNN) using a more limited range of feature sets prepared from people without motor impairments. Relatively low average recognition rates were obtained: LDA 51%, ANN 63%, and RNN 42.9%.

Subsequent work focused on the comparison of dynamic programming (DP) and hidden Markov models (HMM). He compared the performance of these algorithms using the various feature sets as input (15 for DP and 10 for HMM). Three training methods were compared: training using data only from the first data collection session and testing on data from all subsequent sessions; training and testing using data from a single data collection session; and dynamically adapting reference templates/models. A recognition rate increase of a few per cent was achieved by training at the beginning of each data collection session rather than training once on data from the first session. In general the adaptive methods resulted in poorer performance. For disabled subjects dynamic programming consistently outperformed hidden Markov models resulting in best recognition rates of 76.8% and 70.2% respectively. Interestingly, the relative performance of feature sets varied between disabled subjects. The feature set comprising relative x,y,z performed consistently well. Future work will include implementation in

real-time and addressing the segmentation issue.

Harrington et al. (1995) at University of Oxford and the Oxford Orthopaedic Engineering Centre reported using six single axis accelerometers attached to the forearm to transduce arm gesture data from two subjects with athetoid cerebral palsy. The acceleration signals were pre-processed by a moving horizon estimator. The output of the estimator provided the input to a classification algorithm based upon dynamic programming. The gestural repertoire comprised five to eight gestures. These were "swim", "hammer", "wave", "run", "wind-up", "shake", "pluck" and "paint". Preliminary computer recognition rates ranged from 52% to 63%.

Rogers et al. (1992) from the Wright-Patterson Airforce Base, Ohio, USA in the second part of their paper on the application of artificial neural networks for the processing of raw pixel data for segmentation, tracking and identification include an application to recognise the facial expressions of a young girl with cerebral palsy. They report the use of the Karhunen-Loeve transform (KLT) feature extraction and normalized k-nearest neighbour classification for the recognition of three facial expressions: mouth closed, mouth open, and tongue out. Using from two to nine KLT coefficients, recognition rates ranged from 82% to 94% respectively. They state that their current efforts are to eliminate false alarms caused by scale rotations of the child's head with respect to the prototype (Goble et al., 1993).

1.3.3 Gestural Human-Machine Interaction for Other Populations

Pen/Mouse-Based Gestures

One category of research into gestural HMI strongly linked to handwriting recognition focuses on pen-based gestures. Using pen and paper simulation, proof reading type gestures for text editing were shown to have efficiency advantages over keyboard input (Wolf & Morrel-Samuels, 1987). This was implemented by Kim (1988). Four subjects each produced gestures from a repertoire of 32 gesture classes. 73% were correctly recognised, 14% were mis-recognised and 13% were not recognised. The recogniser used a feature called "direction change". Directions were quantised into one of 12 directions (as in the 12 directions of a clock-face). After smoothing the input, each gesture was segmented into a sequence of strokes and then quantised. It was suggested that this method was relatively insensitive to natural variations that appear in gestures such as non-linear scaling, mirror images, rotation and production with reverse directions. The gesture recogniser was further refined by Lipscomb (1991).

Kurtenbach's "VirtualStudio", a graphical interface used mouse gestures control connections among audio devices (1988 reported in Kurtenbach & Hulteen, 1990). "Paper and pen" types of gestures were used to move, copy, delete, connect, and encapsulate icons and arcs.

Rubine (1991) developed GRANDMA gesture based drawing program. He examined both single path and multiple path gestures produced with a mouse or stylus. A vocabulary of 30 single stroke gestures were recognised at 97% using a gesture recognition algorithm based on linear discriminant analysis. In this system, the start and finish of gestures had to be specified (e.g. mouse button press, stylus contact). Of particular interest was the proposed "eager" recognition strategy: Once the gesture starts, the recogniser continually tries to recognise it. As soon as the system has had enough of the gesture and is confident which gesture is being produced, the command is issued

rather than wait for the whole gesture to be completed.

All the above research projects involved movement across a planar surface, either directly with the tip of the finger or using an artifact e.g pen or mouse. Without modification, the applicability of these interfaces to people with SSMICP is questionable. The effort required to maintain contact with a plane is likely to be high for many individuals with SSMICP. An exception would be the case of hypotonicity (low muscle tone), where a planar surface such as table or wheelchair lap-tray may perform an assistive role in countering gravity. Freedom to move in any plane or any number of planes is more likely to achieve the objective of maximally utilising motor ability. This requires automatic sensing and recognition of unconstrained movement in three dimensions. Much less is known about the recognition of three dimensional movement. Obviously, the problem is more complex due to the increase in dimension. Also, unlike pen-gesture recognition, gesture cannot be segmented by removal of the pen from a tablet.

Coverbal Gestures

The early work of Schmandt and Hultheen (1982) as part of Richard Bolt's group within the Architecture Machine Group, Massachusetts Institute of Technology (MIT), USA, (Bolt, 1980) involved combining the automatic recognition of speech and deictic gesture to allow phrases accompanied by pointing such as "put that there" to be interpreted by identifying the associated screen objects. Conceptually they took the approach that "all functions [of the computer] should be controllable by all modes of input".

More recently, this work has been extended by Bolt's Advanced Human Interface Group (AHIG) at the Media Lab, MIT. The AHIG team have extended the concept to include a range of gestures that naturally accompany speech (Bolt & Herranz, 1992, Sparrow, 1993). This work was inspired by research into the integrated nature of gesture and speech production (Rime & Schiaratura, 1991). They recognised that there was not a clear one to one mapping between speech and gesture. The interpretation of the gestural data relies on the semantic content of speech and the temporal relationship between the gestures and speech. In order to solve this problem, they are taking a feature-based approach to gesture analysis and recognition (Wexelblat, 1994). The work represents one approach to solving the problem of connected gesture recognition where the user can create continuous unrestricted motion. In his thesis, Wexelblat describes a scheme for mapping raw movement data from two Cybergloves and magnetic trackers from various parts of the body to a data stream of a higher level of abstraction, independent of any particular set of biosensors. The demonstrator system called ICONIC is described by Koons (1994). Continuous speech is converted into LISP-like "semantic frames". An "interpreter module" parses buffers of speech semantics and gestural features. An object-base module manages a set of objects and their appearance on a large-format display and supplies the interpreter with information. The instruction "move the chair like this", plus an appropriate gesture moves a graphical image of a chair as depicted by the gesture. The action in gesture-space is mapped to the chair in graphics space and is used to construct a corresponding object-base manipulation command.

Although the speech component of this interface is not accessible to people with SSMICP, it is conceivable that their gestures could be translated from raw movement data to higher level representations before gesture interpretation or gesture recognition. The advantage would be the ease with which it can be combined with contextual information. However, it is not yet known how such a system would perform when

presented with “noisy” gestures.

Finger Spelling/Sign Language

Kramer & Leifer (1989) at Stanford University, USA, developed a system that is now marketed as commercial software available to accompany Virtual Technologies' CyberGlove®. The original system was designed as part of a communication aid to allow people who are non-vocal deaf and deaf-blind to communicate with hearing people who can't sign. It involved the use of a glove instrumented using strain gauges to monitor finger and thumb joint angles. A neural network was used to recognise the American finger spelling alphabet.

In the GloveTalk I pilot study, Fels and Hinton at the University of Toronto, Canada (1990) used the VPL DataGlove® and a Polhemus Isotrak® magnetic tracker connected to a DECTalk® speech synthesiser via five neural networks trained using backpropagation. Using a 203 item gestures-to-word vocabulary, they report that the wrong word was chosen less than 1% of the time, and no word was produced 7% of the time. Each network was ascribed a separate task. Dedicated networks determined: Root-word from hand-shape, word-ending from hand-direction, word-rate from hand-speed, and word-stress from hand-displacement. The gesture was segmented using a fifth network referred to as the “strobe network” which continuously monitored scalar velocity and acceleration from the magnetic tracker. A set of handshapes was devised loosely based on those of American finger spelling. Each handshape was mapped to a word. The interface was novel in that it used motion in five directions (up, down, forward, back, left, right) to control word endings and to indicate that a handshape was ready for detection.

Maurakami and Taguchi (1991) at the Human Interface Laboratory, Fujitsu Laboratories, Japan, investigated the application of feedforward and recurrent neural networks to Japanese sign language recognition. They investigated the computer recognition of 42 Japanese finger alphabet as static signs and the recognition of ten dynamic signs using an instrumented glove and magnetic tracker. The first experiment used feedforward neural networks trained by backpropagation of errors. Ten finger angles plus yaw, pitch and roll were normalised and used as input to the network. The system was instructed when the handshape had been made. Using 206 training exemplars, they reported an initial recognition rate of 98% for signs from the user who supplied the training data and 77% when used by other signers. Using training data from six users reportedly yielded recognition rates of 94.3% for the group involved in training the network and 92.9% for other users. In the second experiment, they looked at the computer recognition of ten signs from Japanese sign language. The signs were segmented by recognition. The feedforward network was used from the first experiment to determine when a sign had been produced. This was signalled by the activation level at an output node reaching a threshold value. This technique also separated sign movement from non-sign movement. Sign data including x,y,z position was then input to an Elman recurrent neural network to classify the sign. The ten signs represented a range of sign-types selected from Japanese Sign Language: “skilled/unskilled” - differing only in the direction of movement; “father/mother” - similar for the first movement, but differing in the second; “memorise/forget” - same movement but in opposite directions; “brother/sister”, and “like/hate” - same movement and similar hand postures. Recognition rates of 96% were reported.

Ohki et al. (1994) at the Central Research Laboratory, Hitachi, Japan, report using dynamic programming to recognise 17 signs from Japanese Sign Language produced

continuously. It was reported that this system can translate simple sentences e.g. "I have a stomach ache". They also report that another system can recognise 100 signs but ignores sentence structure.

Kurokawa (1992) addressed the problem of gesture coding and the creation of a gesture dictionary for gestural and sign language HMI. A generalised gesture interface architecture was proposed that recognises signs, indication, illustration and manipulation with corresponding synthesis in the form of a "gesture display". He suggests that pattern recognition techniques such as neural network and template matching that simply map a gesture pattern to meaning are unlikely to scale up. His solution is to map predefined units of gesture to "symbols similar to alphabets in an early stage of processing". In this respect the approach is similar to that of the AHIG at MIT. Kurokawa suggests "gestures have four kinds of representation in the interface: quantified, code, meaning and graphics representations". A "gesture dictionary" is presented that can be used to convert one form into another. The coding scheme is based on Stokoe (1960) for handshape and Shibata et al. (1984) and Hirsbrunner et al. (1987) for body-shape.

1.4 Overview of Thesis

Chapter 2: Systems Approach to Human Machine Interaction

The first half of this chapter examines the proposition that human-machine interaction can be conceptualised as an emergent property. The relevance to people with severe speech and motor impairment is discussed particularly in relation to the possibility of eliciting new behaviour potentially useful for HMI. This argument is related to the methodology developed for the gestural ability pilot study (detailed in chapter 4).

The second half of the chapter examines the machine perception of human behaviour in the context of gestural HMI for people with SSMICP. A number of key issues relating to the problem are detailed and related to the studies presented in the thesis. A system architecture design is presented and related to the system components investigated in the computer recognition study detailed in chapters 6 to 8.

Chapter 3: Human Factors

This chapter begins with detailing the profile of the users of the proposed interface and the nature of their communicative interactions. The problems that this group have accessing technology are described. Technology currently used for augmentative and alternative interaction is detailed. Implications for user requirements and design issues for gestural HMI are discussed.

Chapter 4: Gestural Ability Pilot Study

Twelve subjects aged 5yrs 9m to 18yrs 1m with severe speech and motor impairment took part in this study to explore emergent gestural ability. Existing methods of augmentative and alternative communication were documented. Therapeutic, educational and medical records were examined for documentation of gestural ability. Results are presented.

Ten subjects were video-taped while performing gestures elicited using human-human interaction in the form of a student-centred "charades-like" game. Video-tapes of the interactions were analysed to determine the range of body parts involved in gesture production. Results indicated that multiple body parts and multiple modalities were involved in the gestural responses. The most frequently used body part was either the

right or left arm for nine out of ten subjects. Transcription of gestural repertoires for each subject are presented in appendix A.

Chapter 5: Gestural Data Collection for Pattern Analysis

A subset of 27 gestural classes was chosen from the gestural repertoires elicited during the gestural ability pilot study for one subject with spastic-athetoid cerebral palsy classed as having no functional use of upper extremities. The gestures involved one arm as a principal component. A six-degree-of-freedom magnetic tracker was attached distally to one forearm. Dynamic arm gestural data were recorded at 100 samples per second in 10 minute blocks. 720 gestures were elicited in random order at a rate comfortable for the subject. The gestural sequences were co-articulated i.e. there were no pauses between gestures and the transition from one gesture to the next could follow a wide range of paths. The average rate of production is determined.

Chapter 6: Examination and Processing of Gestural Data Stream

Gestural data was animated on a Silicon Graphics² Workstation. A body model was developed using MATLAB³ scripts and animated on a Silicon Graphics Workstation. An animation tool was developed and used to play-back the gestures at different speeds and view the movement from different angles. Gestures were examined qualitatively using the computer graphics animation. A strategy for gesture recognition is developed. Key factors relating to the recognition problem are identified. A fixed-time window scheme using feedforward neural networks in a time-delay scheme is proposed. The advantages and disadvantages of such an approach are discussed. Gestures were manually segmented using the animation tool. The scheme adopted for signal conditioning and data reduction involving the use of a low-pass filter is detailed and discussed.

Chapter 7: Gesture Classification using Neural Networks

The ability of time-delay feedforward neural networks to classify the gestural data was investigated. Neural networks were coded using MATLAB and the MATLAB Neural Network Toolbox. The neural networks were trained using backpropagation of errors with momentum and adaptive learning rate. The optimum number of hidden neurons for the problem was determined experimentally. Neuronal activation functions are compared. Neural network classification is compared with two types of k-nearest neighbour methods. Results are summarised and discussed.

Chapter 8: Feature Set Comparison using Neural Networks

First, the recognition results of seven feature sets comprising gesture segments of xyz data varying in duration from 160ms to 1120ms (four to seven time samples re-sampled every 160ms) were compared. Optimum recognition rates were obtained with a gesture segment length (time window) of 640ms.

Twenty additional feature sets extracted within this time window were examined involving features that encoded forearm orientation (8), scalar and vector velocity (7), curvature and plane of motion (5). Results indicated that a number of feature sets gave similar results. Four time samples of x,y,z position gave the highest recognition rates.

² Silicon Graphics is a trademark of Silicon Graphics Incorporated.

³ MATLAB and MATLAB Neural Network Toolbox are trademarks of The MathWorks Inc.

Chapter 9: Conclusions

The main results are summarised and discussed in relation to the aims and objectives of the thesis. Ways in which the research has contribution to the field are proposed. Recommendations for future work are detailed.

Appendix A: Gesture Elicitation Sessions: Transcripts

Transcriptions of the gestural ability pilot study are presented for 10 subjects.

Appendix B: Neural Network Training Algorithm

The backpropagation of errors training algorithm and weight initialisation methods are described.

Appendix C: Confusion Matrices

The gesture recognition results for all experiments are documented in the form of confusion matrices that relate the actual gesture class to the decision of the pattern classifier. Results are presents for training and test data.

Appendix D: Hinton Diagrams of W_1 and W_2

Hinton diagrams are a graphical representation of the neural network weights and bias magnitudes. They are documented for each feature set.

Chapter 2

Systems Approach to Human Machine Interaction

2.1 Introduction

The systems approach to modelling traditionally highlights the need for adequately considering both the part and the whole (Thome, 1993:15). More recently, this approach has been further underpinned by the developing area of complexity theory that has found wide application in describing phenomena in the real world from thermodynamics to economics (Nicolis & Prigogine, 1989, Waldrop, 1992). This paradigm is applied to the field of human-machine interaction and used as a framework for describing the problem domain of this thesis. Within this framework, the arguments are made for adopting a human-factors driven approach and the development of the methodology involving human-human interaction used in the gestural ability pilot study.

In the second half of the chapter, issues relating to the machine perception of human behaviour are discussed. A basic system architecture is proposed. The system components and associated research issues investigated in this thesis are described in relation to that architecture.

2.2 Human-Machine Interaction as Emergent Behaviour

Human-machine interaction is a complex process. A number of theories and principles have been applied to the analysis of HMI (Foley & Wallace, 1974, Card et al. 1980, Shneiderman, 1980, Kieras & Polson, 1985, and Norman, 1988 reviewed in Shneiderman, 1992 and Eberts, 1994). Newell's critique of the application of the goals,

operators, methods, selection rules (GOMS) technique to HMI involving AAC users highlights the some of the problems encountered when attempting to apply methods of formal analysis to people for who are severely speech and language impaired (1992). Olson and Olson (1990:223 reported in Newell, 1992) conclude “the GOMS model fails to capture the user’s fatigue, individual differences, or mental workload”. Newell argues: *“There is very little quantified knowledge of the ways disabled people use technology, but clinicians are well aware that different users adopt a wide range of strategies to cope with the individual nature of their disabilities. It is thus unlikely that a single model can be used to characterize any non-trivial task”* (1992:89).

In other words, we know that disabled people adapt to their environments but we cannot adequately model such behaviour, at least at this present time.

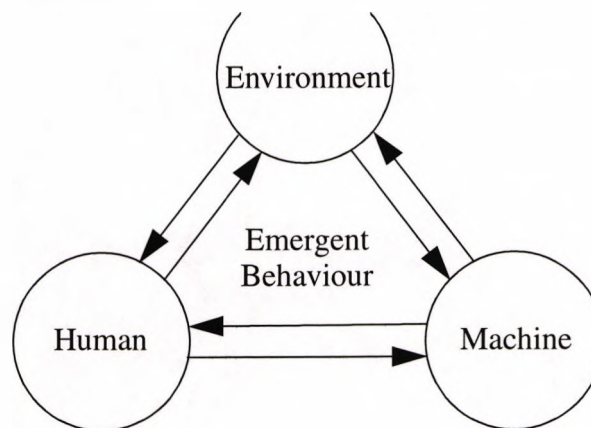


Figure 2.1 Proposed conceptualisation of human-machine interaction. Human, machine, and environment are considered as a system of dynamically interacting non-linear sub-systems. Such systems are known to exhibit emergent properties.

In light of this state of affairs, an alternative approach is proposed. A conceptualisation of HMI is proposed in figure 2.1 which anticipates limits of predictability of human-machine interaction. Human, machine, and environment are considered as interacting dynamic non-linear complex sub-systems. Information flows back and forth between components changing the state of each component. The human component is self-organising and self-adaptive i.e. a system that can modify itself and adapt to its environment. As machines become more “intelligent”, they can be expected to become increasingly self-organising and self-adaptive. The immediate environment is represented by an open circle denoting that it is comprised of components from the environment at large that enter and exit the sub-system.

A numbers of important propositions are consistent with this representation:

- *Human-machine interaction can be considered as an emergent property¹ of the system.*

¹ “Emergent property” is used in the conventional broad sense that stems from the difficulty in predicting system behaviour from examination of its constituent parts. To what degree this infers irreducibility as in Churchland’s definition (1989:51) is difficult to ascertain as we know so little about the neuropsychological and cognitive processes of people with SSMICP.

from this follows:

- *NEW human expressive behaviour may emerge as a result of the interaction.*

In other words the environment and the machine have the potential for facilitating new emergent behaviour².

This is particularly important in the context of HMI for people with severe expressive impairment. The objective of developing a gestural human-machine interface is to enhance and augment "natural" communication. To merely harness observed expressive ability would not be very useful. In addition, it is generally acknowledged that there is a tendency for people with severe expressive impairment to be rather passive in their patterns of social interaction (Selgman, 1975, von Tetchner, 1988, reported in Basil, 1992, Kraat, 1985). Typically, they initiate communicative acts much less frequently than their communicatively unimpaired communication partners. This may at least in part be due to "learned futility"³. That is they tend to be passive and seldom initiate interaction because they have learned that their efforts have a high probability of failure and are likely to lead to frustration. If this is the case for the subjects participating in the study, conclusions pertaining to gestural ability based purely on observations of subjects communicating in everyday settings is likely to be an unreliable indicator of potential.

However, it is proposed that human behaviour harnessed for human-machine interaction need not be constrained to that which has been observed. From a human-factors perspective, the objective of developing gestural HMI for people with SSMICP can be interpreted as developing a system that gives rise to the emergence of latent expressive ability.

In addition:

- *Expressive human behaviour can be critically dependent on the environment and components within that environment, particularly if physical impairment restricts self-adaptive ability.*

People with SSMICP have an impaired ability to adapt to the machine and the environment. The usual remedial course of action commonly adopted in rehabilitation engineering is to adapt or augment the environment and/or machine to compensate. However, people with SSMICP have highly varied profiles of physical and cognitive ability (see chapter 3). For a single machine to meet the needs of a large constituency, it has to have the capacity to be tailored to suit the individual characteristics, needs and preferences of the user. This potentially increases operational and design complexity. Increasing the "intelligence" or self-adaptive ability (maintaining user control) of the HMI, while a challenging goal itself, offers a promising method of dealing with increasing complexity at the interface.

² A good example is the keyboard. Before music or typewriter keyboards existed it would have been hard to imagine that it was possible to perform such rapid ballistic movements of the fingers to perform Bach or type 100 words per minute with little conscious effort for considerable periods of time.

³ I am indebted to Dr. Michael Alexander, Director of Rehabilitation Medicine, A. I. duPont Institute for the suggestion that "learned futility" was more appropriate than the more common terms "learned helplessness" (Seligman, 1975) or "learned dependency" (von Tetzchner, 1988).

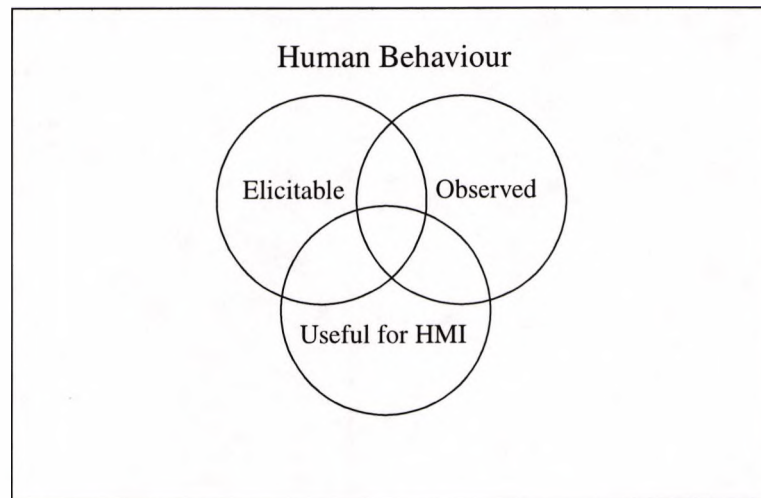


Figure 2.2 Venn Diagram of Human Behaviour. Only a subset of behaviours can be identified by observation. Further behaviours can be elicited through control of the environment. The aim is to eventually harness these for HMI.

Research aims of this thesis can be further conceptualised using the Venn diagram in figure 2.2. The three overlapping circles represent sub-sets of human behaviour that are categorised “observed”, “elicitable”, and/or “useful for HMI”. In the context of HMI for people with SSMICP, “observed” refers to behaviour at any time, past or present, in any setting e.g. during school lessons, therapy sessions, play, sports, clinical examination/assessment, and at home. “Elicitable” refers to behaviour that could notionally be emergent due to exposure to new environments, people and/or machines. “Useful for HMI” refers to the set of behaviours that can be currently harnessed using computer technology.

The aim of the human-factors section of this research project is to increase the intersection between “observed” and “elicitable” by eliciting new behaviour. All observed behaviour is potentially useful for HMI.

The aim in the machine perception section of the project is to work towards increasing the intersection between behaviour “useful for HMI” and “observable” behaviour. At the moment this intersection is small. Only simple behaviours are used in human-computer interaction (HCI) e.g. targeting actions needed to operate a keyboard or planar arm movements needed to operate a mouse. The expansion of the set “useful for HMI” will be a continuing objective for many years to come.

This diagram can be used to contrast other approaches to intelligent HMI development. Approaches that propose that computers should be more like people often infer that HMI should be closer to human-human communication. For example, Negroponte states “the best metaphor for I can conceive of for a human-computer interface is that of a well-trained English butler” (1995:150). This goal would be similar to increasing the intersection between “observed” and “useful for HMI”. While this may be a viable goal for HMI, in itself, this approach would not result in maximally harnessing human behaviour as it ignores the emergent behaviour between human and machine which notionally may be as sophisticated and as complex as human-human interaction, but not necessarily the same. This emergent behaviour may well be more appropriate for many

computer mediated tasks. In the context of HMI for people with SSMICP, as previously stated, the goal of AAC is to go beyond artifact-free “natural” human-human communication and evolve technology mediated enhanced interaction. Thus, the goal of this research is to develop HMI where the human and computer interact to produce behaviour that enhances and extends human ability. As such, it seems reasonable to think of the computer more like an automobile, as an “extension of self”. This is similar to the approach advocated by Shneiderman (1992:546). Inspired by Lewis Mumford (1934), he suggests that preoccupation with anthropomorphic HMI is likely to restrict HMI development as anthropomorphism restricted the development of the aeroplane.

2.3 Eliciting and Identifying Candidate Behaviours for HMI

In the process of attempting to design radically new ways for people to interact with machines there is a “chicken and egg” type of dilemma. How does the designer of a machine that does not yet exist design for emergent behaviour that is unpredictable?

High bandwidth efferent HMI of any sophistication is technologically challenging. As a result lead-time on prototyping is high. However, the resulting computer system hardware and software is likely to be highly dependent on the human behaviour that it harnesses. Information about the user and the way they may be able to use the system is needed at an early stage. In order to address this dilemma, an approach is suggested that involves the parallel investigation of human-factors and technological issues.

In order to develop technology that addresses people’s needs efficiently it is useful to adopt a human-factors driven approach where human-factors and technological issues are considered equally from the onset of the project. The research reported in this thesis is conducted within a methodological framework where issues relating to the human drive the technological development. In respect to this and future work in the area of high bandwidth efferent HMI, it is proposed that this human-factors driven approach:

- is most likely to lead to an HMI design that closely meets the needs of its users;
- assists in establishing clear initial goals for the HMI design;
- establishes a base-line interaction against which prototype HMIs can be compared;
- enables human-factors and technological issues and their inter-relationship to be determined at an early stage;
- facilitates the concurrent development of human-factors and technological components.

Probably the most common methodology for addressing the “chicken and egg” dilemma is the “Wizard of Oz” technique where humans are used to simulate part of a machine. For example, Newell et al. (1990) used a human to simulate computer speech recognition in order to investigate the human-factors issues of speech driven human computer interaction (HCI) before the availability of reliable speech recognition, and thereby evaluate the feasibility and set goals for usable speech recognition. In the Newell study, some sessions involved leading subjects to believe that they were talking to a computer. In other sessions, the subjects were informed that speech recognition was simulated by a human.

In this thesis, the goal was to identify behaviour potentially useful for HMI in a generative manner. In this context, the assumptions built into a “Wizard of Oz” design may have constrained emergent behaviour. It was more appropriate to employ human-

human interaction (HHI). A methodology was developed that involved using human-human interaction in carefully managed environments to elicit behaviour that was a candidate for human-machine interaction. Interpersonal interaction was thought to be the most effective method of engaging the subject cognitively. However, an important distinction should be drawn between using HHI to elicit behaviour that then can be used as the basis for HMI and attempting to imitate HHI. As previously discussed, this research was not concerned with imitating human-human communication within HMI (Ebert, 1994:454-467). The objective was to be generative with a view to *enhancing and extending* existing modalities through the application of technology.

A methodology was developed for the gestural ability pilot study that drew upon techniques from the performance arts. Drama and mime were employed to create a cognitive framework drawing upon mental imagery that facilitated the exploration of a wide range of candidate behaviours from children and adolescents with severe expressive impairment and a range of cognitive ability. This process is presented in systems diagram in figure 2.1.

It was proposed that these techniques could be used to minimise the masking effects of “learned futility” and facilitate the exploration of a subject’s potential. A protocol for human-human interaction was developed that was designed to:

- engage the imagination and encourage creative responses;
- defocus from the subject’s disability;
- defocus from any negative and constraining concepts of self.

A facilitator with experience of working with severely disabled children interacted with each subject in a way that attempted to sustain motivation while minimising fatigue through management of the interaction.

This methodology was used to explore and document emergent gestural behaviour in the gestural ability pilot study (detailed in chapter 4).

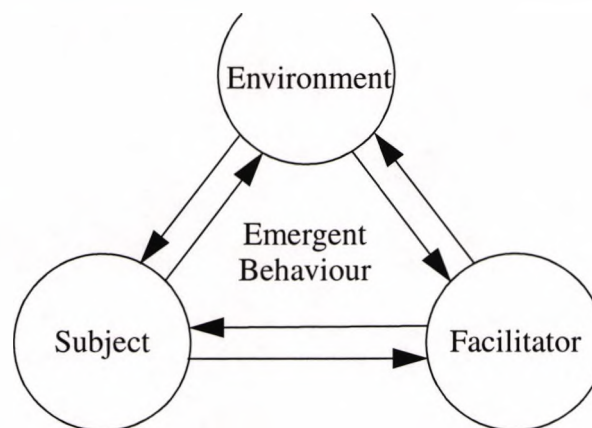


Figure 2.3 Systems model of human-human interaction used to elicit candidate behaviour for HMI. Drama and mime were used to minimise the effects of “learned futility” and to create a cognitive framework that encourage the production of a wide range of behaviours.

2.4 Machine Perception of Human Behaviour

2.4.1 High Bandwidth Efferent Human Machine Interaction

Increasing the physical bandwidth of interaction requires the use of a method of transducing physiological parameters into streams of digital data. Notionally these parameters could be any signal or combination of signals that contain information relating to the user's intention. Categories include:

- Neuro-muscular action leading to movement e.g. of the hand, arm, torso, legs, head, facial expression, eye-gaze, eye-blink.
- Neuro-muscular action with minimal movement e.g. isometric muscular contraction, muscular force applied to a fixed object.
- Signals directly from central nervous system e.g. electroencephalography. A number of studies have looked at this possibility (Hiraiwa et al., 1990, Granger, 1993).
- Signals from the vocal apparatus: Vocalisations, speech, tongue movement, breath.

For future HMI to maximise the bandwidth of expression for people with SSMICP it will be necessary to harness multiple signals and their corresponding expressive modalities.

An important consideration in the fusion of sensor data is the parallel nature of multimodal expression. The problem is not as simple as simultaneous occurrence. Vocalisation, facial expression, arm gesture, and head gesture are likely to be related in time. Relative timing and phase can be salient. Also, although multimodal expressive acts may combine to form a single entity, the onset and finish of each component is not necessarily coincident e.g. in the case of coverbal gestures from unimpaired people, gestures are often produced ahead of the corresponding speech act. The investigation of this area was outside the scope of the research project, although it is revisited in the concluding chapter.

2.4.2 System Architecture for Gestural HMI

Although the creation of a real-time demonstrator system is beyond the scope of this study, it is useful to consider the system architecture required for gestural HMI. One possible architecture is presented in figure 2.4. One or more sensors monitor physical parameters from the body and convert the signals to digital data streams. These data streams are first pre-processed to extract feature vectors that are presented as inputs to a pattern classifier e.g. an artificial neural network (ANN). In this example, sensor data fusion takes place at the pre-processing stage (note: fusion could take place at subsequent stages if this were more appropriate for a particular combination of physiological signals). The extracted features are then continuously sent in a stream to the classifier which has previously been trained by example gestural patterns. These may be whole gestures or gesture segments.

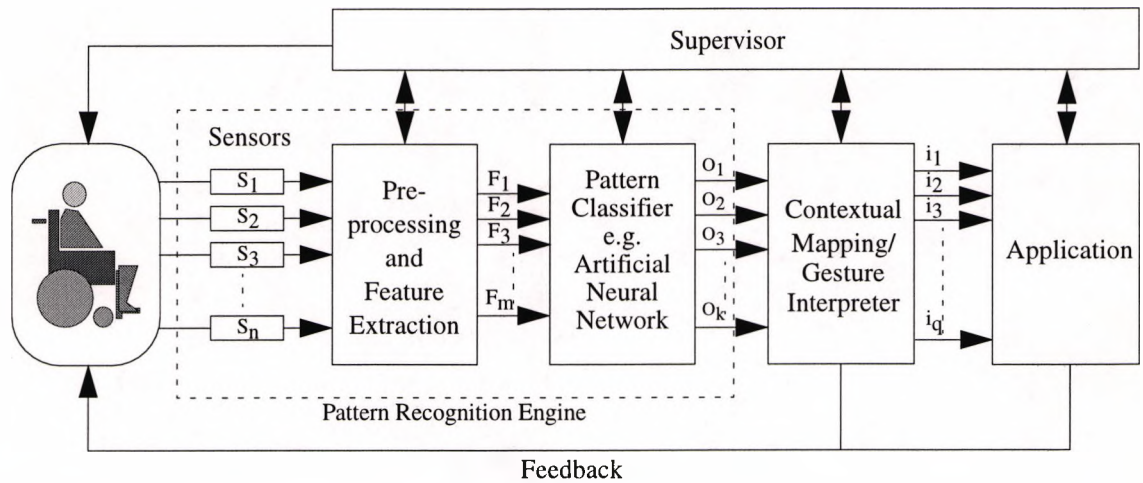


Figure 2.4 Example architecture for gesture recognition system

Each classifier output node corresponds to a possible gesture or gesture segment class. The classifier outputs form a data stream that is interpreted by a gesture parser. The output of the gesture parser provides input to the application. The application may be a communication aid with synthesized speech output, desktop computer, electric wheelchair, environmental, or robot controller. Feedback could take many forms but needs to be adequate to facilitate appropriate gestural behaviour and indicate the state of the system. The whole operation is monitored and controlled by a supervising module which will apply contextual knowledge of the state of the application to determine which inputs are expected at any one time. In this way it will be possible to maximise the recognition rate by only attempting to recognise the set of appropriate gestural actions given the state of the application. The supervisor will also control the training of the neural network and will periodically retrain the neural network using recent gestural data. In this way it should be possible to make the system adaptive so that it can automatically improve and maintain its recognition performance. The potential also exists for it to keep and report summary statistics on performance. This could be used as a monitoring and assessment tool for the user, therapist, educator, and clinician. Only the gesture recognition engine subsystem contained within the dashed rectangle is investigated in this study.

2.4.3 Automatic Gesture Recognition Study Objectives

The gestural ability pilot study (chapter 4) involved eliciting diverse expressive behaviour involving multiple parts of the body. The purpose of the study was to indicate the range of behaviour that could potentially be harnessed by technology. To create a task achievable within the resources of the research project, the problem was constrained to the investigation of automatic gesture recognition using data collected from one body site using a single magnetic tracker (chapters 5 to 8). The purpose of the automatic gesture recognition study was to show that some of the expressive behaviour elicited in the gestural ability pilot study is automatically recognisable and to explore the application of artificial neural networks to gesture recognition.

Extracting movement patterns from streams of data remains a considerable challenge. There are a number of problems that are common to many pattern recognition problems

(Tou and Gonzalez, 1974) e.g.:

Feature Extraction/Data Reduction

It is important to be able to extract the information that is needed and reject the rest, so some form of data reduction/feature extraction is generally required. Ideally a combination of features needs to be found that possesses the minimum variance between exemplars of the same class while showing good class separation in feature space. (chapter 8).

Segmentation/Pattern Spotting

Another problem is the classic pattern recognition problem of recognising patterns in time series data: segmentation - how do you determine the start and finish of each gesture? and pattern spotting - how do you spot gestural movement from a stream of gestural and non-gestural data? Research into the computer recognition of gestures (Harwin, 1991, Cairns, 1993) illustrated the difficulty in finding reliable features for gesture segmentation in the movement of people with CP. A similar difficulty is encountered in connected speech recognition (Lippman, 1989, Grayden & Scordilis, 1993).

Pattern Classification

There are many pattern classification techniques that could be applied to this problem. As the gestural repertoires and gestural forms are likely to be idiosyncratic, it is necessary for the gesture recognition algorithm to be able to learn the individual gestures of each individual. Theoretically, supervised learning ("learning with a teacher") and unsupervised learning ("learning without a teacher") are both applicable to this problem. In practice, it is much more difficult to build a machine that does the latter and is not considered further in this thesis.

This study focused on investigating the use of time-delay feedforward neural network classifiers. These were compared with k-nearest neighbour methods (chapter 7).

In the study reported in this thesis, a fixed time window approach was investigated with a view to segmentation by recognition. After low pass filtering and re-sampling, features were extracted from the data stream using a relatively small amount of data over a fixed time period. The effect of the size of the time window/number of time samples was investigated. A variety of feature vectors were compared (chapter 8).

Assuming gestures can be reliably detected and classified from the raw data stream, then the gesture sequences themselves need to be processed and interpreted and mapped to input to applications. Detailed consideration of this aspect, while a key component in any gestural HMI was not investigated.

Chapter 3

Human Factors

3.1 Introduction

Human factors issues are always important when designing human-machine interaction (Shneiderman, 1992). In the case of technological design for people with severe speech and motor impairment it is all too easy to make bad design decisions based on false assumptions (Newell, 1993). This makes it all the more vital to acquire adequate knowledge relating to the user group, their needs, and their environment.

This chapter documents the user profile, describes the nature of cerebral palsy, and considers communicative interaction for people with severe speech and motor impairment. A review of human-machine interfaces used in electronic assistive technology is presented. In particular, assistive technology for communication is described. The implications for gestural human-machine interaction design are discussed.

3.2 User Profile

3.2.1 Definition of Target Population of Primary Users

The primary user for the purposes of this study is defined as having severe speech and motor impairment due to cerebral palsy (SSMICP). This typically means:

- Speech is severely dysarthric or absent, precluding the use of automatic isolated word speech recognition (e.g. using the Dragon Dictate® system). Although familiar communication partners may be able to understand dysarthric speech relatively well, unfamiliar individuals are likely to have considerable difficulty.
- Impaired fine motor control makes activities of daily living (ADLs) difficult, or impossible without a high degree of personal assistance.

- Motor impairment precludes efficient access to traditional human-machine interfaces. In the computer domain examples include keyboards and mice. In the general domain of machines, examples include buttons, levers, dials, wheels, and knobs.
- A powered wheelchair is required for independent mobility.

3.2.2 Specifying the Secondary Users of the HMI

A child with a developmental disability is likely to follow an early intervention program. Often, a number of health care professionals are involved in this intervention program. Each of these people are potential “users” of assistive technology. For example, in the case of a voice output communication aid (VOCA), the primary user is clearly the person with a disability, but a host of secondary users can be identified including: parents, personal care assistants, friends, peers, paediatrician, neurologist, teachers, occupational therapist, physical therapist, social worker, vendor, maintenance engineer, other members of the general public.

Key secondary users can also be identified as people who require operational knowledge of the assistive device (although not necessarily the same knowledge): therapist or special educator providing VOCA training, using the device as part of language therapy, or within an educational program; clinicians conducting assessments; parents and care assistants who need to set up the system each day and be able to trouble-shoot when something stops working.

As a result, the failure of the interface to accommodate the characteristics or needs of any one of the users or the failure to integrate into the resulting social and physical environment can lead to major operational problems resulting in “technological abandonment” (Phillips, 1993).

3.2.3 Cerebral Palsy: Nature of the Condition

Definition

Cerebral - [*L. cerebrum, brain*]. *Pertaining to the cerebrum*

Palsy - [*ME palsie, from L. for paralysis*]

Few simple definitions of cerebral palsy are adequate without qualification due to the complexity of the problem (Cruikshank, 1976). Cerebral palsy can be viewed as a broad term used to describe a variety of conditions (Gersh, 1991, Levitt, 1995) where the aetiologies and underlying neural mechanisms are only partially understood.

A reasonably comprehensive definition inspired by several sources would be:

Cerebral palsy is a broad term used to describe a variety of conditions caused by damage to the developing brain, usually occurring before, during or shortly after birth. The damage is such that it affects neuromotor development resulting in a continuum of characteristic motor disorders affecting a child’s movement, speech, and posture. The condition can be mild to severe. Although it is considered to be non-progressive, i.e. the initial brain lesions or abnormalities do not get worse, the degree or type of exhibited motor dysfunction can change as a child’s nervous system develops. It is considered permanently disabling although therapeutic intervention is thought to have a beneficial effect on a child’s motor abilities.

Incidence and Prevalence

Studies in a number of countries indicate that the incidence of cerebral palsy is in the region of 2 to 2.5 per 1000 births. Prevalence in the USA is 400,000 children, 700,000 including adults in the USA. In the UK it is reported that one in 400 children have cerebral palsy (Scope, 1995). Interestingly, these figures do not seem to exhibit much variation either temporally or geographically. The incidence of CP seems to have varied little over the years even though the incidence of common causes of the past, e.g. rubella, has been drastically reduced. It is thought that this is because any potential incidence reduction has been offset by a corresponding decrease in infant mortality, particularly the mortality of premature infants.

Causes of Cerebral Palsy

The causes fall into two categories:

1. Developmental brain malformation. Failure of the brain to develop properly in the areas of the brain that are concerned with voluntary motor activity. This includes genetic disorders and faulty blood supply to the fetal brain.
2. Neurological damage to the developing brain. Often associated with premature births, difficult deliveries, neonatal medical complications, and trauma to the brain.

Typical problems that can lead to brain injury include:

1. Lack of oxygen before, during or after birth.
2. Haemorrhaging in the brain.
3. Toxic injuries, or poisoning, from alcohol or drugs used by the mother during pregnancy.
4. Head trauma resulting from a birth injury, fall, car accident, or other cause.
5. Severe jaundice, very low glucose levels, or other metabolic disorders.
6. Infections of the nervous system such as encephalitis and meningitis.

It should be noted that in around 40% of all cases of CP the aetiology has not been determined.

Diagnosis

Diagnosis of CP is seldom instant or straight forward. CP is a developmental disability, and as such, the developmental indicators may not emerge until a considerable time after birth or the time of trauma. Often, a body of evidence gradually grows until finally there is little doubt.

Initially there are risk factors, events that occur or are observed during pregnancy, during birth, or neonatally that indicate that a child has increased risk of CP (and often a host of other conditions). Then, as the child develops, other indicators emerge, developmental milestones are delayed, unusual motor patterns develop, and/or primitive reflexes persist. Early diagnosis can be hard due to the high variance in the patterns of normal development at this age.

Once CP has been diagnosed, appropriate early intervention services can be prescribed. This often does not happen until the later part of the first or even the second year. At this stage, the severity of the various aspects of the condition are still hard to predict. The way the condition manifests itself and the associated problems are likely to change as the child develops. One type of CP may be evident at an early age, then some years later the motor dysfunction may appear to fall into a different category. However, sometimes the initial label sticks, resulting in a need to be cautious when extracting data from medical,

therapeutic, and educational records (Cauley et al., 1989).

Classification of CP

Understanding the different manifestations of CP is important from an HMI perspective as the movement disorders are quite distinct in nature. Although the movement disorder classifications are distinct it should be noted that individual cases are classified in terms of the predominant condition exhibited at the diagnosis. There is increasing recognition that in many cases multiple categories exist to a greater or lesser extent. One type may be predominant in certain parts of the body while other body parts exhibit different characteristics.

There are a number of classification systems that vary mainly in the number of sub-categories (Phelps, 1950, Cruikshank, 1976, Hardy, 1983, Levitt, 1995). There is some disagreement as to whether certain sub-categories are really one of the same class. This highlights the difficulty that exists in attempting precise classification.

Classification usually consists of two components a) muscle and movement behaviour correlated with site of brain lesion and b) location and extent of dysfunction.

The following categorisations have been extracted from McDonald (1987) in Beukelman and Mirenda (1992: 241) and Gersh (1991).

Predominant movement dysfunction/site of brain lesion:

Spastic (50%)

Hypertonic muscle tone, exaggerated stretch reflex, myoclonus, Babinski reflex present after six months, tendency to develop contractures, persistent primitive reflex, lack of inhibition of antagonist muscles. It is attributed to damage to the pyramidal system.

Athetoid¹ (10%)

Uncontrollable and involuntary movement. It is attributed to damage to the extrapyramidal system, in particular the basal ganglia (or more accurately, basal nuclei).

Ataxic (5%)

Difficulty in maintaining balance, clumsy or uncoordinated voluntary movement. It is attributed to damage to the cerebellum.

Mixed (25%)

Combinations of two or more of the basic types.

Rarer types include:

Tremor (<1%)

Repetitive involuntary actions.

Atonia (1%)

No or reduced muscle tone.

¹ Many references to athetosis will describe "slow writhing purposeless movements". In relation to the numerous individuals that the author has observed, this is an inadequate description steeped in historic attitudes. For instance, how do you determine that a movement is "purposeless"? Athetoid movement often results from attempts to communicate, to interact with the environment and to express emotion. In this respect, the movement may well be distinctly expressive.

Location and extent:**Monoplegia**

Only one limb on one side of the body is affected. Movement impairment is usually mild and often disappears over time. It is very rare.

Diplegia

The lower extremities are mainly affected causing the individual to stand on their toes and to scissor their legs due to muscle spasticity. There may be mild movement dysfunction of the upper body, but there is adequate control over the trunk, arms, and head for most daily activities.

Hemiplegia

One side of the body is affected. The arm is often more affected than the legs, trunk, or face. Typically the arm is held in flexion.

Quadriplegia

The whole body is affected. Usually the lower extremities are most affected. The extent of the motor impairment is likely to affect breath control and orofacial muscles used in feeding and speaking.

Double Hemiplegia

Similar to quadriplegia, in that the whole body is severely affected. The difference is that the arms are more affected than the lower extremities.

Associated Disorders:

Associated disorders are common in people with CP. These include:

- Cognitive impairments, learning difficulties (60%-70%)
- Mental Retardation: 25% (3% general population)
- Vision problems (40%), e.g. Strabismus (crossed eyes), refractive errors (near or far sightedness), amblyopia (lazy-eye), cortical blindness.
- Attention Deficit Hyperactivity Disorder (20%)
- Hearing impairments (20%)
- Seizure activity (35%-45%)
- Speech Impairment - oral-motor movements of jaw, lips, tongue and facial muscles used for speaking. Trunk muscle control affecting breath control.
- Dyspraxia.
- Sensory impairments e.g. agnosia, impaired proprioception or vestibular system.
- Tactile hypersensitivity (tactile defensiveness).
- Tactile hyposensitivity.

Changing nature of CP

Improved medical provision is thought to be responsible for a change in the relative incidence of the types of CP. In the Northern Hemisphere, incidence of pure athetoid CP has decreased and incidence of mixed types of CP has increased due to advancements in medical care. Also, there has been an increase in the number of children with multiple disorders including CP (Hagberg et al., 1975). It is worth noting however, that there is probably also an increased propensity on the part of clinicians to recognize and to classify cases that fall into multiple categories (Hardy, 1983).

3.2.4 Communicative Interaction of People with SSMICP

Severe speech and motor impairment due to cerebral palsy is likely to have a profound effect on the development of expressive communication and greatly limit the ability to interact and explore the environment. This often affects receptive and expressive language acquisition (Cauley et al., 1989, van Balkom, 1991, Heim, 1994). It is common for people with SSMICP to initiate communicative acts infrequently, often limiting themselves to answering or responding to the requirements of their speaking communication partners (Yoder, 1984, Kraat, 1985, Light, 1988, van Balkom, 1991). Voice output communication aids (VOCAs) are often used as a last resort, giving way to other methods such as eye gaze, gestures, facial expression and/or vocalisation (Murphy et al., 1995). A comprehensive understanding of the reasons for this phenomena has yet to be ascertained. Plausible contributing factors may include that the users acquire "learned helplessness" or "learned dependency" where a passive outlook is encouraged by their environment (Basil 1992), interaction styles of communication partners (Rowland, 1987, van Balkom & Heim, 1990), issues of technology abandonment particularly due to the mismatch between user requirements and available technology (Phillips, 1993, Scherer, 1993).

3.3 Human-Machine Interfaces used by People with SSMICP

3.3.1 Electronic Assistive Technology

People with SSMICP find it difficult or impossible to access regular HMIs due to their limited motor control. This has inspired many ingenious adaptations or assistive technology specifically designed for this population (Vanderheiden, 1978).

One of the simplest adaptations is the keyguard which is a plate that fits over a computer keyboard, with holes drilled over each key. This prevents more than one key from being pressed simultaneously and also allows the hand to rest on the guard while the user focuses on fine motor control.

If the user's motor control is insufficient to use a conventional keyboard in this manner, extended keyboards are available. These are keyboards with larger than normal sized keys that replace the conventional keyboard. A number of designs use membrane switches that allow the area corresponding to each key to be tailored to the needs and ability of the individual user (figure 2.5).

If motor control is insufficient to reliably target such keyboards (often referred to as "direct selection"), then attempts are made to find parts of the body that can reliably target single switches. A variety of designs of switches are available for this purpose (figure 2.6).

If the applications require anything more than a simple on or off control, some form of "indirect selection" strategy (e.g. scanning, coding) is employed to map the presses of a few switches to a much larger number of actions. A common method is to arrange each item available for selection in a matrix on a rectangular board, each with a corresponding light-emitting diode (LED). All the LEDs in successive columns are illuminated in sequence. The user presses a switch when she/he sees that the column containing the desired item is illuminated. The system, knowing which column has been selected, then

proceeds to sequentially illuminate the individual LEDs in the selected column. The user presses the switch when the desired item is illuminated. The system then returns to scanning columns and is ready for the next selection. In this way, one switch can be used to access a large number of items (figure 2.7). The limiting factors are the increased scanning time for each item and the greater cognitive load as the number of items grows. Typically the maximum rate of selection achievable using this type of selection method is between five and ten words per minute (Foulds, 1985). There are variations on this scheme e.g. initially scanning rows rather than columns, step scanning for two or more switches and directed scanning for five switches. An extensive review of AAC techniques and technology can be found in Beukelman & Mirenda (1992).

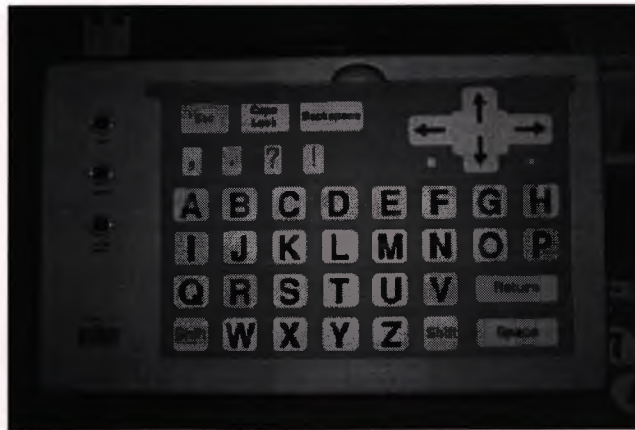


Figure 2.5 Intellikeys® expanded membrane keyboard.

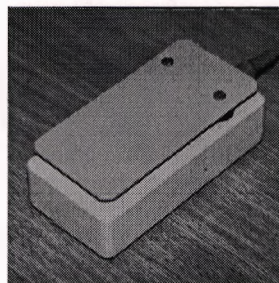


Figure 2.6 Two examples of proprietary switches used to harness movement from people with SSMICP.

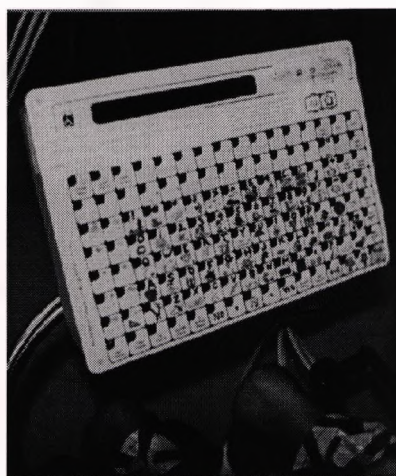


Figure 2.7 An example of a VOCA using indirect selection using single switch scanning of up to 128 items.

3.3.2 Assistive Technology for Communication

Assistive technology for augmentative and alternative communication (AAC) is one of the most challenging areas for a gestural human-machine interface. Electronic communication aids, particularly the voice output communication aid (VOCA) which outputs synthetic speech have made a substantial difference to the lives of many people with SSMICP. Electronic communication aids have offered the opportunity to express thoughts with much greater precision and depth (Nolan, 1981, Murphy, 1994). However, as mentioned previously, many people with SSMICP tend to use their electronic aids to initiate a communicative act infrequently. In addition, the use of such technology does not integrate easily with natural expressive modalities. In the case of indirect selection using scanning, the user has to constantly look at the scanning matrix, disrupting sustained eye contact with any communication partner.

There is still considerable debate as to the relative effectiveness of various AAC options: a summary is given in Beukelman and Mirenda (1992:66-67) taken from a number of sources. Typically, rates for non-speakers using existing AAC options are 2-10 words per minute compared with that of unimpeded speech that proceeds at 150-200 words per minute, a significant order of magnitude of difference (Alm et al. 1992). Many schemes of varying sophistication have been developed to enhance the communication rate. A selection of examples are illustrated.

AAC device developers and researchers have attempted to associate more meaning to each bit of information transferred from user to machine. Scanned word lists, word prediction, coding schemes, semantic compaction techniques have all shared this goal. Semantic mapping schemes such as Minspeak® attempt to associate short sequences of a small number of familiar symbols to a much larger number of words and phrases (Baker, 1982). The cognitive association is based largely on mnemonic principles. This can also increase the cognitive load of the user unless the necessary semantic associations can be easily recalled from memory. Although potentially this gives efficient access to a large vocabulary (Baker & Nyberg, 1990), many users only manage to memorise a small number of sequences (Levelt, 1994). Strategies such as these may increase the rate of communication and give access to more sophisticated linguistic constructs, however it

can also make the system hard to learn if the user has to learn mappings that are arbitrary and abstract. An example of this is the early BlissApple communication software which required a four digit number to be entered for each of the 2000 or so Bliss symbols.

Researchers at the University of Dundee MicroCentre have taken a very different approach. They have developed AAC software which capitalises on the fact that much of human discourse is repetitive and stereotyped in nature (Alm et al., 1992). Through the application of conversational analysis and dialogue design, a model of conversational patterns was developed to predict conversational moves. The system was implemented using augmented transition networks. The aim is to provide the user with "conversational momentum, that is, a small effort can initiate relatively long conversational moves". Features of the system have been commercially incorporated into Talk:About (Don Johnson, 1995).

3.4 Summary and Implications for Gestural HMI Design

A number of key user issues are summarised below (*italics*) followed by suggestions as to the challenges they present for the design of gestural HMI together with possible accommodations:

Cerebral palsy covers a range of movement disorders with different characteristics: This is likely to affect the type of pre-processing needed to enhance the signal-to-noise ratio of the "noisy" gestural movement. Ideally the system should be optimised for each individual.

Each primary user is likely to exhibit a highly individual profile of abilities: The GHMI system will need to be modular. Just as people with SSMICP undergo assessments to identify sites on the body that can reliably operate a switch (Beukelman & Mirenda, 1992), a similar process could be employed to identify parts of the body that are involved in gesture production. Sensors could initially be attached to many body sites, then the system itself may be able to determine the minimum sensor configuration for robust gesture recognition. In order to maximise the size of each user's gestural repertoire, individual repertoires are likely to be unique. The system should be trainable to recognise each user's particular gestural repertoire.

Movement characteristic and abilities can change substantially as a child's neurology develops: It should be possible to accommodate for this by periodically retraining the system using recent gestural data.

User's physical performance is likely to exhibit significant fluctuations on a daily or hourly basis due to factors such as emotional state, fatigue, illness: This is potentially a problem for HMI based on recognition of movement patterns. Conceivably, it may be possible to accommodate this effect by anticipating changes in performance (e.g. changes in range of motion, duration) and adapt the system accordingly.

People with SSMICP communicate multimodally: This implies that the maximum interface bandwidth will most likely be achieved by harnessing multiple modalities possibly involving multiple parts of the body e.g. gesture from arms, hands, torso, legs, head, eye-gaze, and vocalisations.

Potential users' ages range from infants to adults: The interface will need to be adaptable and expandable as the user's abilities, needs, and interests change and develop. Gestural

HMI incorporated into AAC technology offers a number of advantages. Unimpaired children are typically producing deictic gestures by 12 months (Bates et al., 1975) and begin to produce iconic gestures from 12 months onwards (Acredolo & Goodwyn, (in press) reported in McNeill, 1992). On this basis, gestural HMI would appear to be appropriate for use by infants with cerebral palsy as part of an early intervention strategy at an age when complex single-switch scanning technology would be too cognitively demanding.

CP can be accompanied by cognitive impairment ranging from mild to severe: Gesture is currently used as a method of AAC and as an aid to language development with cognitively impaired populations (Musselwhite & St. Louis, 1982, Fuller & Wright, 1994, Grove & Dockrell, 1994). Thus, gestural HMI could be integrated into existing therapeutic language intervention strategies involving gesture.

People with SSMICP often exhibit residual infantile reflexes e.g. asymmetric tonic neck reflex, startle reflex: The system must be designed so that it does not confuse a reflex with gestural input. Fortunately, reflexive movement tends to be very stereotypical. In principle it should be possible to train the system to reliably recognise this type of movement and thereby minimise any spurious input.

The user is likely to use a wheelchair for independent mobility: If the system is to be mobile it would most likely be mounted on the wheelchair. It would have to be robust enough to withstand daily use in a variety of environments. The wheelchair seating and postural restraints would have to allow for appropriate gestural movement, while still offering the user adequate support.

Need to design for secondary users as well as primary user: The system will have to be simple and quick to set-up and maintain requiring little technical expertise. Although its internal operation will necessarily be complex, the user's mental model of the system should be as intuitive and transparent as possible.

Chapter 4

Gestural Ability Pilot Study

4.1 Introduction

In order to develop a gestural human machine interface for people with severe speech and motor impairment it is necessary to investigate the gestural ability of this population. Rather than attempt to document gestural ability in randomly selected daily activities, this pilot study aimed to elicit gestural behaviour above and beyond any commonly used gestural repertoire.

A cognitive framework was constructed using performance arts techniques to elicit a wide variety of volitional expressive behaviour that was potentially useful for human-machine interaction.

The gestural ability of twelve children and young adults between the ages of 5 and 18 were studied. The gesture sessions were video-taped and the video material was reviewed to determine which body parts were involved in the gesture. The summary results for each subject are presented in this chapter. The transcriptions of the elicited gestural repertoires for each subject are presented in appendix A.

4.2 Subject Selection

Subjects were selected with the assistance of therapists, special educators and clinicians from John G. Leach School, Newcastle, Delaware, HMS School for Children with Cerebral Palsy, Philadelphia, Pennsylvania, Widener Memorial School, Philadelphia, and the A. I. duPont Institute Children's Hospital. Meetings were held to describe the research and the subject selection criteria, after which participants were requested to identify potential subjects for the study. The criteria were:

- 1) Need: Subjects should be severely motor and communicatively impaired due to cere-

bral palsy or cerebral palsy-like symptoms. They should have difficulty targeting switches commonly used in human-machine interaction.

2) Cognitive Ability: Subjects should have sufficient receptive language and demonstrated cognitive ability to interact in the proposed sessions.

4.3 Subject Profile

Eleven of the twelve chosen subjects met the above criteria. In the case of subject S8, cognitive abilities had only been informally assessed due to her age (5yrs 9m) and her difficulty with expressive communication. She was included in the study upon recommendation of her therapist who thought that it may be possible to elicit gestural behaviour.

Subject details are listed in table 4.1. All were considered quadriplegic, six spastic, five athetoid and one spastic-athetoid. All had cerebral palsy except subject S4 who had CP-like symptoms due to traumatic brain injury. Ages ranged from five years nine months to eighteen years one month. The group comprised six females and six males. Cognitive age indicated by the Peabody picture vocabulary test (PPVT-R form L, described in McLaughlin & Lewis, 1986) ranged from 3 years six months to 11 years (excluding subject S8). At least eight had persistent asymmetric and/or symmetric tonic neck reflex and at least six had diagnosed vision or visual tracking problems. Where noted in their records, the quality of volitional motor ability is detailed.

Table 4.2 details the twelve subjects' present methods of expressive communication. Both electronic and non-electronic AAC systems are listed including the input method for each. The primary method of communication has been noted when it has appeared on the speech therapy report. Finally any other relevant details have been included.

It should be noted that eleven of the twelve subjects were using or training to use electronic assistive technology in the form of a VOCA or computer system. Six used eye gaze as the selection method with their non-electronic AAC system. Ten subjects used "indirect selection" and one subject used "direct selection".

For those that used an electric wheelchair, the method used to access powered mobility is documented in table 4.3. Four subjects were using powered mobility and five subjects were being assessed for their ability to access powered mobility. Of those using an electric wheelchair, two used a four position joystick and two used from three to five distributed switches to control their wheelchair.

Subject	Diagnosis	Age year, month	Sex	Cognitive level ^a	ATNR/STNR ^b	Vision/Visual tracking problem	Reported quality of volitional movement (extracted from therapists and clinician reports)
S1	Spastic quadriplegia (CP)	13,8	F	7,6 (13,2)	Y	Y	"severe rigid spasticity, very limited active movement"
S2	Spastic quadriplegia (CP)		M	7,9 (14,6)			
S3	Spastic quadriplegia (CP)	12,0	F	3,6 (12,5)	Y	Y	"active range of motion very limited in both upper extremities"
S4	Spastic quadriparesis and hydrocephalus (TBI)	9,4	M	3,11 (8,4)	Y	Y	"slow active movements left-side for functional movements"
S5	Athetoid quadriplegia (CP)	17,9	M	10,5 (16,10)	Y		
S6	Athetoid quadriplegia (CP)	9,9	F		Y		"constant athetoid type movement patterns upon any effort or excitement" "preference for use of right upper extremity"
S7	Athetoid quadriplegia (CP)	10,7	F	3,7 (9,8)	Y	Y	"isolated finger movement", "active range of motion within functional limits although often exhibits excessive end of range of motion"
S8	Athetoid quadriplegia (CP)	5,9	F		Y		"imitation of orofacial expressions" "volition movement of both upper and lower extremities but has poor control"
S9	Spastic-Athetoid quadriplegia (CP)	16,9	M	11,0 (17,4)		Y	"active range of motion limited to flailing type movements of upper and lower extremities" "marked fluctuations in muscle tone with choreoathetotic movements" "unable to functionally use hands secondary to athetoid movements"
S10	Spastic quadriplegia (CP)	18,1	F	7,7(16,6)	Y	Y	"active range of motion limited to involuntary flailing of extremities with high muscle tone"
S11	Athetoid quadriplegia (CP)	10,10	M	6,4 (10,1)			"Active movements: very large poorly graded athetoid motions, usually extensor patterns". "Head control: limited range", "right hand dominant" "all fine motor movements require extreme effort and time"
S12	Spastic quadriplegia (CP)	12,11	M	8,10 (12,9)			"athetoid quality present in all his movements" "left hand dominant"

Table 4.1: Details of subjects chosen for gestural ability pilot study

a. As indicated by Peabody picture vocabulary test (PPVT-R form L) year, month. Age at testing indicated in parenthesis.

b. Asymmetric tonic neck reflex (ATNR) / Symmetric tonic neck reflex. These are both involuntary primitive reflexes.

Subject	Electronic Input method	Selection Strategy	Electronic AAC system	Non-Electronic	Non-Electronic Selection Method	Primary Method of Communication	Other Relevant Details
S1	Pad switch with right elbow	Linear step scanning and auditory scanning	VOCA (Dynavox)	Rebus picture/word language board ~30 items	Right index finger or knuckle	Yes/No head-shake, vocalizations, some word approximation Speech is supplemented using Rebus board	"takes 5 seconds to move arm to picture symbols"
S2	Four direction foot-joystick	Directed scanning	VOCA (Tailor made)	Rebus picture/word language board 250 items	Eye gaze localized to numbers and colours	yes/no eye gaze and head shake	
S3	Single right side-mounted head switch	Scanning	Undergoing assessment for VOCA - Uses computer with keyboard emulator			Single word approximated speech	
S4	Left index finger	Direct selection	VOCA Touch Talker	Picture board - 88 items (8x11)	Left index finder	Combines modalities	
S5	Single side mounted head switch. Being evaluated for multiple switches	Row-column scanning 128 location	VOCA Light Talker	English orthographic board with adapted Fitzgerald key.	Eye gaze to large number on board perimeter	Combines modalities	"constant poorly graded gross movements characteristic of athetosis are hard on equipment"
S6	Single switch operated with the hand mounted vertically at distal right edge of lap-tray	Row-column scanning	VOCA (Liberator)	Picture board built into lap-tray	Targeting of colour coded numbers on rail around edge of lap-tray	Multiple modalities	"accesses computer, battery operated toys and environmental control unit via switch.
S7	Pad switch with left hand	Linear Scanning 32 locations	VOCA (Light Talker) (Under evaluation)	Language board with keyguard-like grid. 40 items	direct pointing using left index finger		"direct selection using hand-held optical indicator was problematic due to increased athetoid movements as selections were attempted"
S8			None (under assessment)	Object selection E-tran	Eye gaze		
S9	Right knee-switch	Row-column scanning	VOCA (Light Talker)	Past use of E-tran	Eye gaze	Speech, Light Talker	
S10	Head switch	Row-column scanning 128 locations	VOCA (Light Talker)	Past use of coded eye-pointing system	Eye gaze	Vocalization, limited facial and eye pointing	
S11	Single left side mounted head switch	Row-column scanning 128 locations	VOCA (Light Talker)	Past use of E-tran - 100 number/colour coded items	Eye gaze		
S12	Knee- switch	Row-column scanning	Light Talker (VOCA)			Speech and Light Talker	

Table 4.2: Existing Methods of Expressive Communication

Subject	Mobility	Mobility Access Method
S1	Manual wheelchair, being evaluated for powered mobility	problem finding three reliable switch sites
S2	Powered wheelchair	Four direction foot joystick with left foot
S3	Manual wheelchair, being evaluated for powered mobility	Looking for four switch sites (not achieved)
S4	Manual wheelchair, being evaluated for powered mobility	Aiming for joystick control
S5	Powered wheelchair	Three head-switches, 1 knee-switch, 1 elbow switch
S6	Manual wheelchair Being evaluated for powered mobility	Vertically mounted switches at edge of tray (just under full elbow extension)
S7	Manual wheelchair evaluated for powered mobility	Three pad switches with left upper extremity
S8	Manual wheelchair and walker	
S9	Powered wheelchair	Three head switch 1 knee switch, 1 elbow switch
S10	Manual wheelchair	
S11	Manual wheelchair	
S12	Powered wheelchair	Left hand operated four position joystick

Table 4.3: Mobility and Powered Mobility Access Method

4.4 Experimental Design

A set of approximately 140 concepts represented by words and phrases was created based on notions that the investigators could easily express non-verbally themselves. A set of flash cards was created with one concept written on each card. The cards were sorted into the categories listed in table 4.4. The individual words and phrases are listed together with the analysis of the response in appendix A.

Considerable effort was put into creating an environment where each subject could feel at ease while performing gestures. With this in mind, all sessions took place in familiar surroundings with familiar people involved. Gestural elicitation sessions were scheduled to take place within each subject's regular school therapy session. Each subject's therapist was invited to participate. At the beginning of the session the facilitator took time to explain the nature and purpose of the session. In addition, the facilitator and investigator had previously met briefly with each subject and their therapist to explain the project. The facilitator ensured that she was familiar with each subject's yes/no response.

Categories	
Actions	Musical instruments
Animals	Objects
Communication	Outlines
Description	People
Events	Senses
Fantasy characters	Sport
Feelings	Travel
Food	Weather
Movement	Miscellaneous

Table 4.4: Concept Categories used to Elicit Gestures

Each subject was told that they would be video-taped so that their gestural responses could be studied and that the video would remain part of their school record. She/he was offered the opportunity to decline to be video recorded. The subject was positioned so that she/he had sufficient space to gesture with the arms without feeling they were likely to hit anything. Where appropriate, lap-trays, VOCAs, and arm-rests were removed, restraints loosened with the agreement of the subject and therapist to enable the subject to move more freely. The therapist was invited to sit next to the subject (but not close enough to constrain the gestures). The facilitator sat opposite the subject to enable good eye contact to be maintained.

The concept of mime with some examples was presented to each subject.

The main part of the session comprised a charade-like game where the therapist was a member of the subject's team. The facilitator would select a flash card and read aloud the word or phrase. Time was then allowed for the subject to produce a "mime" to express the concept. It was explained to each subject that they were in control in that they did not have to produce a mime if they did not wish to and that they could take a rest or stop at any time. No constraint or direction was placed on the type of response required, although it was explained to the subject that they could ask for a clue if they so wished. Similarly, the facilitator took care not to convey any judgement as to the nature of the gestural response. Once the facilitator felt the subject had produced a response, the next concept was introduced. The facilitator carefully managed the interaction so that motivation was maintained using banter and changing intonation of voice. From time to time the subject was asked whether they were happy to continue. Each gesture session lasted around 40 minutes or until the subject asked to stop.

Each video was subsequently reviewed and the gestural responses analysed. The video was reviewed by one investigator. The purpose of the analysis was to ascertain which body-parts were involved in the gestural responses. As each gesture was only produced once it was not possible to make any conclusions as to the repeatability of each gesture. Inter-observer reliability testing was not applied in this pilot study. Each gesture was logged, identified by the eliciting word/phrase, recording the part or parts of the body

observed to be involved in the gestural response. The body-part categories were devised specifically for the study. They were: vocalisation, head, facial expression, eye-gaze, mouth, tongue, trunk, left-arm, right-arm, left-wrist, right-wrist, left-hand, right-hand, legs, feet, whole body. The results for each subject are presented in appendix A.

Histograms of the frequency of involvement of each body part for each subject are presented in the second part of the following results section.

4.5 Results

4.5.1 Previously Documented Gestural Ability

The subjects' unaided expressive communicative ability previously documented in therapists' reports and medical records are listed in table 4.5. All subjects had severely dysarthric or no speech. At least six subjects used non-speech vocalisation as a means of expression. At least six subjects used orofacial expression. At least five used deictic arm and/or hand gestures. Other gestures included head-shaking, the "OK" hand sign, hand up for yes, hand down for no. Except for one subject (S4), subject records only contained details about two types of gesture: gestures symbolizing yes and no and deictic gestures e.g. the use of eye-gaze, or arm/hand pointing to body parts or objects or people in the environment, the meaning of which presumably has to be guessed at from context. Seven gestures were recorded for S4. These are listed in table 4.5. One can only speculate as to how accurately any documented gestural repertoire reflected each subject's actual use of gesture. During the investigators' contact with the subjects in their school settings, a number of the subjects used a gestural method of indicating needs such as wishing to go to the bathroom. This was not documented in their reports.

Subject	Speech ^a	Non word Vocalization	Eye Gaze	Orofacial Expression	Deictic Arm/hand Gestures	Other	Reported Comments
S1	SD	Y			Y		
S2	N		Up - yes Down - no	Y		Head shake	
S3	SD						
S4	SD	Y			Y, 5 "I", point to eye; "love", hand to heart; "you", points to you; "hand to mouth", eat/drink; "hand to ear", hear/listen	2 "big", raises hand at side of head; "mad (angry)", holds fist up	
S5	SD		up - yes no - look down deictic				
S6 ^b	N				Y	Hand up - yes Down by side for no	
S7	N	Y		Y	Y, hand	Sign approximations "thumb and index finger together for yes (OK sign)"	"strong ATNR restricts reliable use of eye gaze, switch scanning, deictic hand gestures"
S8	N	Y	Y	Y			"imitation of orofacial expressions, emergence of cognitive verbal precursors"
S9	SD		Y	Y			
S10	N	Y	up - yes side to side - no	Y - limited			"limited volitional movements"
S11	N	Y	up - yes down - no	Y	"gross pointing with arm and hand"		
S12	SD						"gestures to the best of ability"

Table 4.5: Unaided Modalities of Expression

a. SD = severely dysarthric, N = no documentation of any speech ability in reports.

b. Speech therapy report not available for this subject.

4.5.2 Elicited Gestural Ability

Ten subjects participated in this part of the study. Two subjects, S8 and S10, did not participate. In the case of S8, the protocol would have had to be substantially modified to accommodate her level of receptive language. S10 chose not continue her participation in the study.

From 30 to 141 gestural responses were produced from the remaining ten subjects. The elicitation protocol allowed for subjects to ask for clues. Most of the time the gestural responses were produced readily and spontaneously from the verbal stimuli.

Gestures were articulated from different and multiple body sites depending on the mime. For example, the static mime for umbrella involved holding the hand stationary above the head level, the mime for monster often involved the whole body movement and posture including animated facial expressions and vocalisation and the dynamic mime for violin involved the co-ordination of head, torso, and arm. Minimal prompting in the form of clues was necessary from either the therapist or facilitator. The ease of elicitation and consistency of concept over time suggests that existing kinaesthetic abilities were being harnessed involving a low cognitive load. Examples of gestures from subjects from

different schools in different States often appeared to originate from common concepts. The mime for rainbow typically involved one hand moving in an arc above and across torso or head. The mime for snake was performed by rapid protrusion of the tongue. Mimes were spontaneously enacted often with a sophisticated and creative appreciation of movement in time and space. The subjects were able to convey concepts for weight, emotion, character formation and object visualisation.

Appendix A contains transcripts of session video-tapes indicating the body-parts involved in the gestural responses from ten subjects. Histograms indicating frequency of body-part involvement extracted from these transcripts are presented in figures 4.1 and 4.2. Inspection of the transcripts reveals that, more often than not, multiple parts of the body are involved in gesture production. The histograms show clearly the relative involvement of body parts. The arm is involved more frequently than any other body-part for nine out of ten subjects. All except two subjects exhibited a significant preference for one arm over the other.

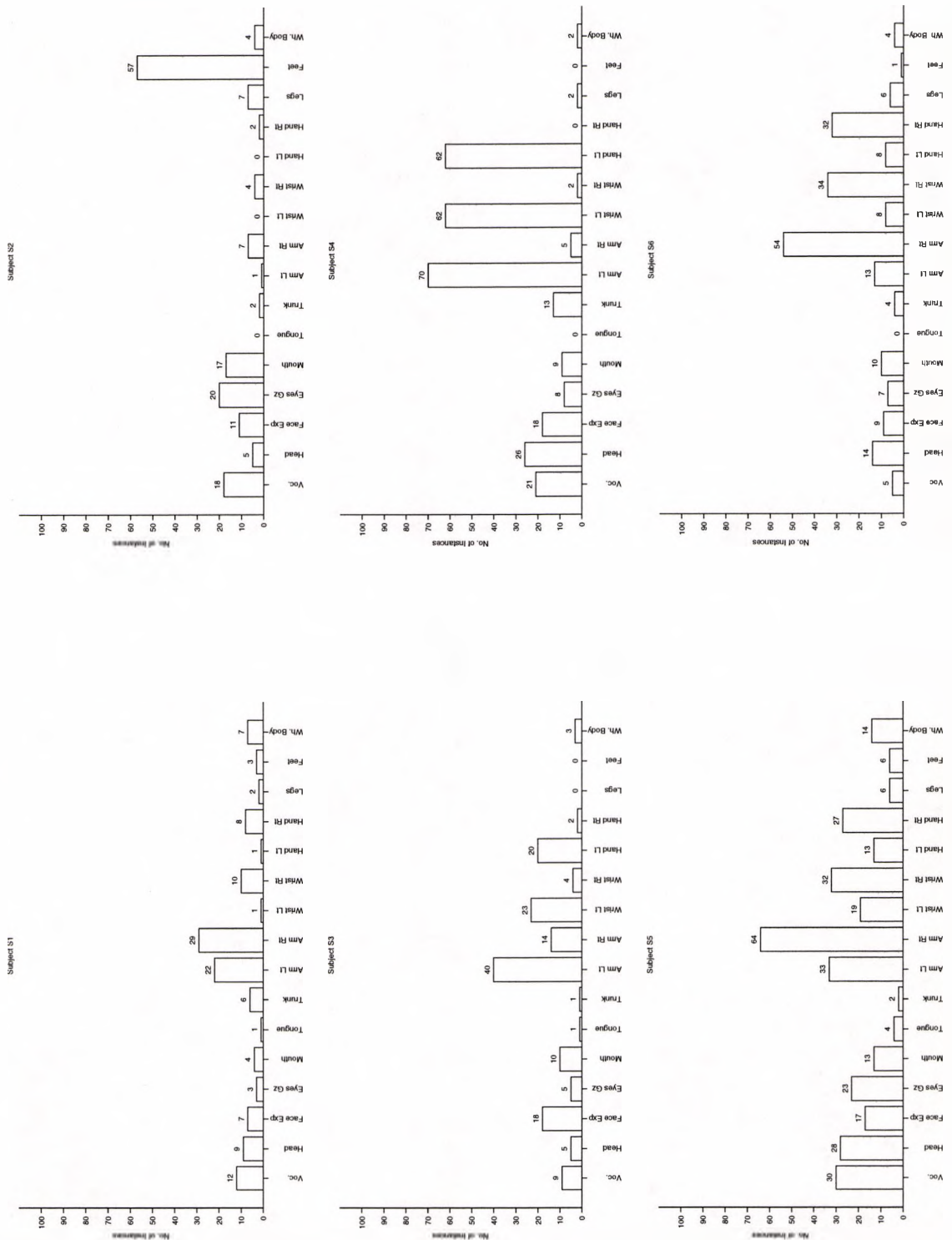


Figure 4.1 Histograms indicating number of instances of use of body parts involved in gestural repertoires for subjects S1 to S6.

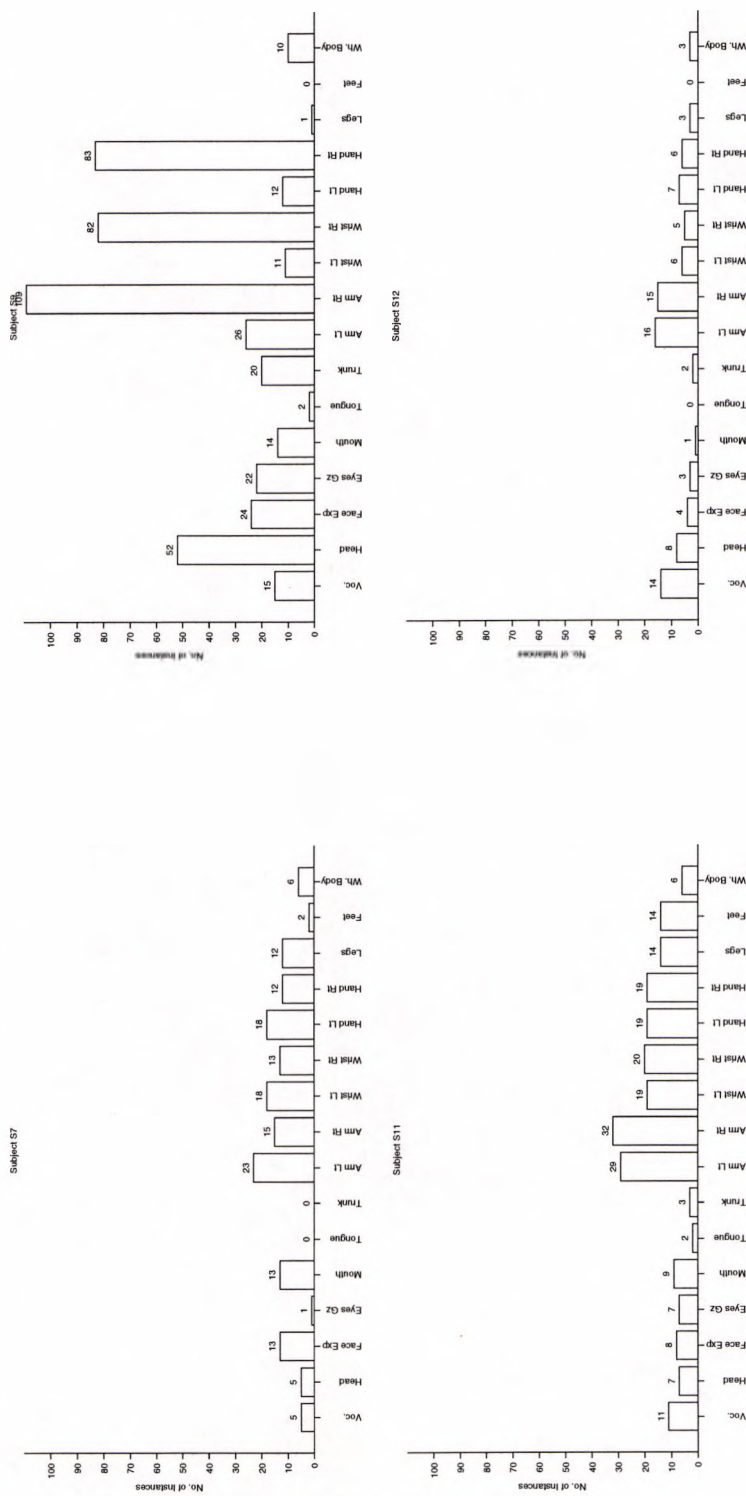


Figure 4.2 Histograms indicating number of instances of use of body parts involved in gestural repertoires for subjects S7, S9, S11, S12.

4.6 Discussion

The gestural ability pilot study results indicate that people with little ability to interact physically with the environment nevertheless can translate their knowledge of the world into appropriate physiographic gestures. Since access to play or opportunity to interact with the environment is severely restricted for this group, it is proposed that the observed gestural ability may have been acquired without practice. The level of gestural ability exhibited by subjects in this study was not anticipated by participating therapists, clinicians, teachers and parents. There was no evidence of appreciable gestural ability based on analysis of prior documentation in educational, therapeutic, and medical records.

As far as familiar communication partners were aware, the subject had never been exposed to similar activities before. The apparent consistency of the underlying form of production is important as it is much easier to turn the gestures into a method of AAC if there is consistency across subjects. This, and the ease with which gestures were elicited suggests that there is a high degree of transparency in the gestural production. In other words, little effort is needed to learn and remember the gestures. In an HMI system, the concepts used to elicit the gestures could easily be mapped to input commands of an application. One of the simplest schemes is to use the conceptual nature of gesture in a similar way to that of graphic icons used in graphical HCIs. The use of gesture offers an additional advantage over the use of graphic icons in that gesture draws upon kinaesthetic recall.

The histograms show that a variety of body parts are involved in production of the gestures. There was considerable variation between subjects as regards which body-parts are used to articulate the gestures. More often than not, multiple parts of the body were involved in gesture production. However the arm was the most commonly used articulation for nine out of ten subjects.

In addition to documenting gestural ability, the purpose of this part of the study was to determine which parts of the body to instrument in order to capture gestural movement that is likely to be recognisable by computer.

From the results presented in this chapter, the most promising single site of the body to instrument would appear to be the subject's dominant arm.

Chapter 5

Gestural Data Collection for Pattern Analysis

5.1 Introduction

The utility of the gestures elicited in the gestural ability pilot study for HMI depends on whether they can be produced consistently and whether they can be reliably recognized by computer. The research reported in chapters 5 to 8 address these issues by looking at a sub-set of gestures that could be transduced using a single magnetic tracker attached to the body. The arm was chosen as it was involved in gestural expression more than any other part of the body.

This chapter details subject selection, the experimental rig, and the methodology for collection of dynamic arm gesture data. Details of the movement tracking are presented. Results relating to the nature of the gestures and their rate of production are presented and discussed.

5.2 Subject Selection

Five subjects were chosen for instrumented data collection. Subjects were selected based on their ability to produce a range of gestures involving the arm. Manual segmentation of the gestural data stream took a considerable amount of time (Chapter 6). As a result, data from just one subject was used in this part of the study.

Subject S9 was chosen from the group of participants in the gestural ability pilot study. The subject had a relatively good range of motion using the upper extremities and had a wide and imaginative range of gestural responses. However, he also had the highest degree of choreoathetosis in the movement involved in executing those gestures.

Although there was plenty of movement to monitor, the level of noise in the movement was high. This subject's diagnosis of mixed spastic-athetoid CP represents a category of cerebral palsy that is increasing in prevalence.

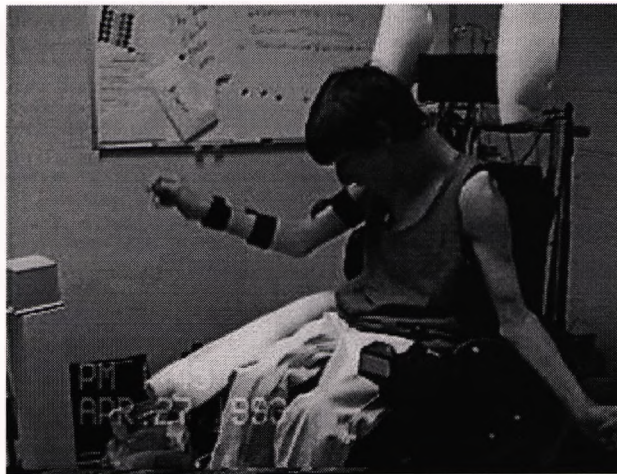


Figure 5.1 Subject performing dynamic arm gesture with magnetic tracker receiver attached distally to the right forearm. The transmitter was mounted on a wooden post in-front and to the side of the subject's wheelchair

5.3 Subject Description

Subject S9's age at the initial data collection was 16 years 9 months. His cerebral palsy was classified as spastic-athetoid quadriplegic.

The motor impairment presented itself as spasticity, particularly in the lower extremities with contractures in both lower extremities. His upper extremities, head, neck and face exhibited athetoid movement.

The subject was highly motivated to participate in the study and showed interest in the research.

His speech was severely dysarthric and usually limited to one or two words. He had been using electronic communication aids for around 11 years. Prior to this, he used a non-electronic eye-gaze "E-tran" system.

The following details were extracted from speech/occupational/physical therapy reports:

Gross Motor control:

Passive range of motion: "Difficult to assess accurately due to marked athetoid movements".

Active range of motion: "Limited to flailing type movements of upper and lower extremities".

Tone/Spasticity/Clonus: "Marked fluctuations in muscle tone with choreoathetoid-type movements".

Fine Motor control: "Unable to functionally use hands secondary to athetoid movements. He effectively uses switches with his head and legs".

Communication:

“Essentially nonverbal due to severe dysarthria”.

“Attempts to communicate verbally, although he is rather difficult to understand”.

“Frequently attempts to communicate using speech at a one to two word level: Approximately 25% intelligible to a familiar listener in known contexts”.

Electronic scanning devices:

Previously used: E-tran boards, Zygo instruments device.

Current system: Light Talker with MinSpeak semantic compaction. Inputs information into the VOCA using a single knee switch.

While working with the subject we were able to observe that he used a mixture of dysarthric speech/vocalisations, gestures, eye gaze, facial expression and knee-switch operated VOCA to expressively communicate on a day-to-day basis.

5.4 Gestural Subset

Although gestures were produced using varied and often multiple parts of the body, the task of transducing let alone integrating and recognizing movement data is anything but trivial. Attempting to investigate the computer recognition of the complete gesture sets documented in the last chapter was beyond the scope of the project, requiring unavailable technology and resources.

The gestural ability pilot study described in chapter four indicated that for the majority of subjects the arm featured more frequently than any other part of the body. This fact provided the rationale for limiting the study to dynamic arm gestures.

Although the ultimate goal is to be able to transduce movement from any part or multiple parts of the body, the computer recognition of CP movement can be investigated in some depth using data from a single body site.

It was hypothesised that *sufficient pertinent information could be transduced from such gestures using a single six-degree of freedom magnetic tracker attached distally to one forearm.*

A sub-set of gestures produced in the gestural ability pilot study were selected on the basis that they involve one arm as a principal component. This gesture set together with example verbal prompts are listed in table 5.1. The set comprises twenty-seven gestures (reduced to twenty-six during segmentation for reasons discussed in section 6.6). The size of the gesture sub-set was chosen rather arbitrarily on the basis that it was one greater than the number of letters in the English alphabet.

A variety of gestural forms were chosen for inclusion in the sub-set. The set comprised gestures that were clearly distinct in their use of space, gestures that used similar areas of space, but had different form, and gestures that were rather similar to each other (see section 6.4).

N c o d e ^a	Gesture	Verbal Prompt
1	bird	Show me a bird
2	cards	Pretend to play cards
3	cut throat	Pretend to cut your throat
4	drive the car	Pretend to drive the car
5	drums	Pretend to play the drums
6	heavy weight	Pretend to hold a heavy weight
7	helicopter	Show me a helicopter
8	hot	Show me it's hot
9	ice-cream	Pretend to eat at an ice-cream
10	ironing	Pretend to do the ironing
11	knock on the door	Pretend to knock on the door
12	lasso	Pretend to lasso the steer
13	light feather	Pretend to hold a light feather
14	rainbow	Show me a rainbow
15	rock a baby	Pretend to rock a baby
16	rock guitar	Pretend to play a rock guitar
17	scratch your knee	Scratch your knee
18	shake hands	Pretend to shake hands
19	shave	Pretend to shave
20	spank	Pretend to spank your brother
21	spider	Show me a spider
22	stroke the cat	Pretend to stroke the cat
23	surrender	Surrender!
24	whistle	Pretend to pull the whistle on the train
25	umbrella	Pretend to hold an umbrella
26	violin	Pretend to play the violin
27	waiter ^b	Pretend to be a waiter in a restaurant

Table 5.1: Gesture Sub-set

- a. This code is used as a numerical identifier for each gesture class
 b. Discarded from gesture set due to reasons discussed in section 6.6

5.5 Experimental Rig

The experimental rig was designed so that it could be transported to the subjects' school (see figure 5.3). A quiet room was provided by the school and the rig installed for the duration of the data collection process. The floor was marked with tape to ensure that equipment remained in the same position for each data collection session. A wooden post and base was constructed and the transmitter attached to the top using self-adhesive

velcro. A wheelchair registration plate was constructed that when placed on the floor ensured that the wheelchair could be located in a similar position for each data collection session. The rig was designed so that data could be collected from either the right or left arm. The transmitter of the magnetic tracker was positioned so that it was as close as possible to the receiver without any possibility of the hand or arm coming into contact with it. It was positioned so as to minimize the possibility that the distance between the centres of the receiver and transmitter exceeded the range-limit of 36 inches.

The receiver was attached via an elastic wristband that fastened around the forearm using velcro (see figure 5.2). Velcro was used to secure the receiver to the wristband on the underside of the receiver and an additional velcro band across the top minimised movement of the receiver relative to the wrist band. The wristband was carefully adjusted so that it was comfortable to wear and minimised movement between the receiver and the forearm. The design shown in figure 5.2 was arrived at after some experimentation.



Figure 5.2 The "Flock of Birds" magnetic tracker receiver was attached distally using a velcro and elastic wristband developed for the study.

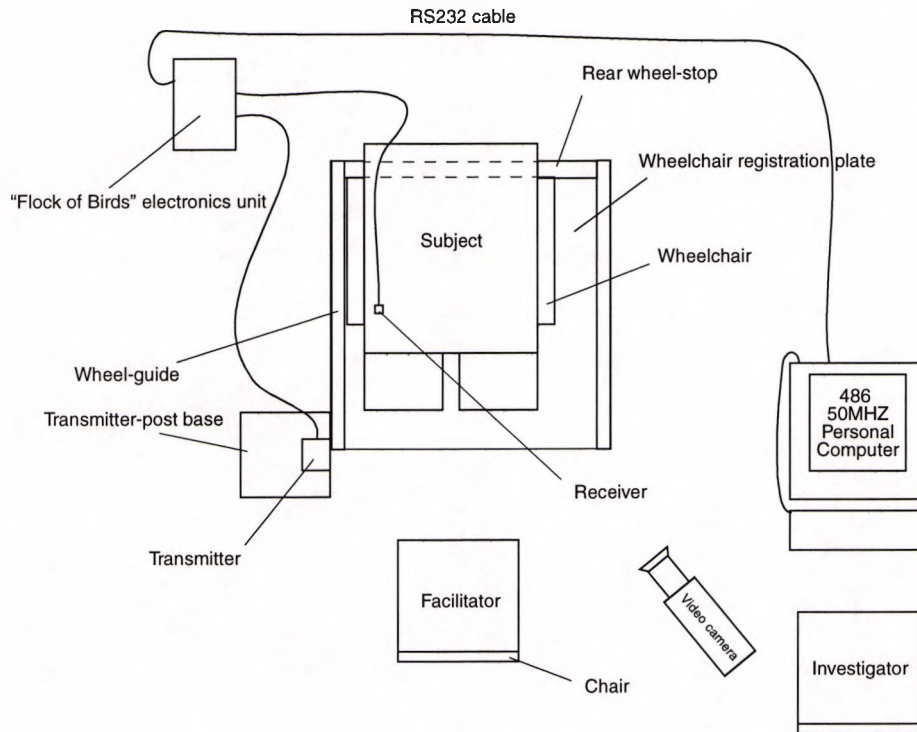


Figure 5.3 Plan view of experimental rig. The facilitator and subject sat facing each other. The wheelchair registration plate ensured that the subject's wheelchair was positioned in the same position relative to the transmitter and the video camera. The gesture tracking system and videocamera were controlled by the investigator. The receiver was attached distally to the subject's right forearm.

5.6 Gesture Tracking System

The "Flock of Birds" six-degrees of freedom magnetic tracker is a self-contained measurement system. The host computer receives data and sends control commands via an RS232C serial interface (see fig. 5.5). Both the transmitter and receiver consist of three large coils of wire wound on perpendicular axes, enclosed in plastic casings. The electronics unit sends short DC pulses to each transmitter coil in rapid succession. This generates a brief electromagnetic field along each successive axis which is sensed by the receiver. The microprocessor in the electronics unit uses the magnitude of these signals to calculate the 3D position and orientation up to 100 times per second, sending the measurement to the computer in real-time with a lag of around 24ms (Pauch, 1991, reported in Meyer et al., 1992). The magnetic tracker was operated in the position/matrix output format mode. The data record output from the magnetic tracker in this mode comprised of three position co-ordinates and nine elements of the square rotation matrix that define the position and orientation of the receiver relative to the transmitter. This output record per sample is described in table 5.2. Each field comprises two bytes sent across the RS232 port in low-high byte pairs.

Software was written in C++ code to poll the unit for a measurement record 100 times per second, the magnetic tracker's nominal measurement rate. The computer's real-time clock (RTC) was temporarily reprogrammed to provide the required timing interrupts

and the measurement was stored in real-time on the computer's hard-disk¹. An interrupt driven serial driver supplied by Ascension Technology Corporation was modified to allow for the increased frequency of RTC timer ticks. As soon as the gestural data were received at the serial port they were transferred to a circular buffer of two kilobytes and then written to a file on the hard-disk of the PC.

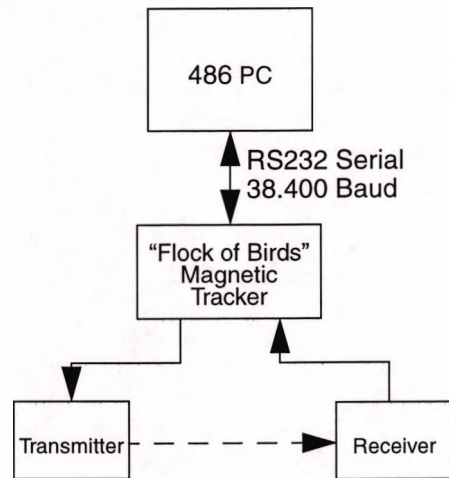


Figure 5.4 Schematic of "The Flock of Birds" magnetic tracker interfaced to a 486 PC.

The manufacturer's specifications supplied with the device were as follows (Ascension Technology Corporation, 1993):

- Positional range: +/- 36 inches from transmitter over one hemisphere
- Angular range: +/- 180 degrees azimuth and roll, +/- 90 elevation
- Static positional accuracy: 0.1 inch RMS averaged over the translational range
- Positional resolution: 0.03 inches at 12 inches
- Static angular accuracy: 0.5 degrees RMS averages over the translational range
- Static angular resolution: 0.1 degrees RMS at 12 inches
- Update rate: 100 measurements/sec
- Environment: 30 degrees C +/- 10 degrees in an environment void of large metal objects and electromagnetic frequencies other than the power line.

Note that measurements are static. Dynamic performance data are not provided. Static system accuracy was verified in the laboratory to be within specification at the extremes of range along each axis. It was beyond the resources of the research project to investigate the dynamic accuracy of the tracker. System functioning was verified by inspection of the data stream using MATLAB graphics and by comparing the Silicon Graphics computer animation of tracker movement with the NTSC video.

Position			Elements of Rotation Matrix, M								
X	Y	Z	M(1,1)	M(2,1)	M(3,1)	M(1,2)	M(2,2)	M(3,2)	M(1,3)	M(2,3)	M(3,3)

Table 5.2: "Flock of Birds" output record per sample

¹ Thanks to John Gray of ASEL for the code for the RTC timer function

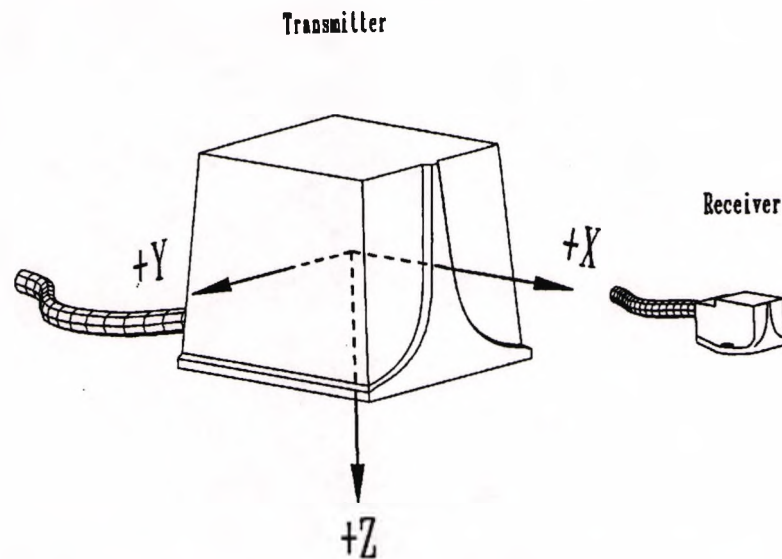


Figure 5.5 Diagram showing the transmitter and receiver modules and the co-ordinate system used by the "Flock of Birds" six-degree of freedom magnetic tracker. (After Ascension Technology, 1993).

The correspondence with video appeared good within the limits of visual inspection (see figures 6.3 to 6.6). The device was found to occasionally produce "glitches" during periods of rapid arm movement (detailed in section 6.6) but these were usually short-lived and the effects were reduced by low pass filtering (see section 6.7 and figure 6.9).

5.7 Gestural Data Collection Sessions

Each session took place in the familiar surroundings of subject S9's school. A room was provided by the school for this purpose that offered a quiet environment with the minimum likelihood of interruptions. Three sessions over three days were scheduled into the subject's regular school timetable, each session taking the place of a regular therapy session. Prior to the commencement of data collection, the investigators organised meetings with the subject, therapists, and teaching staff. During these meetings the purpose and requirements of the data collection sessions were discussed and feedback invited. This ensured that objectives were understood, that the investigators appreciated school and subject needs and that potential problems could be anticipated and accommodated.

Each session lasted 50 to 60 minutes with 33 to 43 minutes of data collection divided into a number of shorter data collection periods separated by breaks of 1 to 5 minutes.

Session schedule:

- | | |
|---|---------|
| 1. Positioning subject, attaching receiver, social interaction with subject | 10 mins |
| 2. First gesture sequence comprising one gesture of each class | ~3 mins |
| 3. Four to five randomly elicited gesture sequences, each lasting exactly | 10 mins |

As with the gestural ability pilot study, the subject was made to feel at ease and in control. It was made clear to the subject that he could take a rest at any time during data

collection if he felt tired (he never exercised this option). Between sequence data collection, the subject was permitted to rest for a few minutes. The exact duration depended on the indicated preference of the subject. The subject was asked whether he was ready to resume before the next series was started.

The gesture sequences were determined by the phrases written on cards that were shuffled before each session. A separate pack was used for the first data collection period at the beginning of each session. This comprised just one gesture per gesture class. This pack was also shuffled before each session.

The subject was warned that the data collection period (DCP) was about to begin and the facilitator waited for confirmation that he was ready before each DCP was initiated. The video camera was started at the beginning of the session just before the first DCP and was allowed to continue running until the end of the session. The investigator verbally announced the name of the computer data file that would store the DCP. The data collection software caused the computer to emit a beep signifying that data recording had begun. This was used as a cue by both the subject and the facilitator so that they could start the interaction. The facilitator read aloud the contents of each card in turn. The subject did not see the contents of the card, but responded to the verbal elicitation and produced a gestural response. In order to produce a connected stream of gestures in a manner that approximated that of a viable gestural HMI, the facilitator attempted to manage the interaction in a number of ways:

- Establishing and maintaining a comfortable and even pace of gesture production in a way that was likely to lead to sustained performance.
- Maintaining an appropriate level of cognitive engagement and interaction by varying the style of elicitation, e.g. tone of voice, the type of engagement, verbal “banter” connecting elicitations.
- Monitoring the subject’s response and estimate fatigue and state of arousal and adjust the pace and style of elicitation appropriately.

Data was collected over three sessions, one per day over three consecutive days.

5.8 Results and Discussion

5.8.1 Gesture co-articulation and Timing

All gestures were naturally co-articulated in as far as they were produced without requesting the subject to rest between gestures or asking the subject to move to a certain position. The next gesture was elicited as soon as the facilitator saw that the gesture had been completed. She based her decision on her perception of the gesture. It was usual for the subject’s arm to still be moving when the next gesture was verbally elicited. Close examination indicated that often movement from the previous gesture was still present for a short period after the beginning of the next verbal elicitation. As gestures were elicited in random order, the transition from one gesture to the next could follow a wide range of path and distances.

5.8.2 Rate of Production of Gestures

Table 5.3 documents the number of gestures produced during each session, the total data collection time per session and the average time per gesture. The rate of production during the last session was rather less than the other two. The subject appeared to show visible signs of fatigue during this session. Although it was not possible to attribute

cause, this is consistent with day-to-day variability in physical ability often observed in people with cerebral palsy.

The average rate of production was one gesture per 9.5 seconds or 6.3 gestures per minute. This measure includes the time for the facilitator to say the next word or phrase and the time taken for the subject to respond. The rate is very similar to the rate of selection for indirect scanning (typically 5 to 10 words per minute, Foulds, 1985 reported in Beukelman & Mirenda, 1992). This rate was maintained over three or four ten minute periods with only a few minutes rest in between during each session. So, from this result it has been shown that it is possible for a person with spastic-athetoid quadriplegia to produce gestures at a sustained rate for a considerable length of time.

Session	No. of Gestures	Duration (min)	Average Time per Gesture (sec)
1	233	34.81	9.0
2	225	34.66	9.2
3	262	44.73	10.2
Total	720	114.20	9.5

Table 5.3: Summary statistics for instrumented gestural data collection sessions

Chapter 6

Examination and Processing of Gestural Data Stream

6.1 Introduction

This chapter describes the process of examining the gestural data stream using computer graphics animation and the subsequent signal conditioning and manual segmentation. The body model and the approximations needed to create an animation of the arm from the single point data are described. The computer animation is compared with the videographic record. Gestures are categorised in terms of their movement characteristics. The co-articulated nature of the gestures is discussed. Examples of sensor noise from the magnetic tracker are presented. With a view to developing a strategy for automatic gesture recognition, the nature of the recognition problem is discussed together with the advantages and disadvantages of applying feedforward neural networks. A fixed-time window scheme involving the use of feedforward neural networks in a time-delay scheme is described and a rationale for adopting this method presented. The process of manual segmentation is described together with details of gestures that could not be segmented. It is shown how preprocessing the signal using low pass filtering reduced sensor noise and reduced the “jerkiness” in the spastic-athetoid movement and reduced the size of the input feature vector to a practical size.

6.2 Gesture Animation

In order to review and segment the data it was necessary to develop a tool that allowed the gestures to be visually reconstructed using computer graphics. MATLAB's graphics routines were too slow to animate an arm model in real-time. Instead, intermediate files of frame-by-frame polygon data were created from the magnetic tracker data using a

body model¹ coded in MATLAB script. The polygon data files were then used to create a real-time animation using a Silicon Graphics (SGI) Iris Indigo Elan workstation. This was achieved through the development of the ANIM² program. The program used the SGI hardware graphics kernel to animate the body model at speeds up to 30 frames per seconds (dependent on the size of the image). The program rendered polygons using Gouraud shading and enabled the gesture to be played back at varying speeds and viewed from any angle. The mouse could be used to manually scan or step through frames forward or backwards. Individual frames could be examined, gesture segments chosen and repeatedly played back and viewed from various angles. Comparisons between gestures could be made by running more than one copy of the program on the workstation. The current frame number, segment start frames and stop frames were continuously displayed. An example of six copies of ANIM being used to display three exemplars of the gesture “cards” is shown in figure 6.1. In this example, each gestural exemplar is displayed using two different viewpoints.

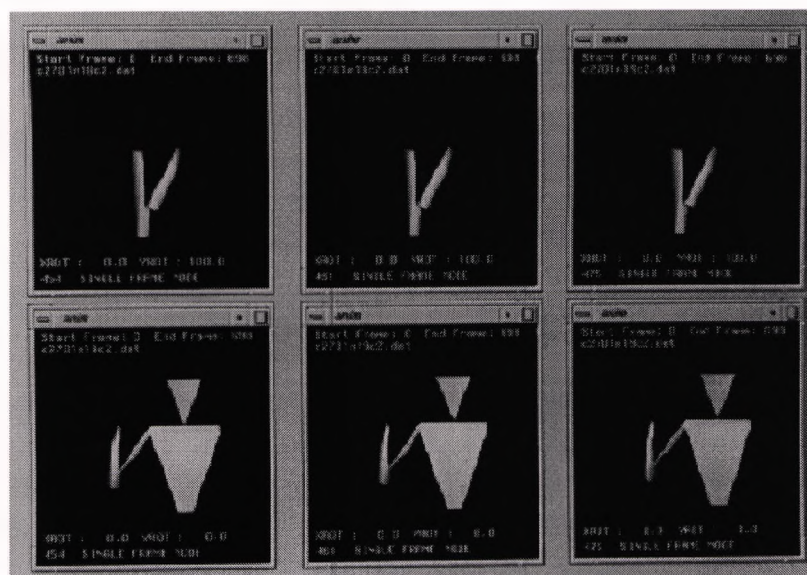


Figure 6.1 Dynamic CP arm gestures animated using body model coded in MATLAB and displayed and viewed from different angles using six copies of the ANIM program running simultaneously on a SGI Iris Indigo Elan workstation. Three exemplars of the “static” gesture “cards” are displayed from the side (top row) and the front (bottom

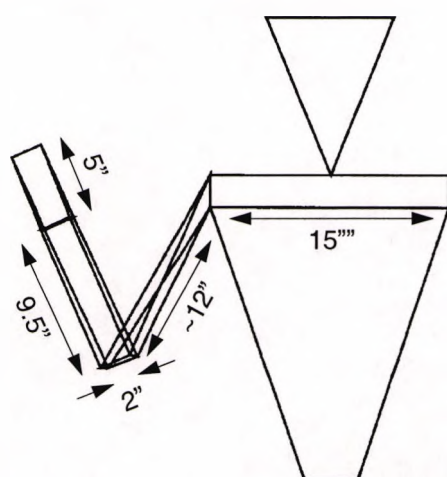
6.3 Body Model

One of the problems of graphically reconstructing arm movement from the magnetic tracker data was that the data was incomplete. Mechanically the shoulder is rather complex, with many degrees of freedom. Since only one tracker was available it was not possible to continuously monitor shoulder position. However, the minimum useful body representation needed to include forearm, upper arm, head and torso. These were needed in order to visually interpret the gesture. Treating the forearm as a rigid body constructed

- ¹ Thanks to Marilyn Panayi for suggesting a suitable level of abstraction for the body model. Thanks to Marion Harrington for coding the initial body model in MATLAB script.
- ² Thanks to Randy Glass and Dr. Garland Stein for the initial coding ANIM and numerous modifications.

from six polygons it was possible to calculate the global position of each vertex by using the magnetic tracker position data to transform the forearm model vertices in local co-ordinates (origin centre of tracker's receiver) to the global co-ordinates (origin at the centre of the magnetic tracker's transmitter). Creating the forearm and body presented more of a challenge since the shoulder position was unknown. Use was made of the fact that, for most arm gestures in the gesture set, the path traced by the hand (and hence the distal end of the forearm) conveyed much of the meaning of the gesture. The body and shoulder were fixed to the position at the start of the first session. The upper arm was constructed by creating six polygons from the vertices of the forearm elbow and the shoulder. This meant that on the occasions when the shoulder moved forward or the torso bent to one side, the length of the upper-arm increased. This seemed to only minimally affect perception of the gesture and was certainly adequate for gesture segmentation. Studying the video record, it was clear that some of the gesture involved significant torso and shoulder movement. However, the base of the trunk was static as it was restrained by the wheelchair. Allowances had to be made when viewing gestures where the relationship of the arm to the head was important. The head was most likely to be angled away from the vertical due to the difficulty that the subject had in maintaining it in the mid-line position (a common problem with CP). This meant that gestures such as "shave" and "(eat an) ice-cream" appeared laterally offset relative to the head of the animation. Gestures involving movement at thoracic level were less of a problem. Despite these approximations, the representation was clear enough to enable the gestures to be readily identified and segmented. Another problem was the representation of the hand. Early experiments developing the model revealed that it was important to represent the hand. However, wrist flexion data was not available. A wedge-shape was employed to represent the hand.

This rather abstract representation of the body was appropriate for the limited data available. Too little detail in the depiction would be detrimental to conveying the sense of the gesture. Too much detail would lead to the model approximations becoming apparent resulting in a negative effect on gesture interpretation. The model comprised 32 polygons and 34 vertices with the following dimensions (see figure 6.2):



- Forearm length = 9.5 inches
- Wrist dimension 1 = 2 ins
- Wrist dimension 2 = 1.2 inches
- Width across shoulders = 15 inches
- Torso thickness = 2 inches
- Shoulder position (inches) = [29.10 11.77 -9.64]

Figure 6.2 Body model comprising 32 polygons used to animate the "Flock of Birds" magnetic tracker arm movement data. Forearm length, wrist dimensions, and shoulder width were based on physiometric data of subject S9.

6.4 Qualitative Examination of Gesture Set using Animation

Figures 6.3 to 6.6 show comparisons between the computer graphics animation of the magnetic tracker data (sampled at 100hz) and nearest corresponding frames of the videographic record (NTSC frame rate 29.7hz). Figures 6.3 and 6.4 depict different exemplars of the gesture class "rainbow". This gesture involved the arm starting on the left side of the body at approximately head level tracing an arc above the head to the right side of the body. Looking at the first exemplar in figure 6.3, shoulder and torso movement is clearly apparent from the first frame of the video (top left). Although this information was not available to the animation, the form of the gesture and its relationship can still be ascertained. Notice there is considerable head movement. In general, the head position does not correspond to the mid-line "neutral" head position depicted in the animation. Similarly, finger and wrist flexion and wrist rotation can be seen to vary with different gestures. Noticeable variation of wrist rotation can be distinguished between the two exemplars. The gesture occurs as a single movement along a path.

Figure 6.5 depicts an exemplar of the gesture "surrender". The gesture involved raising the arms rapidly upwards to a position high over the head. This is also classed as a single movement towards a static pose³. Figure 6.6 depicts an exemplar for "stroke the cat". This gesture involves the extension of the right arm to a position over the left knee and then traces an elliptical path downwards across the lap and then repeats for a variable number of cycles. This gesture is classed as a periodic movement.

Within the natural variation of each gesture, all gestures appeared to be produced along consistent paths and or in consistent regions of space except for the gesture "waiter". This gesture involved dramatic interaction with the facilitator to such a high degree that it was produced in a variety of ways. Although each form was consistent with the "mime", it did not result in consistent movement patterns.

Through study of the graphics animation it was found that it was possible to classify the gestures into three groups based on the movement characteristics of each gesture: static pose, single movement, and periodic movement. The results of this categorisation are described in table 6.1. Eight gestures involved static pose, four involved a single movement and sixteen involved periodic movement. Of these gestures one was performed with a single movement leading to a static pose ("surrender") and one was a periodic movement leading to a single movement ("lasso"). "Waiter" was not categorised because of inconsistency of form.

³ "Static pose" included gestures that included movements in a small region e.g. "shave", "ice-cream".

Ncode	Gesture	Static Pose	Single Movement	Periodic Movement
1	bird			x
2	cards	x		
3	cut throat		x	
4	drive the car			x
5	drums			x
6	heavy weight	x		
7	helicopter			x
8	hot			x
9	ice-cream	x		
10	ironing			x
11	knock on the door			x
12	lasso		x	x
13	light feather	x		
14	rainbow		x	
15	rock a baby			x
16	rock guitar			x
17	scratch your knee			x
18	shake hands			x
19	shave	x		
20	spank			x
21	spider	x		
22	stroke the cat			x
23	surrender	x	x	
24	whistle			x
25	umbrella	x		
26	violin			x
27	waiter	-	-	-
	Total	8	4	16

Table 6.1: Gestures categorised in terms of movement characteristics



Figure 6.3 Corresponding video frames (left) and computer animated frames of magnetic tracker data (right) for the gesture “rainbow” exemplar 1. Frames were chosen to convey the form of the gesture.



Figure 6.4 Successive corresponding video frames (left) and computer animated frames of magnetic tracker data (right) for the gesture “rainbow” exemplar 2. Frames were chosen to convey the form of the gesture.



Figure 6.5 Successive corresponding video frames (left) and computer animated frames of magnetic tracker data (right) for the gesture "surrender". Frames were chosen to convey the form of the gesture.

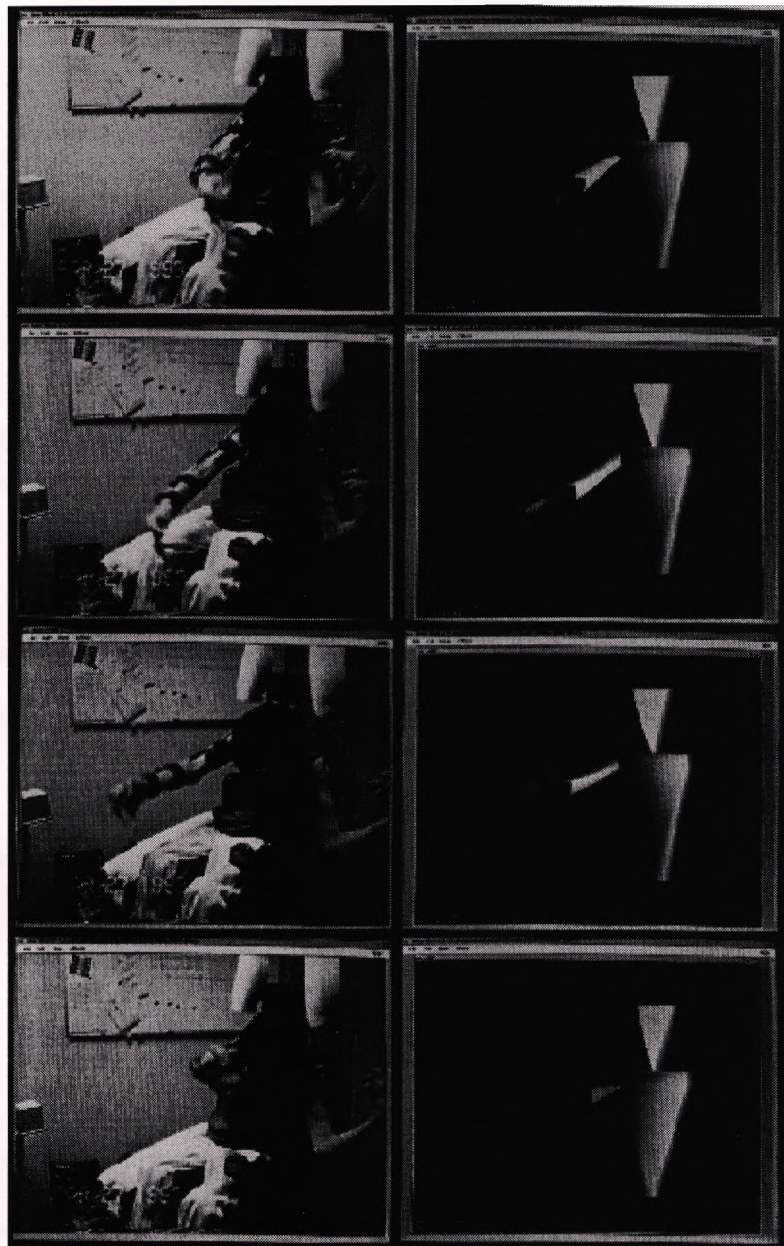


Figure 6.6 Successive corresponding video frames (left) and computer animated frames of magnetic tracker data (right) for the gesture “stroke the cat”. Frames were chosen to convey the form of the gesture.

6.5 Developing a Strategy for Gesture Recognition

6.5.1 Key Factors Affecting Gesture Recognition

A number of key factors that affect the design of automatic gesture recognition can be determined from the nature of the gesture set. These are listed below together with their implications:

Idiosyncratic gesture patterns

This suggests that the system should be trainable using sample gestures from the user.

Gesture patterns were co-articulated i.e produced as a connected sequence

This increases the pattern recognition difficulty substantially. Known previous attempts by others to automatically recognise the gestures from people with SSMICP either involved gestures produced with a pause between each gesture (e.g. Cairns, 1993, Harrington, 1995) or gestures segmented by thresholding scheme such as a "tremor filter" (Harwin, 1990, Perricos, 1994). In the latter case, gestures require the user to have sufficient movement control to inhibit their movement between gestures. Connected speech recognition has proved considerably more challenging than isolated word recognition. It is a similar situation for handwriting recognition, cursive script is much more difficult to recognise than single characters.

Low signal-to-noise ratio data

Signals can often be usefully processed to improve the signal-to-noise ratio. However, to do this optimally is a challenge as the relationship between the signal (volitional movement component) and the noise (uncontrollable movement component) is complex, and poorly understood.

Small number of gesture classes

This increases the chance of attaining practical real-time recognition. So far gesture recognition algorithms have only been successful on small gesture sets.

Gesture set comprises gestures that are static, single movements, or periodic movements.

The presence of static and dynamic gestures in the gesture set prevents the use of movement level for the differentiation of gesture from non-gesture. As described, the periodic gestures varied in the number of cycles between exemplars although the form of each cycle seemed relatively consistent. This is mindful of Rubine's "eager recognition" strategy (1991) where gestures were recognised as soon as possible rather than wait for the gesture to finish.

Addition requirements for a practical system:

- System should be able to learn initially from a relatively small training set of training exemplars with the potential to retrain itself as more gestural data becomes available through use.
- The system should have the potential to advise the user or therapist of the set of gestures that are recognised well. In this manner, it could contribute to the development of a gestural repertoire that can be recognised robustly.

6.5.2 Fixed Time Window Recognition Algorithms

As previously stated, for a gesture recognition system to be of use to people with SSMICP, it must be able to recognise co-articulated gesture sequences. One gesture must be able to flow into another so that the interface can be used without undue effort on the part of the user. As gestures can be produced in any sequence and can conceivably start and end anywhere in gesture space, movement from the end of one gesture to the beginning of the next is likely to be highly variant. This, coupled with the presence of atetotic movement, sheds doubt on the possibility of automatically segmenting the gesture based on features extracted at the beginning and end of gestures.

Manual segmentation was attempted using the computer animation tool ANIM. This process is described in detail in section 6.6. As noted in this section, it was necessary to

observe the movement pattern as a whole in order to achieve segmentation. If this process could be automated, it would correspond to automatic segmentation by recognition. As already noted, although the gestures were produced over variable periods, the gestures of long duration were either totally static or periodic. It was found that the salient repeated part of all single movement gestures fitted with a time-span of around one second. At least one cycle of all periodic gestures also fitted within the same time interval. Thus, the variation in gesture duration was not as high as appeared at first examination.

Although there exists temporal variation between members of the same class and different classes, it is hypothesised that at minimum a subset of these gestures are automatically recognisable using a time-window approach. A scheme is proposed where features are extracted from data over a fixed interval or window back in time. It is proposed that these features could be applied to the input nodes of a feedforward neural network in a time-delay scheme. Conceptually, this involves connecting the input nodes of the neural network to a tapped delay-line through which the sequence of feature vectors passes. In this way, identical features produced over a fixed period of time get mapped to the spatial layout of the neural network input nodes, one feature per node. Although this scheme of recognition has limited ability to accommodate to temporal variation, it has proved effective in practice (Giles, 1994).

It is proposed that in the fully implemented gesture recognition engine, features will be continuously extracted from the gestural data stream. Activation levels of the output layer of the neural network continuously offer a measure of the confidence level that a particular gesture is being produced at any moment in time. A further algorithm determines when a gesture is occurring and rejects spurious network predictions if their duration is too short. Development of this part of the algorithm is left for future work.

As a first step towards implementing this scheme, gestures were manually segmented into gesture segments of 1120ms duration. This was chosen on the basis that the duration of all single movement gestures and at least one cycle of periodic movement gestures was less than this time period.

Manually segmented gestural data were divided up into data sets for training and testing of neural network classifier performance. Although recognition rates are likely to be higher than those obtained in the fully implemented gesture recognition engine, the results are useful for comparing feature sets, comparing classifiers, determining network architecture, and indicating the size of the time window that yields optimum performance for the gestural data set. Time windows of 160ms to 1120ms are compared in section 8.2 using filtered and re-sampled x,y,z position data as input features.

6.5.3 Advantages and Disadvantages of Feed Forward Neural Networks

In addition to the previous rationale, feedforward neural networks arranged in a time-delay scheme were chosen as they offer the following advantages:

- A priori knowledge of the relationship between input and output is not required
- The neural network architecture is relatively well understood, therefore it is an appropriate place to start when considering neural methods
- As pattern classifiers, they are capable of constructing non-linear feature surfaces of complex topology e.g. meshed classes
- Training can be achieved with an initial small set of data and then retrained adap-

tively as more data becomes available

However, disadvantages include:

- Network training using backpropagation of errors is a gradient search method. As such, it can get stuck in a local minimum far from the optimum solution (although the addition of “momentum” and “adaptive learning rate” helps minimise this risk).
- The final weight matrix can be dependent on the initial weights which are initialised randomly.
- There are a number of problem dependent parameters to adjust
- Training can be a relatively long process.
- It is difficult to examine the neural network’s internal representation of a problem
- The feature vector of a time-delay neural network contains data sampled over a fixed time period.

6.6 Segmentation

The gesture type and elicitation time were obtained by logging the information from the sound-track videographic record. The longest data recording period lasted ten minutes. The start and finish of each magnetic tracker data recording period was captured on the video soundtrack by a series of computer generated tones. The gesture type and time of verbal elicitation were obtained by listening to the soundtrack. The time was obtained logging the time-display of the video recorder. The available equipment did not have a time-stamping capability, so some error in the elicitation time log was expected. This was estimated to be less than 1s.

These data were used to automatically segment the data stream into gestural data segments. The time duration of each segment was the time from one verbal elicitation to the next or 15 seconds in the instances where there was a break in gesture production or no gesture followed. Polygon animation data was then generated for each segment using MATLAB. These data were animated and studied using ANIM to determine a more precise start and end time for each gesture based on gestural features.

More often than not it was necessary to view the whole gestural segment a number of times to appreciate the gesture in its entirety, and then back-up frames to a point that could be considered to be the start of the gesture.

It was impossible to identify one common feature in the co-articulated gesture stream that could be used to signify either the start or the end of all gestures. This is consistent with the work of Harwin in the study of the computer recognition of head gestures. He encountered considerable difficulty in automatically segmenting CP head gestures (Harwin, 1990). This problem also plagues speech recognition. Although we perceive the words of connected speech as distinct sounds, actually the sounds all run together and it proves very difficult for a computer to determine the word boundaries automatically (Mammone, 1994).

One pass was made through all the data to determine the start and end of each gesture within each gestural segment. This defined a time window within which gestural movement appeared to correspond to the gesture in question.

The gesture animation revealed that on occasion the magnetic tracker produced “glitches” where the tracker data would freeze either at the last value or jump to an unlikely position and then jump back again. The “glitches” were often associated with rapid arm movement. An example is shown in figure 6.9 (a) and (c). With some

exemplars the amount of sensor “glitching” was considerable.

Some gestural exemplars were impossible to segment due to sensor “glitching”, others were not segmentable due to there not being anything that clearly represented the gesture. Other gestures were not segmented due to the manner in which they were produced. For example, the gesture may have been produced in a number of different ways. This was the case with the gesture for “waiter”. The gesture for “waiter” became a mime sequence with interaction with the facilitator. Sometimes the mime involved holding a tray at shoulder height, sometimes it involved passing a tray or food to the facilitator. Although this was perfectly acceptable as a mime, its inconsistent form made it impossible to segment. As a result this gesture was removed from the gesture set for the purposes of computer recognition. This reduced the total number of gestural exemplars from 720 to 694.

The proportion of gestures of each type that were not segmentable either because of sensor noise or poor production is described in table 6.2 and as a histogram in figure 6.7. Another manual pass was made to determine a starting point for a window of around one second that contained a representative portion of each gestural exemplar. All gesture set features used in this study were derived from these gestural segments.

Figures 6.10 and 6.11 show graphs of x, y, z position data of gestural segments for exemplars of the gestures “rock a baby” and “hot” respectively. The two vertical lines marked on each plot indicate a gestural segment of 640ms. The artificial neural network input features were extracted from these segments. Zero time represents the time the gesture was verbally elicited. The graphs end at the sample before the elicitation of the next gesture. The region before the first vertical line represent the co-articulated portion of the gesture. For “rock a baby” the previous gesture was “umbrella”. For “hot”, the previous gesture was “[pull the train] whistle”. Plots (b) and (d) represent the data after low-pass filtering. The circles on plot (d) represent the data re-sampled after filtering at 6.25 s^{-1} .

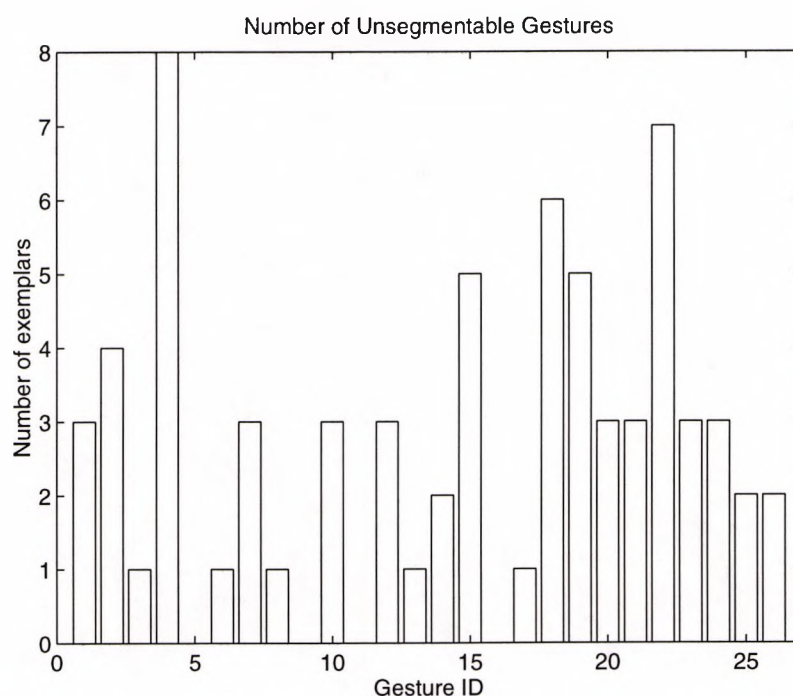


Figure 6.7 Histogram showing the number of gestures that were unsegmentable due to an inconsistent gestural form or severe “glitching” in the magnetic tracker data stream.

ID	Gesture	No. not segmented	Percentage	Total No. Exemplars
1	bird	3	11.1	27
2	cards	4	16.0	25
3	cut throat	1	3.8	26
4	drive the car	8	25.8	31
5	drums	0	0.0	26
6	heavy weight	1	3.7	27
7	helicopter	3	11.5	26
8	hot	1	3.6	28
9	ice-cream	0	0.0	26
10	ironing	3	11.1	27
11	knock on the door	0	0.0	26
12	lasso	3	10.3	29
13	light feather	1	3.8	26
14	rainbow	2	8.3	24
15	rock a baby	5	20.0	25

Table 6.2: The number and proportion of gestures of each class that were not manually segmentable.

ID	Gesture	No. not segmented	Percentage	Total No. Exemplars
16	rock guitar	0	0.0	27
17	scratch your knee	1	4.3	23
18	shake hands	6	22.2	27
19	shave	5	15.6	32
20	spank	3	10.7	28
21	spider	3	12.0	25
22	stroke the cat	7	24.1	29
23	surrender	3	11.1	27
24	whistle	3	11.5	26
25	umbrella	2	7.4	27
26	violin	2	8.3	24
	All 26 gestures	70	9.87	694

Table 6.2: The number and proportion of gestures of each class that were not manually segmentable.

6.7 Signal Conditioning/Data Reduction

Figure 6.9 (a) shows the position data for the gesture “stroke the cat” sampled at 100 samples per second. The jagged component in the signal is due to a combination of “jerky” movement characteristic of spastic-athetose CP and sensor noise. The benefit of low-pass filtering the gestural stream was investigated by choosing a cut-off frequency of 2.8125Hz. A Chebychev IIR filter type I of order nine and 0.5 dB ripple was designed. The data stream was then filtered using the FILTFILT function of MATLAB. This performed a forward and then backward filtering pass over the data to produce an output with zero phase shift. Start-up and end transients were minimised by matching initial conditions (MathWorks Inc., 1992). The frequency response of the filter is plotted in figure 6.8. Data was re-sampled at 6.25Hz or every 16th sample. Note that at half the sampling frequency (3.125Hz) the response is attenuated by over 20dB. Since the FILTFILT function passed twice over the data the effective filter response drop-off is somewhat steeper.

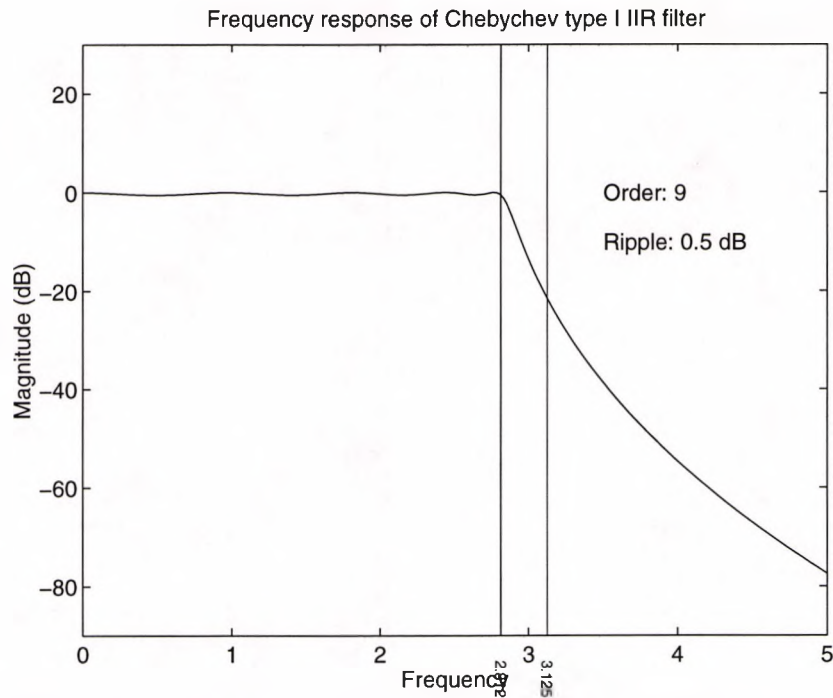


Figure 6.8 Frequency response of low-pass chebychev IIR filter type I used to filter the gestural data stream. The cut-off frequency was 2.8125Hz. Data was re-sampled at 6.25Hz (every 16th sample).

As can be seen from figure 6.9, low-pass filtering the magnetic tracker data smooths out the gestural movement, removing the “jerkiness”. In addition, it removes or reduces high frequency sensor noise. LP filtering with a 2.8125 Hz cut-off was additionally justified by inspecting the resulting computer animation. When the filtered data were animated, the gestures appeared considerably more intelligible (i.e. closer to the investigator’s stereotype of the gesture). Figure 6.12 shows the power spectral density of the complete gestural stream of position data comprising one exemplar of each gesture. These plots show that the dominant frequency components of the signal reside below the filter cut-off.

Re-sampling reduced the effective sample rate from 100s^{-1} to 6.25s^{-1} representing a data reduction ratio of 16:1. This made it feasible to construct input feature vectors comprising sequential frames of position data and present these and other feature vectors in a time-delay fashion to feedforward neural networks.

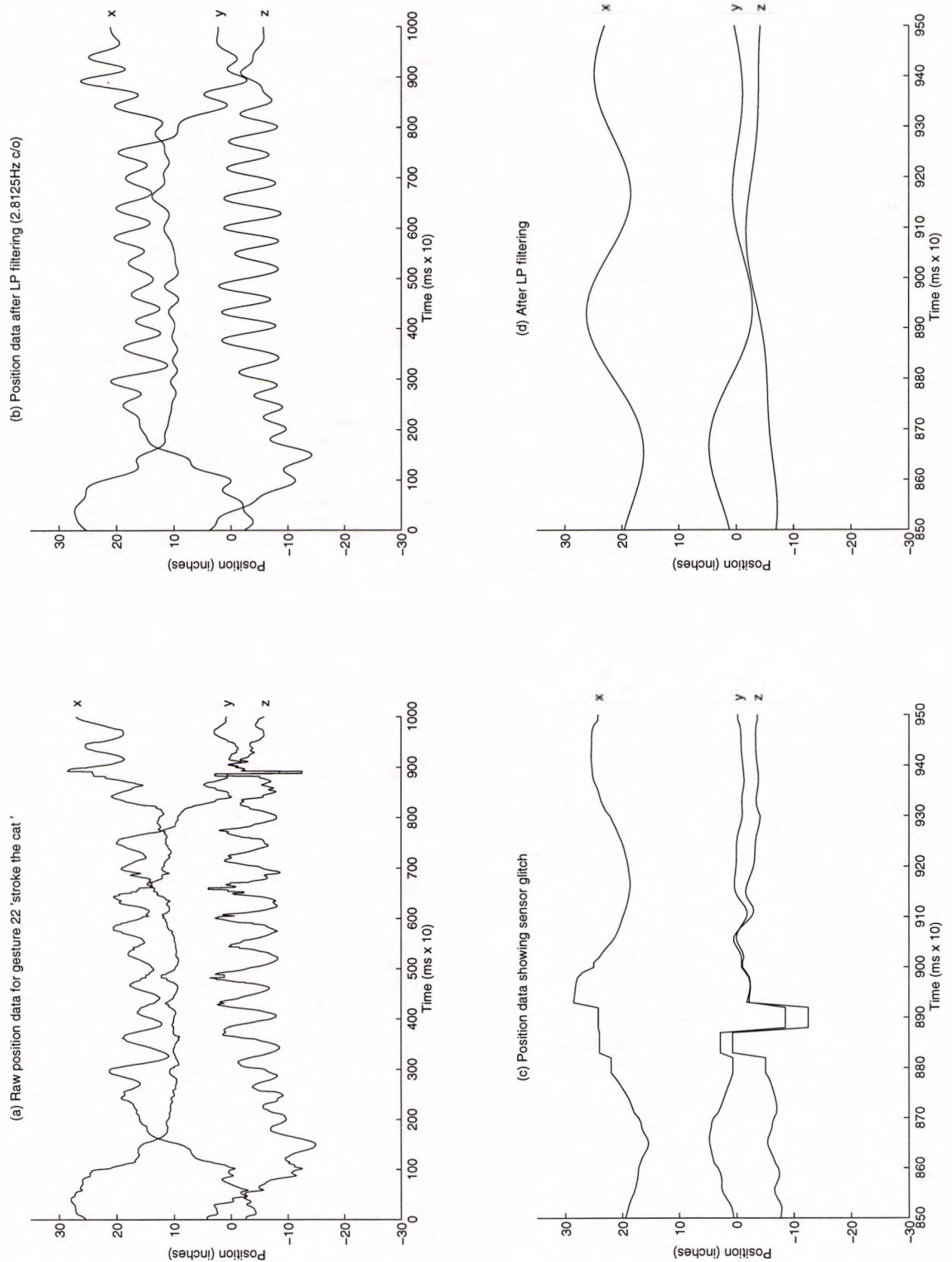


Figure 6.9 Gestural stream position data before ((a), (c)) and after ((b), (d)) filtering. The graphs show how low pass filtering reduces the jerky component of spastic-athetoid movement, high-frequency sensor noise, and sensor "glitches" (between 7416 and 7563 ms).

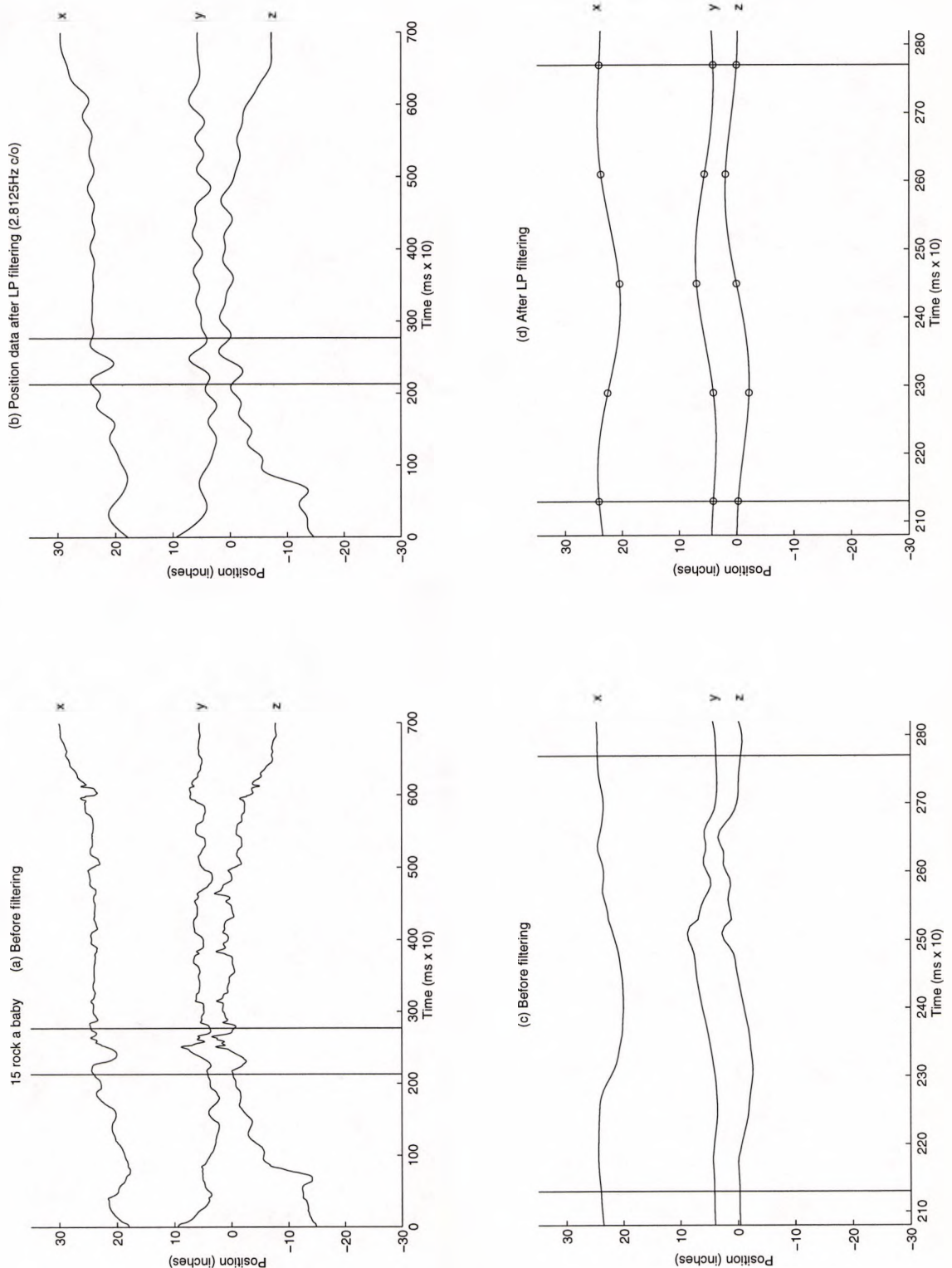


Figure 6.10 Plots of position data for one exemplar of “rock a baby” showing manual segmentation. The vertical lines represent the start and finish of the 640ms time window used to create many of the feature sets. (a) and (b) are plots of raw x, y, z magnetic tracker data. (c) and (d) are plots of the signal after low-pass filtering. The circles on plot (d) represent the re-sampled data points. The portion of the gesture before the first vertical line represents the end of the previous gesture (“umbrella”) plus the movement between gesture.

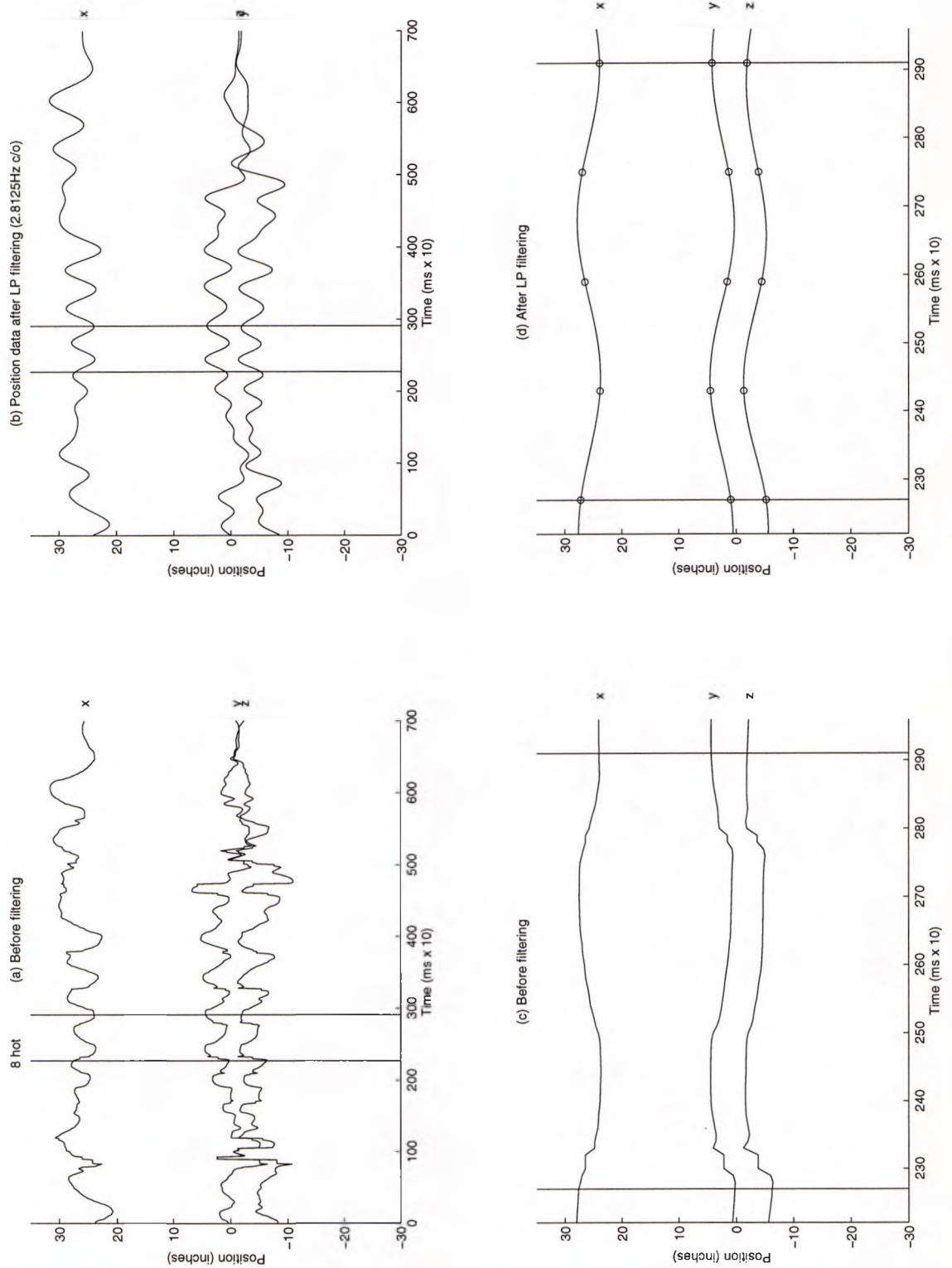


Figure 6.11 "Plots of position data for one exemplar of "hot". The vertical lines represent the start and finish of the 640ms time window used to create many of the feature sets. The circles on plot (d) represent the re-sampled data after low-pass filtering. The preceding gesture was "(pull the train) whistle". The portion of the gesture before the first vertical line represents the co-articulated portion of the gesture.

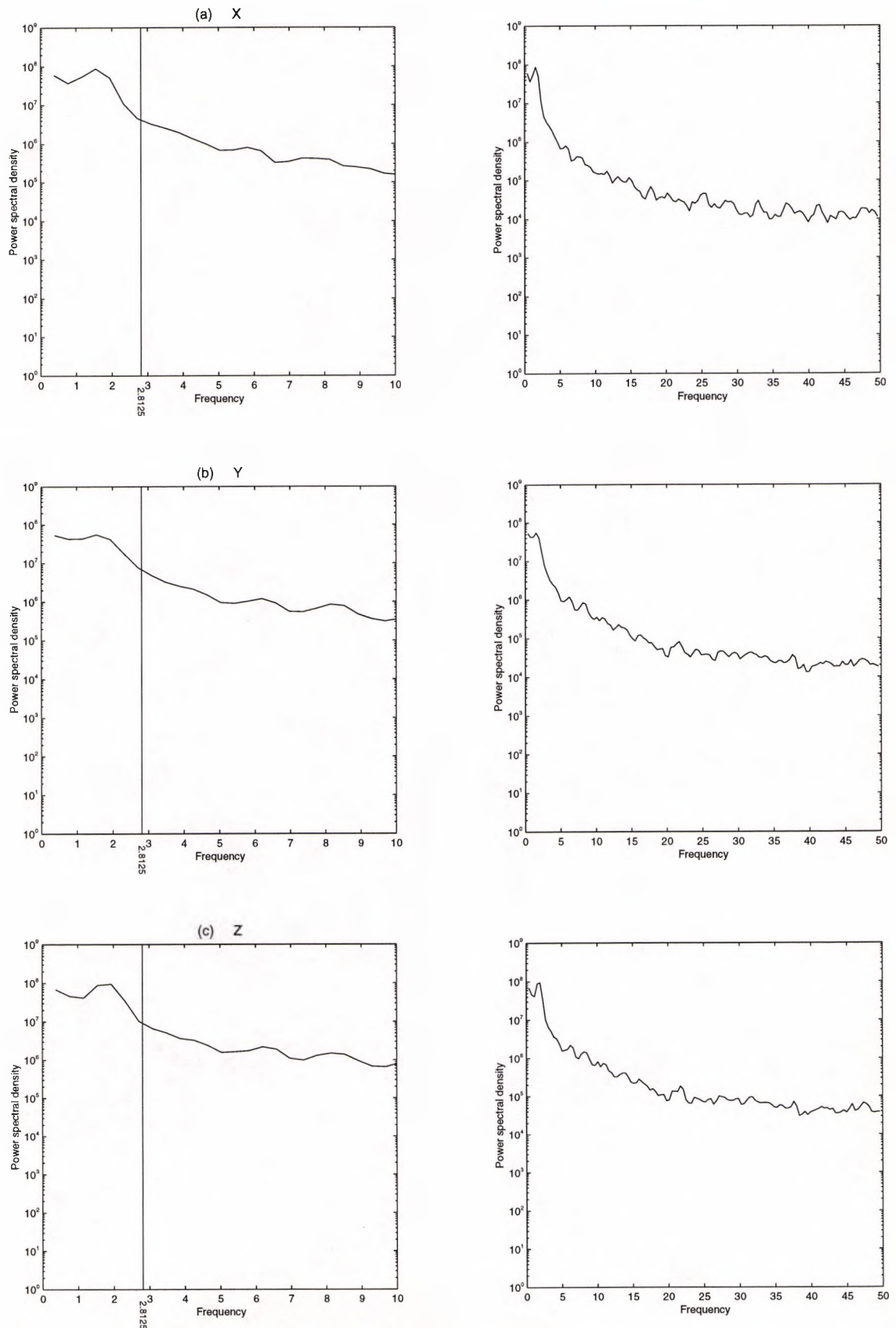


Figure 6.12 Power spectral density of the gestural data stream for x, y, and z positional data showing the cut-off frequency of low-pass filter used to smooth the gestural motion and to remove sensor noise

Chapter 7

Gesture Classification using Neural Networks

7.1 Introduction

The following two chapters describe and document the use of feedforward neural networks to classify CP arm gestures. In this chapter, the neural network architectures, initialisation and training methods are described.

The gesture recognition performance measures used to compare neural networks and feature sets is described and justified.

The methods used to prepare the gesture set for training and validation are described together with the inherent limitations.

Network complexity is discussed and the process of determining the number of hidden-layer neurons required to solve the problem documented.

Network training behaviour using gestural data is described.

Various combinations of activation functions in the hidden and output layers of the neural network are compared. Finally, gestures are classified by two k-nearest-neighbour methods using one feature set for purposes of comparison.

All experiments were based upon the fixed-time window approach involving successive time frames of gestural data described in section 6.5. Feature set descriptions and comparisons are described in chapter 8.

Algorithms were coded using MATLAB v4.2 script utilising functions from the MATLAB Neural Network Toolbox v2.0b where appropriate.

7.2 Neural Network Description

Mathematical descriptions of feedforward neural networks can be found in most introductory neural network texts. Fausett (1994), Zurada (1992) and the MATLAB Neural Network user guide (Demuth and Beale, 1994) were consulted during the writing of this chapter.

7.2.1 Elementary Neuron Model

The elementary neuron model used in this study comprised a summation stage (where weighted inputs and bias are summed) and a differentiable activation function. The neuron in figure 7.1 is shown with R inputs. Each input is associated with an appropriate weight, w_{ij} . The weighted inputs are summed together with bias b to form the input to an appropriate activation function to produce activation a_j (eqn. 7.1).

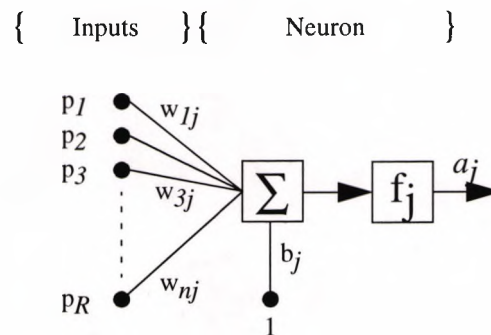


Figure 7.1 Elementary backpropagation neuron with inputs \mathbf{p} and synaptic weights \mathbf{w} and bias b . Weighted inputs and bias are summed to form the input to an appropriate activation function.

$$a_j = f_j \left(\sum_{i=1}^R p_i w_{ij} + b_j \right) \quad \text{Eqn [7.1]}$$

7.2.2 Activation Function

Notionally the activation function could be any differentiable transfer function f_j .

Commonly used functions are the tan-sigmoid¹

$$f(x) = \frac{2}{1 + \exp(-\rho x)} - 1 \quad \text{Eqn [7.2]}$$

and its derivative

$$f'(x) = \frac{\rho}{2} (1 + f(x)) (1 - f(x)) \quad \text{Eqn [7.3]}$$

or log-sigmoid

$$f(x) = \frac{1}{1 + \exp(-\rho x)} \quad \text{Eqn [7.4]}$$

¹ Sigmoid refers to the "S" shape of the curve.

$$f'(x) = \rho f(x) (1 - f(x)) \quad \text{Eqn [7.5]}$$

These functions act as “soft-threshold” functions and have the effect of constraining the outputs no matter how large the magnitude of the input sum² (figures 7.2 and 7.3). The main difference between the two families of curves is that log-sigmoid has asymptotes at $f(x)=0$ and $f(x)=1$ while tan-sigmoid has asymptotes at $f(x)=-1$ and 1 .

Another function often used in the output layer is the identity function

$$f(x) = x \quad \text{Eqn [7.6]}$$

The following tan-sigmoid function was used as the activation function for the hidden layer neurons in most of the experiments in this study:

$$f_1(x) = \frac{2}{1 + \exp(-2x)} - 1 \quad \text{Eqn [7.7]}$$

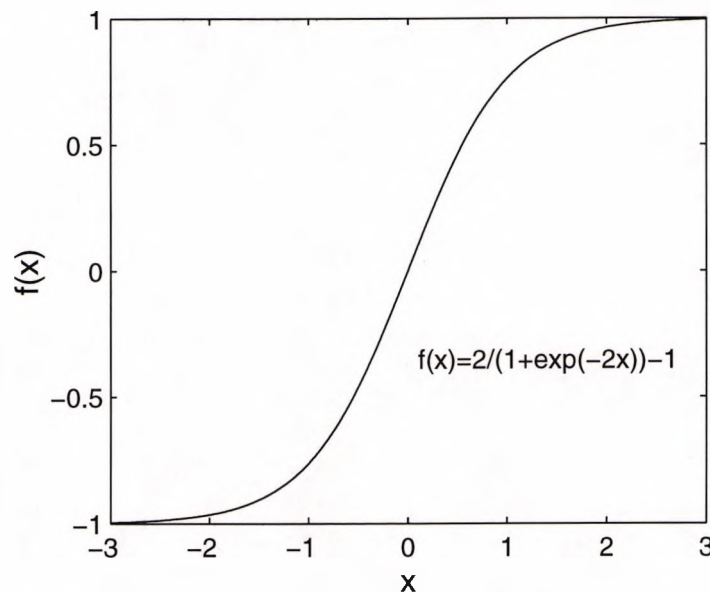


Figure 7.2 Graph of tan-sigmoid activation function for $\rho = 2$

The log-sigmoid function with $\rho = 1$ (eqn. 7.8, fig. 7.3) was also used, but with this particular pattern recognition problem, either failed to converge to a solution during training or resulted in significantly poorer performance than tan-sigmoid (see section 7.7).

$$f(x) = \frac{1}{1 + \exp(-x)} \quad \text{Eqn [7.8]}$$

² They are sometimes referred to as “squashing functions” because of this property.

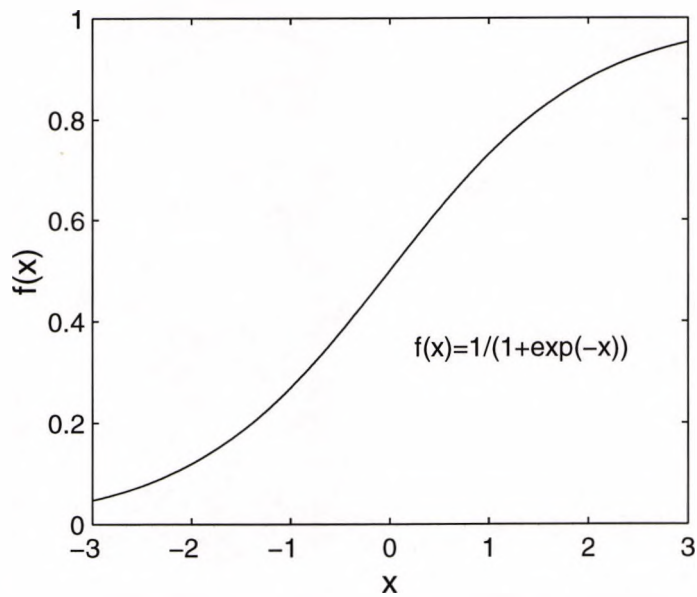


Figure 7.3 Graph of log-sigmoid activation function for $\rho = 1$

7.2.3 Network Architecture

A fully connected feedforward artificial neural network (FFNN) comprises one input layer, one or more hidden neuronal layers, and an output neuronal layer. Figure 7.4 depicts a FFNN with one hidden layer. This is usually referred to as a two layer network. k

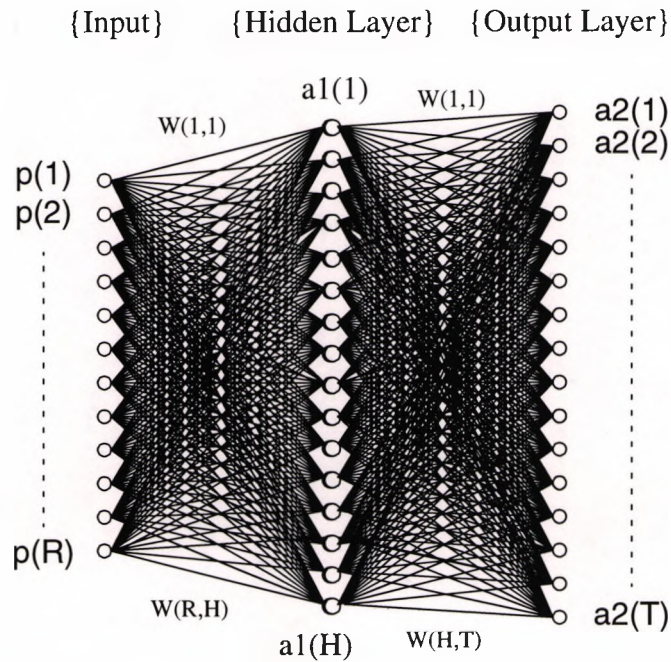


Figure 7.4 Architecture of a fully connected feedforward neural network with a single hidden layer. The diagram shows the weights assembled as matrices W and V and neuronal input p and outputs $a1$, $a2$ as vectors for the MATLAB simulation. Biases are not shown.

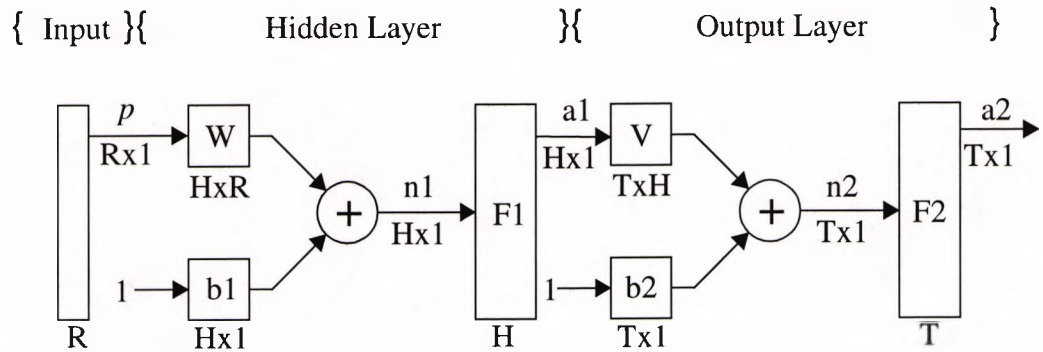


Figure 7.5 Schematic representation of two layer feedforward neural network architecture showing bias vectors and matrix dimensions (Adapted from Demuth & Beale, 1994)

as it comprises two neuronal layers (the input layer contains no neurons).

In this study, only two layer feedforward networks were studied. Kolmogorov's theorem predicts that theoretically any continuous function $f(x_1, \dots, x_n)$ of several variables

defined on I^n ($n \geq 2$), where $I = [0, 1]$, can be represented in the form

$$f(x) = \sum_{j=1}^{2n+1} \chi_j \left(\sum_{i=1}^n \psi_{ij}(x_i) \right) \quad \text{Eqn [7.9]}$$

where χ_j and ψ_{ij} are continuous functions of one variable and ψ_{ij} are monotonic functions that do not depend on f (Kolmogorov, 1963). Thus theory supports the notion that two neuronal layers are likely to be sufficient to model most problems.

A schematic of the general network architecture is presented in figure 7.5. The network architecture involved an input layer where each feature vector \mathbf{p} had dimension R . The summed weighted inputs plus bias formed the input to the activation function of each neuron. The weighted sum of neuron activations from the hidden layer a_1 plus a bias for each neuron form the inputs to each output layer activation function. One output neuron was assigned to each gesture class.

This can be expressed as:

$$a_{1j} = f_1 \left(\sum_{i=1}^R p_i v_{ij} \right) + b_{1j} \quad \text{Eqn [7.10]}$$

$$a_{2k} = \sum_{j=1}^H a_{j1} w_{jk} + b_{2k} \quad \text{Eqn [7.11]}$$

where $k = \{1, 2, \dots, 26\}$ corresponding to output nodes, one per gesture class

a_{1j} , the output activation of the j th hidden neuron

a_{2k} , the activation of the k th output of the network

w_{jk} , the value of the weight connecting the j th hidden neuron to the k th output

v_{ij} , the value of the weight connection the i th input to the j th hidden neuron

p_i , the i th input

b_{1j} , the bias on the j th hidden neuron

b_{2k} , the bias on the k th output neuron

R , the total number of input nodes

H , the total number of hidden neurons

Convention for describing neural network architecture

The following terminology is used through this thesis to concisely describe the neural network architecture used in each experiment:

a-bf-cg

a , no. of nodes in input layer, equivalent to dimension of feature space

b , no. of nodes in hidden layer

c , no. of nodes in output layer

f , activation function used in the hidden layer

g , activation function used in the output layer

$f, g \in \{t, l, p\}$

where:

t , tan-sigmoid fn.

l , log-sigmoid fn.

p , identity fn.

e.g. 12-16t-26p represents a network of 12 input nodes, 16 nodes each with tan-sigmoid activation functions in the hidden layer, and 26 nodes with the identity activation function.

Tan-sigmoid neurons in the hidden layer (eqn. 7.7) and identity function neurons in the output layer (eqn. 7.6) were used in preference to tan-sigmoid neurons in both layers partly on the basis that simpler is better and partly on the results of experiments that suggested that the former architecture yielded a marginally higher recognition rate. Since it is not possible to control the number of nodes in the output layer, 26 non-linear neurons in this layer can be expected to result in over-fitting and poor generalisation.

7.2.4 Network Learning using Back-propagation of Errors

The multilayer feedforward architecture, although fairly simple in structure was once thought to be untrainable (Minsky & Papert 1969). A training method was found by Werbos (1974) but failed to become widely publicised. It was also independently re-discovered by Parker (1985) and by LeCun (1986). The algorithm was also very similar to a yet earlier optimal control algorithm (Bryson & Ho, 1969). Although there is some

argument as to who discovered the training method, it was not until it was refined and publicised by David Rumelhart and James McClelland that the method of “backpropagation of errors” became widely recognised as a method of training multilayer neural networks (Rumelhart, Hinton, & Williams, 1986a, 1986b; McClelland & Rumelhart, 1988).

Backpropagation is a gradient search method of optimization and potentially suffers from the danger of getting stuck in local minima. Whether this phenomenon is observed or not depends on the problem. It depends on whether there are deep local minima in regions far from the desired solution. It should be noted that when training a feedforward neural network (FFNN), we do not want the weight vector that represents the true global minimum of the sum-squared error. This is because the global minimum would in fact represent the weight vector of an overtrained network. Ideally, the search should be converging towards the global minimum. In practice there is no easy way to establish that one is heading for that global minimum rather than a local minimum. However, if the network has trained well enough for a particular task it is of little practical importance, although it is always conceivable that by starting with a different set of initial weights a better solution could be found.

In this study, momentum and adaptive learning rate were added to the delta rule used for the backpropagation of error. Momentum lessens the danger of getting stuck in a local minimum and is added by linearly combining the most recent gradient with the previous gradient (see appendix B, section B.1).

Adaptive learning rate increases the size of the incremental weight change if the previous weight change reduced the sum-squared error. This has the effect of increasing training speed.

It is possible to choose when to update the weight vector e.g. incrementally after each training exemplar or at the end of a complete pass of input-output training pairs. In this study the weights were updated at the end of each epoch backpropagating the network errors produced for the complete training set as a batch (often referred to as “batch mode”).

7.2.5 Weight and Bias Initialization

The weights and biases are usually initialized using small random values between certain limits. Weights and biases between the input and hidden layer \mathbf{v} were initialized using the Nguyen-Widrow method (see appendix B, section B.2). This method is designed to improve the learning ability of the hidden neurons. This is accomplished by distributing the initial weights and biases so that for each input pattern it is likely that the net input to one hidden neuron will be in a range that is conducive to rapid learning.

The weights between the hidden layer and the output layer \mathbf{w} were initialized to random values between +1 and -1.

7.3 Determining a Performance Measure for Gesture Recognition

The average recognition rate of all 26 gestures gives an indication of gesture recognition performance for the complete set of gestures. However, in real-life, poorly recognised gestures would be discarded from the set as their continued inclusion would most likely

be counter productive. Gestures recognised at a low rate would be frustrating to say the least. So, the average recognition rate of individual gestures is important. It is useful to define a recognition rate threshold below which gestures are not considered useful. For the purpose of this study, a threshold of 80% was chosen, i.e. any one gesture recognised correctly every four out of five times. This threshold was based upon the assumption that this rate was the lowest that could be tolerated by the user.

In addition to recognition performance being measured in terms of recognition rate of the complete set, the size of the set of gestures recognised at or above the 80% level was calculated together with the average recognition rate for that set.

7.4 Preparation of Gesture Sets for Training and Validation

The size of the complete gesture set containing gestures of all 26 classes was 624 resulting in an average of 24 gestures per class. It was necessary to divide this set up into data for training the neural network with enough representative exemplars of each gesture for the network to make reasonable generalisations, and leave enough data for a test set to validate the results. To this end, the data was divided by randomly selecting 260 gestures, 10 exemplars of each class for the test set, leaving 364 exemplars, an average of 14 gestures per class for the training set.

In order to train a FFNN using back-propagation it is necessary to determine where to stop training the network, particularly with small data sets. This was achieved in this study by training each network for 100 epochs then testing the network using the test data and storing the resulting confusion matrix together with the weights and biases. The network which yielded the maximum number of gestures recognised at or above 80% was chosen and reported as the test results. It should be noted that this does have implications as regards validation. Although the neural networks were not trained on the test data, knowledge of the test data was used to choose the network with the best performance. Ideally the data set should be divided into three rather than two so that one set can be used to train, the second set used to determine where to stop training, and the third set for validation. However, with an average of only 24 gestures per class available, this was not possible.

7.5 Network Complexity

The complexity of the problem as a pattern recognition exercise was not known a priori, so this needed to be determined empirically. In general, overall problem complexity involves:

- i) Dimension of the pattern-space
- ii) The number of classes
- iii) The topology of the decision surface required to accurately classify the gestures.

In a FFNN, the dimension of feature space corresponds to the number of nodes in the input layer. The upper bound on the complexity of the decision surface is determined by the number of nodes in the hidden layer(s), the number of hidden layers, the type of sigmoid function used at each node in the hidden and output layers. The following subsections describe how the number of hidden nodes was determined experimentally. In

this part of the study, the tan-sigmoid function was used in the hidden layer and the identity function in the output layer.

7.5.1 Method

In order to determine the size of the hidden layer, neural networks with one to twenty hidden layer neurons were compared.

A feature set comprising four time frames of distal arm position was used for this procedure (tr/te2i14p, see section 8.2). The ANN architectures compared comprised:

$$12 - nt - 26p \text{ where } n \in \{1, 2, 3, 4, 6, 8, 10, 12, 14, 15, 16, 17, 18, 20\}$$

The neural networks were trained using momentum of 0.95 and an initial learning rate of 0.01. After each 100 epochs the network was tested and the confusion matrix and weights and biases stored. Training was terminated after 10,000 epochs.

The recognition rate was calculated for the r gestures with the highest recognition rate where $r=1$ to 26. This was calculated for each network every 100 epochs.

Using the test data, both the average recognition rate for 26 gestures and details of the gestures recognised at or above 80% were calculated. The results were then analysed to determine the number of hidden layer neurons that should be used for subsequent experiments.

7.5.2 Results and Discussion

Figure 7.6 is a histogram of average recognition rate for 26 gestures. Results are presented for training and test data. Note how recognition rate increased until the number of nodes in the hidden layer, $n=16$. The recognition rate then fell significantly for $n=17$. This is consistent with the network converging and getting trapped in a local minimum far from the global minimum. This network was retrained once more using different weight initialisation and gave similar results. The maximum recognition rate for 26 gestures was achieved using 18 hidden tan-sigmoid nodes in the hidden layer.

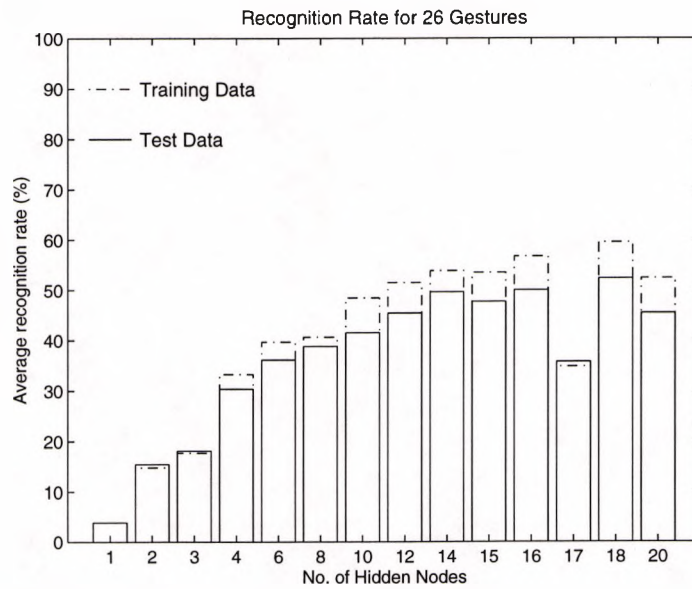


Figure 7.6 Average recognition rate for 26 gestures comparing FFNNs with from 1 to 20 nodes in the hidden layer

The results for gestures recognised at or above 80% are presented in figure 7.7. Looking at graph (b), the number of gestures recognised at or above 80% increases with the number of hidden layer nodes in a similar fashion to average recognition rate of 26 gestures. In this case $n=16$ resulted in the largest number of gestures (12) recognised at an average rate of 90%.

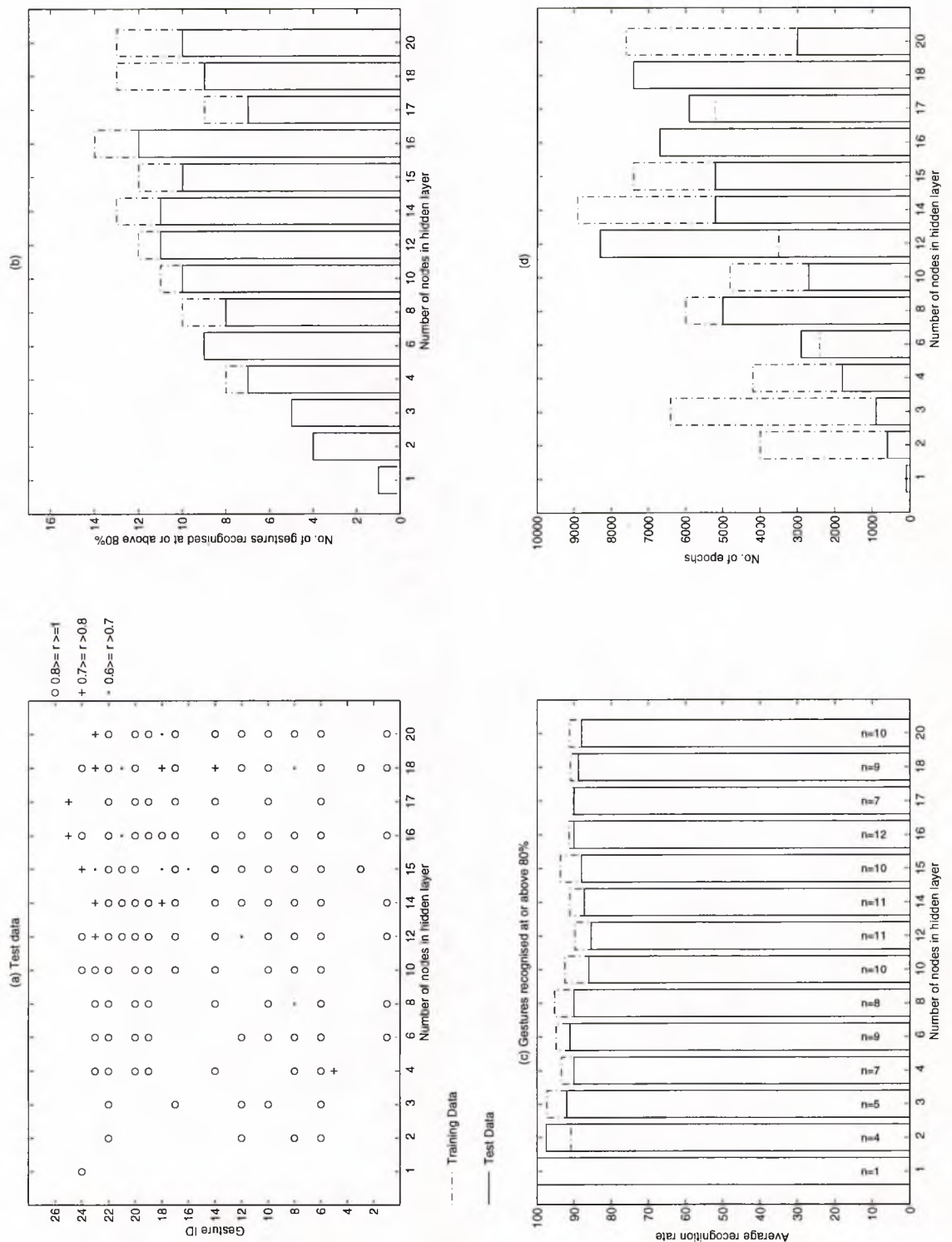


Figure 7.7 Results for gestures recognised at or above 80% for one to 20 nodes in the hidden layer. (a) shows which gestures were recognised at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 \geq r < 0.8$ are marked with '+'. Gestures recognised $0.6 \geq r > 0.7$ are marked with '.'. Histogram (b) shows the maximum number of gestures recognised at $r \geq 0.8$. (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognised at or above 80%).

Tables 7.1 and 7.2 contrast the recognition rates gestures for gesture between the 12-4t-26p and the 12-16t-26p architecture.

Four hidden nodes:

For the network with four hidden nodes, seven gestures are recognised at over the 80% level. These are listed in table 7.1 together with the average recognition rate for the individual gesture for the four node and 16 node networks. The 12-4t-26p architecture comprised 182 weights and biases.

16 hidden nodes:

Adding 12 more hidden nodes results in only five³ more gestures being recognised at or above the 80% level (table 7.2) while the 12-16t-26p architecture comprised 650 weights. Thus, in order to increase the number of gestures recognised at or above 80% by 71% it was necessary to increase the number of weights and biases by 357%.

While most of these gestures increase in recognition rate with more nodes, “heavy weight”, “stroke the cat” and “surrender” decrease.

c o d e	Gesture	4 hidden nodes (%)	16 hidden nodes (%)
6	heavy weight	100	80
22	stroke the cat	100	90
20	spank	90	90
8	hot	80	100
14	rainbow	80	90
19	shave	80	100
23	surrender	80	0

Table 7.1: Seven gestures recognised at or above 80% with only 4 hidden nodes

³ Note that the “surrender” was classified by 12-4t-26p but not 12-16t-26p

c o d e	Gesture	4 hidden nodes	16 hidden nodes
1	bird	10	100
24	whistle	0	100
18	shake hands	0	90
10	ironing	0	80
12	lasso	30	80
17	scratch you knee	10	80

Table 7.2: Additional 6 gestures recognised at or above 80% with 16 hidden nodes

The results indicate that there is a principle of diminishing return as regards the effect of adding more nodes to the hidden layer. An increasing number of hidden nodes are required to recognise just a few more gestures. This suggests that for the feature set comprising 4 time-frames of xyz position (tr/te2i14p), once the decision surface divides the feature space in such a way that 6 gestures are recognised well (defined as 80% or above), the complexity of the feature surface has to increase substantially before more gestures can be recognised. In other words, it is possible to find seven gestures that can be easily distinguished by a simple decision surface, while in order to create decision boundaries for a larger set of gestures, increasingly complex topologies are required due to the presence of complex class boundaries e.g. meshed classes.

The results indicate that the optimum network size is around 16 hidden nodes. Above this, the disparity between the recognition rates of training and test data sets become noticeably greater indicating a loss in generalisation. With such a small data set (13-15 exemplars of each gesture in the training set and 10 exemplars in the test set) it is not surprising that it is relatively easy to over-fit the data.

7.6 Training Behaviour of Neural Networks using Gestural Data

This section documents the training behaviour of a neural network of architecture 12-16t-26p using backpropagation with momentum (0.95) and adaptive learning rate (initially 0.01). The network was trained using the tr/te2i14p feature set comprising four frames of xyz position data. The network was trained until 20000 epochs.

Figure 7.8 shows the change in sum-squared error and adaptive learning rate as training progressed. The sum-squared error is defined in appendix B, equation B.3. The value was relatively high when first calculated at the beginning of training due to the randomly chosen weights. This rapidly fell as the network begins to learn (1 to 100 epochs). From then on the sum-squared error continued to decrease as the backpropagation gradient

search algorithm attempts to converge to a minimum. From time-to-time the sum-squared error increases slightly, often oscillating slightly before returning to its original path. Adaptive learning rate can be seen to increase at an increasing rate until a cusp is reached where it falls rapidly and then starts to increase again. This represents an acceleration down the path of steepest decent until it overshoots. The spikes every 100 epochs are due to the fact that the learning rate was reset to 0.01 after the weights were saved every 100 epochs.

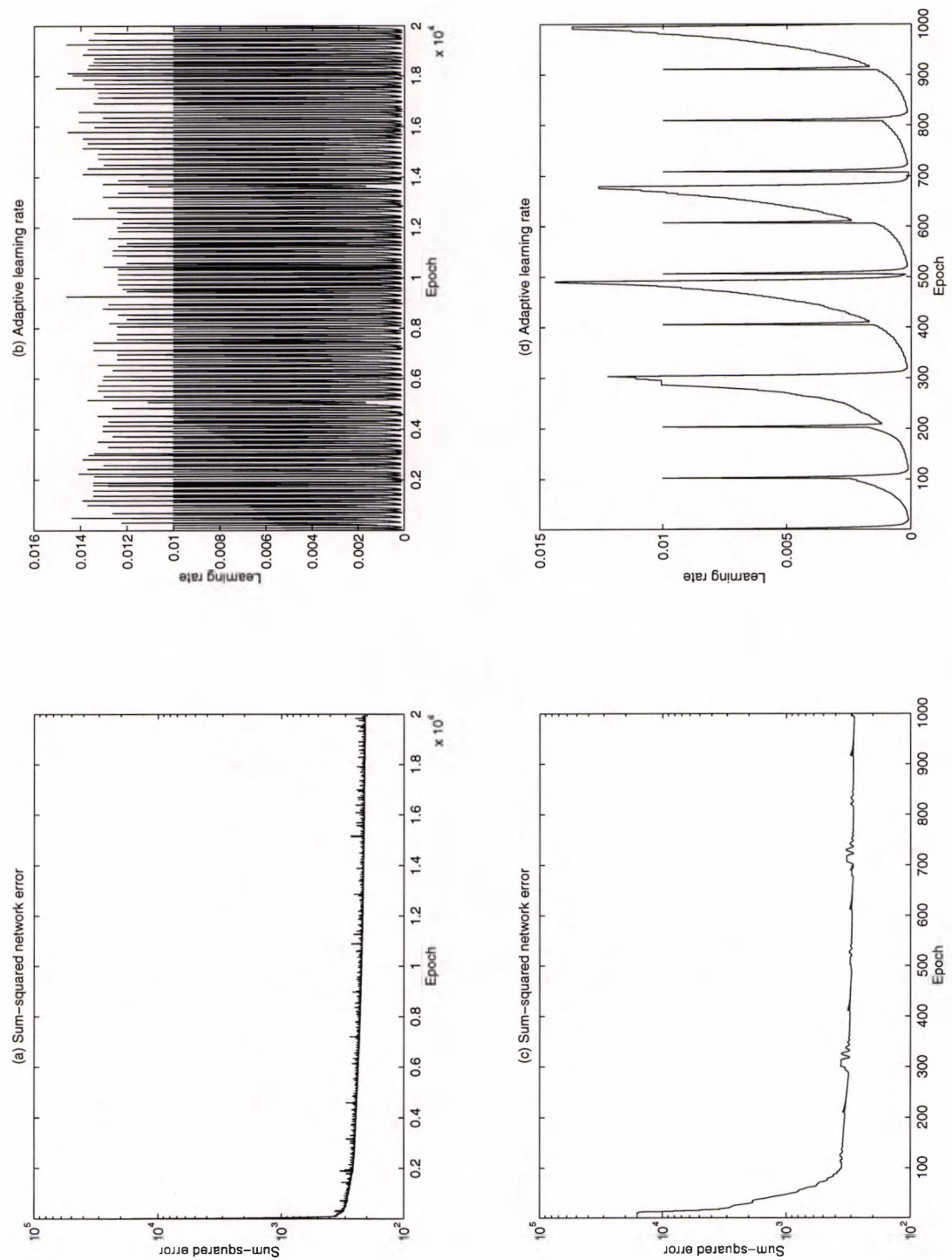


Figure 7.8 Plot of neural network backpropagation training parameters for 12-16t-26 network. (a) sum-squared error, (b) adaptive learning rate during training until 20000 epochs. (c) sum-squared error, (d) adaptive learning rate during training until 1000 epochs

Figure 7.9 shows the neural network's recognition rate from training and test data at 100 epoch intervals until 20000 epochs. Results are shown for all 26 gestures and for the most highly recognised 12 gestures. Note the difference in recognition rate between the training data and the test data. The network has learnt features in the training data that are not good generalisations and are not present in the test data.

Note that although sum-squared error during training (figure 7.8 (a) and (c)) continues to fall up to 20000 epochs, there is no corresponding continued increase in recognition rate of the original training data. This is not fully understood, but it is possible to see how this can happen. The sum-squared error was calculated from the difference between the actual activation level at the output nodes and the target value. The output activation could vary between -1 and +1. However, the decision was based on the output node that exhibited the maximum activation. The same decision would be made whether the differences between the highest valued output node activation and the activation levels of other nodes were large or small. The value of sum-squared error on the other hand could conceivably decrease due to a decrease in some or all of the other nodes, while not affecting the network decision.

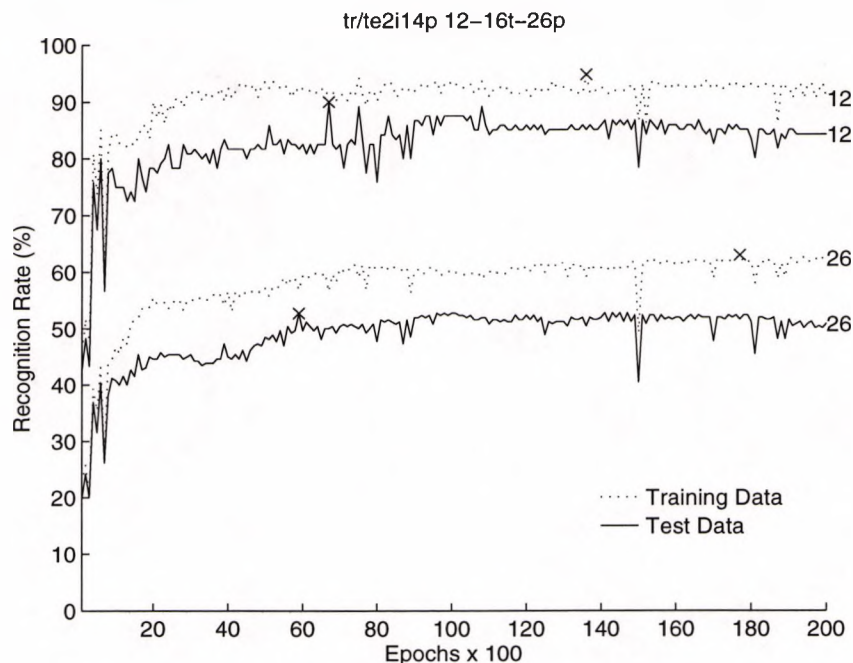


Figure 7.9 Gesture recognition rate against number of training epochs for all 26 gestures and for the best 12 gestures. The maximum recognition rates in each case is indicated by a cross

7.7 Comparison of Activation Functions

7.7.1 Method

In order to ascertain the performance of various combinations of activation functions used in the hidden and output layers, six combinations were compared using the tr/te2i14p feature set. The neural network architectures used for each experiment are detailed in table 7.3. Note that in this table t,p, and l refer to the activation functions used in the corresponding layer. t=tan-sigmoid, p=identity function, l=log-sigmoid.

Ex ID	ANN Architecture
ex16uk	12-16t-26p
ex60uk	12-8t-26t
ex61uk	12-16t-26t
ex62uk	12-16l-26l
ex63uk	12-8l-26p
ex64uk	12-16l-26p
ex65uk	12-16p-26p

Table 7.3: Experiments with varying combinations of activation function

Each network was trained using backpropagation with momentum (0.95) and adaptive learning rate (initially 0.01). The networks were trained and tested in an identical fashion to that described in section 7.5.1.

7.7.2 Results and Discussion

The results for the average recognition rate are presented in figure 7.11. The confusion matrices are documented in appendix C, sections 5 and 2. Different architectures resulted in considerably different recognition results even when the number of weights and biases were identical. All the log-sigmoid networks yielded lower average recognition rates for 26 gestures than the tan-sigmoid networks for the same size of network. In fact the log-sigmoid networks performed either worse or about as well as the network that used only linear identity functions in the hidden and output layers. The worst average recognition rate was produced by the 12-16l-26l network and the best was produced by the 12-16t-26t network (marginally better than 12-16t-26p). Possible explanations include:

- 1) The log-sigmoid networks consistently converge to local minima at some distance from the global minimum.
- 2) The log-sigmoid functions are making poorer generalisations given the small amount of training data available.

The results for gestures recognised at or above 80% are presented in figure 7.10.

As regards the number of gestures recognised at or above 80%, the 12-16t-26p outperformed the 12-16t-26t. Again, the 12-16l-26l performed worst. All log-sigmoid performed worse than 12-16p-26p.

These results indicate that choice of activation function can have significant effects on recognition results. If the plots of the tan-sigmoid and log-sigmoid functions are inspected (see figures 7.2 and 7.3), there are two apparent differences. The first is that output values of a tan-sigmoid neuron lie between +1 and -1 while output values of a log-

sigmoid neuron has output values between 0 and 1. The second is that for the particular forms of the functions used in this study, the linear part of the curve was steeper in the case of the tan-sigmoid function.

The results of this part of the study were used to determine the activation functions used in the FFNNs subsequent experiments. All other neural network experiments used the tan-sigmoid function in the hidden layer and the identity function in the output layer.

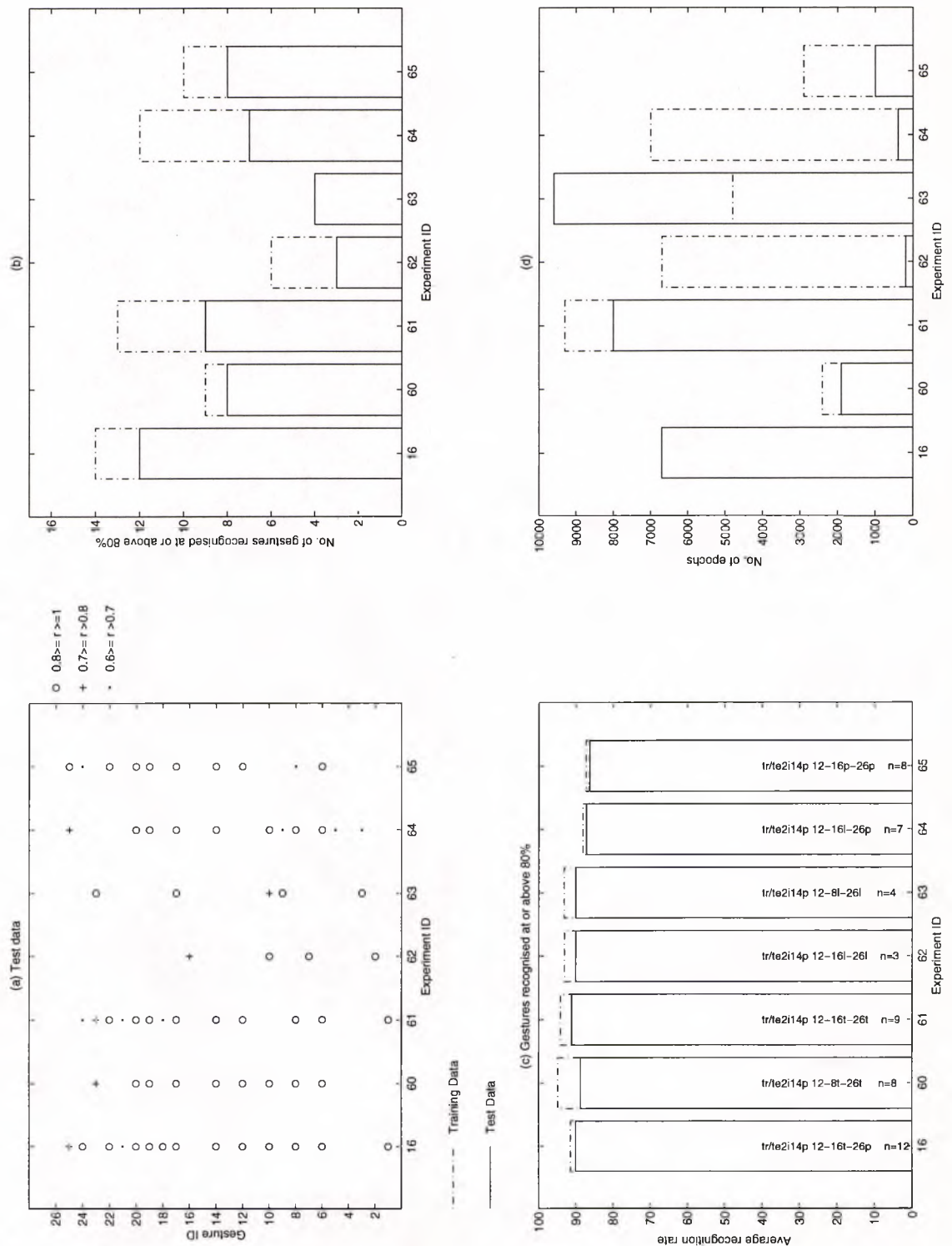


Figure 7.10 Gesture recognition results for six different combinations of activation functions in the hidden and output layers. The tr/te2i14p feature set was used to train and test each network. (a) shows which gestures were recognised at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 \geq r < 0.8$ are marked with '+'. Gestures recognised $0.6 \geq r < 0.7$ are marked with '*). Histogram (b) shows the maximum number of gestures recognised at $r \geq 0.8$. (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognised at or above 80%).

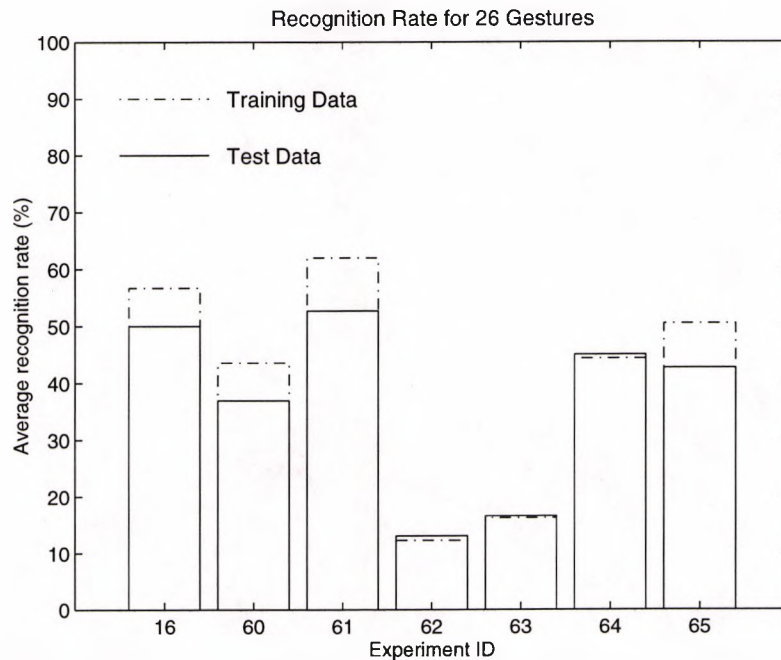


Figure 7.11 Comparison of Activation Functions: Average recognition rate for 26 gestures for experiments 60 to 65

7.8 Comparison with k-Nearest Neighbour Method

7.8.1 Method

In order to obtain a comparative measure of the ANN classifier performance, two k-nearest neighbour type classifiers (kNN) were constructed using MATLAB. Their ability to classify gestures was investigated. The tr2i14p training data set was used as the set of prototypes and te2i14p used to test the classifiers. The two varieties of kNN were a) euclidean distance kNN and b) euclidean distance kNN with standard normalisation. These algorithms are described in most standard texts on pattern recognition (Tou and Gonzalez, 1974).

k was varied from one to nine in each case. The second classifier functioned in an identical fashion to the first except the standard deviation of training prototypes was calculated for each input dimension and the data normalised using this value. This had the effect of compressing space in dimensions where the training set was more widely dispersed so that overall dispersion was approximately hyperspherical.

7.8.2 Results and Discussion

Results are presented for the euclidean kNN classifier in figures 7.12. The corresponding confusion matrices are presented in appendix C, section 3. The maximum average recognition rate for 26 gestures was 55.77% for k=8. The maximum number of gestures recognised at a rate greater or equal to 80% was 6 with an average recognition rate of 86.67. The number of gestures recognised at or above 80% for k=1 was only three. The fact that k=8 gives considerably improved results compared with k=1 means that a large

number of prototypes are required to construct representative piece-wise linear class boundaries. The resultant smoothing of the feature surface improved generalisation.

This suggests that patterns belonging to each class are not tightly clustered compared to the distance between class centres. The results are consistent with there being a lack of representative exemplars and/or the existence of overlapping class boundaries.

Results are presented for the euclidean kNN classifier with standard normalisation in figures 7.13. The corresponding confusion matrices are presented in appendix C, section 4. The maximum average recognition rate for 26 gestures was 55.38% for $k=7$. The maximum number of gestures recognised at a rate greater or equal to 80% was 7 with an average recognition rate of 85.71%.

The results for the euclidean kNN classifier with and without standard normalisation were similar. KNN with standard normalisation recognised one more gesture at or above 80%.

The results are summarized and compared with recognition results using the same feature set for a feedforward neural network with architecture 12-16t-26p (see section 8.2) in table 7.4. The FFNN showed a significantly improved performance over both schemes of kNN. The FFNN recognised 12 gestures at or above 80% with an average recognition rate of 90%.

Interestingly, the kNN results are similar to those of a 12-4t-26l network reported in section 7.5.2.

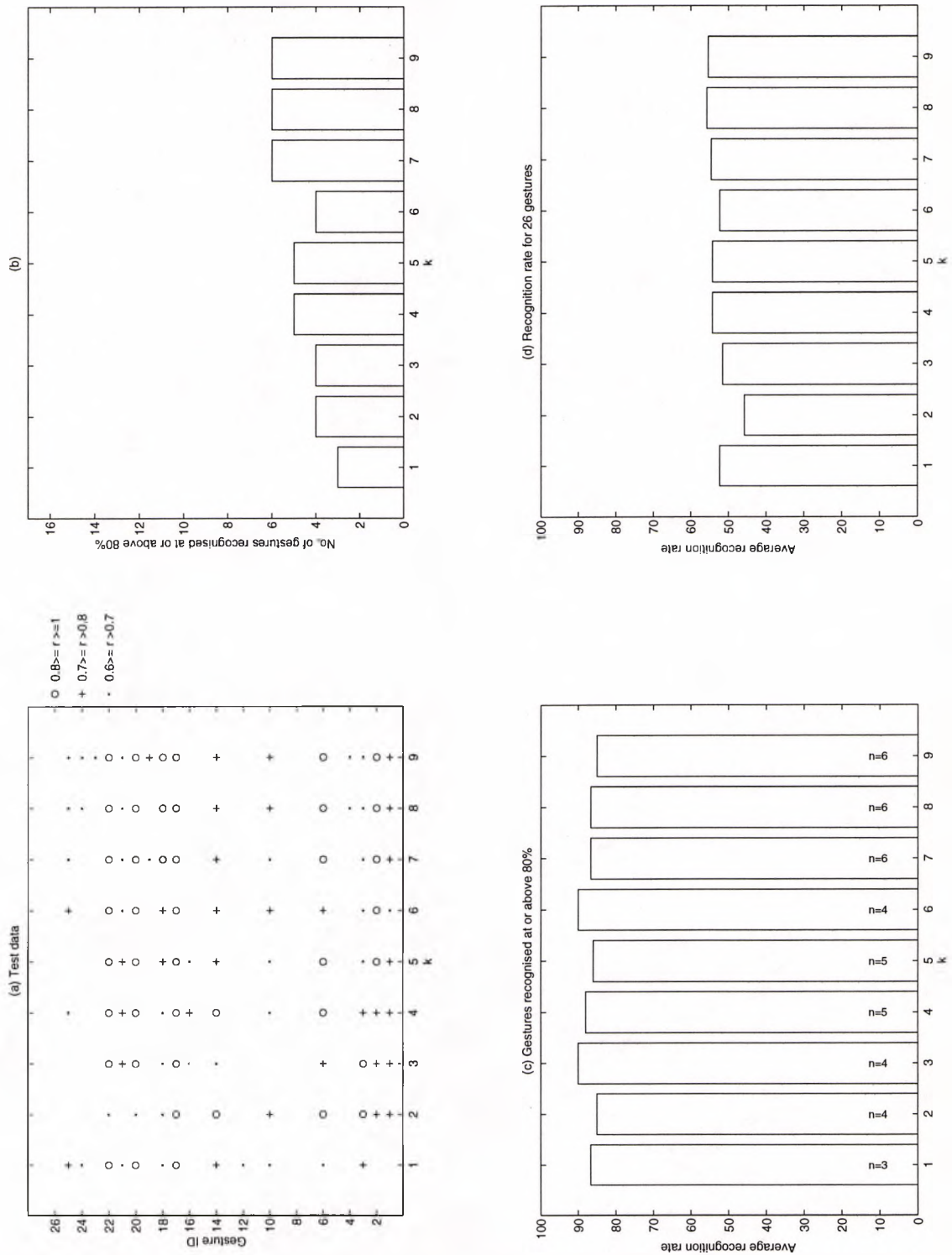


Figure 7.12 Gesture recognition results for Euclidean distance k-nearest neighbour classifier with $k=1$ to 9 using *trte2i14p* feature set. (a) shows which gestures were recognised at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 \leq r < 0.8$ are marked with '+'. Gestures recognised $0.6 \leq r < 0.7$ are marked with '.'. Histogram (b) shows the maximum number of gestures recognised at or above 80% . (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognised at or above 80%).

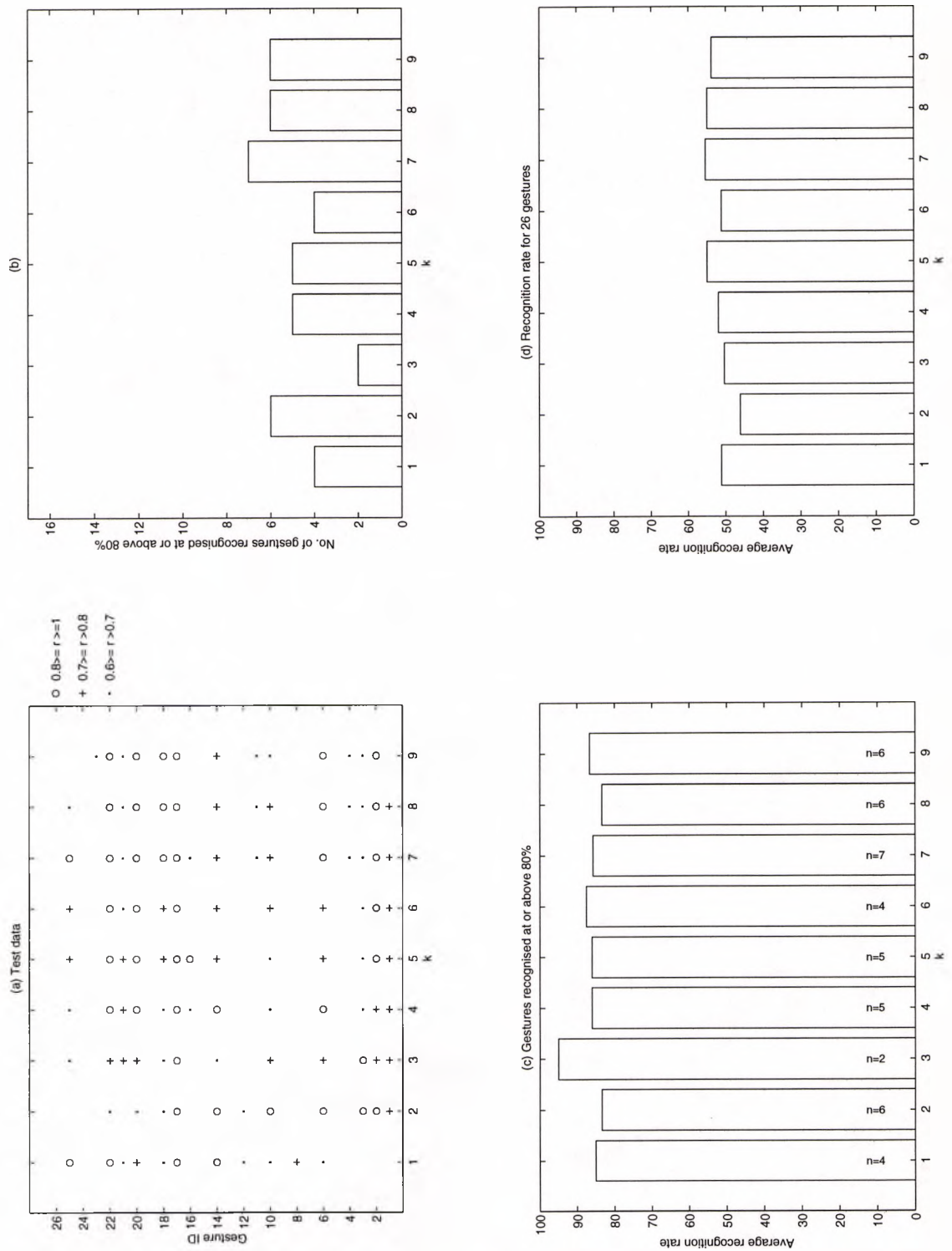


Figure 7.13 Gesture recognition results for Euclidean distance k-nearest neighbour classifier with standard normalisation with $k=1$ to 9 using `trte2i14p` feature set. (a) shows which gestures were recognised at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 > r < 0.8$ are marked with '+'. Gestures recognised $0.6 > r > 0.7$ are marked with ':). Histogram (b) shows the maximum number of gestures recognised at $r \geq 0.8$. (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognised at or above 80%).

Classifier	$n_{r_i \geq 80}$	$\bar{r}_{r_i \geq 80}$ (x100%)	\bar{r}_G (x100%)
Euclidean k-nearest neighbours k=8	6	86.67	55.77
Euclidean k-nearest neighbours with standard normalisation k=7	7	85.71	55.38
Two layer feedforward neural network 12-16t-26p	12	90.00	50.00

Table 7.4: Summary results showing best recognition rates based on $\max(n_{r_i \geq 80})$ for k-nearest neighbours and comparison with best results using ANNs using the tr/te2i14p feature set (four frames of xyz position data)

7.9 Summary and Discussion

Feedforward neural networks with a single hidden layer were trained using backpropagation of errors with momentum (set at 0.95) and adaptive learning rate (set initially to 0.01). Weights connecting the input layer to the hidden layer were initialised using the Nguyen-Widrow method the input-hidden layer. Weights connecting the hidden layer to the output layer were initialised randomly between 1 and -1.

Average recognition rate alone is not a good indicator of practical classifier performance as there is a level below which it would be too frustrating to contemplate including the gesture in the gesture set. Classifier performance was measured by the number of gestures recognised at or above the 80% level plus the average recognition rate for that set of gestures.

The manually segmented gestural set of 624 gestures comprised 26 gestural classes. The gesture set was divided randomly into a training set (364) and a test set (260) for training and validating the neural network classifiers.

The number of hidden layers needed to solve the problem was determined experimentally. A feature set comprising four time frames of x,y,z arm position data (12 features) re-sampled at intervals of 160ms was used to train and test the ANNs involved in this process (feature set tr/te2i14p described in section 8.2). Tan-sigmoid activation functions were used in the hidden layer. The identity function was used in the output layer. Networks with up to 20 nodes were compared. 16 nodes in the hidden layer yielded the optimum number of gestures recognised at or above 80%. This result was used to determine the number of hidden nodes for most neural network experiments. The assumption was made that this result will be valid for other input feature sets and other neuronal activation functions.

A network architecture of 12-4t-26p with four hidden nodes recognised 6 gestures at or above 80%. It took the addition of another 12 to recognise a further 5 gestures.

The effect of choice of neuronal activation was investigated by constructing ANNs with

various combinations of tan-sigmoid, log-sigmoid, and identity function in the hidden and output layers. Substantial variation in recognition rates was found dependent on the choice of activation function. Tan-sigmoid in the hidden layer and identity function in the output layer gave the best results (12 gestures). Log-sigmoid in both layers gave the worst result (3 gestures). Whether this is a local minima problem or another phenomena is not known. In light of these results, all other ANN experiments used tan-sigmoid functions in the hidden layer and the identity function in the output layer.

Two k-nearest neighbour classifiers (kNN) were constructed using MATLAB. These were a) euclidean distance kNN and b) euclidean distance with standard normalisation. The best result obtained using kNN was 7 gestures recognised at or above 80% with an average recognition rate of 85.7%. This was considerably lower than that achieved using ANNs (12 gestures recognised at or above 80% with an average recognition rate of 90%).

Chapter 8

Feature Set Comparison using Neural Networks

8.1 Introduction

This chapter documents the part of the study that compared the performance of feedforward neural networks presented with a variety of feature sets. The feature sets were grouped into four categories: gesture segment length, forearm orientation, scalar and vector velocity, curvature of plane of motion. Results are presented in the form of summary figures and tables. The corresponding confusion matrices can be found in appendix C and the Hinton diagrams in appendix D.

8.2 Gesture Segment Length

8.2.1 Method

Training and test feature sets containing 3D position data representing varying gesture segment lengths were compared. Raw position data sampled at 100 s^{-1} was low-pass filtered and resampled at 6.25 s^{-1} or every 16 samples as described in section 6.7 and the data segmented into the 7 feature sets (see table 8.1) containing 624 exemplars. These data were further randomly divided into a training set of 324 exemplars (prefix tr) and a test data set of 240 exemplars (prefix te) and used as input data to train and test the neural network (see section 7.4).

Feature Set ID	Points in 3Space	ANN architecture	Equivalent Time Window (ms)	No. of Weights and Biases
tr/te2i1p	1	3-16t-26p	160	506
tr/te2i12p	2	6-16t-26p	320	554
tr/te2i13p	3	9-16t-26p	480	602
tr/te2i14p	4	12-16t-26p	640	650
tr/te2i15p	5	15-16t-26p	800	698
tr/te2i16p	6	18-16t-26p	960	746
tr/te2i17p	7	21-16t-26p	1120	794

Table 8.1: Feature sets of increasing sample size representing increasing gesture segment length

Each feature set was used to train and test a neural network with architecture x-16t-26p. Each network was trained for a total of 10000 epochs and the network was tested every 100 epochs, and the results and weights stored.

8.2.2 Results and Discussion

Results are presented in the form of confusion matrices (CFM) together with summary statistics in section 5 of appendix C. These represent the ANN decision based on the output node with maximum activation. ANN decisions for training and test data sets are presented for gesture segment lengths (GSL) 160ms to 1120ms. The accompanying tables and figures summarise the data from these CFMs.

The histograms in figure 8.1 show the recognition rates for all 26 gestures. The highest recognition rate of 51.54% for a gesture segment length of 800ms is rather poor. However, from 7 to 12 gestures were recognised with $r \geq 0.8$ dependant on GSL, representing a good result given that the composition of the gesture set was not chosen or optimised based on ease of human visual recognition.

In general, the difference between the average recognition rate for test data and training data increases with GSL. Factors that are likely to contribute to an increasing generalisation error are:

- i) Finite data set size. As the dimension of the input space increases, the data points become more disperse. The decision surfaces are of increasing dimension, but are fitted to the same number of points in a hyperspace of increasing dimension.
- 2) Variance increases with each time frame. The first point was manually segmented using human recognition, thus this point can be expected to have low variance relative to subsequent time frames. Variation in the speed and duration of a gesture will have an increasing effect on subsequent time frames.

Considering only gestures recognized with $R \geq 0.8$, figure 8.2 (b) shows that seven gestures were recognised well with $r \geq 0.8$. This increases with GSL until a maximum is reached at four points in space equivalent to a time-window of 640ms and an input-space of twelve dimensions. With this gesture segment length, a further five gestures have been recognized making a total of twelve gestures. Histogram (d) plots the number of epochs that yielded the maximum number of $r \geq 0.8$ gestures. For gesture segment length of 160ms the network trained in only 300 epochs, while it took 6700 epochs to train the neural network to recognise twelve gestures. Note, however that with 5 time-frames (800ms) the network trained in 2700 epochs. This, coupled with a decreased overall recognition rate (figure 8.1) for both training and test data, suggests that the ANN converged towards to a local minimum. In addition, only ten rather than twelve gestures were recognised. These results exhibit the typical variability that might be expected with gradient search optimisation with randomly chosen initial conditions (Fausett, 1994). Histogram (c) shows that the average recognition rate of the set of gestures recognized at $r \geq 0.8$ is between 85% and 90% for GSL of 160ms to 1120ms.

The Hinton diagrams for the associated networks are presented in appendix D figures D.1 to D.3 (a), (b), (c), and (d) and figure D.4 (a) and (b). The size of the rectangles in these diagrams are proportional to the magnitude of the corresponding weight. (a) and (c) represent weights $W1$ between the input layer and the hidden layer and (b) and (d) represent weights $W2$ between the hidden layer and the output layer.

Looking at $W1$ for each network, the relative strength of connections associated with each input is fairly evenly distributed across the input dimensions. This indicates that the network has used the information present at each input node in some way in the recognition process.

Looking at $W2$ for each network, the networks for four, six, and seven frames of xyz position seem to have a small number of weights with high magnitude with most weights much lower, while networks for one, two, three, and five have more weights with relatively large activations. In the former case, some of the hidden neurons have one weight connected to them that are much stronger than all the others. This suggests that the activation level of those particular hidden neurons plays a major part in the activation of the output/gesture classes on the other end of the respective weights. Thus the network to some degree has tended to associate individual hidden nodes with individual gesture classes.

In table 8.2 gestures were grouped to show the minimum gesture segment length (GSL) presented to the ANN that yielded $r_i \geq 0.8$ together with those gestures that were not recognised at this level using position co-ordinates in 3-space as features. GSL of greater than 800ms failed to recognise any new gestures. Also, although with a GSL of 800ms the 'Spider' gesture was recognised with $r \geq 0.8$, the number of gestures recognised at this level fell.

These results show that by using a fixed time window approach it is possible to classify CP arm gestures at a level potentially usable for HMI. Segment length is an important parameter in the resultant recognition rate. The results are consistent with the notion that for a particular set of gestures there is an optimum window length when using the time-delay scheme of presenting data to a ANN. For this set of CP gestures, a time window of 640ms results in the greatest number of gestures recognised at $r \geq 0.8$. For all 26 gestures, the 960ms GSL gave marginally higher results. Given the small number of

exemplars of each gesture, the results need to be interpreted with some caution, but clearly a time window of around 600 to 960ms is appropriate for dynamic arm gestures, at least for this individual.

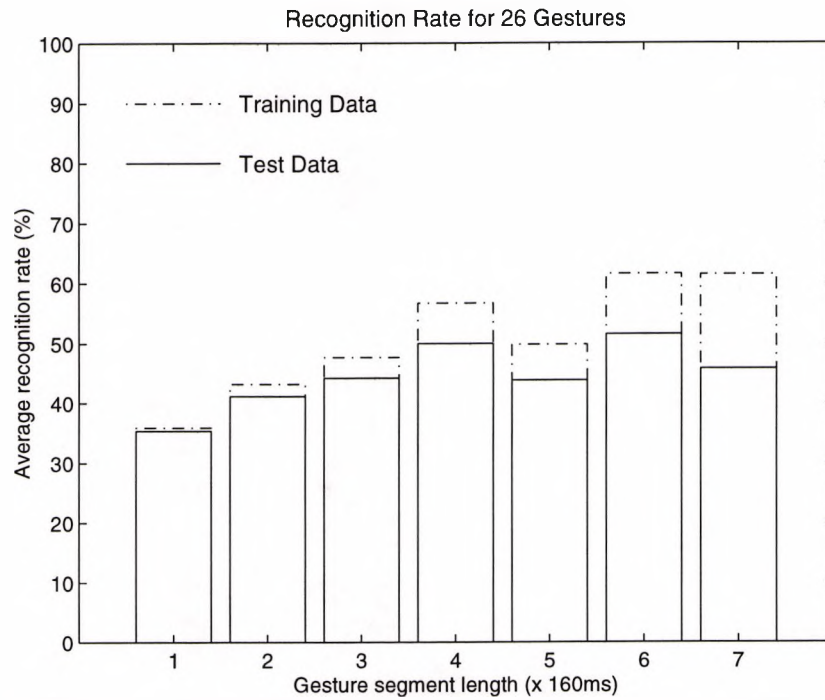


Figure 8.1 Average recognition rate for 26 gestures for feature sets involving gesture segment length

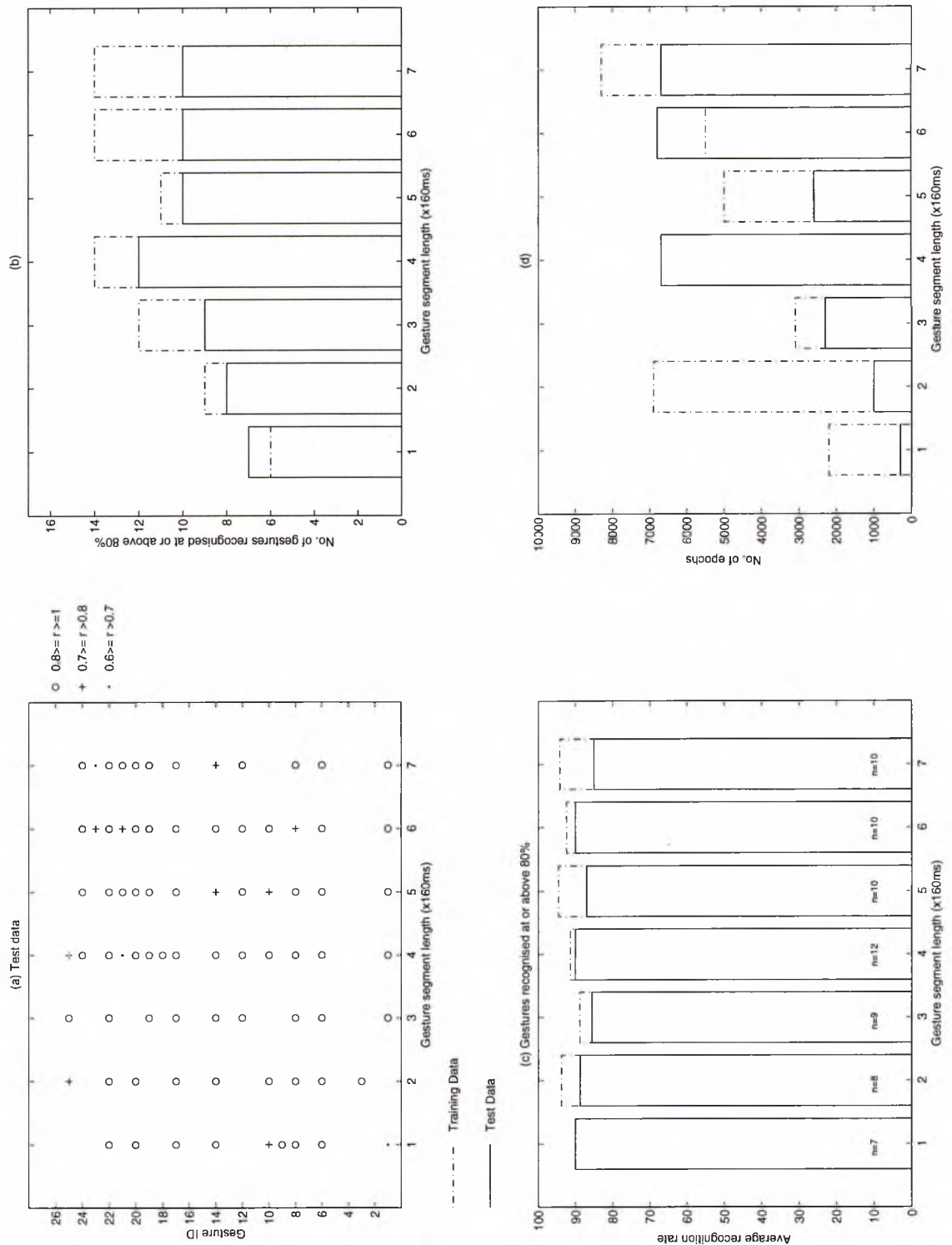


Figure 8.2 Recognition results of gesture segments from 160ms to 1120ms for gestures recognized at or above 80%. (a) shows which gestures were recognized at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 \geq r < 0.8$ are marked with '+'. Gestures recognized $0.6 \geq r < 0.7$ are marked with '.'. Histogram (b) shows the maximum number of gestures recognized at $r \geq 0.8$. (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognized at or above 80%).

GSL	Code	Gesture
160ms	6	heavy weight
	8	hot
	9	ice-cream
	14	rainbow
	17	scratch your knee
	20	spank
	22	stroke the cat
320ms	3	cut throat
	10	ironing
480ms	1	bird
	12	lasso
	19	shave
	25	umbrella
640ms	18	shake hands
	24	whistle
800ms	21	spider
Not Recognised ($r < 0.8$)	2	cards
	4	drive the car
	5	drums
	7	helicopter
	11	knock on the door
	13	light feather
	15	rock a baby
	16	rock guitar
	23	surrender
26	violin	

Table 8.2: Gestures grouped to show the minimum gesture segment length (GSL) needed to recognise each gesture at or above 80%.

It was useful to inspect the confusion matrix results for a GSL of 640ms (appendix C.5)

and to study which gestures the network confuses. Table 8.3 details gestures that have been recognised as other gestures with $r \geq 0.6$. Each one of these mistakes is a 'reasonable' error in that they take place in similar areas of space and/or have a similar morphology.

ANN decision	Actual gesture	r_m	r_i	Actual gesture	r_m	r_i
hot	cards	0.8	0	rock guitar	0.6	0.1
shave	cut-throat	1	0	ice-cream	0.7	0
(train) whistle	helicopter	0.6	0			
umbrella	surrender	0.6	0			

Table 8.3: Gestures misrecognised at or greater than 60% for GSL of 640ms. r_m is the proportion misrecognised as the gesture in the first column. r_i is the recognition rate for the mis-recognised gesture

The ability of the network to learn the "easier to recognise" gestures in the presence of more poorly formed gestures and similarly formed gestures is potentially very useful, in that it could provide feedback to the clinician and/or therapist to aid in the development of gestural repertoires containing gestures that maximise the overall recognition rate.

These results highlight the fact that gestures with very different meaning can appear similar in form. When this occurs the gestural HMI could use additional contextual information to determine which gesture was intended e.g. disambiguation using the knowledge of application program state or previous gestures of a gestural sequence.

8.3 Forearm Orientation

8.3.1 Method

In section 8.3, the use of successive time frames of xyz position as a feature vector was shown to yield a best result of 12 gestures recognised at or greater than 80%, with an average of 90% for those 12 gestures. The average recognition rate for all 26 gestures was only 50%. In order to explore whether other feature vectors exist that are capable of improving the recognition rate, a number of feature sets were derived from the position and combined in various ways to create input feature vectors.

The FFNNs were trained in an identical fashion to the previous studies, training the network by backpropagation of the sum-squared error for 10,000 epochs, testing and storing the results every 100 epochs, then searching for the best results.

Feature sets were constructed following the definitions in table 8.4 in order to explore the effect of adding forearm orientation to the feature vector.

Ex45 to ex48 and ex51 to ex54 all used a FFNN architecture of xx-16t-26p. In the case of the three feature sets of 24 dimensions, the experiments were repeated using an ANN architecture of more hidden nodes to account for the increased dimension of the feature

space (experiments ex55 to ex57).

Orientation information was extracted from the raw data from the "Flock of Birds" sensor by applying the rotation matrix transformation to calculate the elbow position (in 3-space) and also a point approximately radially distant from the elbow. Euler angles were not used because the resulting signal is discontinuous and the azimuth and roll become very noisy and exhibit large errors as the elevation approaches $\pm 90^\circ$.

By considering the wrist relative to the elbow, it is possible to simply derive a pure direction vector (as elbow length is fixed) that conveys four frames of forearm direction (tr/te2i4per). Wrist rotation is conveyed by deriving a vector from the radial point relative to the elbow (tr/te2i4par).

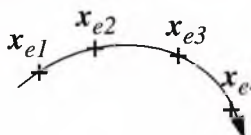
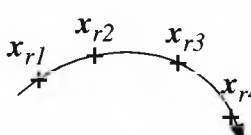
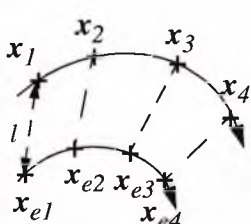
Experim. ID/ Feature Set ID	Dimension of Feature Vector	No. of Weights and Biases	Description
Ex45 tr/te2i14e	12	650	 <p>x_{e1}, \dots, x_{e4} Four frames of elbow position</p> <p>$p = \{x_{e1}, x_{e2}, x_{e3}, x_{e4}\}$</p>
The approximate elbow position was calculated from the wrist position and the rotation matrix assuming the distance from the distal forearm sensor to the elbow was 9.5 inches (no allowance was made for sensor offset),			
Ex46 tr/te2i14a	12	650	 <p>x_{r1}, \dots, x_{r4} Four frames of radial position</p> <p>$p = \{x_{r1}, x_{r2}, x_{r3}, x_{r4}\}$</p>
Radial position is defined as a point on an approximate forearm radius from the elbow. This exhibits large changes with wrist rotation.			
Ex58,59 tr/te2i14pe	24	842	 <p>x_1, \dots, x_4 Four frames of wrist position</p> <p>x_{e1}, \dots, x_{e4} Four frames of elbow position</p> <p>Forearm length $l = 9.5$ inches</p> <p>$p = \{x_1, x_2, x_3, x_4, x_{e1}, x_{e2}, x_{e3}, x_{e4}\}$</p>
Four frames of wrist position were combined with elbow position. (The forearm assumed to be a rigid body)			

Table 8.4: Description of feature vectors p involving forearm orientation and the associated feedforward neural network

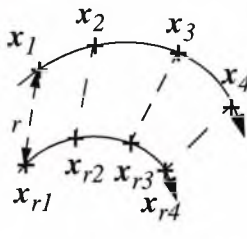
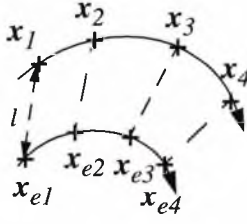
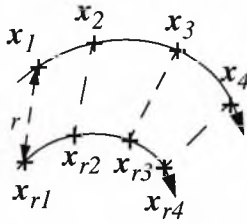
Experim. ID/ Feature Set ID	Dimension of Feature Vector	No. of Weights and Biases	Description
Ex48 tr/te2i14pa	24	842	 <p>x_1, \dots, x_4 Four frames of wrist position</p> <p>x_{r1}, \dots, x_{r4} Four frames of radial position</p> <p>$r = 13.8$ inches</p> <p>$p = \{x_1, x_2, x_3, x_4, x_{r1}, x_{r2}, x_{r3}, x_{r4}\}$</p>
Ex51 tr/te2i14er	12	650	 <p>$x_{re1} = x_1 - x_{e1}$</p> <p>$x_{re2} = x_2 - x_{e2}$</p> <p>$x_{re3} = x_3 - x_{e3}$</p> <p>$x_{re4} = x_4 - x_{e4}$</p> <p>Forearm length $l = 9.5$ inches</p> <p>$p = \{x_{re1}, x_{re2}, x_{re3}, x_{re4}\}$</p>
Ex52 tr/te2i14ar	12	650	 <p>$x_{rr1} = x_{r1} - x_1$</p> <p>$x_{rr2} = x_{r2} - x_2$</p> <p>$x_{rr3} = x_{r3} - x_3$</p> <p>$x_{rr4} = x_{r4} - x_4$</p> <p>$r = 13.8$ inches</p> <p>$p = \{x_{rr1}, x_{rr2}, x_{rr3}, x_{rr4}\}$</p>

Table 8.4: Description of feature vectors p involving forearm orientation and the associated feedforward neural network

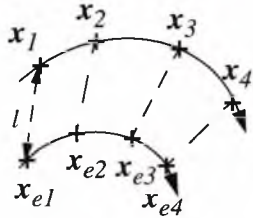
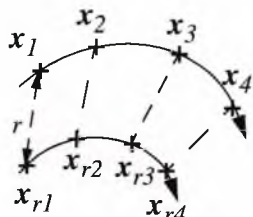
Experm. ID/ Feature Set ID	Dimension of Feature Vector	No. of Weights and Biases	Description
Ex53,56 tr/te2i4per	24	842	 $\begin{aligned}x_{re1} &= x_1 - x_{e1} \\x_{re2} &= x_2 - x_{e2} \\x_{re3} &= x_3 - x_{e3} \\x_{re4} &= x_4 - x_{e4}\end{aligned}$ <p>Forearm length $l = 9.5$ inches</p> $p = \{x_1, x_2, x_3, x_4, x_{re1}, x_{re2}, x_{re3}, x_{re4}\}$
Ex54,57 tr/te2i4par	24	842	 $\begin{aligned}x_{rr1} &= x_1 - x_{r1} \\x_{rr2} &= x_2 - x_{r2} \\x_{rr3} &= x_3 - x_{r3} \\x_{rr4} &= x_4 - x_{r4}\end{aligned}$ <p>$r = 13.8$ inches</p> $p = \{x_1, x_2, x_3, x_4, x_{rr1}, x_{rr2}, x_{rr3}, x_{rr4}\}$

Table 8.4: Description of feature vectors p involving forearm orientation and the associated feedforward neural network

8.3.2 Results

The average recognition rates for 26 gestures are documented in figure 8.3. These were extracted from the confusion matrices detailed in appendix C section 6. The best average recognition rates were achieved by tr/te2i4per (four frames of wrist position plus four frames of wrist direction) at 52.69%. This was closely followed by tr/te2i4pa - four frames of wrist position plus four frames of radial position, and tr/te2i4pe - four frames of wrist position plus four frames of elbow position. Both yielded a recognition rate of 51.15%. The recognition results for gestures recognised at or above 80% are presented in figure 8.4. The number of gestures recognized at or above 80% is detailed in plot (b). tr/te2i4pa, four frames of position plus four frames of radial position, and tr/te2i4per, four frames of wrist position plus four frames of wrist direction produced the highest figure of 12 gestures. Although in the case of the average recognition rate for 26 gestures the tr/te2i4pe gesture set was the second highest, only six gestures were recognised at or above 80%.

The Hinton diagrams associated with these data are presented in appendix D, figure D.10 (c) and (d) and figures D.11 to D.15 (a), (b), (c), (d). Inspecting the weights W_1 , between the input and hidden layers, each input feature had weights of strength comparable to

other input features. This suggested that all features were involved in the pattern classification process. This was the case for all feature sets in involving forearm orientation.

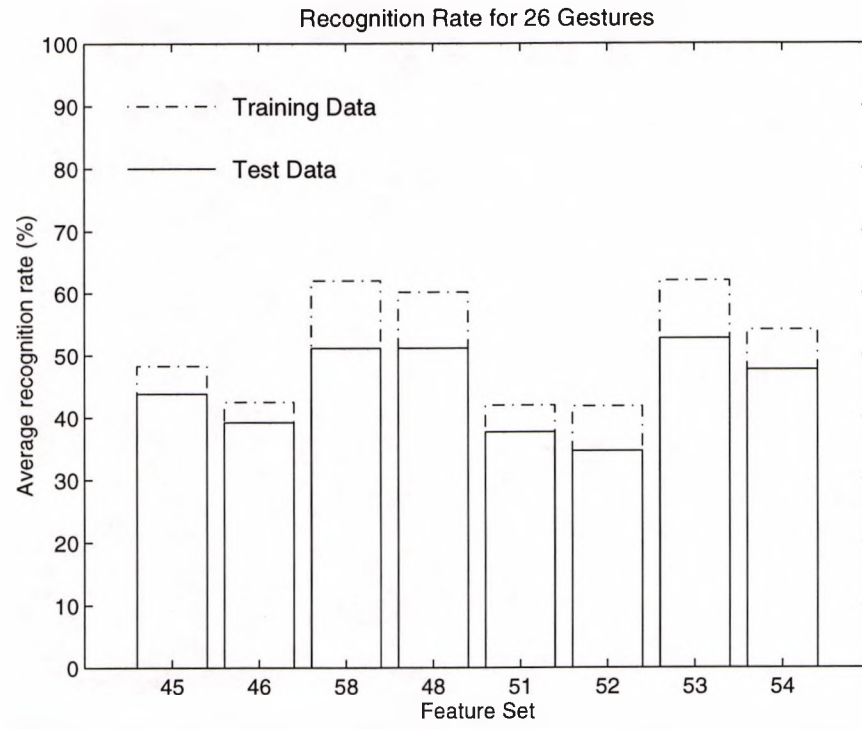


Figure 8.3 Recognition rate for 26 gestures for feature sets involving forearm orientation

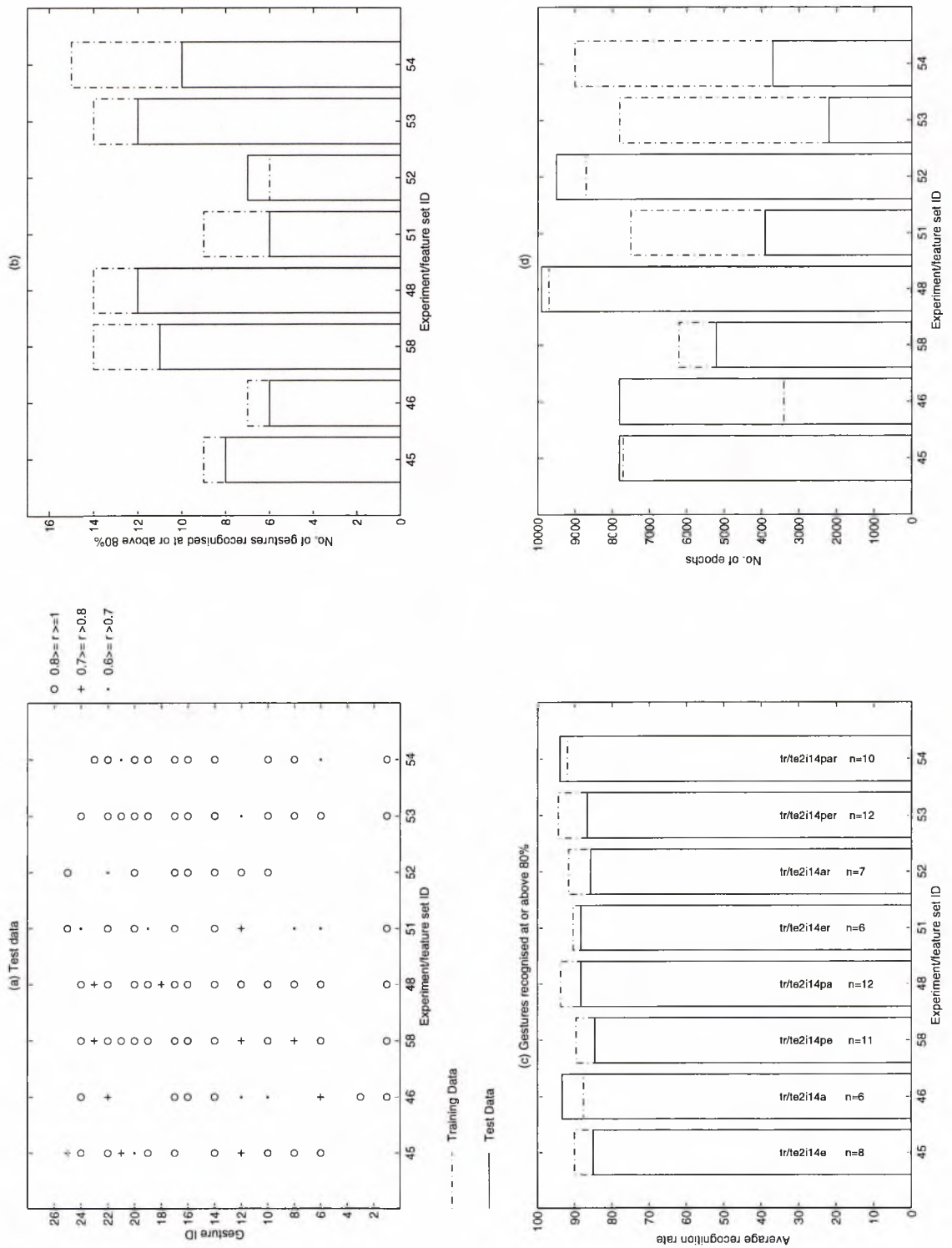


Figure 8.4 Recognition results feature sets containing forearm orientation information for gestures recognized at or above 80%. (a) shows which gestures were recognized at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 \leq r < 0.8$ are marked with '+'. Gestures recognized $0.6 \leq r < 0.7$ are marked with '•'. Histogram (b) shows the maximum number of gestures recognized at $r \geq 0.8$. (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognized at or above 80%).

In the case of three of the feature sets with 24 dimensions (tr/te2i4per, tr/te2i4par, tr/te2i4pe), experiments were conducted training networks of architecture 24-20t-26p (twenty hidden nodes). The results are compared in figures 8.5 and 8.6. The histogram in figure 8.5 details the average recognition rate for 26 gestures. In the case of tr/te2i4pe, 16 hidden nodes performed slightly better than twenty. With tr/te2i4per and tr/te2i4par, the situation is reversed. In fact, the use of the tr/te2i4per with a 24-20t-26p architecture gives the highest recognition rate for 26 gestures of 58.85%.

The recognition results for gestures recognised at or above 80% is described in figure 8.6. Looking at the number of gestures recognised at or above the 80% level, 20 hidden nodes results in worse performance than 16 nodes for tr/te2i4pe and tr/te2i4per. Performance is the same for tr/te2i4par.

Although the average recognition rate did increase in one case by increasing the number of nodes in the hidden layer, this performance was not reflected in the number of gestures recognised at or above 80%. Thus, overall FFNNs with 16 hidden nodes performed better than FFNNs with 20 nodes.

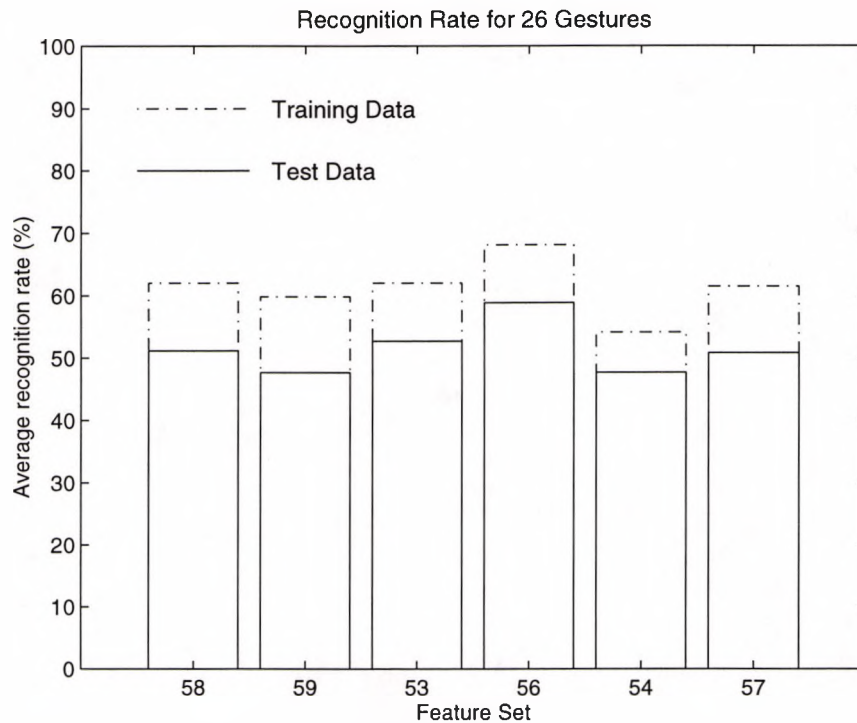


Figure 8.5 Average recognition rate for 26 gestures. Comparison of results for networks of 16 and 20 nodes in the hidden layer for feature sets containing forearm orientation.

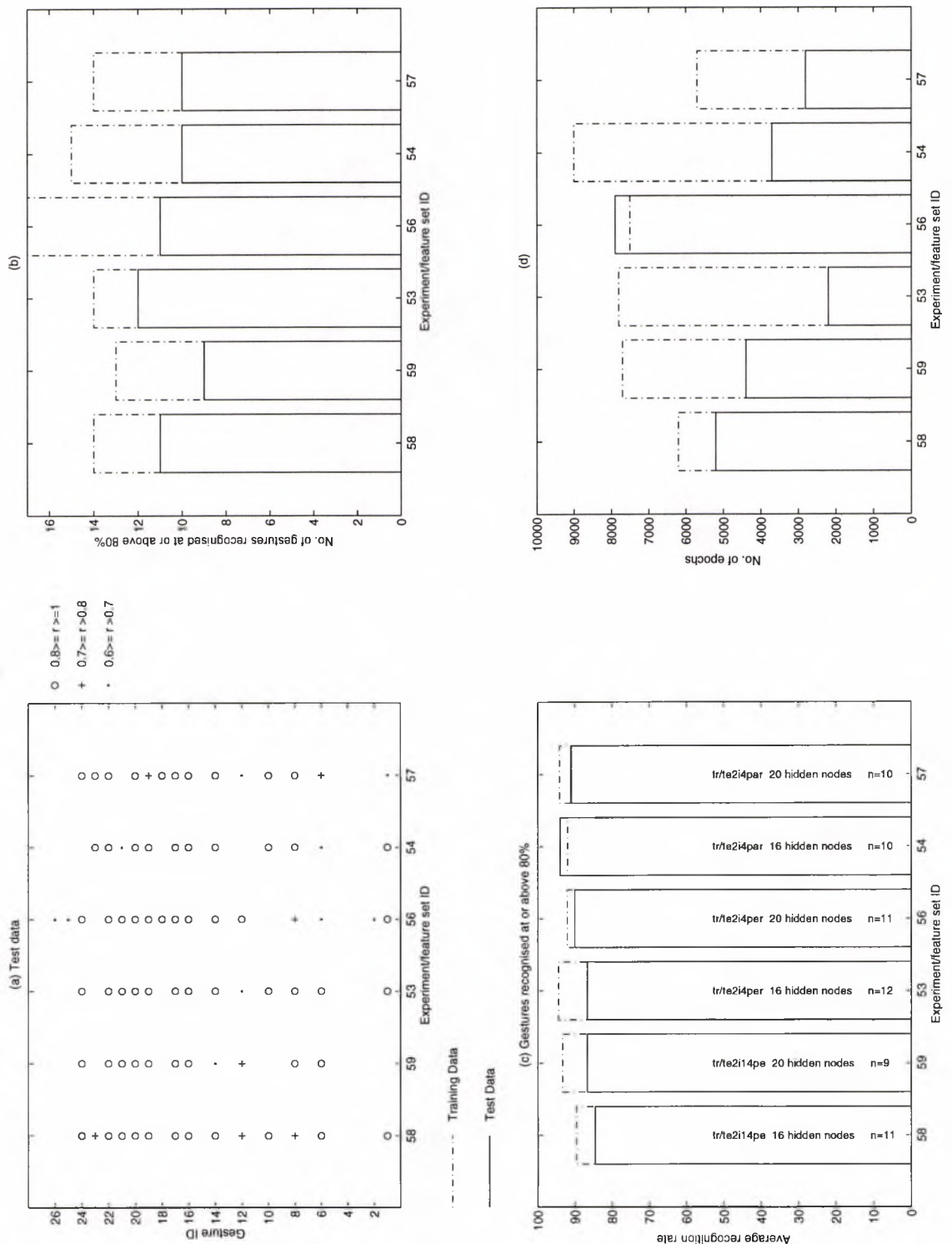


Figure 8.6 Comparison between 16 and 20 hidden nodes ANN architectures for feature sets containing forearm orientation information. (a) shows which gestures were recognised at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 \leq r < 0.8$ are marked with '+'. Gestures recognised $0.6 \leq r < 0.7$ are marked with '.'. Histogram (b) shows the maximum number of gestures recognised at or above 80%. (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognized at or above 80%).

8.4 Scalar and Vector Velocity

8.4.1 Method

Feature sets were created that contained qualities of vector velocity (tr/te2i1p3r, tr/te2i4r, tr/te2i4p1r,) and scalar velocity (tr/te2i4s, tr/te2i4p3s, tr/te2i4ps, and tr/te2i4psa). These are described in table 8.4. Since position was sampled at regular intervals every 1/16 sec., the vector and scalar distance between samples represented a measure of vector and scalar velocity respectively.

The feature sets were used to train FFNNs of architecture xx-16t-26p in the previously described manner.

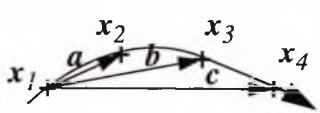
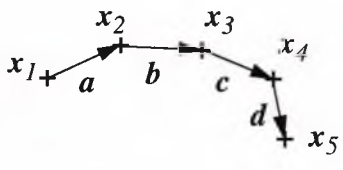
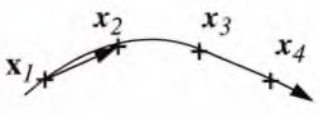
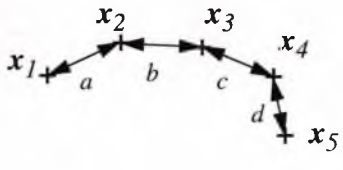
Feature Set ID	Dimension of Feature Vector	No. of Weights and Biases	Description
ex38 tr/te2i1p3r	12	650	 $a = x_2 - x_1$ $b = x_3 - x_1$ $c = x_4 - x_1$ $p = \{x_1, a, b, c\}$
xyz position of first point plus the position of the next four points relative to the first.			
ex39 tr/te2i4r	12	650	 $a = x_2 - x_1$ $b = x_3 - x_2$ $c = x_4 - x_3$ $d = x_5 - x_4$ $p = \{x_1, a, b, c\}$
xyz position of first point plus relative position between four consecutive pairs of points.			
ex40 tr/te2i4p1r	15	698	$p = \{x_1, x_2, x_3, x_4, x_2 - x_1\}$ 
Four frames of position plus the position of the second point relative to the first.			
ex41 tr/te2i4s	4	522	 $a = x_2 - x_1 $ $b = x_3 - x_2 $ $c = x_4 - x_3 $ $d = x_5 - x_4 $ $p = \{a, b, c, d\}$
The scalar distance between four consecutive pairs of points.			

Table 8.5: Description of feature vectors p involving scalar and vector velocity and the associated feedforward neural network

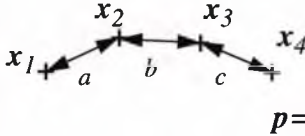
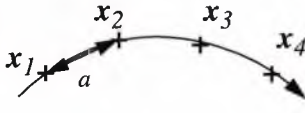
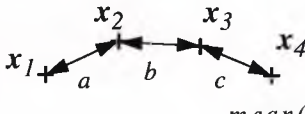
Feature Set ID	Dimension of Feature Vector	No. of Weights and Biases	Description
ex42 tr/te2i4p3s	15	698	 $a = x_2 - x_1 $ $b = x_3 - x_2 $ $c = x_4 - x_3 $ $p = \{x_1, x_2, x_3, x_4, a, b, c\}$
ex43 tr/te2i4ps	13	666	 $a = x_2 - x_1 $ $p = \{x_1, x_2, x_3, x_4, a\}$
ex44 tr/te2i4psa	13	666	 $a = x_2 - x_1 $ $b = x_3 - x_2 $ $c = x_4 - x_3 $ $\text{mean}(a, b, c) = (a + b + c) / 3$ $p = \{x_1, x_2, x_3, x_4, \text{mean}(a, b, c)\}$
Four frames of position plus the distance between neighbouring points.			
Four frames of xyz data plus the scalar distance between the first two points			
Four frames of position plus the mean scalar distance between the four points.			

Table 8.5: Description of feature vectors p involving scalar and vector velocity and the associated feedforward neural network

8.4.2 Results and Discussion

The histogram in figure 8.7 summarises the average recognition rate for 26 gestures. The highest recognition rate for 26 gestures of 55% was obtained using the tr/te2i4ps feature set. The tr/te2i4s, the scalar distance between four consecutive pairs of points, resulted in the lowest recognition rate of 14.62%.

The recognition results for gestures recognised at or above 80% are presented in figure 8.8. The largest number of gestures recognized at or above 80% was 11. This was produced by the tr/te2i4ps and tr/te2i4psa feature sets.

Again, tr/te2i4s performed poorly resulting in the recognition of only three gestures. However, it was very interesting that there was sufficient information in four frames of scalar velocity (i.e. no position or direction information) to recognise three gestures "ironing", "shave", and "spank" with an average recognition rate of 87%.

The associated Hinton diagrams are presented in appendix D, section D.7 to D.9 (a), (b), (c), (d), and D.10 (a), (b). With 1p3r, the strength of the weights connected to the first xyz point are somewhat stronger than the other relative direction features. However this feature set recognized 11 gestures as compared to 7 gestures in the case of tr/te2i1p, a single frame of xyz position. With 4p1r, four frames of position and one frame of relative

position, the three relative input dimensions gave rise to somewhat stronger weights.

In the case of 4p3s, four frames of xyz position plus the scalar distance between them, strong weights were associated with the three scalar dimensions, indicating that significant use was being made of this information in the attempt to classify the gestures.

Again in the case of 4ps, four frames of xyz position plus one feature of scalar velocity, the strongest weights were associated with scalar velocity.

Finally, in the case of 4psa, four frames of position, and one feature of mean scalar velocity, the weight strengths associated with the scalar feature are significantly stronger than the other twelve features.

This suggests that these networks have extracted information in a significantly different way from the 14p, or four frames of xyz position. However, in none of the cases where scalar or vector velocity was added to four frames of position (4p1r, 4p3s, 4ps or 4psa) were there any associated increase in recognition rate or in the number of gestures recognised at or above 80%.

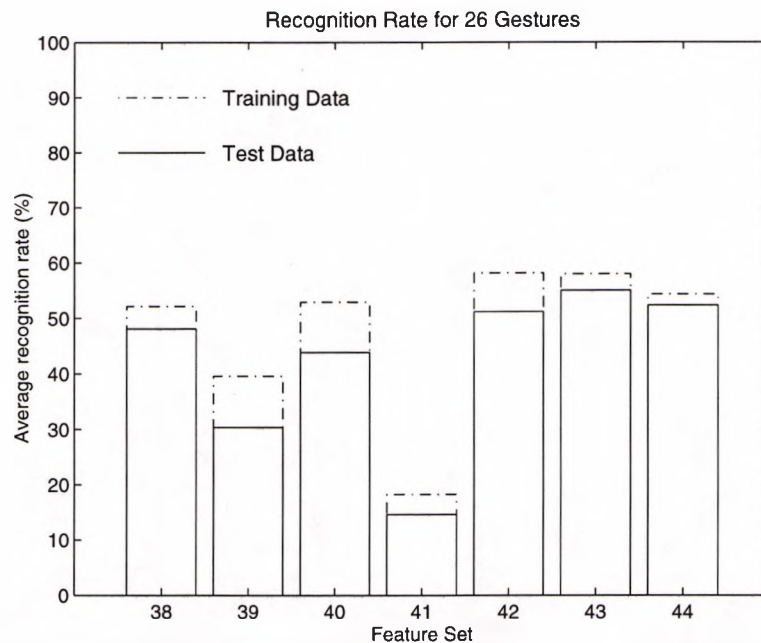


Figure 8.7 Recognition rate for 26 gestures for feature sets involving scalar and vector velocity

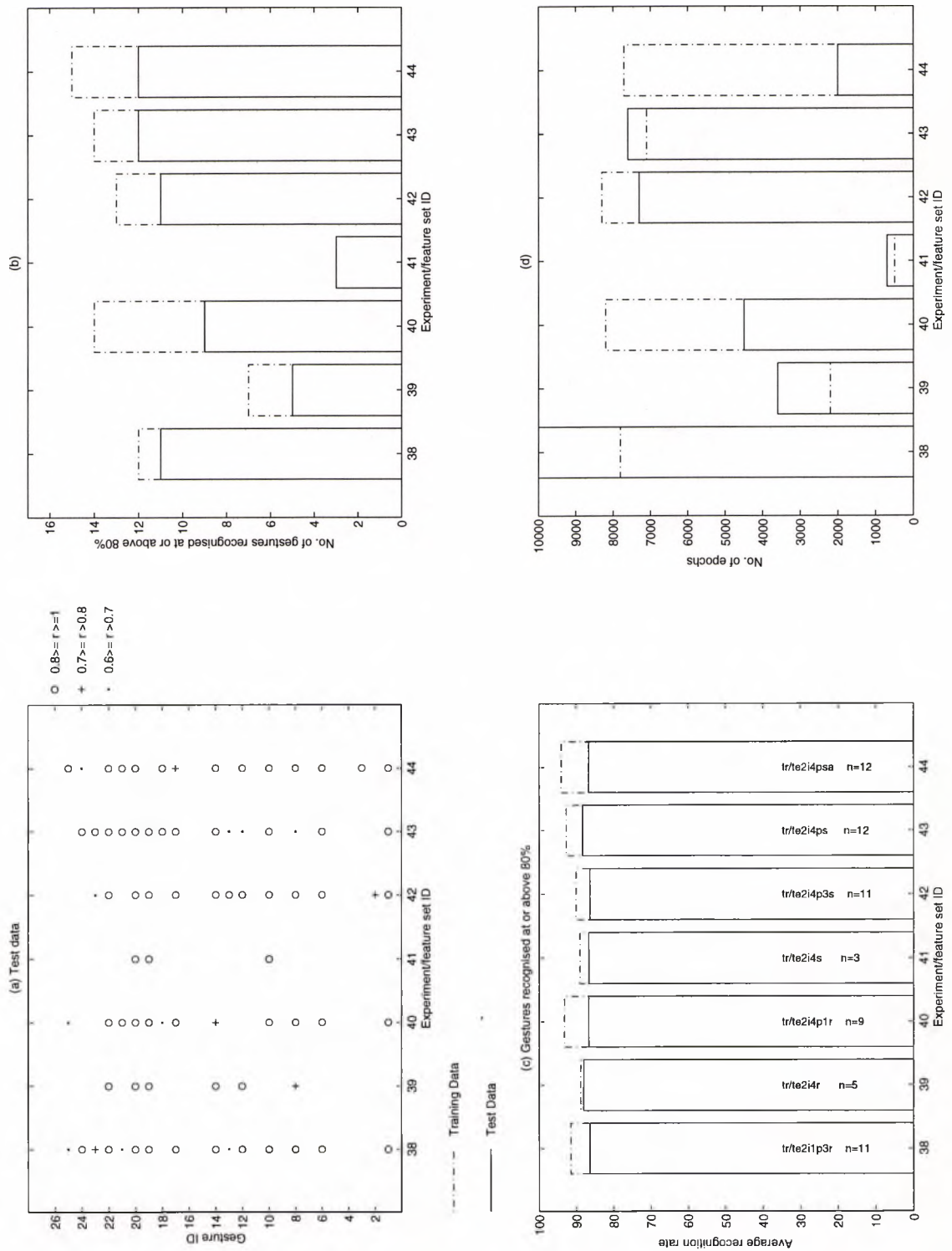


Figure 8.8 Recognition results feature sets containing scalar and vector velocity information for gestures recognized at or above 80%. (a) shows which gestures were recognized at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 \leq r < 0.8$ are marked with '+'. Gestures recognized $0.6 \leq r < 0.7$ are marked with '.'. Histogram (b) shows the maximum number of gestures recognized at $r \geq 0.8$. (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognized at or above 80%).

8.5 Curvature and Plane of Motion

8.5.1 Method

Feature sets were constructed that contained features involving the curvature of the gestural path and also the plane of motion. The scalar product was used to create features with qualities of path curvature in the form of the cosine of the angle between the two direction vectors formed by the three points. The vector product was used to create features which encoded the plane of motion. Each feature set is described in table 8.4.

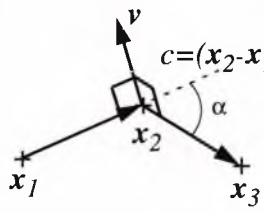
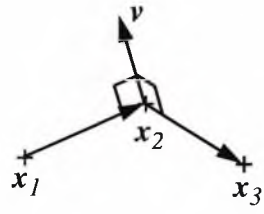
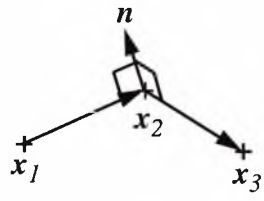
Feature Set ID	Dimension of Feature Vector	No. of Weights and Biases	Description
ex33 tr/te2i13vc	13	666	 $v = (x_2 - x_1) \wedge (x_3 - x_2)$ $c = \frac{(x_2 - x_1) \cdot (x_3 - x_2)}{ x_2 - x_1 x_3 - x_2 }$ $p = \{x_1, x_2, x_3, v, c\}$
ex34 tr/te2i13v	12	650	 $v = (x_2 - x_1) \wedge (x_3 - x_2)$ $p = \{x_1, x_2, x_3, v\}$
ex35 tr/te2i13n	12	650	 $v = (x_2 - x_1) \wedge (x_3 - x_2)$ $n = v / v $ $p = \{x_1, x_2, x_3, n\}$

Table 8.6: Description of feature vectors p involving curvature and plane of motion and the associated feedforward neural network

Feature Set ID	Dimension of Feature Vector	No. of Weights and Biases	Description
ex36 tr/te2i13nc	13	666	$c = (x_2 - x_1) \cdot (x_3 - x_2) / x_2 - x_1 x_3 - x_2 $ $v = (x_2 - x_1) \wedge (x_3 - x_2)$ $n = v / v $ $p = \{x_1, x_2, x_3, n, c\}$
ex37 tr/te2i13c	10	618	$c = (x_2 - x_1) \cdot (x_3 - x_2) / x_2 - x_1 x_3 - x_2 $ $p = \{x_1, x_2, x_3, c\}$

Table 8.6: Description of feature vectors p involving curvature and plane of motion and the associated feedforward neural network

8.5.2 Results

The histogram in figure 8.9 summarises the average recognition rates for 26 gestures. Recognition rates were all between 46% and 52%. The highest recognition rate for 26 gestures was produced by 13n, three frames of position plus normalised vector product, at 52.31%.

The recognition results for gestures recognised at or above 80% are summarised in figure 8.10. The number of gestures recognised at or above 80% spans from 8 to 11. The maximum number of gestures at 11 was achieved by 13c, three frames of position plus the cosine of angle α .

The Hinton diagrams are presented in appendix D, figures D.4 (c), (d) and D.5 to D.6 (a), (b), (c), (d). Looking at the network trained on the 13vc feature set (three frames of position plus vector product, plus cosine), The weights associated with the vector product features are considerably stronger than the position or cosine features. All but one of the weights connected to the cosine feature are zero. This suggests that the vector product information has played a dominant part in the gesture classification process. The situation is similar for the 13v (three frames of position plus vector product) case.

However, in the case of the 13n feature set (three frames of position plus normalised vector product), the weights connected to the normalised vector product features are extremely small in comparison to the three frames of xyz position. Since the normalised vector product represents the plane of motion, in the network trained using the 13n feature set, plane of motion has not played such an important part in network decisions as the position information.

In the case of the 13nc feature set (three frames of position plus normalised vector product plus cosine) and the 13c feature set (three frames of position plus cosine), the weights connected to the normalised vector product and the cosine features are small in comparison to the weights connected to position features.

This suggests that the vector product is a potentially useful feature, but not for the way in which it encodes plane of movement as might be expected, but possibly more for the way it encodes velocity in a particular plane.

Networks associated with 13vc, 13v both used the vector product information in preference to the position information. In comparison with 13p, three frames of position, (see section 8.2.2) the addition of vector product (13v) increased the number of gestures recognised at or above 80% by one from 9 to 10. Interestingly, so did the addition of normalised vector product (13n). Possibly even more surprising is that adding the cosine of angle α increased the number of gestures from 9 to 11. The addition of normalised vector product plus cosine in the case of 13nc seems to have had the effect of reducing the number of gestures recognised at or above 80% by one.

Whether this a real consequence of adding the feature or an artifact due to the training of the neural network (local minima effects) is difficult to ascertain.

Although three feature sets apparently improved the number of gestures recognised at or above 80%, none offered any improvement over the 14p feature set.

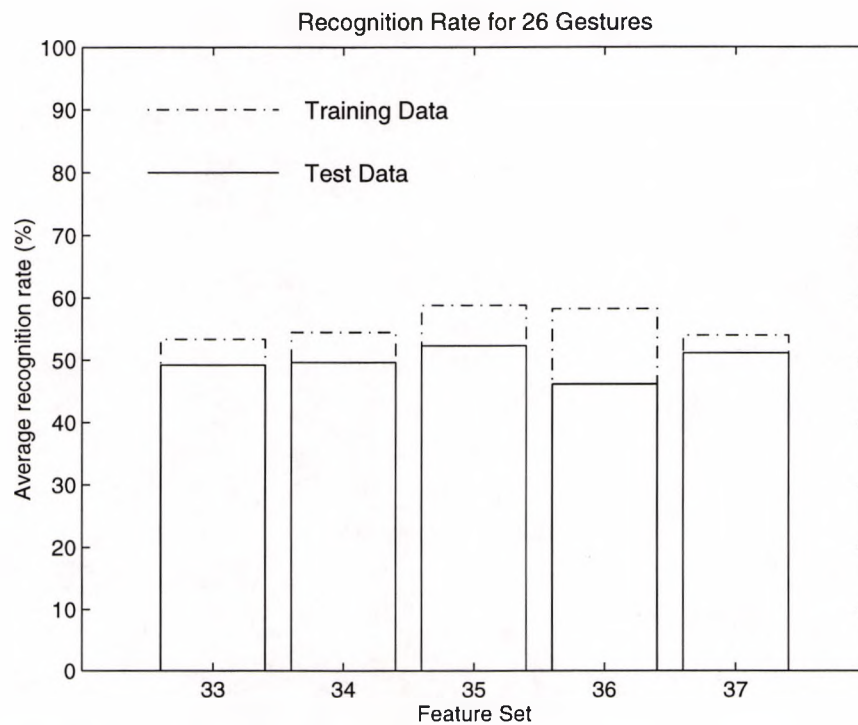


Figure 8.9 Recognition rate for 26 gestures using curvature and plane of motion features

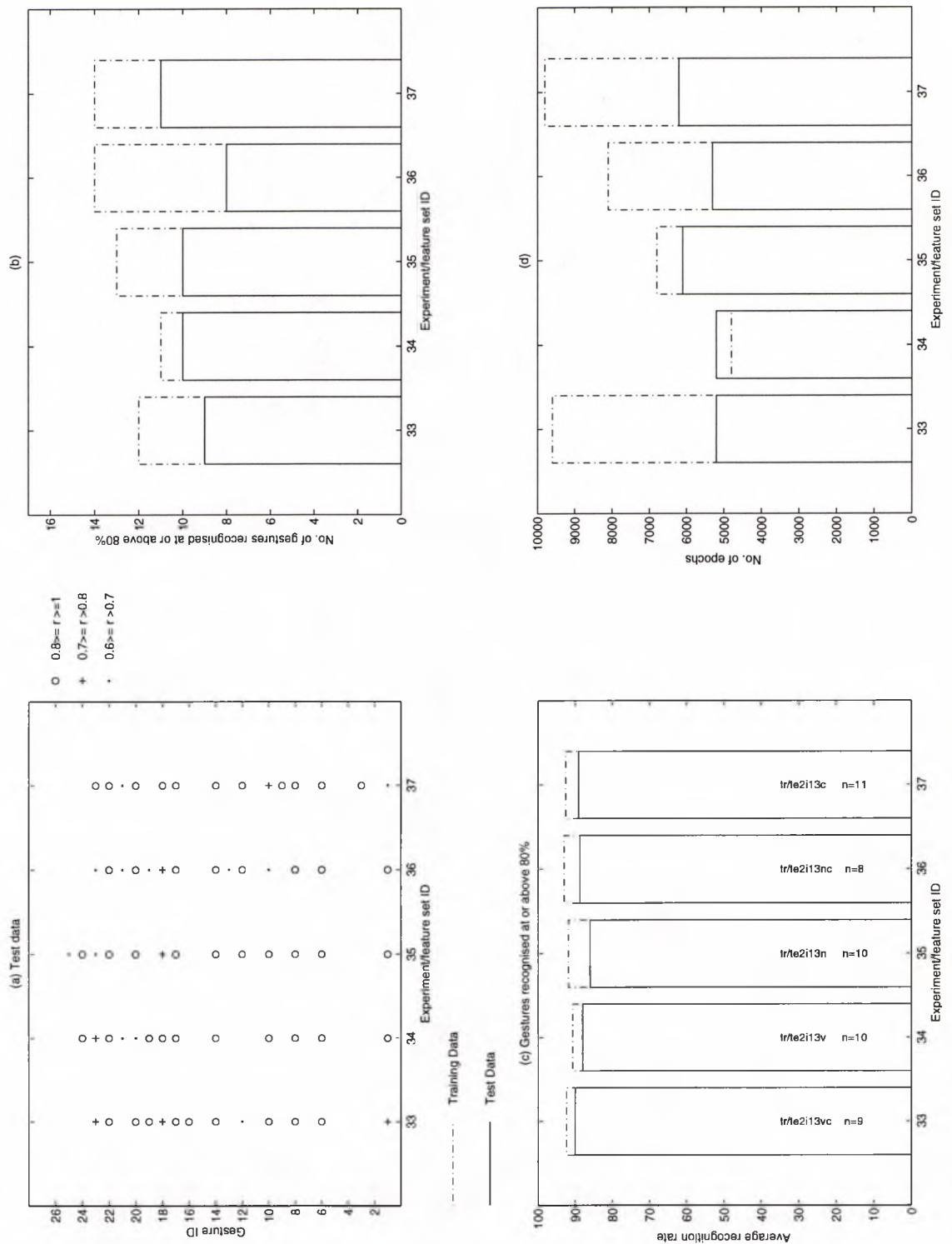


Figure 8.10 Recognition results feature sets containing curvature and plane of motion information for gestures recognized at or above 80%. (a) shows which gestures were recognized at this level $r \geq 0.8$. These are marked with an 'o'. For comparison, gestures with $0.7 \geq r < 0.8$ are marked with '+'. Gestures recognized $0.6 \geq r > 0.7$ are marked with '·'. Histogram (b) shows the maximum number of gestures recognized at $r \geq 0.8$. (c) shows the average recognition rate (%) at the respective gesture sub-set. (d) indicates the amount of training (epochs) the network required to reach maximum performance (defined by the maximum number of gestures recognized at or above 80%).

Examining the confusion matrices of the 13c and 13p networks (appendix C, and respectively) table 8.7 was constructed. The table examines three gestures that occur in a similar region of space, “cut throat”, “ice-cream” and “shave”. In the case of the network associated with 13p, 90% of test exemplars of “shave” are recognised correctly. 90% of “ice-cream” exemplars are recognised as “shave”. Only 40% for the gesture “cut-throat” are recognised correctly, the other 60% being recognised as “shave”.

In the case of the network associated with 13c, all test exemplars of “cut throat” are recognised correctly. 80% of “ice-cream” test exemplars are recognized correctly. However “shave” is never recognised correctly, being recognised either as “cut throat” (80%) or “ice-cream” (20%).

The average recognition rate of the three gestures in the case of 13p is 43% and in the case of 13c is 60%. From the table it is clear that the decision surfaces of these two networks differ considerably.

Actual gesture	Network Decision (%) te2i13c			Network Decision (%) te2i13p		
	cut throat	ice-cream	shave	cut throat	ice-cream	shave
cut throat	100			40		60
ice-cream		80				90
shave	80	20		10		90

Table 8.7: Confusion matrix of three visually similar arm gestures recognised by using tr/te2i13c, three frames of xyz position plus dot product, 9-16t-26p architecture (extracted from appendix C.8 ex37uk and C.5 GSL 480ms)

8.6 Summary and Discussion

These results show that by using a fixed time window approach it is possible to classify CP arm gestures at a level potentially usable for HMI. The GSL feature set study demonstrated how segment length is an important parameter in the resultant recognition rate. For this set of CP gestures, a time window of 640ms results in the greatest number of gestures recognised at or above 80%. For all 26 gestures, the 960ms window gave very marginally higher results. Given the small number of exemplars of each gesture, the results need to be interpreted with some caution, but clearly for this individual, a time window of around 600 to 960ms is appropriate for dynamic arm gestures.

Table 8.8 summaries the recognition results documented in this chapter. The table lists the feature set which yielded the highest number of gestures recognised at or above 80% for each feature set category together with the results for 13p.

From the table, it can be seen that 14p, 14per, 14pa, 4ps recognised the same number of gestures at or above 80% at similar average recognition rates. Thus, in this study, the network associated with four frames of position marginally exhibits the best performance with the test data.

It is quite likely that certain feature combinations offer advantages that have not become apparent in this study. A much larger set of gestural exemplars would have been useful in enabling feature sets to be compared with higher resolution. This would have enabled the recognition rates to be validated with more certainty.

This study has shown that FFNNs can classify gestures at a potentially useful level, using a variety of feature sets and with quite different internal representations. This insensitivity to gestural input features is encouraging as this property is an important element in the construction of a robust gesture recognition system.

Feature Set	$n_{r_i \geq 80}$	$\bar{r}_{r_i \geq 80}$	\bar{r}_G
tr/te2i14p Four frames of xyz position data over 640ms	12	90.00	50.00
tr/te2i14per Four frames of position and four frames wrist relative to the elbow	12	86.67	52.69
tr/te2i14pa Four frames of position and four frames radial point	12	88.33	51.15
tr/te2i4ps Four frames of xyz data plus the scalar distance between the first two points	12	88.33	55.00
tr/te2i13c	11	89.09	51.15
tr/te2i13p	9	85.56	44.24

Table 8.8: Summary results showing best recognition rates based on $\max(n_{r_i \geq 80})$

Chapter 9

Conclusions

9.1 Summary

Chapter 2

Human factors and technological issues of gestural human-machine interaction are closely linked. An argument is made for the need to give equal emphasis to both aspects. A human-factors driven approach was adopted on the basis that human issues should shape technological development. A conceptualisation of HMI was proposed where the human, machine, and environment are considered as a system of dynamically interacting non-linear sub-systems. Such systems are known to exhibit emergent properties. It was proposed that human machine interaction can be considered as an emergent property where new expressive human behaviours may emerge as a result of the interaction. However, it was also proposed that expressive behaviour can be critically dependent on the environment and components within that environment, particularly if physical impairment restricts self-adaptive ability. The research aims of the thesis were related to a conceptualisation where a Venn diagram represented three components of human behaviour: "elicitable", "observed", and "useful for HMI". The aim of the human factors part of the study was to increase the intersection between "observed" and "elicitable" by eliciting new behaviours that were potentially useful for human-machine interaction. In contrast, the aim of machine recognition part of the study was to increase the intersection between "observed" and "useful for HMI" by developing automatic gesture recognition algorithms. The relevance of this theoretical framework to human-machine interaction involving people with severe speech and motor impairment is discussed.

The second half of this chapter discussed issues relating to the machine perception of human behaviour. It is suggested that for high bandwidth efferent HMI to maximally

harness the expressive behaviour of people with SSMICP, it will be necessary to harness multiple signals from the body. An architecture for a complete gesture recognition system involving multiple sensors is proposed. However, the work in this thesis was restricted to the investigation of issues relating to the pattern recognition engine.

Chapter 3

This chapter focused on defining the target users of the proposed gestural HMI and examining human issues that impact on HMI design. The complex nature of cerebral palsy and its effect on communicative interaction are described and discussed. Existing HMI for this population is described including HMI used for augmentative and alternative communication. Key factors considered to be of particular relevance to gestural HMI include: each user is likely to exhibit a highly individual profile of physical and cognitive abilities; movement characteristics can change as a child's neurology develops; physical performance can change as a result of fatigue, emotional state or illness; infantile reflexes often persist into adulthood.

Chapter 4

The theoretical framework established in chapter 2 together with the user issues described and discussed in chapter 3 led to the development of a methodology to elicit candidate behaviour for gestural HMI involving human-human interaction. This methodology was applied in the design of a gestural ability pilot study. The study investigated the gestural ability of twelve children and young adults between the ages of 5 to 18 years. A cognitive framework was constructed using performance arts techniques to elicit a wide variety of volitional expressive behaviour. Subjects readily produced a repertoire of gestural movements far greater than that anticipated from the study of their therapeutic, medical and educational records. Video-tape recordings of the gestural repertoire were analysed in respect to the body parts involved to produce each gesture. Histograms of body part involvement showed that gestures were produced using movement of the head, arms, torso, leg, facial expression, eye-gaze, and vocalisation. Substantial individual variation is apparent. However, the arm was the most frequently used articulator for nine out of ten subjects.

Chapter 5

The utility of the gestures elicited in the gestural ability pilot study for HMI depend on whether they can be produced consistently and whether they can be reliably recognised by computer. Five subjects were chosen for instrumented data collection. A sub-set of the gestures was selected for each subject. It was hypothesised that sufficient pertinent information could be transduced from such gestures suitable for gesture recognition. The six-degree of freedom magnetic tracker was attached distally to one forearm. The arm was selected as it was found to be significantly involved in gestural expression (chapter 4). Data was collected in three sessions over three days, each session lasted 50-60 minutes with 33-43 minutes of data collected divided into a number of shorter data collection periods separated by breaks of 1-5 minutes. The process of manual segmentation of the gestural data stream needed to create training and test data for the neural networks was labour intensive (chapter 6). As a result, data from just one subject was used. Results illustrated that the gestures were co-articulated in so far as they were

produced without requesting the subject to rest between gestures or asking the subject to move to a certain position. Close examination indicated that movement from the previous gesture was often still present for a short period after the beginning of the next verbal prompt. The gestures were elicited in random order, the transition from one gesture to the next could follow a wide range of paths and distances. The average rate of production was one gesture per 9.5 seconds or 6.3 per minute. This rate was maintained over three or four ten minute periods with only a few minutes rest in between during each session. These results indicate that it is possible for a person with spastic-athetoid quadriplegia to produce gestures at a sustained rate for a considerable length of time at a rate similar to the rate of selection for indirect scanning.

Chapter 6

A computer graphics animation tool was developed to "play-back" and examine the dynamic arm gesture data collected using the magnetic tracker. This examination revealed that most gestures fell into one of three categories: static, single movement and periodic movement. In addition, single movement gestures fitted within a time window of around one second. The duration of static gestures was at least one second long. The periodic gestures all had a periodic length of less than one second. A computer recognition scheme was proposed that involved the use of time-delay feedforward neural networks. The scheme would involve continually extracting features from a finite time window of around one second past which all the gestural data flowed. The network would be trained on manually segmented data. The final system would use the neural network output neuron activation levels as a measure of confidence that a particular gesture is being produced at an instant in time. A further algorithm would be used to examine activation level and activation level duration to make the final classification and hence reject network decisions where the confidence level was low or duration too short. While the implementation of this scheme was left for further work, the scheme provided the rationale for the investigation of recognition algorithm performance using manually segmented test data.

Gestural data was animated and viewed after low pass filtering with a cut-off of around 3Hz. Filtering resulted in removal of much of the "jerkiness" of the movement due to cerebral palsy, while preserving the overall form of the gesture. As a result, the gestural data was preprocessed using a low pass filter of 2.8125 Hz. This process also reduced high frequency sensor "glitches" that occurred occasionally during periods of rapid movement and enabled the data to be re-sampled at 3.125 Hz. This meant that 1120ms of gestural data could be represented by only seven samples in time.

Chapter 7

Data was divided randomly into training and test sets of gestural data exemplars. Feed forward neural networks with one hidden layer were trained using backpropagation with momentum and adaptive learning rate. The network complexity, i.e. the number of nodes required for the hidden layer, was investigated experimentally using a feature vector comprising four time samples (frames) of x,y,z re-sampled position data (12 features) as input (representing a time window of 640ms). The results showed that 7 gestures could be recognised using only four nodes in the hidden layer, while it took another 12 nodes to recognise an additional 5 gestures. Optimum recognition rates were obtained using 16 nodes in the hidden layer. The effect of varying the type of activation function used in the

hidden and output layers was examined. There was a surprisingly large difference between the use of tan-sigmoid and log-sigmoid activation functions. The reason for this was not determined. Tan-sigmoid activation functions in the hidden layer, and the identity function in the output layer yielded the highest recognition rates. The classification performance of two k nearest neighbour algorithms (kNN) (euclidean distance kNN and euclidean distance kNN with standard normalisation) was ascertained using the training set as prototypes. The kNN methods recognised 6 or 7 gestures compared to the best NN result of 12. However the recognition rates for all 26 gestures were similar (55%).

Chapter 8

Feature vectors comprising from one to seven time frames of x,y,z position data (3 to 21 input features) corresponding to time windows spanning from 160ms to 1120ms were compared. Optimum recognition was obtained for 4 time frames, equivalent to a time window of 640ms.

A further 20 feature vectors were created that included measures of forearm orientation (8), scalar and vector velocity (7), curvature and plane of motion (5). However, none performed as well as the feature vector containing 4 frames of x,y,z position data.

9.2 Interpretation and Implications of Research

The outcome of the gestural pilot study indicates that the subjects have expressive abilities above and beyond that which can be harnessed using existing technology. In addition, they have expressive ability above and beyond that which is regularly observed in every day settings. The application of techniques from the performing arts proved to be highly appropriate for engaging the imagination and eliciting a wide range of behaviours based on the subject's knowledge of the world. The use of "generative" methods were necessary to facilitate the emergence of the subjects' latent gestural ability. This outcome is consistent with the theoretical arguments posed in chapter 2.

The gesture recognition study suggests that a magnetic tracker attached distally to one forearm can transduce sufficient gestural movement to recognise gestures even when other parts of the body are involved. Results from the feature vector study in chapter 8 suggest that adding features containing forearm orientation decreased the overall recognition rate. This could presumably be due to the presence of a high level of variance in forearm orientation, an important result as it indicates that some components of movement may be so variable that the recognition rate may be reduced. It is common for people with SSMICP to have great difficulty pronating and supinating the hand and arm, so conceivably this lack of volitional control is reflected in these results. In principle, it should be possible to increase the range of expressive behaviours useful for human machine interaction by transducing more of the human body. However, if this is attempted without care, the process may be counterproductive as was found to be the case with forearm orientation.

A number of feature sets gave similar recognition results even though examination of the Hinton diagram indicated that in some instances they had based their decisions upon very different feature subsets. It is possible that these feature sets possessed similar properties as regards class separation and clustering. Another interpretation is that the neural network is exhibiting its ability as a non-linear classifier to construct feature

surfaces with a wide range of topologies. If the latter is the case, then the results are a demonstration of the flexibility of neural networks. This is a useful property in the case of a practical recognition system. The user, clinician, or therapist cannot be expected to derive the optimal set of features for a particular body site and gesture set.

However, the experiments that looked at the effects of the number of hidden nodes demonstrated one of the disadvantages of a neural architecture as simple as a fully connected feedforward neural network with a single hidden layer. While only four hidden nodes were required to recognize 7 gestures, another 12 nodes were needed to recognise 12 gestures. This increase in complexity makes it difficult for such a simple architecture to scale up. In order to recognize a large number of gestures e.g. 1000, it would be necessary to employ such a large network that it would be impractical to implement. However, gestural HMI for people with SSMICP is likely to involve a relatively small set of gestures that can be reliably distinguished from each other. In this situation, artificial neural networks in their current forms seem more favourable.

9.3 Future Work

Human Factors Issues

The gestural elicitation methodology was successful at eliciting the subjects' latent gestural ability. However, for these gestures to be useful for gestural HMI for AAC they have to be turned into a method of communication. This is an area that can be investigated further independent of technology using human-human interaction. The gestures elicited in this research are like many graphic symbol systems in that they relate to concrete notions that draw upon the user's knowledge of the world. Thus, it is fairly easy to envisage the adaptation of AAC techniques involving graphic symbols for gestural input. It is possible that gesture has a distinct advantage over graphic symbols in the ease with which they can be remembered due to kinaesthetic recall. This may be of particular help in the recall of sequences of symbols.

This study has only looked at the gestural abilities of a small number of subjects. More studies need to be carried out in order to learn more about the gestural abilities of people with severe speech and motor impairment due to cerebral palsy, particularly from a developmental perspective. Questions arise such as how do the findings in this study relate to the population as a whole? As the people with SSMICP have difficulty physically interacting with their environment, how has this gestural ability been acquired? What does an infant with SSMICP need to experience to acquire gestural ability, or is it innate? To what degree do the gestural abilities of people with SSMICP correspond to or differ from the general population? Answers to these basic developmental and cognitive questions are needed to see where gestural HMI can be integrated into early intervention programmes.

The instrumented gestural data collected in this study were collected from subjects who had no appreciable practice or training in the activity prior to this research study. Two questions arise. Does daily practice improve gestural ability? If gestural ability is developed as an infant, does it improve gestural ability physically and cognitively as an adult?

Technological Issues

The next stage in this research is to develop an algorithm that can recognise gestures by

passing over a continuous stream of data. If feedforward networks are used in the first stage then a second stage would be needed to interpret the changing output activation levels with respect to time. If such a system can be created with acceptable recognition rates then it should be fairly straightforward to implement in real-time on an 133 MHz Pentium PC. Once this stage is reached, it will be possible to evaluate the system through user trials.

As regards the interpretation of the thesis results, it would be useful to investigate the human recognition rate of the gestural repertoires. These data would be useful as a baseline to determine targets for automatic recognition. In particular, it would be useful to compare human recognition rates with the modest 50-55% automatic recognition rate achieved for all 26 gestures. It was not possible to conduct this study within the resources of the project.

The low-pass filter employed in this research project is likely to be far from optimal. Also, signal enhancement requirements are likely to vary for each individual. Different types of cerebral palsy are characterised by different movement characteristics, so each is likely to need processing differently. Thus, the development of a more optimal filter is a useful line of investigation that could lead to improved recognition performance. Self-adaptive filtering may be particularly useful in this respect.

There are other pattern recognition methods that are superior to FFNNs in the way that they model dynamic processes. In theory, these should be better able to deal with variance in time. Suitable neural methods include the recurrent neural network architecture in its various forms. Some of the most successful speech recognition algorithms combine the use of neural networks with hidden Markov model (HMM) methods. They are reported to offer the discriminating powers of ANNs with the temporal modelling abilities of HMMs (Mammone, 1994). A similar strategy may have value in gesture recognition.

As previously discussed in the thesis, for people with SSMICP, inhibition of movement can take as much effort as initiation. Algorithms that do not take this into account are likely to be of limited application for this population. As illustrated by the gestural repertoires elicited in this study, the salience of gesture can be in either movement, path, or static position. For this reason, segmentation through simple feature extraction followed by thresholding and the pattern recognition techniques that rely on such a strategy are rejected as a future path of investigation. "Segmentation by recognition" seems the line of investigation most likely to produce robust automatic recognition that can recognise a wide range of gesture categories. Although usually these algorithms are significantly more computationally intensive than other methods, they are rapidly becoming more feasible as more computational power becomes available for the interface.

Closely linked to the underlying processes of gesture production, another important area of investigation is the way in which gestures can be symbolically represented in a machine. This impacts on both the type of features that are extracted from raw movement data and at the higher level of gesture interpretation. A number of coding schemes have been developed for transcribing human movement but few have been investigated as to their value in gesture recognition. High level symbolic processing of gestural features is attractive as the gestural data is then in a form that can easily be manipulated by computer and combined with other information. However there are potential problems associated with the application of simple coding schemes to translate body cerebral palsy

movement to gesture symbols. The problem lies in their sensitivity to the "noise" in the movement signal. There are possibly two questions here. Can coding features be found that are relatively "noise" insensitive? Can "noisy" gestural data at the symbolic level be processed effectively?

Gestural Human-Machine Interaction

The area of gestural human-machine interaction is still in its infancy although interest in the area continues to grow rapidly. This research confirms the potential value of gestural human-machine to people with speech and motor impairment due to cerebral palsy. In addition, it has the potential to enhance the quality of interaction for all users of computer technology.

The frontiers of the problem are likely to expand as more and more human activities become computer mediated. More research is needed in both human factors and machine perception areas. Knowledge gained in the cognitive aspects of human gesture promises to offer insight into the creation of machines that can exploit such ability. In the endeavour to create machines that understand gesture, it is likely that we will also learn more about its production.

Appendix A

Gesture Elicitation Sessions: Transcripts

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Pull									x		x		x			
2	Yes																
3	Stop																
4	Hungry	x		x		x											
5	Bye								x								
6	No		x														
7	Pizza			x													
8	Sip Soda																
9	Ice-cream					x			x	x							
10	Fast Car																
11	Drive Bus								x	x							
12	Helicopter								x								
13	Train	x															
14	Aeroplane								x	x							
15	Spider									x							
16	Pig																
17	Alligator																
18	Caterpillar											x		x			
19	Lion	x															
20	Bird								x	x							
21	Butterfly								x	x							
22	Elephant	x							x								
23	Snake	x	x					x	x	x					x		x
24	Fish	x															x
25	Mickey Mouse		x							x		x		x			
26	Spank								x	x							
27	Knock								x								
28	Throw								x	x							
29	Scratch									x							
30	Jump																
31	Ballerina								x	x							
32	Open Box																
33	Open Door																
34	Door Bell								x								
35	Deal Cards																
36	Fishing	x	x		x				x								
37	Bowling								x	x					x	x	
38	Canoeing								x	x							
39	Swimming							x	x	x							
40	Baseball																
41	Basketball								x	x							
42	Cold							x		x							
43	Hot																
44	Drum																x
45	Rock Guitar							x		x		x					
46	Shake Hand									x							
47	Sad		x	x													
48	Kiss																
49	Angry		x	x													
50	Hug								x	x							

Table A.1: Transcript of gesture elicitation session: Subject S1

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
51	Barf (vomiting)																
52	Yummy							x		x		x		x			
53	Soft							x		x		x		x			
54	Listen	x															
55	Sticky															x	
56	Smell		x							x		x		x			
57	Ouch!	x															
58	Love				x												
59	Yuk!	x															
60	Stir																
61	Wash Face		x							x		x		x			
62	Cut throat																
63	Picture																
64	Salute		x							x		x		x			
65	Dig Hole															x	
66	Crawl																x
67	Jump																
68	Knitting																x
69	Toss Pancake																
70	Dance			x													x
71	Fly Kite				x				x	x	x	x	x	x			
72	Milk Cow	x															
73	Poison									x							
74	Naughty			x													
75	Witch			x		x	x										
76	Type				x				x	x							
77	Mosquito Bite																
78	Rip Paper			x													
79	Earthquake																
80	Explosion		x														
81	Witch																
82	Dragon																
83	Monster	x															x

Table A.1: Transcript of gesture elicitation session: Subject S1

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Yes				x												
2	No		x														
3	Bye															x	
4	Eat					x											
5	Hello					x											
6	Stop															x	
7	Foot-shake															x	
8	Ice-cream					x											
9	Sip Soda					x											
10	Pizza	x														x	
11	Cowboy															x	
12	Tall									x							
13	Baby				x												
14	Large														x	x	
15	Giant	x															
16	Bathroom				x												
17	Binocular				x												
18	Mountain															x	
19	Book				x												
20	Beard				x					x		x					
21	Money									x		x					
22	Waiter									x		x		x			
23	Umbrella															x	
24	Necklace				x												
25	Wave															x	
26	Cowboy									x		x		x			
27	Racing Car	x														x	
28	Car																
29	Aeroplane															x	
30	Helicopter															x	
31	Stripes															x	
32	Square															x	
33	Circle																
34	Triangle																
35	Take Picture				x												
36	Cut Throat																
37	Wash Face (of puppet)															x	
38	Stir															x	
39	Hammer															x	
40	Spank															x	
41	Salute															x	
42	Drums															x	
43	Guitar															x	
44	Trumpet															x	
45	Piano															x	
46	Violin															x	
47	Flute															x	
48	Throw Dice															x	
49	Bowling															x	
50	Fishing				x												

Table A.2: Transcript of gesture elicitation session: Subject S2

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
51	Deal Cards															x	
52	Swimming															x	
53	Make of Basket																
54	Grand Slam	x															
55	Touchdown															x	
56	Tennis																
57	Hockey															x	
58	Smell			x													
59	Sticky															x	
60	Listen				x												
61	Hot				x												
62	Cold																
63	Soft																
64	Bright Light				x												
65	Smooth															x	
66	Fish															x	x
67	Snake	x		x		x		x							x	x	
68	Elephant		x	x	x												
69	Butterfly														x	x	
70	Bird														x	x	
71	Lion			x		x											
72	Caterpillar															x	
73	Mickey Mouse																
74	Mouse	x															
75	Alligator					x											
76	Pig			x													
77	Spider					x										x	
78	Explosion				x												
79	Earthquake																x
80	Don't Know		x														
81	Thinking				x												
82	Tired	x				x											
83	Yummy	x															
84	Barf (vomiting)	x															
85	Hug	x								x							
86	Angry					x			x	x							
87	Sad			x	x												
88	Asleep		x														
89	Kiss																
90	Yuk!	x				x											
91	Love	x			x												
92	Ouch!	x															
93	Hungry					x											
94	Rainbow				x												
95	Rain																
96	Knock on Door														x	x	
97	Toss a Pancake				x												
98	Knit															x	
99	Crawl														x	x	
100	Throw															x	
101	Dance																

Table A.2: Transcript of gesture elicitation session: Subject S2

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
102	Scratch Nose															x	
103	Saw																
104	Fly Kite				x												
105	Jump	x															x
106	Door Bell															x	
107	Open Box															x	
108	Open Door															x	
109	Dig Hole														x	x	
110	Push Door Closed															x	
111	Pull Rope															x	
112	Shave																
113	Shampoo (puppet)															x	
114	Phone															x	
115	Smoke Cigar					x											
116	Steal															x	
117	Ironing															x	
118	Mosquito Bite	x		x												x	
119	Type															x	
120	Lick			x	x	x											
121	Naughty	x						x									
122	Milk Cow	x															
123	Poison	x	x	x		x											
124	Ghost																
125	Witch			x													
126	Dragon					x											
127	Monster			x		x											x

Table A.2: Transcript of gesture elicitation session: Subject S2

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Pull								x		x		x				
2	Yes	x			x												
3	Stop	x															
4	Hungry	x															
5	Bye								x	x							
6	No	x															
7	Pizza			x		x			x								
8	Sip Soda			x													
9	Ice-cream					x	x										
10	Fast Car	x							x		x						
11	Drive Bus								x		x		x				
12	Helicopter								x		x		x				
13	Train	x															
14	Aeroplane								x	x							
15	Spider								x		x		x				
16	Pig	x		x													
17	Alligator			x		x											
18	Caterpillar								x		x		x				
19	Lion			x													
20	Bird								x	x	x	x	x				
21	Butterfly								x	x	x	x	x				
22	Elephant		x						x								
23	Snake			x													
24	Fish			x		x											
25	Mickey Mouse		x						x		x		x				
26	Spank								x		x		x				
27	Knock								x		x		x				
28	Throw					x											
29	Scratch								x		x		x				
30	Jump																
31	Ballerina								x	x							
32	Open Box								x								
33	Open Door								x								
34	Door Bell								x		x		x				
35	Deal Cards				x				x		x		x				
36	Fishing					x											
37	Bowling								x								
38	Canoeing								x		x		x				
39	Swimming								x		x		x				
40	Baseball								x	x							
41	Basketball								x								
42	Cold																
43	Hot								x		x		x				
44	Drum																
45	Rock Guitar							x	x		x		x				
46	Shake Hand								x		x		x				
47	Sad			x													
48	Kiss			x		x											
49	Angry			x					x	x							
50	Hug								x	x							

Table A.3: Transcript of gesture elicitation session: Subject S3

Appendix A: Gesture Elicitation Sessions - Transcription

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
51	Barf (vomiting)			x		x											
52	Yummy	x															
53	Soft																
54	Listen				x												
55	Sticky								x		x		x				
56	Smell																
57	Ouch!	x															
58	Love			x													
59	Yuk!			x													
60	Stir																
61	Wash Face		x						x		x		x				
62	Cut throat																
63	Picture				x												
64	Salute																
65	Dig Hole																
66	Crawl																
67	Jump																x
68	Knitting								x	x	x	x	x	x			
69	Toss Pancake																
70	Dance																x
71	Fly Kite			x													
72	Milk Cow	x							x	x	x	x	x	x			
73	Poison	x	x	x		x			x								
74	Naughty			x													
75	Lick (chocolate)					x											
76	Type		x														
77	Mosquito Bite				x												
78	Rip Paper			x					x	x							
79	Earthquake																
80	Explosion																
81	Witch																
82	Dragon								x	x							x
83	Monster			x					x	x							

Table A.3: Transcript of gesture elicitation session: Subject S3

Appendix A: Gesture Elicitation Sessions - Transcription

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Yes		x														
2	No		x														
3	Bye								x								
4	Elephant	x	x						x								
5	Caterpillar	x							x		x		x				
6	Butterfly	x							x		x		x				
7	Bird								x								
8	Lion	x	x	x					x		x		x				
9	Ice-cream		x			x			x		x		x				
10	Yummy	x						x	x		x		x				
11	Explosion	x							x	x							
12	Fly Kite								x		x		x				
13	Catch								x		x		x				
14	Spank								x		x		x				
15	Swim		x					x	x	x							
16	Make a Basket (basketball)							x	x	x							
17	Gram Slam (baseball)								x		x		x				
18	Touchdown (American football)								x		x		x				
19	Fishing								x		x		x				
20	Pizza		x			x			x		x		x				
21	Big							x	x	x	x	x	x		x		
22	Short								x		x		x				
23	Tall								x		x		x				
24	Giant								x		x		x				
25	Cowboy (on a horse)	x															
26	Lasso	x							x								
27	Milk Cow								x								
28	Helicopter	x							x		x		x				
29	Train	x															
30	Racing Car								x		x		x				
31	Car	x							x		x		x				
32	Aeroplane							x	x		x		x				
33	Handshake		x	x													
34	Sleep		x														
35	Earthquake		x														
36	Guitar								x		x		x				
37	Drum		x						x		x		x				
38	Flute	x	x	x					x		x		x				
39	Sip Soda		x			x		x	x		x		x				
40	Bathroom		x		x				x		x		x				
41	Throw		x		x				x		x		x				
42	Itch							x	x		x		x				
43	Alligator	x		x					x		x		x				
44	Smoke Cigar		x			x			x		x		x				
45	Wash Face		x						x		x		x				
46	Handshake (with puppet)								x		x		x				
47	Stir Soup								x		x		x				
48	Barf		x	x		x		x	x		x		x				
49	Triangle (shape)																

Table A.4: Transcript of gesture elicitation session: Subject S4

Appendix A: Gesture Elicitation Sessions - Transcription

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
50	Hot		x						x		x		x				
51	Cold							x	x		x		x				
52	Snowflake							x	x		x		x				
53	Rain							x	x		x		x				
54	Rainbow								x		x		x				
55	I Love You				x			x	x		x		x				
56	Kiss			x		x											
57	Hug							x	x	x	x	x	x				
58	Sleep		x						x		x		x		x		
59	Baby								x		x		x				
60	Eat					x			x		x		x				
61	Money								x		x		x				
62	Waiter								x		x		x				
63	Umbrella		x						x		x		x				
64	Binocular		x		x				x		x		x				
65	Necklace																
66	Waves		x		x				x		x		x				
67	Mountain								x		x		x				x
68	Book		x						x		x		x				
69	Hungry	x															
70	Ouch!	x															
71	Yuk!	x															
72	Angry			x													
73	Sad		x	x													
74	Excited		x	x													
75	Press Door Bell								x		x		x				
76	Open Door								x		x		x				
77	Open Box								x		x		x				
78	Smell			x													
79	Sticky								x		x		x				
80	Lick					x											
81	Phone		x						x		x		x				
82	Shampoo		x		x				x		x		x				
83	Shave		x		x				x		x		x				
84	Salute		x		x				x		x		x				
85	Poison		x	x		x			x		x		x				
86	Cut Throat		x						x		x		x				
87	Steal								x		x		x				
88	Ghost	x															
89	Dragon	x															
90	Witch	x															
91	Monster	x							x		x		x				x
92	Pig	x	x														

Table A.4: Transcript of gesture elicitation session: Subject S4

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Yes				x												
2	No		x														
3	Bye									x		x		x			
4	Ice-cream			x		x	x			x		x		x			
5	Bathroom				x												
6	Wash Face		x							x		x		x			
7	Short	x			x										x	x	
8	Tall				x				x	x							x
9	Large								x	x							
10	Giant	x															
11	Cowboy	x													x	x	
12	Lasso		x		x				x	x		x					
13	Wave																
14	Necklace	x								x							
15	Umbrella									x							
16	Waiter								x		x		x				
17	Money									x		x		x			
18	Beard		x						x		x		x				
19	Baby	x															
20	Book								x	x							x
21	Mountain								x	x							x
22	Binocular		x		x					x		x		x			
23	Pizza		x			x	x		x		x		x				
24	Sip Soda	x	x	x		x											
25	Radio (music)		x		x				x	x		x		x			
26	Flute			x		x											
27	Violin								x	x							
28	Keyboard (music)														x		
29	Trumpet		x	x	x	x											
30	Rock Guitar								x	x		x		x			
31	Hot																x
32	Cold																x
33	Rain									x							
34	Snowflake									x							
35	Rainbow									x							
36	Smooth															x	
37	Bright Light	x			x					x		x		x			
38	Soft									x							
39	Listen								x		x		x				
40	Sticky	x															
41	Smell	x															
42	Train	x															
43	Drive								x	x	x	x	x	x			
44	Racing Car								x	x	x	x	x	x			
45	Aeroplane								x	x							
46	Helicopter									x		x		x			
47	Earthquake																x
48	Explosion	x		x													
49	Triangle (shape)														x		
50	Handshake									x							

Table A.5: Transcript of gesture elicitation session: Subject S5

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
51	Circle									x							
52	Square																
53	Stripes									x							
54	Zig-zag									x							
55	Stop								x								
56	Hello	x															
57	Eat			x		x											
58	Caterpillar																x
59	Lion	x	x	x	x												
60	Bird	x															
61	Butterfly								x	x							
62	Elephant		x		x					x							
63	Snake						x										
64	Spider								x		x		x				
65	Pig	x															
66	Alligator	x															
67	Mouse	x															
68	Tennis									x		x		x			
69	Touchdown																
70	Grand Slam	x							x	x							
71	Swimming								x	x							
72	Dealing Cards									x		x					
73	Fishing									x		x					
74	Bowling														x	x	
75	Canoe								x	x							
76	Throw Dice		x		x					x		x		x			
77	Mosquito Bite	x				x											
78	Whistle	x															
79	Tear Up									x		x					
80	Cross									x		x					
81	Sewing		x		x					x		x		x			
82	Ironing									x		x		x			
83	Pick-pocket, Steal		x		x				x		x		x				
84	Itch								x		x						
85	Smoke Cigar		x			x				x		x					
86	Phone	x	x						x		x		x				
87	Shampoo								x		x						
88	Shave		x		x	x				x		x		x			
89	Salute		x							x		x					
90	Milk Cow				x					x		x					
91	Naughty									x		x					x
92	Poison	x	x	x		x											
93	Cut Throat		x							x		x					
94	Lick						x										
95	Press Bell									x		x		x			
96	Pattercake									x		x		x			
97	Saw		x		x					x		x		x			
98	Hold bunch of balloons				x					x		x		x			
99	Type		x							x		x		x			
100	Fly a Kite									x	x	x	x	x			
101	Pretend to knit									x		x		x			

Table A.5: Transcript of gesture elicitation session: Subject S5

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
102	Fly a Kite																
103	Spank								x		x		x				
104	Crawl																x
105	Catch		x		x				x	x							
106	Jump																x
107	Throw Hammer								x	x	x	x					
108	Stir								x		x						
109	Knock on Door																
110	Open Box									x		x					
111	Open Door									x		x					
112	Pull Rope								x		x						
113	Kick/Push																
114	Dig Hole														x	x	
115	Dance		x		x					x		x		x			
116	Climb																x
117	Scratch Nose																x
118	Take a Picture									x		x		x			
119	Toss a Pancake								x	x	x	x	x	x			
120	Asleep									x		x		x			
121	Sad		x														x
122	Mad			x													
123	Angry			x													
124	Hug								x								
125	Barf	x		x													
126	Yummy							x	x		x						
127	Tired	x		x													
128	Don't Know				x												
129	Think		x							x		x					
130	Excited																x
131	Hungry			x		x											
132	Love	x						x		x		x					
133	Ouch!	x															
134	Yuk!	x															
135	Kiss			x		x											
136	Fish			x		x											x
137	Ghost	x															
138	Witch	x	x	x	x						x		x				
139	Dragon		x	x	x												
140	Monster	x															x

Table A.5: Transcript of gesture elicitation session: Subject S5

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Yes									x							
2	No								x	x							
3	Ice-cream					x				x							
4	Pizza					x				x							
5	Sip Soda					x				x							
6	Lion			x				x	x	x	x	x	x	x			
7	Bird	x								x		x		x			
8	Butterfly									x							
9	Caterpillar									x		x		x	x		
10	Fly Kite				x												
11	Spank									x		x		x			
12	Crawl		x					x									x
13	Jump									x		x		x	x		x
14	Hammer									x		x		x			
15	Train									x							
16	Helicopter									x							
17	Aeroplane									x							
18	Racing Car									x		x		x			
19	Drive Car (slow)								x	x	x	x	x	x			
20	Explosion	x							x	x							x
21	Earthquake		x														x
22	Drums		x							x		x		x	x		
23	Guitar							x	x	x							
24	Trumpet																
25	Piano									x		x		x	x		
26	Violin									x		x		x			
27	Cowboy							x	x	x					x		
28	Large									x							
29	Tall				x												
30	Short									x							
31	Wash Face		x							x							
32	Mosquito									x		x		x			
33	Bathroom								x								
34	Ocean									x		x		x			
35	Smell		x							x		x		x			
36	Sticky								x	x	x	x	x	x			
37	Listen		x						x		x		x				
38	Hot		x						x		x		x				
39	Cold								x	x	x	x	x				
40	Yummy					x				x		x		x			
41	Barf (vomit)			x		x			x	x	x	x	x	x			
42	Hug									x		x		x			
43	Sad			x													
44	Angry			x													
45	Yuk!	x			x												
46	Love								x	x	x	x	x	x			
47	Hungry		x		x					x		x		x			
48	Don't Know				x			x	x	x							
49	Tired		x							x		x		x			
50	Necklace		x														

Table A.6: Transcript of gesture elicitation session: Subject S6

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
51	Milk Cow									x							
52	Salute		x							x		x		x			
53	Shave									x		x		x			
54	Shampoo									x		x		x			
55	Phone		x														
56	Kick																
57	Throw									x		x					
58	Lick					x											
59	Cut Throat		x							x		x		x			
60	Poison		x			x											
61	Kiss					x											
62	Rainbow					x											
63	Snowflake									x		x		x			
64	Rain									x		x		x			
65	Elephant																
66	Alligator									x							
67	Pig			x		x											
68	Spider									x		x		x			
69	Snake									x		x		x			
70	Knock on Door															x	
71	Open Door									x		x		x			
72	Open Box									x							
73	Ghost	x		x													
74	Witch			x													
75	Dragon			x	x					x						x	
76	Monster	x		x	x					x		x		x			

Table A.6: Transcript of gesture elicitation session: Subject S6

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Bye	x							x		x		x				
2	Yes								x		x		x				
3	No		x														
4	Bird								x	x							
5	Butterfly								x	x	x	x	x	x			
6	Caterpillar									x		x		x			
7	Elephant		x						x								
8	Explosion																x
9	Sip Soda			x		x											
10	Tall														x		
11	Short														x		
12	Handshake								x		x		x				
13	Wash Face								x		x		x				
14	Bathroom								x								
15	Train								x	x	x	x	x		x		
16	Helicopter									x		x		x			
17	Racing Car									x		x		x			
18	Pizza	x		x		x											
19	Ice-cream					x											
20	Asleep		x														
21	Barf					x											
22	Lion			x		x			x		x		x		x		x
23	Yummy			x		x			x		x		x				
24	Catch								x	x	x	x	x	x			
25	Crawl																x
26	Climb Mountain	x			x										x		
27	Cut Throat								x	x	x	x	x	x			
28	Dance														x		
29	Dig Hole														x	x	
30	Fly Kite									x		x		x	x		
31	Swimming		x						x	x					x		x
32	Cold																
33	Hot			x		x											
34	Rainbow																
35	Raining								x	x	x	x	x	x			
36	Snowflake									x		x		x			
37	Sad			x													
38	Kiss			x		x											
39	Hug															x	
40	Hungry								x		x		x				
41	Ouch!			x		x											
42	Yuk!					x											
43	Angry		x	x					x		x		x				
44	Itch								x		x		x				
45	Throw								x	x	x	x	x	x	x		
46	Spank								x		x		x				
47	Witch			x		x											
48	Dragon			x					x						x		x
49	Ghost	x		x													
50	Monster	x		x		x			x	x	x	x	x	x	x		x

Table A.7: Transcript of gesture elicitation session: Subject S7

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
51	Alligator					x			x	x	x	x	x	x			
52	Triangle																

Table A.7: Transcript of gesture elicitation session: Subject S7

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Yes	x															
2	Good-bye	x															
3	Don't Know																
4	Hello																
5	No	x															
6	Stop	x															
7	Kiss		x			x				x		x					
8	Mickey Mouse		x						x		x		x				
9	Waiter									x		x		x			
10	Giant									x				x			
11	Open Box									x		x		x			
12	Cowboy/Horse Ride								x	x	x	x	x	x			x
13	Lasso									x		x		x			
14	Baby			x				x		x		x		x			
15	Bathroom																
16	Money		x		x					x		x		x			
17	Necklace		x							x		x		x			
18	Umbrella									x		x		x			
19	Binoculars		x	x	x					x		x		x			
20	Trumpet		x			x				x		x		x			
21	Violin		x							x		x		x			
22	Guitar			x						x		x		x			
23	Piano									x		x		x			
24	Saxophone		x	x						x		x		x			
25	Flute		x	x						x		x		x			
26	Drum									x		x					
27	Explosion							x	x	x							x
28	Earthquake																x
29	Pizza		x			x				x		x		x			
30	Ice Cream		x			x	x			x		x		x			
31	Yuk!	x	x														
32	Sip Soda		x							x		x		x			
33	Eat		x	x	x	x											
34	Yummy			x	x					x		x		x			
35	Triangle									x		x		x			
36	Mountain									x		x		x			
37	Square									x		x		x			
38	Circle									x		x		x			
39	Stripes									x		x		x			
40	Hungry		x	x	x	x											
41	Excited			x						x							x
42	Tired			x		x				x							
43	Hug		x						x	x	x	x	x	x			
44	Sad			x													
45	Love	x	x														
46	Ouch!	x															
47	Angry		x	x													
48	Fast Car (Racing)			x				x	x	x	x	x	x	x			
49	Train/Pull Whistle									x		x		x			
50	Helicopter		x	x						x		x		x			

Table A.8: Transcript of gesture elicitation session: Subject S9

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
51	Car (slow)								x	x		x		x			
52	Aeroplane							x	x	x							
53	Listen		x	x	x				x		x		x				
54	Captured (Surrender)								x	x							
55	Bright Light		x							x			x				
56	Hot									x		x		x			
57	Smell		x							x		x		x			
58	Smooth									x		x		x	x		
59	Cold							x	x	x							
60	Soft				x					x		x		x			
61	Ten Pin Bowling									x		x		x			
62	Cards		x	x	x					x		x		x			
63	Fishing									x		x		x			
64	How Big? (fish)								x		x		x				
65	Canoe									x		x		x			
66	Swimming (Crawl)							x	x	x							
67	Grand Slam																
68	Make a Basket		x		x					x							
69	Tennis									x		x		x			
70	Throw Dice									x		x		x			
71	Football/Touchdown								x	x							x
72	Rain									x		x		x			
73	Cold		x					x	x	x							
74	Hot									x		x		x			
75	Sunny		x		x					x		x		x			
76	Rainbow									x							
77	Snowflake									x		x		x			
78	Lion	x		x		x											
79	Pig	x															
80	Caterpillar									x				x			
81	Butterfly							x	x	x							
82	Alligator		x			x		x	x	x							
83	Elephant		x					x									
84	Snake		x														x
85	Fish		x							x							x
86	Bird		x						x	x							
87	Spider		x		x				x		x		x				
88	Beard								x		x		x				
89	Poison		x			x	x	x									
90	Naughty									x		x		x			
91	Large									x							
92	Tall									x							
93	Short									x							
94	Milking a Cow		x		x					x		x		x			
95	Mosquito					x				x		x		x			
96	Steal									x							
97	Waves (Sea)									x							
98	Think		x	x													x
99	Toss a Pancake									x		x		x			
100	Shampoo									x		x		x			
101	Cigar		x	x	x	x				x		x		x			

Table A.8: Transcript of gesture elicitation session: Subject S9

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
102	Balloons			x						x		x		x			
103	Kite		x							x		x		x			
104	Pattercake									x		x		x			
105	Salute		x							x		x		x			
106	Press Door Bell									x		x		x			
107	Open Door									x		x		x			
108	Close Door									x		x		x			
109	Jump							x									x
110	Itch																
111	Wash Face	x	x							x		x		x			
112	Dig Hole							x		x		x		x			
113	Crawl							x		x		x		x			
114	Pull Rope							x		x		x		x			
115	Asleep		x	x													x
116	Take a Picture	x	x		x					x		x		x			
117	Handshake									x		x		x			
118	Dance							x	x	x	x	x	x	x			
119	Sticky		x							x		x		x			
120	Knit		x						x	x	x	x	x	x			
121	Cut Throat		x							x		x		x			
122	Sewing		x		x				x	x	x	x	x	x			
123	Whistle	x															
124	Stir		x		x					x		x		x			
125	Cup		x		x					x		x		x			
126	Type (Typewriter)									x		x		x			
127	Climb							x	x	x							
128	Tear Up								x	x		x		x			
129	Throw									x		x		x			
130	Knock									x		x		x			
131	Saw		x		x					x		x		x			
132	Bring! bring! (Phone)		x							x		x		x			
133	Catch		x		x					x		x		x			
134	Hammer									x		x		x			
135	Push									x							
136	Shave		x	x	x	x				x		x		x			
137	Ironing		x							x		x		x			
138	Dragon	x		x				x	x	x							
139	Witch			x				x		x		x		x			
140	Ghost	x						x	x	x							
141	Monster	x						x	x	x							

Table A.8: Transcript of gesture elicitation session: Subject S9

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Yes				x												
2	Bye									x							
3	Stop	x							x								
4	Hello	x							x								
5	Don't Know				x												
6	Hungry					x											
7	Yummy							x	x		x		x				
8	Sip Soda			x		x											
9	Eat			x													
10	Ice-cream					x	x										
11	Pizza					x											
12	Rain									x		x		x			
13	Rainbow				x												
14	Snowflake								x		x		x				
15	Hot		x							x		x		x			
16	Cold							x	x	x							
17	Ironing									x		x		x			
18	Shave									x		x		x			
19	Push															x	
20	Hammer														x	x	
21	Saw														x	x	
22	Throw/Catch								x		x		x				
23	Phone									x		x		x			
24	Open Door								x		x		x				
25	Shampoo								x		x		x				
26	Wash Face								x		x		x				
27	Close Door								x		x		x				
28	Dig Hole														x	x	
29	Crawl								x	x					x	x	x
30	Excited	x		x					x	x					x	x	x
31	Pull Rope									x		x		x			
32	Whistle			x													
33	Sewing								x		x		x		x		
34	Itch									x		x		x			
35	Fast Car									x					x	x	
36	Train														x	x	
37	Helicopter								x		x		x				
38	Car														x	x	
39	Type				x				x	x	x	x	x	x			
40	Aeroplane								x	x							
41	Spank									x		x		x			
42	Tear Paper				x				x	x	x	x	x	x			
43	Climb														x	x	
44	Knock									x		x		x			
45	Throw									x		x		x			
46	Cut Throat								x		x		x				
47	Knitting																
48	Sticky									x		x		x			
49	Dance														x	x	
50	Handshake								x		x		x				

Table A.9: Transcript of gesture elicitation session: Subject S11

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
51	Take Picture		x		x					x		x		x			
52	Asleep		x														x
53	Giant				x												
54	Waiter								x		x		x				
55	Mickey Mouse		x							x							
56	Cowboy								x	x					x	x	x
57	Lasso								x	x	x		x		x	x	x
58	Barf					x											
59	Pay Attention/Listen	x															
60	Salute									x		x					
61	Bird								x	x							
62	Butterfly								x	x							
63	Fish					x											
64	Snake						x										
65	Elephant	x															
66	Lion	x															
67	Caterpillar								x		x		x				
68	Alligator					x			x	x							
69	Pig	x		x													
70	Spider	x															
71	Baby								x	x	x	x	x	x			
72	Kiss the Baby		x	x		x			x	x	x	x	x	x			
73	Love							x	x	x	x	x	x	x			
74	Ouch!	x															
75	Angry														x	x	
76	Sad			x													
77	Monster			x		x											x
78	Hug	x								x		x		x			
79	Open Box									x		x		x			

Table A.9: Transcript of gesture elicitation session: Subject S11

	Gesture	Vocalisation	Head	Facial Exp.	Eye Gaze	Mouth	Tongue	Trunk	Arm Left	Arm Right	Wrist Left	Wrist Right	Hand left	Hand Right	Legs	Feet	Whole Body
1	Pull Tail									x		x		x			
2	Yes	x															
3	Hello/Hi	x							x	x							
4	No	x															
5	Kiss			x		x											
6	Handshake								x		x		x				
7	Don't Know	x															
8	Bye	x							x		x		x				
9	Stop	x							x								
10	Yuk	x															
11	Love	x	x		x												
12	Ouch	x								x		x		x	x		
13	Angry	x						x	x	x					x		x
14	Sad		x	x					x	x							
15	Tired		x	x					x	x							
16	Hungry	x															
17	Excited	x		x					x	x							
18	Hug								x	x							
19	Sticky								x	x	x	x	x	x	x		
20	Type		x		x					x		x		x			
21	Lick		x		x				x		x		x				
22	Dance																x
23	Knit								x	x							
24	Crawl	x	x					x	x	x							
25	Wash Face		x						x	x	x		x				
26	Cut Throat		x						x		x		x				
27	Pattercake								x	x			x	x			
28	Saw									x		x		x			
29	Mouse	x															
30	Elephant																

Table A.10: Transcript of gesture elicitation session: Subject S12

Appendix B

Neural Network Training Algorithm

B.1 Backpropagation Training Algorithm

The backpropagation training algorithm was provided by the “trainbpx” function of the MATLAB Neural Network toolbox (Demuth & Beale, 1994). This employed backpropagation of error with adaptive learning rate to decrease the training time and momentum to decrease the likelihood of getting stuck in local minima. The backpropagation learning rule involves minimising the sum-squared error (SSE) of the network. This is achieved by incrementally changing the network’s weights and biases in the direction of steepest descent with respect to error. The derivatives of error (called delta vectors) are calculated for the network’s output layer, and then backpropagated through the network. The basic backpropagation algorithm is described in most introductory neural network texts (Zurada, 1992, Fausett, 1994). Momentum and adaptive learning rate were implemented in a manner similar to that described in Vogl et al. (1988). The sum-squared error was calculated from the individual errors of all training input-output pairs (i.e. “batch mode” training).

If the SSE had decreased from the last pass through the training data (epoch), the adaptive learning rate lr was increase by a factor of 1.05 and the weights incremented using the equation:

$$\Delta \mathbf{W}(i, j) = mc \Delta \mathbf{W}_{\text{old}}(i, j) + (1 - mc) lr \cdot \mathbf{d}(i) \mathbf{p}(j) \quad \text{Eqn [B.1]}$$

Momentum term mc was set to 0.95. \mathbf{d} is the delta vector and \mathbf{p} the input vector.

If the SSE had increased by more that a factor of 1.04 the learning rate was decreased by multiplying by 0.7 and the effects of the momentum term temporarily disabled.

B.2 Nguyen-Widrow Initialisation

This scheme of weight initialisation typically gives rise to faster learning. It is used to initialise weights between the input and the hidden-layer and is designed to improve the learning ability of neurons in the hidden-layer. Weights and biases are initialised so that the resultant hidden layer neuron activation is in the linear region (Nguyen & Widrow, 1990). For the particular case of inputs lying on the interval -1 to 1 it is computed as follows: n , number of input units, p , number of hidden units, β , scale factor

$$b = 0.7 (p)^{\frac{1}{n}} \quad \text{Eqn [B.2]}$$

For each hidden unit ($j=1, \dots, p$):

Initialise the associated input weight vector

$$v_{ij}(\text{old}) = \text{random number between } \pm 1 \quad i = (1, \dots, n) \quad \text{Eqn [B.3]}$$

Compute $\|v(\text{old})\|$

Re-initialise weights:

$$v_{ij} = \frac{\beta v_{ij}(\text{old})}{\|v(\text{old})\|} \quad \text{Eqn [B.4]}$$

Set bias:

$$v_{0j} = \text{random number between } \pm \beta \quad \text{Eqn [B.5]}$$

End

Appendix C

Confusion Matrices

Note: Upper confusion matrix corresponds to training data, lower confusion matrix correspond to test data

C.1 Confusion Matrices: Network Complexity

Experiment ex21uk

12-1t-26p

cfmname: ctrl100

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1																										100
cards	2																										100
cut throat	3																										100
drive the car	4																										100
drums	5																										100
heavy weight	6																										100
helicopter	7																										100
hot	8																										100
ice-cream	9																										100
ironing	10																										100
knock on the door	11																										100
lassou	12																										100
light feather	13																										100
rainbow	14																										100
rock a baby	15																										100
rock guitar	16																										100
scratch your knee	17																										100
shake hands	18																										100
shake	19																										100
spank	20																										100
spider	21																										100
stroke the cat	22																										100
surrender	23																										100
whistle	24																										100
umbrella	25																										100
violin	26																										100

 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
 Number of Epochs = 100
 No of gestures recognised at or above 80% = 1
 Average rec. rate of best 1 gestures = 100.00%
 Average recognition rate of all gestures = 3.85%

cfmname: cte100

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1																										100
cards	2																										100
cut throat	3																										100
drive the car	4																										100
drums	5																										100
heavy weight	6																										100
helicopter	7																										100
hot	8																										100
ice-cream	9																										100
ironing	10																										100
knock on the door	11																										100
lassou	12																										100
light feather	13																										100
rainbow	14																										100
rock a baby	15																										100
rock guitar	16																										100
scratch your knee	17																										100
shake hands	18																										100
shake	19																										100
spank	20																										100
spider	21																										100
stroke the cat	22																										100
surrender	23																										100
whistle	24																										100
umbrella	25																										100
violin	26																										100

 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
 Number of Epochs = 100
 No of gestures recognised at or above 80% = 1
 Average rec. rate of best 1 gestures = 100.00%
 Average recognition rate of all gestures = 3.85%

Experiment ex18uk

12-2t-26p

cfmname: ctr600

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1											100															
cards	2						55	45																			
cut throat	3						33	67																			
drive the car	4											100															
drums	5											100															
heavy weight	6					100																					
helicopter	7											100															
hot	8						29	71																			
ice-cream	9						50	38			6															6	
ironing	10						71	21																		7	
knock on the door	11											13	88														
lassou	12												100														
light feather	13						73	27																			
rainbow	14										8	33								8							50
rock a baby	15						60	20					10													10	
rock guitar	16						18	29				12	29													12	
scratch your knee	17																									100	
shake hands	18											18	36								18					27	
shave	19						29	65					6														
spank	20											20	7														73
spider	21											17	67														17
stroke the cat	22																									100	
surrender	23												100														
whistle	24												100														
umbrella	25													100													
violin	26																										100

Number of Epochs = 600
 No of gestures recognised at or above 80% = 3
 Average rec. rate of best 3 gestures = 100.00%
 Average recognition rate of all gestures = 14.73%

cfmname: cte600

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1											100															
cards	2						40	60																			
cut throat	3						10	90																			
drive the car	4												90							10							
drums	5											20															
heavy weight	6					100																					
helicopter	7												100														
hot	8								90				10														
ice-cream	9						90	10																			
ironing	10						60	30					10														
knock on the door	11								10				90														
lassou	12												100														
light feather	13						70	30																			
rainbow	14											10	10							20						60	
rock a baby	15						80	10					10														
rock guitar	16						50	20					20													10	
scratch your knee	17																									100	
shake hands	18											10	60												30		
shave	19						30	70																			
spank	20											10	10							10						70	
spider	21											10	60							20						10	
stroke the cat	22																									100	
surrender	23												100														
whistle	24												100														
umbrella	25												100														
violin	26												100														

Number of Epochs = 600
 No of gestures recognised at or above 80% = 4
 Average rec. rate of best 4 gestures = 97.50%
 Average recognition rate of all gestures = 15.38%

Experiment ex23uk

12-3t-26p

cfmname: ctr900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1							7		14		29							7	7		21				14
cards	2					9				91																
cut throat	3									87									13							
drive the car	4									8	8	15								23	31		15			
drums	5									25		38					6			19		13				
heavy weight	6					94				6																
helicopter	7											92														8
hot	8									100																
ice-cream	9									81									13	6						
ironing	10					21				79																
knock on the door	11									6	75								6	6		6				
lassou	12											100														
light feather	13					53				47																
rainbow	14											67												33		
rock a baby	15					30				50						10				10						
rock guitar	16									65		6				12				6		12				
scratch your knee	17											100														
shake hands	18											18											82			
shave	19									94									6							
spank	20																20								80	
spider	21											92								8						
stroke the cat	22																25						75			
surrender	23											100														
whistle	24											85							8							8
umbrella	25											100														
violin	26											92							8							

Number of Epochs = 900
 No of gestures recognised at or above 80% = 3
 Average rec. rate of best 3 gestures = 97.92%
 Average recognition rate of all gestures = 17.67%

cfmname: cta900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1							10			10	30							30	10		10				
cards	2									100																
cut throat	3									50									50							
drive the car	4					10				10		10							20	40		10				
drums	5					10				20	10						10		30	10		10				
heavy weight	6					100																				
helicopter	7											100														
hot	8									100																
ice-cream	9									100																
ironing	10					10				90																
knock on the door	11											80							10						10	
lassou	12											100														
light feather	13					40				60																
rainbow	14											80											20			
rock a baby	15					30				60		10														
rock guitar	16					10				60								20			10					
scratch your knee	17																	90					10			
shake hands	18											10								10			80			
shave	19									90									10							
spank	20																								100	
spider	21											100														
stroke the cat	22																20						80			
surrender	23											100														
whistle	24											90														10
umbrella	25											100														
violin	26											100														

Number of Epochs = 900
 No of gestures recognised at or above 80% = 5
 Average rec. rate of best 5 gestures = 92.00%
 Average recognition rate of all gestures = 18.08%

Experiment ex19uk

12-4t-26p

cfmname: ctr1800

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	14				71				7																7		
cards	2						36		64																			
cut throat	3			13						7										80								
drive the car	4	8					54														38							
drums	5							88														6						
heavy weight	6								100																			
helicopter	7										8	23										8			31		15	
hot	8												88															
ice-cream	9																											
ironing	10																											
knock on the door	11																											
lassou	12																											
light feather	13																											
rainbow	14																											
rock a baby	15																											
rock guitar	16																											
scratch your knee	17																											
shake hands	18																											
shave	19																											
spank	20																											
spider	21																											
stroke the cat	22																											
surrender	23																											
whistle	24	15																										
umbrella	25																											
violin	26																											

Number of Epochs = 1800
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 94.50%
 Average recognition rate of all gestures = 33.30%

cfmname: cte1800

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	10																										
cards	2																											
cut throat	3																											
drive the car	4																											
drums	5																											
heavy weight	6																											
helicopter	7	10																										
hot	8																											
ice-cream	9																											
ironing	10																											
knock on the door	11																											
lassou	12																											
light feather	13																											
rainbow	14																											
rock a baby	15																											
rock guitar	16																											
scratch your knee	17																											
shake hands	18																											
shave	19																											
spank	20																											
spider	21																											
stroke the cat	22																											
surrender	23																											
whistle	24	10																										
umbrella	25																											
violin	26																											

Number of Epochs = 1800
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 90.00%
 Average recognition rate of all gestures = 30.38%

Experiment ex20uk

12-6t-26p

cfmname: ctr2900

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	86				7																				7	
cards	2			45		27				9										18							
cut throat	3			20																80							
drive the car	4	69				8															23						
drums	5	38				38			13												13						
heavy weight	6					94				6																	
helicopter	7											54									8		23		15		
hot	8								100																		
ice-cream	9			25				6											63			6					
ironing	10						7			93																	
knock on the door	11	25			25		6				31										13						
lassou	12										94																
light feather	13					53				33		13															
rainbow	14										8		8				8					17	42	17			
rock a baby	15						20		40		20									10	10						
rock guitar	16	6				12			65											6		12					
scratch your knee	17																		42			58					
shake hands	18					9												9			82						
shave	19																			100							
spank	20																7				93						
spider	21											67							8			25					
stroke the cat	22																					100					
surrender	23																						100				
whistle	24	31				8			8															38		15	
umbrella	25																							73		27	
violin	26											83								8	8						

Number of Epochs = 2900

No of gestures recognised at or above 80% = 9
Average rec. rate of best 9 gestures = 95.49%
Average recognition rate of all gestures = 39.69%

cfmname: cte2900

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80			20																						
cards	2			20				30		10										40							
cut throat	3																		100								
drive the car	4	30				10		10		10		10									30						
drums	5	40				20		30													10						
heavy weight	6					100																					
helicopter	7	20						10				40												30			
hot	8	10						80		10																	
ice-cream	9			10															90								
ironing	10						10				90																
knock on the door	11	20						10				50												20			
lassou	12											90														10	
light feather	13			20			50		10		10			10													
rainbow	14											20		10								20	50				
rock a baby	15						10		60			10									20						
rock guitar	16								90																		
scratch your knee	17																	50			10		40				
shake hands	18																			100							
shave	19																				100						
spank	20																				100						
spider	21											90														10	
stroke the cat	22																					100					
surrender	23																						80		20		
whistle	24	60										10												20		10	
umbrella	25											10												60		30	
violin	26											80										20					

Number of Epochs = 2900

No of gestures recognised at or above 80% = 9
Average rec. rate of best 9 gestures = 91.11%
Average recognition rate of all gestures = 36.15%

Experiment ex14uk

12-8t-26p

cfmname: ctr5000

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	86			7																					7		
cards	2					27														73								
cut throat	3																			100								
drive the car	4	69			31																							
drums	5	31			56					6											6							
heavy weight	6					100																						
helicopter	7	8								23	31								8							31		
hot	8					6		94																				
ice-cream	9																			94	6							
ironing	10					21				79																		
knock on the door	11	6			38					31	13		13															
lassou	12									25	56											6		6		6		
light feather	13					100																						
rainbow	14													100														
rock a baby	15				10	40				20										30								
rock guitar	16				35			29		6	6									12	6				6			
scratch your knee	17																	42								58		
shake hands	18					18				9																18		
shave	19																					100						
spank	20					20																73				7		
spider	21										8											42						
stroke the cat	22													8												92		
surrender	23																									100		
whistle	24	23			8					8										8						46		8
umbrella	25									7		7														80		7
violin	26									8	67		8							8		8						

Number of Epochs = 5000
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 95.93%
 Average recognition rate of all gestures = 40.66%

cfmname: cte5000

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	80			20																							
cards	2					40														60								
cut throat	3																			100								
drive the car	4	40			40					20																		
drums	5	40			40	20																						
heavy weight	6					100																						
helicopter	7	20			10					30	10															30		
hot	8							60		30										10								
ice-cream	9																			100								
ironing	10					20				80																		
knock on the door	11	10			10			10			40															30		
lassou	12										10	40									10	10			10		20	
light feather	13					90		10																				
rainbow	14													100														
rock a baby	15					30		20			10									40								
rock guitar	16				20	10		30												30	10							
scratch your knee	17																		40								60	
shake hands	18					10					30											30						
shave	19																					100						
spank	20					20																				80		
spider	21																									50		
stroke the cat	22																									100		
surrender	23																									80		20
whistle	24	40			10						10														10		30	
umbrella	25																									60		20
violin	26										10	70		10								10						

Number of Epochs = 5000
 No of gestures recognised at or above 80% = 8
 Average rec. rate of best 8 gestures = 90.00%
 Average recognition rate of all gestures = 38.85%

Experiment ex13uk

12-10t-26p

cfmname: ctr2700

	Network Decision																										
Gesture:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	50							7		21															7	
cards	2		27			18				27			9							18							
cut throat	3			7																	93						
drive the car	4	23			8	31				15				8					8					8			
drums	5	6				44			13		6	13	6					6				6					
heavy weight	6						88				13																
helicopter	7			15							8	15												38	8	15	
hot	8									100																	
ice-cream	9			13					6											75				6			
ironing	10									14																	
knock on the door	11	6							13			25	31		13						6				6		
lassou	12												69											13		13	
light feather	13									33						47				20							
rainbow	14													100													
rock a baby	15								20		10		20						10	20	10					10	
rock guitar	16					6					53	6							6	12		6	6		6		
scratch your knee	17																		100								
shake hands	18												9	9		27				9	36		9				
shave	19																			100							
spank	20																	20					73		7		
spider	21															25									75		
stroke the cat	22											8						8							75		
surrender	23																							93		7	
whistle	24										8														23	69	
umbrella	25																								73		27
violin	26			8									25												50		8

Number of Epochs = 2700
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 95.15%
 Average recognition rate of all gestures = 48.44%

cfmname: cte2700

	Network Decision																												
Gesture:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26			
bird	1	30				40					10	10														10			
cards	2			10						30		20									40								
cut throat	3																			100									
drive the car	4					20			10		20		10				20				10								
drums	5	10				20	10			20		10								10						20			
heavy weight	6						80									20													
helicopter	7								10			10	30												30	20			
hot	8									80		20																	
ice-cream	9				10																90								
ironing	10										90																		
knock on the door	11					10					20	30												10		10			
lassou	12												20										40		20	10			
light feather	13									40		10		30						20									
rainbow	14																										100		
rock a baby	15								20		70																10		
rock guitar	16						10				70												10	10					
scratch your knee	17																			80									
shake hands	18						10								10						10	40				30			
shave	19																				100								
spank	20																										90		
spider	21												10			10										50		30	
stroke the cat	22																									80			
surrender	23																									80		20	
whistle	24																									10	80	10	
umbrella	25																									60		20	
violin	26												10	30		10										40			10

Number of Epochs = 2700
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 86.00%
 Average recognition rate of all gestures = 41.54%

Experiment ex32uk

12-12t-26p

cfmname: ctr8300

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	57				21		7		7														7			
cards	2						9		91																		
cut throat	3			40				7												53							
drive the car	4	23				54				8								8			8						
drums	5	19				69		6													6						
heavy weight	6						94			6																	
helicopter	7											46							8				31		15		
hot	8								100																		
ice-cream	9			19					25											50	6						
ironing	10						7			93																	
knock on the door	11	6				25						38		13		6										13	
lassou	12											88										6		6			
light feather	13						40		13		33		13														
rainbow	14													92												8	
rock a baby	15						30		10		20					10			20	10							
rock guitar	16								41							24				29		6					
scratch your knee	17																100										
shake hands	18					9						9	36			9	9				27						
shave	19			12																	88						
spank	20																					93		7			
spider	21													25								75					
stroke the cat	22																						100				
surrender	23																							93		7	
whistle	24											8				8								38	38	8	
umbrella	25																								20		73
violin	26												33							8			42				17

Number of Epochs = 8300

No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 94.02%
 Average recognition rate of all gestures = 51.49%

cfmname: cte8300

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80				20																					
cards	2								80		10									10							
cut throat	3			20																80							
drive the car	4	10				20			10		10					20				20							
drums	5	40				20			20							20				10							
heavy weight	6						90						10														
helicopter	7	20										40												20	10	10	
hot	8								90		10																
ice-cream	9			30					20											50							
ironing	10						20				80																
knock on the door	11				10				20				50											20			
lassou	12												60								10	30					
light feather	13						40		40		10			10													
rainbow	14													80								20					
rock a baby	15						20		60			10							10								
rock guitar	16					10			50							10			10	20							
scratch your knee	17																90						10				
shake hands	18					10						40		10			10			30							
shave	19			10															90								
spank	20																			90		10					
spider	21																					80					20
stroke the cat	22																10						90				
surrender	23																							70		30	
whistle	24												10											10	80		
umbrella	25																						20		30		50
violin	26												70		10								20				

Number of Epochs = 8300

No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 85.45%
 Average recognition rate of all gestures = 45.38%

Experiment ex15uk

12-14t-26p

cfmname: ctr5200

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	79			14																				7		
cards	2					9		73	9											9							
cut throat	3							7	7											87							
drive the car	4	54			15	8													8		15						
drums	5	38			44			6			6										6						
heavy weight	6					94				6																	
helicopter	7										8	38												15	8	8	23
hot	8								100																		
ice-cream	9								25											69	6						
ironing	10								7		93																
knock on the door	11	19			13			6			25	13		6					6							13	
lassou	12											81										6		6			6
light feather	13		7			40		7		13			33														
rainbow	14													92								8					
rock a baby	15					30				20						30			20								
rock guitar	16							35								41				18		6					
scratch your knee	17																100										
shake hands	18					9									9		9	64		9							
shave	19																	100									
spank	20														13						87						
spider	21											8		33								58					
stroke the cat	22																8						92				
surrender	23																							93	7		
whistle	24	8						8		23														31	31		
umbrella	25										7														40		53
violin	26											25		8						8		33					25

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 93.07%
 Average recognition rate of all gestures = 53.81%

cfmname: cte5200

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	90			10																						
cards	2		10					50		10										30							
cut throat	3																			100							
drive the car	4	10			20			10		10		10				10		10		20							
drums	5	30			30			30								10											
heavy weight	6					80								20													
helicopter	7	20									10	30												20	10	10	
hot	8	10						80								10											
ice-cream	9							10		10									90								
ironing	10					10		10		80																	
knock on the door	11	20						20		10	30												20				
lassou	12										10	80										10					
light feather	13					40		50					10														
rainbow	14													90								10					
rock a baby	15					10		20			10				30			30									
rock guitar	16							50							30			20									
scratch your knee	17																90					10					
shake hands	18													10			10	70		10							
shave	19																		100								
spank	20																			100							
spider	21											10										80				10	
stroke the cat	22																10						90				
surrender	23																							70		30	
whistle	24	20								20														10	50		
umbrella	25													10								10		40		40	
violin	26										10	40		10				20				10					10

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 87.27%
 Average recognition rate of all gestures = 49.62%

Experiment ex24uk

12-15t-26p

cfmname: ctr5200

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	79				7					7														7	
cards	2						9	91																		
cut throat	3		73					13											13							
drive the car	4	15			15					31								23	8					8		
drums	5	19			38			13		6	19									6						
heavy weight	6					94				6																
helicopter	7									8	46												15	8	23	
hot	8					6	94																			
ice-cream	9		50				25											19	6							
ironing	10					7		7	86																	
knock on the door	11						6		31	31	13								6				13			
lassou	12									88										6		6				
light feather	13					47	13	27			13															
rainbow	14											100														
rock a baby	15					30	10	30							30											
rock guitar	16						47			6					29				6		12					
scratch your knee	17															100										
shake hands	18										18							64	18							
shave	19		76				6												18							
spank	20																			93		7				
spider	21									8	25										67					
stroke the cat	22																				100					
surrender	23																					100				
whistle	24						8		15													15	54	8		
umbrella	25									7													33	7	53	
violin	26										50	17								8				8		17

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 9
 Average rec. rate of best 9 gestures = 94.94%
 Average recognition rate of all gestures = 53.45%

cfmname: cte5200

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	30			20						40													10		
cards	2		10				80	10																		
cut throat	3		90																10							
drive the car	4			20			20	10		20									10					20		
drums	5	30			20		20	20		20					10											
heavy weight	6				100																					
helicopter	7									30	60												10			
hot	8						90	10																		
ice-cream	9		60				30											10								
ironing	10				20			80																		
knock on the door	11			10			10		10	50													20			
lassou	12									10	80										10					
light feather	13				60		40																			
rainbow	14											80									20					
rock a baby	15				50		30			10				10												
rock guitar	16				10		30							60												
scratch your knee	17															90						10				
shake hands	18															10	60		30							
shave	19		100																							
spank	20																			100						
spider	21										20											80				
stroke the cat	22																10						90			
surrender	23										10												60		30	
whistle	24									10			10										10	70		
umbrella	25																				20		40		40	
violin	26										70	10						10								10

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 88.00%
 Average recognition rate of all gestures = 47.69%

Experiment ex16uk

12-16t-26p

cfmname: ctr6700

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	86									7														7	
cards	2						9	91																		
cut throat	3							13											87							
drive the car	4	38		31						8									23							
drums	5	63			19		6			6										6						
heavy weight	6					94				6																
helicopter	7	8										31						8						31	15	8
hot	8							100																		
ice-cream	9							25											69							6
ironing	10							7		93																
knock on the door	11	19		6	6						25	6						6						25		6
lassou	12											81									6				6	6
light feather	13					33		7		27			33													
rainbow	14												92									8				
rock a baby	15					30		10		20					10	10			10	10						
rock guitar	16	6		12				53								12				12			6			
scratch your knee	17																100									
shake hands	18										9									91						
shave	19							6													94					
spank	20																		7		93					
spider	21											8		8								83				
stroke the cat	22																					100				
surrender	23			7																			7	29	57	
whistia	24																							92	8	
umbrella	25												7											13	80	
violin	26											25										8		8		58

Number of Epochs = 6700
 No of gestures recognised at or above 80% = 14
 Average rec. rate of best 14 gestures = 91.37%
 Average recognition rate of all gestures = 56.71%

cfmname: cta6700

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	100																								
cards	2							80		10									10							
cut throat	3																		100							
drive the car	4	30		10				10			10								10						30	
drums	5	30		10	20			20		10															10	
heavy weight	6					80							20													
helicopter	7	20										20												60		
hot	8							100																		
ice-cream	9							30											70							
ironing	10									80		20														
knock on the door	11	20						10			40													30		
lassou	12											80									10				10	
light feather	13					30		50				20														
rainbow	14												90								10					
rock a baby	15					10		50			10			10						20						
rock guitar	16				10			60						10		10			10	10						
scratch your knee	17																80			10		10				
shake hands	18																10	90								
shave	19																		100							
spank	20																		10	90						
spider	21											10		10							60			10	10	
stroke the cat	22																		10			90				
surrender	23	10																			10			20	60	
whistle	24																							100		
umbrella	25																				20			10	70	
violin	26											20									20			10		20

Number of Epochs = 6700
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 90.00%
 Average recognition rate of all gestures = 50.00%

Experiment ex25uk

12-17t-26p

cfmname: ctr5900

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	21				7					57														7		
cards	2			9				9		36		27														18	
cut throat	3					7																				93	
drive the car	4	46						8																			23
drums	5	13						19																			25
heavy weight	6									94																	6
helicopter	7												23														8
hot	8				12																						24
ice-cream	9					6																					88
ironing	10							14																			86
knock on the door	11											6	25	38												13	6
lassou	12													38													13
light feather	13									40																	53
rainbow	14																										7
rock a baby	15														92												92
rock guitar	16																										10
scratch your knee	17																										6
shake hands	18																										53
shave	19																										9
spank	20																										9
spider	21																										27
stroke the cat	22																										100
surrender	23																										9
whistle	24	8																									45
umbrella	25																										9
violin	26																										9

Number of Epochs = 5900
 No of gestures recognised at or above 80% = 6
 Average rec. rate of best 6 gestures = 92.97%
 Average recognition rate of all gestures = 34.80%

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	30				10						50	10														
cards	2													10													40
cut throat	3																										100
drive the car	4	10												10													10
drums	5									10																	10
heavy weight	6																										100
helicopter	7	10																									40
hot	8																										30
ice-cream	9																										20
ironing	10																										100
knock on the door	11	10																									10
lassou	12													40													10
light feather	13																										10
rainbow	14																										100
rock a baby	15																										10
rock guitar	16																										70
scratch your knee	17																										60
shake hands	18																										10
shave	19																										20
spank	20																										80
spider	21																										30
stroke the cat	22																										90
surrender	23																										20
whistle	24																										80
umbrella	25																										10
violin	26																										30

Number of Epochs = 5900
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 90.00%
 Average recognition rate of all gestures = 35.77%

Experiment ex17uk

12-18t-26p

cfmname: ctr7400

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	86			7																			7		
cards	2			9					45	18			27													
cut throat	3		100																							
drive the car	4	23		31							8				8			15		8			8			
drums	5	56			25						6									6			6			
heavy weight	6					94					6															
helicopter	7										8	31											15	31	8	8
hot	8							94									6									
ice-cream	9		81						13												6					
ironing	10				7					93																
knock on the door	11	6									56	13		6											19	
lassou	12											81									6		13			
light feather	13					33				20			47													
rainbow	14												92									8				
rock a baby	15			20			20			20					20	10					10					
rock guitar	16	6		6				24		6						29	6			18				6		
scratch your knee	17																100									
shake hands	18										9							82		9						
shave	19			88					6										6							
spank	20																		7		93					
spider	21											8		8								83				
stroke the cat	22																					100				
surrender	23											7											86		7	
whistle	24			8							8												15	62	8	
umbrella	25										7												33		60	
violin	26											58								8		8		8		17

Number of Epochs = 7400
 No of gestures recognised at or above 80% = 13
 Average rec. rate of best 13 gestures = 91.04%
 Average recognition rate of all gestures = 59.55%

cfmname: cte7400

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	80									10												10			
cards	2			20					10	50			10		10											
cut throat	3		100																							
drive the car	4	30		20											10					10			10	20		
drums	5	10			30	10		10		10										10			20			
heavy weight	6					90							10													
helicopter	7	20									30												20	30		
hot	8							60		30				10												
ice-cream	9		60						30										10							
ironing	10									90			10													
knock on the door	11	10						10			20	50											10			
lassou	12											80									10			10		
light feather	13					40		10					50													
rainbow	14												70								30					
rock a baby	15			10		20		30			10			20					10							
rock guitar	16			10		20		20			10			10	20				10							
scratch your knee	17																100									
shake hands	18										10						10	70		10						
shave	19			100																						
spank	20																		10		90					
spider	21											10									60		10		20	
stroke the cat	22																	10				90				
surrender	23																						70		30	
whistle	24										10												10	80		
umbrella	25																				10		50		40	
violin	26											90		10												

Number of Epochs = 7400
 No of gestures recognised at or above 80% = 9
 Average rec. rate of best 9 gestures = 88.89%
 Average recognition rate of all gestures = 52.31%

Experiment ex22uk

12-20t-26p

cfmname: ctr3000

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	86				7																				7	
cards	2						9		36					18							9						
cut throat	3																				73						
drive the car	4	23				31	15													15				8			
drums	5	50				6	25																			13	
heavy weight	6								94																		
helicopter	7																										
hot	8								6					94													
ice-cream	9																										
ironing	10	7																			56	6					
knock on the door	11	6																									
lassou	12																										
light feather	13																										
rainbow	14																										
rock a baby	15	10																									
rock guitar	16	6																									
scratch your knee	17																										
shake hands	18																										
shave	19																										
spank	20																										
spider	21																										
stroke the cat	22																										
surrender	23																										
whistle	24	8																									
umbrella	25																										
violin	26																										

Number of Epochs = 3000
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 91.89%
 Average recognition rate of all gestures = 52.36%

cfmname: cte3000

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	90				10																					
cards	2																										
cut throat	3																										
drive the car	4	20																									
drums	5	30																									
heavy weight	6																										
helicopter	7	20																									
hot	8	10																									
ice-cream	9																										
ironing	10																										
knock on the door	11	20																									
lassou	12																										
light feather	13																										
rainbow	14																										
rock a baby	15																										
rock guitar	16																										
scratch your knee	17																										
shake hands	18																										
shave	19																										
spank	20																										
spider	21																										
stroke the cat	22																										
surrender	23																										
whistle	24	10																									
umbrella	25																										
violin	26																										

Number of Epochs = 3000
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 88.00%
 Average recognition rate of all gestures = 45.38%

C.2 Comparison of Activation Functions

ex60
cfmname: ctr1900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	43			7	21			7																	21
cards	2			18				9	55	18																
cut throat	3			27																73						
drive the car	4	23				62									8									8		
drums	5	13				75			6												6					
heavy weight	6						94			6																
helicopter	7										54		8										15		23	
hot	8								100																	
ice-cream	9			6					25										63	6						
ironing	10							7	14	79																
knock on the door	11	13				13					44	19	13													
lassou	12											75										6		6		13
light feather	13						53	13		33																
rainbow	14												100													
rock a baby	15						30	30	20										10	10						
rock guitar	16					6		59			6				6	6		6	12							
scratch your knee	17															100										
shake hands	18					27							45							18		9				
shave	19			6															94							
spank	20					7											13			80						
spider	21											8	25								67					
stroke the cat	22												17			67						17				
surrender	23																						100			
whistle	24	15				8					15	23											31		8	
umbrella	25											13											53		33	
violin	26											75									25					

Number of Epochs = 1900
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 95.41%
 Average recognition rate of all gestures = 43.55%

cfmname: cte1900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	50				40																			10	
cards	2			10					40	20										30						
cut throat	3																			100						
drive the car	4	10				20		20	10	20			10							10						
drums	5	30				40	10	10												10						
heavy weight	6						100																			
helicopter	7										20	40											30		10	
hot	8								80	20																
ice-cream	9			40				20		80										40						
ironing	10						20				80															
knock on the door	11					10			10			50											30			
lassou	12											80														20
light feather	13						50	50																		
rainbow	14													100												
rock a baby	15						20	70				10														
rock guitar	16					10		70											10	10						
scratch your knee	17																		100							
shake hands	18					20							30		10					40						
shave	19			10																90						
spank	20					10											10				80					
spider	21										20	20										50				10
stroke the cat	22																70						30			
surrender	23																							70		30
whistle	24	10				20						20												50		
umbrella	25											10	10								10		60		10	
violin	26											60	10									30				

Number of Epochs = 1900
 No of gestures recognised at or above 80% = 8
 Average rec. rate of best 8 gestures = 88.75%
 Average recognition rate of all gestures = 36.92%

Appendix C.2 Comparison of Activation Functions

ex61
cfmname: ctr8000

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	93																						7			
cards	2		18										73														
cut throat	3			7																93							
drive the car	4	46				8						23							8				8		8		
drums	5	38				38						19	6														
heavy weight	6						94						6														
helicopter	7												38							8				15		15 23	
hot	8						6		94																		
ice-cream	9								19				6							69	6						
ironing	10								7		93																
knock on the door	11	13				6						50	13						6					6		6	
lassou	12												75									6		6		6	
light feather	13											7		53													
rainbow	14													92									8				
rock a baby	15						30								40	20							10				
rock guitar	16											6				41				6	6		12			6	
scratch your knee	17																100										
shake hands	18																	91				9					
shave	19														12					88							
spank	20																				93			7			
spider	21												8		17								75				
stroke the cat	22																							100			
surrender	23																								93		7
whistle	24	8										15												15	23	38	
umbrella	25																							13		87	
violin	26												17										8			75	

Number of Epochs = 8000
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 93.10%
 Average recognition rate of all gestures = 62.01%

cfmname: cte8000

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80				10						10															
cards	2		10										50								20						
cut throat	3																				100						
drive the car	4	30				10					10	30	10													10	
drums	5	30				20					20	10															
heavy weight	6						90								10												
helicopter	7	10										10	30											20	10	20	
hot	8								80					10													
ice-cream	9													20		10					70						
ironing	10						10					50				20											
knock on the door	11	10										20	40											10		10	
lassou	12												10	80													
light feather	13													50													
rainbow	14														100												
rock a baby	15						30			30			10			10	20										
rock guitar	16						10	10		10						50					10	10					
scratch your knee	17																	100									
shake hands	18																	10	60			30					
shave	19																			100							
spank	20																				100						
spider	21												10											60		30	
stroke the cat	22																							90			
surrender	23																								70		30
whistle	24	10											10												10	60	10
umbrella	25																								30		50
violin	26												40														40

Number of Epochs = 8000
 No of gestures recognised at or above 80% = 9
 Average rec. rate of best 9 gestures = 91.11%
 Average recognition rate of all gestures = 52.69%

Appendix C.2 Comparison of Activation Functions

ex62uk
cfmname: ctr200

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	14					29			29						29											
cards	2	100																									
cut throat	3	100																									
drive the car	4						38			8					8	46											
drums	5	13					31			13					13	31											
heavy weight	6	50								25					13	13											
helicopter	7						92			8																	
hot	8	100																									
ice-cream	9	94					6																				
ironing	10	7								93																	
knock on the door	11						88									13											
lassou	12	6					94																				
light feather	13	60								40																	
rainbow	14						100																				
rock a baby	15	20								20					10	50											
rock guitar	16	47					6			6					18	24											
scratch your knee	17														100												
shake hands	18						73								27												
shave	19	94					6																				
spank	20														40	60											
spider	21						100																				
stroke the cat	22						17			8					75												
surrender	23	7					93																				
whistle	24	15					85																				
umbrella	25	7					80			13																	
violin	26						100																				

Number of Epochs = 200
 No of gestures recognised at or above 80% = 3
 Average rec. rate of best 3 gestures = 95.05%
 Average recognition rate of all gestures = 12.26%

cfmname: cte200

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	10					40			10					10	30											
cards	2	100																									
cut throat	3	100																									
drive the car	4						50			10					20	20											
drums	5	20								30					10	40											
heavy weight	6	40								10					50												
helicopter	7						90			10																	
hot	8	70								20					10												
ice-cream	9	100																									
ironing	10	20								80																	
knock on the door	11	20					80																				
lassou	12						100																				
light feather	13	60								40																	
rainbow	14						100																				
rock a baby	15	60													40												
rock guitar	16	20								10					70												
scratch your knee	17														100												
shake hands	18						70								20	10											
shave	19	100																									
spank	20														70	30											
spider	21						100																				
stroke the cat	22						40								60												
surrender	23						100																				
whistle	24						100																				
umbrella	25	20					70			10																	
violin	26						100																				

Number of Epochs = 200
 No of gestures recognised at or above 80% = 3
 Average rec. rate of best 3 gestures = 90.00%
 Average recognition rate of all gestures = 13.08%

Appendix C.2 Comparison of Activation Functions

ex63uk
cfmname: ctr9600

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1																									100
cards	2								100																	
cut throat	3		80						20																	
drive the car	4								38	8																
drums	5								63																38	
heavy weight	6		6						69	13							6							6		
helicopter	7								46																54	
hot	8								94																6	
ice-cream	9		44						56																	
ironing	10								7	86															7	
knock on the door	11								63																38	
lassou	12								69																31	
light feather	13								87	13																
rainbow	14								83	8								8								
rock a baby	15								100																	
rock guitar	16								82	12								6								
scratch your knee	17																	100								
shake hands	18								82									9							9	
shave	19		53						47																	
spank	20								80									20								
spider	21								100																	
stroke the cat	22								50	8								42								
surrender	23																								100	
whistle	24								23																77	
umbrella	25								13																87	
violin	26								92	8																

Number of Epochs = 9600
 No of gestures recognised at or above 80% = 4
 Average rec. rate of best 4 gestures = 91.43%
 Average recognition rate of all gestures = 16.23%

cfmname: cte9600

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1								10																	90
cards	2		10						90																	
cut throat	3		100																							
drive the car	4								50	10															40	
drums	5								50	30															20	
heavy weight	6		10						80								10									
helicopter	7								20																80	
hot	8								80	10															10	
ice-cream	9		20						80																	
ironing	10								20	70															10	
knock on the door	11								50																50	
lassou	12								90																10	
light feather	13								100																	
rainbow	14								90																10	
rock a baby	15		10						90																	
rock guitar	16								100																	
scratch your knee	17																	100								
shake hands	18								90								10									
shave	19		60						40																	
spank	20								90								10									
spider	21		10						80																10	
stroke the cat	22								90								10									
surrender	23								10	10															80	
whistle	24								20																80	
umbrella	25								40																60	
violin	26								90	10																

Number of Epochs = 9600
 No of gestures recognised at or above 80% = 4
 Average rec. rate of best 4 gestures = 90.00%
 Average recognition rate of all gestures = 16.54%

Appendix C.2 Comparison of Activation Functions

ex64uk
cfmname: ctr400

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	14			14	29			36			7														
cards	2						9		64	27																
cut throat	3			47						27											27					
drive the car	4	8			8	46					8			8					8		15					
drums	5					69					13			6		6								6		
heavy weight	6						100																			
helicopter	7						8						15		8									8		62
hot	8			6							94															
ice-cream	9			6					6	19											63	6				
ironing	10							21				79														
knock on the door	11						31			6		25	25		6										6	
lassou	12											50												6		25
light feather	13							53		7		13			27											
rainbow	14														100											
rock a baby	15					10	30				30						10				10	10				
rock guitar	16			6		6				41		6					12	6				24				
scratch your knee	17																	100								
shake hands	18						18					9			36			9	18			9				
shave	19											6									94					
spank	20						7												7						87	
spider	21															33							8			42
stroke the cat	22										8				8			50	8		17		8			
surrender	23														7										50	
whistle	24										8	8	8		15											62
umbrella	25														7									7		87
violin	26												25		8					8						58

Number of Epochs = 400
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 94.51%
 Average recognition rate of all gestures = 44.34%

cfmname: cte400

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	20				30			30			20														
cards	2								50	30	10									10						
cut throat	3			60																40						
drive the car	4				20	40					10			10												10
drums	5	10				10	60				20															
heavy weight	6									90					10											
helicopter	7						20								30								20		20	10
hot	8									80	10	10														
ice-cream	9										10	60									30					
ironing	10							20				80														
knock on the door	11									30		30	20											20		
lassou	12												40											10		40
light feather	13							60		40																
rainbow	14														100											
rock a baby	15			10							20	10						20	40							
rock guitar	16			10			10												40			10	10			
scratch your knee	17																				100					
shake hands	18						10										10				20	40		20		
shave	19			20																		80				
spank	20							10												10						80
spider	21															10										60
stroke the cat	22																		30	10			30			
surrender	23																							40		50
whistle	24											10														60
umbrella	25																							10		70
violin	26												40		10											30

Number of Epochs = 400
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 87.14%
 Average recognition rate of all gestures = 45.00%

Appendix C.2 Comparison of Activation Functions

ex65uk
cfmname: ctrl000

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	64			7	29																				
cards	2			9		36						18							36							
cut throat	3			27															73							
drive the car	4	23			38	8				8									15			8				
drums	5	38			38					13						6					6					
heavy weight	6					88				13																
helicopter	7										38							8	23					8	23	
hot	8			6		6	71	12				6														
ice-cream	9			13		6	6					6						69								
ironing	10	7				7				86																
knock on the door	11	13			25	6	6			19	6	6							6		6	6	6	6		
lassou	12										69										6		13	6		6
light feather	13			7		53						40														6
rainbow	14											100														
rock a baby	15					60				10					10	10			10							
rock guitar	16				12	6	18							41	6			18								
scratch your knee	17														100											
shake hands	18			9	9									9		18	18		18		18					
shave	19			6		12												82								
spank	20															7					93					
spider	21											8	33				8				50					
stroke the cat	22												17				8					75				
surrender	23											7										50			43	
whistle	24						8	8			23													46	15	
umbrella	25											7														93
violin	26											33					8			8	17					8 25

Number of Epochs = 1000
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 91.75%
 Average recognition rate of all gestures = 50.49%

cfmname: ctrl000

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	50			10	30																			10	
cards	2			10		30		20					10						30							
cut throat	3			10															90							
drive the car	4			10	50			10	10						10	10										
drums	5	20			10	40	10		10			10														
heavy weight	6					100																				
helicopter	7					20		10		10	20												20	10		10
hot	8	20					60						20													
ice-cream	9					10													90							
ironing	10					50			40			10														
knock on the door	11	10				20		10			10	30											10			10
lassou	12											80									10			10		
light feather	13					60							40													
rainbow	14													100												
rock a baby	15			20		50		20				10														
rock guitar	16	10				10	10		30	10						20										
scratch your knee	17																									
shake hands	18					10						10	10		10		30	10		20						
shave	19							10											90							
spank	20	10										10									80					
spider	21											20										40				20
stroke the cat	22												10				10					80				
surrender	23											10											30			60
whistle	24	10								10	10														60	10
umbrella	25													20												80
violin	26										10	50					10			30						

Number of Epochs = 1000
 No of gestures recognised at or above 80% = 8
 Average rec. rate of best 8 gestures = 86.25%
 Average recognition rate of all gestures = 42.69%

C.3 Confusion Matrices: k-Nearest Neighbours

Appendix C.3 Confusion Matrices: k-Nearest Neigh-

Experiment ex49uk
cfmname: cte1

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	50			40					10																
cards	2		50	10					30							10										
cut throat	3			70					20										10							
drive the car	4				30	20					10				10		10							10		
drums	5					30	30				10		10			10					10					
heavy weight	6						60									40										
helicopter	7							10			40	20	40											10	10	10
hot	8		30						50								20									
ice-cream	9		10	20						40										30						
ironing	10			20							60					20										
knock on the door	11	30	10								30	10														20
lassou	12							10				60										10				20
light feather	13		30					20					50													
rainbow	14													70								10				
rock a baby	15		10						10	20		10			20	20			10							
rock guitar	16					10			10						10	50			20							
scratch your knee	17															100				10						
shake hands	18				10									10			10	60			10					
shave	19			60						10								20		30						
spank	20																		20		80					
spider	21													10								60				30
stroke the cat	22																10				10		80			
surrender	23							10																40	50	
whistle	24							10				20													60	10
umbrella	25																						10	10	70	10
violin	26											10							40							50

k = 1

No of gestures recognised at or above 80% = 3
Average rec. rate of best 3 gestures = 86.67%
Average recognition rate of all gestures = 52.31%

cfmname: cte2

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	70			10	10					10															
cards	2		70	20						10																
cut throat	3			80						10										10						
drive the car	4	20			30	20					10				20											
drums	5	20	10			40	20									10										
heavy weight	6						80								20											
helicopter	7				10			30			40	10												10		
hot	8	10	30						50							10										
ice-cream	9		30	50						20																
ironing	10			20			10				70															
knock on the door	11	30	10		10			10			20	10													10	
lassou	12							20				50										10			10	10
light feather	13		50					40					10													
rainbow	14													80					10			10				
rock a baby	15		20						20	10		10			20	20										
rock guitar	16					10			20	20					20	30										
scratch your knee	17															100										
shake hands	18				10	10								10			10	60								
shave	19			80						10								10								
spank	20															10		30		60						
spider	21											10		10							50					30
stroke the cat	22															20	10				10		60			
surrender	23							10				20												40	10	20
whistle	24	10						20				50	10												10	
umbrella	25																					20		20	20	40
violin	26											30	20							20						30

k = 2

No of gestures recognised at or above 80% = 4
Average rec. rate of best 4 gestures = 85.00%
Average recognition rate of all gestures = 45.77%

cfmname: cte3

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	70			10	10		10																		
cards	2		70	10						20																
cut throat	3			90																10						
drive the car	4	20			50	10					10					10										
drums	5	40	20			20																				
heavy weight	6						70								30											
helicopter	7				20			30				30	20													
hot	8	10	20						50					10		10										
ice-cream	9		20	40						20											20					
ironing	10				10	10					50			30												
knock on the door	11	30	10					10				30	20													
lassou	12							10					40									10				
light feather	13		50					40						10												
rainbow	14														60							40				
rock a baby	15		30	10						20		10				30										
rock guitar	16								20							10	60			10						
scratch your knee	17																100									
shake hands	18				10	10								10			10	60								
shave	19			60															40							
spank	20																	10		90						
spider	21																					70				30
stroke the cat	22															10	10						80			
surrender	23							10																40		50
whistle	24	30						30				20													20	
umbrella	25															10								30		50
violin	26				10								30						20							40

k = 3

No of gestures recognised at or above 80% = 4
Average rec. rate of best 4 gestures = 90.00%
Average recognition rate of all gestures = 51.54%

Appendix C.3 Confusion Matrices: k-Nearest Neigh-

cfmname: cte4

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	70			10	10						10														
cards	2		70	10							20															
cut throat	3			70							10									20						
drive the car	4	20			50	10					10						10									
drums	5	30	10			20	30													10						
heavy weight	6						90								10											
helicopter	7				10			40			40												10			
hot	8		30						50							10	10									
ice-cream	9			30	40					10									20							
ironing	10										60			20		20										
knock on the door	11	30						10			40	10											10			
lassou	12								10			40	10									20			10	20
light feather	13		50					40					10													
rainbow	14													80							20					
rock a baby	15		20	10	10					40					20											
rock guitar	16								20	10							70									
scratch your knee	17																	100								
shake hands	18				10	10									10			10	60							
shave	19			60																40						
spank	20																	10		90						
spider	21																				70					30
stroke the cat	22															10	10					80				
surrender	23							10															50		40	
whistle	24	20						10				20											10	30	10	
umbrella	25																				20		20		60	
violin	26				10					10	30	10							10							30

k = 4
 No of gestures recognised at or above 80% = 5
 Average rec. rate of best 5 gestures = 88.00%
 Average recognition rate of all gestures = 54.23%

cfmname: cte5

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	70			20						10															
cards	2		80	20																						
cut throat	3			60						20									20							
drive the car	4	10			50	20					10					10										
drums	5	20	20		20	30														10						
heavy weight	6						80						20													
helicopter	7				10			40			40												10			
hot	8	10	40						40						10											
ice-cream	9		30	50						10									10							
ironing	10							10			60			20		10										
knock on the door	11	40									50												10			
lassou	12											40										20		10		20
light feather	13		40				40			10			10													
rainbow	14													70							30					
rock a baby	15		20	10	10					40					20											
rock guitar	16		10						10	20						60										
scratch your knee	17																	100								
shake hands	18				10									10			10	70								
shave	19			30							20							10	50							
spank	20																		10		90					
spider	21																					70				30
stroke the cat	22															10	10						80			
surrender	23							10																40		50
whistle	24	20						10				10												10	50	
umbrella	25																					20		30		50
violin	26									10		20	10						20							40

k = 5
 No of gestures recognised at or above 80% = 5
 Average rec. rate of best 5 gestures = 86.00%
 Average recognition rate of all gestures = 54.23%

cfmname: cte6

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	60			30						10															
cards	2		90	10																						
cut throat	3			60															40							
drive the car	4	30			40	10					10					10										
drums	5	20	20		20	20											10			10						
heavy weight	6						70						30													
helicopter	7				10			20			30	10												10		10
hot	8	10	20						50					10		10										
ice-cream	9		30	50						10										10						
ironing	10										70			20		10										
knock on the door	11	20				10					50												10		10	
lassou	12								10			30										10		10		30
light feather	13		40				40			10			10													
rainbow	14													70								30				
rock a baby	15		10	10	10						50				20											
rock guitar	16								10	20						10	50			10						
scratch your knee	17																	100								
shake hands	18				10									10			10	70								
shave	19			50															10	50						
spank	20																			10		90				
spider	21											10											60			30
stroke the cat	22																10				10		80			
surrender	23							10																40		50
whistle	24	20						10				20												10	40	
umbrella	25																						10		70	
violin	26									10		20	10						20							40

k = 6
 No of gestures recognised at or above 80% = 4
 Average rec. rate of best 4 gestures = 90.00%
 Average recognition rate of all gestures = 52.31%

Appendix C.3 Confusion Matrices: k-Nearest Neigh-

cfmname: cte7

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	70			20																							
cards	2		90	10																								
cut throat	3			60																40								
drive the car	4	20			50	20																						
drums	5	20	20		30	10																						
heavy weight	6					80																						
helicopter	7						20						40	20										10	10			
hot	8		20			10				50				10	10													
ice-cream	9		30	50					20																			
ironing	10						10					60			20	10												
knock on the door	11	10				10							50												10		20	
lassou	12							10						30											20	10		30
light feather	13		50					30							20													
rainbow	14																70											
rock a baby	15		20	10							40						20	10										
rock guitar	16					10				10	20							50										
scratch your knee	17																		100									
shake hands	18				10												10	80										
shave	19			40															60									
spank	20																	10		90								
spider	21																				60						30	10
stroke the cat	22																	10			10					80		
surrender	23							10																		50		40
whistle	24	20						10					10												10	60		
umbrella	25													10											10		20	60
violin	26								10			10	10															50

k = 7
 No of gestures recognised at or above 80% = 6
 Average rec. rate of best 6 gestures = 86.67%
 Average recognition rate of all gestures = 54.62%

cfmname: cte8

Gesture:	Network Decision																												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26			
bird	1	70			20																								
cards	2		90	10																									
cut throat	3			60																40									
drive the car	4	10			60	20							10																
drums	5	20	20		20	20																							
heavy weight	6						80									20													
helicopter	7							20						40	10														
hot	8		20			10			50						10		10												
ice-cream	9		30	40						30																			
ironing	10				10																								
knock on the door	11	20						10					50												20				
lassou	12							10						20											20	10	10	30	
light feather	13		50					40	10						10														
rainbow	14																70										30		
rock a baby	15		20	10						40								20	10										
rock guitar	16					10			10	20									50								10		
scratch your knee	17																			100									
shake hands	18				10													10	80										
shave	19				50															10	50								
spank	20																				10						90		
spider	21																					60					30	10	
stroke the cat	22																		10				10				80		
surrender	23							10																			50		40
whistle	24	10						10					10													10	60		
umbrella	25													10												10		20	60
violin	26				10							10	10															50	

k = 8
 No of gestures recognised at or above 80% = 6
 Average rec. rate of best 6 gestures = 86.67%
 Average recognition rate of all gestures = 55.77%

cfmname: cte9

Gesture:	Network Decision																														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26					
bird	1	70			10	10																									
cards	2		80	20																											
cut throat	3			60																	40										
drive the car	4				60	30																									
drums	5	30	20		10	20																									
heavy weight	6						80																								
helicopter	7							20																							
hot	8		30			10			50																						
ice-cream	9		30	50						10																					
ironing	10												70																		
knock on the door	11	20						10																							
lassou	12							10						20																	
light feather	13		50					40	10																						
rainbow	14																	70										30			
rock a baby	15		20	10						40																					
rock guitar	16					10				10	10																				
scratch your knee	17																														
shake hands	18				10																10	80									
shave	19				30																	10	70								
spank	20																						10					90			
spider	21																							60				30	10		
stroke the cat	22																											80			
surrender	23							10																				60		40	
whistle	24	10						10						10														10	60		
umbrella	25														10													10		20	60
violin	26											20	10																	50	

k = 9
 No of gestures recognised at or above 80% = 6
 Average rec. rate of best 6 gestures = 85.00%
 Average recognition rate of all gestures = 55.38%
 Experiment ex50uk

C.4 Confusion matrices: k-Nearest Neighbour with Standard Normalisation

Appendix C.4 Confusion matrices: k-Nearest Neigh-

cfmname: cte1

Gesture:	Network Decision																												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26			
bird	1	50																								10			
cards	2		50	10															10										
cut throat	3			50																30									
drive the car	4				50	10										10			10							10			
drums	5	10				30	20								10		10				10								
heavy weight	6						60									40													
helicopter	7							20						30	20									10	10			10	
hot	8		10						70											20									
ice-cream	9			10	20																								
ironing	10				20								60																
knock on the door	11	30								10																		20	
lassou	12										10															60		10	
light feather	13			40																						30			
rainbow	14																										80		
rock a baby	15				10						10	20				10										20	30		
rock guitar	16								10											10	50							20	
scratch your knee	17																				100								
shake hands	18								10												10	60						10	
shave	19									70																		20	
spank	20																										10	20	
spider	21																										60	70	
stroke the cat	22																										10	80	
surrender	23										10															30	20	40	
whistle	24																											50	10
umbrella	25																											10	80
violin	26												10	10														40	

k = 1
 No of gestures recognised at or above 80% = 4
 Average rec. rate of best 4 gestures = 85.00%
 Average recognition rate of all gestures = 51.15%

cfmname: cte2

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	70			10	10																						
cards	2		80	10																								
cut throat	3			80																								
drive the car	4	30			30	20																						
drums	5	20	10			30	30																					
heavy weight	6							80								20												
helicopter	7								20							40	30											10
hot	8	10	40							40																		
ice-cream	9			30	50						20																	
ironing	10				20							80																
knock on the door	11	30								10	10																	
lassou	12											10	60														10	
light feather	13				50																						10	
rainbow	14																											
rock a baby	15			30	10	10					10	10																
rock guitar	16																20	40										
scratch your knee	17																											
shake hands	18																											
shave	19																											
spank	20																											
spider	21																											
stroke the cat	22																											
surrender	23																											
whistle	24	10																										
umbrella	25																											
violin	26																											

k = 2
 No of gestures recognised at or above 80% = 6
 Average rec. rate of best 6 gestures = 83.33%
 Average recognition rate of all gestures = 46.15%

cfmname: cte3

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	70			10	10																						
cards	2		70	20																								
cut throat	3			90																								
drive the car	4	30			40	10																						
drums	5	30	10			30	20																					
heavy weight	6							70																				
helicopter	7								30																			
hot	8	10	20							50																		
ice-cream	9			30	30						30																	
ironing	10				10																							
knock on the door	11	30																										
lassou	12																											
light feather	13																											
rainbow	14																											
rock a baby	15			30	10	10																						
rock guitar	16																											
scratch your knee	17																											
shake hands	18																											
shave	19																											
spank	20																											
spider	21																											
stroke the cat	22																											
surrender	23																											
whistle	24	40																										
umbrella	25																											
violin	26																											

k = 3
 No of gestures recognised at or above 80% = 2
 Average rec. rate of best 2 gestures = 95.00%
 Average recognition rate of all gestures = 50.38%

C.5 Confusion Matrices: Gesture Segment Length

Gesture Segment Length: 160ms
Time frame: 1
ANN architecture: 3-16t-26p

(a) Training Set

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	71								7																
cards	2					27		45	27																	
cut throat	3							7	87											7						
drive the car	4	15			23						8								23		31					
drums	5	19			19			6		31							6		13							
heavy weight	6					100																				
helicopter	7	23						8				31									8		8			23
hot	8						12		82		6															
ice-cream	9								6	81											6	6				
ironing	10						21				79															
knock on the door	11	31								13	6	25		13												13
lassou	12	6						6			6	31							13							25
light feather	13						73		13		13															
rainbow	14													75				8					17			
rock a baby	15							20		10	60						10									
rock guitar	16								18	12	41						12						6			
scratch your knee	17																	100								
shake hands	18					9								64		9		9		9						
shave	19										82								18							
spank	20										7			7		7					67		13			
spider	21																				42					17
stroke the cat	22														8		17						75			
surrender	23	7			21							71														
whistle	24	23									8	8	31													31
umbrella	25	7											27									7				60
violin	26											8	17	8							8					58

Number of Epochs = 300

No of gestures recognised at or above 80% = 4
Average rec. rate of best 4 gestures = 90.90%
Average recognition rate of all gestures = 35.94%

(b) Test Set

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	60			10								20													10
cards	2								30	50	10					10										
cut throat	3									50									50							
drive the car	4				10						30			10		30					20					
drums	5				20	10		10			50										10					
heavy weight	6					100																				
helicopter	7	20						10					30										10			30
hot	8							80		10							10									
ice-cream	9								10	90																
ironing	10					30					70															
knock on the door	11	30			10			20				20										10				10
lassou	12							10				10	30										10			40
light feather	13					60		40																		
rainbow	14														90								10			
rock a baby	15					20		10	30	30							10									
rock guitar	16					10		10	20	30							30									
scratch your knee	17																		100							
shake hands	18									10				40			10	10		30						
shave	19									80									20							
spank	20																					80				
spider	21														20											40
stroke the cat	22																						90			
surrender	23											60									10					30
whistle	24	10						10				10	40													30
umbrella	25												50									10				20
violin	26	10										10										40				40

Number of Epochs = 300

No of gestures recognised at or above 80% = 7
Average rec. rate of best 7 gestures = 90.00%
Average recognition rate of all gestures = 35.38%

Confusion Matrices for (a) training and (b) test sets for GSL=160ms

Appendix C.5 Confusion Matrices: Gesture Segment

Gesture Segment Length: 320ms
 Time frame: 2
 ANN architecture: 6-16t-26p
 (a) Training Set

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	64			29																					7
cards	2		27	9			9	9					36							9						
cut throat	3			80				7												13						
drive the car	4	8				69					8										15					
drums	5	6				63			13												13		6			
heavy weight	6						100																			
helicopter	7									8	15	8		8												62
hot	8						6	88						6												
ice-cream	9			50				25												19	6					
ironing	10						21				79															
knock on the door	11	19			6	19		13				25									6		6		6	
lassou	12							13				6		6								6				56
light feather	13						33		7	20			40													56
rainbow	14												100													
rock a baby	15						20			20				10	30				10	10						
rock guitar	16					18		24		6					24					24		6				
scratch your knee	17															100										
shake hands	18					18				9				73												
shave	19			88																12						
spank	20																13				87					
spider	21													75							17				8	
stroke the cat	22																8					92				
surrender	23					7																	7		86	
whistle	24							8			8	8	8		8									8	54	
umbrella	25													7												93
violin	26			8						8				8							8	17		8	33	8

Number of Epochs = 1000
 No of gestures recognised at or above 80% = 8
 Average rec. rate of best 8 gestures = 92.49%
 Average recognition rate of all gestures = 43.18%

(b) Test Set

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	30			50			10			10															
cards	2		10	10				20	10	10			20							20						
cut throat	3			90																10						
drive the car	4				10	40					20										20		10			
drums	5					50			10		20			10								10				
heavy weight	6						80				10			10												
helicopter	7	20						10				10												20	40	
hot	8		10						80																	
ice-cream	9			40					10											50						
ironing	10										90			10												
knock on the door	11	10							10			30		10											40	
lassou	12								10															20	60	10
light feather	13		10				30		10					50												
rainbow	14														100											
rock a baby	15			10		10	30		30					10						10						
rock guitar	16					30	10		10	10						30					10					
scratch your knee	17																	100								
shake hands	18						20							40			10				30					
shave	19				90															10						
spank	20					10															80		10			
spider	21													20								30				50
stroke the cat	22																						90			
surrender	23													10												90
whistle	24																									100
umbrella	25													30												70
violin	26							10				20									30					10

Number of Epochs = 1000
 No of gestures recognised at or above 80% = 8
 Average rec. rate of best 8 gestures = 88.75%
 Average recognition rate of all gestures = 41.15%

Confusion Matrices for (a) training and (b) test sets for GSL=320ms

Appendix C.5 Confusion Matrices: Gesture Segment

Gesture Segment Length: 480ms
 Time frame: 3
 ANN architecture: 9-16t-26p

(a) Training Set

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	86						7																	7	
cards	2		9					27	9				45						9							
cut throat	3			20				7	7										67							
drive the car	4	46			8	31														15						
drums	5	31				38					19	6						6								
heavy weight	6	6					88							6												
helicopter	7			8								31						8						23		31
hot	8							94												6						
ice-cream	9			6					19					6					63	6						
ironing	10								7		93															
knock on the door	11	6				38					19	19								13					6	
lassou	12											81												13		6
light feather	13						27						73													
rainbow	14													58			8				17					17
rock a baby	15					20	30				20					10				20						
rock guitar	16				6	29			29		6					12	6		6	6						
scratch your knee	17																	100								
shake hands	18					9											18	18		18			36			
shave	19																		100							
spank	20					13										7	7		53		20					
spider	21											25		17				8			42					8
stroke the cat	22																17			8		75				
surrender	23																						57			43
whistle	24	8						8			8	8											8	23		38
umbrella	25																							7		93
violin	26											50								8		25				8

Number of Epochs = 2300
 No of gestures recognised at or above 80% = 8
 Average rec. rate of best 8 gestures = 91.85%
 Average recognition rate of all gestures = 47.68%

(b) Test Set

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	80			10																					10
cards	2		10					30					30						30							
cut throat	3			40															60							
drive the car	4	10			10	40						20					10								10	
drums	5	20				30			20		10						10								10	
heavy weight	6						90							10												
helicopter	7	10			10	10						10	30												10	20
hot	8	10				10			80																	
ice-cream	9								10											90						
ironing	10						20				30		40				10									
knock on the door	11		10				20			10		10	30											20		
lassou	12												90												10	
light feather	13						30		20				50													
rainbow	14													80							10					10
rock a baby	15						20		10				10							60						10
rock guitar	16			10		40			10				10							30						
scratch your knee	17																	90								10
shake hands	18					10						10	20					10	10							40
shave	19			10																90						
spank	20										10										50			40		
spider	21											10		10								50				30
stroke the cat	22																	10					90			
surrender	23																							50		50
whistle	24	10																						10	40	40
umbrella	25																							10		80
violin	26												90													10

Number of Epochs = 2300
 No of gestures recognised at or above 80% = 9
 Average rec. rate of best 9 gestures = 85.56%
 Average recognition rate of all gestures = 44.23%

Confusion Matrices for (a) training and (b) test sets for GSL=480ms

Appendix C.5 Confusion Matrices: Gesture Segment

Gesture Segment Length: 640ms
 Time frame: 4
 ANN architecture: 12-16t-26p

(a) Training Set

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	86								7															7	
cards	2						9		91																	
cut throat	3								13																	
drive the car	4	38			31						8									23						
drums	5	63				19			6			6									6					
heavy weight	6						94				6															
helicopter	7											31							8					31	15	8
hot	8								100																	
ice-cream	9								25											69						6
ironing	10								7		93															
knock on the door	11	19			6	6					25	6							6					25		6
lassou	12											81										6			6	6
light feather	13						33		7		27		33													
rainbow	14													32								8				
rock a baby	15						30		10		20				10	10				10	10					
rock guitar	16	6				12				53							12					6				
scratch your knee	17																100									
shake hands	18											9							91							
shave	19									6										94						
spank	20																		7		93					
spider	21											8		8								83				
stroke the cat	22																					100				
surrender	23					7																	7	29	57	
whistle	24																							92	8	
umbrella	25													7										13	80	
violin	26											25										8		8		58

Number of Epochs = 6700
 No of gestures recognised at or above 80% = 14
 Average rec. rate of best 14 gestures = 91.37%
 Average recognition rate of all gestures = 56.71%

(b) Test Set

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	100																								
cards	2								80		10										10					
cut throat	3																				100					
drive the car	4	30			10				10			10								10					30	
drums	5	30			10	20			20		10													10		
heavy weight	6						80						20													
helicopter	7	20										20												60		
hot	8								100																	
ice-cream	9								30												70					
ironing	10										80			20												
knock on the door	11	20							10				40											30		
lassou	12												80									10			10	
light feather	13						30		50				20													
rainbow	14													90							10					
rock a baby	15						10		50				10		10					20						
rock guitar	16					10			60						10					10	10					
scratch your knee	17																		80			10		10		
shake hands	18																		10	90						
shave	19																			100						
spank	20																		10		90					
spider	21											10		10							60			10	10	
stroke the cat	22																		10			90				
surrender	23	10																				10			20	60
whistle	24																								100	
umbrella	25																							10	70	
violin	26												20							30			20		10	20

Number of Epochs = 6700
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 90.00%
 Average recognition rate of all gestures = 50.00%

Confusion Matrices for (a) training and (b) test sets for GSL=640ms

Appendix C.5 Confusion Matrices: Gesture Segment

Gesture Segment Length: 800ms
Time frame: 5
ANN architecture: 15-16t-26p

(a) Training Set

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	79			21																						
cards	2		18			9	18					45							9								
cut throat	3		13																87								
drive the car	4	31		8	15					8			8					8		23							
drums	5	50			31											6				13							
heavy weight	6					94				6																	
helicopter	7		23								23													23	31		
hot	8					6	88												6								
ice-cream	9		19				6											69	6								
ironing	10						7	93																			
knock on the door	11	13			13		6				19	6							19	6			19				
lassou	12										75									6			13	6			
light feather	13					40	7		13		40																
rainbow	14											92								8							
rock a baby	15					50		10	20																		
rock guitar	16	6			12		24		6						12				29		6		6				
scratch your knee	17																100										
shake hands	18				9								27					45	18								
shave	19																	100									
spank	20																		100								
spider	21																			100							
stroke the cat	22								8			17									75						
surrender	23									7												14	14	64			
whistle	24	8					8				8	8												62	8		
umbrella	25										13	7								13					67		
violin	26										50									25					17	8	

Number of Epochs = 2600
No of gestures recognised at or above 80% = 8
Average rec. rate of best 8 gestures = 95.81%
Average recognition rate of all gestures = 49.82%

(b) Test Set

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80			20																						
cards	2		30				30					30							10								
cut throat	3																		100								
drive the car	4	10			40		10	10	10			10							10								
drums	5	20			20			20	10										20			10					
heavy weight	6					80		10				10															
helicopter	7	30						10				10								10			20	20			
hot	8							80	10			10															
ice-cream	9		30					10											60								
ironing	10						10			70		20															
knock on the door	11	30					20					20								10			20				
lassou	12											80								10	10						
light feather	13		10			30	30					30															
rainbow	14												70							30							
rock a baby	15					30	20	10									10		20							10	
rock guitar	16			10		10	20	30											10	20							
scratch your knee	17																		100								
shake hands	18												10			10	40		40								
shave	19																		100								
spank	20																			100							
spider	21											10									90						
stroke the cat	22												10			10						80					
surrender	23										10										10					70	
whistle	24	10																						80	10		
umbrella	25											20	10								20				10	40	
violin	26											40	10							10	10	10				20	

Number of Epochs = 2600
No of gestures recognised at or above 80% = 10
Average rec. rate of best 10 gestures = 87.00%
Average recognition rate of all gestures = 43.85%

Confusion Matrices for (a) training and (b) test sets for GSL=800ms

Appendix C.5 Confusion Matrices: Gesture Segment

Gesture Segment Length: 960ms
 Time frame: 6
 ANN architecture: 18-16t-26p

(a) Training Set

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	86				7																				7	
cards	2		45								9			27							18						
cut throat	3			13																87							
drive the car	4	46			15	8												8		8		15					
drums	5	44				31																19				6	
heavy weight	6						94				6																
helicopter	7										8	31												15	31	8	8
hot	8					6		94																			
ice-cream	9		6						6	6					6						75						
ironing	10					14			7		79																
knock on the door	11	6				6					25	31							6			6				19	
lassou	12											88										6		6			
light feather	13		7				33		7		20			33													
rainbow	14														100												
rock a baby	15						30		10		10					20	20				10						
rock guitar	16								24										53	6		6	12				
scratch your knee	17																		100								
shake hands	18											9		9						64		9		9			
shave	19			6																	94						
spank	20																					93		7			
spider	21																						100				
stroke the cat	22																							100			
surrender	23																								100		
whistle	24																									92	8
umbrella	25											7													20	7	67
violin	26												58												25		8

Number of Epochs = 6800
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 95.07%
 Average recognition rate of all gestures = 61.58%

(b) Test Set

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	90				10																					
cards	2		10						20	10	10			20								30					
cut throat	3																					100					
drive the car	4	10				30					10		10													10	
drums	5	40				20										20											
heavy weight	6						80		10					10													
helicopter	7	20						10					30													10	30
hot	8								70		10						20										
ice-cream	9		10							10												80					
ironing	10										80			20													
knock on the door	11	20								10				40											10	10	10
lassou	12											10	80														10
light feather	13						40		20					40													
rainbow	14														80												
rock a baby	15						20		30				10			20	10				10						
rock guitar	16					10			30						10	20					10	20					
scratch your knee	17																	100									
shake hands	18					10													10	50		20	10				
shave	19																				100						
spank	20																					100					
spider	21											10											70				20
stroke the cat	22																	10						90			
surrender	23																							70		30	
whistle	24																									100	
umbrella	25																								20	30	50
violin	26												90														

Number of Epochs = 6800
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 90.00%
 Average recognition rate of all gestures = 51.54%

Confusion Matrices for (a) training and (b) test sets for GSL=960ms

Gesture Segment Length: 1120ms
 Time frame: 7
 ANN architecture: 21-16t-26p

(a) Training Set

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	93				7																						
cards	2		18				9		27				36							9								
cut throat	3			27																73								
drive the car	4	23			54											15					8							
drums	5	38			13	44															6							
heavy weight	6						94				6																	
helicopter	7											23										15		23	15	15	8	
hot	8						6		94																			
ice-cream	9			19						25					6						50							
ironing	10								7		93																	
knock on the door	11	6				6		6				13	44		6										13	6		
lassou	12												88									6		6				
light feather	13							33		13		7			47													
rainbow	14														100													
rock a baby	15								20			10				40	10											
rock guitar	16					6						35					41					12						
scratch your knee	17																	100										
shake hands	18						9	9						9		36						27			9			
shave	19					6																94						
spank	20																					100						
spider	21																						100					
stroke the cat	22																							100				
surrender	23																								86		14	
whistle	24	8								8																69	15	
umbrella	25																								13	7	80	
violin	26												8		8							33				17	8	25

Number of Epochs = 6700
 No of gestures recognised at or above 80% = 13
 Average rec. rate of best 13 gestures = 93.92%
 Average recognition rate of all gestures = 61.46%

(b) Test Set

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80				20																					
cards	2		10						40				20							30							
cut throat	3			20																80							
drive the car	4	10			10	30				10						10					10					20	
drums	5	30				30				20																	
heavy weight	6						80								20												
helicopter	7	10											40									10		20	20		
hot	8								20		10						10										
ice-cream	9									80										80							
ironing	10										40			60													
knock on the door	11	10							10				50												20	10	
lassou	12												80								10			10			
light feather	13		10				40		30					20													
rainbow	14														70							30					
rock a baby	15						20		30				10			10	10				20						
rock guitar	16						10	10		40							10				10	20					
scratch your knee	17																	90						10			
shake hands	18						10									40	10				30	10					
shave	19			10																	90						
spank	20																					100					
spider	21																						80		10		10
stroke the cat	22																	10						90			
surrender	23																								60		40
whistle	24																									80	20
umbrella	25																						20		40		40
violin	26												40				10					10	20				20

Number of Epochs = 6700
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 85.00%
 Average recognition rate of all gestures = 45.77%

Confusion Matrices for (a) training and (b) test sets for GSL=1120ms

C.6 Confusion Matrices: Forearm Orientation

Experiment ex45uk

cfmname: ctr7800

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	57			7	14										7									7	7		
cards	2			9			18		73																			
cut throat	3			27				7											67									
drive the car	4	8			15	23										8	8		8	23					8			
drums	5	19				56				6							13								6			
heavy weight	6						94			6																		
helicopter	7											8						8						23	31	31		
hot	8					6		88		6																		
ice-cream	9		19					25												50	6							
ironing	10					14			79										7									
knock on the door	11	13			13					25	13		6						6			6			19			
lassou	12										63							6				6		6	13		6	
light feather	13					33		20		7						7				33								
rainbow	14												100															
rock a baby	15					10		20		30						40												
rock guitar	16							53		12						24									12			
scratch your knee	17																92						8					
shake hands	18			9	9					18								9						27		27		
shave	19			6			12												82									
spank	20				7												7				73		7	7				
spider	21													8									75	8		8		
stroke the cat	22									8													92					
surrender	23																							43	7	50		
whistle	24									8															77	15		
umbrella	25													13											27		60	
violin	26											17		8								8	8		25		8	25

Number of Epochs = 7800
 No of gestures recognised at or above 80% = 6
 Average rec. rate of best 6 gestures = 91.28%
 Average recognition rate of all gestures = 48.27%
 cfmname: cte7800

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	50			30																				20		
cards	2					10		60	10											20							
cut throat	3			30															70								
drive the car	4	10			10	10										10				20		10			30		
drums	5	40				10	10			20						10				10							
heavy weight	6						80				10									10							
helicopter	7	10									10	20												10	40	10	
hot	8								90							10											
ice-cream	9							20		10									70								
ironing	10					10					80					10									20	10	
knock on the door	11	30						10			10	20													10	10	
lassou	12											70										10				10	10
light feather	13						40		30	10									20								
rainbow	14													90							10						
rock a baby	15										100																
rock guitar	16										60					40											
scratch your knee	17																90						10				
shake hands	18										30	30		10				10							20		
shave	19				20															80							
spank	20											10									60		10		10		
spider	21																					70		10		10	10
stroke the cat	22																						90	10			
surrender	23																							30	10	60	
whistle	24													10										10	80		
umbrella	25														20										10	70	
violin	26											40						10						40			10

Number of Epochs = 7800
 No of gestures recognised at or above 80% = 8
 Average rec. rate of best 8 gestures = 85.00%
 Average recognition rate of all gestures = 43.85%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Experiment ex46uk
cfmname: ctr7800

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	86						7																	7		
cards	2			45					55																		
cut throat	3			100																							
drive the car	4	31			23	8					8	8				23											
drums	5	63				19					6					13											
heavy weight	6						94		6																		
helicopter	7									15		46				8							8	8	15		
hot	8	6		6		6		82																			
ice-cream	9			56		13		19		13																	
ironing	10			7		14				79																	
knock on the door	11	38			6	13				13		6		6									13		6		
lassou	12	13								6		69						6						6			
light feather	13					60		33		7																	
rainbow	14													100													
rock a baby	15	10								20						70											
rock guitar	16								12							82		6									
scratch your knee	17																83								17		
shake hands	18					9				9		9						55		9		9					
shave	19				71			18		6							6										
spank	20	13				13													7		47		20				
spider	21											17		58			8					17					
stroke the cat	22									8							25							67			
surrender	23									14		21		7		7								50			
whistle	24									8		8		8		8							8	8	54		
umbrella	25									27		13		7		13						27		13			
violin	26					8				17		17		25		8	25										

Number of Epochs = 7800
 No of gestures recognised at or above 80% = 7
 Average rec. rate of best 7 gestures = 89.64%
 Average recognition rate of all gestures = 42.50%
 cfmname: cta7800

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	90				10																					
cards	2			50		10		30		10																	
cut throat	3			100																							
drive the car	4	20		10	20		10			20						10									10		
drums	5	30		20	10	10										20								10			
heavy weight	6						70			10		10								10							
helicopter	7	40			20						10	10												10	10		
hot	8			30				50								20											
ice-cream	9			90			10																				
ironing	10								10		60		20											10			
knock on the door	11	20		10	10	10					10		20			10									10		
lassou	12						10					60		10		10		10									
light feather	13				30			40		20		10															
rainbow	14														100												
rock a baby	15								30		10		10				50										
rock guitar	16					10										90											
scratch your knee	17																100										
shake hands	18					10											10	50		20		10					
shave	19					80		10		10																	
spank	20	10					30										10	10		30		10					
spider	21											20		30		10						30			10		
stroke the cat	22											10					10						70				
surrender	23									10		40		10		30								10			
whistle	24										10			10											80		
umbrella	25										10		50		10			10						10	10		
violin	26	10										50				10	20	10									

Number of Epochs = 7800
 No of gestures recognised at or above 80% = 6
 Average rec. rate of best 6 gestures = 93.33%
 Average recognition rate of all gestures = 39.23%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Ex58uk
cfmname: ctr5200

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	86			7	7																				
cards	2		36	9				9		27	9										9					
cut throat	3			27																	73					
drive the car	4	46			46																8					
drums	5	38				44			6			6								6						
heavy weight	6						100																			
helicopter	7											31							8					23	15	23
hot	8									100																
ice-cream	9								6	19											75					
ironing	10						7				86				7											
knock on the door	11	19			6	13						19	19								6		6		13	
lassou	12											81										6		13		
light feather	13						33				7			60												
rainbow	14														92							8				
rock a baby	15									10		20					70									
rock guitar	16									18							76	6								
scratch your knee	17																92						8			
shake hands	18				9	18									9			45		9		9				
shave	19																		100							
spank	20																7				93					
spider	21											8		8								83				
stroke the cat	22										8												92			
surrender	23																							86		14
whistle	24																			8					77	15
umbrella	25														13										47	40
violin	26												25			8						33				33

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 90.84%
 Average recognition rate of all gestures = 62.03%

Ex58uk
cfmname: cts5200

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80				20																					
cards	2		20				10		20	40											10						
cut throat	3																				100						
drive the car	4	30			10	10			10			10					20										
drums	5	30				20			20							20					10						
heavy weight	6						80				10			10													
helicopter	7	20											40											20	10		
hot	8	10	10							70		10															
ice-cream	9			10							10										80						
ironing	10											90			10												
knock on the door	11	20											50											10	10		
lassou	12													70										10	10		
light feather	13							40		40					20												
rainbow	14															80						10	10				
rock a baby	15										10							90									
rock guitar	16										10								90								
scratch your knee	17																			100							
shake hands	18															10			10	40		30					
shave	19																				90						
spank	20																					10		80		10	
spider	21																						80		10		
stroke the cat	22																							80			
surrender	23																								70	20	
whistle	24																									10	80
umbrella	25																								10	30	10
violin	26																										40

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 84.55%
 Average recognition rate of all gestures = 51.15%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Experiment ex48uk
cfmname: ctr9900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	100																								
cards	2		18			9	55		9			9														
cut throat	3																			100						
drive the car	4	38		15	15														8		8				15	
drums	5	31		6	44					6							6			6						
heavy weight	6					100																				
helicopter	7											31												8	31	31
hot	8							100																		
ice-cream	9			25																69	6					
ironing	10					7	7	86																		
knock on the door	11	13			6					19	31		6							6	6				13	
lassou	12										81											6			13	
light feather	13					33	20	13				33														
rainbow	14												100													
rock a baby	15								10						30	60										
rock guitar	16							12								88										
scratch your knee	17																100									
shake hands	18					9												82		9						
shave	19																		100							
spank	20																			100						
spider	21										8		8										83			
stroke the cat	22																							100		
surrender	23																								93	7
whistle	24							8																	15	69
umbrella	25											7													53	40
violin	26								8		8	42		8										33		

Number of Epochs = 9900
 No of gestures recognised at or above 80% = 14
 Average rec. rate of best 14 gestures = 93.80%
 Average recognition rate of all gestures = 60.14%
 cfmname: cte9900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	90			10																					
cards	2		30					20	10				20							20						
cut throat	3																			100						
drive the car	4	10		20				10	10								20				20				10	
drums	5	30			20			20								20						10				
heavy weight	6					80		10	10																	
helicopter	7	20									10	30												10	20	10
hot	8							80									20									
ice-cream	9			20																80						
ironing	10								90																	
knock on the door	11	30						10				50													10	
lassou	12											10	80										10			
light feather	13					30	40	20							10											
rainbow	14													80												
rock a baby	15														10	80	10									
rock guitar	16															100										
scratch your knee	17																									
shake hands	18																10	70		20						
shave	19			10															90							
spank	20																			100						
spider	21											10		10									50			10
stroke the cat	22																							100		
surrender	23	10																							70	20
whistle	24																									10
umbrella	25													20												90
violin	26											50												20	30	

Number of Epochs = 9900
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 88.33%
 Average recognition rate of all gestures = 51.15%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Experiment ex51uk
cfmname: ctr3900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	86																			14					
cards	2					9		45					27							18						
cut throat	3							13					7						53				7			20
drive the car	4	38			8			15				8							8	23						
drums	5	50			6			25			6									13						
heavy weight	6					56							38				6									
helicopter	7	8						15			8	38												8	8	15
hot	8				6			71				6	6	6						6						
ice-cream	9		13					19	6				6	6					25				6		13	6
ironing	10					36					14															
knock on the door	11	19						6			44			6							13					13
lassou	12											75								13		6				
light feather	13					13							80											7		
rainbow	14												75			8						8			8	
rock a baby	15						20						20		30	10				10						
rock guitar	16							18			6			6	24					35		6				6
scratch your knee	17															100										
shake hands	18															27	9				64					
shave	19																		76				6		12	6
spank	20																			100						
spider	21											8		25		8					25					33
stroke the cat	22															8				8		83				
surrender	23																			50			29	7	14	
whistle	24										15				8					8			8	38	23	
umbrella	25							7												27				7	53	
violin	26				8							25		8							8					50

Number of Epochs = 3900
 No of gestures recognised at or above 80% = 5
 Average rec. rate of best 5 gestures = 89.81%
 Average recognition rate of all gestures = 41.99%
 cfmname: cte3900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	80									20															
cards	2							20					20						40				20			
cut throat	3																	40				10			50	
drive the car	4	20			20			10					20						20							10
drums	5	20						50								10				20						
heavy weight	6					60		10					20			10										
helicopter	7										40	20		10					10					20		
hot	8	10			10			60				10							10							
ice-cream	9											40							60							
ironing	10					50				10			30				10									
knock on the door	11	10						30					50			10										
lassou	12												70							10					20	
light feather	13					30		20					50													
rainbow	14													90												10
rock a baby	15													10		30	30			20		10				
rock guitar	16					10								10		30				50						
scratch your knee	17																100									
shake hands	18															20	10			70						
shave	19												10						60		10				10	10
spank	20																			100						
spider	21								10					10							10				30	40
stroke the cat	22																			20		80				
surrender	23									10		10	10							10			20		40	
whistle	24											10								10			20	60		
umbrella	25													10						10					80	
violin	26											30		10		20				10	10					20

Number of Epochs = 3900
 No of gestures recognised at or above 80% = 6
 Average rec. rate of best 6 gestures = 88.33%
 Average recognition rate of all gestures = 37.69%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Experiment ex52uk
cfmname: ctr9500

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	43	---	7	---	---	---	---	---	---	14	---	---	7	---	---	---	---	---	21	---	---	7	---	---	---
cards	2	---	---	27	---	18	---	---	---	18	---	---	36	---	---	---	---	---	---	---	---	---	---	---	---	---
cut throat	3	---	---	40	---	---	---	---	---	---	7	13	---	---	---	---	---	---	40	---	---	---	---	---	---	---
drive the car	4	8	---	---	8	8	---	---	---	---	8	15	15	---	---	---	23	---	---	15	---	---	---	---	---	
drums	5	13	---	---	---	19	6	---	---	---	---	19	---	---	---	19	---	6	---	19	---	---	---	---	---	
heavy weight	6	---	---	---	---	69	---	---	---	---	---	---	13	---	6	---	13	---	---	---	---	---	---	---	---	
helicopter	7	8	---	---	---	---	---	8	---	---	---	46	---	---	---	---	8	---	8	---	---	---	8	15	---	
hot	8	---	---	6	---	---	---	41	---	6	6	---	6	---	---	12	---	---	24	---	---	---	---	---	---	
ice-cream	9	---	6	25	---	---	25	---	---	---	---	6	13	6	---	---	---	6	---	---	---	---	---	13	---	
ironing	10	---	---	---	---	---	---	14	---	64	---	---	7	7	---	---	7	---	---	---	---	---	---	---	---	
knock on the door	11	25	---	---	6	---	---	---	---	---	56	6	---	---	---	---	---	---	6	---	---	---	---	---	---	
lassou	12	6	---	---	---	---	---	---	---	---	---	88	---	---	---	---	---	---	---	---	---	6	---	---	---	
light feather	13	7	---	7	---	---	27	---	---	---	13	---	---	27	---	---	---	20	---	---	---	---	---	---	---	
rainbow	14	---	---	---	---	---	17	---	---	---	---	---	---	67	---	8	---	---	8	---	---	---	---	---	---	
rock a baby	15	---	---	---	---	---	---	---	---	---	---	---	10	20	40	---	20	---	---	---	10	---	---	---	---	
rock guitar	16	---	---	---	---	---	---	---	---	---	---	---	---	94	---	---	---	6	---	---	---	---	---	---	---	
scratch your knee	17	---	---	---	---	---	---	---	---	---	---	---	---	---	---	100	---	---	---	---	---	---	---	---	---	
shake hands	18	---	---	---	---	---	---	---	---	9	9	---	---	---	---	9	9	18	---	18	---	27	---	---	---	
shave	19	---	---	12	---	---	---	---	---	---	---	---	6	---	---	---	---	53	---	6	---	---	---	18	6	
spank	20	7	---	---	---	---	---	---	---	---	---	---	---	---	7	7	---	73	---	7	---	---	---	---	---	
spider	21	---	---	---	---	---	---	---	---	---	---	---	42	---	---	---	---	8	---	25	---	---	---	8	17	
stroke the cat	22	---	---	---	---	---	---	---	---	---	---	---	---	---	---	8	---	---	8	---	83	---	---	---	---	
surrender	23	---	---	21	---	---	---	---	---	---	---	7	---	---	7	---	---	7	---	---	---	7	7	43	---	
whistle	24	---	---	8	---	---	---	---	---	---	---	---	31	---	---	---	---	---	---	---	8	8	46	---	---	
umbrella	25	---	---	---	---	---	---	---	---	---	---	---	---	7	---	---	---	33	---	---	---	---	60	---	---	
violin	26	8	---	17	---	8	---	---	---	---	---	17	---	---	8	---	---	---	8	---	---	---	8	25	---	

Number of Epochs = 9500

No of gestures recognised at or above 80% = 4
Average rec. rate of best 4 gestures = 91.24%
Average recognition rate of all gestures = 41.82%

cfmname: cte9500

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	50	---	---	30	---	---	---	---	---	10	---	10	---	---	---	---	---	---	---	---	---	---	---	---	---
cards	2	---	20	---	10	---	---	---	30	---	---	20	---	---	---	---	---	20	---	---	---	---	---	---	---	---
cut throat	3	---	20	---	---	---	---	---	---	---	10	---	---	---	---	---	---	60	---	---	---	---	---	---	10	---
drive the car	4	10	---	---	10	10	---	---	---	20	---	---	---	---	---	50	---	---	---	---	---	---	---	---	---	---
drums	5	10	---	---	20	20	---	10	---	---	---	---	---	10	10	---	---	10	10	---	---	---	---	---	---	
heavy weight	6	---	---	---	---	50	---	---	---	10	---	---	10	---	10	---	20	---	---	---	---	---	---	---	---	
helicopter	7	10	---	40	---	---	---	---	---	10	30	---	---	---	---	---	10	---	---	---	---	---	---	---	---	
hot	8	---	10	---	10	---	10	---	20	---	10	---	10	10	---	10	---	---	---	---	---	---	---	---	---	
ice-cream	9	---	60	---	---	---	---	---	---	10	---	20	---	---	---	---	---	---	---	---	---	---	---	10	---	
ironing	10	---	---	---	---	10	---	---	---	80	---	---	---	---	---	---	10	---	---	---	---	---	---	---	---	
knock on the door	11	20	---	10	---	---	---	---	---	20	40	---	---	---	---	---	---	---	10	---	---	---	---	---	---	
lassou	12	---	---	---	---	---	---	---	---	---	80	---	---	---	---	---	---	---	---	---	---	---	---	10	10	
light feather	13	---	---	10	---	60	---	10	---	10	---	10	---	---	---	---	---	---	---	---	---	---	---	---	---	
rainbow	14	---	---	---	---	---	---	---	---	---	---	---	90	---	---	---	---	---	---	---	10	---	---	---	---	
rock a baby	15	---	---	---	---	---	---	---	---	---	---	---	---	10	40	30	---	---	20	---	---	---	---	---	---	
rock guitar	16	---	---	---	---	---	---	---	---	---	---	---	---	---	90	---	---	---	---	---	10	---	---	---	---	
scratch your knee	17	---	---	---	---	---	---	---	---	---	---	---	---	---	---	100	---	---	---	---	---	---	---	---	---	
shake hands	18	---	---	---	---	10	---	10	---	10	10	10	---	10	10	---	10	---	30	---	---	---	---	---	---	
shave	19	---	---	40	---	10	---	---	---	---	---	---	---	---	---	---	30	---	---	---	---	---	---	20	---	
spank	20	---	---	---	---	---	---	---	---	---	---	---	---	---	10	---	---	---	80	---	10	---	---	---	---	
spider	21	---	---	10	---	---	---	---	---	---	10	---	10	---	---	---	10	---	---	20	---	---	---	20	20	
stroke the cat	22	---	---	---	---	---	---	---	---	---	---	---	---	---	---	10	10	10	---	10	---	60	---	---	---	
surrender	23	---	---	10	---	---	---	---	---	10	10	10	---	---	---	---	---	10	---	---	---	---	---	50	---	
whistle	24	---	---	---	---	---	---	---	---	10	30	---	---	---	---	---	---	10	---	---	---	10	---	30	10	
umbrella	25	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	10	---	---	---	---	80	10	
violin	26	---	---	20	---	---	---	---	---	10	30	---	10	---	20	---	---	---	---	10	---	---	---	---	---	

Number of Epochs = 9500

No of gestures recognised at or above 80% = 7
Average rec. rate of best 7 gestures = 85.71%
Average recognition rate of all gestures = 34.62%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Experiment ex53uk
cfmname: ctr2200

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	86			14																					
cards	2					9		55	27				9													
cut throat	3			20					27												53					
drive the car	4	23			46	15																15				
drums	5	25			6	50						13										6				
heavy weight	6						81						19													
helicopter	7					8		8				8	31				8							23	8	8
hot	8								100																	
ice-cream	9		6	13					6	6						6			63							
ironing	10	7						7				86														
knock on the door	11				6	19						44	6							6		13		6		
lassou	12												75								13		6		6	
light feather	13						40		7		7			47												
rainbow	14													100												
rock a baby	15						10				20					70										
rock guitar	16									13						82						6				
scratch your knee	17																100									
shake hands	18					9										9	18	27		18		18				
shave	19																	100								
spank	20																		100							
spider	21														8					8		83				
stroke the cat	22																					100				
surrender	23																						93		7	
whistle	24																8						8	77	8	
umbrella	25																						27	7	60	
violin	26					8							25				8					8			8	42

Number of Epochs = 2200
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 92.60%
 Average recognition rate of all gestures = 62.02%
 cfmname: cte2200

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80			20																						
cards	2		10					30	10				10							40							
cut throat	3			40															60								
drive the car	4	10			10	20					10										10		20	10	10		
drums	5	20				20			20		10	10				20											
heavy weight	6						80						20														
helicopter	7	10										50	10										20	10			
hot	8	10								80							10										
ice-cream	9									30									70								
ironing	10						10				80		10														
knock on the door	11	10							10			30	30			10							10				
lassou	12												60							10				10	20		
light feather	13		10				40		20					30													
rainbow	14														90												
rock a baby	15								10							80	10										
rock guitar	16								10							90											
scratch your knee	17																100										
shake hands	18														10			20		70							
shave	19			10														90									
spank	20																			80		10					
spider	21														10						80		10				
stroke the cat	22																					100					
surrender	23												10										40		50		
whistle	24																							90	10		
umbrella	25																							40	10	40	
violin	26											10	50		10		10									20	

Number of Epochs = 2200
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 86.67%
 Average recognition rate of all gestures = 52.69%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Experiment ex54uk
cfmname: ctr3700

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	93						7																				
cards	2							82	18																			
cut throat	3			13																87								
drive the car	4	38		46					8	8																		
drums	5	25			25			19		19						6				6								
heavy weight	6					100																						
helicopter	7							8	31	15														31	8	8		
hot	8							100																				
ice-cream	9			13				6		6										75								
ironing	10						7		7	86																		
knock on the door	11	6		6	6					75										6								
lassou	12									19	63										6		13					
light feather	13						53		20		20		7															
rainbow	14												92															
rock a baby	15									10						90												
rock guitar	16								12							88												
scratch your knees	17																100											
shake hands	18				9	9					9						9		45		18							
shave	19																		100									
spank	20																			100								
spider	21												17								75		8					
stroke the cat	22																					100						
surrender	23																						100					
whistle	24	8									8	23												31	31			
umbrella	25																								87		13	
violin	26										8	8								8		75						

Number of Epochs = 3700
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 96.22%
 Average recognition rate of all gestures = 54.09%
 cfmname: cte3700

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	90			10																						
cards	2			10					20		40									30							
cut throat	3																		100								
drive the car	4	10		10	20					20						30								10			
drums	5	10			20				20	20						20				10							
heavy weight	6					60		10	20								10										
helicopter	7	20							10	40	10													10	10		
hot	8								90								10										
ice-cream	9			10						10									80								
ironing	10								10		90																
knock on the door	11	20							10		40	20												10			
lassou	12										10	50									10		10	10	10		
light feather	13						30		40		20						10										
rainbow	14															100											
rock a baby	15																90	10									
rock guitar	16																100										
scratch your knees	17																	100									
shake hands	18												10				10			80							
shave	19																			100							
spank	20																10				90						
spider	21												10									60		30			
stroke the cat	22																						100				
surrender	23	10																						80		10	
whistle	24	40			10																			10	40		
umbrella	25																								70		20
violin	26											20	20				20					20	20				

Number of Epochs = 3700
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 94.00%
 Average recognition rate of all gestures = 47.69%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Ex59uk
cfmname: ctr4400

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	50			7	21					7									7				7		
cards	2			9			18		55	9							9									
cut throat	3			47																53						
drive the car	4	23			62	8															8					
drums	5	13			6	50			6		13								6			6				
heavy weight	6						94				6															
helicopter	7							23				38						8						8	8	15
hot	8								100																	
ice-cream	9			13					19	13							6			50						
ironing	10						14				86															
knock on the door	11				6	25					50						6								6	6
lassou	12											63									6			13	19	
light feather	13						67			13			20													
rainbow	14													83								17				
rock a baby	15										20						80									
rock guitar	16										18						76	6								
scratch your knee	17																100									
shake hands	18				9	9											9	9	36		18		9			
shave	19				6														94							
spank	20																7					87		7		
spider	21															8			8			83				
stroke the cat	22										8														92	
surrender	23														7									79		14
whistle	24																							8	85	8
umbrella	25																							40		60
violin	26			8				17					25				8						17			25

Number of Epochs = 4400
 No of gestures recognized at or above 80% = 10
 Average rec. rate of best 10 gestures = 90.32%
 Average recognition rate of all gestures = 59.84%

Ex59uk
cfmname: cta4400

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	40			10	30					10	10														
cards	2			10			30		40	10										10						
cut throat	3			40																60						
drive the car	4	10			10	40			10								30									
drums	5	30				20			20								20				10					
heavy weight	6						90		10																	
helicopter	7							20			30	20												10	10	10
hot	8								80								20									
ice-cream	9			30															70							
ironing	10						20			40			10	20			10									
knock on the door	11			10		10			10			10	30				10							20		
lassou	12												70										20			10
light feather	13						40		40				20													
rainbow	14													60								30	10			
rock a baby	15									20							70	10								
rock guitar	16									10							90									
scratch your knee	17																	100								
shake hands	18					20							10					10	40		20					
shave	19				10					10									80							
spank	20																	10		80		10				
spider	21													10								90				
stroke the cat	22																20						80			
surrender	23												10											50		40
whistle	24	10																							90	
umbrella	25																				20			30	10	40
violin	26			10		10						10	40		10											20

Number of Epochs = 4400
 No of gestures recognized at or above 80% = 9
 Average rec. rate of best 9 gestures = 86.67%
 Average recognition rate of all gestures = 47.69%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Experiment ex56uk
cfmname: ctr7900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	93																			7					
cards	2		91				9																			
cut throat	3			13															87							
drive the car	4	69			23													8								
drums	5	44				13			6			25						6		6						
heavy weight	6						100												6							
helicopter	7	8										31						8						15	8	31
hot	8												100													
ice-cream	9		31																63							6
ironing	10						14				06															
knock on the door	11	31									44							6			6		6		6	6
lassou	12											81									6		6		6	6
light feather	13						13							87												
rainbow	14													83									17			
rock a baby	15						10				20					30	40									
rock guitar	16								12							76					6					6
scratch your knee	17																100									
shake hands	18																	100								
shave	19		6																94							
spank	20																			100						
spider	21													8							92					
stroke the cat	22																					100				
surrender	23																						71	7	21	
whistle	24	8																						77	15	
umbrella	25											7												7	7	80
violin	26											8	8									33				50

Number of Epochs = 7900
 No of gestures recognised at or above 80% = 15
 Average rec. rate of best 15 gestures = 92.43%
 Average recognition rate of all gestures = 68.10%
 cfmname: cte7900

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	80				10					10															
cards	2	10	60						10				10						10							
cut throat	3																	100								
drive the car	4	40				10			10									10								30
drums	5	40	10			20			10								10			10						
heavy weight	6						60							30			10									
helicopter	7	20									30	20												10	10	10
hot	8	30							70																	
ice-cream	9		50										10						40							
ironing	10									50			40		10											
knock on the door	11	30							10			20	30									10				
lassou	12												80											10	10	
light feather	13		10				30		10					50												
rainbow	14														80									20		
rock a baby	15								10							50	30					10				
rock guitar	16								10								80					10				
scratch your knee	17																	100								
shake hands	18																	10	90							
shave	19																			100						
spank	20																				100					
spider	21													10								80				10
stroke the cat	22																						100			
surrender	23	10																						40	10	40
whistle	24																								100	
umbrella	25																					10		20	10	60
violin	26											10							10		10	10				60

Number of Epochs = 7900
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 90.00%
 Average recognition rate of all gestures = 58.85%

Appendix C.6 Confusion Matrices: Forearm Orienta-

Experiment ex57uk
cfmname: ctr2800

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	79						7												7					7	
cards	2		9				9	82																		
cut throat	3			40															60							
drive the car	4	15		23	8					8					8		23						8	8		
drums	5	13			25				19		13				25	6										
heavy weight	6					100																				
helicopter	7						8	8			23				8		8						31	8	8	
hot	8								100																	
ice-cream	9		6	6			6	13							6	6		50						6		
ironing	10						14	7	79																	
knock on the door	11	6						6		56	6							6					6	13		
lassou	12										94												6			
light feather	13					40		20		13		27														
rainbow	14												100													
rock a baby	15									10					10	80										
rock guitar	16								12							88										
scratch your knee	17															100										
shake hands	18												9				91									
shave	19																	94						6		
spank	20																13	7	80							
spider	21											8	17							67						8
stroke the cat	22																				100					
surrender	23																						100			
whistle	24							8															23	69		
umbrella	25																							67	7	27
violin	26		8								8	25			8			8			8		8			33

Number of Epochs = 2800
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 95.18%
 Average recognition rate of all gestures = 61.45%
 cfmname: cta2800

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	60			20						10														10	
cards	2		10	20			10	40					20													
cut throat	3																		100							
drive the car	4	20			10	10									40								10	10		
drums	5	30				10			20													10				
heavy weight	6						70	10	10		10															
helicopter	7									20	10												30	20		
hot	8										90						10									
ice-cream	9			50					20										30							
ironing	10						10			90																
knock on the door	11	20						10			10	40						10					10			
lassou	12												60									10		20		10
light feather	13		20				40	30	10																	
rainbow	14													90								10				
rock a baby	15						10								80	10										
rock guitar	16															100										
scratch your knee	17																100									
shake hands	18												10				10	80								
shave	19			30														70								
spank	20															10		10		80						
spider	21											10	10								40		20		20	
stroke the cat	22																					100				
surrender	23																							90		10
whistle	24																						10	90		
umbrella	25																			10				60		30
violin	26										10	30				20										40

Number of Epochs = 2800
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 91.00%
 Average recognition rate of all gestures = 50.77%

C.7 Confusion Matrices: Scalar and Vector Velocity

Experiment ex38uk

cfmname: ctr10000

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	93				7																					
cards	2							9		9				55							18						
cut throat	3			7																	93						
drive the car	4	46				31																15			8		
drums	5	44				44				6																	
heavy weight	6						94																				
helicopter	7										8	23									8		23	15	23		
hot	8								100																		
ice-cream	9								13					6							75	6					
ironing	10							14																			
knock on the door	11	38				6						31	13										6		6		
lassou	12												81									6		13			
light feather	13						33					13		53													
rainbow	14													92							8						
rock a baby	15						50					20		10				10			10						
rock guitar	16	6								59		6									24		6				
scratch your knee	17																	100									
shake hands	18												18	9		9					36		18				
shave	19																			100							
spank	20																				93		7				
spider	21												8		17							75					
stroke the cat	22																					100					
surrender	23																							79		21	
whistle	24	38																							54	8	
umbrella	25														7									20		73	
violin	26												25								8		67				

Number of Epochs = 10000
 No of gestures recognized at or above 80% = 10
 Average rec. rate of best 10 gestures = 93.86%
 Average recognition rate of all gestures = 52.09%
 cfmname: cta10000

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80				20																					
cards	2								20	10				30						40							
cut throat	3																		100								
drive the car	4	10				40			10		10										20				10		
drums	5	30				40			20												10						
heavy weight	6						80							20													
helicopter	7	20				10						10	30											20	10		
hot	8								80		20																
ice-cream	9								10											90							
ironing	10									10				10													
knock on the door	11	10									20		10	40											20		
lassou	12												80									10			10		
light feather	13						20			20				60													
rainbow	14																		100								
rock a baby	15							20		40							10				30						
rock guitar	16						10			60												20	10				
scratch your knee	17																		90				10				
shake hands	18						10						40		10						30						
shave	19																				100						
spank	20																					90		10			
spider	21												10		10								60			20	
stroke the cat	22																		10					90			
surrender	23																								70		30
whistle	24	20																								80	
umbrella	25														20									20		60	
violin	26												50								10			40			

Number of Epochs = 10000
 No of gestures recognized at or above 80% = 11
 Average rec. rate of best 11 gestures = 86.36%
 Average recognition rate of all gestures = 48.08%

Appendix C.7 Confusion Matrices: Scalar and Vec-

Experiment ex40uk
cfname: ctr4500

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	93			7																					
cards	2						9	45					45													
cut throat	3			47				7												47						
drive the car	4	23			46	23													8							
drums	5	31				31						6									31					
heavy weight	6						100																			
helicopter	7	8		8	8							15										8		8	46	
hot	8						6		94																	
ice-cream	9			31						25										38	6					
ironing	10						14					86														
knock on the door	11	31			6	13						13	6		6						6	6	6		6	
lassou	12											6	44									25				25
light feather	13						33		13		7			47												
rainbow	14													100												
rock a baby	15							50		10		20									20					
rock guitar	16					6				47											41		6			
scratch your knee	17																		100							
shake hands	18	9			9														82							
shave	19			18																82						
spank	20																				100					
spider	21																17					83				
stroke the cat	22																	8					92			
surrender	23					7																	43		50	
whistle	24	15			8				8			15							8					8		38
umbrella	25																									93
violin	26												33		8					8			33			17

Number of Epochs = 4500
No of gestures recognised at or above 80% = 12
Average rec. rate of best 12 gestures = 92.10%
Average recognition rate of all gestures = 52.89%
cfname: cta4500

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	90				10																				
cards	2									40	10			30							20					
cut throat	3			50																	50					
drive the car	4	20				30					10									20		20				
drums	5	20				30				20		20										10				
heavy weight	6						80				10		10													
helicopter	7	20				20						20												20	20	
hot	8					10				90																
ice-cream	9			10						20										70						
ironing	10						10				90															
knock on the door	11	20								20				30										10	20	
lassou	12													30								10	40			20
light feather	13						30		50					20												
rainbow	14													70								30				
rock a baby	15						30		50				10							10						
rock guitar	16			10		10	10		40												30					
scratch your knee	17																			100						
shake hands	18				10														10	60		20				
shave	19			20																	80					
spank	20																			10		80		10		
spider	21													10								80				10
stroke the cat	22																	10					90			
surrender	23	10																						40		50
whistle	24	30										30												10		30
umbrella	25													30										10		60
violin	26												10		20			10	10			30				20

Number of Epochs = 4500
No of gestures recognised at or above 80% = 9
Average rec. rate of best 9 gestures = 86.67%
Average recognition rate of all gestures = 43.85%

Appendix C.7 Confusion Matrices: Scalar and Vec-

Experiment ex41uk
cfmname: ctr700

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	21		7	7			14		29	14						7									
cards	2																		100							
cut throat	3					7													93							
drive the car	4				8	8		15		23	15					15										15
drums	5				6	25		13		19	6					13										19
heavy weight	6						19				25															56
helicopter	7	8					8		8		15		23	8	8		8									15
hot	8				6			18		24	18															6 24
ice-cream	9													6												88 6
ironing	10										93	7														
knock on the door	11				6			6		38	38					6										6
lassou	12				13			13		25		13			19											19
light feather	13					7		7		20			7			7										40 13
rainbow	14							8						33		8										17 33
rock a baby	15						20		10		40	10														20
rock guitar	16							12		12					6		12									6 53
scratch your knee	17						8		17		8	8				8										33 17
shake hands	18					9		9		18					9		9									45
shave	19																									100
spank	20										7			7		7										80
spider	21						25			8			8													42 17
stroke the cat	22								17					17		8										8 50
surrender	23					14		14		29	7		7													7 21
whistle	24					8	8			31	8															15 31
umbrella	25						13		7		7			7												67
violin	26					8		25		17			8	8												17 17

Number of Epochs = 700
 No of gestures recognized at or above 80% = 3
 Average rec. rate of best 3 gestures = 90.95%
 Average recognition rate of all gestures = 18.24%
 cfmname: cte700

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	30			10			20		40																
cards	2																		100							
cut throat	3																		100							
drive the car	4									20			10		30						10	30				
drums	5				10	10		10		30							10					30				
heavy weight	6					30				10												60				
helicopter	7				10	10		10		10	20	10										10	20			
hot	8							30		40		10										10	10			
ice-cream	9																					100				
ironing	10					20					80															
knock on the door	11	10						20		70																
lassou	12				10					30	10	10		20		10										10
light feather	13					20		10		30												70				
rainbow	14					10		20						10								10	50			
rock a baby	15							30		20			10									20	20			
rock guitar	16					20			30	10	10			10									70			
scratch your knee	17								20														10	70		
shake hands	18								10			10											20	60		
shave	19																						100			
spank	20					10								10										80		
spider	21						10			30	10		10										30	10		
stroke the cat	22																10							90		
surrender	23							20		30	10		20										20			
whistle	24						20		20		20	20											20			
umbrella	25						30		20	10													40			
violin	26						20				10		10		10									50		

Number of Epochs = 700
 No of gestures recognized at or above 80% = 3
 Average rec. rate of best 3 gestures = 86.67%
 Average recognition rate of all gestures = 14.62%

Appendix C.7 Confusion Matrices: Scalar and Vec-

Experiment ex42uk
cfmname: ctr7300

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	79				14																			7		
cards	2		82										18														
cut throat	3			20					7	7										67							
drive the car	4	54			8	15															15						
drums	5	25				56															6						
heavy weight	6						94							6													
helicopter	7											23	38											15		15	8
hot	8							6		94																	
ice-cream	9		25	6																	63	6					
ironing	10						7		7			86															
knock on the door	11	6			6	6						44	19							6		6		6			
lassou	12													69									6				13
light feather	13							13							73												
rainbow	14												8		92												
rock a baby	15						10			20							20	30									
rock guitar	16					6				24							41				24		6				
scratch your knee	17																	100									
shake hands	18					9						9			9				18		27		18				
shave	19																			100							
spank	20																				100						
spider	21											17		33								50					
stroke the cat	22																	8						92			
surrender	23																								86		14
whistle	24	15								8			23												31	8	15
umbrella	25																								20		80
violin	26												42								8		8				42

Number of Epochs = 7300
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 91.31%
 Average recognition rate of all gestures = 58.14%
 cfmname: cte7300

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80				20																					
cards	2		70										10								20						
cut throat	3			40																	60						
drive the car	4	30							10				10	10						10	10	20					
drums	5	40				30			20								10										
heavy weight	6						80				10			10													
helicopter	7	30											50											20			
hot	8	10											10														
ice-cream	9		20																		80						
ironing	10										80			20													
knock on the door	11	10							10			20	40											20			
lassou	12													90								10					
light feather	13						10		10					80													
rainbow	14														90												10
rock a baby	15						10		60				10					10		10							
rock guitar	16					20	10		40								20				10						
scratch your knee	17																	90			10						
shake hands	18					10								10					10	20		50					
shave	19			10																90							
spank	20																				100						
spider	21												30										50				20
stroke the cat	22																			10					90		
surrender	23													10											60		30
whistle	24	20											10							10					30	20	10
umbrella	25													10											50		30
violin	26												70									10					20

Number of Epochs = 7300
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 86.36%
 Average recognition rate of all gestures = 51.15%

Appendix C.7 Confusion Matrices: Scalar and Vec-

Experiment ex43uk
cfmname: ctr7600

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	93			7																						
cards	2							9				82							9								
cut throat	3						7	7											87								
drive the car	4	69									8								23								
drums	5	38			25			6			25										6						
heavy weight	6					94							6														
helicopter	7											38						8			8		15	8	23		
hot	8					6		94																			
ice-cream	9		6					6				13							69	6							
ironing	10							7		93																	
knock on the door	11	13									56	6		6					13						6		
lassou	12										6	69									6		13		6		
light feather	13					20							80														
rainbow	14													100													
rock a baby	15					20				30					10		10		10	20							
rock guitar	16	6						35			6					35				12		6					
scratch your knee	17																100										
shake hands	18					9												82		9							
shave	19																		100								
spank	20																		7		93						
spider	21													8								92					
stroke the cat	22																						100				
surrender	23																							100			
whistle	24	8						8			15													15	38	15	
umbrella	25													7											40		53
violin	26											17	42											42			

Number of Epochs = 7600
 No of gestures recognised at or above 80% = 13
 Average rec. rate of best 13 gestures = 93.88%
 Average recognition rate of all gestures = 57.98%
 cfmname: cte7600

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80			10						10																
cards	2		20					10	20				40						10								
cut throat	3																		100								
drive the car	4	30						10			10	10						10		10		10			10		
drums	5	30				40		20				10															
heavy weight	6						90						10														
helicopter	7	10									30	30													10	10	10
hot	8	10						60		10			20														
ice-cream	9		10						20										70								
ironing	10									90			10														
knock on the door	11	10						10			20	30							10					20			
lassou	12											60							10			20				10	
light feather	13						30		10				60														
rainbow	14													90							10						
rock a baby	15					40		20				10			10	10			10								
rock guitar	16				10	10		20								40				10		10					
scratch your knee	17																100										
shake hands	18												18				10	80		10							
shave	19																		100								
spank	20																			100							
spider	21													10							80					10	
stroke the cat	22																10			10			80				
surrender	23																						80		20		
whistle	24																							10	90		
umbrella	25													10								10		40		40	
violin	26												60									10					

Number of Epochs = 7600
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 88.33%
 Average recognition rate of all gestures = 55.00%

Appendix C.7 Confusion Matrices: Scalar and Vec-

Experiment ex44uk
cfmname: ctr2000

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	71				14					7															7		
cards	2		18					27					9															
cut throat	3			60							7																	
drive the car	4	38			31	8					8																	
drums	5	31				31					6																	
heavy weight	6						100																					
helicopter	7											54															8	31
hot	8						6				94																	
ice-cream	9			31								25																44
ironing	10							7			7		86															
knock on the door	11	25				13					6			19	19		6									6	6	
lassou	12														81												6	
light feather	13							33			7			20														
rainbow	14																	83									17	
rock a baby	15										40																	
rock guitar	16							6			47			6				18								6		6
scratch your knee	17																											75
shake hands	18																			9								91
shave	19				53						6																	41
spank	20																											100
spider	21																										58	8
stroke the cat	22																											100
surrender	23																										50	50
whistle	24	23									8			8														38 23
umbrella	25																											100
violin	26												8	58													8	25

Number of Epochs = 2000
 No of gestures recognised at or above 80% = 9
 Average rec. rate of best 9 gestures = 92.81%
 Average recognition rate of all gestures = 54.28%
 cfmname: cta2000

Gesture:	Network Decision																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
bird	1	80				10						10																
cards	2		10					20		10	30																	30
cut throat	3			100																								
drive the car	4	10						10					10		10		10									10		10
drums	5	30																										10
heavy weight	6																											10
helicopter	7	20																										30
hot	8																											90
ice-cream	9				10																							30
ironing	10																											90
knock on the door	11	10																										10 40
lassou	12																											80
light feather	13																											10 10
rainbow	14																											40
rock a baby	15																											90
rock guitar	16																											10
scratch your knee	17																											70
shake hands	18																											10 80
shave	19																											40
spank	20																											90
spider	21																											80
stroke the cat	22																											100
surrender	23																											10 70
whistle	24	20																										10 60
umbrella	25																											20 80
violin	26																											70 10

Number of Epochs = 2000
 No of gestures recognised at or above 80% = 12
 Average rec. rate of best 12 gestures = 86.67%
 Average recognition rate of all gestures = 52.31%

C.8 Curvature and Plane of Motion

Experiment ex33uk
cfmname: ctr5200

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	79			7								7											7		
cards	2							82					9						9							
cut throat	3							13					9						87							
drive the car	4	38		8	8			15										15	8	8						
drums	5	19			31			13								13		6		19						
heavy weight	6					94												6								
helicopter	7											15	15					8					31		31	
hot	8						94											6								
ice-cream	9							25										75								
ironing	10						7			86			7													
knock on the door	11	6			13			13			44	6				6		6		6						
lassou	12										6	69										6		13		6
light feather	13					27					7		67													
rainbow	14													100												
rock a baby	15		10			20		10								40			20							
rock guitar	16							24								53				18		6				
scratch your knee	17															100										
shake hands	18				9									9				73		9						
shave	19																		100							
spank	20																	7		93						
spider	21											8	58								33					
strokes the cat	22																8					92				
surrender	23																						86		14	
whistle	24	8						15			15	8										15		38		
umbrella	25																						13		87	
violin	26											33								17	8	42				

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 93.10%
 Average recognition rate of all gestures = 53.33%
 cfmname: cte5200

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	70			20																			10		
cards	2							50					10						40							
cut throat	3																		100							
drive the car	4	10		10	10	10										30		10		10		10				
drums	5	20			20			20								40										
heavy weight	6					80									20											
helicopter	7	10						10			10	40											10	10	10	
hot	8							90								10										
ice-cream	9							20											80							
ironing	10									80			20													
knock on the door	11	20						10			30	20										10		10		
lassou	12											60									10	10				20
light feather	13					40		50					10													
rainbow	14													100												
rock a baby	15							30							30			40								
rock guitar	16							10							80			10								
scratch your knee	17															100										
shake hands	18												10				10	70		10						
shave	19																		100							
spank	20																	10		90						
spider	21												10								50					40
strokes the cat	22																10					90				
surrender	23																						70		30	
whistle	24	10						10										10					10	30	30	
umbrella	25														20								30		50	
violin	26											20	30		10			10		20						10

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 9
 Average rec. rate of best 9 gestures = 90.00%
 Average recognition rate of all gestures = 49.23%

Appendix C.8 Curvature and Plane of Motion

Experiment ex34uk
cfnname: ctr5200

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	71			14					7														7		
cards	2					9	64	9				18														
cut throat	3		13																87							
drive the car	4	46		23						8									23							
drums	5	31			19			6	6	6									6	25						
heavy weight	6					94				6																
helicopter	7											31							8		8		23	23	8	
hot	8							100																		
ice-cream	9		6					19	6										63							6
ironing	10						7	7	86																	
knock on the door	11	6		6						63	6								6	6	6			6		
lassou	12										81											6		13		
light feather	13					40		13	7			40														
rainbow	14												75									25				
rock a baby	15					20		10	30	10						20			10							
rock guitar	16							47		6						24	6		6	12						
scratch your knee	17																100									
shake hands	18					9								9				82								
shave	19																		100							
spank	20					7																				93
spider	21													17									75			8
stroke the cat	22																	8					92			
surrender	23																							93		7
whistle	24	15										8												15	54	8
umbrella	25																							60	7	33
violin	26										8	17							8		58				8	

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 92.04%
 Average recognition rate of all gestures = 54.48%
 cfnname: cte5200

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	80			20																					
cards	2		10					40	10				20							20						
cut throat	3																		100							
drive the car	4	10		10						20	20					10		10		10		10				
drums	5	30			20			20		10	10													10		
heavy weight	6					100																				
helicopter	7	10									30	30												10	20	
hot	8							90									10									
ice-cream	9							20											80							
ironing	10									80			20													
knock on the door	11							10			50	20											10	10		
lassou	12											50									10	20			20	
light feather	13					50		30					20													
rainbow	14													80								20				
rock a baby	15					10		60				10							20							
rock guitar	16				10	10		40								30					10					
scratch your knee	17																100									
shake hands	18																10	80		10						
shave	19																		100							
spank	20																		40		60					
spider	21																					60		10		30
stroke the cat	22																		10				90			
surrender	23																							70		30
whistle	24																							10	80	10
umbrella	25																						20	40		40
violin	26										30	20										40		10		

Number of Epochs = 5200
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 88.00%
 Average recognition rate of all gestures = 49.62%

Appendix C.8 Curvature and Plane of Motion

Experiment ex35uk
cfmname: ctr6100

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	71			7	7					7															7	
cards	2		18				18		36		9			9						9							
cut throat	3			53						7										40							
drive the car	4	46			31	8													15								
drums	5	25				44			6			6									13						
heavy weight	6						94				6																
helicopter	7											38							8					15	8	31	
hot	8									100																	
ice-cream	9			19					13	31										31	6						
ironing	10									7	93																
knock on the door	11				6	19			6			6	31								13					19	
lassou	12												88									6					6
light feather	13						40		7		20			33													
rainbow	14													92								8					
rock a baby	15							20			20					10	40			10							
rock guitar	16				6				35								47				6		6				
scratch your knee	17																	100									
shake hands	18						9												82		9						
shave	19			18						6											76						
spank	20															7			7		80		7				
spider	21												8		25								67				
stroke the cat	22																							100			
surrender	23											7															86
whistle	24	8			8				8				15												15	38	8
umbrella	25												7												13		80
violin	26												42							8	8	33					8

Number of Epochs = 6100
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 90.30%
 Average recognition rate of all gestures = 58.79%
 cfmname: cte6100

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80			10																					10	
cards	2		20				10		30	20	10									10							
cut throat	3			50																50							
drive the car	4	10			20				10		10					20		10			10						
drums	5	20				30			20		10					20											
heavy weight	6						80							20													
helicopter	7	10										10	40											10	20	10	
hot	8									90		10															
ice-cream	9								20	10											70						
ironing	10						10				90																
knock on the door	11								20			10	50												20		
lassou	12												100														
light feather	13						40		50				10														
rainbow	14													80											20		
rock a baby	15							20		30	10		10			10	20										
rock guitar	16					10	10		30							40				10							
scratch your knee	17																	80						20			
shake hands	18											10							10	70		10					
shave	19			50																	50						
spank	20																										
spider	21												10		10									50			30
stroke the cat	22																								100		
surrender	23																									60	40
whistle	24																					10				10	80
umbrella	25													10										10		20	60
violin	26												80														10

Number of Epochs = 6100
 No of gestures recognised at or above 80% = 10
 Average rec. rate of best 10 gestures = 86.00%
 Average recognition rate of all gestures = 52.31%

Appendix C.8 Curvature and Plane of Motion

Experiment ex36uk
cfmname: ctr5300

Gesture:	Network Decision																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
bird	1	71			7	7								7											7	
cards	2								27				64							9						
cut throat	3			33																67						
drive the car	4	31			38	8													8			8				
drums	5	31				38			6					13	6							6				
heavy weight	6						94							6												
helicopter	7											38							8			8		15		23
hot	8								100																	
ice-cream	9			38					19											38	6					
ironing	10									7		93														
knock on the door	11	13			6	25						44	6									6				
lassou	12											6	88													6
light feather	13						13					7			80											
rainbow	14													92												8
rock a baby	15						20				10			10		30	20			10						
rock guitar	16				6		6		29		6						35				12		6			
scratch your knee	17																				100					
shake hands	18						9														91					
shave	19						18															82				
spank	20						7																87		7	
spider	21												8		25									67		
stroke the cat	22																								100	
surrender	23																								79	21
whistle	24	15							8			15													23	38
umbrella	25												7												20	73
violin	26												33										8		58	

Number of Epochs = 5300

No of gestures recognised at or above 80% = 11
Average rec. rate of best 11 gestures = 91.43%
Average recognition rate of all gestures = 58.23%
cfmname: cte5300

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	80				20																					
cards	2			20			10		40					20						10							
cut throat	3																			100							
drive the car	4	10				30			10		10			10								10	10				
drums	5	20				20			20		10					10	10					10					
heavy weight	6						80							20													
helicopter	7	10										30	40											10		10	
hot	8								80					10		10											
ice-cream	9			10					20												70						
ironing	10										60			30												10	
knock on the door	11	20							10			10	30							10				10		10	
lassou	12												90									10					
light feather	13						30		10					60													
rainbow	14															100											
rock a baby	15						20		10	10			10			20	10			20							
rock guitar	16					20	10		20				10				20				10	10					
scratch your knee	17																				100						
shake hands	18												10								10	70		10			
shave	19						40															60					
spank	20											10											90				
spider	21													10									60			30	
stroke the cat	22																							90			
surrender	23	10																							60	10	
whistle	24	40										20	10												20	10	
umbrella	25													10										10		40	
violin	26											20	50											20			10

Number of Epochs = 5300

No of gestures recognised at or above 80% = 8
Average rec. rate of best 8 gestures = 88.75%
Average recognition rate of all gestures = 46.15%

Appendix C.8 Curvature and Plane of Motion

Experiment ex37uk
cfmname: ctr6200

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	79			7													7		7							
cards	2								36				64														
cut throat	3		60						40																		
drive the car	4	31		8					8									46		8							
drums	5	31			6			6	6									19		19							
heavy weight	6					88				6			6														
helicopter	7			8								31							8				23			8	23
hot	8							100																			
ice-cream	9		50						38				6														6
ironing	10								7		93																
knock on the door	11	19			6						19	25						25	6								
lassou	12											81										6		13			
light feather	13					20								80													
rainbow	14												92														
rock a baby	15					20			20	20			10				10										
rock guitar	16							41		6							24	12									18
scratch your knee	17																	100									
shake hands	18																		100								
shave	19		76						24																		
spank	20																		7								93
spider	21												8		17												75
stroke the cat	22																										100
surrender	23																										100
whistle	24	23		8								15	23														31
umbrella	25																										80
violin	26												25							8		17					50

Number of Epochs = 6200
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 93.33%
 Average recognition rate of all gestures = 54.00%
 cfmname: cte6200

Gesture:	Network Decision																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
bird	1	60			20													10		10							
cards	2									70			30														
cut throat	3		100																								
drive the car	4	20		10				10										30		10		10	10				
drums	5	20			10			20		10							10		20		10						
heavy weight	6					80							20														
helicopter	7	20										40												30		10	
hot	8	10	10					80																			
ice-cream	9		10						80				10														
ironing	10									70			30														
knock on the door	11	10							20			20	40												10		
lassou	12												80									10		10			
light feather	13					30		20					50														
rainbow	14													100													
rock a baby	15					20		30	40				10														
rock guitar	16				10			30	20								10										30
scratch your knee	17																	90				10					
shake hands	18																		100								
shave	19		80							20																	
spank	20																		20			80					
spider	21													10									60		20		10
stroke the cat	22																										100
surrender	23																										90
whistle	24	40										20															10
umbrella	25																										20
violin	26												30						10								10

Number of Epochs = 6200
 No of gestures recognised at or above 80% = 11
 Average rec. rate of best 11 gestures = 89.09%
 Average recognition rate of all gestures = 51.15%

Appendix D

Hinton Diagrams of W1 and W2

This appendix contains the Hinton diagrams for each neural network experiment in chapter 7. The Hinton diagram shows graphically the magnitude and sign of each ANN weight. Each rectangle represents a weight, the size being proportional to its magnitude. Rectangles with light shading are positive and those with dark shading negative. The largest rectangle corresponds to the largest weight in the weight matrix and all others are drawn with sizes relative to this rectangle. W1 is the weight matrix that contains the weights connecting the network inputs to the hidden layer. W2 is the weight matrix that contains the weights connecting the hidden layer (labelled input) to the output neurons. The Hinton diagram is useful in determining which weights have a large effect on neuronal output. If all the weights connected to a particular input feature are small it suggests that this feature is having relatively little effect on the network decisions compared to other features.

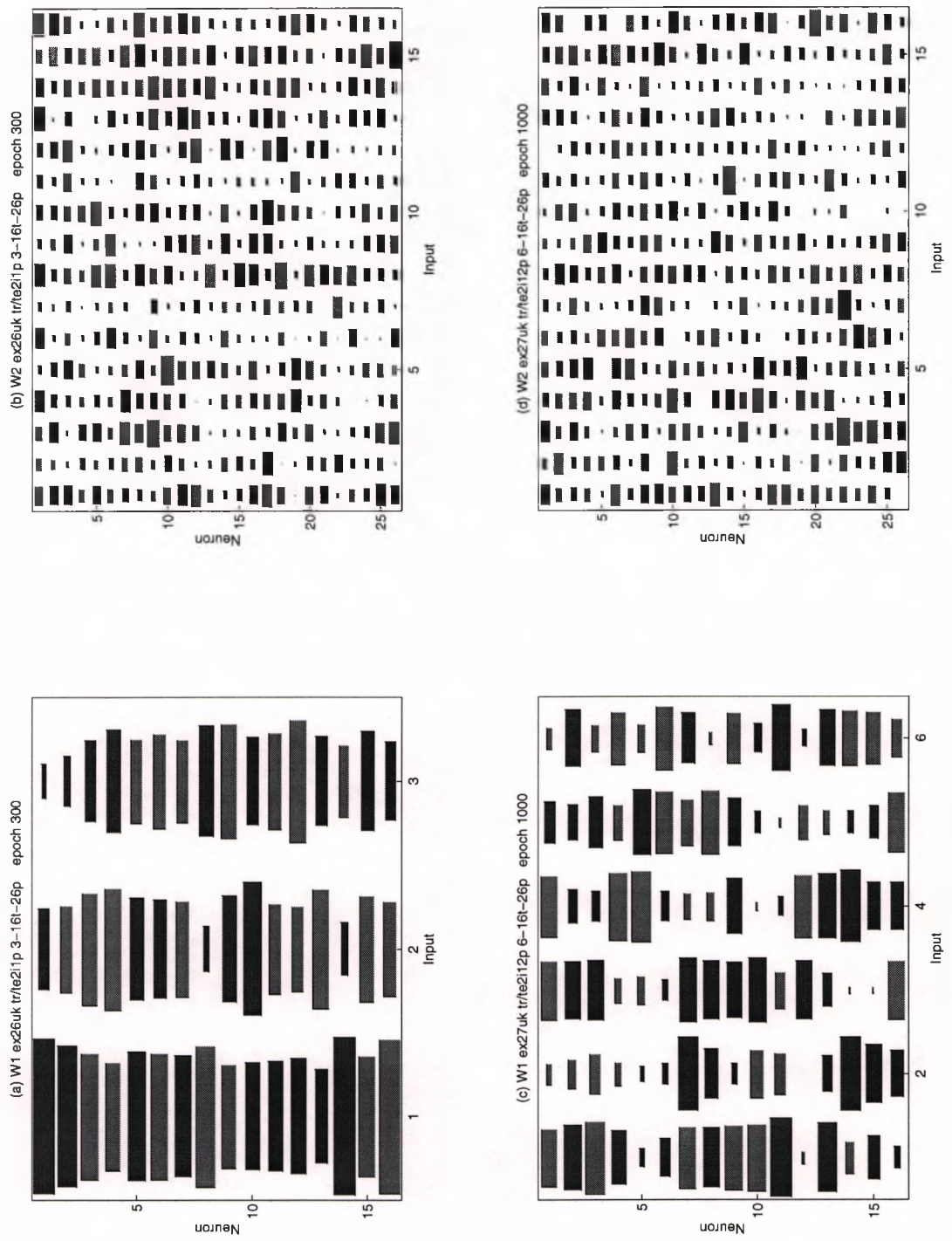


Figure D.1

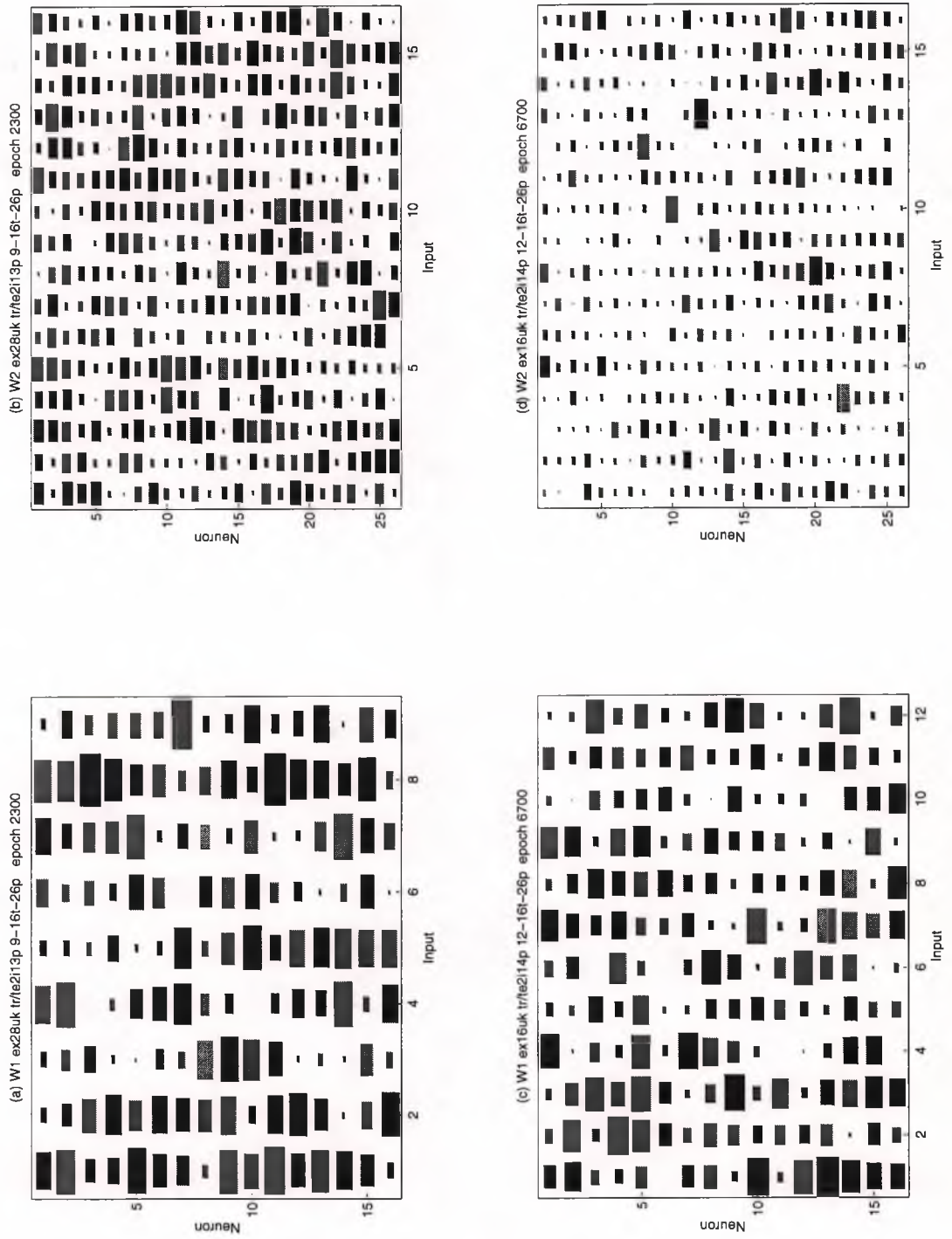


Figure D.2

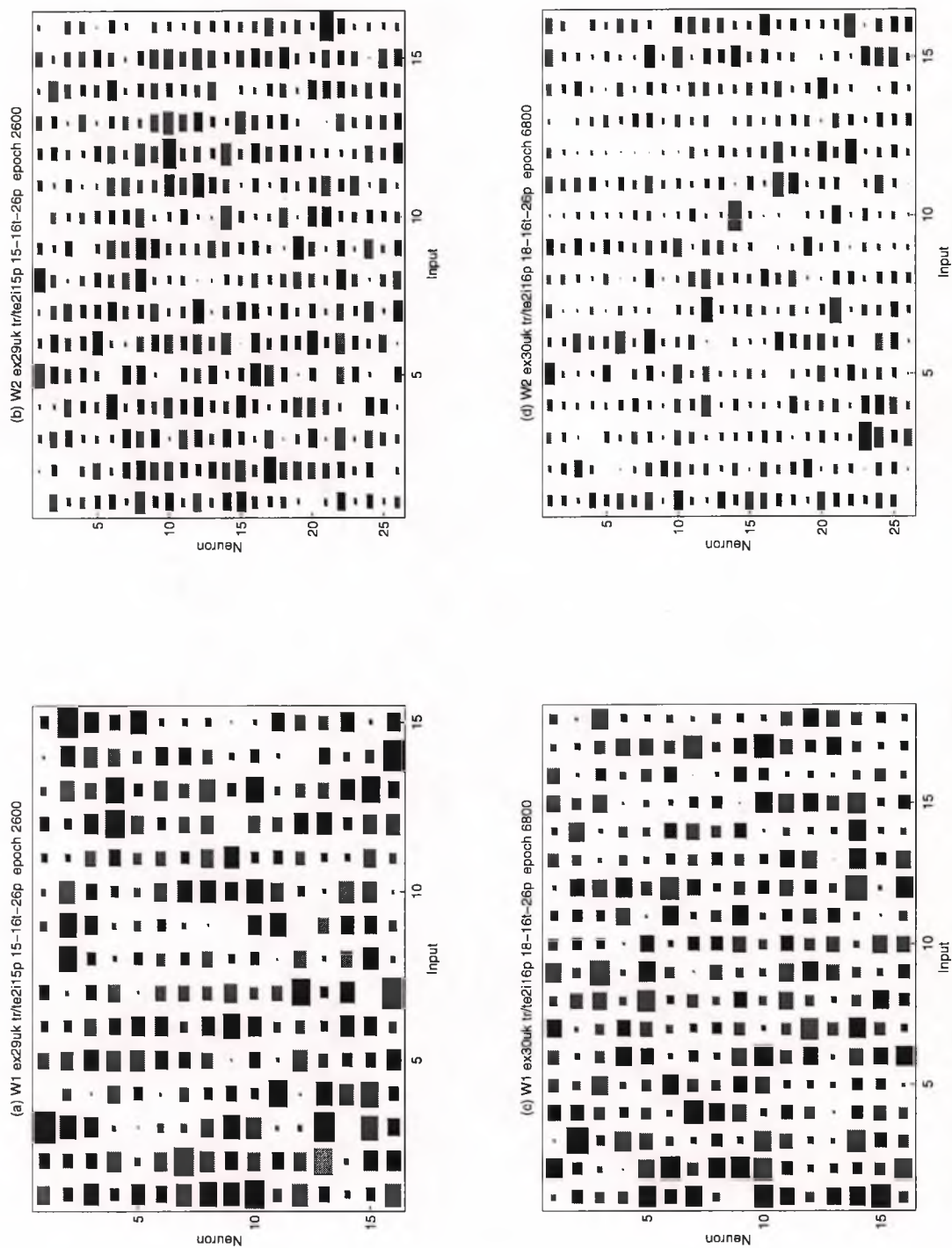


Figure D.3

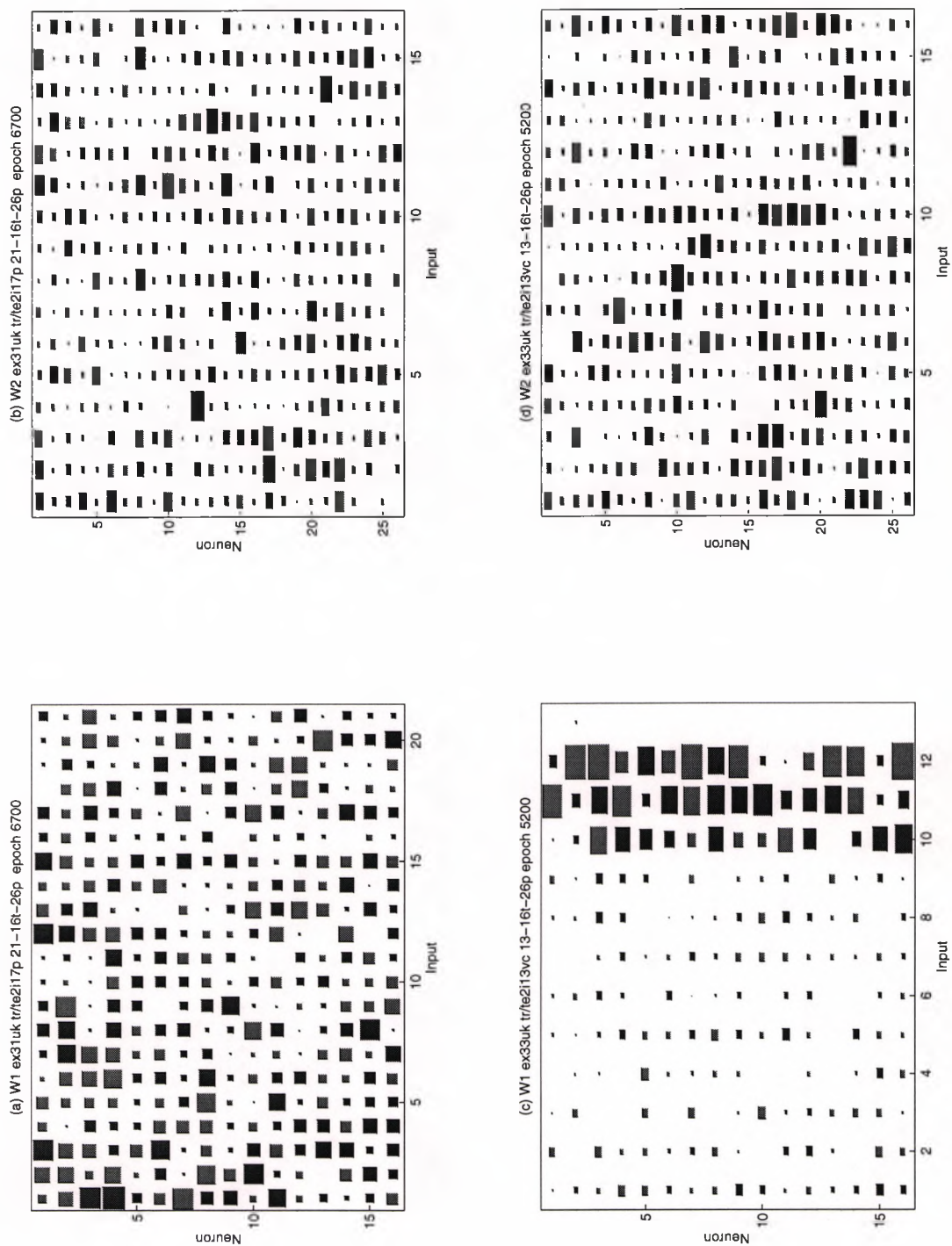


Figure D.4

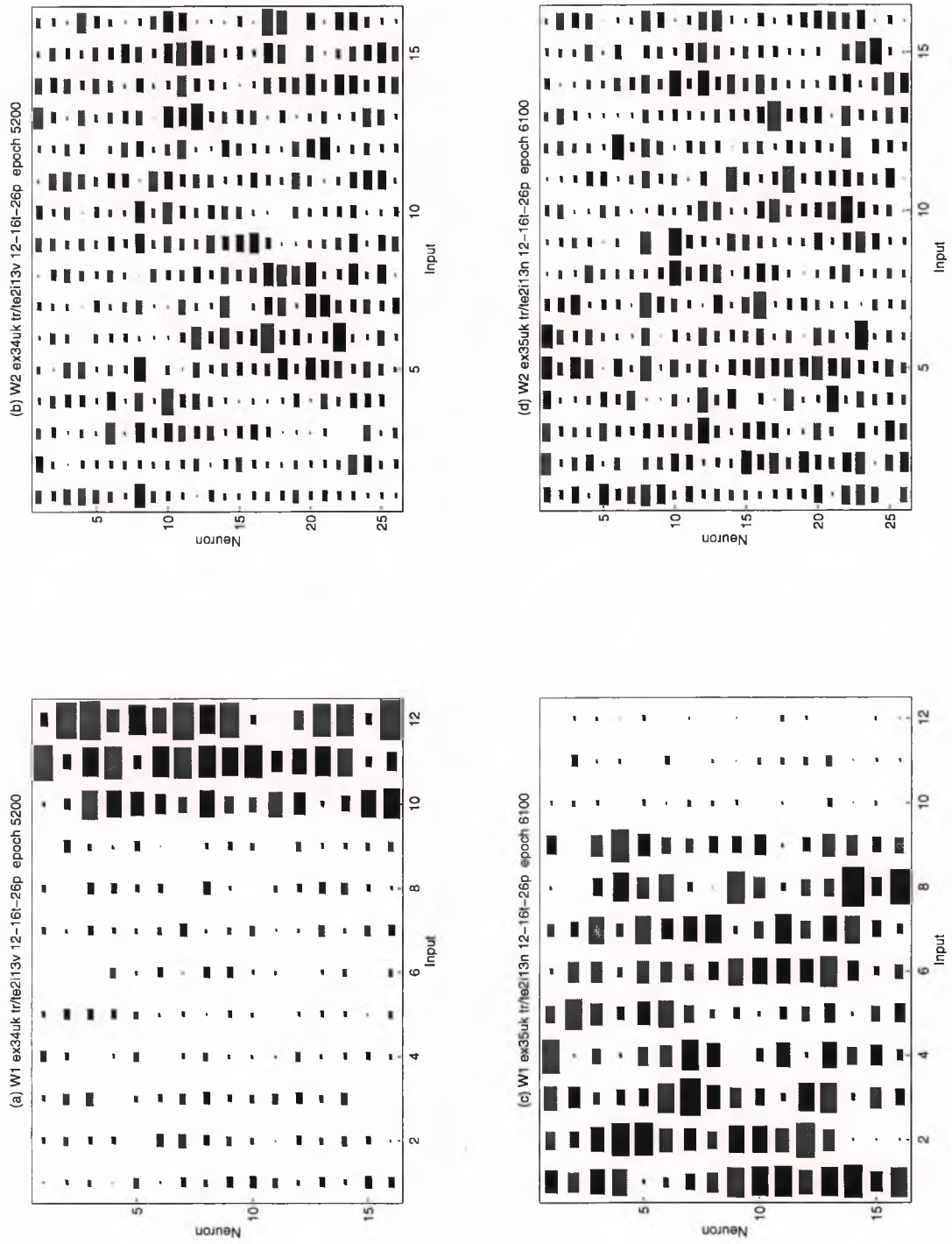


Figure D.5

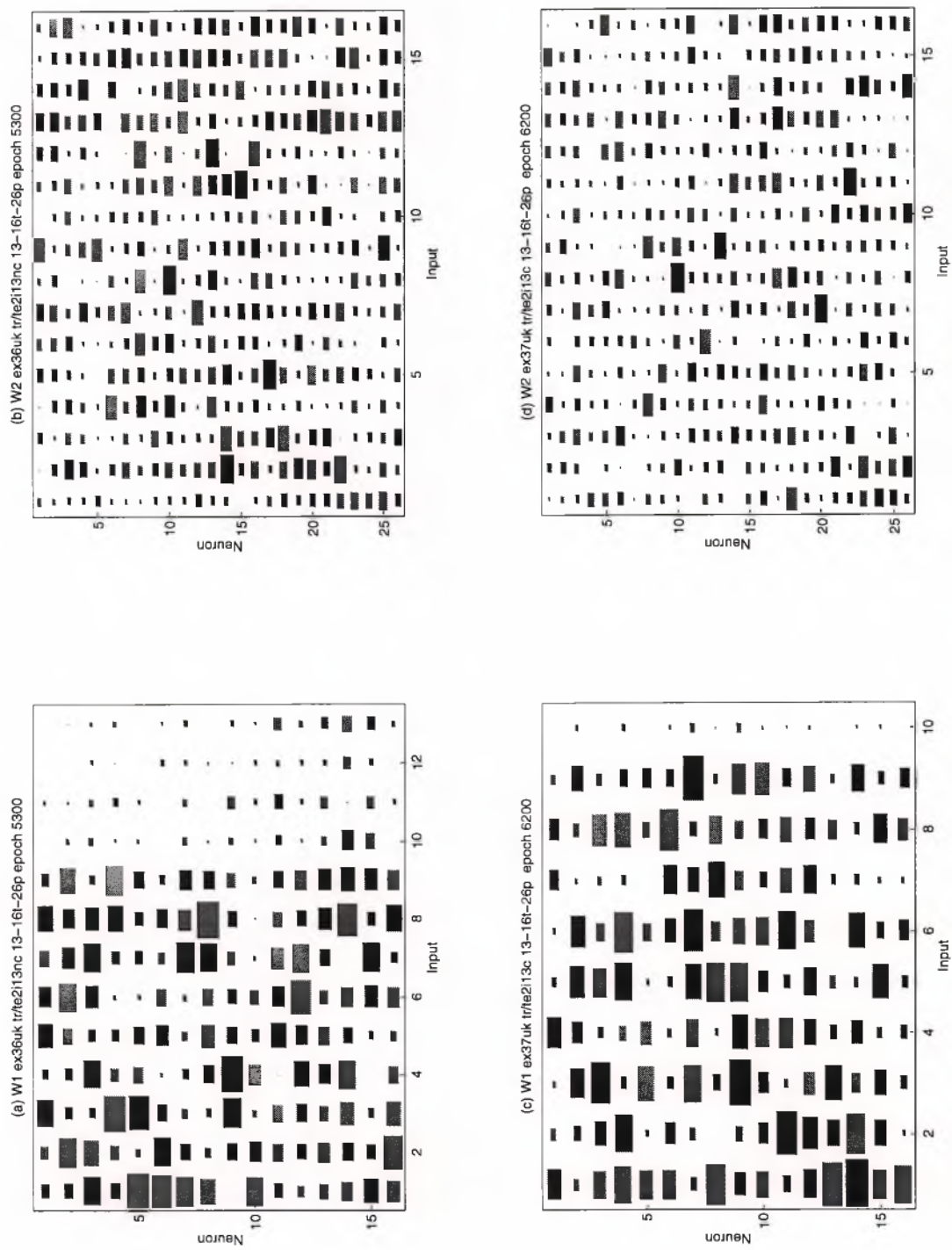


Figure D.6

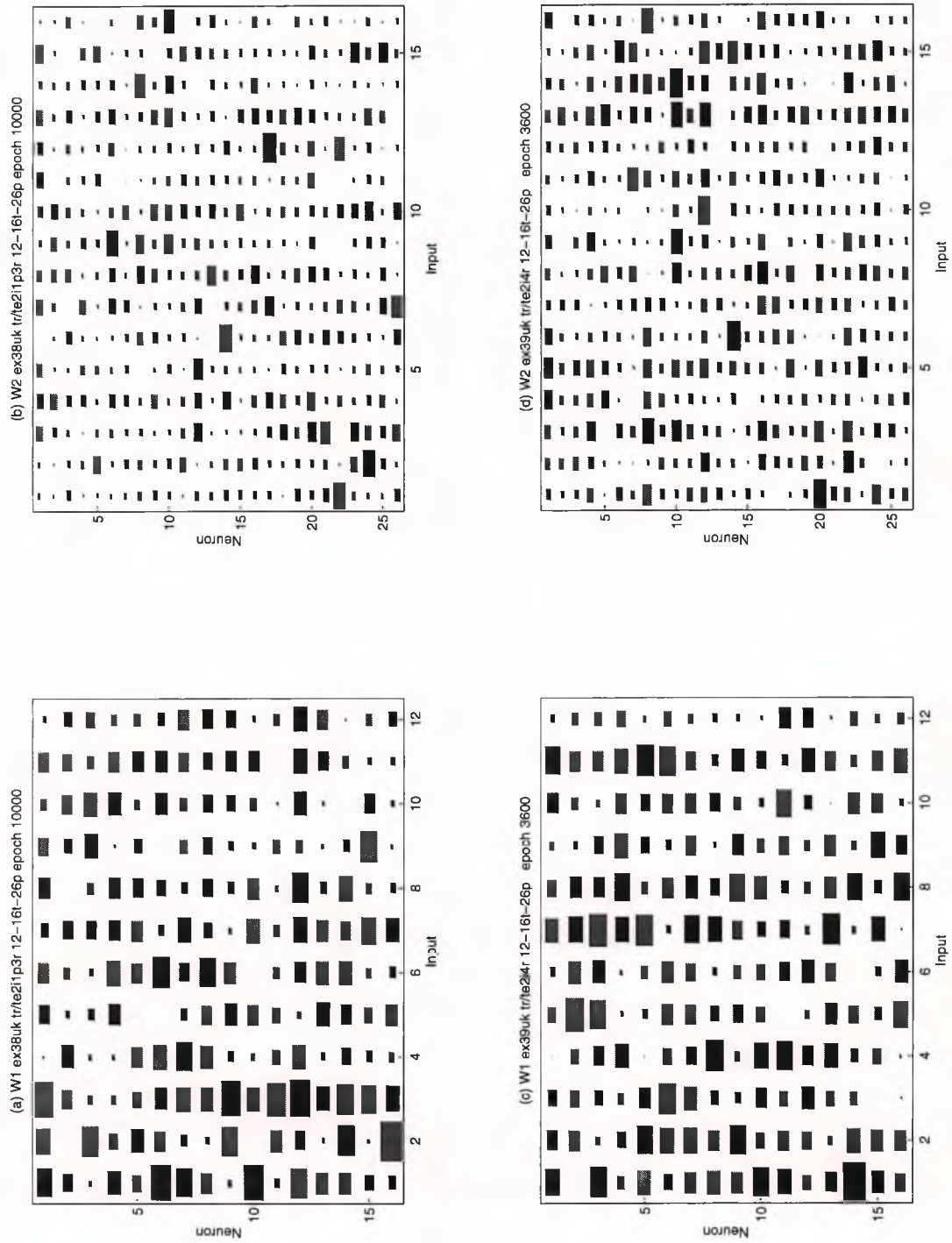


Figure D.7

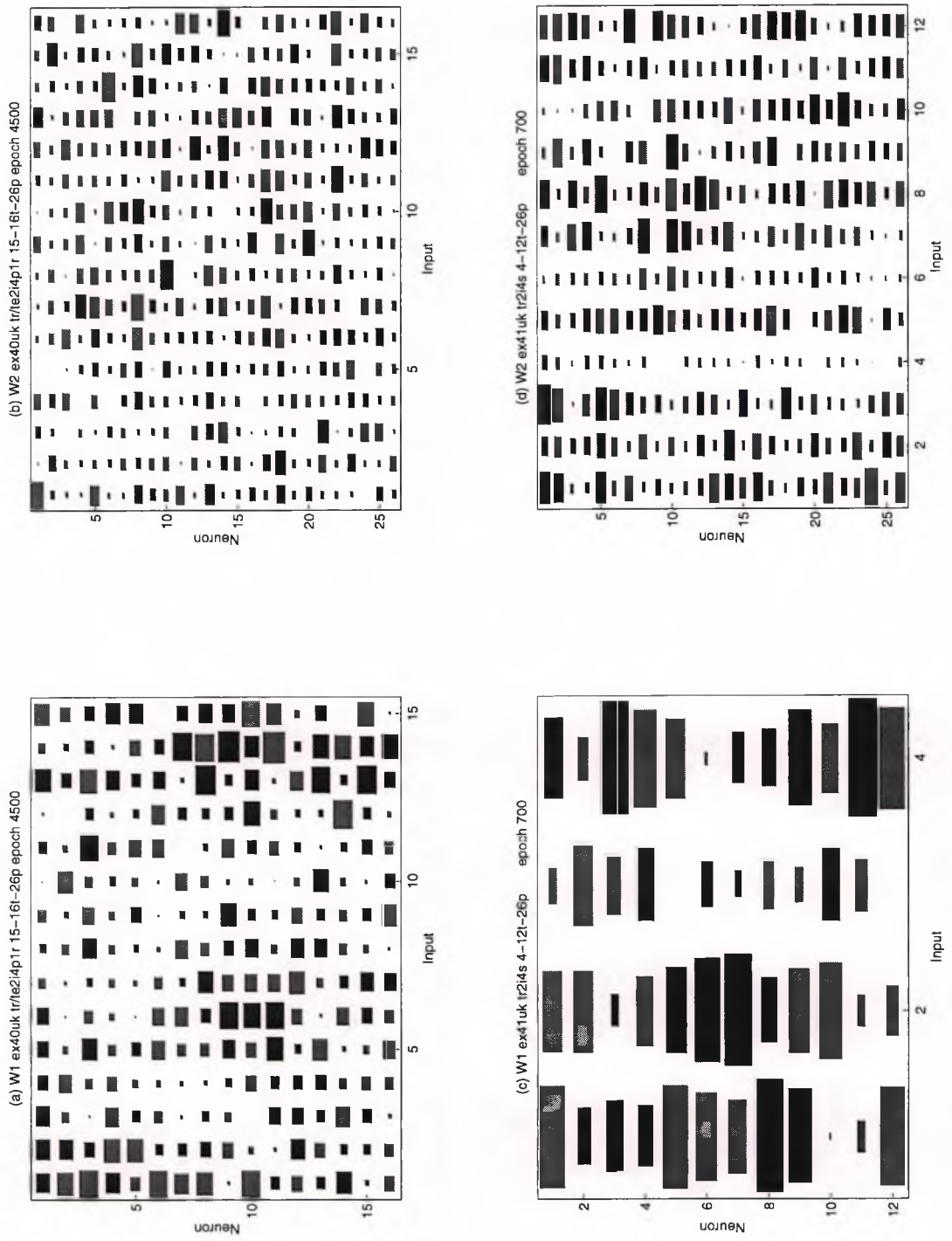


Figure D.8

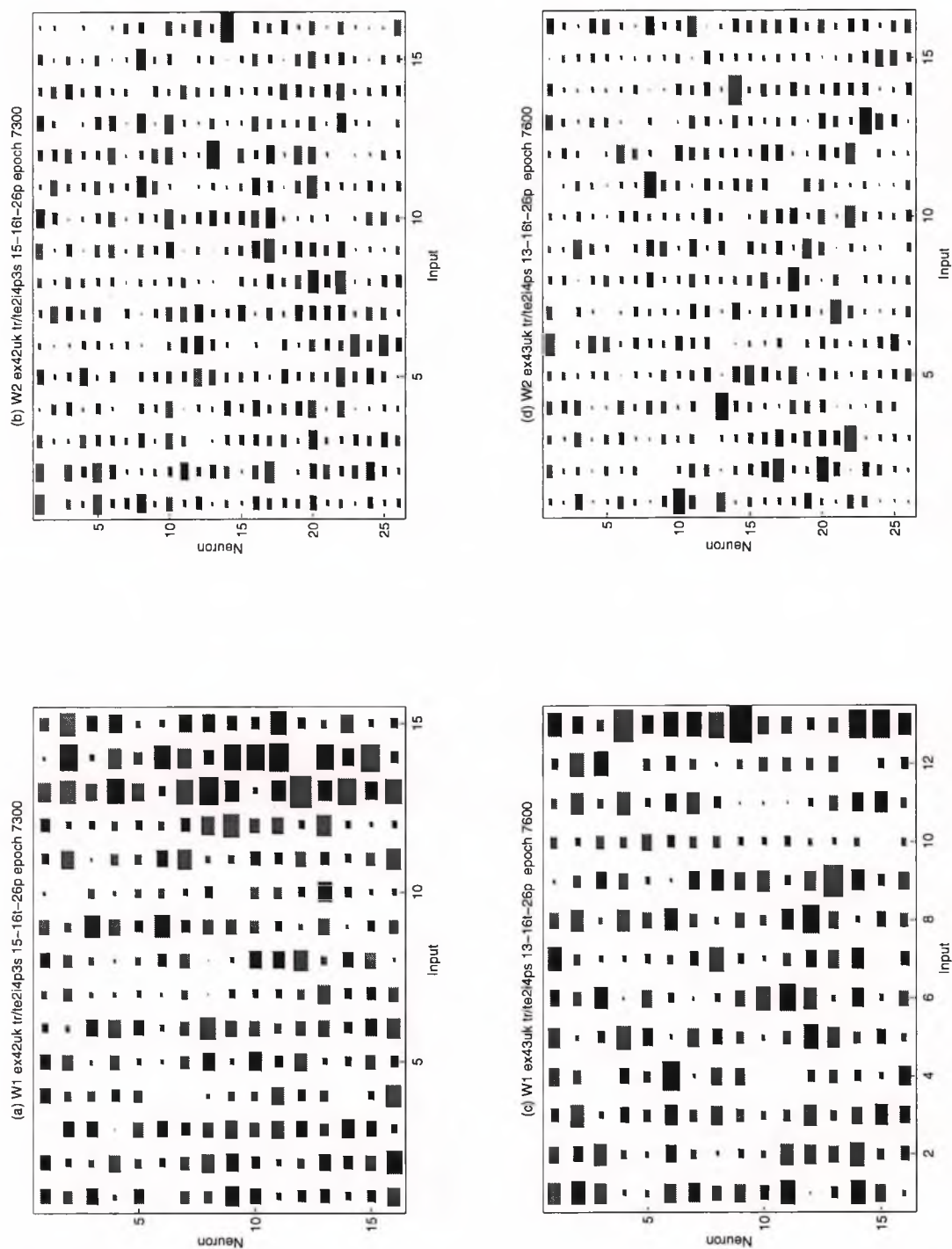


Figure D.9

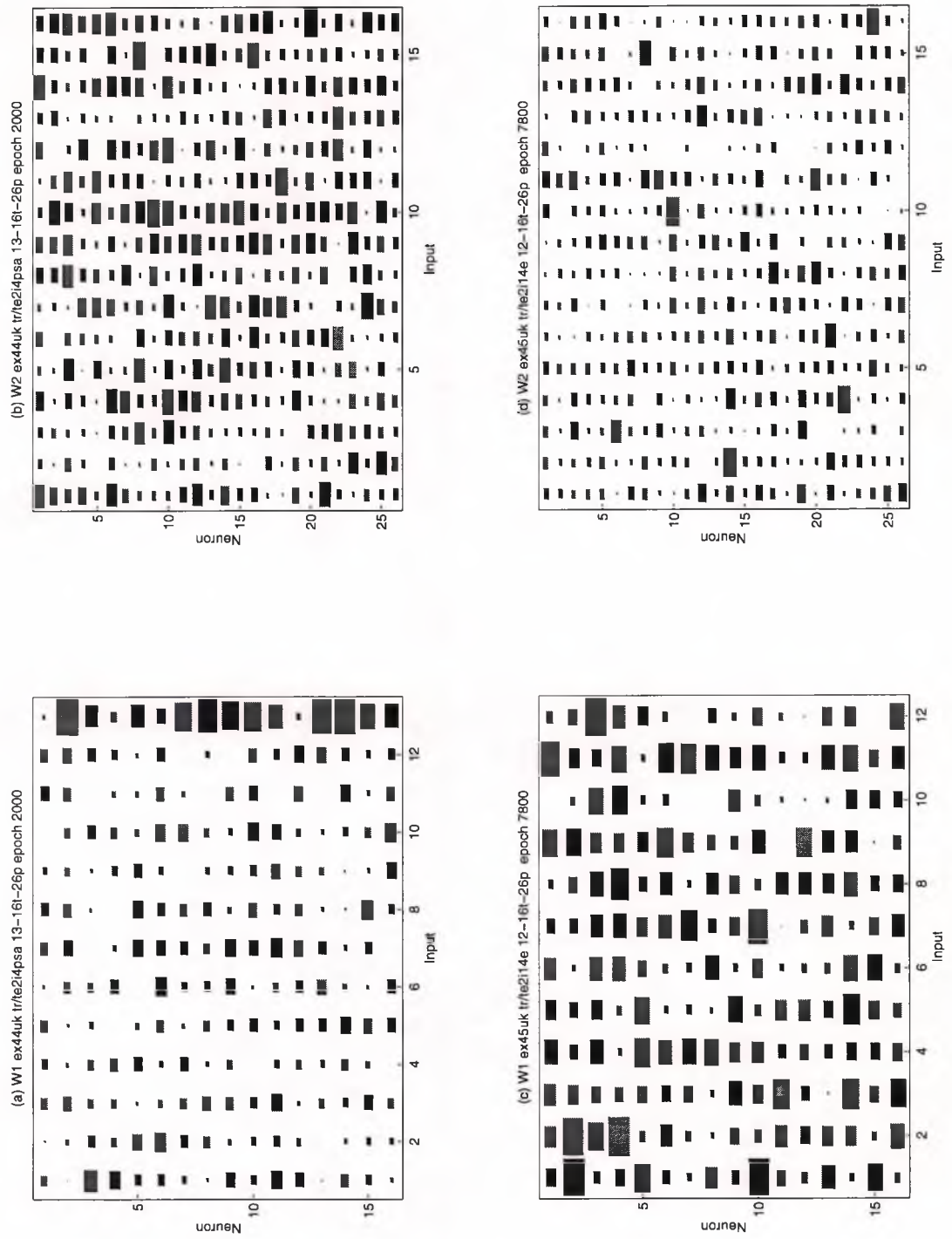


Figure D.10

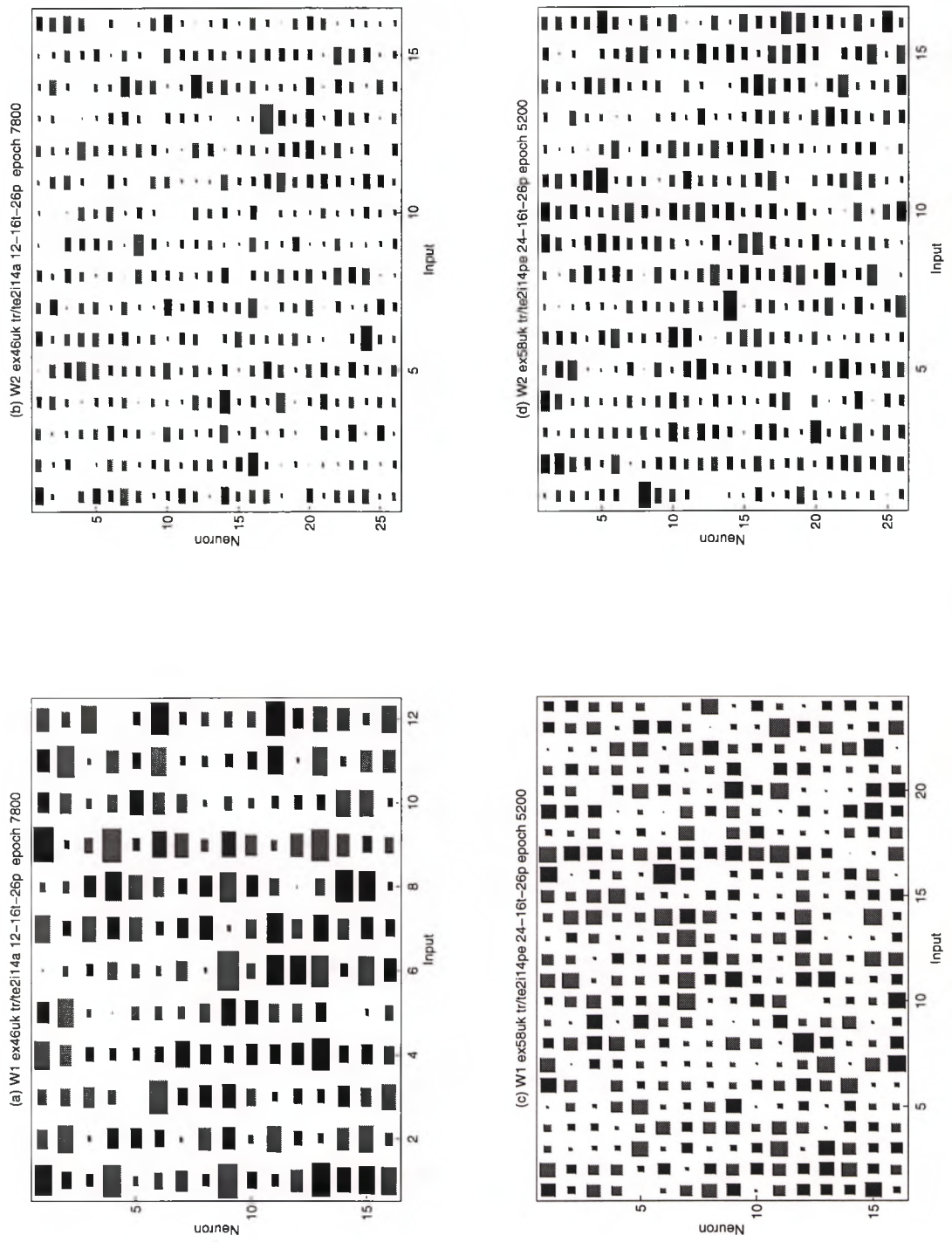


Figure D.11

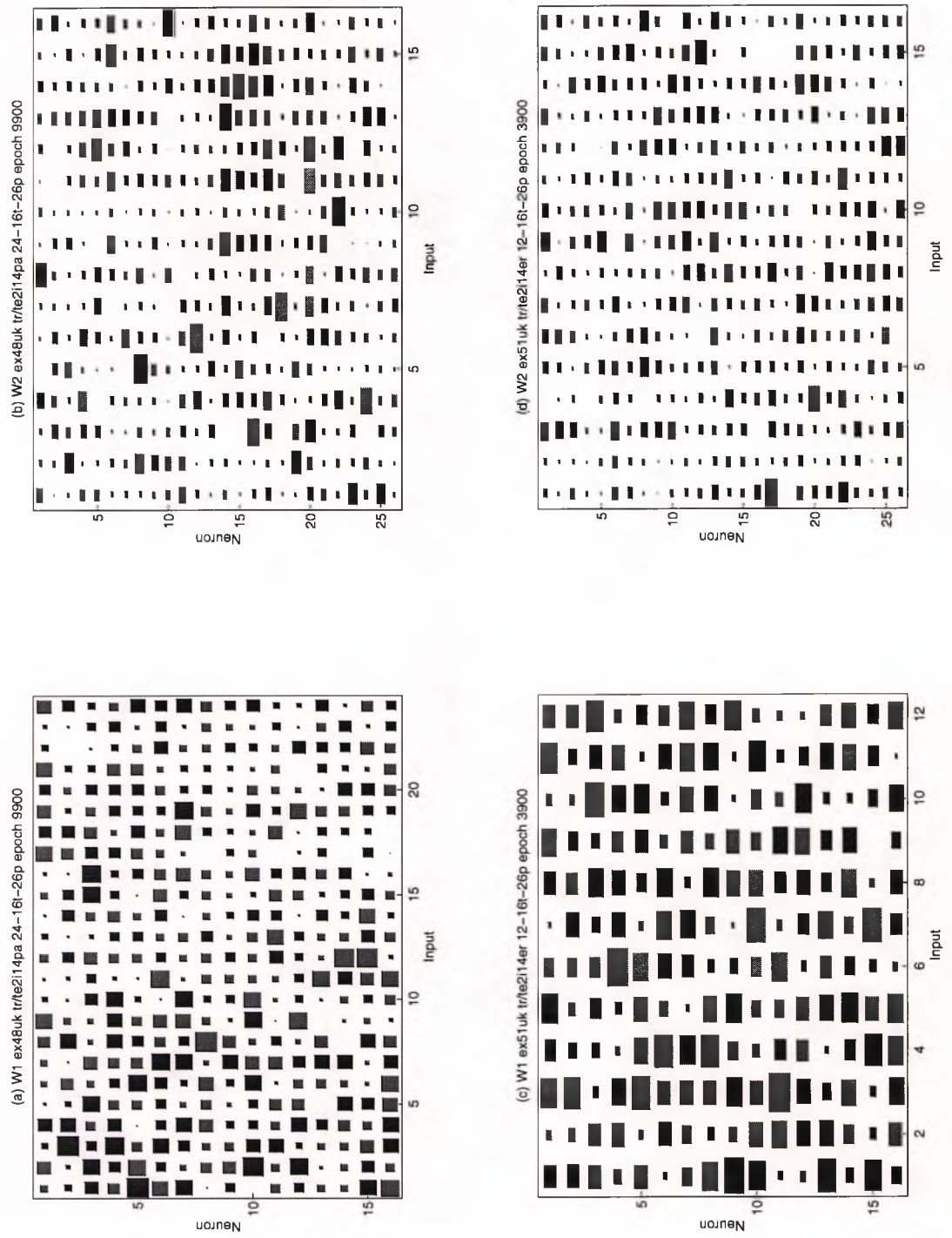


Figure D.12

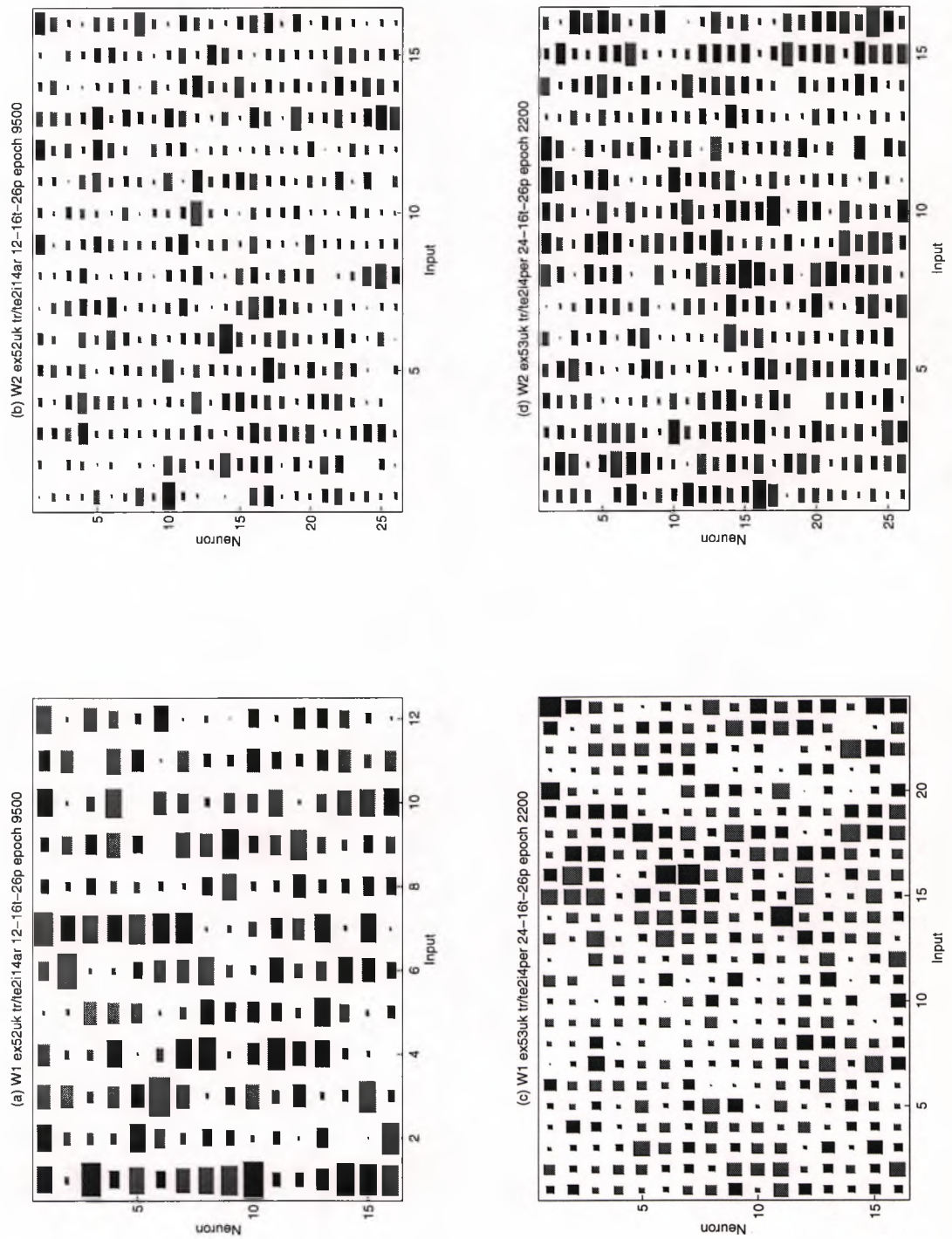


Figure D.13

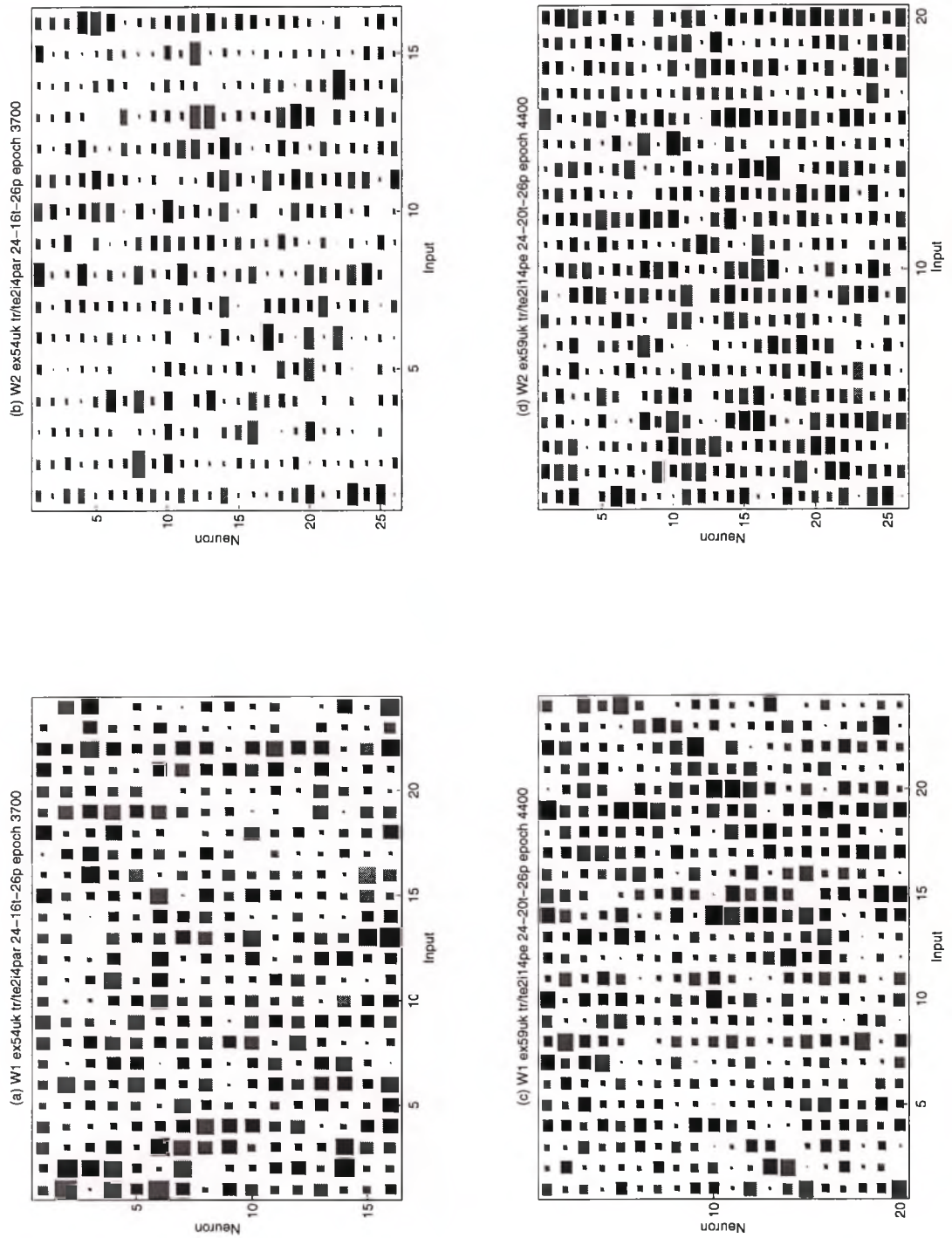


Figure D.14

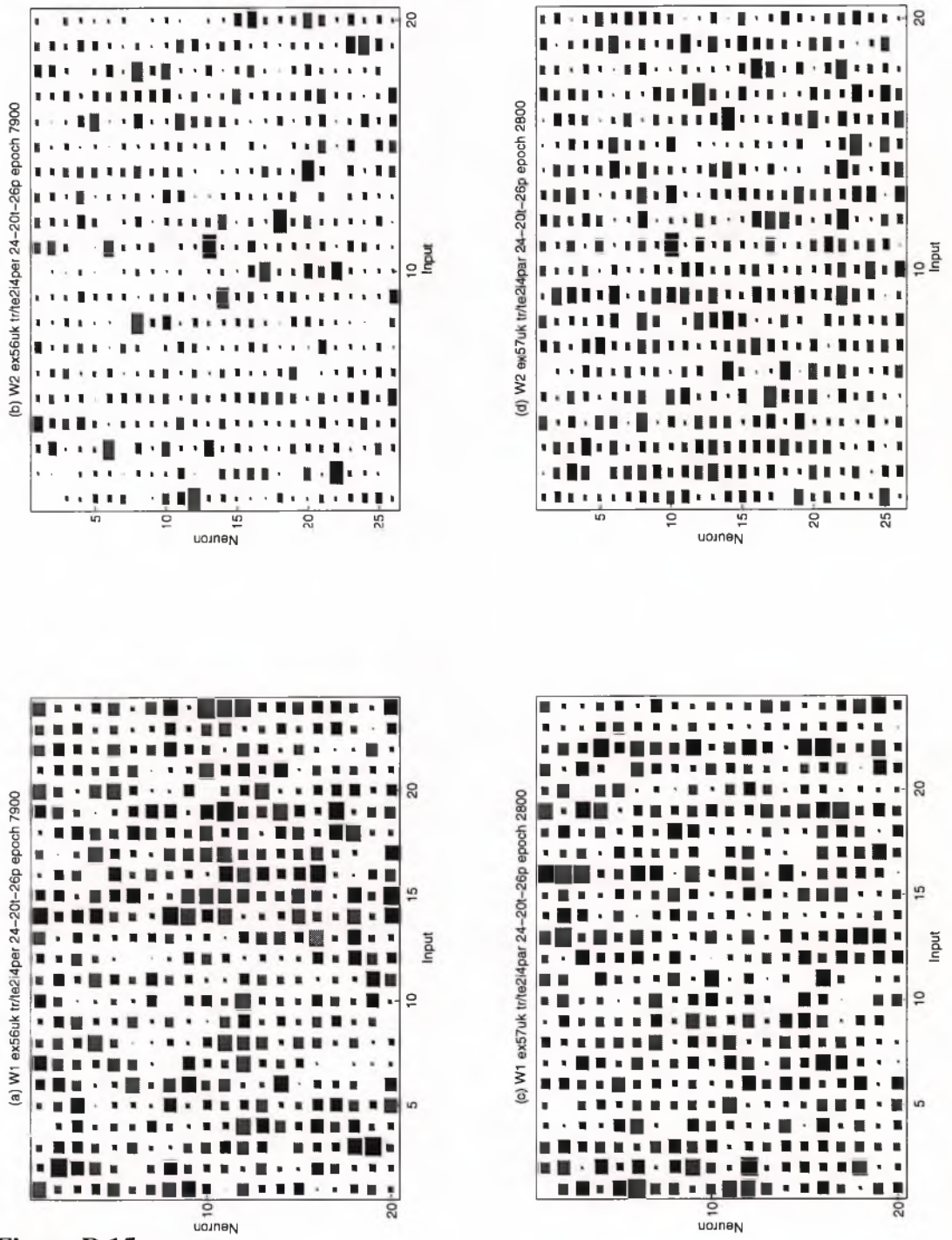


Figure D.15

References

- Acredolo, L., Goodwyn, S. (In Press). "Sign Language in babies: The significance of symbolic gesturing for understanding language development", In R. Vasta (ed.) *Annals of Child Development*, vol. 7, Greenwich, CN: JAI Press.
- Alm, N., Arnott, L. J., Newell, A. F. (1992). "Prediction and Conversational Momentum in an Augmentative Communication System", *Communications of the ACM*, vol. 35, no. 5.
- Anderson, J. A., E. Rosenfeld, (eds.). (1988). *Neurocomputing: Foundations of Research*. Cambridge, MA: MIT Press.
- Ascension Technology Corporation (1993). *The Flock of Birds Position and Orientation Measurement System installation and operation guide*, Ascension Technology Corporation, POB 527, Burlington, Vermont 05402.
- Bates, E., Camaioni, L., Volterra, V. (1975). "The Acquisition of Performatives Prior to Speech", *Merrill-Palmer Quarterly*, vol. 21, no.3. July. Merrill-Palmer Institute, Detroit, Michigan 48202, USA.
- Baker, R. B. (1982). "Minspeak: A semantic compaction system that makes self-expression easier for communicatively disabled individuals", *Byte*, pp.186-202.
- Baker, R. B., Nyberg, E. (1994). "Semantic Compaction: A Basic Technology for Artificial Intelligence in AAC", *5th Annual Minspeak Conference*, Nov 14-15.
- Basil, C. (1992). "Social Interaction and Learned Helplessness in Severely Disabled Children" *Augmentative and Alternative Communication*, vol. 8, no. 3 pp188-199.
- Beukelman, D. R., Mirenda, P. (1992). *Augmentative and Alternative Communication: Management of Severe Communication Disorders in Children and Adults*, Baltimore, Maryland, USA: Paul Brookes.
- Bichel, M. (ed.). (1995). *Proceedings of the International Workshop on Automatic Face and Gesture Recognition*, Zurich, Switzerland.
- Birdwhistell, R. L. (1952). *Introduction to Kinesics: An annotation system for analysis of body motion and gesture*, Kentucky, USA: University of Louisville.

- Bloom, L., Lahey, M. (1978). *Language Development and Language Disorders*, New York: John Wiley & sons.
- Bloomberg, K., Karlan, G., Lloyd, L. (1990). "The comparative translucency of initial lexical items represented by five graphic systems and sets" *Journal of Speech and Hearing Research*, vol.33, pp.717-725.
- Bolt, R. A., (1980). "Put-That-There": Voice and Gesture at the Graphic Interface", *Computer Graphics, Proceedings of SIGGRAPH' 80*, vol.14, no. 3, pp. 262-270. New York: ACM Press.
- Bolt, R. A., (1987). "The Integrated Multi-Modal Interface" *The Transactions of the Institute of Electronics, Information and Communication Engineers*, Japan, vol. J70-D, no. 11, pp. 2017-2025.
- Bolt, R. A., Herranz, E. J. (1992). "Two-handed Gestures with Speech in Multi-Modal Natural Dialogue" *Proceedings of UIST'92*. ACM, New York.
- Bryson, A. E., Ho, Y-C. (1969) *Applied Optical Control*, New York: Blaisdell.
- Buxton, W., Flume, E., Hill, R., Lee, A., Woo, C. (1983). "Continuous Hand-Gesture Input". *Proceedings of Graphics Interface*, pp.191-195.
- Buxton, W., Card, S. (1994). "Revolutionary Concepts in Human Computer Interaction" *The 61st Videoconference Seminar via Satellite*, May, IEEE, Inc, 445 Hoes lane, Piscataway, NJ 8855-1331, USA.
- Card, S., Moran, T. P., Newell, A. (1980). "The Keystroke-level model for user performance with interactive systems", *Communications of the ACM*, 23, pp. 396-410.
- Cauley, K. M., Golinkoff, R. M., Hirsh-Pasek, K., Gordon, L. (1989). "Revealing Hidden Competencies: A New Method for studying Language Comprehension in Children with Motor Impairments", *American Journal on Mental Retardation*, vol. 94, no. 1, pp.53-63.
- Chang, S. K., Polese, G., Abate, A. F., Baker, R. B. (1994) "An Iconic language Design Environment for Augmentative Communication". *Proceedings Sixth Biennial Conference of the International Society for Augmentative and Alternative Communication*, Maastricht, The Netherlands, pp 108-110.
- Churchland, P. M. (1989). *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*, Cambridge, Massachusetts, USA: MIT Press.
- Cruickshank W. M. (1976). "The Problem and its Scope" in Cruickshank W. M. (ed.) *Cerebral Palsy - A Developmental Disability. 3rd Revised Edition*, pp. 1-27, New York, USA: Syracuse University Press.
- Demuth, H., Beale, M. (1994). *Neural Network Toolbox: For use with MATLAB*, The MATHWORKS Inc., MA, USA.
- Doherty, J., Daniloff, J., Lloyd, L. (1985). "The Effect of Categorical Presentation on Amer-Ind Transparency", *Augmentative and Alternative Communication*, vol. 1, pp.10-16.
- Duffy, L. (1977). *An innovative approach to the development of communication skills of severely handicapped cerebral palsied children*. Master's Thesis. University of Nevada, Las Vegas.
- Efron, D. (1941). *Gesture and Environment*. Morningside Heights, New York: King's Crown Press.

- Ehn, P., Bodker, S., Greenbaum, J. (1993). Chapters in D. Schuler, A. Namioka (eds.), *Participatory Design: Principles and Practices*, New Jersey, USA: Lawrence Erlbaum.
- Ekman, P., Friesen, W. V. (1969). "The repertoire of nonverbal behavioural categories: Origins, usage and coding", *Semiotica*, 1, pp 49-98.
- Ekman, P., Huang, T.S., Sejnowski, J. T., Hager, C. J. (eds.) (1992) *Facial Expression Understanding*, Final Planning Workshop Report to NSF, Human Interaction Laboratory, 401 Parnassus Avenue, Box HIL-0984, University of California, San Francisco, CA 94143.
- Emmorey, K., Reilly, J. S. (1995). *Language, Gesture and Space*, Hillside, New Jersey, USA: Lawrence Erlbaum.
- Fausett, L. (1994) *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. New Jersey, USA: Prentice-Hall International.
- Feallock, B. (1958). "Communication for the Non-Vocal Individual" *American Journal of Occupational Therapy*, vol. 12, pp. 60-63.
- Fels, S. S., Hinton, E. G. (1990). "Building Adaptive Interfaces with Neural Networks", D.Diaper et al. (eds.), *Human-Computer Interaction - INTERACT'90*, Elsevier Science Publications B.V. North- Holland.
- Fels, S. S. (1994). Personal Communication.
- Foley, J. D., Wallace, V. L. (1974). "The Art of Natural Graphic Man-Machine Conversation", *Proceedings IEEE*, vol. 63, no.4. April, pp.462-470.
- Foulds, R. (1985). "Observations on interfacing nonvocal communication", In C. Barry & M. Byrne (Eds.), *Proceedings of the Fourth International Conference on Communication Through Technology for the Physically Disabled*, pp.46-51, International Cerebral Palsy Association, London.
- Fried-Oken, M., Creech, D. R., Baker, R. B. (1994). "Pragmatics issues for Augmented and Natural Speakers in Public Venues ". *Proceedings Sixth Biennial Conference of the International Society for Augmentative and Alternative Communication*, Maastricht, The Netherlands, pp 186-187.
- Fuller, P., Wright, A. (1994). "The Beauty of the Unspoken: The Development of Language without either recognition or production of speech. *Proceedings Sixth Biennial Conference of the International Society for Augmentative and Alternative Communication*, Maastricht, The Netherlands, pp. 89-90.
- Funahashi, K. (1989). "On the Approximate Realization of Continuous Mappings by Neural Networks" *Neural Networks*, vol. 2, no.3, pp.183-192.
- Gerber, S., Kraat, A. (1992). "Use of a Developmental Model of Language Acquisition: Applications to Children using AAC Systems", *Augmentative and Alternative Communication*, vol. 8, no. 1, pp. 19-32.
- Gersh, E. S. (1991). "What Is Cerebral Palsy" in Geralis, E. (ed.) *Children with Cerebral Palsy: A Parents' Guide*, pp. 1-32. USA: Woodbine House Inc.
- Giles, C. L.(1994). Personal Communication.
- Goble, R., Suarez, P. F., Rogers, S. S., Ruck, W.D., Arndt, C., Kabrisky, M. (1993). "A Facial Feature Communication Interface for the Non- Vocal" *IEEE Engineering in Medicine and Biology*, Sept. pp. 46-48, cont. p82.
- Goldberg, H. R., Fenton, J. (1960). *Aphonic Communication for those with Cerebral Palsy: Guide for the Development and Use of a Communication Board*. New York:

United Cerebral Palsy of New York State.

Golinkoff, R. M., Hirsh-Pasek., Cauley, K., Gordon, L. (1987). "The eyes have it: Lexical and semantic comprehension in a new paradigm", *Journal of Child Language*, vol. 14, pp. 23-45.

Gould, J. D., Conti, J., Hovanyecz, T. (1983) "Composing Letter with a Simulated Listening Typewriter", *Communications of the ACM*, vol. 26, pp. 295-308.

Granger, A. (1993), Personal Communication, Wright Patterson Air Force Base, Dayton, Ohio, USA.

Grayden, D. B., Scorilis, M. S. (1993). "Phonemic Segmentation of Fluent Speech", *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 1, pp. 73-76.

Greenbaum, J., Kyng, M. (1995). "The Design Challenge: Creating a Mosaic out of Chaos", *Proceedings ACM CHI'95*, Denver, Colorado, USA. pp.195-196.

Grove, N., Dockrell, J. (1994). "Multi-sign utterances: Strings of Necklaces? Analysing the linguistic skills of signers with learning disabilities", *Proceedings Sixth Biennial Conference of the International Society for Augmentative and Alternative Communication*, Maastricht, The Netherlands, pp 285-286.

Hagberg, B., Hagberg, G., Olow L. (1975). "The Changing Panorama of Cerebral Palsy in Sweden 1954-1970. I-II", *Acta Paediatrica Scandanavica*, vol. 64, pp. 187-199.

Hamre- Nietupski, S., Stoll, A., Holtz, K., Fullerton, P., Ryan-Flottum, M., Brown, L. (1987). "Curricular strategies for teaching selected nonverbal communication skills to nonverbal and verbal severely handicapped students ". in L.Brown, J. Nietupski, S. Lyon, S. Hamre-Nietupski, T. Crouner, L Gruenewald (eds.), *Curricular Strategies for teaching functional object use, nonverbal communication, problem solving, and meal-time skills to severely handicapped students*, vol. 2, part 1, pp. 94-250. University of Wisconsin-Madison and Madison Metropolitan School District, Madison, USA.

Hardy, J. C. (1983). *Cerebral Palsy*, New Jersey, USA: Prentice-Hall.

Harrington, M. E., Daniel, R. W., Kyberd, P. J. (1995). "Gesture Recognition of Arm Movements using Accelerometers", In *The European Context for Assistive Technology: Proceedings of the 2nd TIDE Congress*, Porrero, I., Puig de la Bellacasa, R. (eds.), pp. 432-435, Amsterdam, Netherlands: IOS Press.

Harwin, W. (1990). *Computer Recognition of the Unconstrained and Intentional Head Gestures of Physically Disabled People*. Ph.D. thesis: University of Cambridge, Department of Engineering.

Heim, M, J., Mills, E. A. (1994). "Early Development of Symbolic Communication and Linguistic Complexity Through AAC", *Proceedings Sixth Biennial Conference of the International Society for Augmentative and Alternative Communication*, Maastricht, The Netherlands, pp. 224-226.

Hiraiwa, A., Shimohara, K., Tokunaga, Y. (1990). "EEG Topography Recognition by Neural Networks", *Engineering in Medicine and Biology*, September.

Hirsbrunner, H. P., Frey, S., Crawford, R. (1987). "Movement in human interaction: description, parameter formation, and analysis", In A.W. Siegman and S. Feldstein (eds.), *Nonverbal Behaviour and Communication*, pp.99-140, Lawrence Erlbaum Associates, Hillsdale, USA.

Hjelmquist, E., Sandberg, A. D. (1994). "The Development of Communicative and

- Metacommunicative skills among non-speaking persons", *Proceedings Sixth Biennial Conference of the International Society for Augmentative and Alternative Communication*, Maastricht, The Netherlands, pp 227-228.
- Houghton, J., Bronicki, B., Guess, D. (1987). "Opportunities to express preference and make choices among students with severe disabilities in classroom settings", *Journal of the Association for Persons with Severe Handicaps*, vol. 11, pp. 225-265.
- Huer, M. (1987) "White's Gestural System for the Lower Extremities", *Communicating Together*, 5, pp. 3-4.
- Johnston, D. (1995) Personal Communication
- Kates, B., McNaughton, S. (1975). *The First application of Blissymbolics as a communication medium for non-speaking children: History and development, 1971-1974*. Easter Seals Communication Institute, Ontario, Canada: Don Mills.
- Kendon, A. (1988). "How Gestures can become more like words", In: F. Poyatos (ed.), *Cross Cultural Perspectives in Non-verbal Communication*, pp. 131-141, Toronto, Canada: Hogrefe.
- Kim, J. (1988). "On-Line Gesture Recognition by Feature Analysis". *Proceedings of Vision Interface*, June, pp. 51-55.
- Kohonen, T. (1989). *"Self-Organization and Associative Memory"* 3rd edition, Berlin: Springer-Verlag.
- Kolmogorov, A. N. (1963). "On the Representation of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Addition" *Doklady Akademii Nauk SSSR*, vol. 144 pp. 679-681. (*American Mathematical Society Translation*, vol. 28 pp. 55-59).
- Koons, D. B., Sparrell, J. C. (1994). "Iconic: Speech and Depictive Gestures at the Human-Machine Interface," *ACM CHI'94*, Conference Companion, Boston, Massachusetts, USA.
- Kraat, A. W. (1985). *Communication Interaction Between Aided and Natural Speakers: An IPCAS study report*. Toronto: Canadian Rehabilitation Council for the Disabled.
- Kraat, A. W. (1991). "Methodological issues in the study of language development for children with AAC systems", In J. Brodin & E. Bjorck- Akesson (eds.), *Methodological issues in research in augmentative and alternative communication: Proceedings from the First ISAAC Research Symposium in Augmentative and Alternative Communication* pp. 118-123. Stockholm, Sweden: Swedish Handicap Institute.
- Kruger, M. A. (1993) "The Experience Society", *Presence*, vol.2, no.2, Spring, pp. 162-168. USA: MIT Press.
- Kurokawa, T. (1992) "Gesture Coding and a gesture Dictionary for a Nonverbal Interface", *IEICE Transactions on Fundamentals of Electronics and Computer Science*, February, vol. E75-A, no. 2.
- Kurtenbach, G, Buxton, W. (1991). "GEDIT: A Test Bed for Editing by Contiguous Gesture". *SIG-CHI Bulletin*, vol. 3, no. 2, April, pp. 22-26.
- Kurtenbach, G, Hulteen, E. A. (1990). "Gestures in Human-Computer Communication". in Laurel, B. (Ed.), *The Art of Human-Computer Interface Design*, pp. 309-496, Reading, Massachusetts, USA: Addison-Wesley Inc.
- Kyng, M., Muller, M. J., Suchman, L. (1995) "Representations of Work", *Communications of the ACM*, USA: ACM, September, vol.38, no.9.

- LeCun, Y (1986). "Learning Processes in an Asymmetric Threshold Network" In E. Bienenstock, F. Fogelman-Souli, & G. Weisbuch, eds. *Disordered Systems and Biological Organization*. NATO ASI Series, F20, Berlin: Springer-Verlag.
- Leroi-Gourhan, A (1993) *Gesture and Speech*, English version translated by Berger, A. B. MIT Press, MA, USA from (1963) original in French: *Le Geste et la parole*, Paris: Albin Michel.
- Letto, M., Bedrosian, J. L., Skarakis-Doyle, E. (1994). "Application of Vygotskian Developmental Theory to Language Acquisition in a Young Child with Cerebral Palsy", *Augmentative and Alternative Communication*, vol.10. no. 3. pp. 151- 159.
- Levelt, W. J. M. (1994) "What Can A Theory of Normal Speaking Contribute to AAC", *Proceedings Sixth Biennial Conference of the International Society for Augmentative and Alternative Communication*, Maastricht, The Netherlands, pp 18-20.
- Levitt, S. (1995) *Treatment of Cerebral Palsy and Motor Delay. Third Edition*. Oxford, UK: Blackwell Scientific Publications.
- Light, J., Collier, B., Parnes, P. (1985). "Communicative Interaction between Young Non-speaking Physically Disabled Children and their Primary Caregivers: Part III - Modes of Communication", *Augmentative and Alternative Communication*, vol. 1, no. 4, pp.125-133.
- Light, J. (1988). "Interaction Involving Individuals using Augmentative and Alternative Communication Systems: State of the Art and Future Directions". *Augmentative and Alternative Communication*, vol.2, pp. 98-107.
- Lippman, R. P. (1989). "Review of Neural Networks for Speech Recognition", *Neural Computation*, vol. 1, pp. 1-38.
- Lipscomb, J. S. (1991). "A Trainable Gesture Recogniser". *Pattern Recognition*, vol. 24, no. 9, pp. 895-907.
- Lloyd, L. L., Quist, R.W., Windsor, J. (1990). "A proposed AAC model", *Augmentative and Alternative Communication*, vol. 6, pp. 172- 183.
- Mammone, R. J. (Ed.) (1994) *Artificial Neural Networks for Speech and Vision*, London: Chapman & Hall.
- MathWorks Inc. (1992). *MATLAB Reference Guide*, Natick, Mass., USA: MathWorks Inc.
- Maybury, M. T. (1993). *Intelligent Multimedia Interfaces*, Menlo Park, Massachusetts, USA: AAAI/MIT Press.
- McClelland, J. L., D. E. Rumelhart (1988). *Explorations in Parallel Distributed Processing*. Cambridge, MA, USA: MIT Press.
- McDonald, E. T. (1987) "Cerebral Palsy: Its nature, pathogenesis, and management". In E.T. McDonald (ed.), *Treating cerebral palsy: For Clinicians by clinicians*, pp. 1-20, Austin, TX: PRO-ED.
- McLaughlin, J. A., Lewis, R. B. (1986). *Assessing Special Students*, Ohio, USA: Merrill.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago, USA: The University of Chicago Press.
- Meyer, K., Applewhite, H. L., Biocca, A. F. (1992). "A Survey of Position Trackers", *Presence*, Vol.1, No. 2, pp. 173-200, MA, USA: MIT Press.
- Minsky, M. L., & S. A. Papert (1969). *Perceptrons*. Cambridge, MA, USA: MIT Press.

- Morrel-Samuels, P. (1990). "Clarifying the Distinction between Lexical and Gestural Commands". *International Journal of Man-Machine Studies*, vol. 32, pp. 581-590.
- Mumford, L. (1934). *Technics and Civilization*, New York, USA: Harcourt Brace and World.
- Murphy, A. (1994). "Event to Remember", *Communicating Together, Sharing to Learn*, Vol. 12, No.3, September, Communicating Together, P.O. Box 989, Ontario, Canada L3T, 4A5: Thornhill, ISSN No. 822-0683.
- Murphy, J., Markova, I., Moodie, E., Scott, J., Boa, S. (1995) "Augmentative and Alternative Communication Systems Used by People with Cerebral Palsy in Scotland: Demographic Study". *Augmentative and Alternative Communication*, vol.11. no. 1. pp. 26- 36.
- Musselwhite, C. R., St. Louis K. W. (1982). *Communication Programming for the Severely Handicapped: Vocal and Non-Vocal Strategies*, San Diego, California, USA: College Hill Press.
- Negroponte, N. (1995) *Being Digital*, London: Hodder & Stoughton.
- Newell, A. F. (1992) "On the Validity of User- Modelling in AAC: Comments on Horstmann and Levine (1990)", Forum, Letter to the Editor, *Augmentative and Alternative Communication*, vol. 8. no. 2. pp. 89- 92, with response from Horstmann and Levine, pp. 92- 97.
- Newell, A. F., Arnott, J. L., Carter, K., Cruickshank, G. (1990). "Listening Typewriter Simulation Studies", *International Journal of Man-Machine Studies*, vol. 33, pp. 1-19.
- Newell, A.F., Arnott, J. L., Booth, L., Beattie, W., Brophy, B., Ricketts, I. W. (1992). "Effect of 'PAL' Word Prediction System on the Quality and Quantity of Text Generation". *Augmentative and Alternative Communication*, vol.8. no. 4. pp. 304-311.
- Newell, A. F. (1993). "Designing for Extraordinary Users", *Ergonomics and Design*, pp.10-16, Santa Monica, CA, USA: Human Factors and Ergonomics Society.
- Nguyen, D., B. Widrow (1990). "Improving the Learning Speed of Two-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights" *International Joint Conference on Neural Networks*, San Diego, CA, vol 3. pp. 21-26.
- Nicolis, G., Prigogine, I. (1989). *Exploring Complexity: An Introduction*, New York, New York, USA: W. H. Freeman and Company.
- Nolan, C. (1981). *Dam-Burst of Dreams*, London: Weidenfeld and Nicolson.
- Norman, D. A. (1988). *The Psychology of Everyday Things*, USA: Harper Collins.
- Ohki, M., Sagawa, H., Sakiyama, T., Oohira, E., Ikeda, H., Fujisawa, H. (1994) "Pattern Recognition and Synthesis for Sign Language Translation System", *ACM, ASSETS'94*, Marina Del Rey, CA, USA, pp.1-8.
- Parker, D. (1985). *Learning Logic*, Technical Report TR-87, Center for Computational Research in Ergonomics and Management Science, Cambridge, MA, USA: MIT Press.
- Pausch, R. (1991). "Virtual Reality on five dollars a day", *ACM SIGCHI Conference Proceedings*, New Orleans, pp.265-270.
- Pelachaud, C., Badler.N. (1993). *Facial Animation Workshop (NSF)*. Graphics Laboratory, University of Pennsylvania, Philadelphia, USA.
- Perricos, C. (1993). Personal communication, Department of Engineering, University of Cambridge.
- Perricos, C., Jackson. R. D. (1994). "A Head Gesture Recognition System for Computer

- Access", *Proceedings RESNA '94*, June, Nashville, Tennessee, USA, pp.92-94.
- Phelps, W. M. (1950). "Etiology and diagnostic classification of cerebral palsy", *Proceedings of the Cerebral Palsy Institute*, New York: Association for Aid of Crippled Children, Inc.
- Phillips, B. (1993). "Technology Abandonment form the Consumer Point of View", NARDIC (National Rehabilitation Information Center), Quarterly, vol. 3, nos. 2 & 3, Washington, D.C, USA.
- Rime, B., Schiaratura, L. (1991). "Gesture and Speech." in Feldman & Rime (eds.), *Fundamentals of Non-verbal Behavior*, New York, USA: Press Syndicate of University of Cambridge.
- Riseborough, M. G. (1982). "Meaning in Movement: An investigation into the Interrelationship of Physiographic Gestures and Speech in Seven-year-olds" *British Journal of Psychology*, vol. 73, pp. 497-503.
- Rogers, S. K., Ruck, D. W., Priddy, K. (1992). "Neural Network for Segmentation, Tracking and Identification", *Applications of Artificial Neural Networks III SPIE*, vol. 1709 pp. 2-8.
- Rowland, C. (1990). "Communication in the classroom for children with dual sensory impairments. Studies of teacher and child behaviour", *Augmentative and Alternative Communication*, vol.6. pp. 262-274.
- Rubine, D. H. (1991). *The Automatic Recognition of Gestures*. Ph.D. Thesis, Carnegie-Mellon University, Pittsburg, USA.
- Rumelhart, D. E., G. E. Hinton, & R. J. Williams (1986). "Learning Internal Representations by Error Backpropagation" In D. E. Rumelhart & J. L. McClelland, (eds) (1988). *Parallel Distributed Processing*, vol. 1 chapter 8 pp. 318-362.
- Sacks, O. (1989). *Seeing Voices, A journey into the World of the Deaf*, University of California Press, Berkeley.
- Sayre, M. J. (1963). "Communication for the Non-Vocal." *Cerebral Palsy Review*, pp. 3-8.
- Schmandt, C. M., Hulteen, E. A. (1982). "The Intelligent Voice- Interactive Interface", *Proceedings of Human Factors in Computing Systems*, March15-17, pp. 363-366.
- Schuler, D., Namioka, A.(1993). (Eds.) *Participatory Design: Principles and practices*, Hillside, New Jersey, USA: Lawrence Erlbaum.
- Scherer, M. J. (1993) *Living in the State of Stuck*, Massachusetts, USA: Brookline Books.
- Scope, (1995) Personal communication. Scope, Library, London.
- Seligman, M. E. P. (1975). *Helplessness*. New York: W. H. Freeman.
- Shibata, A., Yunogawa, T., Kurokawa, T. (1984). "Measurement of three dimensional posture from two dimensional human image", *Proceedings Kansai Branch Conference of Ergonomics Society of Japan*, pp. 17-24.
- Shneiderman, B. (1980). *Software Psychology: Human Factors in Computer and Information Systems*, Boston, MA, USA: Little Brown,
- Shneiderman, B. (1992). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. USA: Addison Wesley.
- Skelly, M., Schinsky, L., Smith, R., Donaldson, R., Griffin, P. (1975). "American Indian

- Sign: A Gestural Communication for the Speechless", *Archives of Physical and Rehabilitation Medicine*, vol. 56, pp. 156-160.
- Skelly, M. (1979). *Amer-Ind Gestural Code Based on Universal American Indian Hand Talk*, New York, USA: Elsevier North Holland, Inc.
- Siegel-Causey, E., Guess, D. (1989). *Enhancing non-symbolic communication interactions among learners with severe disabilities*, Baltimore, USA: Paul Brookes Publishing Company.
- Soto, G., Olmstead, W. (1993). "A Semiotic Perspective for AAC". *Augmentative and Alternative Communication*, vol.9, no. 2, pp.134- 141.
- Sparrow, C. J. (1993). *Coverbal Iconic Gestures in Human-Computer Interaction*. Masters Thesis, MIT Media Arts and Sciences Section.
- Stokoe, W. (1960). *Sign language structure: An outline of the visual communication system of the American deaf*, Galludet College, Washington, DC, USA.
- Thome, B. (ed.) (1993). *Systems Engineering: Principles and Practice of Computer-based Systems Engineering*, West Sussex, UK: John Wiley and Sons Ltd.
- Tou, J. T., Gonzanlez, R. C. (1974). *Pattern Recognition Principles*, Massachusetts, USA: Addison-Wesley.
- van Balkom, H. (1991a). *The Communication of Language Impaired Children: A Study of discourse coherence in conversations of Specific Language Impaired and Normal Language Acquiring Children and their Primary Caregivers*. IRV Series in Rehabilitation Research. Amsterdam/Lisse: Swets & Zeitlinger.
- B.V. van Balkom, H., Heim, M.J.M. (1991b). "The Methodological challenge of interaction research in AAC", In J, Brodin & E. Bjorck-Akesson (eds.), *Methodological Issues in Research in Augmentative and Alternative Communication*, Stockholm, Sweden: Swedish handicapped Institute.
- van Dijk, J. (1966). "The first steps of the deaf-blind child towards a language", *International Journal for the Education of the Blind*, vol. 15, no.4, pp. 112-114.
- Vanderheiden, G. (ed.) (1978). *Non-vocal communication resource book*, Baltimore, USA: University Park Press.
- Vogl, T.P., Mangis, J.K., Rigler, A.K., Zink, W.T., Alkon, D.L. (1988). "Accelerating the Convergence of the Backpropagation Method", *Biological Cybernetics*, vol. 59, pp. 257-263.
- Verplank, W., Fulton, J., Black, A., Moggridge, W. (1993). "Observation and invention the use of scenarios interaction design", CHI Tutorial, *ACM CHI'94 Human Factors in Computing Systems*, Boston, Conference Companion, pp.119-120.
- von Tetzchner, S. (1988). "Becoming an Aided Speaker". Paper presented at the ISAAC conference, Anaheim, CA.
- von Tetzchner, S., Grove, N., Loncke, F., Barnetts, S., Woll, B., Clibbens, J. (In Press). "Towards a Theoretical Model of Augmentative and Alternative Communication" in von Tetzchner, S., Jensen, M. H. (eds.) *European Perspectives on Augmentative and Alternative Communication*, London: Whurr.
- Vygotsky, L. S. (1986). *Thought and Language*, A. Kozulin, (ed.), Massachusetts, USA: MIT Press.
- Waldrop, M. M. (1992). *Complexity: The Emerging Science at the Edge of Order and Chaos*, New York: Simon and Schistose Inc.

- Werbos, P. (1974). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioural Sciences* (Ph.D. thesis). Cambridge, MA, USA: Harvard University Committee on Applied Mathematics.
- Werner, H., Kaplan, B. (1963). *Symbol Formation*. New York, Wiley & Sons.
- Wexelblat, A. D. (1994). *A Feature-Based Approach to Continuous-Gesture Analysis*. Masters Thesis. MIT Media Arts and Sciences Section.
- Wolf, C. G., Morrel-Samuels, P. (1987). "The Use of Hand-Drawn Gestures for Text Editing", *International Journal of Man-Machine Studies*, vol. 27, pp. 91-102.
- Wundt, W. (1921). *The Language of Gesture* (Reprinted 1973) The Hague: Mouton.
- Yoder, D. E. (1984). "Conversational Interaction of Augmentative Systems Users: What Does Our Research Tell Us?" Paper presented at the *Third International Conference on Augmentative and Alternative Communication*, Boston, MA.
- Zangari, C., Lloyd, L. L., Vicker, B. (1994). Augmentative and Alternative Communication: An Historic Perspective. *Augmentative and Alternative Communication*, vol.10, March, pp.27-59.
- Zurada, J. M. (1992). *Introduction to Artificial Neural Systems*, St. Pauls, MN, USA: West Publishing Company.

Glossary

ADL. Activities of Daily Living: Tasks associated with self-care and independent living.

Assistive Technology Device - "any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities." An assistive device can be low-tech (mechanical) or high-tech (electromechanical or computerized) and includes products that compensate for sensory and functional losses by providing the means to move (e.g. wheelchairs, lifts), speak (e.g. voice synthesizers, voice recognisers), read (e.g. Opticon systems for persons who are blind), hear (e.g. vibro-tactile aids) and manage self-care tasks (e.g. automatic feeders, environmental control systems). [as defined in "Technology- Related Assistance of Individuals with Disabilities Act 1988" (P.L. 100-407), USA, adapted to include voice recognition systems.]

Augmented and Alternative Communication. An area of clinical practice that attempts to compensate (either temporarily or permanently) for the impairment and disability patterns of individuals with severe expressive communication disorders (i.e., the severely speech-language and writing impaired). [ASHA, 1989, p.107]... "utilize the individual's full communication capabilities, including any residual speech or vocalizations, gestures, signs and aided communication" [ASHA, 1991, p.10].

Augmented and Alternative Communication System. (AAC). An integrated group of components, including the symbols, aids, strategies, and techniques used by individuals to enhance communication" [ASHA, 1991, p.10].

Augmented and Alternative Communication Devices.

Electronic- Technologies that enable a person with limited communicative modalities i.e speech, motor control, hearing, vision, cognitive ability to visually or auditorially access and display their communicative intentions e.g. Voice activated communication aid (VOCA).

Non-Electronic- Technologies that enable a person with limited communicative modalities i.e speech, motor control, hearing, vision, cognitive ability to visually or auditorially access and display their communicative intentions. Examples include: flat surface communication word board (e.g. Fitzgerald Key, Bliss board) or book (e.g. Rebus, Makaton) that contains the letters of the alphabet, numbers, key phrases and/or symbols that the users is able to access.

Backpropagation. A method of training a feedforward artificial neural network with at least one hidden layer.

Cerebral Palsy. A broad term used to describe a variety of conditions caused by damage to the developing brain, usually occurring before, during or shortly after birth. The dam-

age is such that it affects neuromotor development resulting in a continuum of characteristic motor disorders affecting a child's movement, speech, and posture. The condition can be mild to severe. Although it is considered to be non-progressive, i.e. the initial brain lesions or abnormalities do not get worse, the degree or type of exhibited motor dysfunction can change as a child's nervous system develops. It is considered permanently disabling although therapeutic intervention is thought to have a beneficial effect on a child's motor abilities.

Co-articulated. Typically used to describe the act or mode of joining in speech, in this context used to refer to the linkage of gestures similar to the linking of speech phrases.

Confusion Matrix. A square matrix of numbers. Each row corresponds to a gesture class. Each column corresponds to a classifier decision. Each number represents the proportion of gestures of a particular class that have been classified as belonging to the same or another class as indicated by their position in the matrix.

Coverbal gesture. Gestures produced in the presence of speech, either simultaneously, prior or immediately after speech.

Developmental disability. Term used for conditions due to congenital abnormality, trauma, deprivation, or disease that interrupt or delay the sequence and rate of normal growth, development and maturation. [(1989) *Taber's Cyclopedic Medical Dictionary*, 16th Edition, F.A. Davis Co. Philadelphia.]

Dysarthria. Difficult and defective speech due to impairment of tongue or other muscle essential to speech production. [(1989) *Taber's Cyclopedic Medical Dictionary*, 16th Edition, F.A. Davis Co. Philadelphia.]

Emergent. 1. becoming apparent: emerging. 2. the new qualitative synthesis produced by structures organised in certain patterns that cannot be predicted from examination of the constituent parts of the whole. 1. [(1991) *The Oxford Encyclopedic English Dictionary*, Eds. J.M. Hawkins and R. Allen, Clarendon Press, Oxford.]. 2. [(1992) *Philosophy*, The Harper Collins Dictionary, P.A. Angeles, Harper Perennial.]

Gesture. 1. A significant movement of a limb of the body. 2. the use of such a movement especially to convey feeling or as a rhetorical device. 3. an action to evoke response or convey intention. [(1991) *The Oxford Encyclopedic English Dictionary*, Eds. J.M. Hawkins and R. Allen, Clarendon Press, Oxford.] Gesture are considered both global i.e the whole determines the meaning of the parts and synthetic i.e one gesture can combine many meanings. This is contrast to the combinatoric linear-segmented property of speech and sign language. [(1992) *Hand and Mind*, D. McNeill.]

High bandwidth Afferent HMI. Human-machine interaction where the bandwidth of information transfer from the machine to the human is of appreciable magnitude.

High Bandwidth Efferent HMI. Human-machine interaction where the bandwidth of information transfer from the human to the machine is of appreciable magnitude.

Iconic. 1. of or having the nature of an image or portrait. 2. (of a statue) following a conventional type. 3. *Linguistics* that is an icon. Iconicity: with reference to gesture taxonomy, gesture bearing a close formal relationship to the semantic content of speech e.g.

Learned dependency. Dependency on others for interaction with the world. Term suggested by (von Tetzchner 1988, see learned helplessness.)

Learned futility. Creation of dependency on another for assistance and/or mediation for daily interaction with the world, due to restrictions of a physical and/or cognitive nature and reaction and responses of others to those restrictions, whereby an individual experi-

ences a sense of helplessness

Learned helplessness. Precursor to learned dependency, suggestive of more conscious or unconscious decision to rely on other for interaction in the world. Theory suggests that the pattern of passivity found in AAC users and its possible relationship to particular behaviours of their communication partners could be partially explained as an example of learned helplessness. Could be reinforced by giving rewards that are not dependent on performance. Consequences included e.g. decrease in motivation, effects may persist even when environmental condition may have changed. Seligman (1975), regarded experiences of control and independence in daily life as essential for reducing learned helplessness.

Lexeme. A basic lexical unit of language comprising one or several words, the elements of which do not separately convey the meaning of the whole.

Physiographic. Depiction through bodily movement.

Saliency. That part which carries the meaning e.g., in a gesture of a rainbow the form of the arc traced in space.

Self-Adaptive. A system capable of adjusting its own internal states in response to its environment.

Self-Organising. A system capable of autonomously structuring its own internal states.

Sign Language. The various natural language of deaf communities.

Symbol. (-AAC). Refers to the methods used for "visual, auditory, and/or tactile representation of conventional concepts e.g., gestures, photographs, manual sign sets, systems, picto-ideographs, printed words objects, spoken words, Braille" [ASHA, 1991, p.10].