

Evolving Spiking Neural Networks for Adaptive Audiovisual Pattern Recognition

Simei Gomes Wysoski

A thesis submitted to
Auckland University of Technology
in fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)

2008

Knowledge Engineering & Discovery Research Institute
Faculty of Design and Creative Technologies

Primary Supervisor: Nikola Kasabov
Secondary Supervisor: Lubica Benuskova

Table of Contents

Attestation of Authorship	iv
Dedication	v
Acknowledgements	vi
Abstract	viii
Chapter 1 - Introduction	1
1.1 Definition and motivation	1
1.2 Objective of the research.....	3
1.2.1 Specific objectives	3
1.2.2 Research design.....	3
1.3 Scientific contribution.....	5
1.4 Organisation of the dissertation	6
Chapter 2 - Modelling the biological brain and brain-like computation - A review 7	
2.1 Brain-like information processing units and their organization.....	8
2.1.1 Neurons	8
2.1.2 Organization of neurons into specialized functional ensembles	17
2.2 Information processing in sensory pathways	20
2.2.1 Processing visual information	21
2.2.2 Processing auditory information	26
2.2.3 Integration of audiovisual information.....	29
2.3 Classification of neural-based information processing models	32
2.3.1 Connectionist Systems (COS).....	33
2.3.2 Evolving Connectionist Systems (ECOS).....	35
2.3.3 Brain-like Evolving Connectionist Systems (<i>Brain-like</i> ECOS)	36
2.4 Spiking Neural Networks (SNNs).....	37

2.4.1 SNNs for modelling brain function.....	38
2.4.2 SNNs in artificial intelligence.....	39
2.4.3 Properties of SNNs for pattern recognition.....	40
2.4.4 Implementation of SNN models.....	42
2.4.5 Mathematical facts about SNNs.....	44
2.4.6 Limitations of SNN models	44
2.4.7 Parameter optimization in neuronal models.....	45
2.5 Chapter conclusion.....	46
Chapter 3 - Evolving SNNs for visual pattern recognition	47
3.1 Background and benchmarking	48
3.2 Brain-like networks for visual information processing.....	52
3.3 SNN model for face recognition	55
3.3.1 Model description	55
3.4 A new online learning procedure with structural adaptation	59
3.4.1 General description	59
3.4.2 Learning procedure	60
3.4.3 Experimental setup.....	62
3.4.4 Comparative results.....	64
3.5 A new multi-view face recognition system using evolving SNNs	67
3.5.1 Model description	67
3.5.2 Network dynamics during test	69
3.5.3 Experiments and results	71
3.6 Implementation considerations	82
3.7 Chapter conclusion.....	84
Chapter 4 - Evolving SNNs for auditory pattern recognition.....	87
4.1 Background and benchmarking	88
4.2 SNNs for speech processing.....	91
4.3 A new SNN-based method for text-independent speaker authentication	93
4.3.1 Speech pre-processing.....	94
4.3.2 Evolving spiking neural network models.....	101
4.3.3 Architecture 1 - Integration of binary opinions.....	101
4.3.4 Architecture 2 - Integration of similarity scores	104
4.3.5 Learning procedure	105
4.3.6 Experiments and results	107

4.3.7 Implementation considerations	111
4.4 A new wavelet-based speech feature extraction design with evolving SNNs	111
4.5 Sparse representation of low-dimensional data.....	114
4.5.1 Population encoding of features.....	116
4.6 Chapter conclusion.....	120
Chapter 5 - Evolving SNNs for audiovisual pattern recognition.....	124
5.1 Background and benchmarking	125
5.1.1 Insights from psychology.....	125
5.1.2 Insights from physiology	128
5.1.3 Insights from pattern analysis	128
5.2 A new evolving SNN model for audiovisual integration.....	133
5.2.1 Visual system model	135
5.2.2 Auditory system model	136
5.2.3 Audiovisual integration.....	137
5.2.4 Network dynamics during test	139
5.2.5 Experiments and results	142
5.2.6 Implementation considerations	148
5.3 Chapter conclusion.....	149
Chapter 6 - Conclusion and future directions	153
6.1 Introduction.....	153
6.2 Summary of achievements	155
6.2.1 Biologically inspired processing units	155
6.2.2 Biologically inspired pathways.....	157
6.2.3 Biologically inspired adaptive rules.....	157
6.2.4 Experimental results.....	158
6.2.5 Original contributions	159
6.3 Future directions	160
6.3.1 Scalability.....	160
6.3.2 Information coding methods	161
6.3.3 Olfactory and gustatory sensory modalities.....	161
6.3.4 Hardware implementation of evolving SNNs.....	164
6.3.5 Evolving SNNs for commercial applications.....	164
6.4 Overall conclusion	165
References	166

“I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.”

Simeí Gomes Wysoski

This dissertation is dedicated to my parents

Eleozina Gomes Wysoski

and

Paulo Silas Wysoski (*in memoriam*)

Of all the things I've learnt, what you have taught me is still by far the most important

Acknowledgements

Firstly, I would like to thank my primary supervisor Prof. Nikola Kasabov for providing ideas, inspiration and the necessary support to make this research possible. Special thanks go to my secondary supervisor, Dr. Lubica Benuskova, who devoted a lot of time teaching me the basic principles of neuroscience and for always having her office door open for “a quick discussion”. Undoubtedly, it was from these discussions that most of the ideas for this work came and most of the implementations were designed. Through Dr. Benuskova’s guidance, I was introduced to the modelling of brain functions, for which I am also truly indebted.

I would also like to thank the past and present members of KEDRI - Knowledge Engineering and Discovery Research Institute at the Auckland University of Technology, who directly or indirectly provided me with the conditions for the completion of this work. Dr. Ilkka Havukkalla deserves special acknowledgment for always bringing me challenging research questions. Peter Hwang for always being around for discussions, for providing diligent technical support, and teaching me how to use NeuCom. “Obrigado” to Dr. Paulo C. M. Gottgroy for introducing me to the ontological way of describing the world and for teaching me to look beyond the spike domain. Dr. Michael D. Platel for many discussions over a cup of coffee. Vishal Jain, David Zhang, Akbar Ghobakhlou, Vishwa Shukla, Dougal Greer, Richard Walton, Helene Wall, Yoshitaka Inoue also contributed in different ways to this work. From KEDRI, I am indebted to Joyce D’Mello for always encouraging students at the times we were thinking of giving up and for being ready to help in whatever situation we faced. With her assistance, getting a PhD became a much easier task.

Thanks to Snjezana Soltic for the opportunity to share my thoughts about this research and for returning many intriguing questions that resulted in a deeper understanding of the topic. I am grateful to Dr. Marcia Gottgroy, for kindly reading the draft of this dissertation and particularly, helping me to clarify the research design.

Thanks also to the Tertiary Education Commission of New Zealand, through the Top Achiever Scholarship, for the financial support for this research.

Finally, I would especially like to thank my mother Eleozina Gomes Wysoski and my wife Marisa van Arragon. Without their love and encouragement this work would probably have not even started.

Abstract

This dissertation presents new modular and integrative information methods and systems inspired by the way the brain performs information processing, in particular, pattern recognition. The proposed artificial systems use spiking neurons as basic elements, which are the key components of spiking neural networks. Of particular interest to this research are various spiking neural network architectures and learning procedures that permit different pattern recognition problems to be solved in an evolvable and adaptive way.

Spiking neural networks are used to model human visual and auditory pathways and are trained to perform the specific task of person authentication. The systems are individually tuned and trained to recognize facial information and to analyze sound signals from spoken sentences. The modelling of the integration of different sources of information (multisensory integration) using spiking neural networks is also a subject of investigation. A network architecture is proposed and a model for audiovisual pattern recognition is designed as an example.

The main original contributions of this thesis are:

a) Evaluation and further extension of adaptive learning procedures to perform visual pattern recognition. A new learning procedure that enables the system to change its structure, creating/merging neuronal maps of spiking neurons is presented and evaluated on a face recognition problem.

b) Design of two new spiking neural network architectures to perform person authentication through the processing of speech signals.

c) Design and evaluation of a new architecture that integrates sensory modalities based on spiking neurons. The integrative architecture combines opinions from individual modalities within a supramodal layer, which contains neurons sensitive to multiple sensory information. An additional feature that increases biological relevance is the crossmodal coupling of modalities, which effectively enables a given sensory modality to exert direct influence upon the processing areas typically related to other modalities.

The contributions were published in one journal paper and in four refereed international conference proceedings. The proposed system designs were implemented and, through computer simulations, demonstrated comparable performance with traditional benchmarking methods. The systems have some promising features: they can be naturally optimized in respect to different criteria: accuracy (when very accurate results are expected), energy efficiency (when management of resources play an important role), and speed (when a decision needs to be made within a limited time).

In this thesis, most of the parameters have been exhaustively optimized by hand or by using simple heuristics. As a direction for future work, there is an opportunity to include automated, specially tailored parameters optimization procedures or even general-purpose optimization algorithms, e.g., Genetic Algorithms and Particle Swarm Optimization.

Overall, the results obtained in this thesis clearly indicate that it is indeed possible to have fast and accurate adaptive pattern recognition systems scalable for multiple modalities computing with simple models of spiking neurons. However, it is important to advance the theory of spiking neurons to take advantage of its biological relevance to reach similar or better performance when compared to the human brain, for instance, exploring new neuron models, information coding schemes and network connectivity.