

Virginia Commonwealth University VCU Scholars Compass

Theses and Dissertations

Graduate School

2010

TOPOLOGICAL PROPERTIES OF A NETWORK OF SPIKING NEURONS IN FACE IMAGE RECOGNITION

Joo-Heon Shin Virginia Commonwealth University

Follow this and additional works at: http://scholarscompass.vcu.edu/etd Part of the <u>Computer Sciences Commons</u>

© The Author

Downloaded from http://scholarscompass.vcu.edu/etd/2133

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

School of Engineering Virginia Commonwealth University

This is to certify that the dissertation prepared by Joo-Heon Shin entitled TOPOLOGICAL PROPERTIES OF A NETWORK OF SPIKING NEURONS IN FACE IMAGE RECOGNITION has been approved by his or her committee as satisfactory completion of the dissertation requirement for the degree of Doctor of Philosophy

Dr. Krzysztof J. Cios, Committee Chair, Department of Computer Science

Dr. Kayvan Najarian, Committee Member, Department of Computer Science

Dr. Vojislav Kecman, Committee Member, Department of Computer Science

Dr. Alex Meredith, Committee Member, Department of Anatomy and Neurobiology

Dr. Alessio Plebe, Committee Member, Department of Cognitive Science, University of Messina, Italy

Dr. Roman Swiniarski, Committee Member, Department of Computer Science, San Diego State University

Dr. Sebastián Ventura, Committee Member, Department of Computer Science and Numerical Analysis, University of Cordoba, Spain

Dr. Russell Jamison, Dean of the School of Engineering

Dr. F. Douglas Boudinot, Dean of the Graduate School

March 24, 2010

© Joo-Heon Shin 2010

All Rights Reserved

TOPOLOGICAL PROPERTIES OF A NETWORK OF SPIKING NEURONS IN FACE IMAGE RECOGNITION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

JOO-HEON SHIN

B.S., Hankuk University of Foreign Studies (Seoul, South Korea), 1998M.S., Hankuk University of Foreign Studies (Seoul, South Korea), 2000M.S., San Diego State University (San Diego, CA, U.S.A.), 2006

Director: KRZYSZTOF J. CIOS PROFESSOR AND CHAIR, DEPARTMENT OF COMPUTER SCIENCE

Virginia Commonwealth University Richmond, Virginia March, 2010

Acknowledgement

I would like to express my sincere gratitude to my advisor Professor Krzysztof J. Cios for his continuous help and comments during the entire period of graduate school. It would have been impossible to complete this paper without his guidance. I am especially grateful to all of committee members for their helpful comments regarding this paper. I have always received support from my friends: Dr. Cao Nguyen, Dr. Leslie Keniston, Mr. Hun Ki Lim, Mr. Dat Nguyen, and Mrs. Jane Starrett. I am much grateful to all of them. But most of all, I would like to dedicate this thesis to my father and mother; Dong-Hee Shin and Song-Woel Choi. Also, I would like to give thanks to my sisters (Kyung-Joo Shin, Young-Joo Shin, Hye-Seon Shin, and Hyun-Joo Shin) and their families. I could not have invested the valuable time needed to complete this thesis without their continued support. Especially, for my lovely family, Jae-Man (Elliot) Shin and Sung-Hei Choe, I would not have existed without them. I dedicate this thesis to all of them.

Table of Contents

Acknowledgementsii		
List of Tables		
List of Figures		
Abstractix		
Chapter		
1	Introduction 1	
	1.1 Research Objectives 1	
	1.2 Research Methods	
	1.3 Motivation and Significance	
	1.4 Literature Review	
2	Design Issues in a Network of Spiking Neurons	
	2.1 Artificial Spiking Neuron Models	
	2.2 Synaptic Plasticity (Learning) Rules 14	
	2.3 Network Topology	
3	Recognition of Partially Occluded and Rotated Images	
	3.1 Methods	
	3.2 Datasets	
	3.3 Network Simulation: Organization and Recall	
4	Experimental Results and Discussion	
	4.1 Evaluation	

	4.2 Results and Discussion	37
5	Stopping Criterion	44
	5.1 Issue in Iteration Process	44
	5.2 Entropy-based Stopping Criterion	45
	5.3 Results and Discussion	50
6	Conclusions	59
	6.1 Summary	60
	6.2 Future Work	62
Bibliogra	ıphy	64
Appendie	ces	73
А	Occlusion and Rotation Datasets	74
В	Other Performance Measures	79

List of Tables

Page

Table 2.1: MacGregor's modified neuron model	11
Table 3.1: MacGregor's modified neuron model and its parameter settings	21
Table 3.2: Number of neurons in the recognition layer and the corresponding synaptic connections from the feature extraction layer to the recognition layer for 32×32 image	26
Table 3.3: Pseudo code of an algorithm to compute "signature" patterns	35
Table 5.1: Pseudo code of an algorithm for the stopping criterion	49

List of Figures

Page

Figure 2.1: Membrane electrical circuit of the Hodgkin-Huxley model	10
Figure 2.2: Neuron responses to an external artificial stimulation (<i>SCN</i>), where <i>E</i> is membrane potential, T_h is threshold, and G_k is potassium channel conductance.	12
Figure 2.3: Examples of Izhikevich neuron model and its modeling of several properties	14
Figure 2.4: Learning functions used in the SAPR (a) and in the STDP (b)	17
Figure 3.1: Learning functions used in the system using the SAPR (a) and the STDP (b)	22
 Figure 3.2: Topology of the network: (a) high-level block diagram; (b) recurrent synaptic connections between the excitatory neurons in the feature extraction layer; and (c) synaptic connections between the excitatory neurons in the sensory/feature extraction layer and the inhibitory neurons in the feature extraction layer. 	23
Figure 3.3: Illustration of formation of the "signature" vectors in the recognition layer	25
Figure 3.4: Illustration of how the network recognizes an image using the threshold comparator.	29
Figure 3.5: Examples of the Original and Occlusion & Rotation images	30
Figure 3.6: Generation of an occlusion	31
 Figure 4.1: Performance of the system in terms of accuracy with different number of neurons in the recognition layer using 10-FCV: (a) using SAPR, (b) using STDP, and (c) comparison of the SAPR and STDP rules on the Original (black bars) and the Occlusion & Rotation (gray bars) images	40

Figure 4.2: Graphical comparison of the results generated on the feature extraction layer using the SAPR and STDP rules on one of the images, with inhibition (columns (a) and (b))	
and without inhibition (column (c)).	41
Figure 4.3: Comparison of the network performance in terms of accuracy using the SAPR and the SVM	42
Figure 5.1: The network topology and its corresponding neuronal spike responses: (a) the network topology and its output of the feature extraction layer for an input image in 32×32 resolution, and (b) one example of neuronal spike responses with SAPR rule in the sensory (<i>s/e</i>), feature extraction (<i>f/e</i> and <i>f/i</i>) and recognition (<i>r/e</i>) layers	46
 Figure 5.2: Relationship between an entropy variation and a performance of the network for JAFFE data in 32×32 resolution using SAPR; (a) neuronal spike responses in the sensory (<i>s/e</i>), feature extraction (<i>f/e</i> and <i>f/i</i>) and recognition layers (<i>r/e</i>), (b) the changes of organization of excitatory neuronal spike patterns through simulation time in the sensory (<i>s/e</i>) and feature extraction layer (<i>f/e</i>), (c) an entropy variation at hierarchy through simulation time, and (d) a performance of the network through simulation time 	48
Figure 5.3: Performance of the system with accuracy; (a) with SAPR, and (b) with STDP	52
 Figure 5.4: Changes of the entropy in the sensory (<i>s/e</i>), feature extraction (<i>f/e</i> and <i>f/i</i>) and recognition layers (<i>r/e</i>); (a) with SAPR, and (b) with STDP 	54
Figure 5.5: Changes of the total CPU time in minutes; (a) with SAPR, and (b) with STDP	56
Figure 5.6: Changes of the total computational time in an absolute unit; (a) with SAPR, and (b) with STDP	57
Figure A.1: The Occlusion & Rotation images of JAFFE dataset	75
Figure A.2: The Occlusion & Rotation images of ORL dataset	76
Figure A.3: The Occlusion & Rotation images of CMU dataset	77

Figure A.4: The Occlusion & Rotation images of UMIST dataset	78
 Figure B.1: Performance of the system in terms of precision with different number of neurons in the recognition layer using 10-FCV: (a) using SAPR, (b) using STDP, and (c) comparison of the SAPR and STDP rules on the Original (black bars) and the Occlusion & Rotation (gray bars) images 	79
 Figure B.2: Performance of the system in terms of recall with different number of neurons in the recognition layer using 10-FCV: (a) using SAPR, (b) using STDP, and (c) comparison of the SAPR and STDP rules on the Original (black bars) and the Occlusion & Rotation (gray bars) images 	80
 Figure B.3: Performance of the system in terms of Harmonic mean with different number of neurons in the recognition layer using 10-FCV: (a) using SAPR, (b) using STDP, and (c) comparison of the SAPR and STDP rules on the Original (black bars) and the Occlusion & Rotation (gray bars) images 	81
Figure B.4: Comparison of the network performance in terms of precision using the SAPR and the SVM	82
Figure B.5: Comparison of the network performance in terms of recall using the SAPR and the SVM	82
Figure B.6: Comparison of the network performance in terms of Harmonic mean using the SAPR and the SVM	83

Abstract

TOPOLOGICAL PROPERTIES OF A NETWORK OF SPIKING NEURONS IN FACE IMAGE RECOGNITION

By Joo-Heon Shin, Ph.D.

A dissertation submitted in partial fulfillment of the requirements for the degree of Degree of Doctor of Philosophy in Engineering (Computer Science track) at Virginia Commonwealth University.

Virginia Commonwealth University, 2010

Major Director: Krzysztof J. Cios Professor and Chair, Department of Computer Science

We introduce a novel system for recognition of partially occluded and rotated images. The system is based on a hierarchical network of integrate-and-fire spiking neurons with random synaptic connections and a novel organization process. The network generates integrated output sequences that are used for image classification. The network performed satisfactorily given appropriate topology, i.e. the number of neurons and synaptic connections, which corresponded to the size of input images. Comparison of Synaptic Plasticity Activity Rule (SAPR) and Spike Timing Dependant Plasticity (STDP) rules, used to update connections between the neurons, indicated that the SAPR gave better results and thus was used throughout. Test results showed that the network performed better than Support Vector Machines. We also introduced a stopping criterion based on entropy, which significantly shortened the iterative process while only slightly affecting classification performance.

CHAPTER 1 Introduction

In this chapter the main objective of this research along with its motivation and significance are explained.

1.1 Research Objectives

The main objective of this work is to develop a Network of Spiking Neurons (NSN) for image classification and study its dynamics in the task of face recognition. The work is an extension of previous research in dynamic synaptic plasticity and image segmentation with NSN. The outcome is a novel recognition system that consists of integrate-and-fire spiking neurons and its application to face recognition in gray scale images. Specifically, the system solves computationally challenging task of recognizing partially occluded and rotated images. Computational experiments enable investigating the dynamics of the system performance in the task of image classification by specifying minimum number of excitatory neurons and minimum number of synaptic connections needed for acceptable (user-specified) accuracy of classification. Moreover, an effective stopping criterion is also presented.

1.2 Research Methods

We present a novel image recognition system designed as a network of integrateand-fire spiking neurons that uses the Synaptic Activity Plasticity Rule (SAPR) (Swiercz et al., 2006) and the Spike Timing Dependent Plasticity (STDP) (Song et al., 2000) learning rules. The system consists of three layers of neurons including the input layer, the feature extraction layer, and the recognition layer (Swiercz et al., 2006, 2007). The latter layer consists of excitatory neurons that have random synaptic connections to neurons from the feature extraction layer. We address the usage of spiking neurons in the recognition layer for the purpose of decision making, which is a challenging task (Wang, 2002). In designing our system we were inspired by organization (formation) of neural circuits and retrieval of neural memories in a primate's hippocampus (Yanike et al., 2004).

Many of the neuron connections in the brain appear during initial brain growth and later are continuously organized/reorganized. These changes are due to repeated experiences such as seeing similar images many times (Rosenblatt, 1958; Tsodyks and Gilbert, 2004). Visual input is transported to visual and entorhinal cortex for processing. The results are forwarded to hippocampus where new memories are formed. With time the new memories are transferred to cerebral cortex for long term storage (Kandel et al., 1996). Once the long term memory connections are established, the brain is able to immediately "recall" what the new input image is by comparing it with the stored information, which is exploited in our system. As a result, relatively little processing is required to make a recognition decision.

2

We evaluate our system using gray scale images of human faces. To make the recognition task more difficult (although this would not make it much more difficult for a human) we use occluded images that are presented at a random angle/rotation. Upon presentation of the images, the system forms unique signature patterns that correspond to a person's face in a way vaguely similar to the above described "organization" phenomenon in the brain. When a new face exemplar is presented to the system, the system "recalls" the corresponding person's face "signature" and thus is able to correctly recognize it relatively quickly.

One of the challenges is to decide whether a new image is truly similar (or dissimilar) to the image(s) seen during the organization phase. In our previous work we explored different metrics to perform these comparisons (Cios and Shin, 1995). In this work, however, we design a new similarity measure that operates on the firing patterns of the recognition layer neurons rather than on pixels from the original image. The new image similarity measure, named the "threshold comparator," combines information concerning spike timing and the corresponding transmembrane potential.

1.3 Motivation and Significance

Digital image databases grow exponentially as a result of advances in imaging hardware/software and its proliferation in commercial, medical, and military systems. Unfortunately, the value of such data is not yet fully realized due to relatively slow advancements in automated image recognition, recall, and understanding. Organizations such as the Defense Advanced Research Projects Agency, after spending millions of dollars on research in this area, have come to the conclusion that image recognition performed by a human is far better (in terms of speed and accuracy) than that of the best automatic recognition systems (Chellappa et al., 1995; Samal and Lyengar, 1992). For instance, humans are much quicker in recognizing previously seen objects even if they are partially occluded and rotated. Solving challenging image distortions, such as occlusions and rotations, via automated means calls for development of systems that combine several computational approaches (Kanan et al., 2008; Tan et al., 2005). More specifically, one of the basic problems in image analysis is the underlying assumption of a certain mathematical decomposition hypothesis, say, assuming that different causes of distortions are independent and can be isolated. For instance, assuming that occlusion and rotation are independent, we can first solve the occlusion problem and later address the problem of the rotation, or the other way round. In practice, however, such independency cannot be proven, and therefore such sequential solutions could potentially fail.

Manual analysis of large and ever growing image databases is impractical which motivates research in automated image recognition. Unfortunately, automating this process is difficult and poses many challenges. Recognizing the fact that computer systems will not be on par with humans for some time to come, we propose an approach that may lead to design of systems that improve over the existing solutions. We observe that a relatively low speed of information transmission between biological neurons, when compared with transmission speeds of computers, is compensated by a "smart" organization of the brain's neural circuits that enables the remarkable ability of humans to perform complex image recognition tasks. This fact motivated us to design a system that attempts to model some brain operations, while being fully aware that this attempt is a simplistic approximation of how the brain's neural circuits truly operate.

1.4 Literature Review

Networks of spiking neurons (NSN) have shown to be great modeling tools in fields like neuroinformatics and computational biology (Buzsaki, 2004; Deco and Rolls, 2004; Delorme and Thorpe, 2001; Gerstner and Kistler, 2003; Gold and Shadlen, 2001; Hagmann et al., 2008; Izhikevich, 2006, 2007; Izhikevich et al., 2003; Izhikevich and Edelman, 2008; Markram, 2006; Masquelier and Thorpe, 2007; Mel, 1997; Morgan and Soltesz, 2008; Muller-Linow et al., 2008; Mutch and Lowe, 2006; Olshausen et al., 1993; Serre et al., 2007). Advancements in neuron modeling (Hodgkin and Huxley, 1952; Lovelace and Cios, 2008; MacGregor, 1993), network topology (Cios et al., 2004; Izhikevich, 2006; Sala and Cios, 1998, 1999; Swiercz et al., 2006, 2007), learning/plasticity rules (Konorski and Garry, 1948; Song et al., 2000; Swiercz et al., 2006), and simulators (Delorme and Thorpe, 2003) made it possible to accurately model some biological phenomena (Buzsaki, 2004; Deco and Rolls, 2004; Gold and Shadlen, 2001; Hagmann et al., 2008; Izhikevich, 2007; Izhikevich and Edelman, 2008; Markram, 2006; Morgan and Soltesz, 2008; Muller-Linow et al., 2008; Swiercz et al., 2007) as well as solve practical problems (Bohte et al., 2002; Cios et al., 2004; Delorme and Thorpe, 2001; Sala and Cios, 1998, 1999; Swiercz et al., 2006; Wysoski et al., 2008). These

improvements even allow for building models that exhibit brain-like behavior (Izhikevich and G.M. Edelman, 2008; Markram, 2006; Morgan and Soltesz, 2008).

In the previous work a dynamic learning rule, the Synaptic Plasticity Activity Rule (SAPR) was developed based on studies of the CA3 region of hippocampus and later used in a network of spiking neurons for image edge detection (Swiercz et al., 2006). The rule exhibited the same kind of image segmentation effects that were also reported in (Buhmann et al., 2005; Choe and R. Miikkulainen, 1998; Delorme and S.J. Thorpe, 2001; Wu et al., 2007). Using a similar network topology with the SAPR rule the epileptic bursting of neurons in the CA3 hippocampal region were successfully modeled (Swiercz et al., 2007).

Understanding cognitive object recognition tasks in biological systems resulted in many interpretations by computational neuroscientists. Many theories and experiments of object recognition have been performed in the areas of computational neuroscience (Rolls and Deco, 2002; Deco and Rolls, 2002; Lee and Seung, 1999; Riesenhuber and Poggio, 1999, 2000; Stringer and Rolls, 2002; Plebe, 2007, 2008; Plebe and Domenella, 2006, 2007; Miikkulainen et al., 2005). Two well-known visual recognition models are Rolls and Deco's Visnet, and Riesenhuber and Poggio's modified Neocognitron, which have been tested on face recognition resulting to desirable invariance with respect to rotation (Rolls and Deco, 2002; Deco and Rolls, 2004; Riesenhuber and Poggio, 1999, 2000). While Riesenhuber and Poggio's modified Neocognitron provides a feedforward model by using max-like operation, Visnet is a neurophysiological model for invariant visual object recognition using the top-down approach. Lee and Seung proposed an interesting mathematical approach for learning the parts of objects by nonnegative matrix factorization (Lee and Seung, 1999). They showed its formulation as a method of decomposing images and a neural network implementation for parts-based representation of objects.

CHAPTER 2 Design Issues in a Network of Spiking Neurons

The design of any type of neural networks requires selection of a neuron model to be used in computations, a learning rule for updating the weights/synapses between neurons, and a topology that determines how the neurons are arranged and interconnected.

The first two elements must be specified *a priori*, while the topology can be either static (likewise defined *a priori*) or can be dynamically modified later by adding neurons and/or layers as needed to solve a problem. If the latter approach is used, the networks are referred to as ontogenic (Cios et al., 2007; Cios and Sztandera, 1997; Fiesler and Cios, 1997).

2.1 Artificial Spiking Neuron Models

All neuron models resemble the biological neurons to some degree, and this resemblance is an important distinguishing factor between different models. The most biophysically accurate neuron models mimic almost all key characteristics of biological neurons, resulting in reproduction of their rich firing patterns (their temporal spiking nature). The features include: membrane potential; sodium, potassium, and calcium channels; threshold accommodation; refractory periods; and multi-compartmental structure

(dendrites, soma, axon) etc. While having great applicability, using a highly detailed neuron model is characterized by very high computational cost. However, for accurate modeling it is crucial to preserve the representation of the spiking nature of biological neurons. Biological neurons, both excitatory (those that help other neurons to fire) and inhibitory (those that prevent other neurons from firing), generate a series of action potentials (also called a train of spikes, or just spikes) in response to a stimulus input. The response duration and discharge frequency depend both on the character and strength of the stimulus. There are several types of spiking neuron models; the conductance models (e.g., Hodgkin and Huxley, 1952; FitzHugh and Nagumo, 1961, 1962; Morris and Lecar, 1981; Hindmarsh and Rose, 1982, 1984), leaky integrate-and-fire (e.g., MacGregor, 1993), the Izhikevich model (Izhikevich, 2003, 2004), spike response model (Gerstner and Kistler, 2002), etc.

2.1.1 Conductance-based Neuron Model

Conductance-based neuron models are the biophysical representation of a neuron with the conductances and currents depending on the membrane potential. One of the most well known important models in computational neuroscience is the Hodgkin-Huxley (HH) model of the giant squid axon (Hodgkin and Huxley, 1952; Izhikevich, 2007). The basic components of the HH model are shown in Figure 2.1. Note that, in Figure 2.1, $R_{Na} = 1/g_{Na}$ is membrane sodium conductance, $R_K = 1/g_K$ is membrane potassium



Figure 2.1: Membrane electrical circuit of the Hodgkin-Huxley model (Hodgkin and Huxley, 1952).

conductance, $R_l = 1/g_l$ is membrane leak conductance, E_{Na} is sodium resting potential, E_K is potassium resting potential, E_l is membrane resting potential, I_{Na} is sodium current, I_K is potassium current, I_l is a small leakage current made up by chloride and other ions, and C_M is the membrane capacity (Hodgkin and Huxley, 1952).

2.1.2 Integrate and Fire Neuron Model

Integrate-and-fire neuron models provide some simplification of spike generation, as compared with conductance models, while accounting for the membrane potential and other neuron properties. MacGregor's model closely represents the behavior of a biological neuron in terms of its membrane potential, potassium channel response, refractory properties, and adaptation to stimuli. However, instead of modeling each individual channel (except for potassium), this model imitates the resulting neuron's excitatory and inhibitory properties. For accurate time modeling, it is crucial to preserve representation of the spiking nature of the biological neurons. The modified McGregor integrate-and-fire model is described by the equations given below in Table 2.1, while Figure 2.2 illustrates its operation. Note that, in Figure 2.2, *E* is membrane potential, T_h is threshold, and G_K is potassium channel conductance.

Spike generation	$S = \begin{cases} 1 & \text{if } E \ge T_h \\ 0 & \text{if } E < T_h \end{cases} $ (2.1)
Refractory properties	$\frac{dG_{K}}{dt} = \frac{-G_{K} + B \cdot S}{T_{GK}} $ (2.2)
Threshold accommodation	$\frac{dT_{h}}{dt} = \frac{-(T_{h} - T_{h0}) + c \cdot E}{T_{th}}$ (2.3)
Transmembrane potential	$\frac{dE}{dt} = \frac{-E + G_K \cdot (E_K - E) + G_e \cdot (E_e - E) + G_i \cdot (E_i - E) + SCN + N}{T_{mem}} $ (2.4)

Table 2.1: MacGregor's modified neuron model

The neuron model's membrane potential changes according to the timing of incoming spikes. Spikes influence the neuron through synaptic connections, thereby, increasing the synaptic conductance. This results in postsynaptic potential changes. There are two types of synaptic connections: excitatory and inhibitory with type depending upon the presynaptic neuron connected. The weighted sum of all excitatory and inhibitory synaptic conductances yields the excitatory or inhibitory stimulus values, respectively. If the excitatory stimulus is too weak or the inhibitory stimulus is too strong, the membrane potential cannot reach the threshold and the neuron does not fire. If the stimulus is strong enough for the membrane potential to reach the threshold, the neuron fires (generates a spike train traveling along its axon). For a short time immediately after the spike generation, the neuron is incapable of responding to any additional stimulation. This time interval is referred to as the absolute refractory period. Following the absolute refractory period is an interval known as the relative refractory period, during which the neuron can only respond to very strong stimulation.



Figure 2.2: Neuron responses to an external artificial stimulation (*SCN*), where *E* is membrane potential, T_h is threshold, and G_k is potassium channel conductance.

2.1.3 Izhikevich Neuron Model

The Izhikevich neuron model (Izhikevich, 2003) is almost as realistic at generating firing patterns as Hodgkin-Huxley but in addition it is computationally very simple. The dynamics of the model are based on a two-dimensional system of ordinary differential equations:

$$v' = 0.04v^{2} + 5v + 140 - u + I$$

$$u' = a(bv - u),$$
(2.5)

with the auxiliary after-spike resetting equation being:

if
$$v \ge +30 \text{ mV}$$
, then $\begin{cases} v \leftarrow c \\ u \leftarrow u + d, \end{cases}$ (2.6)

where v is the membrane potential of the neuron and u is a membrane recovery variable based on the activation levels of sodium Na^+ and potassium K^+ ionic currents. Depending upon dimensionless parameters a, b, c and d, the dynamic of the neuron model can mimic several neuro-computational properties, as shown in Figure 2.3.



Figure 2.3: Examples of Izhikevich neuron model and its modeling of several properties (The figure and reproduction permission are available at http://vesicle.nsi.edu/users/izhikevich).

2.2 Synaptic Plasticity (Learning) Rules

The concept of synaptic plasticity rules was stated first by Konorski (Konorski,

1948) and then by Hebb (Hebb, 1949). In short, the relative activity between pre- and post-

synaptic neurons is critical for the synaptic changes; adjustment of the strength of synaptic connections between neurons takes place every time the postsynaptic neuron fires. If firing occurs, the synaptic weight values are updated according to the postsynaptic potential value between the connected neurons with the learning rate. Arrival of the action potential at the synaptic connection changes the synaptic conduction that elicits synaptic current alteration and thus results in the postsynaptic potential (PSP) modification. The two types of learning rules are Spike-Timing Dependent Plasticity (STDP) (Song et al., 2000), and the Synaptic Activity Plasticity Rule (SAPR) (Swiercz et al., 2006).

2.2.1 Spike-Timing Dependent Plasticity (STDP)

The relative activity between pre- and post- synaptic neurons is critical. Adjustment of the strength of synaptic connections between neurons takes place every time the postsynaptic neuron fires (Song et al., 2000). If firing occurs, the synaptic weight values are updated according to eq. (2.7). The learning rate α controls the amount of adjustment; it can assume any value, with 0 meaning that there is no learning. To keep the synaptic connection strength bounded, a sigmoid function is used to produce a smoothly shaped learning curve:

$$w_{ij}(t+1) = sig(w_{ij}(t) + \alpha_{+-} \cdot PSP_{ij}(t)), \qquad (2.7)$$

where *t* is time; w_{ij} is the synaptic weight between neurons *i* and *j*; a_{+-} is the learning rate, which has different values for positive and negative adjustments; PSP_{ij} is the postsynaptic potential value for the connection between neurons *i* and *j*; and sig(x) is a sigmoid function. Arrival of the action potential at the synaptic connection changes the synaptic conduction, which elicits synaptic current alteration and thus results in the postsynaptic potential (PSP) modification. The PSP can be either excitatory (EPSP) or inhibitory (IPSP) and is proportional to the synaptic conductance change. These changes directly affect the neuron's membrane potential.

Various types of learning rules have been extensively studied (Bi and Poo, 1998; Fusi, 2002; Fusi et al., 2000; Lengyel et al., 2005; Sala and Cios, 1998; Song et al., 2000; Swiercz et al., 2006). All but one of the learning rules for spiking neuron networks use artificial functions to evaluate the amount of the synaptic strength adjustment. The prime example is the Synaptic Time-Delayed Plasticity (STDP) rule that is specified by eq. (2.8) and illustrated in Figure 2.4(b):

$$STDP(\Delta t) = \begin{cases} \alpha_{+} \exp(-\Delta t/\tau_{+}) & \text{if } \Delta t > 0\\ -\alpha_{-} \exp(\Delta t/\tau_{-}) & \text{if } \Delta t \le 0, \end{cases}$$
(2.8)

where $\Delta t = (t_{post} - t_{pre})$ is the time delay between the postsynaptic spike and the presynaptic spike; α_{+-} is the learning rate; and τ_{+-} is the time constant (Song et al., 2000). The STDP rule embodies Konorski/Hebb-type plasticity using the concept of relative timing, as can be seen in eq. (2.7).

2.2.2 Synaptic Activity Plasticity Rule (SAPR)

In contrast to the above rules, the Synaptic Activity Plasticity Rule (SAPR) (Swiercz et al., 2006) uses the actual synaptic dynamics to decide amount of adjustment; it



Figure 2.4: Learning functions used in the SAPR (a) and in the STDP (b).

also follows Konorski/Hebb-type plasticity. When modification of the synaptic weight between the pre- and post-synaptic neurons occurs, the SAPR adjusts the synaptic weight depending on the particular synapse type and its recent actual activity. There is no explicit equation or function shape for the synaptic strength adjustment in the SAPR. The adjustment only approximates possible function using a PSP shape (Swiercz et al., 2006). Figure 2.4(a) shows just one example of a learning function using a general PSP shape for excitatory and inhibitory synapses. The actual shape varies depending on the particular synapse parameters, current synaptic strength, and learning rate used. In contrast to the STDP function, the SAPR function is continuous, has a finite range of values, and is dynamic (i.e., it changes from experiment to experiment while the STDP function is static) (Swiercz et al., 2006). Similar to the STDP, the polarity of weight change in the SAPR generally matches the phase of spike times, although it has been shown that this is not necessarily correct since polarity might exhibit frequency dependence (Lisman and Spruston, 2005). In general, frequency dependence of spike timing dependent plasticity is a complex phenomenon that depends on the precise frequency of stimulation, presynaptic depression and postsynaptic desensitization, the amplitude and duration of the neuronal calcium transient determined by calcium-dependent calcium release, calcium buffering, and calcium transport capacities, and the relative sensitivities of CAMKII and calcineurin. At this point, there is not enough information concerning these processes to permit adequate modeling of the frequency dependence of the SAPR.

The advantage of the SAPR is that instead of using an artificial function, as used in the STDP, it uses the actual value present in each synapse. Modification of the synaptic weights between pre- and postsynaptic neurons takes place every time the postsynaptic neuron fires. When the firing occurs, all of the neuron's incoming synapses are evaluated and their synaptic strengths are adjusted depending on the particular synapse type and recent activity. The amount of the adjustment is proportional to the contribution of a particular synapse to the neuron's firing. If a particular excitatory presynaptic neuron spike arrives before the postsynaptic neuron fires, then the related synapse is assumed to have a positive contribution, and thus its synaptic strength increases by an amount proportional to the current postsynaptic potential (PSP) value. When an excitatory presynaptic neuron's spike arrives after the postsynaptic neuron fires, it has no contribution to the recent firing and thus its strength is decreased by an amount proportional to the current PSP value.

18

2.3 Network Topology

To design a computational model that is biologically accurate it is important not only to use realistic models of neurons and dynamic plasticity rules but also have structure similar to the way the nervous system does. Thus, realistic computer models of the nervous system process information in stages before it is projected onto a higher neuronal structure. This hierarchical processing of information is achieved via designing appropriate network topology. Many researchers used hierarchical structure, composed of both excitatory $(75 \sim 80\%)$ and inhibitory $(20 \sim 25\%)$ neurons, to keep the same ratio as present in the brain, to mimic biologically meaningful information processing (Cios et al., 2004; Dehaene et al., 2003; Grill-Spector et al., 2006; Makeig et al., 2002; Risenhuber and Poggio, 1999, 2000; Rybak et al., 1997; Sala and Cios, 1998; Silver et al., 2007; Swiercz, 2006, 2007). However, it is difficult to determine how synaptic connections among neurons in a laminar structure are made, so the randomness of connections plays an important role in a given topology. These, as well as additional features can be modeled using a realistic multilayer network topology, in which each layer deploys both excitatory and inhibitory neurons as well as uses concepts of randomness, ultimately to produce a realistic model of information processing.

CHAPTER 3 Recognition of Partially Occluded and Rotated Images

In this chapter, we introduce a novel recognition system that is capable of recognizing partially occluded and rotated images. The system is based on a hierarchical network of integrate-and-fire spiking neurons, with random synaptic connections, and a novel organization process similar to the one observed in biological systems. The network generates integrated output sequences that are used for image classification. The network uses a fixed number of neurons in the recognition layer, corresponding to a given image size, to obtain good classification performance of the system.

3.1 Methods

As mentioned in Chapter 2, the design of any type of a neural network requires the neuron model used for computations, the learning rule used to update the weights/synapses, and the topology of a network that determines how the neurons are arranged and interconnected. We use the MacGregor's modified integrate-and-fire (I&F) neuron model for designing the network topology, and the STDP and SAPR learning rules.

3.1.1 MacGregor's Modified Neuron Model and Its Parameter Settings

There are several types of artificial spiking neuron models, as mentioned in Section 2.1. We use the modified MacGregor I&F model because it strives to achieve a balance between a detailed biologically-inspired model and a highly abstracted mathematical model. Table 3.1 shows the model and its parameter settings.

	$S = \int 1 \text{if } E \ge T_h$
Snike generation	$\begin{bmatrix} S \\ - \end{bmatrix} 0$ if $E < T_h$
Spike generation	(3.1)
	$dG_K - G_K + B \cdot S$
Definentemy	$\frac{dt}{dt} = \frac{T_{CK}}{T_{CK}}$
Refractory	(3.2)
properties	amplitude of the postfiring potassium conductance decay $B=20$,
	potassium conductance time constant $T_{GK}=15$ msec
	$dT_h - (T_h - T_{h0}) + c \cdot E$
	$\frac{dt}{dt} = \frac{T_{th}}{T_{th}}$
Threshold	(3.3)
accommodation	amplitude of the threshold $C=0.5$
	amplitude of the uneshold $C=0.5$,
	result inference of the cell $T_{h0} = 10 \text{ mV}$,
	time constant for decay of threshold $T_{th}=30 \text{ mV}$
	$\underline{dE}_{-} - \underline{E} + G_{K} \cdot (E_{K} - E) + G_{e} \cdot (E_{e} - E) + G_{i} \cdot (E_{i} - E) + SCN + N$
	dt T_{mem}
Transmembrane	(3.4)
notontial	equilibrium potential of the potassium conductance $E_k = -12 \text{ mV}$
potential	equilibrium potential of the excitatory conductance $E_c=80$ mV
	equilibrium potential of the inhibitory conductance $E_e = 10 \text{ mV}$
	equinorium potential of the initiation y conductance $E_i = 10 \text{ mV}$,
	memorane une constant $I_{mem}=23$ msec

 Table 3.1: MacGregor's modified neuron model and its parameter settings

SCN is an external input current injected to the neuron. G_K , G_e and G_i represent the activity of potassium, excitatory and inhibitory synaptic conductances, respectively.



Figure 3.1: Learning functions used in the system using the SAPR (a) and the STDP (b).

3.1.2 Learning/Plasticity Rules

The dynamic SAPR rule (Swiercz et al., 2006 & 2007) that uses actual postsynaptic potentials vs. a pre-defined function like the one used in the Spike-Timing-Dependent Plasticity (STDP) rule (Song et al., 2000) is used as the plasticity rule in the network. We used both SAPR and STDP rules for the comparison purposes, see Figure 3.1. Note that, in Figure 3.1, $\Delta t = (t_{post} - t_{pre})$ is the time delay between the postsynaptic spike and the presynaptic spike. For STDP, we used $\alpha_{+-} = 1$ and $\tau_{+-} = 20$ ms.



Figure 3.2: Topology of the network: (a) high-level block diagram; (b) recurrent synaptic connections between the excitatory neurons in the feature extraction layer; and (c) synaptic connections between the excitatory neurons in the sensory/feature extraction layer and the inhibitory neurons in the feature extraction layer.

3.1.3 Network Topology

Many methods of computational neuroscience are based on the observation that the brain uses a hierarchical structure in performing cognitive tasks (Riesenhuber and Poggio, 1999, 2000), although details of this structure are yet to be fully understood. Our network also uses a hierarchical organization, from the sensory input layer to the recognition layer. The network topology for the image recognition system consists of three layers, (i) the sensory/receptive layer, which consists of only excitatory neurons; (ii) the feature extraction layer, which consists of both excitatory and inhibitory neurons; and (iii) the recognition layer, which consists of excitatory neurons (see Figure 3.2(a)). The recognition layer uses the output of the feature extraction layer to generate sequences of firings that are applied as specific signatures to recognize new images.

The sensory and feature extraction layers draw on our previous work concerning application of networks of spiking neurons in image recognition (Cios and Shin, 1995; Swiercz et al., 2006). Each sensory layer's dimension is three times larger (i.e., the corresponding area is nine times larger) than the size of the processed image to allow for overlapping between neurons in the sensory layer. The number of excitatory neurons in the feature extraction layer is also three times larger than the number of pixels in the input image, while the number of inhibitory neurons in this layer is equal to the number of pixels in the image (Swiercz et al., 2006). This results in the 9:1 ratio of excitatory to inhibitory neurons, which is consistent with estimates based on hippocampus (Traub and Miles, 1991). The inhibitory neurons provide negative feedback to prevent the network from becoming extremely excited (Kandel et al., 1996). Figure 4.2 demonstrates the importance of using inhibitory neurons in the feature extraction layer; the face shown there is more clearly outlined using inhibition than without using it. The connections from the sensory to the feature extraction layers are shown in Figure 3.2(c). Excitatory neurons in the sensory and feature extraction layers have the same dimension, $3n \times 3m$, which means that each excitatory neuron in the sensory layer is connected to the corresponding (i.e., located at the same position) excitatory neuron in the feature extraction layer. However, since inhibitory neurons in the feature extraction layer are organized as an $n \times m$ matrix, each inhibitory


Figure 3.3: Illustration of formation of the "signature" vectors in the recognition layer.

neuron in the feature extraction layer is connected to the corresponding 3×3 matrix of neurons in the sensory layer; see Figure 3.2(c).

The feature extraction layer includes three types of synaptic connections. Similar to the connections from the excitatory neurons in the sensory layer to the inhibitory neurons in the feature extraction layer, each inhibitory neuron is connected to a 3×3 matrix of excitatory neurons within the feature extraction layer, see Figure 3.2(c), and vice versa. This layer also includes recurrent connections between its excitatory neurons, where each

neuron is connected to the eight neighboring excitatory neurons (Swiercz et al., 2006), as shown in Figure 3.2(b).

The recognition layer, which consists of only excitatory neurons, collects the information coming from the feature extraction layer. It is constructed by randomly (and evenly) partitioning the total number of excitatory neurons in the feature extraction layer. Thus, for *r* neurons in the recognition layer and $n \times m$ size of the input images, each neuron in the recognition layer is randomly connected to $(3n \times 3m)/r$ excitatory neurons in the feature extraction neurons, *r*, and the number of synaptic connections, *c* (the number of synaptic connections to a neuron in the recognition layer from $(3n \times 3m)/r$ neurons in the feature extraction layer). The relation is 9nm = rc, and Table 3.2 shows values of *r* and *c* used.

Table 3.2: Number of neurons in the recognition layer and the corresponding synaptic connections from the feature extraction layer to the recognition layer for 32×32 image

Number of neurons (<i>r</i>) in the recognition layer	9	18	36	72	144	288	576	1152	2304	4608	9216
Number of synaptic connections (c)	1024	512	256	128	64	32	16	8	4	2	1

The best results are obtained using the shaded number of recognition neurons that correspond to the image size (see Figure 4.1(a) and (b)).

The recognition layer, which consists of spiking neurons, is an important part of the network. It takes input from the excitatory neurons in the feature extraction layer and summarizes the features extracted by this layer. The system tracks the excitation of neurons in the recognition layer over a short period of simulation time (300 ms) to generate and store an "organization/signature" pattern that is later used to identify similar images. The latter operation constitutes the "recall."

The "signature" pattern of a recognition neuron is a vector of its transmembrane potentials (E) recorded whenever the neuron fired (see Table 3.1). Neurons that did not fire in the recognition layer, based on their inputs from the feature extraction layer, do not form "signature" patterns.

The recognition layer of the system performs a decision making role based on the formed organization/signature pattern, separately for each class. Notice that the way in which the layer operates vaguely resembles population coding (Gerstner and Kistler, 2002; Dayan and Abbott, 2005).

During training (or organization) phase the system creates an index of the organization patterns that correspond to an image class; see Table 3.3 for the pseudo-code. When presented with a new image, the system passes it through the network and compares the outcome pattern to the "recall" table in which the previously identified "organizations" are stored. The comparison is based on a similarity measure which is defined below.

3.1.4 Image Similarity Measure

We have experimented with several image similarity measures (Yen and Cios, 2008), including the Image Similarity Measure (ISM) (Cios and Shin, 1995), Euclidean and Hamming distances, as well as spiking information. Since none of them performed well in our application, we propose a new similarity measure named the "threshold comparator," in which the vector is the transmembrane potential value of the neurons that fired in the recognition layer. The MacGregor's modified neuron model used here (see Table 3.1) fires a spike of the same absolute voltage when the transmembrane potential (eq. (3.4)) reaches the threshold (eq. (3.3)). Thus, the model is different in this aspect from the Hodgkin-Huxley neuron model, which can vary its spiking voltage due to rate effects and has no threshold value. Therefore, instead of spike voltages, we use transmembrane potential values in the comparator. That is, the comparator uses both the spike timing information (a discrete value) and the corresponding transmembrane potential (a continuous value) (Ohno-Shosaku et al., 2001; Wilson and Nicoll, 2001). The comparator "matches" a new image with the closest "organization" (signature) stored in the recall table, much like a new sensory input will influence a given brain region dedicated to processing the input, see Figure 3.4. Note that only the E values of neurons that fired are used. E_{th} is the threshold value used for finding the best match.

Given a new input image, N_{input} , the threshold comparator finds the best match after the recognition layer generates its output. When the output for the new input image is generated, the comparator compares the transmembrane potential of the neurons that have



Figure 3.4: Illustration of how the network recognizes an image using the threshold comparator.

fired with the signature vectors stored in the recall table. If the difference between the two is less than threshold E_{th} , namely:

$$\left|E_{i,j} - E_{i,j}^{N_{input}}\right| \le E_{th}, \qquad (3.5)$$

where *E* is the transmembrane potential, $1 \le i \le r$ and $1 \le j \le k$, then a match is found. After the match is found, the matching score is calculated by dividing all the matches by the total number of neurons in the recognition layer (for simulation time=300 ms)

$$M_j = \frac{\# \text{of matches}}{300 \cdot r},\tag{3.6}$$

so that $M_j \in [0, 1]$. The winning class for the input N_{input} is:

$$N_{input} \text{ winning class} = \arg \max_{1 \le j \le k} M_j.$$
(3.7)

Figure 3.4 illustrates the operation of the threshold comparator.



Figure 3.5: Examples of the Original and Occlusion & Rotation images.

3.2 Datasets

Four image datasets were used to test the performance of the network. Each dataset has different characteristics, illustrated via a few representative images given in Figure 3.5. The Japanese Female Facial Expression (JAFFE) dataset (Lyons et al., 1998) contains 213 images of 7 facial expressions which include 6 basic facial expressions and 1 neutral expression posed by 10 Japanese models. The faces are of 256×256 resolutions. The ORL (or AT&T) dataset (Samaria and Harter, 1994) contains face images of 40 people. Each



Figure 3.6: Generation of an occlusion.

picture is taken at a different time, with slightly varied lighting, different angles, open/closed eyes, glasses/no-glasses, and smiling/non-smiling facial details. The faces are of 92×112 resolutions and are represented using gray scale (one byte per resolution). The CMU dataset (Mitchell, 1997) contains 640 face images of 20 people. This dataset is characterized by varying the pose (straight, left, right, up), expression (neutral, happy, sad, angry), eyes (wearing sunglasses or not) and resolution. Although this set includes three different scale-resolutions for each image, we used the 128×120 full-resolution images. The UMIST dataset (Graham and Allinson, 1998) consists of 564 face images of 20 people. It covers a wide range of different angles of poses with variations of race, sex, and appearance. The faces are at the 92×112 resolutions.

For each data set, we prepared its occluded and rotated version by introducing occlusions into 50% of images according to the procedure shown in Figure 3.6 and by rotating 25% of images by 90, 180 and 270 degrees each in the clockwise direction. Note that, in Figure 3.6, *n* is # of vertical pixels, *m* is # of horizontal pixels, occ(x,y) is a random occlusion point, occ_{height} is a random height of occlusion rectangle where $n/3 \le occ_{height} \le 2n/3$, and occ_{width} is a random width of occlusion rectangle where $m/3 \le occ_{width} \le 2m/3$. Random numbers are distributed uniformly in the range of random height/width. While Figure 3.5 shows 10 example images for one model from each dataset, the complete Occlusion & Rotation datasets can be found in Appendix B.

3.3 Network Simulation: Organization and Recall

Given an input image N_{input} the network organizes through the sequence of the following three steps, (i) the excitatory neurons in the sensory layer are stimulated; (ii) the excitatory and inhibitory neurons in the feature extraction layer are activated; and (iii) excitatory neurons in the recognition layer are activated.

The grey image pixel values entering the sensory layer are normalized to provide proper firings of neurons, and then a sinusoidal function with amplitude of the normalized pixel intensity and a period of 300 ms in simulation time is used to stimulate the sensory layer (see Tr.4 in Table 3.3) (Barrionuevo and Brown, 1983; Swiercz et al., 2006). The sensory layer receives the input image and passes the signal to the feature extraction layer to extract visual features in terms of the excitatory neurons that fired. Then the recognition layer neurons fire in a certain pattern according to the input image. Each input image is presented to the network for 300 ms of simulation time. For each simulation, the current state of the recognition layer is recorded (the actual membrane potential (E)) for the neurons that fired.

During the organization/training phase, each input image generates its own signature vector in the recognition layer. An average of each recognition neuron's output is used to form a signature vector, for all image presentations from every class (see line Tr.13 in Table 3.3). After an input image is processed, we keep only the synapses strengths that exceed a certain threshold (i.e., 4), see line Tr.12 in Table 3.3. If synapses of some neurons did not have their values above the threshold then they are set to 0. Then, the next image in the training data is processed, etc. The formation of the signature vector is based on the spiking activity of neurons in the recognition layer. Basically, if we kept the synapses strengths after the signature vector was formed by presenting the first image to the network, then the neurons which spiked would be the only ones spiking even when other very different images were input because their synaptic strengths were set to a very high value. Thus, resetting the synaptic strength values to zero allows other neurons in the recognition layer to spike when presented with new images. This is done in order to allow all images to contribute to the formation of the signature vector; otherwise, only the first input image would form this vector. A storage needed for saving the signature vectors

depends on the number of r recognition neurons. For each recognition neuron, we have about 19 doubles, 4 double arrays, and 2 booleans (see Table 3.1).

During the recall/testing phase, the output of the neurons in the recognition layer for a given input image is used to find the best match with the signature vectors stored in the recall table by using the threshold comparator (see line Tt.3 in Table 3.3). The pseudo code for the two phases with actual parameter values is provided in Table 3.3.

Table 3.3: Pseudo code of an algorithm to compute "signature" patterns

Operation $E_{\rm eff} = E_{\rm eff}$
Organization: Computing signature vectors $E = [L_{r,k}]$
Input: - an image training data set $T_{tra} = \{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in R^{min}, i=1, \dots, N_{tra}\}$ with $y_i \in \{I, \dots, k\}$,
where \mathbf{x}_i stands for an image in $n \times m$ resolution
- a number of recognition neurons r
- a choice of synaptic plasticity such as STDP or SAPR
Output: a signature vector for each class $E = [E_{r,k}]$
Ir.1 Initialize $E = [E_{r,k}];$
Ir.2 Initialize the network based on the topology
- Arrange the sensory, feature extraction and recognition layers based on the topology explained in Fig. 3.2
- Initialize synapses in the processor
From sensory to feature extraction layer with fixed synapses of 1.4
In the feature extraction layer with random synapses in $[0.6, 2.5]$
From the feature extraction to recognition layer with random synapses in $[0.6, 2.5]$
Tr.3 For $i \leftarrow 1$ to N_{tra}
Tr.4 For time $\leftarrow 1$ to 300
- Stimulation of the sensory layer
Normalize the intensity of pixels in the image from $[0, 255]$ to $[35, 61]$
Get stimuli for each sensory neuron at the <i>time</i> by a positive-valued sinusoid with amplitude of the normalized pixel intensity and a period of 300ms
- Activation of the feature extraction layer
- Activation of the recognition layer
Tr 5 For $i \leftarrow 1$ to r
Tr 6 If $(t^{\text{th}}$ recognition neuron fired a spike)
Tr.7 $\mathbf{F}^{[time \times r+i]}$ = transmembrane potential value:
Tr 8 End If
Tr 9 End For
Tr.10 Apply synaptic plasticity by $\frac{d}{dt} w_{pq} = sig(w_{pq} + \alpha_{+-} \cdot PSP_{pq})$ between neurons p and q,
where PSP_{pq} is the postsynaptic potential value for the connection between neurons p and q,
and $sig(x)$ is a sigmoid function
IT.11 End FOF
11.12 Save only synapses that are over 4 Γ_{12} by Γ_{12} by
Ir.1.5 Update $[E_{r,k}]$ = average of $[E_{r,k}]$ and $[E_{r,k}]$;
11.15 Keurn $E = [E_{r,k}]$;
Recall: Performance of the system
Input: an image testing data set $T_{tort} = \{(\mathbf{x}_i, v_i) \mathbf{x}_i \in \mathbb{R}^{nm}, i=1, \dots, N_{tort}\}$ and $\mathbf{E} = [E_{r,k}]$
Output: a performance of the system in terms of Accuracy, Precision, Recall, and HM
Tt 1 For $N_{\text{invert}} \leftarrow 1$ to N_{invert}
Generate E^{Ninput} based on the steps 3 through 11 in Organization
Tt 3 Identify image with Threshold Comparator
- Calculate $\left E_{i,j} - E_{i,j}^{N_{input}} \right \le E_{ih}$, where $E_{ih} = 10$
Calculate the metale secret (1, c,, t,

- Calculate the match scores $M_j = \#$ of matches $300 \cdot r$

- Select winning class by N_{input} winning class = arg $\max_{1 \le j \le k} M_j$

- Tt.4End ForTt.5Evaluate the performance of the system based on eq. (4.1)Tt.6Return the performance of the system

CHAPTER 4 Experimental Results and Discussion

This chapter presents experimental results and discusses the system introduced in Chapter 3.

4.1 Evaluation

We evaluated the performance of the system using 10-fold cross-validation (10-FCV) as follows. The entire dataset is randomly partitioned into 10 subsets; 9 subsets are used for training and the remaining one is used for testing. This procedure is repeated 10 times, and the results are averaged. 10-FCV is used on both the Original images (JAFFE, ORL, CMU and UMIST) and the same sets of images after we modified them to the Occlusion & Rotations images.

The results of 10-FCV are analyzed using performance measures specified in eq. (4.1), where TP denotes true positive, TN true negative, FP false positive, and FN false negative predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Harmonic mean (HM) = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
(4.1)

The harmonic mean, which ranges between 0 and 1, is an evenly weighted harmonic mean of precision and recall.

4.2 Results and Discussion

The system is tested using two different settings for the four image datasets (see Section 3.2). For computational efficiency the images are compressed into 32×32 (in pixels) size. The results reported below are based on two experiments. One compares performance of the system using different number of neurons in the recognition layer, with the SAPR and STDP rules. The second compares the results while using only the SAPR (because it performed better) with a recognition system implemented using Support Vector Machine (SVM). The SVM-based solution was used since this classifier was previously shown to provide accurate results on the JAFFE and ORL datasets (Amine et al., 2009; Shih et al., 2008).

4.2.1 Experiment 1: Comparison of Designs with Different Number of Neurons in the Recognition Layer using the SAPR and STDP rules

The network was tested with different numbers of neurons in the recognition layer for each dataset. Figure 4.1 only shows the trend of performance of the system in terms of accuracy.

During the organization phase, the network processes instances of faces and monitors the "signature" vectors generated at the output of the recognition layer. Next, the network is tested using unseen face instances. As described in Section 3.1.4, we use the threshold comparator (with E_{th} =10) to measure similarity between the outputs of the recognition layer.

We observe that increasing the number of the recognition neurons up to a certain level leads to improved performance for both the SAPR and STDP rules; see Figure 4.1(a) and (b). For certain combinations of the number of recognition neurons and the number of synaptic connections they receive from the feature extraction layer, the system performs well. This agrees with the hypothesis that the "brain" allocates different neurons (using a fixed number) for processing a recognition task. This idea is similar to the idea of polychronization (Izhikevich, 2006), or neural Darwinism, embodied in the Theory of Neuronal Group Selection (TNGS) (Edelman, 1987). Figure 4.1(a) and (b) also show that the best results are obtained using the number of synaptic connections between c=4 and c=16. The sharp drop in performance from the number of recognition neurons r=2304 to r=4608 is caused by too few synaptic connections (c=2 and c=1, respectively) to the recognition layer (see Table 3.2). We also observe that the designed network is characterized by a relatively good performance for r between 72 and 2304, especially for the Original images processed with the SAPR rule, which would allow for computationally efficient implementations that use smaller number of neurons.

The performance comparison while using the SAPR and STDP rules is shown in Figure 4.1(c). This is done separately for the Original images and the Occlusion & Rotation images. The average and standard deviation of the difference between using the SAPR and STDP are calculated for each recognition neuron. In general, the SAPR performs better than the STDP, in particular on the Occlusion & Rotation images. The differences are between -0.84% and 36.02% for the Original images and between 1.5%and 23.15% for the Occlusion & Rotation images. Although for r between 144 and 2304 using the Original images both plasticity rules give comparable results (differences are less than 1%), in case of the Occlusion & Rotation images the network using SAPR is always better. Although we do not fully understand why the SAPR performs better than the STDP in our application, we argued that it was more biologically relevant because it relied purely on the PSPs (Swiercz et al., 2006). The significant difference between the two rules is that the SAPR is continuous around zero (see Figure 2.4 and 3.1), which was the reason for its development; we were looking for a synaptic plasticity rule that would not allow for big changes $(STDP(\Delta t) >> SAPR(\Delta t)$ at $\Delta t \rightarrow 0)$ for small time differences. Figure 4.2 shows the quality of the extracted key face features with and without inhibitory neurons in the feature extraction layer. The excitatory neuron firings are shown in red at the output of the feature extraction layer. SAPR is less prone to saturate in areas of low contrast/information than STDP (columns Figure 4.2(a) through (c)). Adjusting the synaptic strength by a small amount (according to SAPR, as opposed to STDP) when pre- and post-synaptic neurons fire very closely in time results in better recognition for the Occlusion & Rotation images. Thus, in all subsequent tests we use only the SAPR learning.



Figure 4.1: Performance of the system in terms of accuracy with different number of neurons in the recognition layer using 10-FCV: (a) using SAPR, (b) using STDP, and (c) comparison of the SAPR and STDP rules on the Original (black bars) and the Occlusion & Rotation (gray bars) images.



Figure 4.2: Graphical comparison of the results generated on the feature extraction layer using the SAPR and STDP rules on one of the images, with inhibition (columns (a) and (b)) and without inhibition (column (c)).

4.2.2 Experiment 2: Comparison with SVM

In this experiment, the system that applies the SAPR is compared with the SVM on the (a) Original and (b) Occlusion & Rotation datasets, without any preprocessing except for rescaling to the 32×32 size, see Figure 4.3(a) and (b).

The Occlusion & Rotation datasets were prepared as described in Section 3.2.

However, in order to calculate the standard deviation error for Figure 4.3(b) we generated



Figure 4.3: Comparison of the network performance in terms of accuracy using the SAPR and the SVM.

ten different variations of the data (Shin et al., 2010). Note that there are no error bars for the Original images because the results shown in Figure 4.3(a) are based on the 10-FCV.

We used 1152 neurons in the recognition layer in all experiments. To encode the images as inputs to the SVM, we converted each image into the 1024-dimentional (which corresponds to the total number of pixels) intensity vector. 10-FCV was performed on each dataset. We report the SVM result for the 2^{nd} degree polynomial kernel with the value of the complexity constant *C*=1. This kernel provides favorable predictive performance when compared with the 1^{st} and 3^{rd} degree polynomial and Radial Basis Function kernels.

As shown in Figure 4.3, SVMs perform slightly better on all Original datasets, except the CMU. However, in the case of the more challenging Occlusion & Rotation datasets, the system using the SAPR outperforms the SVMs in terms of accuracy. Results for other performance measures are provided at Appendix B. To achieve good performance using the SVM on occluded datasets, it has been known that several preprocessing tasks were performed (Déniz et al., 2003; Heisele et al., 2003). In contrast, our system does not require any preprocessing of the images. One of the key challenges in object recognition is to achieve invariance to object transformations. The especially challenging case concerns objects that are transformed in a non-linear way, which is characteristic of the Occlusion & Rotation images that are considered here. We attribute better than SVM performance of our network on the Occlusion & Rotation images to its randomly generated synaptic connections between the feature extraction layer and the recognition layer.

CHAPTER 5 Stopping Criterion

This chapter introduces a criterion for stopping iteration process in the designed system. The criterion, well- suited for real-time implementation, is based on information entropy and allows to significantly reduce the number of iterations while only slightly decreasing accuracy of recognition.

This chapter is organized as follows: we first describe an issue on the iteration process of the network in Section 5.1, the entropy based stopping criterion in Section 5.2, and the simulation results and comparison of results in Section 5.3.

5.1 Issue in Iteration Process

In the previous chapter, we focused on the practical problem of face recognition for partially occluded and rotated images. One of the issues related to the network is the iteration process. To stimulate a neuron, neurophysiologists inject current through an electrode to the transmembrane and record its potential change (Izhikevich et al., 2003; MacGregor, 1993). Depending on the shape or duration of the input currents, it has been shown that there are different spiking patterns of neurons (Aubie et al., 2009; Izhikevich, 2006; Izhikevich et al., 2003). The network for face recognition also follows similar stimulation. The input currents to the sensory layer in the network for face recognition are

based on the grayscale intensities, which are rescaled within a defined range (see tr.4 in Table 3.3). Given the input currents/information to the sensory layer, neurons in each layer generate spikes in response, and then the spike information is passed to the recognition layer and used to perform the classification task. Therefore, this mechanism causes several iteration processes corresponding to the length of input currents. We address this problem by introducing simple and effective criterion to reduce computational time.

The new stopping criterion is based on information entropy calculated from spike responses in the network. Neurons at each layer generate spike responses that carry neural representations in multiple stages of information processing. To quantify the information transmission, information theory has been often used (Borst and Theunissen, 1999; Nelken and Chechik, 2007). The stopping criterion for face recognition network is based on information entropy calculated at the recognition layer when it reaches the peak point during the iteration process. Based on this stopping criterion, the full iteration process of the network requires much less computation while only slightly affecting accuracy of recognition (specifically, less than 5% reduction in performance for threefold reduction of computation time).

5.2 Entropy-based Stopping Criterion

As mentioned Chapter 4, the network topology for face recognition is hierarchical and consists of the sensory/receptive layer, feature extraction layer, and the recognition (decision making) layer.



Figure 5.1: The network topology and its corresponding neuronal spike responses: (a) the network topology and its output of the feature extraction layer for an input image in 32×32 resolution, and (b) one example of neuronal spike responses with SAPR rule in the sensory (*s/e*), feature extraction (*f/e* and *f/i*) and recognition (*r/e*) layers.

Given an input image to the sensory layer of the network, neurons in each layer are activated during a period of 300 ms of simulation time. Figure 5.1(b) shows one example of neuronal spike responses with SAPR rule in the sensory (*s/e*), feature extraction (*f/e* and *f/i*) and recognition (*r/e*) layers for one input image. One point which should be noticed in Figure 5.1(b) is that the spike responses of the feature extraction layer are concentrated in time/iteration steps that are between 50 and 125 ms, out of a total simulation time/steps of 300 ms. Note that, in Figure 5.1, *s/e* stands for excitatory neurons in the feature extraction

layer, f/e and f/i stand for excitatory and inhibitory neurons in the feature extraction layer, and r/e stands for excitatory neurons in the recognition layer.

Neurons in each layer generate spike responses which carry neural representations in multiple stages of information processing. Information theory has been widely used to quantify information transmission (Borst and Theunissen, 1999; Nelken and Chechik, 2007). As illustrated in Figure 5.1(b), the concentration of a response is an important characteristic of the network. The neurons in each layer generate spike responses during a period of 300 ms of simulation time (iteration process). The spiking information of the neurons in the recognition layer is used only for forming a signature vector for each class. To measure this information content we use Shannon's entropy

$$E_{entropy} = -\sum_{i=1}^{n} p_i \log_2 p_i,$$
(5.1)

where *n* is the number of neurons, and p_i is the probability for the neuron *i* spike ($p_i = number of spikes / simulation time$).

Based on the entropy, we experimentally show that entropy decreases as the information is processed in the hierarchical network for each simulation time (iteration process). Figure 5.2 shows performance of the network (Figure 5.2(c)) according to the changes of entropies (Figure 5.2(d)) for one (JAFFE (Lyons et al., 1998)) dataset while the spiking responses (Figure 5.2(a)) and the changes of organization of excitatory neuronal spike patterns through simulation time (Figure 5.2(b)) are based on one image. Note that, in Figure 5.2, *s/e* stands for excitatory neurons in the feature extraction layer, *f/e*



Figure 5.2: Relationship between an entropy variation and a performance of the network for JAFFE data in 32×32 resolution using SAPR; (a) neuronal spike responses in the sensory (*s/e*), feature extraction (*f/e* and *f/i*) and recognition layers (*r/e*), (b) the changes of organization of excitatory neuronal spike patterns through simulation time in the sensory (*s/e*) and feature extraction layer (*f/e*), (c) an entropy variation at hierarchy through simulation time, and (d) a performance of the network through simulation time.

and f/i stand for excitatory and inhibitory neurons in the feature extraction layer, r/e stands for excitatory neurons in the recognition layer, and i=80 and j=20.

Entropy at each layer first increases and then decreases (after a certain amount of simulation time). An interesting observation is that the accuracy of the network also stabilizes at the point when entropies start to decrease, with the exception of the sensory layer. Note that information entropy becomes smaller as it passes through the network such

that at the recognition layer it is the smallest (see Figure 5.2(c)), while at the same time preserving enough information for classification purpose (see Figure 5.2(d)).

Based on the above, we introduce a more efficient stopping criterion based on spiking entropy, which reduces the number of iterations while only slightly affecting accuracy. The algorithm for the stopping criterion is shown in Table 5.1.

Table 5.1: Pseudo code of an algorithm for the stopping criterion

```
Organization: Computing signature vectors E = [E_{r,k}] and a stopping time
Input: - an image training data set T_{tra} = \{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in \mathbb{R}^{nm}, i=1, \dots, N_{tra}\} with y_i \in \{1, \dots, k\},
              where \mathbf{x}_i stands for an image in n \times m resolution
         - a number of recognition neurons 9 \cdot 2^r
         - a choice of synaptic plasticity such as STDP or SAPR
Output: - a signature vector for each class \boldsymbol{E} = [E_{r,k}]
           - a stopping time
Tr.1 Initialize \boldsymbol{E} = [E_{r,k}]; Entropy (r/e) = 0; stopping time = 0;
       Initialize the network based on the topology
Tr.2
Tr.3 While Entropy (r/e) does not reach maximum
          For i \leftarrow 1 to N_{tra}
Tr 4
Tr.5
              - Stimulation of the sensory layer, activation of the feature extraction layer, and activation of the recognition layer
Tr.9
              For j \leftarrow 1 to 9 \cdot 2^r
Tr.10
                 If (j^{\text{th}} recognition neuron fired a spike )
                    \mathbf{E}^{i}[simulation time \times 9 \cdot 2^{r} + j] = transmembrane potential value;
Tr.11
Tr.12
                 End If
Tr.13
              End For
Tr.14
              Apply synaptic plasticity
Tr.16
           End For
Tr.18
          Update [E_{r,k}] = average of [E_{r,k}] and [E_{r,k}^{i}];
Tr.19
           simulation time = simulation time + 1;
Tr.20 End While
Tr.21 Return E = [E_{r,k}]; stopping time = simulation time;
```

Recall: Testing the system

Input: an image testing data set $T_{test} = \{(\mathbf{x}_{i}, y_{i}) | \mathbf{x}_{i} \in R^{nm}, i=1, \dots, N_{test}\}$ and $\mathbf{E} = [E_{r,k}]$ Output: a performance of the system in terms of Accuracy, Precision, Recall, and HM Tt.1 For $N_{input} \leftarrow 1$ to N_{test} Tt.2 Generate E^{Ninput} based on the steps 9 through 16 in **Organization** until the simulation time reaches the stopping time Tt.3 Identify image with Threshold Comparator defined as follows: - Calculate match as $\left| E_{i,j} - E_{i,j}^{N_{input}} \right| \leq E_{th}$, where $E_{th} = 10$ - Calculate match as $M_{j} = \# \text{ of matches} / 300 \cdot r$ - Select winning class by N_{input} winning class = arg $\max_{1 \leq j \leq k} M_{j}$ Tt.4 End For Tt.5 Evaluate the performance of the system based on eq. (4.1)

Tt.6 Return the performance of the system

5.3 Results and Discussion

Simulations results are presented in this section. Four different sets of image datasets (JAFFE (Lyons et al., 1998), ORL (or AT&T) (Samaria and Harter, 1994), CMU (Mitchell, 1997) and UMIST (Graham and Allinson, 1998)) that have been rotated and partially occluded are used to test the performance of the network (see the entire datasets in the Appendix A). The images are compressed into 32×32 (in pixels) size for computational efficiency. We choose 576 neurons in the recognition layer because that much number belongs to the medium performance range (from 576 to 2304 neurons in the recognition layer) of the system. In reporting all results, we use 10-fold cross-validation. All simulations are performed on a computer with a 2.2-GHz Core 2 Duo E4500 processor with 2GB of RAM.

The results shown below are for the SAPR and STDP learning rules. During the simulations, accuracies of the system, changes of computational time (in minutes and in an absolute time unit), and changes of total entropies for each layer are recorded. Because there are no spike responses from neurons in the recognition layer before a certain amount of simulation time (iterations) all results are reported from the simulation/iteration time 80s until 300s.

The key characteristics of the system are described below.

The dynamics of the system performance can be divided into three stages: silent, oscillating, and stable (see Figure 5.3). At the beginning of simulation (iterations)

no recognition neurons fired because neurons in the sensory and feature extraction layers did not fire either. In the silent stage, the neurons in the recognition layer start becoming organized, but none of them exceeded threshold to generate a spike. Once neurons in the sensory layer start responding to the input stimuli, they trigger the neurons in the feature extraction layer to fire, causing the neurons in the recognition layer to also start firing spikes. Since majority of the neurons in the recognition layer still compete with each other to fire a spike, the system is in the unstable oscillating stage. Eventually only a small number of recognition neurons will keep firing and thus the system converges into a stable stage. The three stages are identified based on the simulation (iteration) experiments. In Figure 5.3, we show the three stages when using either the SAPR or the STDP plasticity rule. For the SAPR (Fig. 5.3(a)), we have the silent stage between occurring 0 and 90 iterations, the oscillating stage between 90 and 120 iterations, and the stable stage from 120 to 300 iterations. Similarly, for the STDP rule (Fig. 5.3(b)), the three stages are; 0 to 100 for the silent stage, 100 to 130 for the oscillating stage, and 130 to 300 for the stable stage. Let us note that the system reaches the stable stage with the SAPR rule *faster* than with the STDP rule.





Figure 5.3: Performance of the system with accuracy; (a) with SAPR, and (b) with STDP.

Figure 5.4 shows changes of entropy (average and standard deviation) for the eight datasets, with both SAPR and STDP. For each simulation (iteration) step we observe that entropy of the sensory layer is greater than the entropy of the feature extraction layer, which is still greater than the entropy of the recognition layer. This shows that neural firings at multiple stages of hierarchical information processing become more compact as measured by entropy. An interesting observation is that the entropy at each layer of the network initially increases before it becomes decreased. However, the entropy of the sensory layer decreases significantly more slowly than of the other layers. These phenomena are clearly seen when we look at the neuronal spike responses shown in Figure 5.2(a). Note that the average and the standard deviation of total entropies for all eight datasets are shown in Figure 5.4.



Figure 5.4: Changes of entropy in the sensory (*s/e*), feature extraction (f/e and f/i) and recognition layers (r/e); (a) with SAPR, and (b) with STDP.

Third, the computational time increases linearly as shown in Figures 5.5 and 5.6. Figure 5.5 depicts changes of the total CPU time (in minutes) for the eight datasets, while using either SAPR (Figure 5.5(a)) or STDP (Figure 5.5(b)) rules. In addition, Figure 5.6 illustrates changes of the total computational time (in an absolute unit), while using either SAPR (Figure 5.6(a)) or STDP (Figure 5.6(b)) rules. Because of different number of images in each of the eight datasets the computational times in CPU are different (see Figure 5.5). However, when normalized to 300 iteration steps and averaged over all datasets it shows that the computational times increase linearly (see Figure 5.6). We can easily identify a stopping point for iterations as corresponding to the maximum value of entropy at the recognition layer; see Figure 5.4. For the SAPR it is at 90 iterations step, and for the STDP it is 100. The implications of this result are very significant as we can save much time and still be able to identify the images without a significant decrease in performance. For example (Figure 5.6), the total computational time can be reduced by 2/3 for application of both SAPR and STDP rules. Note that the average and the standard deviation of total computational time in an absolute unit for all eight datasets are shown in the Figure 5.6.





Figure 5.5: Changes of the total CPU time in minutes; (a) with SAPR, and (b) with STDP.



Figure 5.6: Changes of the total computational time in an absolute unit; (a) with SAPR, and (b) with STDP.

From Figures 5.3, 5.4, 5.5 and 5.6, we identify a stopping point of the simulation time as the one when entropy of the recognition layer reaches maximum; 90 for SAPR and 100 for STDP (based on the simulations). However, this holds true for all datasets in terms of absolute computational time unit without much degradation of error/performance. Therefore, the total computational time can be reduced to 0.329 ± 0.001 for SAPR and 0.369 ± 0.010 for STDP, out of 1.00 (for simulation time = 300), respectively. This is the most important characteristic of the network because the simulation time because we can reduce the number of iterations steps, and thus time three-fold (as it does not need to continue until it reaches 300) to achieve good performance system without compromising accuracy.

CHAPTER 6 Conclusions

In spite of the ever-increasing computational power (both in speed and memory) there exist many difficult computational problems, such as weather prediction, optimization problems, genomics, proteomics and neuroinformatics problems, and image recognition, that still are not easily solved by the existing computational tools. One of the reasons is that conventional solving methods are limited in their sequential solving power when they assume that different causes of distortions are independent and can be isolated. For instance, given that an assumption that the occlusion and rotation are independent, we can first solve the first problem and later address the second. However, such independency often cannot be proven, and therefore such sequential solutions fail.

We focused on a difficult task of recognizing partially occluded and rotated images and borrowed the ideas from neuroscience for designing a new classifier. We have shown that a satisfactory performance of the system was achieved using a limited number of resources, in terms of a specific number of recognition neurons and synaptic connections, for a given input image size.

6.1 Summary

We have developed a novel image recognition system, based on a network of spiking neurons, for solving computationally challenging task of recognizing partially occluded and rotated face images. The developed network is hierarchical and consists of three layers: input, feature extraction, and recognition. During organization/training phase, the system generates signature vectors and saves them in the recall table. During the recall/testing phase, the new output is matched with the stored signature vectors. The matching is performed using the introduced image similarity measure, the threshold comparator, based on the transmembrane potential values of neurons that generated spikes.

Although it is known that information transfer between the neurons in the brain is much slower than that in modern computers, its organization allows for superior human performance in difficult tasks like image recognition. Thus, we designed our network using simple clues from the still limited knowledge of how the brain processes information. They included hierarchical organization, random synaptic connections, and the biologically plausible learning rules.

Using the system we compared performance of the SAPR and STDP synaptic plasticity rules on the Original face images and on the more difficult Occlusion & Rotation images. The results showed that the SAPR performed better on both types of images. We attributed it to its continuity around zero that resulted in small adjustments in contrast to big adjustments when using STDP. We have also shown that a satisfactory performance of the system could be achieved using a specific network topology (i.e., a specific number of
the recognition neurons and a specific number of synaptic connections), given an image size.

Next, we compared the system with the SVM classifier. The SVM performed negligibly better on three out of four Original images. However, our system significantly outperformed SVM, between 5% to 10% in terms of accuracy, on the challenging Occlusion & Rotation images. The results showed that it was beneficial to borrow the ideas/mechanisms of how brains process information for designing our system. At the same time we do not claim biological plausibility of our system.

We also introduced a stopping criterion for iterations based on entropy. It efficiently reduced the number of required iterations while only slightly affecting classification performance. The entropy-based stopping criterion significantly shortened iterative process and thus improved computational performance. The criterion is wellsuited for real-time implementation in designing classification systems.

The dynamics of the system's performance were divided into the silent, oscillating, and stabilizing periods. The information entropy decreased as the information was processed in the hierarchical structure of the network. The information was compressed during hierarchical processing so that at the last, recognition layer, the entropy was smaller than at the preceding layer, while preserving enough information to perform accurate classification. An important fact was that the performance of the network was stabilized at the point when entropies start to decrease, except for the sensory layer.

61

The most important characteristic of the new entropy-based stopping criterion was that the simulation time could be reduced three-fold without compromising accuracy. This is important for other applications where it is important to make a quick decision.

We have shown that our network of spiking neurons was successfully applied to a difficult task of recognizing partially occluded and rotated images. Although we do not claim any biological plausibility, we have shown that mimicking the way of how the brain processes information and designing an automated system helped in achieving good recognition results that outperformed state-of-the-art machine learning methods.

The main contributions of the system are its novel hierarchical topology, a new image similarity measure, and an entropy-based stopping criterion.

We think that our general approach can help in solving other difficult engineering and computational problems, such as weather prediction, optimization problems, problems in bioinformatics, and image understanding that still are not easily solved by existing computational tools.

6.2 Future Work

We plan to perform experiments on more datasets with different parameter settings for generating occluded and rotated datasets. Several issues in object recognition, such as scale, rotation and position invariance, will be also deeply studied. Moreover, we plan to further investigate the role of inhibitory neurons to know how crucial their presence is in the feature extraction layer or recognition layer for the performance of the system. There are many issues which have not been solved, for example, whether inhibitory neurons update their weights.

Possible improvement of the system is its computational time efficiency. It could be improved by using multithreaded implementations or other simple spiking neuron models.

It would be also interesting to see if the computational time of the network could be minimized by using other probability functions in the entropy based stopping criteria.

Although the presented results were good a detailed study of all aspects of the network is planned to fully understand fundamental principles governing its behavior.

Bibliography

Bibliography

A. Amine, S. Ghouzali, M. Rziza, and D. Aboutajdine. 2009. An improved method for face recognition based on SVM in frequency domain. *Machine Graphics and Vision*, 18(2):187–199.

B. Aubie, S. Becker, and P.A. Faure. 2009. Computational models of millisecond level duration tuning in neural circuits. *Journal of Neuroscience* 29(29):9255-9270.

P.F. Baldi and K. Hornik. 1995. Learning in linear neural networks: a survey. *IEEE Tr Neural Networks* 6(4):837-858.

G. Barrionuevo, T.H. Brown. 1983. Associative long-term potentiation in hippocampal slices. *PNAS* 80:7347-7351.

G. Bi and M. Poo. 1998. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience* 18(24): 10464-10472.

S. M. Bohte, J. N. Kok, and J. A. La Poutré. 2002. Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing* 48(1-4):17-37.

A. Borst and F.E. 1999. Theunissen, Information theory and neural coding. *Nature neuroscience* 2(11):947-957.

J.M. Buhmann, T. Lange, and U. Ramacher. 2005. Image segmentation by networks of spiking neurons. *Neural Computation* 17(5):1010-1031.

G. Buzsaki. 2004. Large-scale recording of neuronal ensembles. *Nature Neuroscience* 7(5):446-451.

R. Chellappa, C.L. Wilson, and S. Sirohey. 1995. Human and machine recognition of faces: a survey. *Proceedings of the IEEE* 83(5):705-741.

Y. Choe and R. Miikkulainen. 1998. Self-organization and segmentation in a laterally connected orientation map of spiking neurons. *Neurocomputing* 21:1-3.

K.J. Cios, W. Pedrycz, R.W. Swiniarski, and L.A. Kurgan. 2007. *Data Mining: A Knowledge Discovery Approach*. Springer.

K.J. Cios and I. Shin. 1995. Image recognition neural network: IRNN. *Neurocomputing* 7(2):159-185.

K.J. Cios, W. Swiercz, and W. Jackson. 2004. Networks of spiking neurons in modeling and problem solving. *Neurocomputing* 61:99-119.

K.J. Cios and L.M. Sztandera. 1997. Ontogenic neuro-fuzzy algorithm: F-CID3. *Neurocomputing* 14(4):383-402.

P. Dayan and L. F. Abbott. 2005. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press.

G. Deco and E. Rolls. 2004. A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research* 44 (6):621-642.

A. Delorme and S.J. Thorpe. 2001. Face identification using one spike per neuron: resistance to image degradations. *Neural Netw* 14: 795-803.

A. Delorme and S.J. Thorpe. 2003. SpikeNET: an event-driven simulation package for modeling large networks of spiking neurons. *Network: Computation in Neural Systems* 14:513-627.

O. Déniz, M. Castrillón, and M. Hernández. 2003. Face recognition using independent component analysis and support vector machines. *Pattern Recognition Letters* 24(13):2153-2157.

G.M. Edelman. 1987. *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.

G. M. Edelman. 1999. Building a Picture of the Brain. *ANNALS of the New York Academy of Sciences* 882(1):68-89.

E. Fiesler and K.J. Cios. 1997. Supervised ontogenic neural networks. *Handbook on Neural Computation* C1.7.

R. FitzHugh. 1961. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal* 1(6):445–466.

S. Fusi. 2002. Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biological Cybernetics* 87: 459-470.

S. Fusi, M. Annunziato, D. Badoni, A. Salamon, and D.J. Amit. 2000. Spike-driven synaptic plasticity: theory, simulation, VLSI implementation. *Neural Computation* 12: 2227-2258.

W. Gerstner and W. M. Kistler. 2002. *Spiking Neuron Models: Single Neurons, Populations, Plasticity.* Cambridge.

J.I. Gold and M.N. Shadlen. 2001. Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences* 5(1):10-16.

D.B. Graham and N.M. Allinson. 1998. Characterizing virtual eigensignatures for general purpose face recognition. *Face recognition: from theory to applications*, H. Wechsler, P. J. Phillips, V. Bruce, F. Fogelman-Soulie, and T. S. Huang, Ed., 1998, vol. 163, NATO ASI Series F, Computer and Systems Sciences, pp. 446-456.

D.O. Hebb. 1949. The organization of behavior. New York: Wiley.

P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C.J. Honey, V.J. Wedeen, and O. Sporns. 2008. Mapping the strucutral core of human cerebral cortex. *PLOS* 6(7):1479-1493.

B. Heisele, P. Ho, J. Wu, and T. Poggio. 2003. Face recognition: component-based versus global approaches. *Computer Vision and Image Understanding* 91(1-2):6-21.

J. L. Hindmarsh and R. M. Rose. 1982. A model of the nerve impulse using two first-order differential equations. *Nature* 296:162–164.

J. L. Hindmarsh and R. M. Rose. 1984. A model of neuronal bursting using three coupled first order differential equations. *Proceedings of the Royal Society of London, Seies B, Biological Sciences*, 221(1222):87–102.

A.L. Hodgkin and A.F. Huxley. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology* 117:500-544.

E.M. Izhikevich. 2003. Simple model of spiking neurons. *IEEE Transactions on Neural Networks* 14:1569-1572.

E.M. Izhikevich. 2006. Polychronization: computation with spikes. *Neural Computation* 18:245-282.

E.M. Izhikevich. 2007. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Celebral Cortex* 17:2443-2452.

E.M. Izhikevich. 2007. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. MIT Press.

E.M. Izhikevich, N.S. Desai, E.C. Walcott, and F.C. Hoppensteadt. 2003. Bursts as a unit of neural information: selective communication via resonance. *Trends in Neuroscience* 26(3):161-167.

E.M. Izhikevich and G.M. Edelman. 2008. Large-scale model of mammalian thalamocortical systems. *PNAS* 105(9):2593-2598.

E. M. Izhikevich, J. A. Gally, and G. M. Edelman. 2004. Spike-timing Dynamics of Neuronal Groups. *Cerebral Cortex* 14:933-944.

H.R. Kanan, K. Faez, and Y. Gao. 2008. Face recognition using adaptively weighted patch PZM array from a single exemplar image per person. *Pattern Recognition* 41(12): 3799-3812.

E.R. Kandel, J.H. Schwartz, and T.M. Jessell. 1996. *Essentials of Neural Science and Behavior*. McGraw-Hill.

C. Koch and I. Segev. 2000. The Role of Single Neurons in Information Processing. *Nature Neuroscience Supplement* 3:1171-1177.

T. Kohonen. 2006. Self-Organizing Maps. Springer.

J. Konorski. 1948. *Conditioned Reflexes and Neuron Organization*. Cambridge University Press.

M. Lengyel, J. Kwag, O. Paulsen and P. Dayan. 2005. Matching storage and recall: hippocampal spike timing-dependent plasticity and phase response curves. *Nat. Neurosci.* 8(12): 1677-1683.

J. Lisman and N. Spruston. 2005. Postsynaptic depolarization requirements for LTP and LTD: a critique of spike timing-dependent plasticity. *Nat. Neurosci.* 8(7): 839-841.

J. Lovelace and K.J. Cios. 2007. A very simple spiking neuron model that allows for efficient modeling of complex systems. *Neural Computation* 19:1-26.

M.J. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba. 1998. Coding facial expressions with gabor wavelets. *In Proceedings of the Third International Conference on Automatic Face and Gesture Recognition*, 200-205, April 14-16, Nara, Japan.

R.J. MacGregor. 1993. *Theoretical Mechanics of Biological Neural Networks*. Academic Press.

H. Markram. 2006. The blue brain project. Nature Reviews Neuroscience 7:153-160.

T. Masquelier and S.J. Thorpe. 2007. Unsupervised learning of visual features through spike timing dependent plasticity. *PLoS Computational Biology* 3(2):247-257.

W.S. McCulloch and W.H. Pitts. 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematics and Biophysics* 5:115-133.

M. Megias, ZS. Emri, T.F. Freund, and A.I. Gulyas. 2001. Total number and distribution of inhibitory and excitatory synapses on hippocampal CA1 pyramidal cells. *Neuroscience* 102(3):527-540.

B.W. Mel. 1997. SEEMORE: combining color, shape, and texture histogramming in a neurally inspired approach to visual object recognition. *Neural Computation* 9(4) 777-804.

R. Miikkulainen, J. A. Bednar, Y. Choe, and J. Sirosh. 2005. *Computational Maps in the Visual Cortex*. Springer.

T. Mitchell. 1997. Machine Learning. McGraw Hill.

R.J. Morgan and I. Soltesz. 2008. Nonrandom connectivity of the epileptic dentate gyrus predicts a major role for neuronal hubs in seizures. *PNAS* 105(16):6179-6184.

C. Morris and H. Lecar. 1981. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal* 35(1):193–213.

M. Muller-Linow, C.C. Hilgetag, and M.T. Hutt. 2008. Organization of Excitable Dynamics in Hierarchical Biological Networks. *PLOS* 4(9):1-15.

J. Mutch and D. Lowe. 2006. Multiclass object recognition with sparse, localized features. *In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 11-18, June 17-22, in New York, U.S.A.

J. Nagumo, S. Arimoto, and S. Yoshizawa. 1962. An active pulse transmission line simulating nerve axon. *Proceedings of the Institute of Radio Engineers (IRE)* 50(10):2061–2070.

I. Nelken and G. Chechik. 2007. Information theory in auditory research. *Hearing research* 229(1-2):94-105.

T. Ohno-Shosaku, T. Maejima, and M. Kano. 2001. Endogenous cannabinoids mediate retrograde signals from depolarized postsynaptic neurons to presynaptic terminals. *Neuron* 29(3):729-738.

B.A. Olshausen, C.H. Anderson, and D.C. Van Essen. 1993. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience* 13:4700-4719.

A. Plebe. 2008. The Ventral Visual Path: Moving Beyond V1 with Computational Models. in *Visual Cortex: New Research*. T. A. Portocello and R. B. Vellotiin, Eds. Nova Science Publishers.

A. Plebe. 2007. A model of angle selectivity development in visual area V2. *Neurocomputing* 70(10-12):2060-2063.

A. Plebe and R. G. Domenella. 2006. Early development of visual recognition. *BioSystems* 86(1-3):63-74.

A. Plebe and R. G. Domenella. 2007. Object recognition by artificial cortical maps. *Neural Networks* 20(7):763-780.

M. Riesenhuber and T. Poggio. 1999. Hierarchical models of object recognition in cortex. *Nat Neurosci* 2:1019-1025.

M. Riesenhuber and T. Poggio. 2000. Models of object recognition. *Nat Neurosci* 3:1199-1204.

E.T. Rolls and G. Deco. 2002. *Computational Neuroscience of Vision*. Oxford University Press.

F. Rosenblatt. 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review* 65(6):386-408.

D.M. Sala and K.J. Cios. 1998. Self-organization in networks of spiking neurons. *Australian Journal of Intelligent Information Processing Systems* 5(3):161-170.

D.M. Sala and K.J. Cios. 1999. Solving graph algorithms with networks of spiking neurons. *IEEE Tr Neural Networks* 10(4):953-957.

A. Samal and P.A. Lyengar. 1992. Automatic recognition and analysis of human faces and facial expressions: a survey. *Pattern Recognition* 25(1):65-77.

F.S. Samaria and A.C. Harter. 1994. Parameterization of a stochastic model for human face identification. *In Proceedings of the Second IEEE Workshop on Applications of Computer Vision*, 138-142, December 5-7, in Sarasota, U.S.A.

T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio. 2007. Robust object recognition with cortex-like mechanisms. *IEEE Tr on Pattern Analysis and Machine Intelligence* 29 (3):411-426.

JH. Shin, D. Smith, W. Swiercz, K. Staley, T. Rickard, J. Montero, L. Kurgan, and K.J. Cios. 2010. *IEEE Transactions on Neural Networks* accepted for publication in spring 2010. Supplementary Materials to http://www.egr.vcu.edu/cs/dmb/Projects/NSN/TNN_supplementary.html

F. Y. Shih, C.-F. Chuang, and P. S. P. Wang. 2008. Performance comparisons of facial expression recognition in JAFFE database. *International Journal of Pattern Recognition and Artificial Intelligence*, 22(3):445–459.

S. Song, K.D. Miller, and L.F. Abbot. 2000. Competitive hebbian learning through spike timing-dependent synaptic plasticity. *Nat Neurosci* 3:919-926.

S.M. Stringer and E.T. Rolls. 2002. Invariant object recognition in the visual system with novel views of 3D objects. *Neural Computation* 14(11):2585-2596.

W. Swiercz, K.J. Cios, J. Hellier, A. Yee, and K. Staley. 2007. Effects of synaptic depression and recovery on synchronous network activity. *Journal of Clinical Neurophysiology* 24(2):165-174.

W. Swiercz, K.J. Cios, K. Staley, L. Kurgan, F. Accurso, and S. Sagel. 2006. New synaptic plasticity rule for networks of spiking neurons. *IEEE Tr Neural Networks* 17(1):94-105.

X. Tan, S. Chen, Z.H. Zhou, and F. Zhang. 2005. Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft *k*NN ensemble. *IEEE Tr Neural Networks* 16(4):875-886.

R.D. Traub and R. Miles. 1991. Neuronal Networks of the Hippocampus. Cambridge.

M. Tsodyks and C. Gilbert. 2004. Neural networks and perceptual learning. *Nat* 431:775-781.

X.J. Wang. 2002. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36:955-968.

R.I. Wilson and R.A. Nicoll. 2001. Endogenous cannabinoids mediate retrograde signalling at hippocampal synapses. *Nature* 410:588-592.

Q. Wu, M. McGinnity, L. Maguire, A. Belatreche, and B. Glackin. 2007. Edge detection based on spiking neural network model. *Lecture Notes in Computer Science* 4682:26-34.

S.G. Wysoski, L. Benuskova, and N. Kasabov. 2008. Fast and adaptive network of spiking neurons for multi-view visual pattern recognition. *Neurocomputing* 71:2563-2575.

M. Yanike, S. Wirth, and W.A. Suzuki. 2004. Representation of well-learned information in the monkey hippocampus. *Neuron* 42(3):477-487.

C-Y. Yen and K.J. Cios. 2008. Image recognition system based on novel measures of image similarity and cluster validity. *Neurocomputing* 72(1-3):401-412.

G.P. Zhang. 2000. Neural networks for classification: a survey. *IEEE Tr on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 30(4):451-462.

Appendices

APPENDIX A Occlusion and Rotation Datasets

Simulations in this dissertation were performed with four different sets of image datasets; JAFFE, ORL (or AT&T), CMU and UMIST. For each data set, we prepared its occluded and rotated version by introducing occlusions into 50% of images according to the procedure shown in Figure 3.6 and by rotating 25% of images by 90, 180 and 270 degrees each in the clockwise direction (explained in Section 3.2). We generated ten different variations of the occluded and rotated data. One of the entire variations is shown below.



Figure A.1: The Occlusion & Rotation images of JAFFE dataset.



Figure A.2: The Occlusion & Rotation images of ORL dataset.



Figure A.3: The Occlusion & Rotation images of CMU dataset.



Figure A.4: The Occlusion & Rotation images of UMIST dataset.

APPENDIX B Other Performance Measures

The appendix B provides performance results for other measures; precision, recall and F-measure.



Figure B.1: Performance of the system in terms of precision with different number of neurons in the recognition layer using 10-FCV: (a) using SAPR, (b) using STDP, and (c) comparison of the SAPR and STDP rules on the Original (black bars) and the Occlusion & Rotation (gray bars) images.



Figure B.2: Performance of the system in terms of recall with different number of neurons in the recognition layer using 10-FCV: (a) using SAPR, (b) using STDP, and (c) comparison of the SAPR and STDP rules on the Original (black bars) and the Occlusion & Rotation (gray bars) images.



Figure B.3: Performance of the system in terms of Harmonic mean with different number of neurons in the recognition layer using 10-FCV: (a) using SAPR, (b) using STDP, and (c) comparison of the SAPR and STDP rules on the Original (black bars) and the Occlusion & Rotation (gray bars) images.



Figure B.4: Comparison of the network performance in terms of precision using the SAPR and the SVM.



Figure B.5: Comparison of the network performance in terms of recall using the SAPR and the SVM.



Figure B.6: Comparison of the network performance in terms of Harmonic mean using the SAPR and the SVM.

VITA

Joo-Heon Shin received a MS majored in Computer Science from San Diego State University (SDSU) in 2006, and he received his MS and BS majored in Mathematics from Hankuk University of Foreign Studies (HUFS) in 2000 and 1998, respectively. His research interests are in the area of Neuroinformatics, Data Mining, Mathematical Modeling, and Medical Informatics.

EDUCATION

- M.S. San Diego State University (SDSU), San Diego, U.S.A.
 Computer Science, May 2006
 Modeling and Recognition of Breast Cancer Based on Mammographic Images Advisor: Prof. RW Swiniarski
 M.S. Hankuk University of Foreign Studies (HUFS), Seoul, Korea
 - Mathematics, May 2000 *On Real Hypersurfaces of Type A in a Complex Space Form* Advisor: Prof. In-Bae Kim
- B.S. Hankuk University of Foreign Studies (HUFS), Seoul, Korea Mathematics, February 1998

GRANTS

Involved in computational modeling part of the grant:

"Computational Models of Multisensory Processing", NIH-NS064675, 2008-2011

EMPLOYMENT

- Research Assistant of Computer Science, Virginia Commonwealth University (VCU), 2007-present
 - Conducting individual research
- Research Assistant of Robotics and Neural Networks Laboratory, Computer Science, San Diego State University (SDSU), 2005-2006
 - o Responsibilities included research and counselling students
- Teaching Assistant, Mathematics, Hankuk University of Foreign Studies (HUFS), 1999-2000
 - o Responsibilities included teaching, research and counselling students

PUBLICATIONS

Journal Articles

 JH. Shin, D. Smith, W. Swiercz, K. Staley, J.T. Rickard, J. Montero, L. Kurgan, K.J. Cios: Recognition of Partially Occluded and Rotated Images with a Network of Spiking Neurons. *IEEE Transactions on Neural Networks* <u>accepted</u> for publication in spring 2010

- HK. Lim, L. Keniston, JH. Shin, B. Allman, MA. Meredith, K.J. Cios: Connectional Parameters Determine Multisensory Processing in a Spiking Network Model of Multisensory Convergence. Under review for *Journal of Computational Neuroscience* (2010)
- JH. Shin, L. Keniston, N. Cao, HK. Lim, K.J. Cios: Entropy-based Early Stopping Criterion for a Network of Spiking Neurons in Image Classification. Under review for *Neurocomputing* (2010)
- B. Kim, JH. Shin: On Real Hypersurfaces of Type A in a Complex Space Form. J.
 Basic Sci., HUFS, 10: 7-17 (2000)

Conference Articles

- HK. Lim, L. Keniston, JH. Shin, C. Nguyen, AM., K.J. Cios: A Neuronal Multisensory Processing Simulator. <u>Accepted</u> at IJCNN (2010)
- JH. Shin, D. Smith, R.W. Swiniarski, F.E. Dudek, A. White, K. Staley, K.J. Cios: Analysis of EEG Epileptic Signals with Rough Sets and Support Vector Machines. AIME 2009: 325-334 (2009)

Posters

 HK. Lim, JH. Shin, C. Nguyen, L.P. Keniston, B.L. Allman, M.A. Meredith, K.J. Cios: Modeling Multisensory Integration with a Spiking Neuron Network. Neuroscience 2009: G.06.290.23 (2009)

PROFESSIONAL AFFILIATIONS

- Student member of Society for Neuroscience (SFN), 2008-2010
- Student member of Korean-American Scientists and Engineers Association (KSEA) in San Diego, 2001-2005
- Student member of Korean Mathematical Society (KMS) in Korea, 1999-2000

DEPARTMENT SERVICE

Member of the Student Advisory Board, Computer Science Department, VCU, 2008-

2010

Representing doctoral students at the SAB; a student group established for the purpose of involving students in all aspects of the life of the Department