

Progress in Neural Networks

VOLUME 2

edited by
OMID M. OMIDVAR

TP18
P 964
V.2

9561669

PROGRESS IN NEURAL NETWORKS

Volume 2

edited by

Omid M. Omidvar



E9561669



**ABLEX PUBLISHING CORPORATION
NORWOOD, NEW JERSEY**

Copyright © 1994 by Ablex Publishing Corporation

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without permission of the publisher.

Printed in the United States of America

ISSN: 89-696306

ISBN: 0-89381-735-4

Ablex Publishing Corporation
355 Chestnut Street
Norwood, New Jersey 07648

*To my parents
who taught me love and respect*

Preface

Progress in Neural Networks is a series concerned with the advances in neural networks—natural and synthetic. This series reviews state-of-the-art research in modeling, analysis, design, and development of neural networks in software and hardware areas. This series is intended to serve as a platform for detailed and expanded discussion of topics of interest to the neural network and cognitive information-processing communities. We hope the series will help to shape and define academic and professional progress in this area. This series seeks contributions from leading researchers and practitioners to provide its audience with a wide variety of in-depth discussion of active research and presentation of complex ideas in the neural networks field. The second volume consists of chapters that are self-contained and tutorial in nature; however, one must have a background in general information about neural networks to appreciate the depth and complexity of the research presented here. This series is intended for a wide audience, those professionally involved in neural networks research, such as lecturers and primary investigators in neural computing, neural modeling, neural learning, neural memory, and neurocomputers.

The first chapter in this volume focuses on self-organizing neural classifier for complex imagery. The second chapter deals with constraint satisfaction networks for vision. The third chapter is devoted to neural networks for position, scale, and rotation invariant pattern recognition. Chapter 4 is on generalization of back propagation networks. Chapter 5 deals with optical implementation of closest vector selection in neural networks. Chapters 6 and 7 are devoted to the study of the biological basis for artificial neural networks and analysis of neuronal spike trains. Chapter 8 is of interest to scientists in the area of unsupervised learning. Chapter 9 focuses on neural modeling of complex systems. Chapter 10 discusses neural systems for computation and decision making. Chapter 11 details neural feature analysis and is of interest to scientists in neural character recognition. Chapter 12 deals with research in graph theory aspects of neural networks and its relation to cellular automata. Chapter 13 discusses neural networks pattern processing and logical reasoning capabilities. The last chapter is devoted to primacy and recency effects in back propagation learning.

This is an attempt to provide the readers with an in-depth presentation of a specific subject without limitation on the size, shape, and content of the presented work. We adhere to all the valuable principles of integrity in research while pursuing this endeavor. This series is the result of the hard work of about

more than 50 reviewers who have done a great deal of work for which I am grateful for their frank and valuable suggestions and recommendations. I would like to thank the president of Ablex, Mr. Walter Johnson, who trusted my judgment to start a series in neural networks back in the late 1980s; Ms. Carol Davidson for her invaluable advice every step of the way; and Ms. Roxanne Guidice for her constant work on the production of the series. Also, I would like to thank Ms. Sylvia C. Neuman for keeping my correspondence up to date, my wife for her continuous encouragement, and my children for their understanding of the fact that my work has as much importance as their computer games. Last but not least I am grateful to all the authors for their valuable contributions to this volume and the series. The third volume is also available now, and Volumes 4 to 10 will be available in the near future.

Omid M. Omidvar, Ph.D.
Series Editor
Computer Science Department
University of the District of Columbia

Table of Contents

PREFACE	vii
1 PROSPECTS FOR CLASSIFYING COMPLEX IMAGERY USING A SELF-ORGANIZING NEURAL NETWORK <i>MURALI M. MENON and KARL G. HEINEMANN</i>	1
2 CONSTRAINT SATISFACTION NETWORKS FOR VISION <i>RAKESH MOHAN</i>	25
3 HIGHER-ORDER NEURAL NETWORKS FOR POSITION, SCALE, AND ROTATION INVARIANT PATTERN RECOGNITION <i>MAX B. REID and LILLY SPIRKOVSKA</i>	69
4 OPTICAL IMPLEMENTATION OF CLOSEST VECTOR SELECTION IN NEURAL NETWORKS <i>STANLEY C. AHALT, STUART A. COLLINS, JR., ASHOK K. KRISHNAMURTHY, and DANIEL F. STEWART</i>	93
5 BIOLOGICAL BASIS FOR COMPACT ARTIFICIAL NEURAL NETWORKS <i>PHYLLIS L. ROSTYKUS and ARUN K. SOMANI</i>	109
6 A HYBRID TIME-SHIFTED NEURAL NETWORK FOR ANALYZING BIOLOGICAL NEURONAL SPIKE TRAINS <i>DAVID C. TAM</i>	129
7 A SELF-ORGANIZED NEURAL SYSTEM EXTENDING BACK PROPAGATION TO ACQUIRE AND UTILIZE A KNOWLEDGE STRUCTURE WITHOUT A TEACHER <i>BRUCE Y. WU</i>	147

vi TABLE OF CONTENTS

8	ANALYSIS AND SYNTHESIS OF ARTIFICIAL NEURAL NETWORKS FOR MODELING COMPLEX SYSTEMS	163
	<i>JUN WANG and B. MALAKOOTI</i>	
9	GRADIENT-TYPE NEURAL SYSTEMS FOR COMPUTATION AND DECISION MAKING	195
	<i>JACEK ZURADA</i>	
10	COMPUTATION ON GRAPHS: FROM NEURAL NETWORKS TO CELLULAR AUTOMATA	229
	<i>MAX GARZON and STAN FRANKLIN</i>	
11	NON-GAUSSIAN FEATURE ANALYSES USING A NEURAL NETWORK	253
	<i>WEI GONG, HUNG-CHUN YAU, and MICHAEL T. MANRY</i>	
12	PRIMACY AND RECENCY EFFECTS IN BACK PROPAGATION LEARNING	271
	<i>SHELLY D.D. GOGGIN, KRISTINA M. JOHNSON, and KARL GUSTAFSON</i>	
13	NEURAL-LOGIC NETWORKS	299
	<i>S.C. CHAN, L.S. HSU, K.F. LOE, and H.H. TEH</i>	
	AUTHOR INDEX	337
	SUBJECT INDEX	342

Prospects for Classifying Complex Imagery Using a Self-Organizing Neural Network

Murali M. Menon

Karl G. Heinemann

Massachusetts Institute of Technology Lincoln Laboratory
Lexington, MA

1. INTRODUCTION

The Neocognitron of Fukushima [1] is a massively parallel multilevel neural network system which performs visual pattern recognition. Its architecture models the anatomy of the human retina in a qualitative way. This system also resembles the Adaptive Resonance model of Carpenter and Grossberg [2] in that it is self organizing and operates without a "teacher." The Neocognitron has a demonstrated capability to discriminate alphabetical characters stored in a matrix of 16×16 pixels. Performance on handwritten characters in a 19×19 matrix was demonstrated by Fukushima [3]. A more recent study by Stoner and Schilke [4] has confirmed the model's ability to classify dot-matrix characters. While many accurate character recognition algorithms already exist, the Neocognitron is noteworthy because it handles positional shifts and moderate deformations in the shapes of input characters. These properties suggest that Fukushima's model might be very useful in solving more demanding machine vision problems. Work at the Massachusetts Institute of Technology Lincoln Laboratory has produced a simulation of the Neocognitron on a serial machine. This program has operated successfully on wire-frame images embedded in a matrix of 128×128 pixels. The model was able to classify images by extracting features from the input images and retaining only those whose response was above the average. Results from different Neocognitron systems showed that its shift tolerance depends on the number of levels used. A four-level system was unable to classify patterns uniquely and tolerate shifts with an input plane of 128×128 pixels. However, a single-level version was found that did classify properly and provide shift invariance at the same time. The shift-tolerance property can be exploited to cope with other kinds of variation by submitting appropriate transforms of the imagery as

input. This approach has been investigated by applying a polar transform to achieve automatic recognition of rotated images.

2. MODEL DESCRIPTION

The Neocognitron is a structured network of analog processing units which receive and transmit zero or positive valued analog signals. This network contains four distinct kinds of computational elements called S-cells, C-cells, V_s -cells, and V_c -cells. Each class of processor is defined by the types of cells which provide its input and a specific mathematical operation which determines the strength of its output.

The output from an individual processor generates input signals for certain other nodes after passing through a set of weighted connections. Each of these communication channels multiplies the transmitting unit's output by a specific connection strength (weight) and presents that product as an input for the receiving unit. The weight for a given connection can take on any positive value, so the effect of a specific unit's output may vary considerably from one node to the next.

Cells in the Neocognitron generally receive inputs from a number of different nodes and respond to the total received signal, but signals from the different types of processors are summed separately, because they affect the response in different ways. For a given unit, different patterns of output at the source nodes will produce varying levels of total input. This behavior arises, because the specific pattern of connection weights will amplify some of the individual source signals more than others. The total input will be particularly high when the source nodes send strong signals along paths with large weights, and it will decrease as strong signals are shifted to paths with smaller weights or the paths with large weights carry smaller signals. Thus, communication through the weighted connections enables the processors to detect differences in the pattern of transmitted signals. An analog transfer function then produces corresponding variations in the response level.

The Neocognitron's processing elements are organized into a hierarchical series of levels, where units of each type appear at every level. All these levels share a common structure wherein the different types of cells are segregated into distinct layers, and signals traverse these layers in the same order. A schematic representation of this architecture is shown in Figure 1.1, where an image comes in at the left and data flow to the right. A layer of V_s -cells and a layer of V_c -cells also exists at each level, but these have been omitted in order to simplify the diagram. Output from any given level serves as input for the next one, until a layer representing the final classification categories is reached.

The system is strictly a feed-forward network where signals originate at an initial input layer and propagate towards the final output layer. A hierarchical structure is produced by connecting the cells in a "fan-out" pattern, so that the

LEVELS AND LAYERS IN THE NEOCOGNITRON

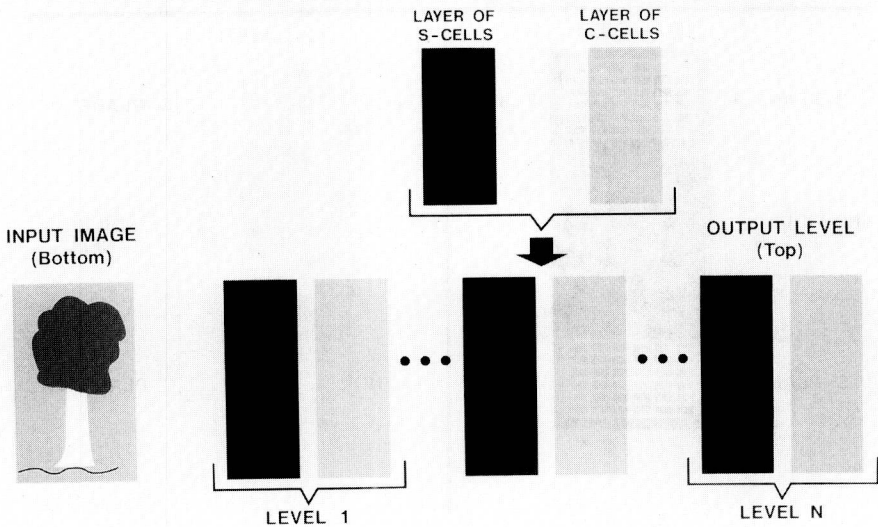


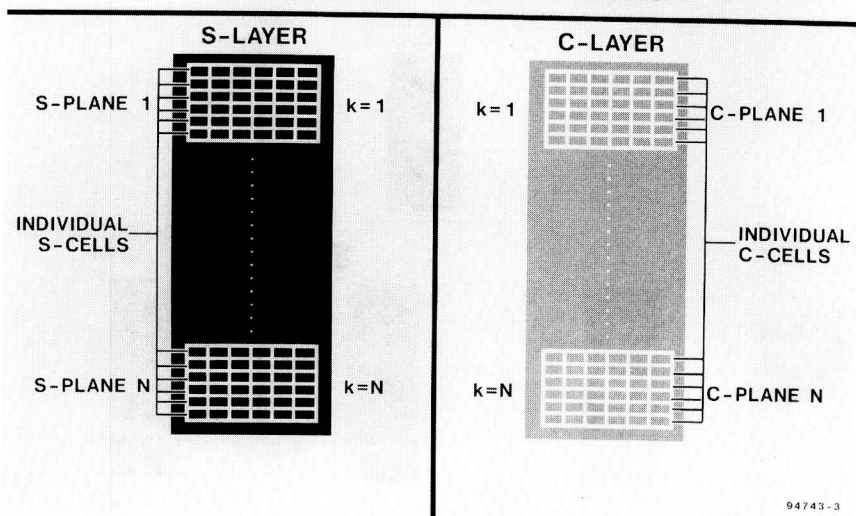
Figure 1.1. Multilevel feed-forward architecture of the Neocognitron.

number of units gradually decreases as signals propagate into deeper levels of the system. Under this connection scheme, each unit receives input signals from specific small regions on the layers which immediately precede it. However, the number of indirect connections between a processor and more distant predecessors grows significantly as the number of intervening layers increases. For any particular cell, the complete set of input sources on an earlier layer will be referred to as the cell's "receptive field" on that layer. Since processors at deeper levels gain access to progressively larger portions of the input patterns, they can respond to progressively more complicated features, and simpler features will be detected over a progressively larger receptive field. The final output layer consists of cells whose receptive field covers the entire input layer. This hierarchical structure contributes to the Neocognitron's capacity for shift invariant pattern recognition.

In order to completely explain the property of shift invariance, one must consider the structure of an individual level. The S-cells and the C-cells on any given level are organized into a number of subgroups which will be called "S-planes" or "C-planes" according to the type of processor which is involved. V_s -cells and V_c -cells also are grouped into planes, but there is only one V_s -plane and one V_c -plane on any level. These cell planes are treated as two-dimensional matrices where the location of an individual element is specified by a pair of column and row coordinates.

The relationships between planes, layers, and individual cells are illustrated in Figure 1.2. All the elements in a given plane share a single pattern of connection

LAYERS AND PLANES



94743-3

Figure 1.2. Detailed organization of the Neocognitron.

weights on their input channels. Consequently, a specific pattern of transmitted signals will elicit the same response from any element which observes that pattern exactly. While the input field of an individual processor covers only a small portion of the source layer, the fields of adjacent cells can be positioned in a way which insures that the entire source layer is covered. If the number of cells in the plane, the size of their input fields, and the offset between these fields are correctly matched, one can guarantee that some cell will show the optimum response when a specific pattern appears anywhere in the source layer. Hence, the behavior of these cell planes provides a massively parallel technique for shift invariant feature detection. This architectural feature is the fundamental mechanism responsible for the Neocognitron's tolerance of positional shifts.

The Neocognitron acquires its ability to classify patterns because each level contains a number of separate S-planes and C-planes. These two structures must always be paired with one another, so a given level has the same number of each type. However, the number of paired cell planes can vary from one level to the next. Each of the S-planes has a distinct pattern of input connection weights, but the C-planes on any particular level share one pattern in common.

The weights which feed into the S-planes have a special role, because they change as the system learns. All the other connection weights are built into the design of a specific Neocognitron architecture, and they cannot be modified. As the system learns to discriminate between diverse input images, the S-planes become sensitive to different spatial arrangements of the source signals. However, units in any given plane will receive small input signals from almost any

pattern that happens to occur. Difficulties could arise if all these signals were allowed to propagate deeper into the system. Some of the very weak signals could be greatly amplified when they pass through connections with large weights, and the results might convey some very misleading information.

In order to avoid this problem, the Neocognitron incorporates mechanisms which suppress the transmission of insignificant input signals. Interactions between the different types of processors work in concert with their particular response functions to provide a form of adaptive filtering. This design prevents the S-cells and C-cells from responding unless the pattern dependent input signal exceeds an independent estimate of the "typical" incoming signal strength.

A brief discussion of the different processors' actual operating characteristics and their interconnections will help to illustrate and clarify these general principles.

2.1. Cells in the S-layer and the V_c -layer

The S-cells in a given level obtain information about the previous one through two separate input mechanisms. Units in the first S-layer respond to the initial input signals, while those on subsequent levels receive input from C-planes on the preceding level. Direct connections from C-cells to S-cells carry excitatory signals which act to increase the S-cell's output. The S-cells also receive an inhibitory input which reduces the output signal through a shunting effect. This inhibitory signal ultimately comes from the same C-cells which produce the excitatory ones, but a layer of V_c -cells intervenes to perform some additional processing.

Any given level contains a number of S-planes and a single V_c -plane which all share the same geometric structure. The units at a given position in any of these planes share the same input fields, which extend over a specific set of adjacent coordinates in the preceding C-planes. Figure 1.3 illustrates the configuration of direct connections going from a C-layer to a particular S-plane. As a result of this connection scheme, S-cells and V_c -cells receive input from small regions on all of the C-planes. This arrangement enables the S-cells to recognize groupings of features that might have been detected in earlier stages of processing. If k refers to the k th S-plane in level l and n refers to a specific position in that S-plane, the response of the corresponding S-cell is given by:

$$U_{S_l}(k_l, n) = r_l \times f(A) \quad (2.1)$$

where:

$$A = \left[\frac{1 + \sum_{k_{l-1}=1}^{K_{l-1}} \sum_{v \in S_l} a_l(k_{l-1}, v, k_l) \times U_{C_{l-1}}(k_{l-1}, n + v)}{1 + \frac{r_l}{l + r_l} \times b_l(k_l) \times V_{C_l}(n)} - 1 \right]. \quad (2.2)$$

RECEPTIVE FIELDS OF S-CELLS AND C-CELLS

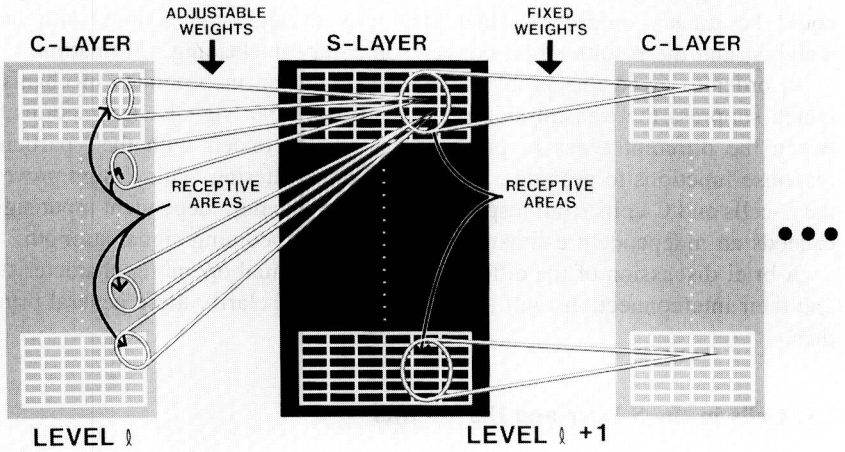


Figure 1.3. Interconnection architecture of the Neocognitron.

Expression $U_{C_{l-1}}(k_{l-1}, n + v)$ represents an excitatory signal coming from the unit at position $n + v$ on C-plane k_{l-1} , and $V_{C_l}(n)$ represents the inhibitory input. In Equation 2.2, $a_l(k_{l-1}, v, k_l)$ are the connection weights for excitatory input, and v designates a relative position inside the region of input. Positional variations of the weight values give rise to the “weight patterns” which were discussed earlier. These weight distributions differ from one source plane (k_{l-1}) to the next, so that the S-cells can recognize combinations of different source patterns. Note that $a_l(k_{l-1}, v, k_l)$ does not depend on the S-cell position, n , because all members of a given plane k_l have the same distribution of input weights. In the Equation 2.2, the inner sum computes the total excitatory input from a specific C-plane, and the outer sum adds together the contributions from different planes.

The inhibitory effect works through this expression’s denominator, where the connection weight $b_l(k_l)$ multiplies the output from a single node on the V_c plane. This inhibitory signal comes from the V_c cell at location n , which corresponds to the S-cell’s position in its own plane. The V_c -cell at those coordinates receives input from the same C-cells which are exciting the S-cell, and it responds by computing a weighted root-mean-square:

$$V_{C_l}(n) = \sqrt{\sum_{k_{l-1}=1}^{K_{l-1}} \sum_{v \in S_l} c_l(v) \times U_{C_{l-1}}(k_{l-1}, n + v)} \quad (2.3)$$

where $c_l(v)$ represents the input connection strength for a particular position v in that cell's input field. These weights can follow any distribution which decreases monotonically as the magnitude of v increases, and they must be normalized so that their sum is exactly equal to unity, that is:

$$\sum_{k_{l-1}=1}^{K_{l-1}} \sum_{v \in S_l} c_l(v) = 1. \quad (2.4)$$

In the present study, the $c_l(v)$ were defined by a decaying exponential distribution:

$$c_l(v) = \frac{1}{C(l)} \alpha_l^{r'_l(v)} \quad (2.5)$$

where $r'_l(v)$ is the normalized distance between location v and the center of the input region ($0 \leq r' \leq 1$). The parameter α_l is a small constant ($\alpha_l < 1$) which determines how quickly these weights fall off as $r'_l(v)$ increases. Consequently, weights at the edge of S_l ($r' = 1$) are equal to a fraction of α_l of the value at the center ($r' = 0$). The expression $C(l)$ is a normalizing constant:

$$C(l) = \sum_{k_{l-1}=1}^{K_{l-1}} \sum_{v \in S_l} \alpha_l^{r'_l(v)} \quad (2.6)$$

which insures that Equation 2.4 will be satisfied. All the V_c -cells in a given plane (and level) use the same pattern of input connections, but α_l is free to assume a different value for each level.

The weighted root-mean-square signal from a V_c -cell propagates to all S-cells at the same coordinates n . However, the weights on these connections, $b_l(k_l)$, are all independent, so the actual inhibitory effect will differ from one S-plane to the next. In addition, the denominator of the S-cell response function contains a factor $r_l/1 + r_l$, which further modulates the inhibition. This factor can provide any degree of attenuation as the parameter r_l goes from 0 to ∞ , and it has great sensitivity at the low end of its dynamic range. Note that r_l also appears as a multiplicative factor in the S-cell response function. It is given this additional role to curb growth in the final output when high attenuation (low r_l) makes the inhibition ineffective. The values for r_l are set by the system designer, and the subscript indicates that these values can be different at each level. Hence, the action of these parameters enables the system designer to control the overall influence of the weighted root-mean-square input at each level of the system. In order to prevent division by zero when inhibitory input is totally absent, the

attenuated signal is incremented by one. This solution conveniently neutralizes the denominator when there is no inhibitory input.

As discussed previously, an S-cell's excitatory input measures the degree of similarity between a particular arrangement of source signals and a feature represented by the distribution of input weights a_i . The quotient in Equation 2.2 compares the actual excitatory input with some fraction of the weighted root-mean-square source signal. The resulting ratio is decreased by one to determine which of the two input signals is greater. A positive difference indicates that the excitatory signal is greater, because the previous ratio exceeded unity, and a negative difference indicates that the inhibitory signal was greater. The function "f" which operates on this result is the linear threshold function:

$$f(x) = \begin{cases} x(x \geq 0) \\ 0(x < 0), \end{cases} \quad (2.7)$$

Consequently, an S-cell responds only if the excitatory input exceeds the inhibitory input, and the transmitted signal is proportional to the relative difference. The double sum in the numerator is incremented by one to produce proper behavior (zero response) when excitatory input is absent.

The Neocognitron learns to discriminate between different patterns of input by updating the adjustable weights $a_i(k_{l-1}, v, k_l)$ and $b_i(k_l)$ in Equation 2.2. Weights for the excitatory connections (a_i) start off with small values that allow different S-planes to produce distinct responses to an arbitrary input pattern. The inhibitory weights (b_i) are set to zero initially. Increments for both types of weight are determined by finding those S-cells which show the greatest response with respect to a certain set of the others. These units are selected by imagining that all S-planes on a given level are stacked vertically.

Many overlapping columns are defined in this stack, where a column goes through the same set of spatial positions in each S-plane. The learning procedure examines each column and records the position and plane where the S-cell response is strongest. This analysis is carried out for all possible columns, so that the entire S-layer is considered. If two or more maxima occur in one S-plane, the strongest one of those is retained and the others are discarded. Hence, this selection procedure locates the strongest response in each S-plane, but the maximum for a given plane can be rejected if it is overshadowed by the output from a nearby cell in some other plane. If this procedure selects a representative for S-plane \hat{k}_l at position \hat{n} , then the input weights for that plane are reinforced according to the rules:

$$\Delta a_i(k_{l-1}, v, \hat{k}_l) = q_l \times c_{l-1}(v) \times U_{c_{l-1}}(k_{l-1}, \hat{n} + v) \quad (2.8)$$

$$\Delta b_i(\hat{k}_l) = q_l \times V_{c_{l-1}}(\hat{n}). \quad (2.9)$$

None of the weights are reinforced if all the columns produce the same response. The parameter q_l is a gain factor that controls the rate of learning at each level, and it usually becomes larger as one progresses to higher levels. Since the increment for a given excitatory weight is proportional to input from the C-cell, only those connections carrying strong input signals are substantially reinforced. Consequently, the most significant modifications occur for connections where the input and the output are both relatively strong. This behavior is similar to Hebbian learning without a decay term. Note that the Neocognitron could be operated in a supervised learning ("with a teacher") mode by specifying the plane k_l , and location \hat{n} to be used at each level for a given input image. Further refinements are possible by including decay terms in the learning rule, but only Equations (2.8) and (2.9) were implemented in the present work.

2.2. Cells in the C-layer and the V_s -layer

The interactions and operational characteristics of the C-cells and the V_s -cells function in a manner that is very similar to the subsystem of S-cells and V_c -cells. These processors also take a given collection of source features and perform a comparison of two metrics. The result again determines whether information about that feature set will be passed on to higher-level classifiers. Units in the C-layers and the V_s -layer have input fields on the preceding S-planes. The exact locations covered by a specific field are related to the position of the receiving unit, just as before. However, elements in a given C-plane receive excitatory inputs from only one particular S-plane, and each S-plane communicates with only one C-plane. This design principle is depicted in Figure 1.3, and it is responsible for the pairing of cell planes that was mentioned above. V_s -cells receive input from all the preceding S-planes and generate an inhibitory signal. The C-cells compare these excitatory and inhibitory inputs by applying the same shunting mechanism which an S-cell uses:

$$U_{C_l}(k_l, n) = g \left[\frac{1 + \sum_{v \in D_l} d_l(v) \times U_{S_l}(k_l, n + v)}{1 + V_{S_l}(n)} - 1 \right] \quad (2.10)$$

where D_l is the region of input on the S-layer, U_s is the S-cell output from position $n + v$ on S-plane k_l , and d_l is the input connection weight at relative position v in the region of input. This expression is quite similar to the S-cell response function given in Equations (2.1) and (2.2), but the excitatory component in the numerator includes contributions from only one S-plane, and the connection weights for inhibitory input are set to unity. The excitatory connection weights d_l have fixed values that are determined according to the same general principles used for the weights $c_l(v)$ in Equation 2.3. In practice, setting the d_l to be uniform across the receptive field has proven to be adequate.