Image Processing Using Pulse-Coupled Neural Networks

T. Lindblad J.M. Kinser

Image Processing Using Pulse-Coupled Neural Networks

Second, Revised Edition

With 140 Figures

Springer

Professor Dr. Thomas Lindblad
Royal Institute of Technology, KTH-Physics, AlbaNova
S-10691 Stockholm, Sweden
E-mail: Lindblad@particle.kth.se

Professor Dr. Jason M. Kinser
George Mason University
MSN 4E3, 10900 University Blvd., Manassas, VA 20110, USA, and
12230 Scones Hill Ct., Bristow VA, 20136, USA
E-mail: jkinser@gmu.edu

Library of Congress Control Number: 2005924953

ISBN-10 3-540-24218-X 2nd Edition, Springer Berlin Heidelberg New York
ISBN-13 978-3-540-24218-5 2nd Edition Springer Berlin Heidelberg New York
ISBN 3-540-76264-7 1st Edition, Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media.

springeronline.com

Printed in The Netherlands

Typesetting and prodcution: PTP-Berlin, Protago-TEX-Production GmbH, Berlin
Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN 10965221 57/3141/YU 5 4 3 2 1 0

Preface

It was stated in the preface to the first edition of this book that image processing by electronic means has been a very active field for decades. This is certainly still true and the goal has been, and still is, to have a machine perform the same image functions which humans do quite easily. In reaching this goal we have learnt about the human mechanisms and how to apply this knowledge to image processing problems. Although there is still a long way to go, we have learnt a lot during the last five or six years. This new information and some ideas based upon it has been added to the second edition of our book

The present edition includes the theory and application of two cortical models: the PCNN (pulse coupled neural network) and the ICM (intersecting cortical model). These models are based upon biological models of the visual cortex and it is prudent to review the algorithms that strongly influenced the development of the PCNN and ICM. The outline of the book is otherwise very much the same as in the first edition although several new application examples have been added.

In Chap. 7 a few of these applications will be reviewed including original ideas by co-workers and colleagues. Special thanks are due to Soonil D.D.V. Rughooputh, the dean of the Faculty of Science at the University of Mauritius Guisong, and Harry C.S. Rughooputh, the dean of the Faculty of Engineering at the University of Mauritius.

We should also like to acknowledge that Guisong Wang, a doctoral candidate in the School of Computational Sciences at GMU, made a significant contribution to Chap. 5.

We would also like to acknowledge the work of several diploma and Ph.D. students at KTH, in particular Jenny Atmer, Nils Zetterlund and Ulf Ekblad.

Stockholm and Manassas, *Thomas Lindblad*
April 2005 *Jason M. Kinser*

Preface to the First Edition

Image processing by electronic means has been a very active field for decades. The goal has been, and still is, to have a machine perform the same image functions which humans do quite easily. This goal is still far from being reached. So we must learn more about the human mechanisms and how to apply this knowledge to image processing problems. Traditionally, the activities in the brain are assumed to take place through the aggregate action of billions of simple processing elements referred to as neurons and connected by complex systems of synapses. Within the concepts of artificial neural networks, the neurons are generally simple devices performing summing, thresholding, etc. However, we show now that the biological neurons are fairly complex and perform much more sophisticated calculations than their artificial counterparts. The neurons are also fairly specialised and it is thought that there are several hundred types in the brain and messages travel from one neuron to another as pulses.

Recently, scientists have begun to understand the visual cortex of small mammals. This understanding has led to the creation of new algorithms that are achieving new levels of sophistication in electronic image processing. With the advent of such biologically inspired approaches, in particular with respect to neural networks, we have taken another step towards the aforementioned goals.

In our presentation of the visual cortical models we will use the term Pulse-Coupled Neural Network (PCNN). The PCNN is a neural network algorithm that produces a series of binary pulse images when stimulated with a grey scale or colour image. This network is different from what we generally mean by artificial neural networks in the sense that it does not train.

The goad for image processing is to eventually reach a decision on the content of that image. These decisions are generally easier to accomplish by examining the pulse output of the PCNN rather than the original image. Thus the PCNN becomes a very useful pre-processing tool. There exists, however, an argument that the PCNN is more than a pre-processor. It is possible that the PCNN also has self-organising abilities which make it possible to use the PCNN as an associative memory. This is unusual for an algorithm that does not train.

Finally, it should be noted that the PCNN is quite feasible to implement in hardware. Traditional neural networks have had a large fan-in and fanout. In other words, each neuron was connected to several other neurons. In electronics a different "wire" is needed to make each connection and large networks are quite difficult to build. The PCNN, on the other hand, has only local connections and in most cases these are always positive. This is quite plausible for electronic implementation.

The PCNN is quite powerful and we are just in the beginning to explore the possibilities. This text will review the theory and then explore its known image processing applications: segmentation, edge extraction, texture extraction, object identification, object isolation, motion processing, foveation, noise suppression and image fusion. This text will also introduce arguments to its ability to process logical arguments and its use as a synergetic computer. Hardware realisation of the PCNN will also be presented.

This text is intended for the individual who is familiar with image processing terms and has a basic understanding of previous image processing techniques. It does not require the reader to have an extensive background in these areas. Furthermore, the PCNN is not extremely complicated mathematically so it does not require extensive mathematical skills. However, the text will use Fourier image processing techniques and a working understanding of this field will be helpful in some areas.

The PCNN is fundamentally unique from many of the standard techniques being used today. Many techniques have the same basic mathematical foundation and the PCNN deviates from this path. It is an exciting field that shows tremendous promise.

Contents

1 Introduction and Theory

1.1 General Aspects

Humans have an outstanding ability to recognise, classify and discriminate objects with extreme ease. For example, if a person was in a large classroom and was asked to find the light switch it would not take more than a second or two. Even if the light switch was located in a different place than the human expected or it was shaped differently than the human expected it would not be difficult to find the switch. Humans also don't need to see hundreds of exemplars in order to identify similar objects. For example, a human needs to see only a few dogs and then he is able to recognise dogs even from species that he has not seen before. This recognition ability also holds true for animals, to a greater or lesser extent. A spider has no problem recognising a fly. Even a baby spider can do that. At this level we are talking about a few hundred to a thousand processing elements or neurons. Nevertheless the biological systems seem to do their job very well.

Computers, on the other hand, have a very difficult time with these tasks. Machines need a large amount of memory and significant speed to even come close to the processing time of a human. Furthermore, the software for such simple general tasks does not exist. There are special problems where the machine can perform specific functions well, but the machines do not perform general image processing and recognition tasks.

In the early days of electronic image processing, many thought that a single algorithm could be found to perform recognition. The most popular of these is Fourier processing. It, as well as many of its successors, has fallen short of emulating human vision. It has become obvious that the human uses many elegantly structured processes to achieve its image processing goals, and we are beginning to understand only a few of these.

One of the processes occurs in the visual cortex, which is the part of the brain that receives information from the eye. At this point in the system the eye has already processed and significantly changed the image. The visual cortex converts the resultant eye image into a stream of pulses. A synthetic model of this portion of the brain for small mammals has been developed and successfully applied to many image processing applications.

So then many questions are raised. How does it work? What does it do? How can it be applied? Does it gain us any advantage over current systems?

Can we implement it with today's hardware knowledge? This is what many scientists are working with today [2].

1.2 The State of Traditional Image Processing

Image processing has been a science for decades. Early excitement was created with the invention of the laser, which opened the door for optical Fourier image processing. Excitement was heightened further as the electronic computer became powerful enough and cheap enough to process images of significant dimension. Even though many scientists are working in this field, progress towards achieving recognition capabilities similar to humans has been very slow in coming.

Emulation of the visual cortex takes new steps forward for a couple of reasons. First, it directly emulates a portion of the brain, which we believe to be the most efficient image processor available. Second, is that mathematically it is fundamentally different than many such traditional algorithms being used today.

1.2.1 Generalisation *versus* Discrimination

There are many terms used in image processing which need to be clarified immediately. Image processing is a general term that covers many areas. Image processing includes morphology (changing the image into another image), filtering (removing or extracting portions of the image), recognition, and classification.

Filtering an image concerns the extraction of a certain portion of the image. These techniques may be used to find all of the edges, or find a particular object within the image, or to locate particular object. There are many ways of filtering an image of which a few will be discussed.

Recognition is concerned with the identification of a particular target within the image. Traditionally, a target is an object such as a dog, but targets can also be signal signatures such as a certain set of frequencies or a pattern. The example of recognising dogs is applicable here. Once a human has seen a few dogs he can then recognise most dogs.

Classification is slightly different that recognition. Classification also requires that a label be applied to the portion of the input. It is possible to recognise that a target exists but not be able to attach a specific label to it.

It should also be noted that there are two types of recognition and classification. These types are generalisation and discrimination. Generalisation is finding the similarities amongst the classes. For example, we can see an animal with four legs, a tail, fur, and the shape and style similar to those of the dogs we have seen, and can therefore recognise the animal as a dog. Discrimination requires knowledge of the differences. For example, this dog may have a short snout and a curly tail, which is quite different than most other dogs, and we therefore classify this dog as a pug.

1.2.2 "The World of Inner Products"

There are many methods that are used today in image processing. Some of the more popular techniques are frequency-based filters, neural networks, and wavelets. The fundamental computational engine in each of these is the inner product. For example, a Fourier filter produces the same result as a set of inner products for each of the possible positions that the target filter can be overlaid on the input image.

A neural network may consist of many neurons in several layers. However, the computation for each neuron is an inner product of the weights with the data. After the inner product computation the result is passed through a non-linear operation. Wavelets are a set of filters, which have unique properties when the results are considered collectively. Again the computation can be traced back to the inner product.

The inner product is a first order operation which is limited in the services it can provide. That is why algorithms such as filters and networks must use many inner products to provide meaningful results for higher order problems. The difficulty in solving a higher order problem with a set of inner products is that the number of inner products necessary is neither known nor easy to determine, and the role of each inner product is not easily identified. Some work towards solving these problems for binary systems have been proposed [8]. However, for the general case of analogue data the user must resort to using training algorithms (many of which require the user to predetermine the number of inner products and their relationship to each other). This training optimises the inner products towards a correct solution. This training may be very involved, tedious, computationally costly and provides no guarantee of a solution.

Most importantly is that the inner product is extremely limited in what it can do. This is a first order computation and can only extract one order of information from a data set. One well known problem is the XOR (exclusive OR) gate, which contains four, 2D inputs paired with 1D outputs, namely (00:0, 01:1, 10:1, 11:0). This system can not be mapped fully by a single inner product since it is a second order problem. Feedforward artificial neural networks, for example, require two layers of neurons to solve the XOR task.

Although inner products are extremely limited in what they can do, most of the image recognition engines rely heavily upon them. The mammalian system, however, uses a higher order system that is considerably more complicated and powerful.

1.2.3 The Mammalian Visual System

The mammalian visual system is considerably more elaborate than simply processing an input image with a set of inner products. Many operations are performed before decisions are reached as to the content of the image. Furthermore, neuro-science is not at all close to understanding all of the operations. This section will mention a few of the important operations to provide a glimpse of the complexity of the processes. It soon becomes clear that the mammalian system is far more complicated than the usual computer algorithms used in image recognition. It is almost silly to assume that such simple operations can match the performance of the biological system.

Of course, image input is performed through the eyes. Receptors within the retina at the back of the eye are not evenly distributed nor are they all sensitive to the same optical information. Some receptors are more sensitive to motion, colour, or intensity. Furthermore, the receptors are interconnected. When one receptor receives optical information it alters the behaviour of other surrounding receptors. A mathematical operation is thus performed on the image before it even leaves the eye.

The eye also receives feedback information. We humans do not stare at images, we foveate. Our centre of attention moves about portions of the image as we gather clues as to the content. Furthermore, feedback information also alters the output of the receptors.

After the image information leaves the eye it is received by the visual cortex. Here the information is further analysed by the brain. The investigations of the visual cortex of the cat [1] and the guinea pig [12] have been the foundation of the digital models used in this text. Although these models are a big step in emulating the mammalian visual system, they are still very simplified models of a very complicated system. Intensive research continues to understand fully the processing. However, much can still be implemented or applied already today.

1.2.4 Where Do We Go From Here?

The main point of this chapter is that current computer algorithms fail miserably in attempting to perform image recognition at the level of a human. The reason is obvious. The computer algorithms are incredibly simple compared to what we know of the biological systems. In order to advance the computer systems it is necessary to begin to emulate some of the biological systems.

One important step in this process is to emulate the processes of the visual cortex. These processes are becoming understood although there still exists significant debate on them. These processes are very powerful and can instantly lead to new tools to the image recognition field.

1.3 Visual Cortex Theory

In this text we will explore the theory and application of two cortical models: the PCNN (pulse coupled neural network) and the ICM (intersecting cortical model) [3, 4]. However, these models are based upon biological models of the visual cortex. Thus, it is prudent to review the algorithms that strongly influenced the development of the PCNN and ICM.

1.3.1 A Brief Overview of the Visual Cortex

While there are discussions as to the actual cortex mechanisms, the products of these discussions are quite useful and applicable to many fields. In other words, the algorithms being presented as cortical models are quite useful regardless of their accuracy in modelling the cortex. Following this brief introduction to the primate cortical system, the rest of this book will be concerned with applying cortical models and not with the actual mechanisms of the visual cortex.

In spite of its enormous complexity, two basic hierarchical pathways can model the visual cortex system: the pavocellular one and the mangnocellular one, processing (mainly) colour information and form/motion, respectively. Figure 1.1 shows a model of these two pathways. The retina has luminance and colour detectors which interpret images and pre-process them before conveying the information to visual cortex. The Lateral Geniculate Nucleus, LGN, separates the image into components that include luminance, contrast, frequency, etc. before information is sent to the visual cortex (labelled V, in Fig. 1.1).

The cortical visual areas are labelled V1 to V5 in Fig. 1.1. V1 represents the striate visual cortex and is believed to contain the most detailed and least processed image. Area V2 contains a visual map that is less detailed and pre-processed than area V1. Areas V3 to V5 can be viewed as speciality areas and process only selective information such as, colour/form, static form and motion, respectively.

Information between the areas flows in both directions, although only the feedforward signals are shown in Fig. 1.1. The processing area spanned by each neuron increases as you move to the right in Fig. 1.1, i.e. a single neuron in V3 processes a larger part of the input image than a single neuron in V1.

The re-entrant connections from the visual areas are not restricted to the areas that supply its input. It is suggested that this may resolve conflict between areas that have the same input but different capabilities.

Much is to be learnt from how the visual cortex processes information, adapts to both the actual and feedback information for intelligent processing. However, a 'smart sensor' will probably never look like the visual cortex system, but only use a few of its basic features.

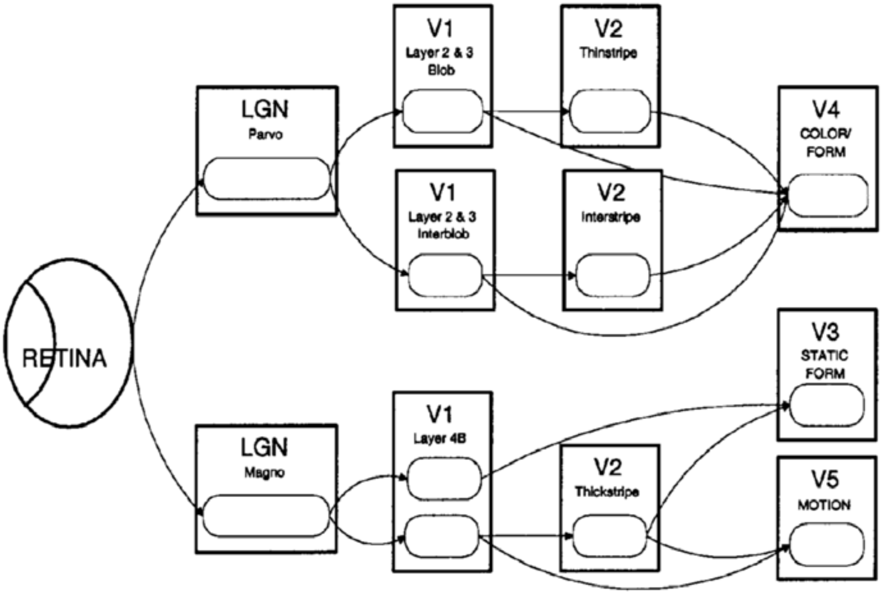

Fig. 1.1. A model of the visual system. The abbreviations are explained in the text. Only feedforward signals are shown

1.3.2 The Hodgkin–Huxley Model

Research into mammalian cortical models received its first major thrust about a half century ago with the work of Hodgkin and Huxley [6]. Their system described membrane potentials as

$$I = m^3 h G_{\mathrm{Na}}(E - E_{\mathrm{Na}}) + n^4 G_{\mathrm{K}}(E - E_{\mathrm{K}}) + G_{\mathrm{L}}(E - E_{\mathrm{L}})\,, \tag{1.1}$$

where I is the ionic current across the membrane, m is the probability that an open channel has been produced, G is conductance (for sodium, potassium, and leakage), E is the total potential and a subscripted E is the potential for the different constituents. The probability term was described by,

$$\frac{\mathrm{d}m}{\mathrm{d}t} = a_m(1 - m) - b_m m\,, \tag{1.2}$$

where a_m is the rate for a particle not opening a gate and b_m is the rate for activating a gate. Both a_m and b_m are dependent upon E and have different forms for sodium and potassium.

The importance to cortical modelling is that the neurons are now described as a differential equation. The current is dependent upon the rate changes of the different chemical elements. The dynamics of a neuron are now described as an oscillatory process.

1.3.3 The Fitzhugh–Nagumo Model

A mathematical advance published a few years later has become known as the Fitzhugh–Nagumo model [5,10] in which the neuron's behaviour is described as a van der Pol oscillator. This model is described in many forms but each form is essentially the same as it describes a coupled oscillator for each neuron. One example [9] describes the interaction of an excitation x and a recovery y,

$$\varepsilon \frac{\mathrm{d}x}{\mathrm{d}t} = -y - g(x) + I\,, \tag{1.3}$$

and

$$\frac{\mathrm{d}y}{\mathrm{d}t} = x - by\,, \tag{1.4}$$

where $g(x) = x(x-a)(x-1)$, $0 < a < 1$, I is the input current, and $\varepsilon \ll 1$. This coupled oscillator model will be the foundation of the many models that would follow.

These equations describe a simple coupled system and very simple simulations can present different characteristics of the system. By using ($\varepsilon = 0.3$, $a = 0.3$, $b = 0.3$, and $I = 1$) it is possible to get an oscillatory behaviour as shown in Fig. 1.2. By changing a parameter such as b it is possible to generate different types of behaviour such as steady state (Fig. 1.3 with $b = 0.6$).

The importance of the Fitzhugh–Nagumo system is that it describes the neurons in a manner that will be repeated in many different biological models. Each neuron is two coupled oscillators that are connected to other neurons.

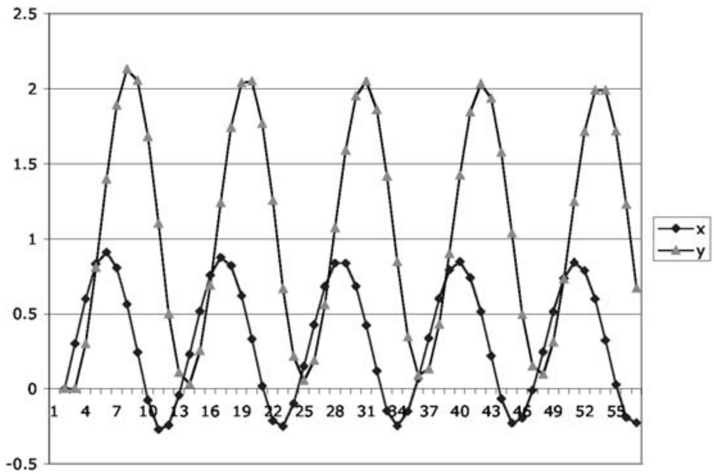

Fig. 1.2. An oscillatory system described through the Fitzhugh–Nagumo equations

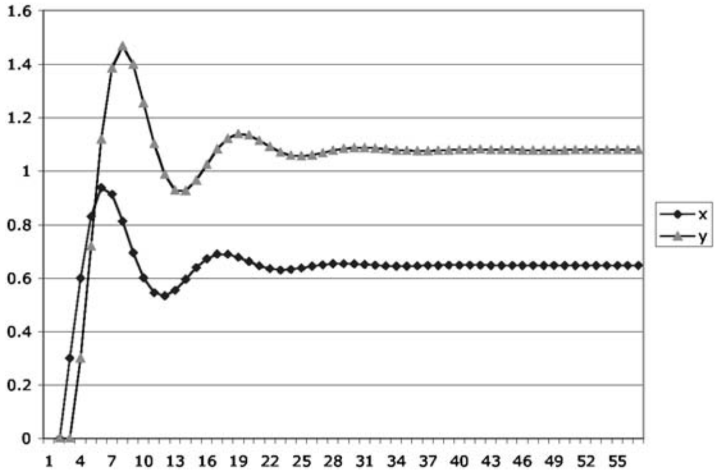

Fig. 1.3. A steady state system described through the Fitzhugh–Nagumo equations

1.3.4 The Eckhorn Model

Eckhorn [1] introduced a model of the cat visual cortex, and this is shown schematically in Fig. 1.4, and inter-neuron communication is shown in Fig. 1.5. The neuron contains two input compartments: the feeding and the linking. The feeding receives an external stimulus as well as local stimulus. The linking receives local stimulus. The feeding and the linking are combined in a second-order fashion to create the membrane voltage, U_m that is then compared to a local threshold, Θ.

The Eckhorn model is expressed by the following equations,

$$U_{m,k}(t) = F_k(t)[1 + L_k(t)] \tag{1.5}$$

$$F_k(t) = \sum_{i=1}^{N} \left[w_{ki}^{f} Y_i(t) + S_k(t) + N_k(t) \right] \otimes I\left(V^a, \tau^a, t\right) \tag{1.6}$$

$$L_k(t) = \sum_{i=1}^{N} \left[w_{ki}^{l} Y_i(t) + N_k(t) \right] \otimes I\left(V^l, \tau^l, t\right) \tag{1.7}$$

$$Y_k(t) = \begin{cases} 1 & \text{if } U_{m,k}(t) \geq \Theta_k(t) \\ 0 & \text{Otherwise} \end{cases} \tag{1.8}$$

where, in general

$$X(t) = Z(t) \otimes I(v, \tau, t) \tag{1.9}$$

is

$$X[n] = X[n-1]e^{-t/\tau} + V\, Z[n] \tag{1.10}$$