

# Chapter 1

## Biological Models

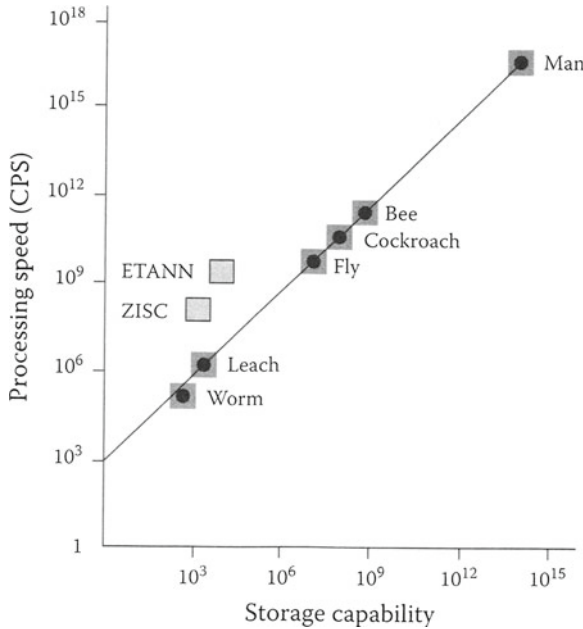
Humans have an outstanding ability to recognise, classify and discriminate objects with extreme ease. For example, if a person was in a large classroom and was asked to find the light switch it would not take more than a second or two. Even if the light switch was located in a different place than the person expected or it was shaped differently than expected it would not be difficult to find the switch. Humans also do not need to see hundreds of exemplars in order to identify similar objects. A person 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 as 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. Von Neumann 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 to the extent that animals and humans do.

Implementations of neural systems in silicon hardware have been tried by companies Intel and IBM. The Electrically Trainable Neural Network (ETANN) chip [43] from Intel had 128 neurons<sup>1</sup> and the first Zero Instruction Set Computer (ZISC36) chip [1] from IBM had 36 neurons. However, these are all “mathematical neurons” based on the back-propagation algorithm [14] and the radial basis function algorithms [73] and really not any implementation of biological systems. The ZISC36 chip could easily be put in parallel [66] to make use of several hundreds of neurons. It has also been further developed and is today available as C1MK with a thousand neurons [21]. This chip is between a fly and a worm with respect to the number of interconnections, although a bit “faster” than both. However, there is still a long way to go before reaching small mammals and humans as illustrated in

---

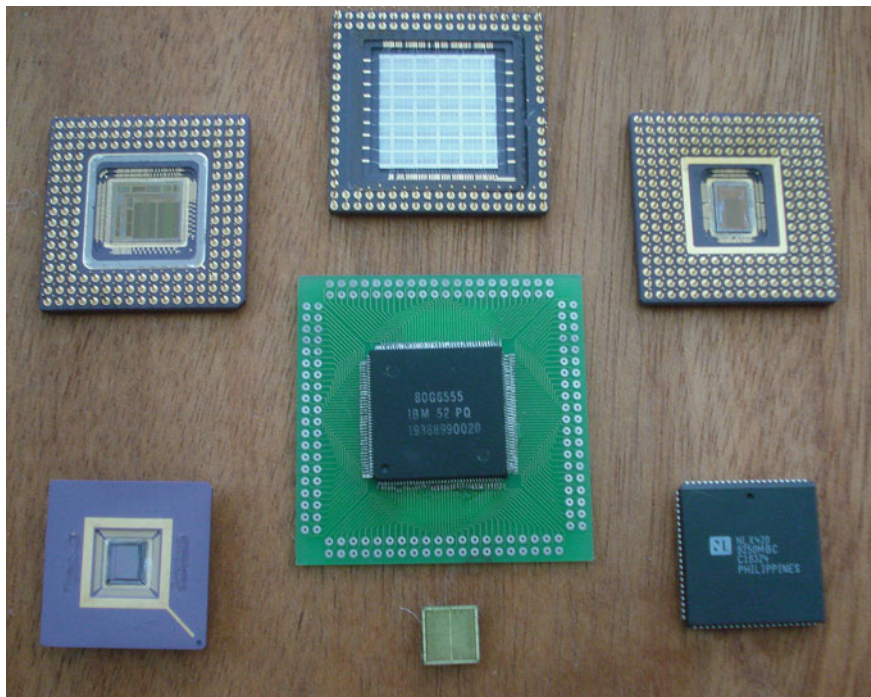
<sup>1</sup> There are several examples of how 128 neurons can be used with one example show in [65].



**Fig. 1.1** Comparing “brains” of some animals in some neural networks systems as discussed in the text

Fig. 1.1. Recent work at Manchester University is to develop SpiNNaker (Spiking Neural Network Architecture) which is a parallel computer specifically designed to model large scale spiking neural networks. The design will allow each computer to contain more than one million cores. Completion of the first machine is expected by the end of 2013. The ETANN, ZISC and a few other neural network chips are shown in Fig. 1.2.

It is often claimed that neuromorphic machines outperform von Neumann machines at certain environmental complexity (e.g., input combinatorics). There is a “break even point” and after this point the machine complexity (e.g., size, power, memory, gates, synapses) increases steeply for von Neumann computers but not so much for the neural architectures. However, there are still many orders of magnitude of complexity before one reaches the “human” level of performance. Besides the above problem of the number of neurons, there are at least two other fundamental items to consider: the neurons designed for specific tasks and the intelligent mammal sensors. This is particularly true for the visual system. The neurons get auxiliary information from adjacent neurons, the information is sent in separate paths in the mid-brain, backward signals are used to prioritise important information. Using computer language one would say that the feature extraction system, its redundancy and the parallel triggering system in the mid-brain ensure that important information reaches the visual cortex. At the same time, there is a tremendous reduction in data volume from perhaps initially  $10^8$  neurons and a bit-rate of 50 Mbits/s to approximate video speed when we become aware of what we are seeing.

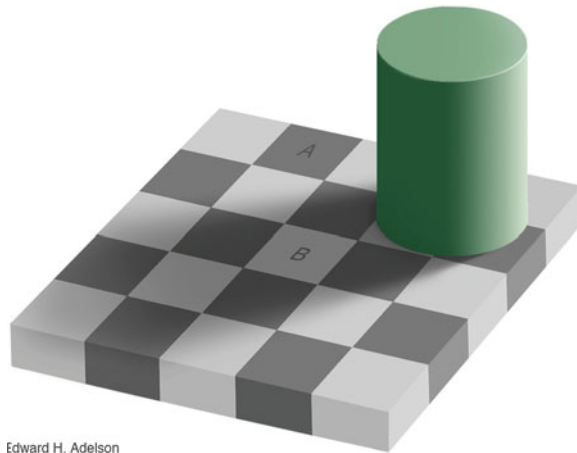


**Fig. 1.2** A few neural network chips. The ETANN (*top right*) and ZISC036 (*middle and lower middle*) are mentioned in the text and plotted in Fig. 1.1

## 1.1 Introduction

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. Synthetic models of this portion of the brain for small mammals have been developed and successfully applied to many image processing applications.

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. Neuro-science does not yet understand all of processes. However, sometimes the visual system is fooled, in particular where we expect colour and shades to follow some rules and patterns. There are very many examples of this, (e.g. the shadow of a cylinder on a board of chess shown in Fig. 1.3). Most people would say that the grey scale of square “A” is darker than that of square “B”, but even the simplest Paint program of a von Neumann computer would say that they are exactly the same.



Edward H. Adelson

**Fig. 1.3** The shadow of a cylinder on a checkerboard. Is square **a** really darker than square **b**?

Another example is the case of hiding a person from a searching adversary. Hiding in an open field may offer advantages over hiding in a ditch in that such a place is unexpected and away from visual edges which are natural attractors in human vision (see Sect. 5.4). “What you see is not always the truth.” An excellent example is the awareness test [4] in which the viewer is asked to count the number of passes of a ball between a set of players wearing a particular jersey. In this video a dancer dressed as a bear moves across the frame of view and most viewers completely miss his presence. 80% or more of our (university) students did not see the bear. Human processing of information is clearly not based solely on the visual input but also highly affected by other processes in the brain.

This chapter 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 investigation of the visual cortex of the cat [26] and the guinea pig [93] have been the foundation of the digital models

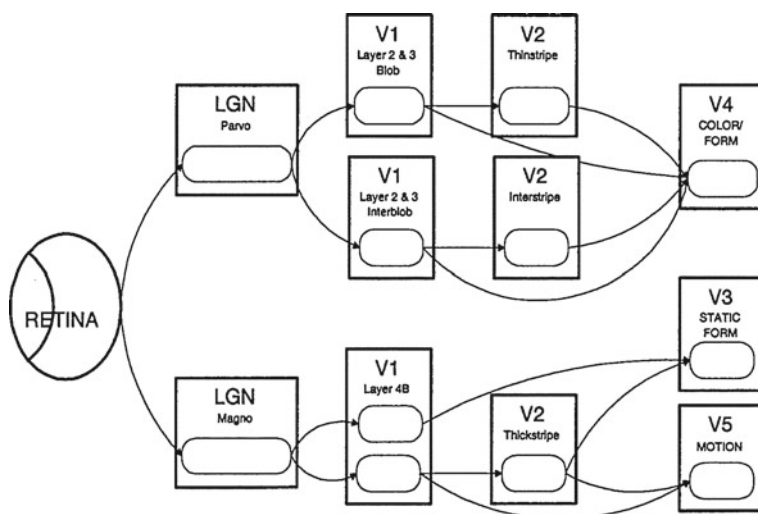
used in this text. Although these models are a big step in emulating the mammalian visual system, they are still very simplified models of a very complicated system. Intensive research continues to understand fully the processing. However, much can still be implemented or applied already today.

## 1.2 Biological Foundation

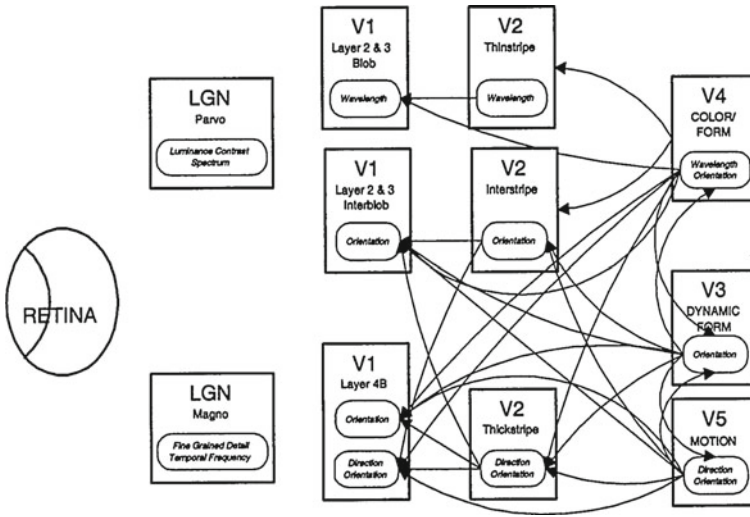
While there are discussions as to the actual cortex mechanisms, the products of these discussions are quite useful and applicable to many fields. In other words, the algorithms being presented as cortical models are quite useful regardless of their accuracy in modelling the cortex. Following this brief 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 parvocellular one and the magnocellular one, processing (mainly) colour information and form/motion, respectively. Figure 1.4 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 the 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. 1.4).

The cortical visual areas are labelled V1–V5 in Fig. 1.4. V1 represents the striate visual cortex and is believed to contain the most detailed and least processed image.



**Fig. 1.4** A model of the visual system. The abbreviations are explained in the text. Only feedforward signals are shown



**Fig. 1.5** Continuation of the model with the reverse connections displayed

Area V2 contains a visual map that is less detailed and pre-processed than area V1. Areas V3–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. 1.4. The processing area spanned by each neuron increases as you move to the right in Fig. 1.4, (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.

The backward connections shown in Fig. 1.5 are not restricted to the areas from which the feedforward signals came. Indeed, the most probable reason for this is that the signals are used to resolve conflicts between areas and to set priorities on the most important paths and processings [120, 121].

Much is to be learned 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.

### 1.3 Hodgkin-Huxley

Research into mammalian cortical models received its first major thrust about half a century ago with the work of Hodgkin and Huxley [39]. Their model is a mathematical (set of nonlinear ordinary differential equations) that describes how action

potentials in neurons are initiated and propagated. The basic components of the model are current sources, conductances and batteries. These are used in the model to represent the biophysical characteristic of cell membranes. Their system describes the membrane potentials ( $E$ ) using the equation:

$$I = m^3 h G_{Na} (E - E_{Na}) + n^4 + G_K (E - E_K) + G_L (E - E_L), \quad (1.1)$$

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,

$$\frac{dm}{dt} = a_m(1 - m) - b_m m, \quad (1.2)$$

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. Neuron activity is represented as oscillatory and dynamic processes. Although the original Hodgkin-Huxley model is regarded as one of the great achievements in biophysics, modern Hodgkin-Huxley-type models have been extended to include additional ion channel populations as well as highly complex dendrite and axon structures.

## 1.4 Fitzhugh-Nagumo

A mathematical advance published a few years later has become known as the Fitzhugh-Nagumo model [29, 75] in which the neuron 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 [61] describes the interaction of an excitation  $x$  and  $y$  using the expressions,

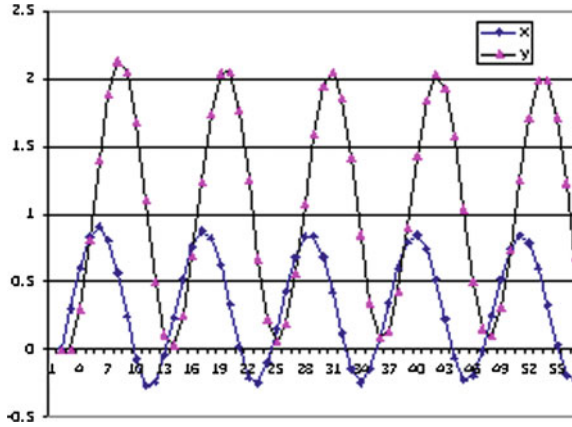
$$\epsilon \frac{dx}{dt} = -y - g(x) - I, \quad (1.3)$$

and

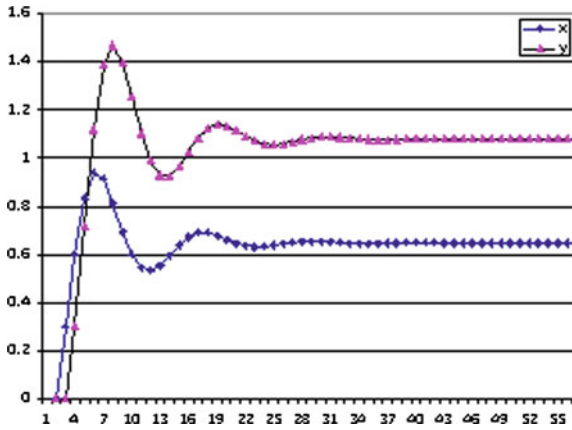
$$\frac{dy}{dt} = x - by, \quad (1.4)$$

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





**Fig. 1.6** An oscillatory system described through the Fitzhugh-Nagumo equations with the  $x$ -axis representing time and the  $y$ -axis representing excitations of  $x$  and  $y$



**Fig. 1.7** A steady state system described through the Fitzhugh-Nagumo equations with the  $x$ -axis representing time and the  $y$ -axis representing excitations of  $x$  and  $y$

These equations describe a simple coupled system and very simple simulations can present different characteristics of the system. By using ( $\epsilon = 0.3$ ,  $a = 0.3$ ,  $b = 0.3$ , and  $I = 1$ ) it is possible to get an oscillatory behaviour as shown in Fig. 1.6. By changing a parameter such as  $b$  it is possible to generate different types of behaviour. For example, setting  $b = 0.6$  will create a steady state which is shown in Fig. 1.7 where both  $x$  and  $y$  reach a constant value.

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.



## 1.5 Eckhorn Model

In 1989, Eckhorn et al. [26] introduced a neural model to emulate the cat visual cortex. Shortly, thereafter, Johnson [46] extrapolated the model to a digital form creating the Pulse-Coupled Neural Network (PCNN). This was the seminal work in transferring the cortical model into the field of digital image processing and recognition. Since then the PCNN has been expanded into a variety of applications in image processing, image segmentation, feature generation, face extraction, motion detection, region growing, noise reduction and image signatures just to name a few. Several of these applications will be discussed in the subsequent chapters.

One of the innovations that the Eckhorn model brought to image processing was a system that relied solely on local connections. Figure 1.8 depicts a neuron which contains two types of input (linking and feeding) which are combined to create the neuron's potential or membrane voltage  $U_m$ . This potential is then compared to a dynamic threshold  $\Theta$  to produce the neuron's output.

The Eckhorn model is expressed by the following equations,

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

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

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

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

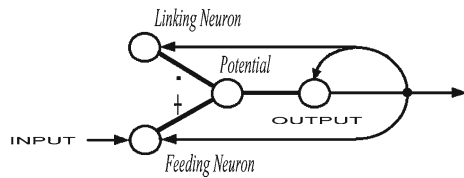
where, in general

$$X(t) = Z(t) \otimes I(v, \tau, t), \quad (1.9)$$

is

$$X[n] = X[n-1]e^{-t/\tau} + VZ[n]. \quad (1.10)$$

**Fig. 1.8** The Eckhorn-type neuron



Here  $N$  is the number of neurons,  $w$  is the synaptic weights,  $Y$  is the binary outputs, and  $S$  is the external stimulus. Typical value ranges are  $\tau_a = [10, 15]$ ,  $\tau_l = [0.1, 1.0]$ ,  $\tau_s = [5, 7]$ ,  $V_a = 0.5$ ,  $V_l = [5, 30]$ ,  $V_s = [50, 70]$ , and  $\Theta_o = [0.5, 1.8]$ . This model represents neural activity as coupled oscillators with two diffusion terms. Furthermore, there is now a second order (1.5) and a non-linear term (1.8) involved.

## 1.6 Rybak Model

Independently, Rybak [93] studied the visual cortex of the guinea pig and found similar neural interactions. While Rybak's equations differ from Eckhorn's the behaviour of the neurons is quite similar. Rybak's neuron has two compartments  $X$  and  $Z$ . These interact with the stimulus,  $S$ , as,

$$X_{ij}^S = F^S \otimes ||S_{i,j}||, \quad (1.11)$$

$$X_{ij}^I = F^I \otimes ||Z_{ij}||, \quad (1.12)$$

and,

$$Z_{ij} = f \left\{ \sum X_{ij}^S - \left( \frac{1}{\tau p + 1} \right) X_{ij}^I - h \right\}, \quad (1.13)$$

where  $F^S$  are local On-Centre/Off-Surround connections,  $F^I$  are local directional connections,  $\tau$  is the time constant and  $h$  is a global inhibitor. In the cortex there are several such networks which work on the input at differing resolutions and differing  $F^I$ . The non-linear threshold function is denoted by  $f\{\}$ .

The neural connections are quite different from the Eckhorn model in which the strength of the neural connections was inversely proportional to the physical distance between the neurons. In image processing terms this is a smoothing operation which would tend to blur the image. Rybak's model uses on-centre/off-centre connections in which there are positive connections neighbouring neurons and negative connections between neurons that are physically farther apart. In image processing terms this type of connections would enhance edges. This is very different from the Eckhorn model in which the smoothing operation would destroy edge information.

## 1.7 Parodi Model

There is still great disagreement as to the exact model of the visual cortex. Parodi [80] presented alternatives to the Eckhorn model. The arguments against the Eckhorn model included the lack of synchronisation of neural firings, the undesired similar

outputs for both moving and stationary targets and that neural modulations in the linking fields were measured considerably higher than the Eckhorn model allowed.

Parodi presented an alternative model, which included delays along the synaptic connections and would require that the neurons be occasionally reset *en masse*. Parodi's system followed these equations,

$$\frac{\partial V(x, y, t)}{\partial t} = -\frac{V(x, y, t)}{\tau} = D\nabla^2 V(x, y, t) + h(x, y, t), \quad (1.14)$$

where  $V_i$  is the potential for the  $i$ th neuron,  $D$  is the diffusion ( $D = a^2/CR_c$ ),  $R_c$  is the neural coupling resistance,  $t = CR_l$ ,  $R_l$  is the leakage resistance, and  $R_c^{-1} < R_l^{-1}$ , and

$$h_i(t) = \sum_j w_{ij} \delta(t - t^s - \tau_{ij}). \quad (1.15)$$

## 1.8 Summary

This book will focus on two digital models the PCNN and the ICM and their applications in the rest of the chapters. The PCNN is based on the Eckhorn model with the only modification being to unify the communication times between neurons to a discrete unit of time. The Eckhorn model is a set of coupled differential equations describing a multi-faceted neuron. This model has its foundation in earlier models of Hodgkin-Huxley and Fitzhugh-Nagumo. The Eckhorn model is not the only visual cortex model that exists and two others (Rybak and Parodi) have also been reviewed but this is not an exhaustive list of the proposed models.

The models of Eckhorn, Rybak and Parodi do have common mathematical foundations which are in congruence with the works from the previous decades. The ICM was developed in an attempt to capture common elements of these biological models and to reduce the components that are unique to each. Unlike the PCNN the ICM is a purely mathematical conjecture and not an attempt to replicate the visual cortex. Instead its purpose is to create a useful image processing engine. Both the PCNN and ICM will be reviewed in detail in Chap. 4. The ensuing chapters apply these two models to a variety of applications.

In several applications Python scripts are provided so that the reader can replicate the results and then pursue other avenues of related research. Before the PCNN and ICM are reviewed two chapters are presented to familiarize the reader with the Python language in support of the scripts that are provided throughout the rest of the chapters.

Image Processing using Pulse-Coupled Neural Networks

Applications in Python

Lindblad, Th.; Kinser, J.

2013, XXIV, 238 p., Hardcover

ISBN: 978-3-642-36876-9