Smart Pixel Optoelectronic Neural Networks

Keith J. Symington

MSc Optoelectronics and Laser Devices

Registration number: 9711710098481

Project Supervisor: M. R. Taghizadeh

External Supervisor: R. P. Webb, BT Research Labs.

Date: 10th April 1998

1   Abstract

Control tasks beyond the scope of serial processors can be performed using parallel systems such as neural networks. Current neural systems suffer serious interconnection problems on silicon but by employing spatial optics to interconnect neurons this limitation can be overcome. This report examines both the neural network architectures and smart pixel interconnection technologies that can be used to construct them. In conclusion, it examines a sample system and considers the potential applications of a smart pixel optical neural network.

2 Contents

3   Introduction

3.1   On Highly Parallel Systems

Traditional computer systems, as found in widespread use today, have always been sequential in nature with one instruction being rigidly performed after the other. As the power of these computers has increased by between 20% and 35% per year over the last few decades, it will not be long before the sequential computer reaches its physical computational limit – the speed of light. To sustain such a rate of growth new computing techniques will need to be developed and the only way

Performance v. Processors

Figure 1

Based on a diagram by Danny Hillis, architect of the connection machines.

forward will be the implementation of *parallel architectures*. Figure 1 shows the approaches taken by various computer systems towards performance "El Dorado" [5].

Turning to nature for inspiration, a good example of a parallel architecture may be found in the brain. In computer science terms a human brain can be considered as containing in the region of 10^{11} processors working in parallel. Each processor, or *neuron*, has its own very simple task to perform and is part of a *highly interconnected* system. It is this extreme interconnectivity within the network that is the attraction of a neural system since noise or a few errors will probably be inconsequential. An entire calculation on a sequential system can be ruined by a single bit error but on a neural system it simply results in graceful degradation. This fault tolerance creates a robust system which is the major attraction of *artificial neural networks* (ANNs).

Applying one of today's most powerful computers to a task such as image recognition begins to highlight the shortcomings of current computer technology: even though the cycle times of a silicon system are 10^6 times that of the brain, the brain is still extremely fast at a task like image recognition. Combining the speed of silicon and the tolerance of a neural system could lead to an evolution in computer architecture.

3.2   Report Outline

It is the intention of this report to examine the component parts of *optoelectronic neural networks* which are implemented using *smart pixels*. This report divides up the system into its two component parts, examines

them separately and then recombines them to consider the possible implementation advantages of a combined solution. This results in the following 3 chapters:

Artificial Neural Networks: An examination of artificial neural networks ranging from simple interconnection schemes to *multi-layer perceptrons* (MLPs) and *learning vector quantisation* (LVQ) including their advantages and disadvantages.

Smart Pixels: Analysis of what a smart pixel is and currently viable implementations. Information on the strengths and weaknesses of each type is also included here for both *monolithic* and *hybrid* approaches.

Optoelectronic Networks: A culmination of the knowledge from the previous two chapters with an examination of how smart pixels could revolutionise neural network implementations. A sample network and various applications are also briefly considered.

4   Artificial Neural Networks

4.1   Introduction

Artificial neural networks were designed to be an analogy (if a rather poor one) of the human brain, which is a massively parallel system, because the human brain is perhaps one of the most robust and fault tolerant pattern recognition systems known. However, ANNs have not quite reached the size and complexity of the human brain yet. Even so, ANNs have proved themselves to be competent at pattern recognition. What makes ANNs so interesting though is their power of discrimination – a characteristic not exhibited by *stochastic* techniques. In addition to this trait, they also have the power to learn.

This analogy has now developed beyond the point of a copy of the human brain into a field in its own right and today ANNs are generally referred to simply as *neural networks* (NN) [57].

4.2   The Node – An Artificial Neuron

A node, as shown in figure 2, is the basic building block of neural networks. The node was developed as an approximation of a neuron in nature.

A node takes n $(X_1,…,X_n)$ inputs and multiplies its strength by a scalar weight W_{nj} known as the synaptic weight. This allows a certain input to have more importance than others. These inputs are then summed by the node and put through a transfer function $f(x)$. The transfer function is node dependent and can be almost anything desired. The result is then returned on a single output Y. Figure 3 overleaf shows some common transfer functions.
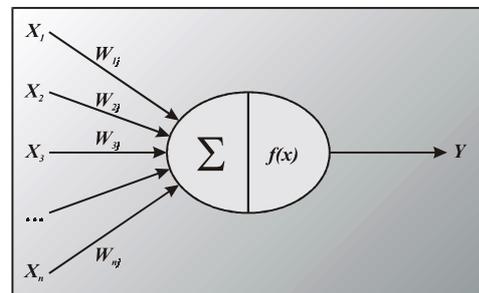
Figure 2

A node sums all incoming weights X_1 to X_n after multiplying them by the appropriate weight W_{1j} to W_{nj} and then determines its response based on the in-built function $f(x)$.

The node's output depends on its transfer function and sum of inputs. For example, the linear threshold transfer function would output either a -1 or 1 dependent on input. On the other hand, a linear decision neuron would smooth the response of the node providing not just the values of -1 and 1 but also all the values in between.

One of the most frequently used transfer functions is the *sigmoid* function, as shown in equation 1, where α determines how rapid the function's response is.

$$f(x) = \frac{1}{1 + e^{-\alpha x}}$$ **Equation 1**
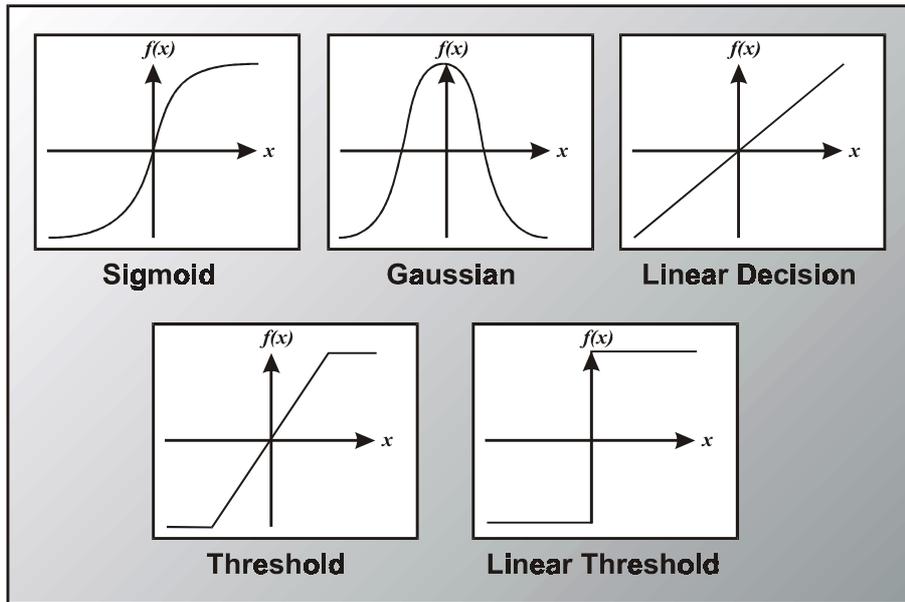
Figure 3

The above are all examples of threshold functions f(x). Please note that the minimum and maximum values shown here are considered to be normalised: i.e. -1 and 1 respectively.

The advantage of the sigmoid function is that it has smooth transition limits. Another similar function of note which is widely used is the tangent hyperbolic function.

The node is the basic building block of neural networks and by combining them correctly, altering their functions and weighting methods a variety of different systems can be built up which may be put to use in a great deal of applications.

4.3   Neural Network Types

There are two basic types of neural networks: ones which take binary input and ones which take continuous valued input.

Neural Network Name	Input Type	Training Method
Hopfield Net	Binary	Supervised
Hamming Net	Binary	Supervised
Carpenter/Grossberg Classifier	Binary	Unsupervised
Perceptron	Continuous	Supervised
Multi-Layer Perceptron	Continuous	Supervised
Kohonen Self-Organising Feature Maps	Continuous	Unsupervised

Table 1

These networks can again be sub-classified dependent on the procedure used to train them: *supervised* or *unsupervised* learning. Table 1 shows the properties of some common neural networks.

Unsupervised learning is where the nodes compete for data: no desired response is applied since there is no target set. Supervised learning is where there is a specific response which must be achieved. A network that has been trained in a supervised manner is good for modelling whereas a network trained in an unsupervised manner is good for exploratory work.

4.4　Simple Networks

This section mentions the basic structures found in all neural networks and how they are classified.

4.4.1　Outstar Networks

An *outstar* network (as shown in figure 4) learns and recalls a spatial input pattern which is impressed onto an array of neurons. To learn, X_0 is activated which turns on the command neuron 0. The pattern which is currently on X_1 to X_n is then learned and stored by the neurons 1 to n. To recall, the command neuron then reactivates the neurons proportionally to the previously learned input. Please see [24], [25] and [26] for further information.
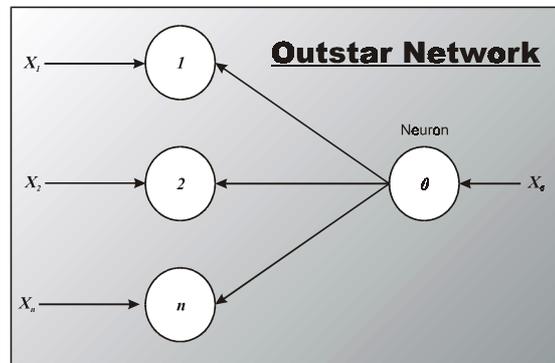
Figure 4

Outstar network for spatial pattern recognition.

4.4.2　Avalanche Networks

An *avalanche* network is designed to recall space-time patterns. It can be though of as a series of outstar networks (see 4.4.1) activated sequentially in time by only one controlling neuron. Read-out produces a short term memory pattern across the neurons proportional to the original pattern. Please see [24], [25] and [26] for further information.

4.4.3　Instar Networks

The *instar* network, as can be seen in figure 5, is used for recognising spatial input patterns. An input pattern is applied to neurons X_1 to X_n which signal the central neuron 0. Activation of the central neuron then causes it to go towards a steady state proportional to its inputs. To recognise, the pattern is replayed to the central neuron and if the inputs
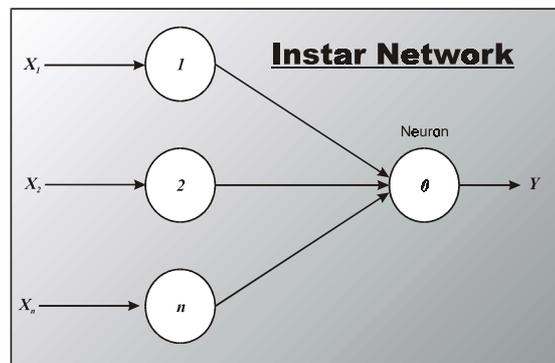
Figure 5

The Instar neural network: used for spatial pattern recognition.

are over a certain threshold then the neuron fires. Please see [24], [25] and [26] for further information.

4.5   The Perceptron

Rosenblatt [27] invented many variations of a simple network which he called the *perceptron*. The perceptron is simply a layer of input neurons connected directly to a layer of output neurons and is especially suited for simple pattern classification problems. Please examine figure 6.
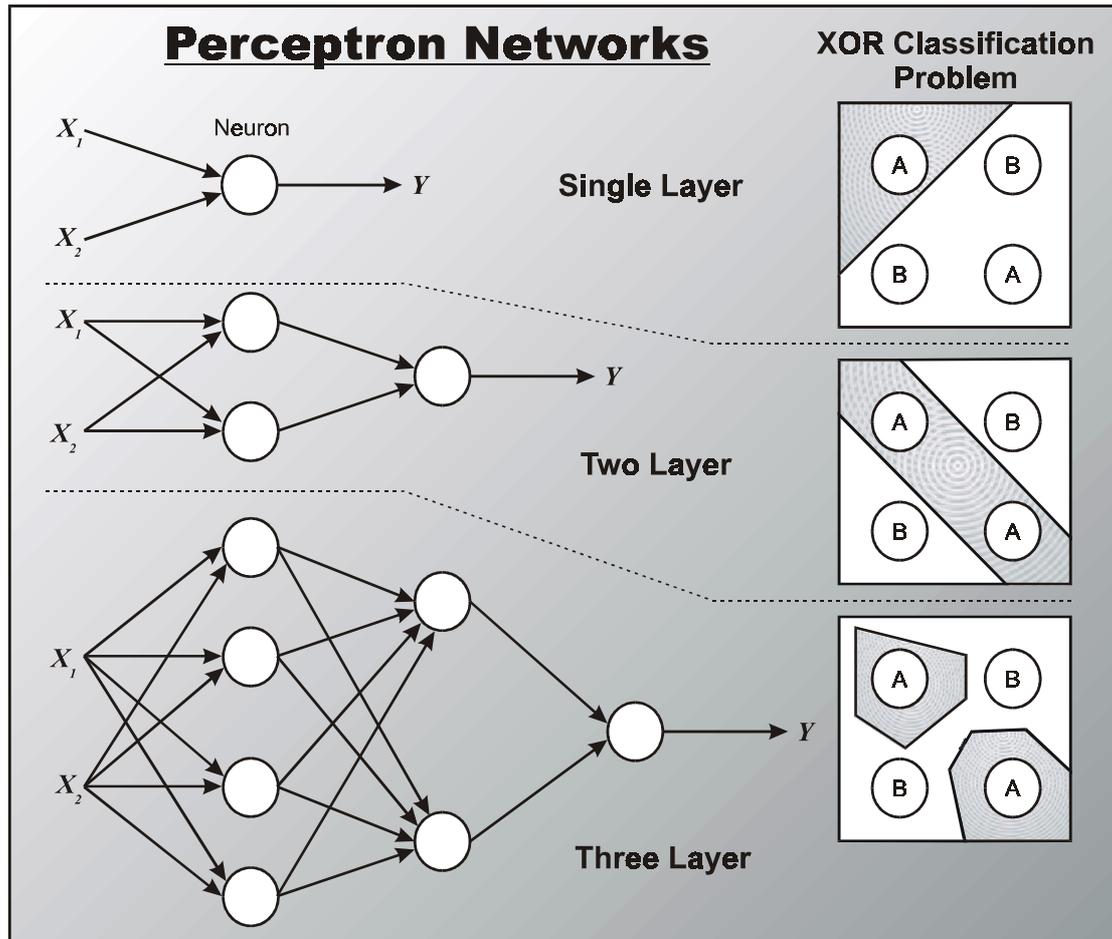

Figure 6

This figure shows the three major types of perceptron network and their response to the XOR classification problem (see text for details).

The XOR (exclusive or) problem used to illustrate network performance here is where two continuous valued inputs X_1 and X_2 are to be classified into a category A or B which is output on Y. In the diagrams, the actual output Y from the neuron is indicated by the two contrastingly shaded regions: the value of which is irrelevant. The target classifications of A and B are also shown here and, if the system is to function correctly, both A circles must be bounded by only one shade.

A *single layer perceptron* is the simplest of all and consists of only one neuron. Given the two inputs X_1 and X_2, it is only capable of classifying information with a single *hyperplane* line. This is not enough to successfully

implement a XOR architecture and as can be seen the neuron is only capable of learning to fire on one instance of A.

The next step up in complexity is the *two layer perceptron*. As can be seen, this network is capable of implementing a XOR operation because the system is capable of remembering more than just one hyperplane: it can model a convex open or closed region.

Finally we have the multi-layer perceptron (MLP) (more than two layers – three layer example given in this case). They differ from a normal perceptron in that they have *hidden layers* - so called because their output is not actually observable. Each layer that is added enables an extra, distinct, non-linear classification line. This allows a decision line of almost any shape to be created. It can be easily seen that this network can implement the XOR problem, if perhaps a bit to specifically. MLPs, unlike single layer perceptrons, are capable of modelling training sets that are not linearly separable. This is their major advantage over normal perceptrons.

An perceptron network has two modes of operation: Recall and Learning mode. Learning mode [2] trains the network to simulate appropriate data by altering each nodes' input weights until the system is modelled to within predefined limits: i.e. a target set. Recall mode simply puts values in and the perceptron will attempt to classify it.

Perceptron networks exhibit many advantages over normal methods:

- A perceptron can be set up to learn through experience from the input data itself.

- They can be applied to a classification, noise reduction or prediction problems.

- They can make a conclusion even if the input data is not well defined.

- Patterns can be extracted even though differences are very subtle.

- Decisions can be made even if the data is chaotic by normal mathematical standards.

For further information please see [26], [27], [28], [29] and [30].

4.6   Learning Vector Quantisation

Learning vector quantisation (LVQ) is a type of supervised training that teaches a *competitive layer* to automatically classify input vectors. Unfortunately, automatic classification means that vectors that are close together will probably be put into the same class. There is no method in a competitive layer to specify whether input vectors are of the same class or not. This is where an LVQ network comes in to play as it allows specification of the target vectors. Please see [31] for more information.

When applying the Kohonen rule ([28] and [29]) the competitive neuron whose weight forms the closest match wins and outputs a 1. However, in an LVQ network a target vector is found using the linear layer weights W2. This means that the Kohonen rule is used only to update the weights if the neuron

target is that target. If the winning neuron is found to be of the target class, then the Kohonen rule is still applied but with a sign change. This moves the offending neuron away from the vector. This results in competitive neurons moving towards vectors of their own class and then competing with each other to form sub-classes.

Figure 7 shows an example of an LVQ network:
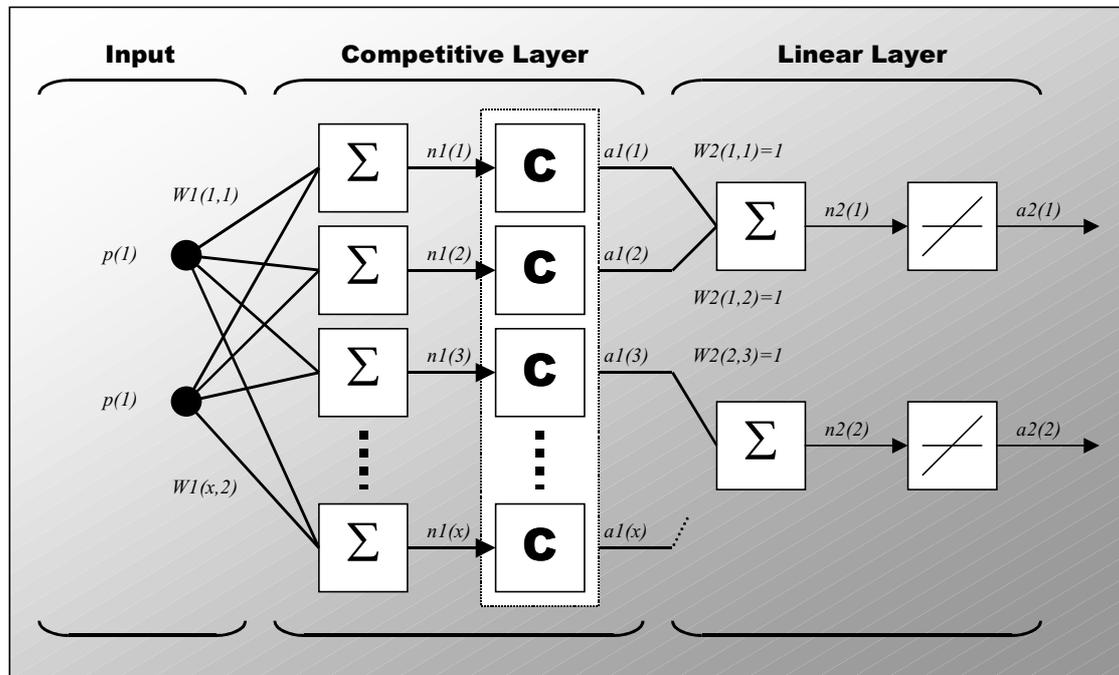
Figure 7

Sample Learning Vector Quantisation neural network.

4.7   Conclusion

The amount of theory in the field of neurocomputing is immense. This chapter gave an overview of the basics of the field, followed by a brief description of some common networks, but even so only scratched the surface. Section 4.3 mentions other interesting networks but does not describe them. For further reading on these (and other) neural networks please consult [24], [26], [28] and [29].

Even though large scale neural computers do not yet exist, the possible applications remain diverse from pattern recognition in both spatial and time domains to the implementation of digital logic with a degree of fault tolerance.

5   Smart Pixels

5.1   The Interconnection Problem

In 1971, Stone developed a sorting algorithm for parallel systems which has, to this day, not been surpassed as far as a minimal rate of growth of computational steps is concerned.
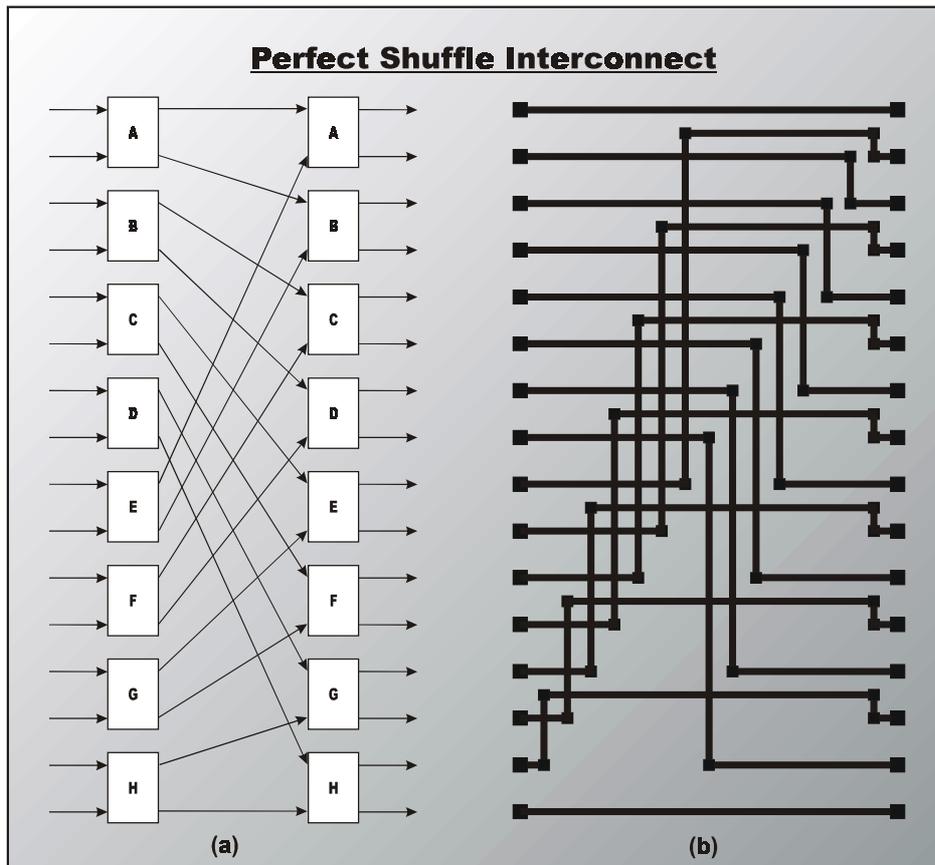
Figure 8

This network is capable of performing a perfect shuffle on 16 elements. Each node in (a) is lettered A to H and is interconnected as shown to two successor nodes. Part (b) shows this network implemented using a 2 metal layout process.

The algorithm is based on work done by Batcher in 1968 called the *bitonic merge-sort*, Stone adapting Batcher's work for a shuffle exchange network ([6], [7] and [23]). The interconnection methodology is generally known as "Stone's perfect shuffle" and can be seen in figure 8.

The major disadvantage is that to implement this system in any scale the amount of interconnection layout becomes prohibitive. It is this very interconnection problem that limits implementations of highly interconnected
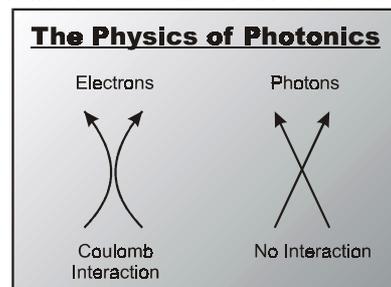
Figure 9

Photons have the advantage of being non-interacting in free space.

concurrent systems such as neural networks: however, this is where optical interconnection comes into play.

In a confined area, electronic connections need to be placed with sufficient spacing so that there is no electronic crosstalk due to capacitance. This leads to physical size limitations in any silicon processor. However, light has the property that it is non-interacting in free space and therefore the interconnects can effectively cross each other (figure 9 and [21]). Since the interconnects can then be more direct, the amount of routing is reduced and skew becomes less of a problem. Figure 10 shows a perfect shuffle network using an optical interconnection strategy. Here the data from each of the input nodes is converted into an optical information stream (e.g. modulated onto a laser beam) and routed through a *Computer Generated Hologram* (CGH). The CGH is computer controlled and can deflect each input beam to a pre-programmed target. The beams are then refocused down onto a detector array which converts the information back into electronic signals for the output nodes. It can be easily seen that this solution is also two dimensional: i.e. a 2D array of emitter and detector elements can be used.

There are various combinations of components that can be used in such a system [22]. For example:

Active Output Devices (require no external light source):

• *Vertical Cavity Surface Emitting Laser* (VCSEL) [8] and [9].

Passive Output Devices (require external light source):

• *Electro-Optic Modulator* (EOM) [10], [11], [12] and [13].

• *Self Electro-optic Effect Device* (SEED) [10] and [14].

Optical interconnection elements are mainly dominated by:

• *Computer Generated Holograms* (CGH) which are also known as a *Diffractive Optic Element*s (DOE) [15], [16], [17] and [18].

Receiver elements are usually some form of:

• *Photodiode* [10].

Figure 10

Example of optical communications channels interconnecting to create a perfect shuffle network.

• *Self Electro-optic Effect Device* (SEED) [10] and [19].

Unfortunately there are a few problems with constructing a practical detector/modulator system: although it is possible to implement a detector on silicon, it is not possible to implement a modulator (insufficient bandgap energy difference in silicon [20]). The choice of modulator is therefore very

relevant since it not only greatly influences the fabrication process but also what the system can be used for thereafter. Table 2 summarises the properties of four common output interfaces.

Technology	On-Chip Drive Power	Inter-Chip Connectivity	Fan-out
Modulators	Inefficient	Good	Poor
VCSELs	High power consumption	Moderate	Excellent
LEDs	Inefficient	Incompatible with holograms	Very Poor
μcavity LEDs	Good	Use fibre bundles	Poor

Table 2

Taken from reference number [21].

Unfortunately each technology has a drawback which makes it undesirable in certain situations. This problem resulted in the development of hybrid technologies which employ a Si detector and GaAs modulator arrays, *flip-chip bonded* (references [21] and [37]) on top of one another (see figure 11) after each has been separately fabricated. This combination of input, processing and output elements is generally known as a *smart pixel*.
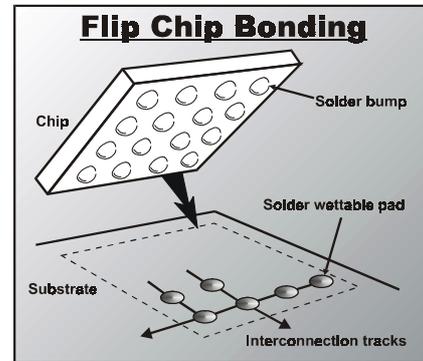
Figure 11

Example of flip-chip bonding as used to create hybrid smart pixels.

5.2   Smart Pixels

Smart pixel refers to picture elements in a display or photodetector array that may have some or all of the following attributes:

- Memory.

- Intra-pixel processing.

- Inter-pixel communication.

- Optical input from photodetectors or SEED elements.

- Optical output in the form of modulators or emitters.

Smart pixel arrays (SPAs) ([3] and [4]) have evolved from passively addressed *spatial light modulators* (SLMs) and displays. SLMs and displays with simple pixels suffer from low frame and refresh rates. This results in image degradation, ghosting, low resolution and limited pixel functionality. By incorporating active circuitry into each element (e.g. VCSEL emitters), smart pixel arrays overcome many of the problems associated with their predecessors.

Any implementation of a smart pixel VLSI technology must carefully consider the following six points [42]:

1. **Conversion efficiency.** Electro-optical, opto-electric and optical coupling efficiencies must be high.

2. **High speed operation.** Any figure in excess of 100Mhz is generally considered to be high speed.

3. **High device reliability.**

4. **High yield fabrication processes.**

5. **Able to handle complex logic functions.**

6. **Allow for logic with a low power∗delay product.**

There are currently two types of fabrication process for smart pixels: hybrid (figure 12) and monolithic (figure 13).

Hybrid approaches use two different fabricates which are then combined using one of three techniques: flip-chip bonding ([37], [38], [39] and figure 11), wafer bonding [40] or epitaxial lift-off [41].
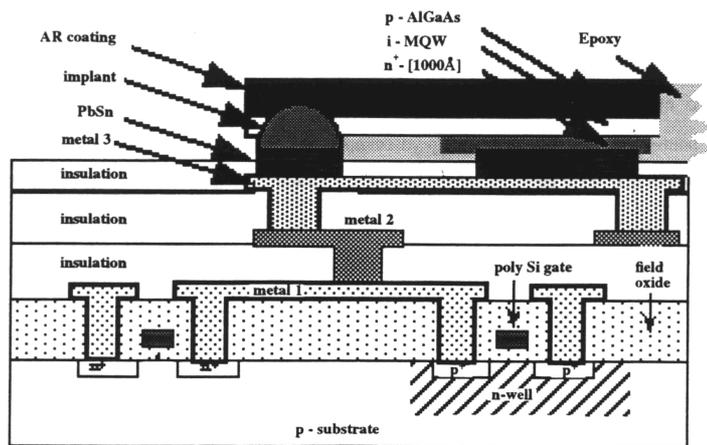
Figure 12

Cross section of a hybrid GaAs SEED/Si-CMOS smart pixel implementation

The monolithic approach uses a single custom grown wafer of III to IV materials ([32], [33], [34] and [35]) or an optical material re-grown on a processed wafer [36].

The next section examines a few promising implementations examining their individual advantages and disadvantages.
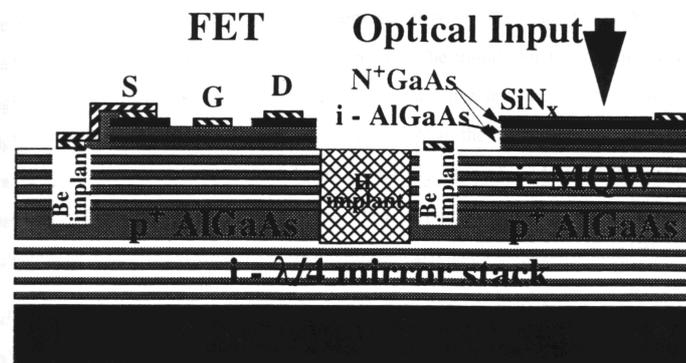
Figure 13

Cross section of a monolithic FET-SEED smart pixel implementation.

5.3   Smart Pixel Implementations

5.3.1   Photo-Thyristor

Photo-thyristor smart pixels are fabricated using a p-n-p-n doped structure and are so-called because their opto-electronic components exhibit thyristor like non-linearity. They are constructed by stacking a phototransistor on top of an optical output element such as LED or VCSEL and basically perform an optical dynamic memory operation. Unfortunately this device requires a reset

pulse (several nanoseconds) to turn off the output element which results in poor operating speeds.

Three examples are Double Heterostructure Optoelectronic Switches (DOES) [43], Light Amplifying Optical Switches (LAOS) [44] and Vertical to Surface Transmission Electro-Photonic (VSTEP) [45] technology. Table 3 shows the respective properties of the devices.

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
DOES (LED)	M	High	Low	Low	No	No	Low	No	High
VSTEP (LED)	M	High	Low	Low	No	No	Low	No	High
LAOS (LED)	M	High	Low	Low	No	No	Low	No	High
DOES (VCSEL)	M	High	High	High	No	No	Low	No	High
VSTEP (VCSEL)	M	High	High	High	No	No	Low	No	High

Table 3

5.3.2   Heterojunction Photo-Transistor – Vertical Cavity Surface Emitting Laser (HPT-VCSEL)

The HPT-VCSEL [46] is an evolution of the photo-thyristor (See 5.3.1). This implementation incorporates a *heterojunction phototransistor* (HPT) to the side of a vertical cavity surface emitting laser thus reducing optical feedback (i.e. no photo-thyristor effect).

This type system is however plagued by a need to dissipate excessive amounts of power. Table 4 shows the respective properties of this device.

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
HPT-VCSEL	M	High	High	High	No	No	Low	No	High

Table 4

5.3.3   Heterojunction Phototransistor-Modulator (HPT-MOD)

The HPT-MOD integrates heterojunction phototransistors with *multiple quantum well* (MQW) reflection modulators. One example implementation is the exciton absorption reflection switch (EARS) [47] which consists of a HPT and modulator which are vertically stacked and connected electrically in series. This arrangement allows the reflectivity of the modulator to be controlled by the optical intensity applied to the HPT.

This architecture has a few problems: firstly, heat extraction. Since both sides of the chip are used, it is not possible to attach an efficient heat sink. Secondly, EARS does not allow an electrical contact to be made to the base of the transistor which prevents the implementation of complex logic. Finally, the system suffers from Miller capacitance of the HPT. This imposes a relatively low modulation limit on the technology. Table 5 shows the characteristics of this device.

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
HPT-MOD (EARS)	M	High	High	High	No	No	Low	No	High

Table 5

5.3.4   GaAs-FET/LED

This technology has been demonstrated in two different implementations.

The first used a GaAs *complimentary heterostructure* FET (CHFET) where vertically emitting LEDs and ion-implanted photoconducting detectors were monolithically combined [48]. The system was limited to a maximum clock speed of 10MHz due to limited mobility of holes and high leakage currents in a p-type GaAs material.

The second integrated LEDs with n-channel MESFETs and p-n photodiodes [49]. The system had a severely limited switching speed (800kHz) due to the fact that large photodetectors were required to couple light from the LEDs of the previous stage.

Table 6 gives the characteristics of both of these devices.

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
MESFET/p-n/LED	M	High	Low	Low	No	No	Low	Yes	High

Table 6

5.3.5   MSM/FET/VCSEL

This is a recently developed technology which monolithically integrates VCSELs with metal-semiconductor-metal (MSM) photodetectors [50]. This implementation has been demonstrated working at above 100MHz, however further improvements are still required on FET transconductance and VCSEL wall-plug efficiency.

Table 7 overleaf shows the advantages and disadvantages of these devices.

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
MSM/FET/VCSEL	M	Low	High	High	Yes	No	Low	Yes	High

Table 7

5.3.6 Self-Electro-Optic Effect Device (SEED) Technologies

A self-electro-optic Effect Device (SEED) [51] is simply a MQW diode connected in series with another element and a voltage source, providing an *optically bistable circuit*. The MQW diode can function as either a detector or modulator.

There are five different types of SEED technologies:

R-SEED: Second element is a resistor.

D-SEED: Second element is a photodiode.

S-SEED: Second element is another SEED forming a Symmetric SEED.

M-SEED: Multi-state SEEDs have additional SEEDs connected in series.

L-SEED: Logic SEEDs are a series/parallel combination of SEED elements.

Many demonstrations of SEED implementations have been made including an optical logic processor, a cellular-logic image processor and switching systems for telecommunications.

The big advantages of SEEDs are high production yields and device reliability. Unfortunately they are also hindered by relatively high switching energies and a lack of true VLSI complex logic.

Table 8 shows the respective properties of the devices.

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
R-SEED	M	High	High	High	No	No	High	No	High
D-SEED	M	High	High	High	No	No	High	No	High
S-SEED	M	High	High	High	No	Yes	High	No	High
M-SEED	M	High	High	High	No	Yes	High	Yes	High
L-SEED	M	High	High	High	No	Yes	High	Yes	High

Table 8

5.3.7   GaAs FET/MQW Modulators

There are three of GaAs FET/MQW modulators, each of which is briefly discussed here.

To overcome the limitations of EARS, a technology was developed which uses *metal-semiconductor-metal* (MSM) photodetectors, MESFETs and MQW modulators [52]. An MSM photodetector as receiver instead of MQW diode allowed the system to use either positive or negative input logic. Unfortunately, low quantum efficiency is a trait of MSM detectors (~50%) and results in a reduced system responsivity.

A *monolithic optoelectronic transistor* (MOET) [53] is a combination of MQW modulators/detectors with MESFETs and *resonant tunnelling diodes* (RTD) (designed to be load elements for the photodiode). The problem with this technology is that the large photodiode used dominates the input capacitance limiting switching speeds to ~1μs.

Finally we have the *field-effect transistor - self electro-optic effect device* (FET-SEED): a technology derived from the S-SEED to improve switching speed and device functionality. The FET-SEED fabricate combines doped channel FETs with GaAs/AlGaAs MQW diodes. In this implementation the MQW diodes serve a dual purpose dependant upon applied bias voltage: they can either function as a high quantum efficiency photodiode or as an electro-optic modulator.

Table 9 shows the respective properties of these three devices:

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
MSM / MESFET / MOD	M	Low	High	High	Yes	No	High	Yes	High
MOET	M	High	High	High	No	No	High	Yes	High
FET-SEED	M	High	High	High	Yes	No	High	Yes	High

Table 9

5.3.8   Hybrid Flip-Chip Bonded InP-MQW-Modulator/Si-CMOS

Hybrid systems are the alternative to monolithic technologies where two different substrata are bonded on top of one another using a controlled volume of solder sandwiched between two wettable pads: one on each chip (see [38] and figure 11).

In this implementation one chip is silicon CMOS logic circuitry and the other an InP MQW modulator (inverted asymmetric Fabry-Perot) or p-i-n photodetectors. The MQW modulator is operated at 1.55μm since the substrate is transparent at this wavelength, allowing optical access once the chip has been flipped over.

The major problem with this technology is the flip-chip bonding. To connect both substrata they must be placed on top of one another and heated to reflow the solder bumps (which are normally 30μm in diameter) thus creating a connection. Such high temperatures can be detrimental to the chips and create a strained connection when attached due to thermal expansion coefficient mismatch. Note that a high chip yield has been demonstrated if larger solder bumps are used (>40μm in diameter).

Table 10 shows the relevant device characteristics.

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
InP-FPMOD / CMOS	H	High	High	High	No	Yes	Low	Yes	Low

Table 10

5.3.9 Hybrid Flip-Chip Bonded GaAs SEED/Si-CMOS

This hybrid approach addresses the limitations of monolithic FET-SEED technology (section 5.3.7) by flip-chip solder bump bonding GaAs MQW diodes to VLSI Si-CMOS. The GaAs substrate is removed by chemical etching to leave MQW diode islands behind (see figure 12 and [39]). Thermal expansion problems are avoided by simply using cold solder tacking with epoxy or low temperature thermal compression bonding.

Table 11 summarises implementation characteristics.

Technology	Monolithic or Hybrid	Opto-electric Conversion Efficiency	Electro-optic Conversion Efficiency	Optical Coupling Efficiency	High Speed Operation (>100 MHz)	Proven Reliability	Fabrication Yield	Handles Complex Logic	Power*Delay Product
GaAs-MOD / CMOS	H	High	High	High	Yes	Yes	High	Yes	Low

Table 11

5.4 Smart Pixel Research

Many manufacturers produce smart pixel technologies. This section simply lists, in the form of table 12 overleaf, a few interesting implementations which are broken down into *Input*, *Logic* and *Output* elements.

Researcher	Input	Logic	Output
AT & T	GaAs SEED	GaAs FET	GaAs-SEED Modulator
NTT, Opticomp	MSM GaAs	HEMT-MESFET	GaAs VCSEL
Zurich	GaAs	GaAs-MESFET	GaAs LED
AT & T	GaAs	Si-CMOS	GaAs Modulator
Heriot Watt	InGaAs	Si-CMOS	InGaAsP Modulator
Georgia Tech	ELO InGaAsP	Si-CMOS	ELO InGaAsP LED
Erlangen, Siemens, Colorado	Si PIN	Si-CMOS	GaAs VCSEL
Edinburgh, Colorado	Si	CMOS	Liquid Crystal Modulator
UC San Diego	Si	MOS	PLZT Modulator

Table 12

5.5   Conclusion

It is apparent from the information in this section that the GaAs MQW diode on Si CMOS hybrid technology is the only implementation to meet all the criterion for an efficient system. Using Si CMOS VLSI logic is a great advantage since it is a technology which is not only very mature but scalable and flexible.

Comparing monolithic to hybrid technologies, it can be seen that monolithic systems cannot compete against new hybrid technologies. These new technologies are more efficient and better developed in almost every aspect.

---

6   Optoelectronic Networks

6.1   Introduction

The main advantage of implementing an optoelectronic neural network is that interconnectivity is no longer a problem in complex systems: a problem that makes it impracticable to implement pure silicon neural systems. This section examines how smart pixels can taken advantage of to build such systems.

6.2   Fan-In and Fan-Out

Fan-in and *fan-out* are both major problems to implement in silicon but are a crucial part of neural networks.

Firstly let us consider fan-in by looking at figure 14. Referring back to the instar network in 4.4.3, we see that this network involves j inputs being collected by a single neuron in a parallel array of i networks. If we take the output from each neuron in P_{ij}, we can see that the input to P_i is a linear summation of all j outputs respectively. Each pixel element need only then be programmed to respond as though it were a neuron.
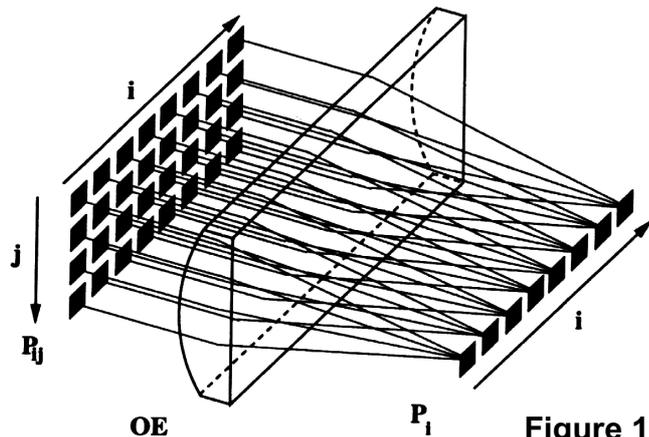
Fan-in example using a cylindrical lens OE.

Figure 14

Fan-out, as shown in figure 15, is used in structures such as the outstar network in 4.4.1. We see that this network involves a single input from i parallel networks in P_i being distributed to j multiple target neurons in P_{ij}. The diffractive optic element (DOE) is used here to split the signal evenly among the targets but this results in the signal power being divided amongst the target pixel elements which obviously results in an increased noise level in the system. However, as long as the noise
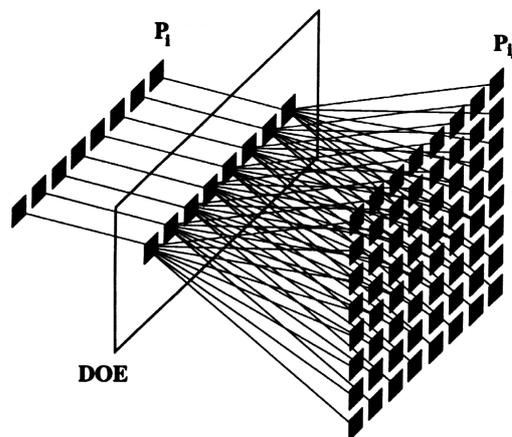

Fan-out example using a diffractive optic element (DOE).

Figure 15

[54] remains within manageable limits the robust nature of neural networks should easily be able to compensate for this problem.

Please refer to [22] for further information on fan-in and fan-out.

6.3 Optoelectronic Neural Network

The combination of fan-in and fan-out in an optical neural network allows large networks to be implemented. Please examine figure 16.
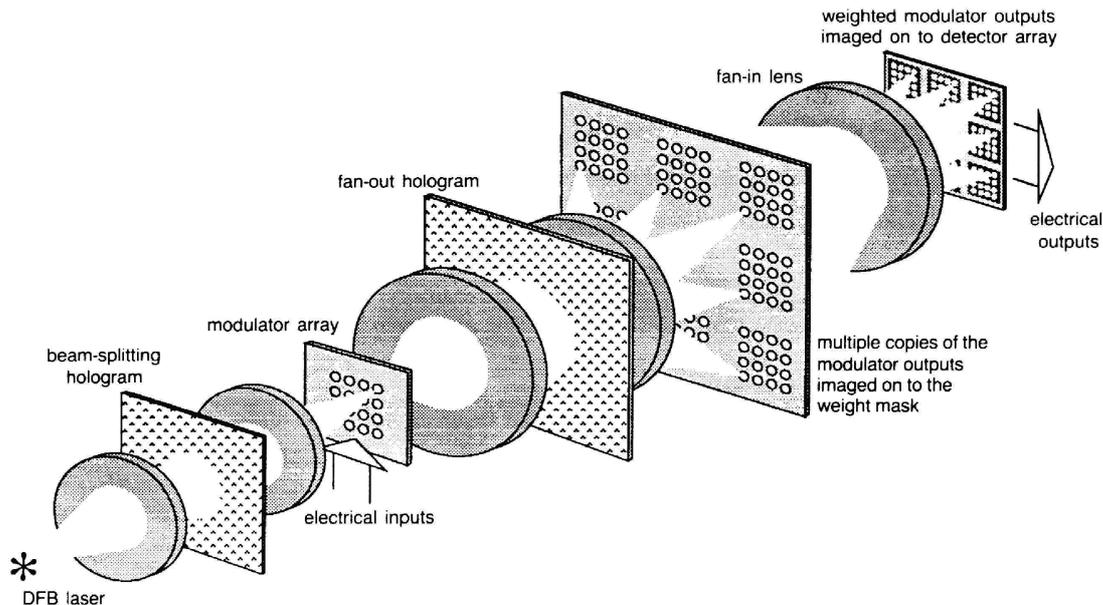

Figure 16

Example of an optoelectronic neural network employing both fan-in and fan-out. Please see text for more details. Diagram taken from [1].

This system takes input straight from a *distributed feedback* (DFB) laser which it splits and distributes evenly to the modulator array. Fan-out is then performed by a hologram to create multiple images of the modulator array on a weight mask. This mask applies an appropriate weight modification on the input signal before it continues on. Another lens is then used to fan-in the weighted information onto the correct detector on the detector array.

6.4 Applications of Optoelectronic Neural Networks

Optoelectronic neural networks can be applied to any situation where a neural network could be applied. Some good examples of such are:

- Telecommunications Switching.

- Smart Control Systems

- Image Processing.

- Character Recognition.

- Speech Processing.

As long as the problem has a lot of inherent concurrency then the system is suited to finding a solution. These systems are not suited to any problems which are sequential in nature: the next instruction needing a result from the previous.

For further information on optoelectronic neural networks, please consult [1], [22], [55] and [56]

6.5   Conclusion

Optics has a major advantage over electronics when implementing highly interconnected systems because of the inherent properties of light. It can be clearly seen that the promise in such systems and their application to neural network implementation is great. There is unfortunately one problem lurking in the background: noise. This is the very problem which limits the degree of fan-out because each signal that the input is divided into gives its own contribution to overall system noise. Too much noise and the system becomes useless.

7   Conclusion

This report has carefully looked at the technologies and theories involved in creating a smart pixel optical neural network. It highlighted the reasons behind an optical implementation but has not yet addressed one issue: what is the demand for such a system?

Optoelectronic neural networks will remain simply a curiosity rather than a useful system unless industrial uses can be found for them. There are however some very promising applications for the systems: telecommunications switching and control tasks. The sheer parallelism of optoelectronic neural networks makes them highly suited to these applications and, given continued development, smart pixel optoelectronic neural networks may one day be as common as the microprocessor.

8   Glossary

ANN	Artificial Neural Network
EOM	Electro-Optic Modulator
CGH	Computer Generated Hologram
CHFET	Complimentary Heterostructure Field Effect Transistor
CMOS	Complimentary Metal Oxide Semiconductor
DOE	Diffractive Optic Element
DOES	Double Heterostructure Optoelectronic Switch
DFB	Distributed Feed Back
D-SEED	Photodiode Self Electro-optic Effect Device
EARS	Exciton Absorption Reflection Switch
ELO	Epitaxial Lift-Off
EOM	Electro Optic Modulator
FET	Field Effect Transistor
FET-SEED	Field Effect Transistor Self Electro-optic Effect Device
HEMT	High Electron Mobility Transistor
HPT	Heterojunction Photo-Transistor
HPT-MOD	HPT – Modulator
HPT-VCSEL	Heterojunction Photo-Transistor – Vertical Cavity Surface Emitting Laser
LAOS	Light Amplifying Optical Switch
LED	Light Emitting Diode
L-SEED	Logic Self Electro-optic Effect Device
LVQ	Learning Vector Quantisation
MESFET	Metal Semiconductor Field Effect Transistor
MFLOPS	Million Floating Operations Per Second
MLP	Multi-Layer Perceptron
MOD	Modulator
MOET	Monolithic Opto-Electronic Transistor
MQW	Multiple Quantum Well
M-SEED	Multi-state Self Electro-optic Effect Device
MSM	Metal-Semiconductor-Metal
NN	Neural Network
OE	Optical Element
OEIC	Opto-Electronic Interconnects
R-SEED	Resistive Self Electro-optic Effect Device
RTD	Resonant Tunnelling Diode
SEED	Self Electro-optic Effect Device
SPA	Smart Pixel Array
SLM	Spatial Light Modulator
S-SEED	Symmetric Self Electro-optic Effect Device
VCSEL	Vertical Cavity Surface Emitting Laser
VLSI	Very Large Scale Integration
VSTEP	Vertical to Surface Transmission Electro-Photonic
XOR	Exclusive Or

9 Bibliography

[1] R. P. Webb and A. W. O'Neill, *"Optoelectronic Neural Networks"*, BT Technology Journal, Volume 10, No. 3, pages 144-154, July 1992.

[2] S. G. Smyth, *"Preprogramming multilayer Perceptrons"*, BT Technology Journal, Volume 10, No. 3, pages 30-37, July 1992.

[3] Song Yu and Stephen R. Forrest, *"Implementations of Smart Pixels for Optoelectronic Processors and Interconnection Systems I: Optoelectronic Gate Technology"*, Journal of Lightwave Technology, Volume 11, No. 10, pages 1659-1669, October 1993.

[4] Song Yu and Stephen R. Forrest, *"Implementations of Smart Pixels for Optoelectronic Processors and Interconnection Systems II: SEED Based Technology and Comparison with Optoelectronic Gates"*, Journal of Lightwave Technology, Volume 11, No. 10, pages 1669-1680, October 1993.

[5] John l. Hennessey and David A. Patterson, *"Computer Architecture a Quantitative Approach"*, 1996 Morgan Kaufmann.

[6] Michael J. Quinn, *"Parallel Computing Theory and Practice"*, 1994 McGraw-Hill.

[7] Harold S. Stone, *"Parallel Processing with the Perfect Shuffle"*, IEEE Transactions on Computers, Volume C-20, No. 2, February 1971

[8] H. Soda, K. Iga, C. Kitahara and Y. Suematsu, *"GaInAsP/InP Surface Emitting Lasers"*, Japanese Journal of Applied Physics, 18, page 2329, 1979.

[9] C. J. Chang-Hasnain, M. and J. P. Harbison, L. T. Florez and C.Lin, *"Monolithic Multiple Wavelength Surface Emitting Lasers"*, Journal of Lightwave Technology 9 (12), pages 1665-1672, 1991.

[10] S. Desmond Smith, *"Optoelectronic Devices"*, 1995 Prentice Hall International Series in Optoelectronics.

[11] Amnon Yariv, *"Quantum Electronics – Third Edition"*, 1989 John Wiley & Sons.

[12] T. C. B. Yu, R. J. Mears, A. B. Davey, W. A. Crossland, M. W. G. Snook, N. Collings and M.Birch, *"Design of Smart Spatial Light Modulators for Neural Networks"*, Optical Memory and Neural Networks 4 (3), pages 223-229, 1995.

[13] R. H. Yan, R. J. Simes and L. A. Coldren, *"Analysis and Design of Surface-Normal Fabry-Perot Electro-optic Modulators"*, IEEE Journal of Quantum Electronics 25 (11), pages 2272-2280, 1989.

[14] D. A. B. Miller, D. S. Chemla, T. C. Damen, T. H. Wood, C. A. Burrus, A. C. Gossard and W. Wiegmann, *"Novel Hybrid Optically Bistable*

Switch: The Quantum-Well Self-Electro-Optic Effect Device", Optics Letters QE-21, 13-15, 1984.

[15] M. Taghizadeh, J. M. Miller, P. Blair and F. A. P. Tooley, *"Developing Diffractive Optic Elements for Optical Computing"*, IEEE Micro, Volume 14, Number 16, December 1994.

[16] M. Taghizadeh and J. Turunen*, "Synthetic Diffractive Elements for Optical Interconnection"*, Optical Computing and Processing, 2 (4), pages 221-242, (1992).

[17] P. Keller and A. Gmitro, *"Design and Analysis of Fixed Planar Holographic Interconnects for Optical Neural Networks"*, Applied Optics 31 (26), pages 5517-5525, 1992.

[18] B. Fracasso and J. L. de Bougrenet de la Tocnaye, *"Node-based Reconfigurable Volume Interconnections"*, Applied Optics 33 (23), pages 5348-5481, 1994.

[19] A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham and L. M. F. Chirovsky, *"Symmetric Self-Electro-Optic Effect Device: Optical Set-Reset Latch, Differential Logic Gate and Differential Modulator/Detector"*, IEEE Journal of Quantum Electronics 25 (8), pages 1928-1936 (1989).

[20] F. A. P. Tooley, *"Optoelectronic Devices"*, Lecture Notes from Heriot-Watt University, 1998.

[21] B. S. Wherret, *"Optics in Computing"*, Lecture Notes from Heriot-Watt University, 1998.

[22] A. J. Waddie, *"Optoelectronic Neural Networks: The Design and Analysis of a Network Based Around Optoelectronic Smart Pixels"*, Phd Thesis, Heriot-Watt University, September 1995.

[23] John J. Soragnan, *"Distributed and Parallel Processing and Neurocomputing"*, Lecture Notes for Strathclyde University, 1996-1997.

[24] Robert L. Harvey, *"Neural Network Principles"*, Prentice Hall International Editions, 1994.

[25] James A. Anderson, *"An Introduction to Neural Nets"*, IT Press, 1995.

[26] Simon Haykin, *"Neural Networks"*, Macmillan Publishing Company, 1994.

[27] F. Rosenblatt, *"Principles of Neurodynamics"*, Washington DC Spartan Press, 1961.

[28] Clifford Lau*, "Neural Networks: Theoretical Foundations and Analysis"*, IEEE Press, 1991.

[29] R. P. Lippmann, *"An Introduction to Computing With Neural Nets"*, IEEE ASSP Magazine, pages 4-22, April 1987.

[30] B. Widrow and M. A. Lehr, *"30 Years of Adaptive Neural Networks: Perceptron, Madaline and Backpropagation"*, Proceedings of IEEE, volume 78, number 9, pages 1415-1442, September 1990.

[31] T. Kohonen, *"Self Organisation and Associative Memory"*, 2nd edition, Berlin: Springer Verlag, 1987.

[32] K. Beyzavi, D. S. Kim, C. P. Onto and P. E. Burrows, *"A Cascadable InGaAsP InP Optoelectronic Smart Pixel with Low Switching Energy"*, IEEE Photon, volume 7 10, pages 1162-1164, October 1995.

[33] J. Cowles, A. L. Gutierrezaitken, P. Bhattacharya and G. I. Haddad, *"7.1GHz Bandwidth Monolithically Integrated In0.53Ga0.47As & InO.52Al0.48As Pin HBT Transimpedance Photoreceiver"*, IEEE Photon, volume 6 8, pages 963-965, August 1994.

[34] D. Nichols, N. K. Dutta, P. R. Berger and P. R. Smith, *"Monolithic GaAs/AlGaAs Pin/MESFET Photoreceiver Using a Single Molecular Beam Epitaxy Growth Step"*, Electronic Letters, volume 29 12, pages 1133-1134, June 10th 1993.

[35] D. A. B. Miller, M. D. Feuer, Y. T. Chang, A. S. C. Shunk, J. E. Henry, D. J. Burrows, D. S. Burrows and D. S. Chemla, *"Field-Effect Transistor Self-Electro-Optic Effect Device: Integrated Photodiode, Quantum Well Modulator and Transistor"*, IEEE PTL, volume I, number 3, pages 62-64, March 1989.

[36] K. V. Shenoy, C. G. Fonstad, A. C. Grot and U. Psaltis, *"Monolithic Optoelectronic Circuit Design and Fabrication by Epitaxial Growth on Commercial VLSI GaAs MESFETs"*, IEEE Photon, volume 7 5, pages 508-510, May 1995.

[37] M. Makiuchi, H. Hamaguchi, T. Kumai, O. Aoki, Y. Oikawa and O. Wada, *"GaInAs pin Photodiode/GaAs Preamplifier Photoreceiver for Gigabit-rate Communications Systems using Flip-Chip Bonding Techniques"*, Electronics Letters, volume 24, number 16, pages 995-996, August 4th, 1988.

[38] M. L. Goodwin, A. J. Mosely, M. Q. Kearley, R. C. Morris, C. J. G. Kirkby, J. Thompson, R. C. Goodfellow and I. Bennion, *"Optoelectronic Component Array for Optical Interconnection of Circuits and Systems"*, Journal of Lightwave Technology, volume 9, number 12, pages 1639-1645, December 1991.

[39] K. W. Goossen, J. A. Walker, L. A. D'asaro and S. P. Hui, *"GaAs MQW Modulators Integrated with Silicon CMOS"*, IEEE Photon, volume 7 4, pages 360-362, April 1995.

[40] H. Wada and T. Kamijoh, *"Room-Temperature CW Operation of InGaAsP Lasers on Si Fabricated by Wafer Bonding"*, IEEE PTL, volume 8, number 2, February 1996.

[41] S. M. Fike, B. Buchanan, N. M. Jokerst and M. A. Brooke, *"8x8 Array of Thin Film Photodetectors Vertically Electrically Interconnected to Silicon Circuitry"*, IEEE Photon, volume 7 10, pages 1168-1170, October 1995.

[42] R. A. Novotny, *"Analysis of Smart Pixel Digital Logic"*, Heriot-Watt University, November 1996.

[43] G. W. Taylor, R. S. Mand, J. G. Simmons and A. Y. Cho, *"Optically Induced Switching in a p-Channel Double Heterostructure Optoelectronic Switch"*, APL 49 (21), pages 1406-1408, 1986.

[44] C. W. Wilmsen and F. R. Beyette, *"Smart Pixels Using the Light Amplifying Optical Switch (LAOS)"*, IEEE Journal of Quantum Electronics, volume 29, number 2, pages 769-774, 1993.

[45] K. Kasahara, Y. Tashiro, N. Hamao, M. Sugimoto and T. Yanase, *"Double Heterostructure Optoelectronic Switch as a Dynamic Memory with Low-Power Consumption"*, APL 52, pages 679-681, 1988.

[46] P. Zhou, J. Cheng, J. C. Zolper, K L. Lear, S. A. Chalmers, G. A. Vawter, R. E. Leibenguth and A. C. Adams, *"Monolithic Optoelectronic Switch Based on the Integration of a GaAs/AlGaAs Heterojunction Bipolar Transistor and a GaAs Vertical-Cavity Surface-Emitting Laser"*, IEEE PTL volume 5, number 9, pages 1035-1038, September 1993.

[47] C. Amano, S. Matsuo and T. Kurokawa, *"Novel Photonic Switch Arrays Consisting of Vertically Integrated Multiple-Quantum-Well Reflection Modulators and Phototransistors: Exciton Absorptive Reflection Switch"*, IEEE PTL, volume 31, number 8, August 1991.

[48] M. K. Hibbs-Brenner, S. D. Mukherjee, B. L. Grung and J. Skogen, *"GaAs OEICs for Optoelectronic Smart Pixels"*, LEOS 1992 Summer Topical Meeting Digest on Smart Pixels, Paper MC4, pages 26-27, August 1992.

[49] U. Kehrli, D. Leipold, K. Thelen, O. J. Homan, H. P. Schweizer, P. Seitz and B. D. Patterson, *"Monolithic Integrated Optoelectronic Smart Pixels"*, Electronic Letters, volume 30, number 24, page 2069-2070, 1994.

[50] S. J. Matsuo, T. Nakahara, Y. Kohama, Y. Ohiso, S. Fukushima and T. Kurokawa, *"Monolithically Integrated Photonic Switching Device Using an MSM PD, MESFETs, and a VCSEL"*, IEEE Photonics Technology Letters, volume 7, number 10, pages 1165-1167, 1995.

[51] A. L. Lentine and D. A. B. Miller, *"Evolution of the SEED Technology: Bistable Logic Gates to Optoelectronic Smart Pixels"*, IEEE Journal of Quantum Electronics, volume 29, number 2, pages 655-669, 1993.

[52] T. Nakahara, S. Matsuo, C. Amano and T. Kurokawa, *"Switchable-Logic Photonic Switch Array Monolithically Integrating MSM's, FET's, and MQW Modulators"*, IEEE PTL, volume 7, number 1, pages 53-55, January 1995.

[53] B. Y. Aull, K. B. Nichols, PA. Maki, S.C. Palmateer, E. R. Brown and T. A. Lind, *"Monolithic Optoelectronic Transistor: A New Smart Pixel Device"*, Applied Physics Letters, volume 63, number 11, pages 1555-1557, 1993.

[54] R. P. Webb, *"Overcoming Noise in a Fast Neural Network"*, BT Laboratories, Martlesham Heath, Ipswitch, IP5 7RE, UK, webb_r_p@bt-web.bt.co.uk.

[55]   A. J. Waddie and J. F. Snowdon, *"A Smart Pixel Optical Neural Network Design Using Customised Error Propagation"*, Department of Physics, Heriot-Watt University.

[56]   D. Huang, J. F. Snowdon and A. J. Waddie, *"Free Space Optical Technology and Neural Network Applications"*, Paper presented at Appl. Opt. Div. Conf. Reading, 16-19 September 1996.

[57]   J. Hertz, A. Krough and R. G. Palmer, *"Introduction to the Theory of Neural Computation"*, Addison-Wesley, 1991.