Sheffield Hallam University

An FPGA based real-time image classification system

Deepayan Bhowmik

M.Sc. in Electronics and Information Technology
2005-06

First supervisor: Dr. Bala P. Amavasai
Second supervisor: Mr. Tim J. Mulroy

This thesis is submitted in partial fulfilment
of the requirements for the degree of
Masters of Science
in Electronics and Information Technology, School of Engineering,
Sheffield Hallam University

Acknowledgement

My sincere thanks and regards should go to my supervisor Dr. Bala Amavasai who was a constant source of inspiration and director throughout this project even before starting it. I am also grateful to my second supervisor Mr. Tim J. Mulroy. I acknowledge the support of EPSRC Nanorobotics project GR/S85696/01. My special thanks should go to Dr. Clay Gloster, RARE project, Howard University, Washington, USA to allow me to use his VHDL code for floating point division.

This thesis is dedicated to my parents for their continuous support. I also extend my acknowledgement to all friends in Sheffield. Last but not the least, I express my sincere gratitude to Machine Vision Lab researchers, lab technician Mr. Misko and all those who have directly or indirectly contributed to the completion of this project work.

Publication

[1] **Deepayan Bhowmik**, Bala P. Amavasai, T. J. Mulroy, "Real-time object classification on FPGA using moment invariants and Kohonen neural networks", Proceedings of the IEEE SMC UK-RI Chapter Conference 2006 on Advances in Cybernetic Systems, September 7-8, Sheffield, UK, ISSN 1744-9189

Abstract

Machine vision is an integral part of machine intelligence. The primary focus of any machine vision system is to recognise and classify objects in the surrounding area. The concept of Artificial Intelligence has been deployed in number of machine intelligence application keeping in mind the fact that a machine will one day be able to imitate a human being. It is also applicable for an object recognition task.

In this thesis we address the issue of object classification using a different approach. Our objective of object classification can be divided into two parts, namely image processing and recognition, or classification of processed image. The image processing stage is handled by a moment calculation of the captured images. An Artificial Neural Network is implemented in the recognition or object classification stage. The main constraint of this work is to reduce the response time and make the system as fast as possible in order to fulfill real-time objectives. The use of dedicated hardware is probably one of the possible solutions. Hence we make use of reconfigurable hardware like Field Programmable Gate Array (FPGA).

In this thesis, the use of moment invariants and Kohonen neural networks for real time object classification is addressed. The implementation of such a scheme using a reconfigurable hardware FPGA (Field Programmable Gate Array) device is described. In the image processing stage, the Hu's moment invariants algorithm has been implemented in hardware and the issues surrounding this implementation is discussed. Following the image processing stage, a neural network is employed for the classification stage. By using the Kohonen unsupervised neural network the system is essentially self-supervised and it is able to perform an all parallel neural computation for classification purposes. A discussion of the concept and real simulation results are provided.

Contents

Chapter 1: Introduction

1.1 Introduction

The requirement for the recognition and classification of objects in real-time is important for many real world tasks, especially in robotics and industrial-type applications. A dedicated hardware for this purpose is highly. The use of of artificial neural network for classification allows the system to adapt to the environment. The aim of the thesis is to explore the possibility of developing a dedicated hardware for image processing and by using artificial neural network to make it adapt to the environment. The following issues have been addressed during the course of this project:

- The use of moment invariants for image classification
- Hu's moment invariant
- The use of Artificial Neural Network (ANN) as efficient mode of image classification
- Kohonen Neural Network as self-adaptive learning system
- The use of Field Programmable Gate Array (FPGA) for designing hardware
- Timing analysis as a constraint of real-time application

1.2 Background

The greatest challenge in current times, of any algorithm and circuitry, is speed. The idea of dedicated hardware for a specific type of application is to handle complexity with better time response. One of the ways to implement and test a circuit is to map and synthesis the algorithm in FPGA (Field Programmable

Gate Array). The domain of application of FPGAs has traditionally been in digital logic and digital signal processing (DSP). DSP-type problems are easily mapped to FPGAs due to the fact that most DSP functions are made up of sum-of-product type operations that consist of simple logic. Since images are essentially 2-D signals, image processing algorithms have also been widely implemented on FPGA.

However the implementation of higher level machine vision algorithms (with ANN optimisation) that consist of a number of decision making stages is somewhat limited. This is largely due to the limited number of directly mapped arithmetic functions available and the complexity in designing algorithms that are able to adapt or optimise online, since FPGA designs are often static. The latter problem can be overcome by parameterising the algorithms and using external memory to store these parameters.

1.3 Motivation

Machine vision is one of the most interesting and current research areas where researchers are trying to develop intelligent machine. The primary focus of machine vision in industry is to classify objects as quick as possible. Real-time classification is important for many real world and industrial applications especially in robotics. A variety of blob and shape based algorithms exist, but many of these do not meet real-time constraints.

In recent times, with the advent of sophisticated software tools, the use of Field Programmable Gate Arrays (FPGAs) has changed from being simply a gluelogic type component to a vehicle for complete delivered solutions. FPGAs are made up of programmable logic components and programmable interconnects. Unlike other technologies, that force the programmer or designer to make critical decisions in the early part of development, FPGAs allow the development of the application to be adapted and improved over time.

Developing FPGA solutions is comparable to developing solutions in software (rather than hard-coding).

The use of ANN is quite interesting to improve the timing performance of the system. However the Van-Neuman computer architecture (the normal computer processor used now-a-days *e.g.* Pentium series from Intel) does not directly support the construction of parallel ANNs. As far as ANN is concerned, it performs activities are performed in parallel on each neuron and that is where a sequential computer lacks. A neuro-chip based processor might solve the problem. But for a specialized application like object classification, probably a dedicated hardware is more useful. And here comes the concept to use of FPGA in this purpose.

The above discussion is the main motivation of this project. The project tries to explore the opportunity how suitable a FPGA board is to perform an object classification with the application of ANN.

1.4 Thesis Description

This thesis, describes an attempt to implement an object classification paradigm that is able to classify objects in real-time using an FPGA solution. The process makes use of two well studied algorithms in the area of machine vision and neural networks.

Hu [1] moment invariants is one of the well proven algorithm for 2-D image processing using a non-orthogonal central moments. Seven Hu descriptors is unique given a specific pattern or shape.

Kohonen [2] unsupervised neural network is a popular self-organising unsupervised learning Neural Network. It is very useful for data classification with an all parallel computation. A K-means clustering algorithm can classify and give identity to a self organizing unsupervised system.

The computation of moment invariants has been implemented in hardware. However the Kohonen neural network algorithm is divided into two

parts, namely training mode, and detection mode. Currently only training of the system is performed using MATLAB. The result of the training is used for object classification and this has been implemented (and simulated) on an FPGA device.

1.5 Programme of work

At the beginning of the project the work programme has been decided. The project follows the work program as a guide line. Following is the work program for the project.

a) Algorithm development for 2-D image recognition (Hu's moment invariant chosen here)
b) Algorithm optimisation using Artificial Neural Network (ANN)
c) Proposed ANN optimisation: Kohonen Neural Network
d) Implementation of Hu's moment invariants in FPGA
e) Hardware realization using VHDL coding
f) VHDL implementation of Kohonen ANN
g) Initial training of the system by MATLAB
h) Simulation and testing of VHDL code
i) Development and simulation the Stratix FPGA board in Altera Quartus software
j) Timing analysis for next pahse

1.6 Deliverables

The deliverables from this project are listed as follows

a) A real time 2-D object classification with very small response time
b) An implementation of Hu's moment invariant in FPGA board
c) An implementation of artificial neural network in FPGA board
d) An algorithm optimisation for image processing using ANN

e) Dedicated hardware for real time image classification

1.7 Thesis Formulation

This section discusses constraints such as schedule, technical limitations, possible hazards etc. and develops a work plan for the entire thesis.

1.7.1 Time/Schedule

Table 1.1 shows the Gantt chart of the thesis time line.

ID	Task Name	Duration	Start	Finish
1	Background Work	27 days	Mon 05/06/06	Tue 11/07/06
2	Background study	4 days	Mon 05/06/06	Thu 08/06/06
3	Literature review	4 days	Tue 06/06/06	Fri 09/06/06
4	Understanding Hu's invarient	4 days	Mon 12/06/06	Thu 15/06/06
5	Understanding ANN	11 days	Wed 14/06/06	Wed 28/06/06
6	Understanding Stratix FPGA	7 days	Mon 19/06/06	Tue 27/06/06
7	Learning VHDL	9 days	Thu 29/06/06	Tue 11/07/06
8				
9	Implementation	24 days	Wed 12/07/06	Mon 14/08/06
10	Moment	5 days	Wed 12/07/06	Tue 18/07/06
11	Hu's invarient	6 days	Wed 19/07/06	Wed 26/07/06
12	Neural Network	13 days	Thu 27/07/06	Mon 14/08/06
13				
14	Testing	7 days	Tue 15/08/06	Wed 23/08/06
15	Phase-I (Hu's Inverient)	4 days	Tue 15/08/06	Fri 18/08/06
16	Phase-II (Training of FPGA, ANN)	4 days	Tue 22/08/06	Fri 25/08/06
17	Phase-III (Testing with unknown obj)	3 days	Mon 28/08/06	Wed 30/08/06
18				
19	Report Writing	14 days	Tue 15/08/06	Fri 01/09/06

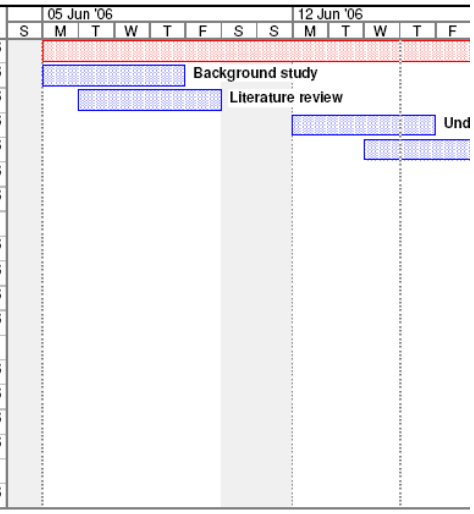

Table 1.1: Gantt chart of the project time line

1.7.2 Technical Limitations

This section focuses on various limitations of the thesis as listed below

- Difficulty on hardware realization of sequential algorithm
- Handling floating point arithmetic in VHDL, specially multiplication and division operation
- Handling function and process statements in VHDL
- Large compilation and simulation time made the experiment time longer

1.7.3 Potential Hazards

The precautions that were strictly followed during the entire project period were

- An erect sitting posture was maintained while working on the computer
- Breaks were taken at regular intervals to avoid cramps and sprains that can be caused due to sitting in front of the computer for long hours
- Hazards related for handling electronic equipment like FPGA (i.e. electro static discharge, ground problem etc)
- Hazards related to main power supply

1.8 Thesis organisation

The thesis has been written in the same way the project has been carried out. Chapter 2 discusses the background of the work. Details of literature review and discussion of similar kind of work has been described here. Chapter 3 is basically the discussion of theories namely Hu's moment invariant, Kohonen Artificial Neural Network *etc.* Based on these well proven theories the next part of the project was carried out. Algorithmic development & implementation is probably the main feature of the project. All details about the methodology and Algorithmic development are written in chapter 4. So far the thesis frames the idea and suggested the implementation method. Chapter 5 shows the results after simulation and does the timing analysis of the same. Finally chapter 6 concludes the project and suggests some future work.

Chapter 2: Literature Survey

2.1 Introduction

The chapter focuses on the previous work by other researchers in the area of this thesis. The literature survey discusses the image classification method on FPGA board and related hardware using Artificial Neural Network concept. In an overall view, the main limitation faced by researcher in the field is the hardware capacity of the FPGA board. However number people have developed an integrated system combined with a computer.

2.2 General Discussion

The domain of application of Field Programmable Gate Arrays has traditionally been in digital logic and digital signal processing. Digital Signal Processing type problems are easily mapped to FPGAs due to the fact that most Digital Signal Processing functions are made up of sum-of-product type operations that consist of simple logic. Since images are essentially 2-D signals, image processing algorithms have also been widely implemented on FPGA.

The implementation of higher level machine vision algorithms that consist of a number of decision making stages is more limited. This is largely due to the limited number of directly mapped arithmetic functions available and the complexity in designing algorithms that are able to adapt or optimise online, since FPGA designs are often static. Parameterising the algorithms and using external memory to store these parameters can overcome this problem.

Another limitation with FPGA is to handle floating point arithmetic. The signal processing calculations are sometimes quite complicated and deal with

lots of floating point number. The precision of floating point values is also important in this regards. The researchers in the field have always faced difficulty dealing with floating point arithmetic. In case of floating point calculation IEEE 754 format is followed and the arithmetic with IEEE 754 is different from a normal arithmetic calculation. This is probably the reason why hardware design engineers use to hate floating point numbers where as software engineers love to play with floating points.

Floating point can be described in IEEE 754 single precision (32 bit) or a double precision (64 bit) or even with an extended single precision (43 bit). As the floating points increase complexity of the design, often it is preferred to deal with single precision number. The next section discussed about different approach followed and explored by previous researchers in this area.

2.3 Relevant Past Work

In recent times Hirai *et. al.* [2] have designed complete single-task vision systems on FPGA, with the objective of detecting the position and orientation of distinctly visible planar objects. Their system consisted of three parts, namely the computation of image gravity centre, the detection of object orientation using radial projection and the computation of the Hough transform, albeit discrete. They reported being able to process images at a rate of around 200 fps, far more than the standard frame-rate of a PAL camera.

In a separate study, Arribas and Maciá [3] implemented the Santos-Victor paradigm to compute motion fields for use in robot guidance applications. The technique required the computation of optical flow fields in real-time. The flow fields were computed using both standard differential and correlation methods.

Recent approaches have been made towards the implementation of real-time object recognition using reconfigurable hardware. Most of these systems have been implemented using multiple FPGA boards or similar devices together

with a sequential computer, due to the limited number of gates available on each device.

For instance, Neema *et. al.* [6] have used multiple Digital Signal Processing processors and a sequential computer for their Automated Image Recognition system. They have described the model-based development of a real-time, embedded Automated Image Recognition (AIR) system. The idea behind a distributed processing using multiple Digital Signal Processing board is to implement a system capable of processing high sampling rate and large input images. DSP processors are used here to handle the complexity of the applied Discrete-classifier correlation filter algorithm. However the full system integration is only possible with a sequential computer along with the DSP boards.

In another paper, Yasin *et. al.* [7] have demonstrated an effective FPGA prototype for iris recognition. However, the system itself is not completely hardware based, since images are preprocessed using MATLAB, and FPGAs are only used for the recognition part. In their design, the authors make use of Multi Layer Perceptrons (MLPs) for classification. The numbers of neurons are limited due to hardware constraints and hence the computation complexity had to be reduced for implementation on an FPGA board.

Jean *et. al.* [8] implemented a system to accelerate the recognition process in infra-red images using an FPGA device. However once again, the system is dependent on a separate computer and in order to reduce complexity, the mathematical calculations have been performed with full precision integer values instead of floating point operations.

Y. Y. Chung *et. al.* [10] describes and implemented a partially conected neural network by Giga-Ops Spectrum G800 FPGAs based custom computer for a high speed neural network based classifier which can be used real time application. Again the custom computer consists multiple (up to 32) Xilinx logic

chips. They demanded that the system can provide a very high speed classifier for real time image recognition purpose.

Neural networks are an ideal example where the use of FPGAs can offer an advantage. This is due to the fact that the operations carried out are largely parallel. A number of efforts have been made for mapping neural networks onto hardware. For example Aibe *et. al.* [9] implemented a probabilistic neural network in FPGA hardware which exhibited a parallel architecture.

2.4 Summary

A brief literature review has been carried out in this chapter. The main focus is on use of FPGA for the implementation of image processing algorithm and Artificial Neural Network. Researcher has tried to implement a parallel and dedicated hardware in order to improve the speed over a sequential computer. But the main limitation faced by almost everybody is the number of gates available in FPGA. Number of times either people used multiple boards and a sequential computer to overcome this limitation. Interestingly everybody tries reconfigurable hardware namely Field Programmable Gate Arrays to design a dedicated hardware due the high cost of manufacturing a customised chip. The interests of researcher's show that the real-time image recognition is one of most lucrative and interesting field for FPGA application.

Chapter 3: Theoretical Discussion

3.1 Introduction

The aim of this chapter is to discuss the theory of the implemented concept and algorithms of the project. The project is clearly separated in two different parts namely moment calculation and object classification. The moment classification is based on Hu's Moment Invariants where as the object classification part has been dealt with Kohonen Artificial Neural Network. Both the algorithms are well proven and quite popular in its won field. The K-means clustering algorithm discussed here is useful for object classification in training mode. A brief discussion about Field Programmable Gate Array (FPGA) and Atera Stratix board is then carried out. One of the main issues faced during the VHDL coding was floating point arithmetic. A brief description on floating point arithmetic hardware, specially the details about IEEE 754 format might be useful at this stage.

3.2 Hu's Moment Invariant

In the field of image processing specially in machine vision and related fields, image moments are popular as it has some unique property or interpretation [16]. The moments are basically certain particular weighted averages (moments) of the image pixels' intensities or function of other moments. Moments can describe the object efficiently after segmentation.

One of the well studied methods of describing 2-D objects using moment invariants is Hu's moment invariant or descriptors [1]. The regular moment of a shape in an M by N binary image is defined in equation 3.1:

$$u_{pq} = \sum_{j=0}^{N-1}\sum_{i=0}^{N-1} i^p j^q f(i,j) \qquad \text{(eqn. 3.1)}$$

Where f(x; y) is the intensity of the pixel (either 1 or 0) at the coordinates (x; y) and p+q is said to be the order of the moment.

The calculation is a function of the distance between shape pixels. So the origin measurements are taken relative to the shapes centroid (x′; y′) to remove translational variability. The coordinates of the centroid are determined using the equation above as described in equation 3.2:

$$i' = \frac{u_{10}}{u_{00}} \quad \text{and} \quad j' = \frac{u_{01}}{u_{00}} \qquad \text{(eqn. 3.2)}$$

Now the relative moments are calculated with respect to the central moments and the same is described in equation 3.3

$$u_{pq} = \sum_{j=0}^{N-1}\sum_{i=0}^{N-1} (i-i')^p (j-j')^q f(i,j) \qquad \text{(eqn. 3.3)}$$

Normally individual moment values do not have the descriptive power to uniquely represent arbitrary shapes, nor do those posses the required invariance characteristics, but, sets of functions based on these moments can be determined which do [17]. Hu derived a set of seven rotational invariant moment functions which form a suitable shape representation (or vector). Hu descriptors are based on non-orthogonalised central moments that are invariant towards rotation, translation and scale. Hu descriptors are thus computed from normalised centralised moments up to the third order, and consist of seven moments in total. The seven formulae given in equation set 3.4 to 3.10 are normally used for algorithmic implementation.

$$I_1 = \eta_{20} + \eta_{02} \quad \text{(eqn. 3.4)}$$

$$I_2 = (\eta_{20} - \eta_{02})^2 + (2\eta_{11})^2 \quad \text{(eqn. 3.5)}$$

$$I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad \text{(eqn. 3.6)}$$

$$I_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad \text{(eqn. 3.7)}$$

$$I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad \text{(eqn. 3.8)}$$

$$I_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})] \quad \text{(eqn. 3.9)}$$

$$I_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad \text{(eqn. 3.10)}$$

The derivation of these seven invariants and the computation of η is available in [1], and so will not be described here. The combination of these seven Hu descriptors is unique given a specific pattern or shape. Hence it can be used in conjunction with a matching scheme to uniquely identify the object being input to the system. Computing these descriptors is straightforward and fast, and hence it is widely used in the area of real-time vision in robotic applications [18].

3.3 Kohonen Artificial Neural Network

Kohonen Artificial Neural Network is a popular unsupervised learning neural network. The Kohonen ANN is essentially a self-organizing unsupervised mapping system that can map input vectors of arbitrary length onto a lower dimension map. It is frequently described as a sheet-like neural network array. Patterns that are close together in Euclidean space will remain close in the final map, and are topologically ordered. The learning process of Kohonen ANNs optimizes the mapping until the weight change becomes negligible [11].

The Kohonen Neural Network is consists of a Kohonen map which has a single layer of neurons. All input vectors are connected to each neuron in the map. Kohonen map can be arranged in different topologies namely rectangular and hexagonal. A neuron map of rectangular shape is quite popular as shown in figure 3.1. The weight vector connected to each neuron is represented by (w_{ji}). The learning algorithm is determined by the Euclidean distance between an input vector and the corresponding weights of each neuron. The weights associated with the neuron which provides the smallest Euclidean distance is considered as the winning node. Then the weights associated with the winning neurons are updated. This weight update method has done in such a way that the winning neuron becomes representative of a specific pattern or class.

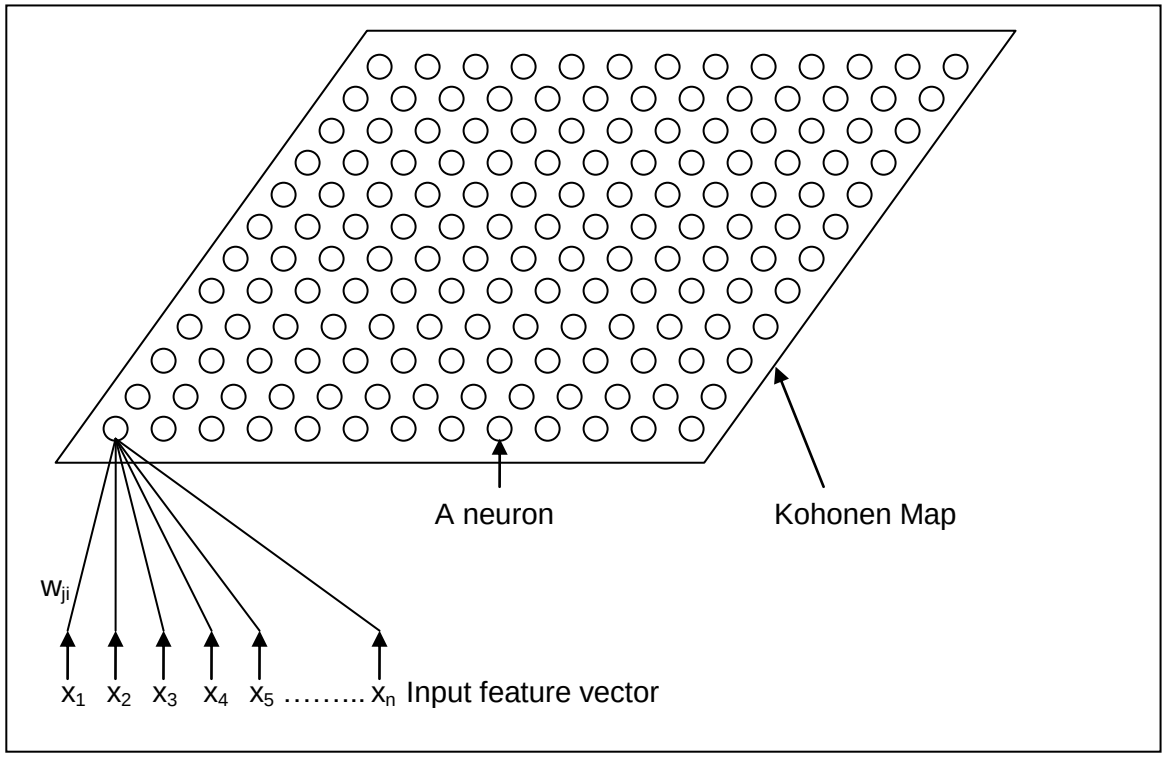

Figure 3.1: A Kohonen network

Though the winning node is main point of interest, in practice the weights associated with the neighborhood region of the winning node has also updated in a similar fashion. This method is quiet useful to specify a region of the

Kohonen map to be associated with different pattern or class. Figure 3.2 has shown a complete idea about the winning node and its neighborhood region.

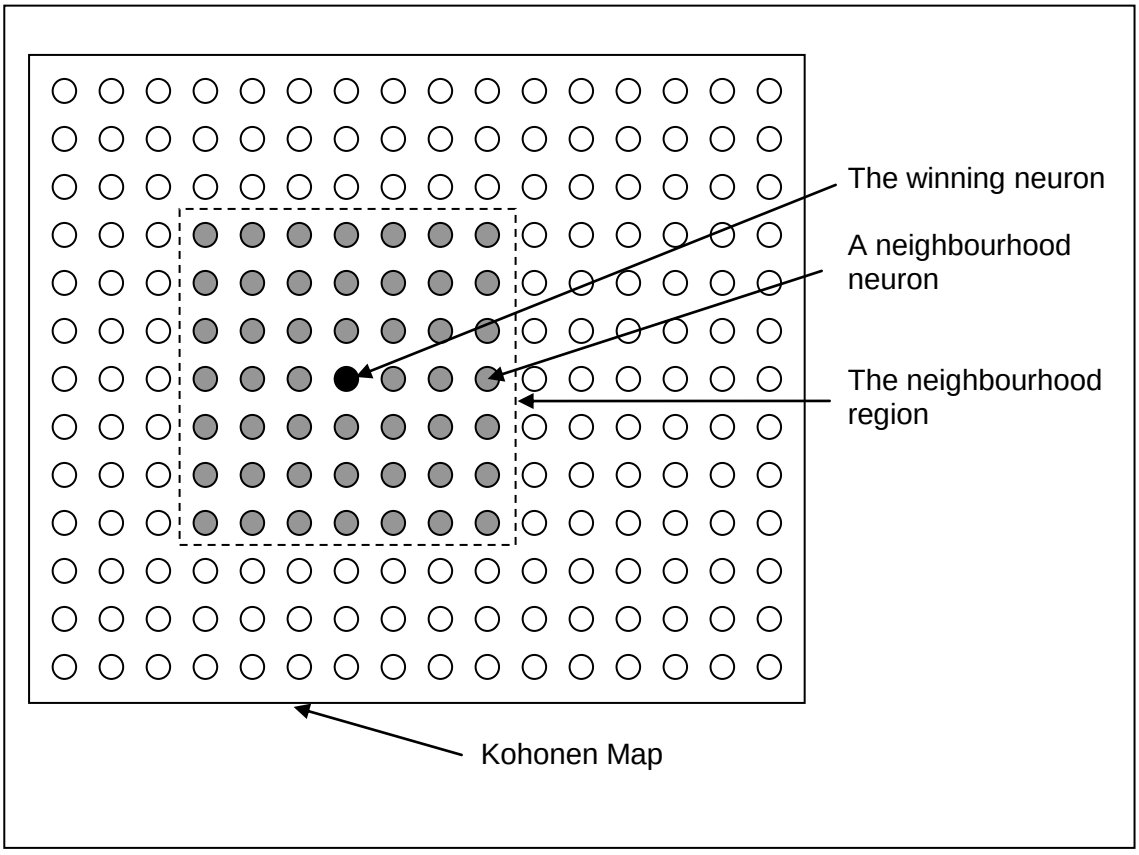

Figure 3.2: A Kohonen network with winning node and neighbourhood region

The normal method to deal with Kohonen map is to initialize weights of the neurons and standard neighborhood region. The initial weight values are chosen randomly with a value between 0.45 and 0.55. A neighborhood size of 5 is fair enough for a Kohonen map of 10 x 10 neurons.

The inputs data is mostly normalize for the application. Normally this normalization method varies in different application. All input data are given to the Kohonen map to all neurons.

The next step is to determine the Euclidean distance between the weight vector and the input vector of each neuron. Equation 3.11 represents the Euclidean distance measurement equation. Basically this computes the Euclidean

distance (d_j) between the input pattern (x) and the network connection weights w_{ij} of each neuron. The connection to j^{th} neuron from i^{th} input feature of the pattern is represented by the suffix *ij*. As described earlier the winning node or winning neuron will be the neuron associated with the smallest Euclidean distance. Equation 3.11 computes the value of (d_j) for j^{th} neuron with each input pattern has N elements.

$$\text{Euclidean distance: } d_j = \sqrt{\sum_{i=0}^{N-1}(x_i - w_{ij})^2} \qquad \text{(eqn. 3.11)}$$

Where x_i is the input to i and w_{ij} is the weight from input node i to output node j.

For simplicity of the calculation sometimes Euclidean distance calculation can be replaced by Manhattan distance. This is basically the summation of the absolute distances (d_j) measured between the input vector x_i and the weight vector w_{ij}. The Equation 3.12 represents the mathematical form of Manhattan distance calculation.

$$\text{Manhattan distance (“L}_1\text{ Norm”): } \sum_{i=1}^{k} | x_i - y_i | \qquad \text{(eqn. 3.12)}$$

Now the weights of the neurons in the Kohonen map needs to be updated. The neuron corresponding to the minimum distance is the winning node. The weights of the winning node and the neighbourhood region around the winning node are updated with the following update rule referred in Equation 3.13:

$$w_{ij}(n+1) = w_{ij}(n) + \lambda(n)(x_i(n) - w_{ij}(n)) \quad for\, 0 \le i \le N-1 \qquad \text{(eqn. 3.13)}$$

Where $w_{ij}(n+1)$ is the updated weight, $\lambda(n)$ is the learning rate and the term $(x_i - w_{ij})$ represents the error. The value of the learning rate normally lies between 0 and 1 and it controls convergence speed and stability. The value of the learning rate is normally determined experimentally. The learning process is iterative and continues until the network has converged satisfactorily. At the same time at predefined intervals the learning rate and the neighborhood size are gradually reduced. Normally a monitoring process has employed at regular intervals for the performance of the network. A network is said to be converged when the patterns of the same category active neurons in the same regions of the Kohonen map [11].

3.4 K-Means Clustering

Clustering is an unsupervised learning technique which allows a set of recorder data to be partitioned into two or more group [11]. The idea has been illustrated in figure 3.3

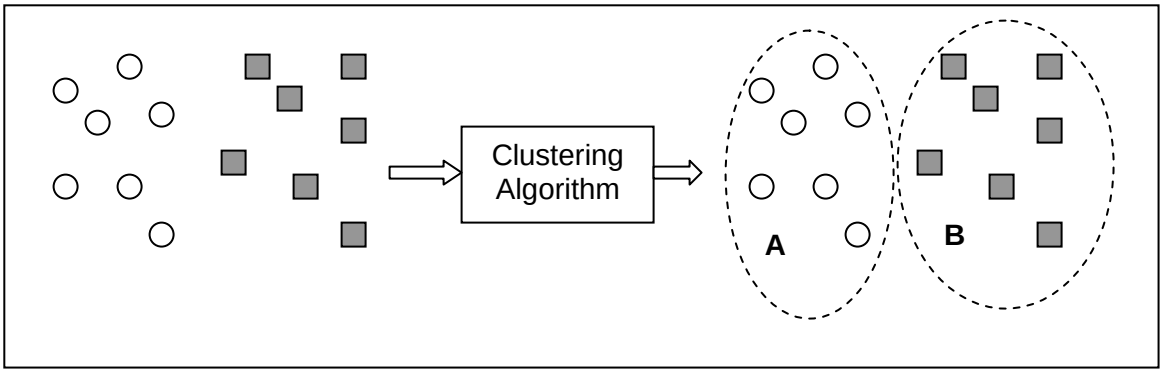

Figure 3.3: The role of clustering algorithm

The K-means algorithm clusters the objects based on attributes into k partitions [15]. The goal of this algorithm is to determine the k means of data generated from Gaussian distributions. The object attributes are asumed to be in

a vector space. It tries to minimize total intra-cluster variance which can be represented by the equation 3.14

$$V = \sum_{i=1}^{k} \sum_{j \in S_i} | x_j - \mu_i |^2 \qquad \text{(eqn. 3.14)}$$

Where there are k clusters S_i, i = 1, 2... k and μ_i is the centroid or mean point of all the points $x_j \in S_i$.

The k-Means algorithm starts by partitioning the input points into k initial sets, either at random or using some heuristic data and calculates the mean point or centroid [15]. A new partition is then formed by associating each point with the closet mean. The algorithm is then repeated until convergence which is obtained when the points now longer switch between clusters.

3.5 Field Programmable Gate Array (FPGA)

A Field Programmable Gate Array (FPGA) is a reconfigurable hardware which containing programmable logic and programmable interconnects [14]. A combinational or sequential logic can be programmed with the help of basic gate functions. The FPGAs are capable of handling a hardware implementation of simple or complex mathematical function and formulae. In most of the FPGAs the programmable logic components (commonly called logic elements [LE]) consists of basic gates like AND, OR, XOR etc. and also include memory element like a simple flip-flop or more complex block of memories.

The programmable interconnects with proper hierarchy allows logic elements (LE) of the FPGA to be interconnected as required by the hardware designer. The logic elements and the interconnections among them can be programmed after the manufacturing process by the designer. A custom made design is possible depending on the application requirement. And hence it is called field programmable and so the FPGA can perform whatever logical function is needed. Though FPGAs are slower than a application specific

integrated circuit (ASIC) and some more pros, designer prefer to use an FPGA because of their reconfiguration facility (probably easy to fix a bug) and lower non-recurring cost.

Now-a-days FPGAs have a very wide range of application field including DSP, software defined radio, aerospace, defence systems, ASIC prototyping, medical imaging, computer vision, bio-informatics etc. [14]. FPGAs especially find applications in any area or algorithm that can make use of the massive parallelism offered by their architecture.

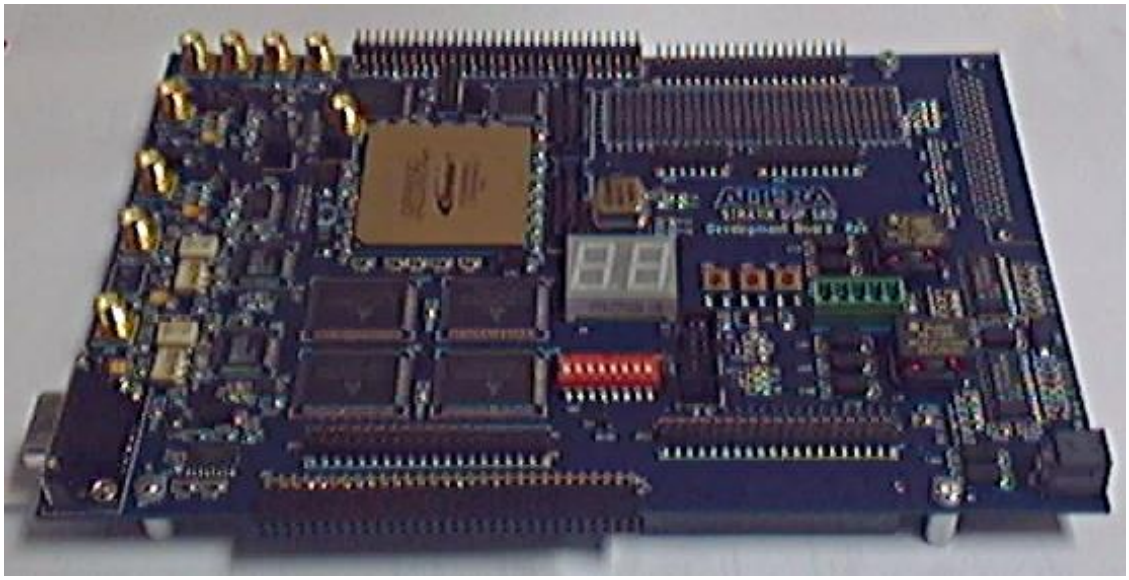

Figure 3.4: the Altera EP1S80 development board

The FPGA hardware used in this thesis is the Altera Stratix EP1S80 (by EPSRC Nanorobotics project GR/S85696/01), shown in Figure 3.4. The EP1S80 device from Altera Corp. is one of the largest 0.13-micron FPGA devices available. This board consists of 79,040 logic elements (LEs), 7.2 Mbits of embedded RAM, and 1,238 user I/Os. This board is a powerful development platform for digital signal Funded processing (DSP) designs, and features the Stratix EP1S80 device in the fastest speed grade (-6) 956-pin package. It consists of two 12-bit 125-MHz A/D converters, two 14-bit 165-MHz D/A converters, single ended or differential inputs and single-ended outputs, 2 MBytes of 7.5-ns

synchronous SRAM configured as two independent 36-bit buses, 64 Mbits of flash memory, dual seven-segment display, one 8-pin dipswitch, three user-definable pushbutton switches, one 9-pin RS-232 connector, two user-definable LEDs, on-board 80-MHz oscillator and a single 5-V DC power supply. For debugging interfaces, two Mictor-type connectors for Hewlett Packard (HP) logic analyzers and several 0.1- inch headers are available.

3.6 IEEE 754 Floating Point Format

IEEE Standard for Binary Floating-Point Arithmetic (IEEE 754) is the most common representation for real numbers in computers. IEEE 754 specifies four formats for representing floating-point values: single-precision (32-bit), double-precision (64-bit), single-extended precision (≥ 43-bit, not commonly used) and double-extended precision (≥ 79-bit, usually implemented with 80 bits) [12]. The main IEEE standard is described under the title of IEEE Standard for Binary Floating-Point Arithmetic (ANSI/IEEE Std 754-1985), and it is also known as IEC 60559:1989, Binary floating-point arithmetic for microprocessor systems (originally the reference number was IEC 559:1989) [12].

This part of the thesis will describe the normal formats used in 32 bit single precision floating point number. Normally IEEE 754 numbers have three basic components:

- Sign
- Exponent
- Mantissa

Usually the Mantissa is composed of the fraction and an implicit leading digit [13].

There are two basic number formats in IEEE 754: single precision and double precision. Table 3.1 shows the formats for single and double precision floating point numbers. The number of bits for each filed are shown below where as bit ranges are in square brackets.

	Sign	Exponent	Fraction
Single Precision	1[31]	8[30-23]	23[22-0]
Double Precision	1[63]	11[62-52]	52[51-0]

Table 3.1: IEEE 754 floating point number format

The sign bit represents the sign of the number. 0 denotes the positive number, where as a 1 denote a negative number.

The next field is the exponent field which needs to be represented both positive and negative exponents. For this purpose a bias is added to the actual exponent to get the stored exponent. A bias value of 127 is added for a single precision number (8 bit) and a double precision number (11 bit exponent) has a bias of 1023. For example a stored value of 167 indicates an exponent of (167 - 127) or 40 in single precision value. On the other hand an exponent of 5 means that (127 + 5) or 132 is stored in the exponent field in a single precision value.

The Mantissa represents the precision bits of the floating points. The Mantissa normally has an implicit bit and 23 fraction bits for a single precision number or 52 fraction bits for a double precision number. The floating point numbers are typically stored in normalized form to put the radix after the first non-zero digit. In a base 2 binary number has only possible non-zero number '1' and thus the IEEE 754 format does the optimization by assuming a leading digit of 1 where it is not represented explicitly. So the Mantissa has effectively 24 bit resolution by way of 23 fraction bits.

There are some special cases available in IEEE 754 format and those are listed below for reference only:

- Zero
- NaN (Not a Number)
- Demoralized number

3.7 Summary

This chapter here has discussed about the related theory of the implemented algorithm. The algorithm is mainly based on two popular and well studied theories namely Hu's moment invariant and Kohonen Neural network. The theories discussed here tried to give a general overview rather than a detail derivation or analysis.

Chapter 4: Methodology and Algorithmic Development

4.1 Introduction

Methodology and Algorithmic Development have been discussed and described in this chapter. The basic system diagram of the complete architecture of the proposed vision system is shown in Figure 4.1. As mentioned previously, classifying of objects consists of two modes, the training mode and the classification mode. The Otsu thresholder is commonly used to automatically binarise images by optimising a threshold level.

The whole implementation procedure has been divided in 3 parts:

i) Moment invariant computation,

ii) Training of the system

iii) Classification mode.

4.2 Moment Invariant Computation

In this mode, sample 2-D objects (patterns) are captured via a composite camera. The gray-valued image is then thresholded and the binarised values are presented to the system. The Hu descriptors that represent the pattern in the image are then extracted.

Floating point calculations have been achieved by using the Altera Megacore library or by using the author's own floating-point implementation. Implementation in hardware has been achieved using VHDL code with Altera Quartus II software.

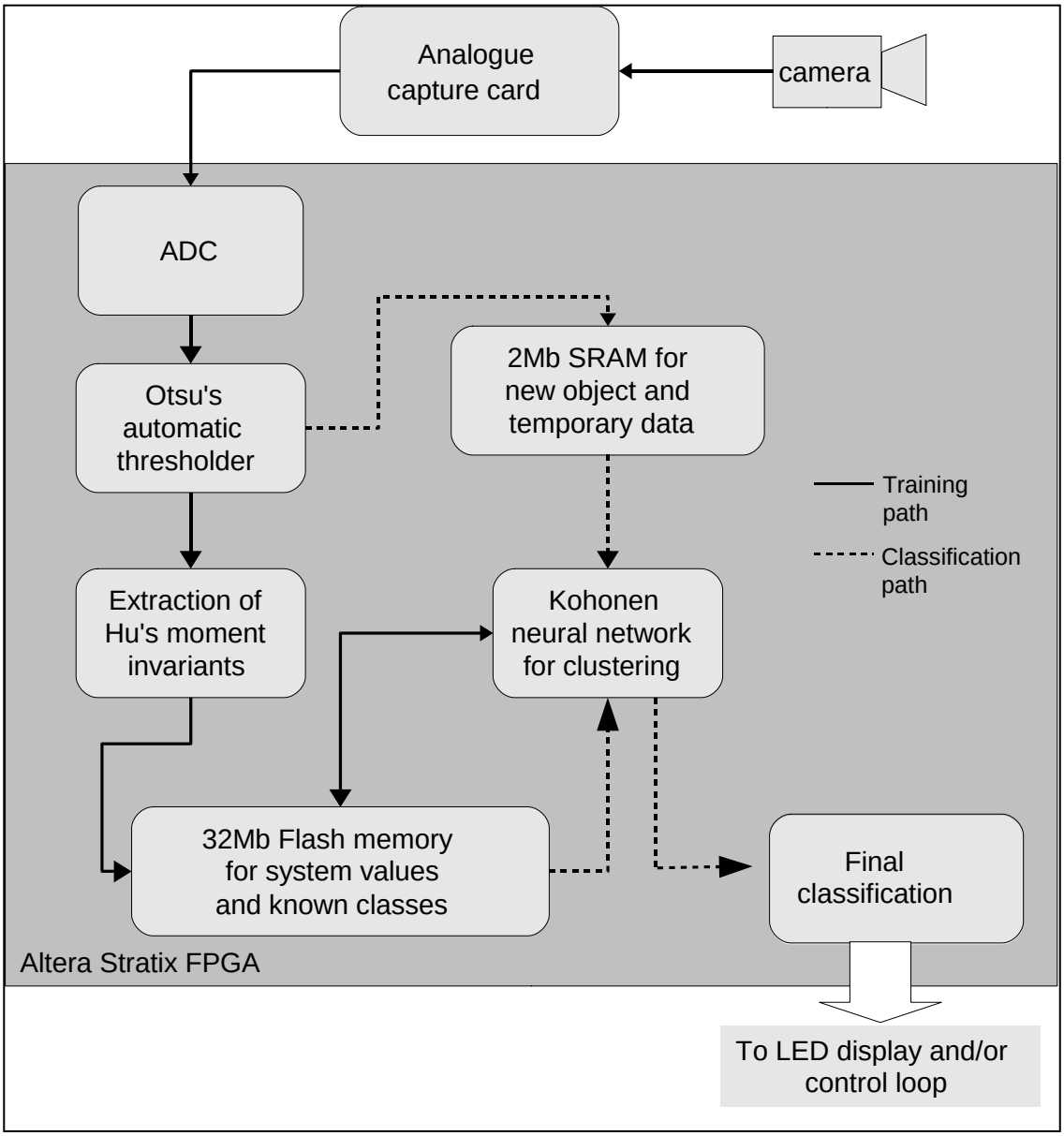

Figure 4.1: Training and classifying objects (the proposed system)

To ease and reduce computation, the moment computation is simplified by replacing all floating point powers with its nearest integer values. This has been experimentally observed to produce good results.

Virtually all the operations of the algorithm have been mapped in parallel. This parallel operation significantly increases the speed of the system. Single

precision floating point values have been used for all computation and conform to the IEEE 754 standard.

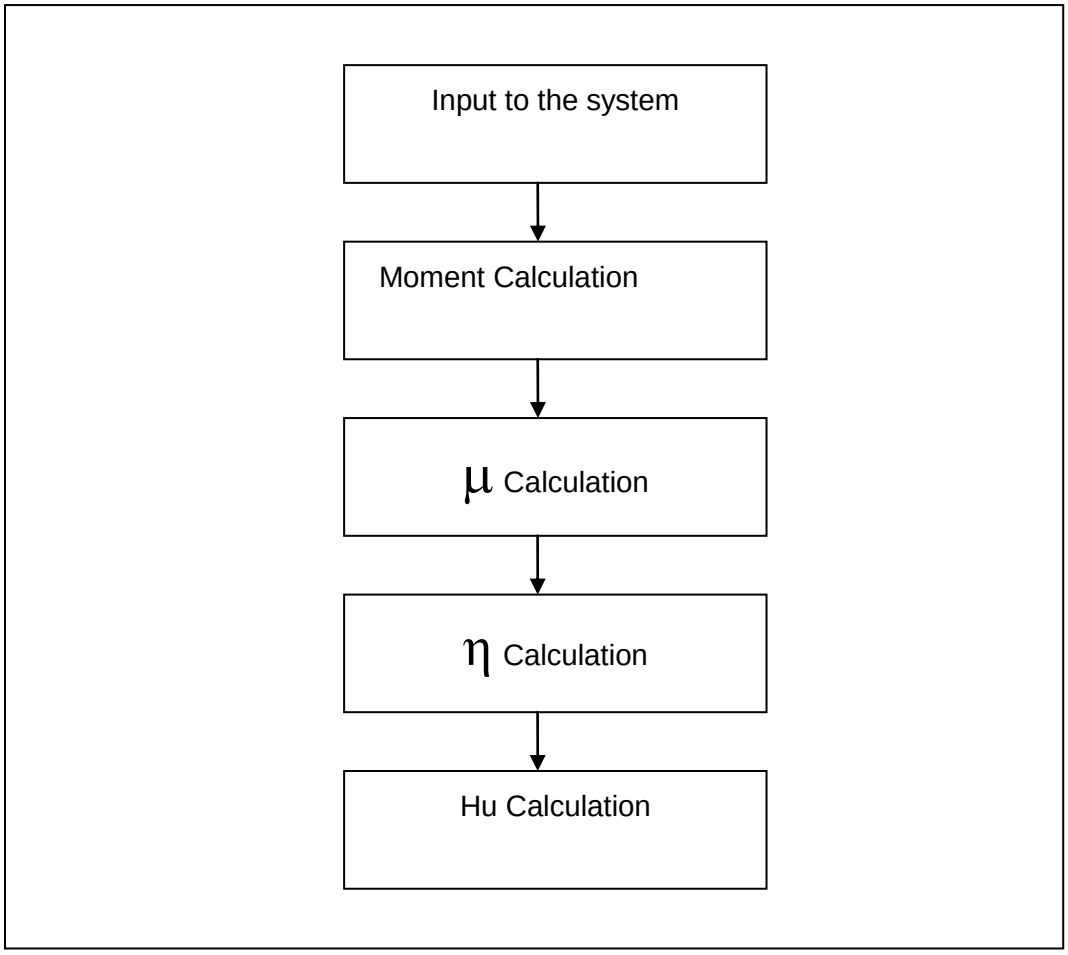

Figure 4.2: Algorithm for moment calculation.

As this work is a proof of a concept implementation, the input image matrix is hardcoded in VHDL. The computed seven moments have been used in the next stage of the system either for training (in the training mode) or classification (detection mode). The algorithm for the moment computation is shown in figure 4.2. The input data is iterated through in a sequential manner and then the different moment values have been calculated. After obtaining all moments the computation of the μ parameters (essentially high order moments) are computed in parallel mode. The μ values are required to compute η as shown in equation [3.4 to 3.10]. The final seven Hu moment invariants are calculated and passed on to the next phase.

4.3 Training of the System

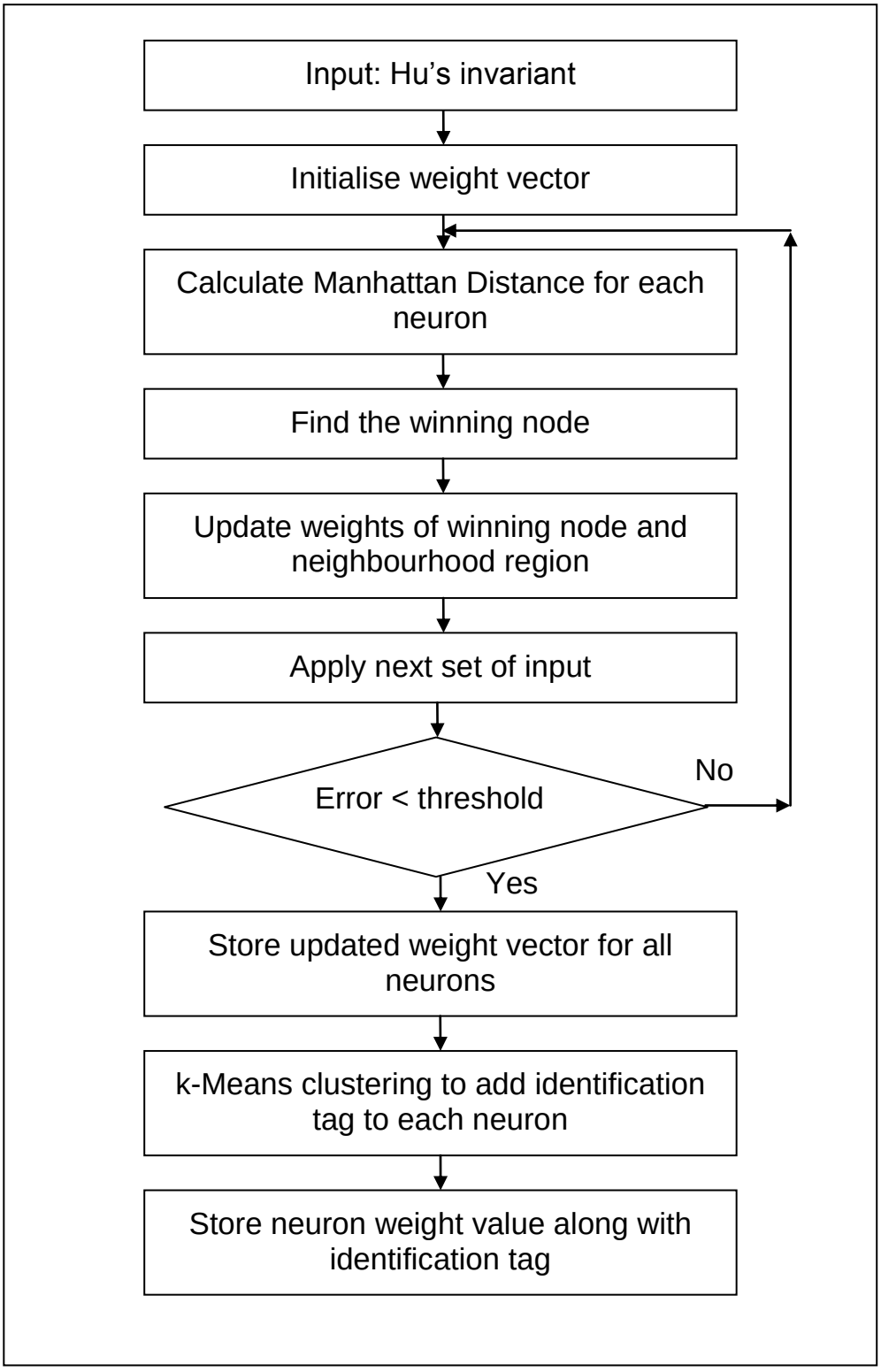

Figure 4.3: Algorithm for Kohonen training and clustering

The training of the Kohonen ANN is currently implemented in MATLAB but the generated weight vectors are used in hardware. The proposed system is organised as follows. The Kohonen neural network will self-organise around the descriptors. The map that is produced is then stored onto non-volatile flash memory, so that it may be used later during the classification. The algorithm for the training part is shown in figure 4.2. A 2-dimensional neuron map is mapped with a random weight initialisation (The random weights are between 0.45 and 0.55). Seven weights are associated with each neuron. The input parameters are then fed to the ANN to find the Manhattan distance (Equation 4.1) for each neuron. The Manhattan distance is given as:

Manhattan distance ("L_1 Norm"): $\sum_{i=1}^{k} | x_i - y_i |$ (eqn. 4.1)

The neuron corresponding to the minimum distance is the winning node. The weights of the winning node and the neighbourhood region around the winning node are updated with the following update rule:

$$w_{ij}(n+1) = w_{ij}(n) + \lambda(n)(x_i(n) - w_{ij}(n)) \qquad \text{for } 0 \le i \le N-1 \qquad \text{(eqn. 4.2)}$$

Where $w_{ij}(n+1)$ is the updated weight, $\lambda(n)$ is the learning rate and the term $(x_i - w_{ij})$ represents the error. The value of the learning rate normally lies between 0 and 1 and it controls convergence speed and stability. The learning process is iterative and continues until the network has converged satisfactorily.

To identify the different sets of neurons, a k-Means clustering algorithm is applied and one identification tag is attached to each neuron. The Manhattan distance measurement is applied again for clustering purpose. The k-Means clustering algorithm has been discussed in chapter 3.

4.4 Classification Mode

In classification mode, the weight vector map is first recalled from flash memory. Incoming patterns are stored onto the high speed SRAM in order to increase performance. The new pattern is then matched against the patterns stored on the Kohonen map and a final classification is produced. As described earlier a set of eight parallel neurons have been used for the detection of an unknown object. After obtaining the moment invariants of the unknown images it is passed to the inputs of the neurons. The Manhattan distance is then measured for each neuron. The minimum distance is now detected and the associated neuron with the minimum distance is declared as the winning node. The identification tag is then simulated and corresponding outputs are activated. Figure 4.4 represents the algorithm for classification mode.

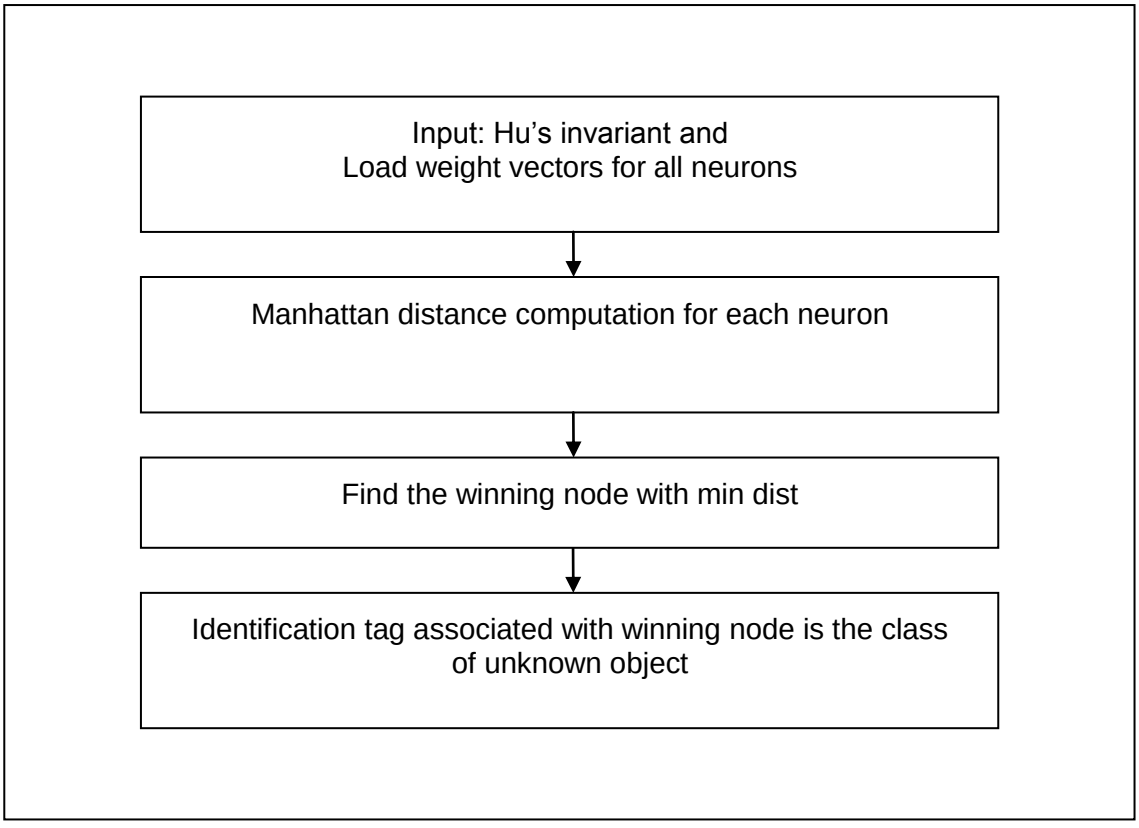

Figure 4.4: Algorithm for classification mode

The complete system response and timing analysis is discussed in the results and discussion section. It has been observed that the response time for the system yields a good result due to the parallel design.

4.5 General Discussion on VHDL Implementation of Algorithm

The project has been coded in VHDL. VHDL is a very powerful high level hardware description language. Some problems have been addressed during the code implementation. These are described below for future reference.

- VHDL code normally generates all parallel hardware unlike a sequential C / C++ compiler. For implementation of a sequential algorithm, it is necessary to handle with the system clock information. A clock transition can control the sequential operation.
- Software engineers probably love to play with floating point numbers. On the contrary a hardware engineer does not like to handle floating points. IEEE 754 numbering formats are followed to implement a floating point number. As off now a floating point calculation is not included in VHDL IEEE library. As a result arithmetic operations need to be designed separately.

  In this project, maximum calculations are based on floating point values. Altera Megacore library provides some of the floating point arithmetic like addition, subtraction, multiplication etc. The project used a floating point division code from RARE project [19]. Other floating point operations like integer to floating conversion etc. are implemented by the author.
- Use of floating point library from Altera (specially Multiplication DSP block) demands more logic elements (LE). Currently the code can not be compiled fully in a Stratix board. In stead we used Stratix II environment. An optimisation of the code and separate floating point multiplication

implementation will be useful to accommodate and compile the full code in a Stratix board only.

- In some cases the same hardware block can be used instead of generating the hardware every time. This is mainly possible where a sequential logic works.

4.6 Explanation of VHDL Code

In this part of the thesis the implemented VHDL code has been explained briefly. For the testing purpose, images have been predefined in the code. The following part of the code has defined a 20x20 matrix of image.

```
----
        pic1(0) <=      "0001111111111111";
        pic1(1) <=      "0011111111111111";
        pic1(2) <=      "0010111111111111";
        pic1(3) <=      "0011111111111111";
        pic1(4) <=      "0011111111111111";
        pic1(5) <=      "0011111111111111";
        pic1(6) <=      "0011011111111111";
        pic1(7) <=      "0011111111111111";
        pic1(8) <=      "0111111111111111";
        pic1(9) <=      "0010111111111111";
        pic1(10) <=     "0011111111111111";
        pic1(11) <=     "0011101111111111";
        pic1(12) <=     "0011111111111111";
        pic1(13) <=     "0011111101111111";
        pic1(14) <=     "0011111111111111";
        pic1(15) <=     "0111111111111111";
----
```

The initial calculation is based on the black pixels available on the white background. Though normally white pixels are represented by '1' and black by '0' , here black is represented by '1' and white is '0' for testing purpose. The following part of the code has shown the sequential operation of the initial calculation and accumulation of moment values. The sequential operation is handled by the system clock:

```
process(clk,en_signal,count,en_count,x1,x2,x3,y1,y2,y3,m11,m12,m13……)
begin
        if(clk'EVENT AND clk='1') then   --the calculations are based on clock cycle
                en_count<=en_count+1;
```

```
                    if(y<16)then
                    if(x<16)then

                    if(pic1(y)(x)='1' AND en_count=1)then
                            en_signal<='1';
                            count<=count+1;         --1 clock cycle is simulated in 10ns

--Calculations start here || Pipeline architecture used here

                            y1<=y;
                            x1<=x;


--First phase data update, assumed clock cycle 4

                            if(count=1)then
                                    y2<=y1*y1;
                                    x2<=x1*x1;

                                    m11<=m11+1;
                                    m12<=m12+y1;
                                    m21<=m21+x1;
                                    m22<=m22+x1*y1;

                            end if;

--2nd phase data update, assumed clock cycle 8

                            --------

--3rd phase data update, assumed clock cycle 8

                                    ----------------

                            end if;

                    end process;
```

The calculation so far has been done with integer number. To calculate eta and Hu's invariant, we need to have floating point number. A function called *int2fp*.vhd converts integer to 32 bit single precision floating point.

```
--------------------------------------------------------------------------------
-       Integer to IEEE 754 Floating Point format conversion
--------------------------------------------------------------------------------
ENTITY int2fp IS

PORT (input : integer range 0 to 100000;

        fp_op : OUT STD_LOGIC_VECTOR(31 downto 0));
```

```
END int2fp ;


ARCHITECTURE Behavior OF int2fp IS

signal int_ip : STD_LOGIC_VECTOR(31 downto 0);
signal sign_bit : STD_LOGIC ;
signal exponent : STD_LOGIC_VECTOR(7 downto 0);
signal mentissa : STD_LOGIC_VECTOR(22 downto 0);

BEGIN

        int_ip<=(conv_std_logic_vector(input,32));

        sign_bit<='0';

        process(int_ip)
        begin
                if(int_ip(31)='1')then
                        exponent<=X"9E";
                        mentissa<=int_ip(30 downto 8);

                elsif(int_ip(30)='1')then
                        exponent<=X"9D";
                        mentissa<=int_ip(29 downto 7);

                ----------
                ----------

                elsif(int_ip(1)='1')then
                        exponent<=X"80";
                        mentissa(22)<=int_ip(0);
                        mentissa(21 downto 0)<=(OTHERS=>'0');

                elsif(int_ip(0)='1')then
                        exponent<=X"7F";
                        mentissa(22 downto 0)<=(OTHERS=>'0');

                end if;

        end process;

fp_op(31)<=sign_bit;
fp_op(30 downto 23)<=exponent;
fp_op(22 downto 0)<=mentissa;


END Behavior;
```

The next part of the code implements the eta calculation and Hu's moment invariants. The code can be found in the appendix.

The next part of the code is for classification mode. All parallel neurons are implemented in VHDL:

```
------------------------------------------------------------
--                    COMPUTATION FOR NEURON 1              --
------------------------------------------------------------
        addr1_1<=X"000";
        addr1_2<=X"001";
        addr1_3<=X"002";
        addr1_4<=X"003";
        addr1_5<=X"004";
        addr1_6<=X"005";
        addr1_7<=X"006";
        im_type(0)<=X"007";


        mem_read1_1: rom
        port map (addr1_1,clk,mem1_1);
        mem_read1_2: rom
        port map (addr1_2,clk,mem1_2);
        mem_read1_3: rom
        port map (addr1_3,clk,mem1_3);
-------
-----
user_add8_6: mega_add
        port map
(clk,add8_27,add8_46,dummy_sub_nan,dummy_sub_overflow,distance(7),dummy_sub_underflow);

------------------------------------------------------------------
--                    CALCULATION END FOR NEURON 8
------------------------------------------------------------------
```

Finally the shortest distance has been found out in order to find the winning node and the corresponding class. The following part of the VHDL code performs that operation.

```
        process(clock,distance(0),distance(1),distance(2),distance(3),distance(4),distance(5),distance(6),dis
tance(7))
                begin

                if(clock'EVENT AND clock='1') then      --the calculations are based on clock cycle
                                count2<=count2+1;
        -------
        --------

                                        temp<=distance(0);
                                        index <= 0;
```

```
                    end if;

                    if (count2 = 21) then
                            if (temp > distance(1))then
                                    temp<=distance(1);
                                    index <= 1;
                            end if;
                    end if;
----------------------
----------------------
                    if (count2 = 26) then
                            if (temp > distance(6))then
                                    temp<=distance(6);
                                    index <= 6;
                            end if;
                    end if;

                    if (count2 = 27) then
                            if (temp > distance(7))then
                                    temp<=distance(7);
                                    index <= 7;
                            end if;
                    end if;
                    end if;

            end process;

            mem_read_final: rom
            port map (im_type(index),clock,test1);
```

4.7 Discussion on MATLAB Coding of Training Algorithm

The training of the system has been implemented with MATLAB coding. A database of seven Hu's moments is created using different predefined images. More the 75 images have been used for training purpose. This data base has been called in the MATLAB program and the training procedure has performed. Finally the updated weight values are stored in a separate database. The updated weight values of the trained neurons have been used for the classification mode which has been implemented in VHDL. The Kohonen Neural network updates the weight for neuron and clustering algorithm gives identification of different classes.

4.8 Summary

This chapter mainly discussed about the algorithmic implementation of the project. As a proof of concept predefined images has been used in place of a real image. Probably the next phase of the project will overcome this and deal with real images. Few problems have been addressed during the VHDL implementation. The success of the concept has been proved in the result and discussion chapter.

Chapter 5: Discussion

5.1 Introduction

So far the thesis described different aspect of the project namely literature survey, theory, implementation etc. This chapter shows the result of the system. It also analysed whether the concept is pragmatic. Different pros and cons of the exploration and implementation are also reflected here.

5.2 Result & Timing Analysis

The algorithm has been implemented in synthesizable VHDL code and a timing analysis has been performed. The simulation results exhibit a very good system timing performance, and the results indicate that the concept design fulfils the requirement for real time object recognition. As the algorithm is divided in three parts, the thesis has also analysed the result separately.

5.2.1 Moment Calculation

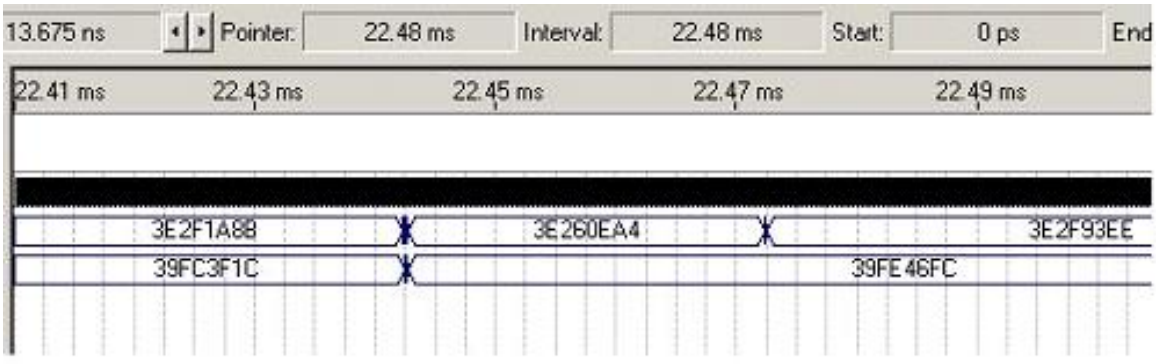

Figure 5.1: Timing analysis of Moment Calculation

The timing analysis of Hu's moment invariant calculation is shown in figure 5.1. The processing of input data consumes the most time in this part. For a 20x20 pixel input image it took an average of 22.47 milliseconds whereas the rest of the computation of Hu's invariant required approximately 1.5 milliseconds to complete (refer to table 5.1).

This is due to the number of sequential operations performed. If the input variables can be processed in parallel, the time response could be significantly improved. The floating point calculations, however, have been implemented efficiently as can be seen from the results.

5.2.2 Training of the System

```
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Basic image for class I

```
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Basic image for class II

Figure 5.2: Predefined test images for two different classes

As said earlier, the training of the system has been performed through a MATLAB programme. During the training phase, Hu's seven moment descriptors have been taken as the input vectors and calculated accordingly. In this project two types of pre-defined images (shown in figure 5.2) have been taken for the training purpose. Around 75 different images of these two classes have been prepared for the training purpose. It has been observed that the classification stage has an accuracy of 95% (with a classification of 20 unknown images of similar kind).

5.2.3 Classification Mode

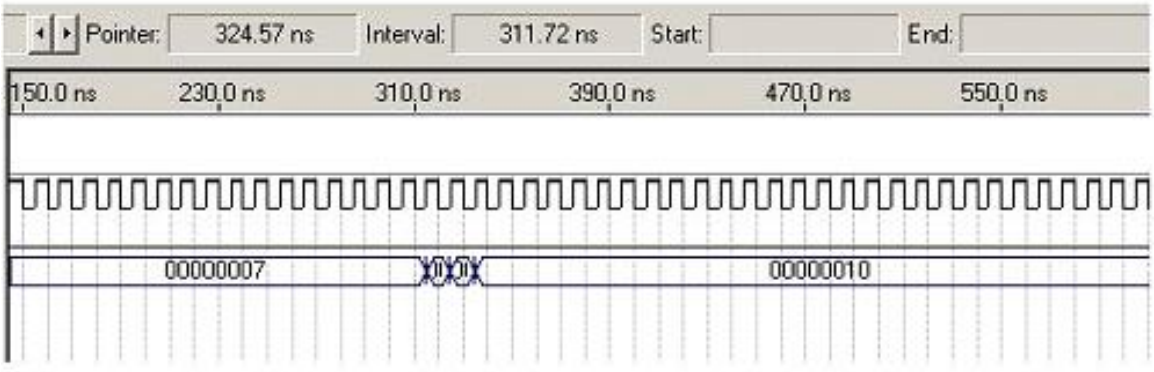

Figure 5.3: Timing analysis of Object Classification mode

The object classification mode indicates a good response time with parallel operation of trained neurons. For this application, eight neurons are designed to run in parallel, providing a response time of 310-360 nanoseconds, as shown in figure 5.3.

5.3 Analysis & Discussion for Real Time Object Classification

The timing analysis shows that the system requires an average of 24 milliseconds (average) for the detection of an unknown object. This indicates that even with the sequential computations in parts of the design, a recognition rate of 40 fps (frames per second) is achievable.

	Hu's Moment			Classification Mode
	i/p cal	other calculation	total	
Time	23 ms	1.5 ms	24.5 ms	310-360 ns

Table 5.1: Average time for different calculations

5.4 Summary

This chapter discussed about the result obtained during the experiment. It also discussed the exploration of the new concept. As a proof concept this project has shown a commendable result with a good future prospect with the given approach. A continuation of the similar work will be able to make a single board FPGA system for a real time image recognition classification with a very good time response.

Chapter 6: Conclusion and Future Work

6.1 Introduction

This chapter concludes the thesis and discusses about the possible future work that may be carried out. This thesis has addressed various aspects of the project. The introduction chapter provided an overview of the work and the motivation of the project. The following sections describe the theoretical aspects and the algorithm development. Finally the result and discussion completed the work.

6.2 Conclusion

In this thesis, we have demonstrated a possible solution of classifying objects in real-time using an FPGA solution. Object features are represented using Hu's moment invariant method, and an onboard Kohonen artificial neural network is used for initial clustering. This implementation uses the Altera Stratix FPGA device, which has the capability of performing complex mathematical operations. The estimated time to recognise an unknown object (20x20 pixels) is 24 milliseconds. We have shown that the classification of objects can be achieved using a single Altera Stratix board. The use of multiple FPGA boards working in parallel, however, can enhance system performance. At present, the input pixels to the system and the initial moment calculations are performed sequentially. As this is the most time consuming aspect of the process, overall system performance can be further improved by using a parallel implementation.

6.2.1 Why FPGA

The motivation of the project was to build an image recognition and classification system with a very good response time. A dedicated and customised hardware is always preferred in place of a software running in generalize computer. Obviously a stand alone and separate equipment for this purpose makes the whole system more robust and reliable, hence the idea of using reconfigurable hardware. The FPGAs offers a solution to the need for reconfigurable hardware. Normally FPGAs are more cost effective than designing a custom made ASIC. Also FPGAs can be reprogrammed to fix a bug or upgrade a system. Hence FPGAs are a natural selection of hardware designer.

6.2.2 Implementation of ANN on FPGA

Artificial Neural Networks are a powerful tool for machine intelligence. The ANN is based on the principle of human neuronal system. The main feature addressed by ANNs is the ability to operate in parallel. This feature of the ANN speeds up the system as compared to a traditional sequential computer. A true parallel system can never be realised in a general purpose sequential computer. The only solution is to use a neural-processor or design a dedicated parallel hardware. Hence the idea of use of a FPGA comes again. Hardware designer can easily accommodate a truly parallel hardware inside a FPGA.

6.2.3 Potential Pros & Cons

The project described has tried to explore the opportunities of image recognition system in FPGA board. During the experiment a few pros and cons have been found out.

Probably a FPGA solution is one of the best possible solutions to address the problem. FPGA can be programmed easily with high level hardware description language like VHLD, Verilog, System C *etc.* A block diagram design

can also be useful. The FPGA operation can be simulated after the compilation of the program. It is easy to debug and test the system with simulation. A complete timing diagram gives the opportunity for analysis of the system.

However some problems exist. Apparently simple arithmetic operations might be complicated when implanted on FPGA. As an example, the use of floating point operation will make the FPGA system complex. In comparison with ASICs, FPGAs are generally slower. An interconnect delay sometimes plays a significant delay in overall system performance. So far the best available FPGAs have their system clocks rated in MHz where as a simple computer has system clocks running in the GHz range.

Now-a-days very advanced FPGAs are available in the market. But still numbers of logic elements are not sufficient enough to address certain problems. For example we have finally simulated the project in the Stratix II environment instead of available Stratix board. By optimising further, we may be able to accommodate the system on a single Stratix board. In the case of onboard online training which is proposed as future work, this limitation may play a vital role and good optimisation is required in order address this factor.

6.3 Future Work

The main aim of the project was to build a stand alone system for image recognition using a single FPGA board. But due time constraints some of the targetted features could not be implemented or explored during this project. Hence some future work has been proposed for the next stage of advancement.

6.3.1 Online Training

So far the *training* of the system has been performed using a MATLAB program. In future work, the onboard and online training of the Kohonen ANN will be addressed. The updated weight vectors of the neurons will be stored in

the memory of the FPGA board, and will be made accessible during the classification mode.

6.3.2 System Integration with Camera Interface

The camera interface is currently simulated by uploading image data directly to the FPGA board. This has been successful as a proof of concept. But in order to deploy the system in a real working environment a camera interface is required in order to enable it to grab and classify images in real time. The Stratix FPGA development board has a number of interface options through high speed A/D converter or a RS 232 serial interface.

6.4 Summary

The current constraint of the system is the capacity of the FPGA board. By using a larger FPGA, the number of neurons can be increased and hence more parallel operations can be performed. Further code optimisation is also possible and some of the implemented sequential operations can be replaced by parallel hardware design. The use of multiple boards for a single system will also enhance the system performance. We have demonstrated that it is possible to build a full imaging system on a single FPGA board. As a proof of concept, the VHDL simulation has been conducted using the Altera Quartus II software environment. Analysis from timing diagrams has shown promising results..

Reference and Bibliography

[1] M-K. Hu, "Visual pattern recognition by moment invariants", IRE Trans. on Information Theory, vol 8, pp. 179-187, 1962.

[2] S. Hirai, M. Zakouji and T.Tsuboi, "Implementing Image Processing Algorithms on FPGA-based Realtime Vision System", Proc. 11th Synthesis and System Integration of Mixed Information Technologies (SASIMI 2003), pp. 378-385, Hiroshima, April, 2003.

[3] P.C. Arribas and F.M-H. Maciá, "FPGA implementation of Santos-Victor optical flow algorithm for real time image processing: an useful attempt", Proceedings of SPIE's International Symposium on Microtechnologies for the New Millennium 2003 - Conference on VLSI Circuits and Systems, pp. 23-32, Canary Islands, Spain, May 19-21, 2003.

[4] P.J. Sanz, R. Marin and J.S. Sanchez, "Including efficient object recognition capabilities in online robots: from a statistical to a Neural-network classifier," IEEE Transactions on Systems, Man and Cybernetics, Part C, vol.35, no.1, pp. 87- 96, February, 2005.

[5] T. Kohonen, "Automatic formation of topological maps of patterns in a self-organizing system", Proceedings of 2nd Scandinavian Conference on Image Analysis, Espoo, Finland, pp. 214--220, 1981.

[6] S. Neema, J. Scott, T. Bapty, "Real time reconfigurable image recognition system", Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference, 2001. Volume 1, 21-23 May 2001, pp. 350 - 355 vol.1

[7] F. Mohd-Yasin, A.L. Tan, M.I. Reaz, "The FPGA prototyping of iris recognition for biometric identification employing neural network", Proceedings of The 16th International Conference on Microelectronics, ICM 2004. 6-8 Dec. 2004, pp. 458 – 461

[8] J. Jean, Liang Xiejun, B. Drozd, K. Tomko, "Accelerating an IR automatic target recognition application with FPGAs", Proceedings of Seventh Annual IEEE Symposium on Field-Programmable Custom Computing Machines, FCCM '99. 21-23 April 1999, pp. 290 – 291

[9] N. Aibe, M. Yasunaga, I. Yoshihara, J.H. Kim, "A probabilistic neural network hardware system using a learning-parameter parallel architecture", Proceedings of the 2002 International Joint Conference on Neural Networks, IJCNN '02., Volume 3, 12-17 May 2002, pp. 2270 – 2275

[10] Yuk Ying Chung, Man To Wong, Neil W. Bergmann, "High speed neural network based classifier for Real-time application", Proceedings of ICSP '98, pp. 506 – 509

[11] Dr. R. Saatchi, Handout given in classroom & laboratory session (16-7206-00S Artificial Intelligence), February - April 2006, Sheffield Hallam University

[12] http://en.wikipedia.org/wiki/IEEE_754 last visited on 14.09.2006

[13] Hsiao-Fen Fu, IEEE 754 Floating Point, CSCI 313 Tutorial, http://ftp.csci.csusb.edu/schubert/tutorials/csci313/w04/HsiaoFenFu_Tutorial_IEEE%20754%20Floating%20Point.pdf last visited on 14.09.2006

[14] http://en.wikipedia.org/wiki/FPGA last visited on 14.09.2006

[15] http://en.wikipedia.org/wiki/K-means_clustering_algorithm last visited on 14.09.2006

[16] http://en.wikipedia.org/wiki/Image_moments last visited on 14.09.2006

[17] A. Ashbrook and N. A. Thacker, Tutorial: Algorithms for 2-Dimensional Object Recognition, Tina Memo No. 1996-003, Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester

[18] P.J. Sanz, R. Marin and J.S. Sanchez, "Including efficient object recognition capabilities in online robots: from a statistical to a Neural-network classifier," IEEE Transactions on Systems, Man and Cybernetics, Part C, vol.35, no.1, pp. 87- 96, February, 2005.

[19] http://www.imappl.org/~cgloster/rare/vhdl/ last visited on 14.09.2006

[20] Timothy Masters, "Signal and Image Processing with Neural Networks, A C++ Sourcebook", John Wiley & Sons, Inc., New York, 1994

[21] William Kleitz, Digital Electronics with VHDL quartus II Version, Prentice Hall, 2006, New Jersey