A VLSI Neural Network for Morphology Classification

Philip H. W. Leong and Marwan A. Jabri
Sydney University Electrical Engineering, NSW 2006 Australia

Abstract

An architecture for a low power analogue VLSI implementation of a three layer perceptron is presented. The chip is designed specifically for the classification of signals obtained from the internal surface of the heart. The neural network is composed of synapses which are implemented as multiplying digital to analogue converters, and neurons which are high resistance off-chip resistors. A test chip with a (3,3,1) neural network and bucket brigade has been fabricated and results from this test chip are presented.

1 Introduction

A typical intracardiac electrogram (ICEG) signal is shown in Figure 1. The interesting features of the signal lie in the periodic sharp peaks of the ICEG, called QRS complexes. Changes in shape (also called morphology) of the QRS complexes can indicate abnormal heart conditions (called arrhythmias), and in Figure 1, the signal can be seen to change from a normal morphology to a widened signal with a different shape as the patient experiences an arrhythmia.

The aim of the VLSI neural network chip is to provide real-time classification of these signals. As the intended application of the chip is in an implantable device, power consumption must be kept to an absolute minimum, a figure of approximately $20\mu W$ being desired.

2 Implementation

A block diagram of the proposed chip architecture is shown in Figure 2. Input is obtained from a cathetor inserted in the internal wall of the heart. From this probe, the electrical activity of the heart can be measured. In order to obtain a time sequence of samples from this probe, a bucket brigade device (BBD) is used as an analogue shift register, the output taps of the BBD being the inputs to the three layer perceptron. An off-chip "QRS detector" circuit which detects the onset of the QRS complex by comparing its derivative with an adaptive threshold is used to synchronize the neural network with the QRS complex. The period of QRS complexes is approximately $1Hz$ and power can be saved by switching the neural network on only when the QRS complex is centered in the neural network.

2.1 Synapse

Storage of the synapse values is achieved using flip-flops which are converted to analogue values via digital to analogue converters (DAC). This allows analogue signal processing techniques to be used whilst maintaining the advantages of digital weight storage. The DAC is constructed through current summing. Each bit of the DAC is controlled by a transmission gate which can be turned on or off depending on a value stored in the input (static) flip-flop (see Figure 3). The entire synapse array appears as a large (write only) register to the controlling digital circuitry which programs the weight values. Each bit of the DAC consists of a current mirror transistor (I0–I4), a pass transistor (B0–B4) and a flip flop.

The current source (Figure 4) is constructed by summing unit current sources. For transistors with uncorrelated matching properties, this improves the matching by a factor of $\sqrt{N}$. Correlated matching

0-7803-0559-0/92 $3.00

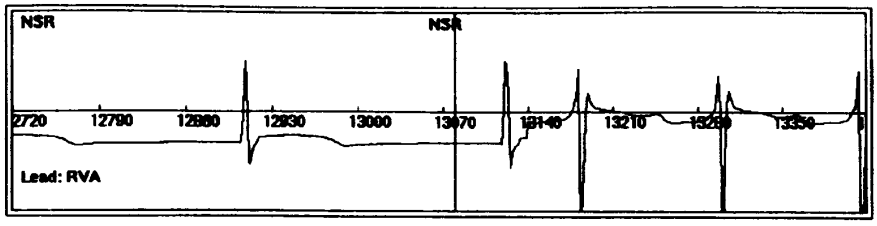

Figure 1: Typical Intracardiac Electrogram. Note changes in shape of the QRS Complex.

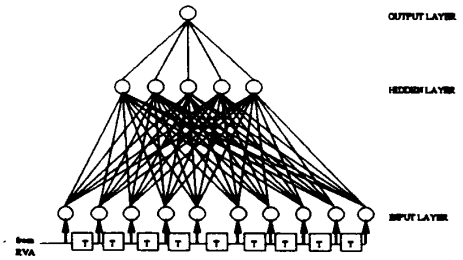

Figure 2: Block Diagram of Neural Network Chip.

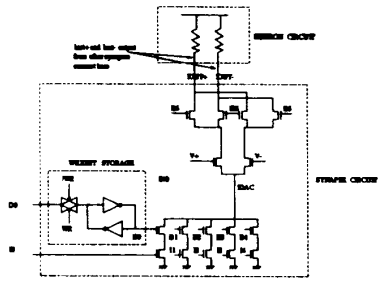

Figure 3: Synapse and Neuron Circuitry.

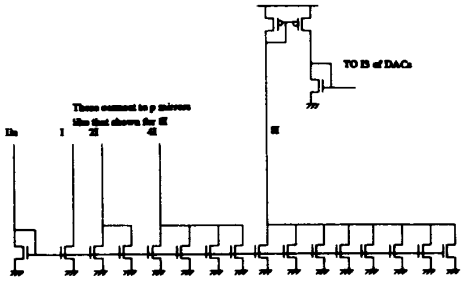

Figure 4: Current Source Circuitry (4 bits shown)

properties such as changes in doping or oxide thickness are addressed by arranging the current sources in a common centroid configuration [1]. Large ($553\mu m^2$) transistors are used for the current source although smaller ($81\mu m^2$) transistors are used inside the DACs in order to keep the total synapse area small. The input bias current Iin was set to a value of $5nA$ for all simulations and experiments.

The DAC is connected to a Gilbert multiplier to form a synapse (Figure 3). The multiplier has a pair of voltage inputs, a pair of current inputs (from the DAC) and a pair of current outputs. The transfer function of this multiplier is given by the relation

$$I_{out+} - I_{out-} = \begin{cases} +I_{DAC} \tanh(\frac{\kappa(V_3 - V_4)}{2}) & \text{if B5} = 1 \\ -I_{DAC} \tanh(\frac{\kappa(V_3 - V_4)}{2}) & \text{if B5} = 0 \end{cases} \quad (1)$$

The multiplier is linear with the current inputs (from the DAC) and nonlinear to the neuron voltage inputs. This is the desired situation as if they were reversed, nonlinear problems could not be solved. If the tanh function were in the weights, it would only serve to compress the range of weight values allowed and would not allow nonlinear problems to be solved. The DAC only produces positive values. Current switching logic controlled by B5 enables the output to be changed in sign if a negative weight is desired. The V3 and V4 inputs are from either neurons or input pins. Output of the multiplier are two current sinks. The area of the synapse is $106 \times 113\mu m$.

2.2 Neuron

The neuron is simply a $1M\Omega$ resistor connected to the supply voltage (Figure 3). This configuration allows the current sink output of the synapse to produce a voltage proportional to its output, and two are used to produce the differential voltage output (Figure 3). The resistors are provided off chip in order to allow easy control of the impedance and transfer characteristics of the neuron. Static RAM processes offer a high resistance polysilicon which would enable such neurons to be placed on chip. These neurons also serve as test points for the chip.

3 Test Chip

A network which uses the above building blocks has been designed using Orbit Semiconductor's $1.2\mu m$ double metal, single poly nwell process. The test chip contained a 10 stage bucket brigade device and a neural network with 3 input units, 3 hidden units and 1 output unit (3,3,1). Other circuitry includes a set of row and column shift registers used to select a synapse for writing, test blocks which consist of single synapses (used to measure the monotonicity of the MDAC) and the current source circuitry. Figure 5 shows a plot of the chip, major blocks being labeled.

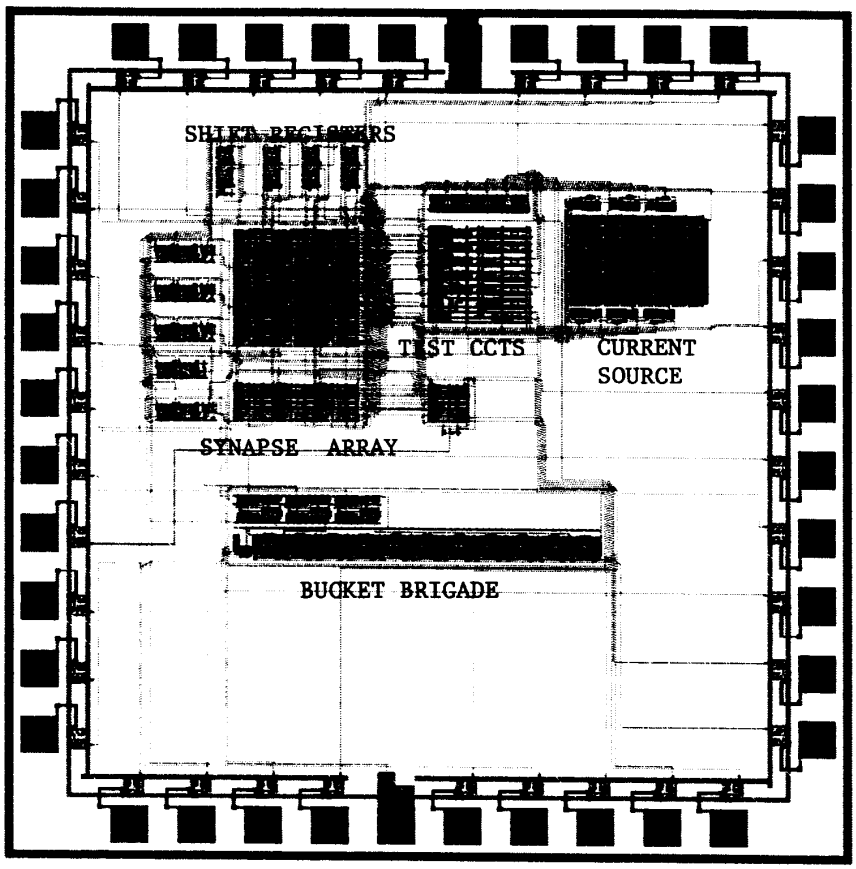

Figure 5: Neural Network Test Chip (3,3,1)

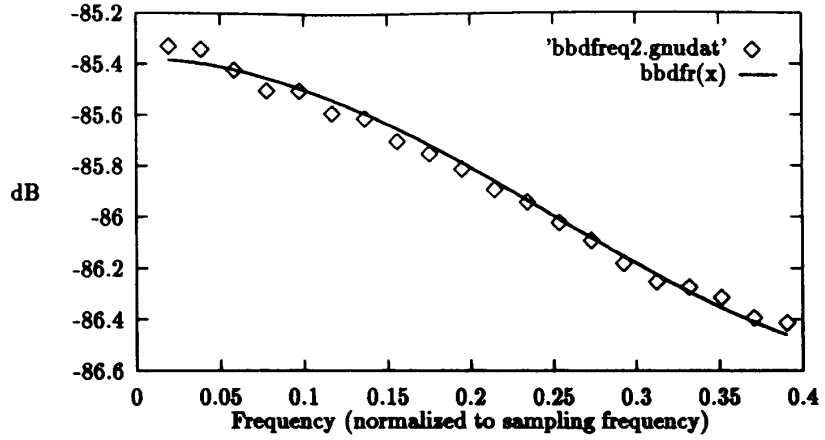

Figure 6: BBD Frequency Response. Dotted line represents theoretical frequency response for CTI = 0.35%.

3.1 Bucket Brigade Device

Incomplete charge transfer in the BBD causes the device to have a nonideal frequency response of [2]

$$H(f) = [(\frac{1-2\alpha-2\beta}{1-2\alpha e^{-j2\pi f/f_s}})e^{-j2\pi f/f_s}]^n \tag{2}$$

where n is the number of stages in the BBD, f_s is the sampling frequency, α is the charge transfer inefficiency and β is a constant which is considered zero. The charge transfer inefficiency of the BBD (configured as a simple BBD) was obtained by measuring the frequency response (sampling frequency $125Hz$) of the BBD and curve fitting it to the theoretical response (see Figure 6). From this, the charge transfer inefficiency was found to be 0.035%. As this figure was adequate for our purpose, the tetrode configuration was not tested.

3.2 Neural Network

A plot of the measured synapse linearity (obtained by sweeping through the weight values of a test synapse) is shown in Figure 7. The "bumps" in the transfer characteristic are caused by the binary encoding of the DAC.

The neural network was tested by solving the XOR problem. This problem can only be solved by a nonlinear system and is thus a good problem with which to test the chip.

Two inputs plus a bias input are used. The test chip was controlled using a custom test jig interfaced to a PC-LabCard "Multi-Lab Card" which provides 12 bit D/A, A/D and digital IO. The starting weight condition were obtained using a training based on optimization of a series of increasingly sophisticated models of the circuit [3], starting from a floating point mathematical model and finishing with a Spice simulation of the circuit. The chip was then trained (in-loop) using Nelder-Mead optimization of the error function until no improvement was seen. The test chip can only be trained using this method if the initial weights of the final training are sufficiently close to the final solution. The problem of training the chip directly from random weights (without the heirarchy of models) has not yet been addressed. The final solution reached by the test chip is shown in Table 1.

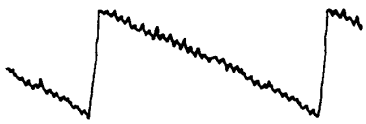

Figure 7: DAC Transfer Function (Measured). Amplitude is $160mV$ pp.

Input			*Desired Output*	Output
0.0	0.0	0.2	-0.1	-0.153809
0.2	0.0	0.2	+0.1	+0.134277
0.0	0.2	0.2	+0.1	+0.119629
0.2	0.2	0.2	-0.1	-0.090332

Table 1: Results of Applying Test Chip to the XOR Problem

4 Conclusion

The design of an analogue VLSI chip for heart signal recognition was presented. By using multiplying digital to analogue converters as synapse elements, multilayer perceptrons can be constructed which feature both a simple design and low power consumption. A test chip with a (3,3,1) neural network (12 synapses and 4 neurons) was successfully applied to the XOR problem.

5 Acknowledgement

Work presented in this paper was supported by the Australian Department of Industry, Technology and Commerce, Telectronics Pacing Systems and the Australian Research Council.

References

[1] C.A.A. Bastiaansen, Wouter J. Groeneveld D., H.J. Schouwenaars, and H.A.H. Termeer. A 10–b 40Mhz 0.8μm CMOS Current–Output D/A Converter. *IEEE Joural of Solid-State Circuits*, 26(7):917–921, July 1991.

[2] C.N. Berglund. Analog performance limitations of charge–transfer dynamic shift registers. *IEEE Joural of Solid-State Circuits*, SC–6(6):391–394, December 1971.

[3] P.H.W. Leong and M.A. Jabri. An analogue low power vlsi neural network. In *Proceedings of the Third Australian Conference on Neural Networks*, pages 147–150, Canberra, Australia, 1992.