IMPLEMENTATION OF A LEARNING SYNAPSE AND A NEURON FOR PULSE–COUPLED NEURAL NETWORKS

Pavol Tikovič — Marián Vörös — Daniela Ďuračková *

A new architecture of a learning synapse with on-chip learning and a neuron for pulse-coupled neural networks is presented. The main advantages of the proposed synapse are: continuous learning, easily adjustable parameters, compact design, low power and area consumption. This paper also contains results obtained from simulations performed with our model of a leaky integrate-and-fire neuron. The proposed neuron works with both constant or slowly changing input voltage level (input layer of a neural network) and pulse input (hidden layers and the output layer). Our circuits were designed in HSPICE and implemented in CAMELEON.

Keywords: pulse-coupled neural networks, continuous learning synapse, synaptic weight, Hebbian learning rule, integrate-and-fire neuron, threshold, action potential timing

1 INTRODUCTION

The massive parallelism predetermines artificial neural networks (NN) for use in non-biological applications. Well known are applications like pattern recognition, edge detection, sound preprocessing, etc. A very important question is the size and topology in those applications. A lot of authors discuss these problems in their articles. In our case we try to design a user-configurable system.

The main goal of this work is to model and implement basic features of biological neurons and synapses. The biggest advantage of our neuron is the possibility to use constant or slowly changing input voltage (input layer of a neural network) as well as spiking input (hidden layers and the output layer). We reached a maximum spiking frequency of 165 kHz for a constant input of 3.3 V. All results presented in this paper are obtained from simulations because our neuron was not fabricated on a chip yet. The main advantages of the proposed synapse are: continuous learning, easily adjustable parameters, compact design, low power and area consumption.

Neural networks contain large amounts of connections that highly increase the area consumption. This is the most serious restriction of using neural networks. The best way to solve this problem is using multi-layer integrated circuits or using bulk for transistors fabrication. This technology would allow us to increase the size of the network and of course its capabilities. Unfortunately this technology is not available at present time.

2 BIOLOGICAL SYSTEMS

In biological networks two (pre-synaptic and post-synaptic) neurons are connected through a synapse and they communicate by firing spikes. There are many different types of neurons with a very complex structure. Neurons and their interconnections build a dense net: more than 10^4 neuron cells per square millimeter and some kilometers of interconnections.

Neurons receive input through a dendrite tree. The body of a neuron (soma) is covered by a membrane. This membrane has different types of ionic channels across it. The output part of a neuron is formed by an axon. The axon can branch to many neurons (axonal arborization) and build connections in form of synapses. In humans the axon can be up to 1 m long.

When a neuron is stimulated, the membrane potential rises. There is a threshold value. When this value is reached, an action potential is generated at the axon hillock and propagated along the axon. Immediately after that, the threshold of excitation rises theoretically to infinity and after a small time interval falls down exponentially. During this interval, called the refractory period, the neuron is not able to fire a spike. For more information see [1].

Synapses define the strength of connection between two neurons. Knowledge, in artificial neural networks, is stored in form of the synaptic weight. The learning rule is very important for learning networks. During learning the synaptic weight is changed. In pulse- coupled (PC) NN Hebbian learning rule is used. The weight increases, when synchronized spikes are generated *ie* a spike fired by a pre-synaptic neuron is followed by a spike of a post-synaptic neuron within a small delay called the learning window. Otherwise the weight decreases. In biological neural systems a grow effect in synapses causes changes of efficacy (see [2]).

* Slovak University of Technology, Faculty of Electrical Engineering and Information Technology, Department of Microelectronics, Ilkovičova 3, 812 19 Bratislava, Slovakia, E-mail: tikovic@elf.stuba.sk, voros@pobox.sk, valehrac@elf.stuba.sk

ISSN 1335-3632 © 2001 FEI STU

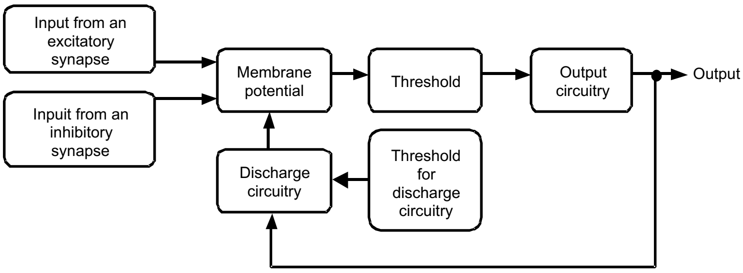
Input from an excitatory synapse
Inpuit from an inhibitory synapse
Membrane potential
Threshold
Output circuitry
Output
Discharge circuitry
Threshold for discharge circuitry

Fig. 1. Block schematic of our integrate-and-fire neuron implementation

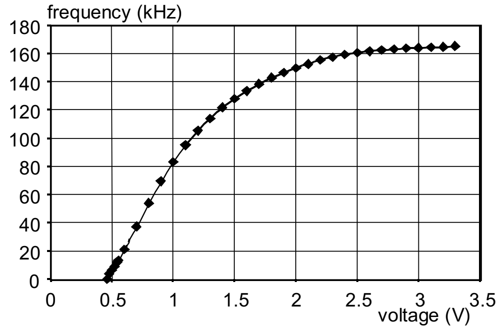

Fig. 2. Relation between spike frequency at the output of the neuron and the constant input voltage level at the excitatory synapse output

3 MATHEMATICAL MODEL OF AN INTEGRATE AND FIRE NEURON

The basic electric circuit of an integrate-and-fire model consists of a capacitor C connected in parallel with a resistor R. If there is a current $I(t)$ that flows through this circuit, it is divided into two parts. One part will charge the capacitor and the second part will flow through the resistor. We will obtain this equation

$$I(t) = \frac{u(t)}{R} + C\frac{\mathrm{d}u}{\mathrm{d}t}\,, \tag{1.1}$$

where u is the voltage across capacitor C. We can define a time constant of a leaky integrator $\tau_m = RC$ and write

$$\tau_m\frac{\mathrm{d}u}{\mathrm{d}t} = -u(t) + RI(t)\,. \tag{1.2}$$

Further we refer to u as membrane potential and τ_m as time constant of a neuron membrane. Equation (1.2) is a first order linear differential equation and cannot describe the behavior of a neuron generating an action potential. To describe the essence of pulse emission we add a threshold condition.

Threshold crossing $u\big(t^{(f)}\big)= \vartheta$ is used to define the firing time $t^{(f)}$. The shape of the action potential is not described explicitly. Immediately after $t^{(f)}$ the potential is reset to a new value of u_r,

$$\lim_{\delta\to 0} u\big(t^{(f)} + \delta\big)= u_r\,. \tag{1.3}$$

For $t > t^{(f)}$ the circuit dynamics is described by (1.2) until a new threshold crossing occurs. The combination of threshold crossing (1.2) and reset (1.3) defines the basic integrate- and-fire model. To show the function of the model we assume a modeled neuron with constant input current I_0 and resting potential $u_r = 0$. We assume the first spike occurred at $t = t^{(0)}$. We can find the trajectory of the membrane potential by integrating (1.2) with the initial condition $u\big(t^{(0)}\big)= u_r = 0$.

The solution is

$$u(t) = RI_0\left[1 - \exp\Big(-\frac{t - t^{(0)}}{\tau_m}\Big)\right]. \tag{1.4}$$

For $RI_0 < \vartheta$ no more spikes can occur. For $RI_0 > \vartheta$ the membrane potential will cross the threshold ϑ at the time $t^{(1)}$ which can be seen from the threshold condition

$$\vartheta = RI_0\left[1 - \exp\Big(-\frac{t^{(1)} - t^{(0)}}{\tau_m}\Big)\right]. \tag{1.5}$$

The solution of (1.5) for time interval $T = t^{(1)} - t^{(0)}$ is

$$T = \tau_m \ln\frac{RI_0}{RI_0 - \vartheta}\,. \tag{1.6}$$

After the spike is fired at time $t^{(1)}$, the membrane potential is reset to $u_r = 0$ and the integration process starts again. With a constant input current I_0 a neuron of the integrate-and- fire type fires regularly with a period of T given by (1.6).

In the next step we add an absolute refractory period. After a spike is fired at time $t^{(f)}$ we clamp the membrane potential to $u_r = 0$ and we hold this value for a time period of δ^{abs}. At time $t^{(f)} + \delta^{abs}$ we continue with integrating (1.2) with an initial value of $u = u_r$. This is a modification of the biological behavior. Other neuron models set the value of the membrane potential u equal to $-K < 0$ or to a value lower than the resting potential (for biological neurons). Our modification is based on the fact that we use only non-negative voltages and the resting potential is set to zero.

Like before, we can derive the dynamics of the model for a constant input current I_0. If $RI_0 > \vartheta$, our neuron will fire regularly. For an absolute refractory period the inter-spike interval will be longer by δ^{abs} compared with (1.6). Usually we determine the average firing frequency as $v = 1/T$, so

$$v = \left[\delta^{abs} + \tau_m \ln\frac{RI_0}{RI_0 - \vartheta}\right]^{-1}. \tag{1.7}$$

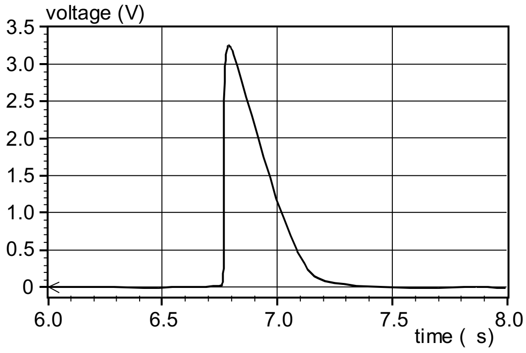

Fig. 3. The shape of a spike generated at the output of our neuron

4 NEURON IMPLEMENTATION

Since in real neural networks synapses are excitatory or inhibitory, we use two different kinds of synapse outputs connected to our neuron. We performed our simulations with one excitatory synapse output and one inhibitory synapse output as it is shown in Fig. 1. The implemented neuron can work with both constant and slowly changing input signals. The value of the membrane potential depends on the input signal. It should be approximately equal to the sum of excitatory and inhibitory inputs. This value is stored as charge on a capacitor. For constant or slowly changing input signals our neuron fires with a frequency given by the input signal level. Figure 2 shows the relation between spike frequency at the output of the neuron and the value of the constant input voltage at the excitatory input.

In a biological neuron the membrane potential decreases to the value of the resting potential when no excitatory spikes arrive. This is achieved by slowly discharging the membrane capacitor. In our case it is a benefit that the parasitic leak-current that discharges the capacitor does not have a bad influence on the function of the circuit. It is obvious that we do not need to add a refresh circuit to refresh the voltage across the membrane capacitor.

The supply voltage of 3.3 V allows low energy consumption of the circuit. Our neuron implementation fires spikes when constant input voltage in the range 0.48 V–3.3 V is connected to the output of the excitatory synapse. Figure 3 shows the shape of the generated spikes. There is no hyper-polarization included here (we assume only nonnegative voltages). Typical action potential in biological neural networks lasts for 1–2 ms [6]. Our action potential lasts approx. 200 ns (half amplitude duration). Amplitude of the spike was chosen to be nearly equal to the supply voltage. Synapses used in artificial neural networks together with our neurons should modify the amplitude of spikes only downwards.

Input voltage lover than 0.48 V is not able to charge the membrane capacitor above the threshold level of 1.5 V, which was the value used for our simulations. There is a minimum frequency for a given amplitude. Spikes can be generated only above this frequency.

For constant input voltage level and for pulse input there is a saturation of the output spikes frequency. The maximum frequency of the output spikes is given by the refractory period.

5 NEURAL CODING

The problem of coding in biological neural networks has not been solved yet. We assume that the signal in a neural network is not only coded by the mean firing rate. Many biological neural systems use the timing of single action potentials to encode information [7, 8]. Results from experiments performed on monkeys show that mainly in the optical nerve there is not enough time to encode information by spike frequency. We performed a simulation to show the relation between the time a neuron needs to fire the first spike (time to first spike, TFS) and the inter-spike interval (ISI).

In biological neural networks the amplitude of spikes arriving at the neuron's membrane is not constant. We performed this simulation only to show the linear character of the relation between inter-spike interval and the time to first spike for a non-changing stimulus.

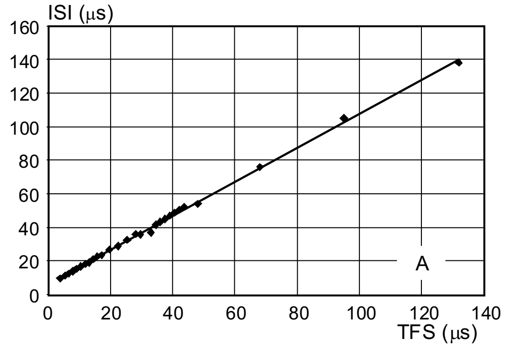

Fig. 4a. The amplitude of input spikes used in simulation was 2 V, input spike frequencies were chosen from the range. 166.7 kHz–3.3 MHz.

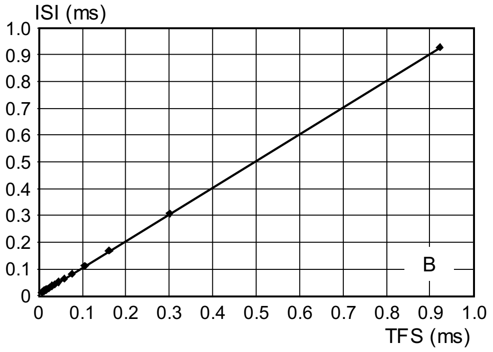

Fig. 4b. The frequency of input spikes used in simulation was 625 kHz, input spike amplitude was swept within the range 1 V–3.3 V (at this frequency, input spikes with an amplitude lower than 1 V are not able to fire the neuron).

We used two different approaches:

- We changed the frequency of input spikes from the excitatory synapse with a constant amplitude (amplitude chosen for our simulation = 2 V)
- We changed the amplitude of input spikes from the excitatory synapse with a constant frequency (frequency chosen for our simulation = 600 kHz)

In Fig. 4 it is easy to see that the relation between ISI and TFS is linear in both cases, which confirms our assumption that the information about the analog signal is encoded in TFS as well as in ISI.

6 MATHEMATICAL DESCRIPTION OF LEARNING

The strength of connections between neurons in the brain is permanently changed according to the Hebbian learning rule. When the axon of cell A has an exciting influence on cell B and it causes cell B to fire repetitively, the efficacy of cell A will rise. We can write this rule as the product of activities of both neurons times the learning rate

$$\Delta w_{ij} = \eta o_i^{k+1} o_j^k$$

where η is the learning rate, o_i^{k+1} is the output of the i-th neuron from the $k+1$-th layer and o_j^k is the output of the j-th neuron from the k-th layer.

Assume that a neuron receives input from $N > 1$ synapses with an efficacy J_i (describing the influence of weight on a signal), where $1 \leq i \leq N$ and the learning window $W(s)$ is a real function. The Hebbian learning rule is described by three basic assumptions:

- t_i^m is the time when the m-th action potential reaches the i-th synapse. The action potential causes a weight change w^{in} that can be positive or negative
- t^n is the n-th action potential at the output of a neuron, that causes a weight change w^{out} that can be also positive or negative
- the time difference $s = t_i^m - t^n$ between an action potential at the input and an action potential at the output of a neuron causes a function change $W(s)$ during the learning window W

If at time t the efficacy is $J_i(t)$, than the difference

$$\Delta J_i(t) = J_i(t+\tau) - J_i(t)$$

during the time interval τ is calculated as a sum of contributions from input and output spikes across the time interval $[t, t+\tau]$. For the spike-train at the input we can write the following equation

$$S_i^{in}(t) = \sum_m \bigl(t - t_i^m\bigr)\,.$$

and similarly for the spike train at the output

$$S^{out}(t) = \sum_n \bigl(t - t_i^n\bigr)\,.$$

Then the efficacy is

$$\Delta J_i(t) = \int_t^{t+\tau} \mathrm{d}t' \Bigl[w^{in} S_i^{in}(t') + w^{out} S^{out}(t') + \int_t^{t+\tau} \mathrm{d}t'' W[t''-t'] S_i^{in}(t'') S^{out}(t') \Bigr].$$

Functions S_i^{in} and S^{out} are called firing rates and τ is the time duration of the learning window.

7 A LEARNING SYNAPSE

We simulated this synapse connected between two neurons in order to know its behavior. We assume that the neuron presented by Ota and Wiliamowski [3] is suitable for our synapse. Figure 5 shows the main idea of this neuron. This neuron works in current mode. The charge stored on capacitor C_1 corresponds with the charge of sodium ions accumulated on the outer side of neuron

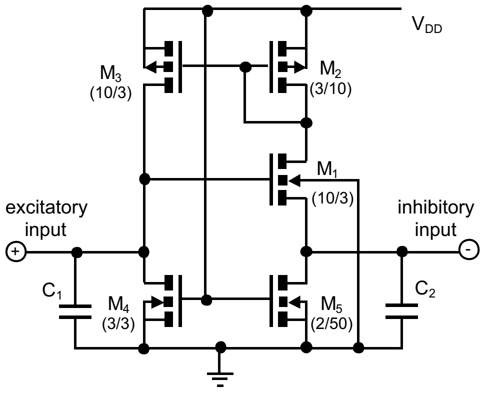

Fig. 5. Principle of the neuron used (redrawn according to [3]

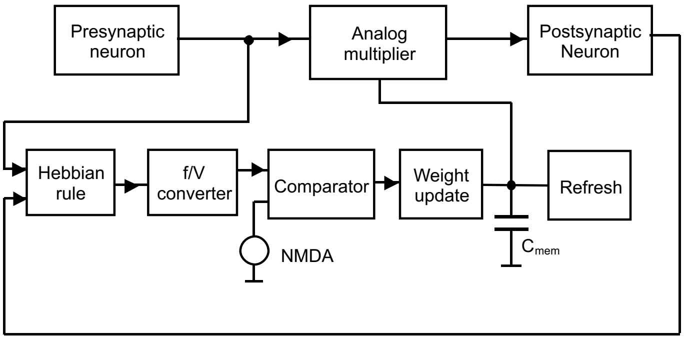

Fig. 6. Block diagram of the designed learning synapse

membrane. The charge stored on capacitor C_2 corresponds with the charge of potassium ions inside the neuron. The time constant C_1R_1 is smaller than time constant C_2R_2, because the potential of sodium ions changes faster than the charge stored on capacitor C_2.

In a stable case transistors $M_1{-}M_3$ are off. When the potential on capacitor C_1 exceeds the potential on capacitor C_2 plus threshold voltage of transistor M_1 (threshold of excitation), transistors M_2 and M_3 that are connected as a current mirror switch to the active region. Positive feedback through transistors $M_1{-}M_3$ causes charging of capacitor C_2 and the potential between emitter and gate of transistor M_1 falls down. This will switch all transistors off. The neuron does not respond to any incoming spikes till the potential on C_1 exceeds the threshold of excitation. Unfortunately, output frequencies of the presented neuron are bigger than frequencies used in biological systems. The response to maximum constant voltage is about 750 kHz (in biological systems it is around hundreds of Hz).

During the learning process synaptic weights are set to get the desired function of NN. In this work the synapse learns during a time interval that is called "learning window". The duration of a learning window was set to 39 μs. If the number of synchronized spikes within the learning window is bigger than the threshold of learning (represented by the NMDA signal), the weight increases. Otherwise it decreases. The weight is modified in both directions in steps of approximately. 15 mV. NMDA and the duration of the learning window are parameters of learning and are adjustable externally.

As a memory element we use a capacitor because floating gate transistors need a high voltage to change the amount of stored charge. It means to design a special circuit which would enlarge our synapse's area consumption. The gate oxide of a floating gate transistor is degraded after a few thousand learning cycles, which is not acceptable for a continuously learning synapse. Certainly we have to design a refresh circuit to hold the charge stored on the memory capacitor. The weight is stored as a charge on the memory capacitor in 256 levels, which corresponds to 8 bit precision. The maximal potential of the weight is 5 V and minimal is 1.2 V. The maximal current through our synapse is 100 μA. We use the maximum supply voltage and therefore we use transistors connected in sub-threshold region as loads. This has reduced the maximal synapse current. Figure 6 shows a block diagram of the presented synapse. It is connected between one pre-synaptic and one post-synaptic neuron.

The learning synapse shown in Fig. 6 has four inputs: activity of pre-synaptic and post-synaptic neuron, NMDA and clock. The period of clock signal defines the weight update rate. The memory capacitor is charged and discharged by current pulses during half amplitude duration.

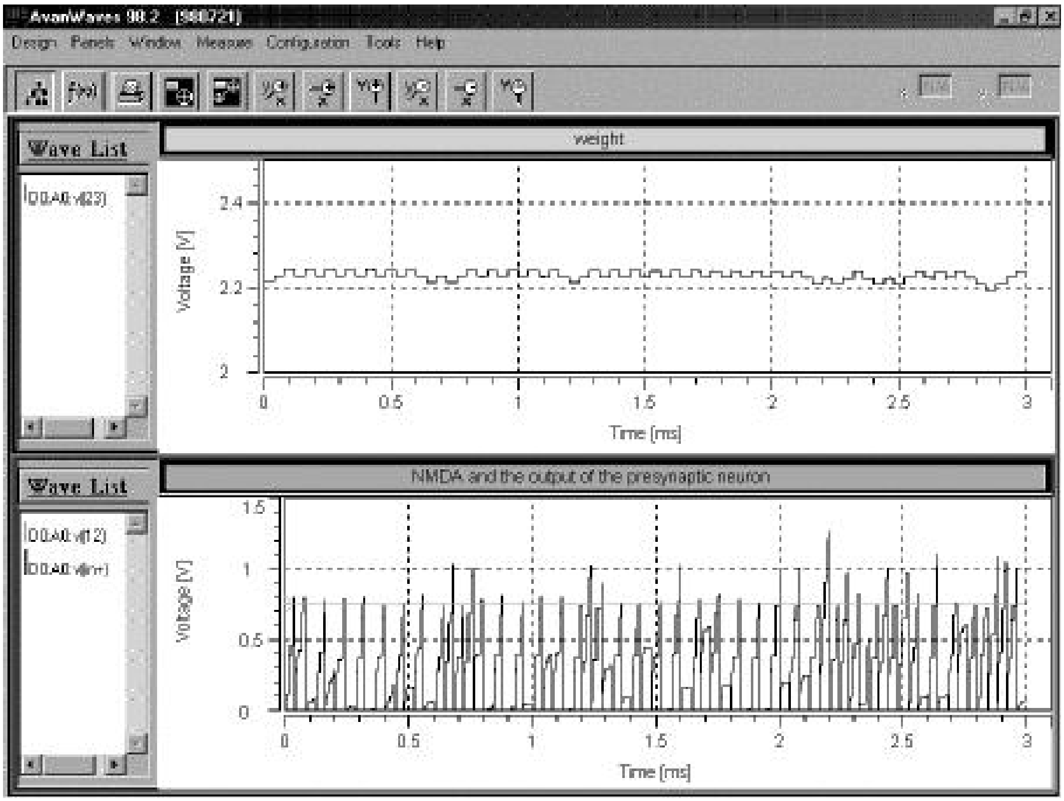

Fig. 7. Simulation results on weight adaptation of a synapse. Pre-synaptic neuron was stimulated with constant voltage level of 5 V. NMDA was set to 0.75 V and the initial voltage across the memory capacitor was 2.2 V.

The function of blocks in the block diagram is explained below.

- **Hebbian rule** is the implementation of the Hebbian learning rule. There is a spike generated at the output of this sub-circuit when the delay between activity of post-synaptic and pre-synaptic neuron is smaller than 350 ns. If there is a bigger delay at the output, spikes are narrower and when the delay is too big (about 700 ns), there are no spikes at the output.
- **f/V converter** converts the spike rate at the output of the Hebbian rule block to the corresponding voltage.
- **Comparator** compares the voltage at the output of the f/V converter with NMDA and determines whether the weight should increase or decrease. The weight increases when voltage from f/V converter is bigger than NMDA.
- **Weight update** changes the weight stored on the memory capacitor. The weight is changed always when the clock signal is active.
- **Refresh** sub-circuit compensates the leakage current.
- **Multiplier** relays effect of weight on signal between neurons by multiplying the output spike of the pre-synaptic neuron with the weight.

Figure 7 shows simulation results of a synapse and two neurons. Parameters for simulations were: half amplitude duration of clock signal $= 1\,\mu$s, the period $= 40\,\mu$s, voltage at the input of pre-synaptic neuron = 5 V and initial voltage on memory capacitor = 2 V. At the bottom of the figure is the output from the f/V converter and the NMDA signal *ie* inputs of the comparator. The voltage change across the memory capacitor is shown in the upper part of the picture. A stable state of weight is shown, which lasts for more than 75 learning windows. Simulations were done in the HSPICE from Metasoft.

The implementation of our learning synapse was realized in $0.7\,\mu$m Alcatel MIETEC technology. The whole area occupied by synapse is about 10^4 square μm, without supply and GND wires.

8 CONCLUSIONS

A new learning synapse and a neuron for pulse coupled neural networks have been presented. The proposed neuron works with both constant or slowly changing input voltage level (input layer of a neural network) and pulse input (hidden layers and the output layer). It is around 10^4-times faster than the biological neuron. Its applications could be in the field of image processing. For medical applications the neuron's parameters need to be changed.

The behavior of the presented synapse is very close to its biological counterpart. Learning is realized using the Hebbian learning rule and the weight is stored on a memory capacitor in more than 256 levels. Advantages of the presented synapse are: continuous learning, low area consumption, low power consumption, easy and fully externally adjustable learning parameters and low cost.

The presented synapse has a smaller area and power consumption than the synapse presented in [7]. Lehman's synapse uses leakage current for its activity and therefore does not need a refresh circuit.

Acknowledgement

The authors wish to thank the Ministry of Education of the Slovak Republic which supported this work under Grant. No. 1/6096/99.

References

[1] ECKHORN, R. : Neural Mechanisms of Scene Segmentation: Recordings from the Visual Cortex Suggest Basic Circuits for Linking Field Model, IEEE Transactions on Neural Networks **10** No. 3 (1999), 480–497.

[2] KVASNIČKA, V.—BEŇUŠKOVÁ, .—POSPÍCHAL, J.– FARKAŠ, I.—TIŇO, P.—KRÁ, A. : Introduction to the Theory of Neural Networks (Úvod do teórie neurónových sietí), IRIS, Bratislava, 1997, pp. 11–30, 142–181. (in Slovak)

[3] OTA, Y.—WILIAMOWSKI, B. M. : Analog Implementation of Pulse-Coupled Neural Networks, IEEE Transactions on Neural Networks **10** No. 3 (1999), 539–543.

[4] GERSTNER, W. : Spiking Neurons, In: W. Maass and C. M. Bishop (Editors), Pulsed Neural Networks, MIT press, 1998, pp. 3–54.

[5] RUF, B. : Computing and Learning with Spiking Neurons - Theory and Simulations, Doctoral Thesis, Institute for Theoretical Computer Science, Technische Universität Graz, Austria, May 1998, pp. 19–20.

[6] MAASS, W. : Networks of Spiking Neurons: the Third Generation of Neural Network Models, In: Proceedings of the 7th Australian Conference on Neural Networks, Canberra, Australia, 1996, pp. 1–10.

[7] LEHMAN, T. : Classical Conditioning with Pulsed Integrated Neural Networks, IEEE Transactions on Circuits and Systems – II: Analog and Digital Signal Processing **45** No. 6 (1998), 720–728.

Received 2 December 2000

Pavol Tikovič (Ing), received his MSc degree from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in 2000. Since October 2000 has he been a PhD student at the Microelectronic Department of the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava. His research interests include VLSI design and on-chip implementation of neural networks.

Marián Vörös (Ing), received his MSc degree from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in 2000. His research interests include VLSI design and on-chip implementation of neural networks.

Daniela Ďuračková (Doc, Ing, CSc), received her MSc and CSc (PhD) from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in 1974 and 1981 respectively. Since 1991 she has been an associate Professor at the Microelectronics Department of the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava. The main field of her research and teaching activities has moved from semiconductor devices towards the design of analog and digital ASICs and neural network implementation on-chip.