ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 *2010. No. 2(98)*
ELEKTRONIKA IR ELEKTROTECHNIKA

T170 *ELECTRONICS*
ELEKTRONIKA

Pulsed Neural Network Cell in Integrated Circuit

V. Paukštaitis, A. Dosinas
Department of Electrical and Control Equipment, Kaunas University of Technology,
Studentų str. 50, LT-51368 Kaunas, Lithuania, tel.: +370 37 300270, e-mail: kat0102@ktu.lt

Introduction

The pulsed neural networks are electronic devices taking in the pulse coded information and outputting the processed information in the pulse stream manner.

The publication [1] has presented the information streams model (ISM), to get an answer whether such artificial neural system is capable to convolute two-dimensional signals. Therefore, two ways of capacitance charging were investigated [1] having in mind that such system is implementable in the integrated circuit (IC). Such devices are applied in real-time signal processing tasks, for example, the transformation to decrease the processing information [2]. The mobile intellectual systems with vision perception (vehicle control, rocket targeting subsystem) are the examples of actual application of such networks.

Various simulations and IC implementations of the artificial pulsed neural networks are presented in multiple papers. For example, the paper [3] presents *Complementary-Metal-Oxide-Semiconductor* (CMOS) circuit of a cell which is capable to be implemented into processors where arithmetic arrays in a double-base number system are necessary. Such cell was not implemented in the IC but simulated. The real cell density and electrical parameters are unknown. Another simulated artificial network is described in [4]. Here, the cell having the dead zone generator and using it with other microelectronic techniques such as wired sum, current mirrors and differential amplifier as the activation function is presented. The authors of [5] publication present the CMOS IC fabricated in 0.5μm process. Such IC is based on an asynchronous cell having of 200x300μm^2 size and 3.3V supply voltage. This artificial neuron is highly complex due to synapses capable to switch the arithmetical sign. The artificial neuron also includes the relative refractoriness circuit. The publication [6] presents the artificial pulsed neurons implementing the system to perform logical operations such as XOR or AND.

The literature describing the pulsed artificial networks do not present clearly the error of the system, its dependency on temperature and the variation of fabrication process parameters, the possibilities of error stabilization. Such inconsistencies motivate the designing of cell for the commercial application or mass production purposes. So, this publication presents the pulsed neural network cell withstanding a large range of the temperature and compensation methods. The simulation results are presented and discussed to prove the commercial viability of the artificial pulsed neural cell in the IC implementation using a commercial design-kit.

Principle of Operation

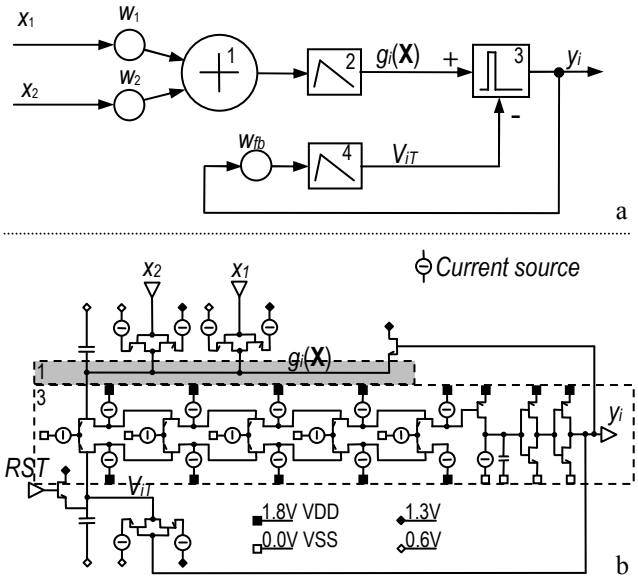

Fig. 1. Information model (a) and (b) electric schematics of artificial pulsed neuron: 1 – adder, 2, 4 – integrators, 3 – comparator and pulse formation unit

The information streams model (Fig. 1a) was built for the IC schematic engineers to be the guide on the performance metrics of the neuron with different transfer functions, to help to handle the correct weight setting and to check the viability of the larger neural networks as in [2] fast. The ISM was defined by two main equations in [1]:

$$g_i(\mathbf{X},t)=\left(\sum_{j=i}^{N} w_j A_0\right)\left(h(\tau_1,t-t_0)-h(\tau_2,t-t_1)\right), \quad (1)$$

$$V_{iT}(t) = (w_{fb_i} y_i)(h(\tau_3, t-t_2) - h(\tau_4, t-t_3)), \qquad (2)$$

where $g_i(\mathbf{X}, t)$ – the signal representing tension built by positive input signals, $V_{iT}(t)$ – the signal representing inner tension of neuron, w_j – the weight of the *j*-th input, A_0 – the amplitude of the pulse, $h(\cdot)$ – the transfer response function, τ_1, τ_2, τ_3, τ_4 – time constants, t_1, t_2, t_3, t_4 – times of pulse appearance and disappearance for input and out pulse respectively, w_{fb_i} – the feedback weight of the *i*-th neuron, y_i – the output of the *i*-th neuron.

The information streams model of the artificial pulsed neuron (fig. 1a) was realized as IC cell called electronic model (EM). The electric schematics of such cell are depicted in Fig. 1b.

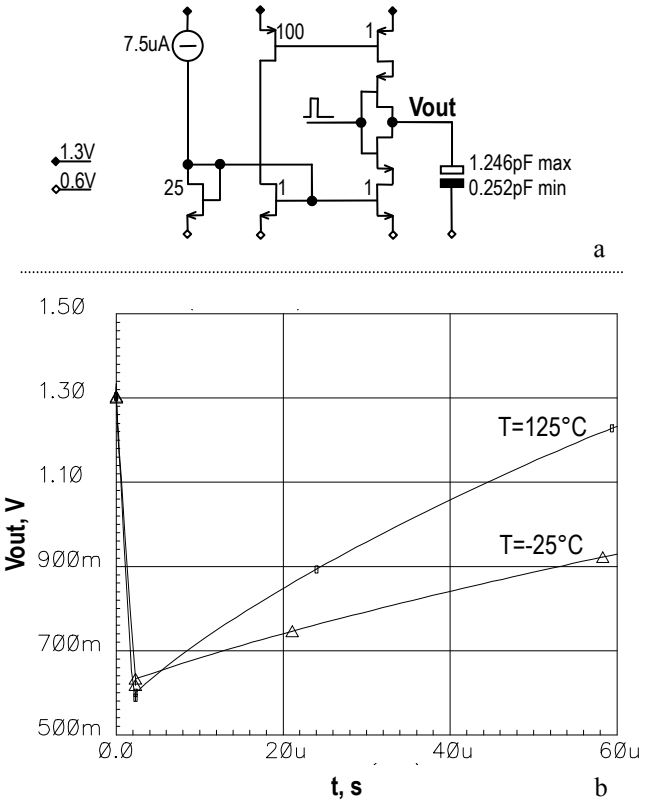

Fig. 2. Charge pump's capacitor voltage variation by unloading and loading: a) testing schematics, b) time diagram

The CMOS input signals x_1 and x_2 as pulse streams are controlling the charge-pump's switches connecting the current source charging the FET capacitor when the input signal is low and connecting the current source to discharge the FET capacitor when the input signal is high (Fig. 2a). The operation of addition is realized by wiring the outputs of the charge pumps together (Fig. 1b). The *RST* signal resets the feedback signal in the beginning of the operation. The charge pumps of EM shapes the transfer function (Fig. 2b) tested at different temperatures.

The voltage on the charge pump's capacitor is supplied to the input of the high gain differential amplifier which performs the subtraction operation of two analog signals. One output of this differential amplifier is directed to the pulse formation unit which shapes pulses to conform to input and output pulse specification. The output stream of the formatted pulses can be directed to the following neural layer or, as depicted in Fig. 1a, to the feedback of the current neural layer.

The actual design of the pulsed neuron is designed entirely for CMOS type process. The cell is nominally powered by VDD=1.8V analog and digital supply. The charge pump's input signal is CMOS 1.8V. The dynamic range of the high-gain amplifier is mitigated by voltage drop on the NFET and PFET which build the functionally necessary tail current source and loading current sources, respectively. This way, the lowest input voltage with respect to common node (ground, VSS) is 0.6V and the highest input voltage with respect to VSS is 1.3V. Although, the high-gain differential amplifier's signal transistors are selected NFET type, so all analog type signals are with respect to VDD rather than VSS and all digital signals are with respect to VSS.

The weights are set by current mirrors referenced by band-gap current source to stabilize the current magnitude in the temperature range [-25,125]ºC. The high-gain amplifier's tail current and load currents are set by current mirrors but these are fed from Proportional-To-Absolute-Temperature current source (PTAT). The latter is not linearly proportional to absolute temperature but varies its output current by the second order law to keep the gain of high-gain amplifier constant throughout the temperature ranging from -25 to 125°C. The PTAT, band-gap, relative and differential setting of the charging and discharging current makes the characteristics of the design theoretically stable against temperature. Therefore, the current design is compensated only partially for the variation of the components' parameters and not tested yet.

Weight Setting

The setting of weights for the artificial neuron is set by current mirrors in the EM. The ISM [1] stated that

$$(\tau_4 >> \tau_3) >> (\tau_2 >> \tau_1), \qquad (3)$$

where τ_3, τ_4 – feedback time constants, τ_2, τ_1 – positive signal time constants.

This equation is not applicable to EM because the relationship of current mirrors defining the current inflow or outflow to the capacitor is set by a finite number of FET fingers. While developing practical realization, relationship (3) shows the incapability to produce the expected transfer response for ISM and EM. The dual input and one feedback ISM was checked while changing the time constants so that the target transfer responses were similar to presented in [1]. The best transfer response error was considered when the mean transfer error function was at the centre to dynamic range and symmetrical according to that centre. The best target transfer response error has been produced when the relationships of the time constants were such:

$$\begin{cases} \tau_2/\tau_1 = 0.02, \\ \tau_4/\tau_3 = 0.0546. \end{cases} \qquad (4)$$

In turn, these numbers show that (3) must be

rewritten as follows:

$$(\tau_3 > \tau_4) > (\tau_1 > \tau_2). \quad (5)$$

Following the obtained relationships it was easy to find an appropriate number of fingers for current sources. Although, the weights vary due to impacts, such as temperature (Fig. 2), the distribution of impurities, gate-oxide thickness and etc. Such variations are partially compensated by the feedback charge-pump and in the practical implementation it is expected that transfer response will be stable in the wide temperature range.

Simulation and Design Results

The described electronic model implemented in factory-calibrated commercial *130µm IBM bicmos8hp SiGe* process is tested by a test bench for ISM [1] adjusted to *Cadence Design System's Spectre* simulator. The obtained test bench (Fig. 3) is simulated for 30 times with the input pulse streams which values differ by 2 to obtain the static response of the artificial neuron.

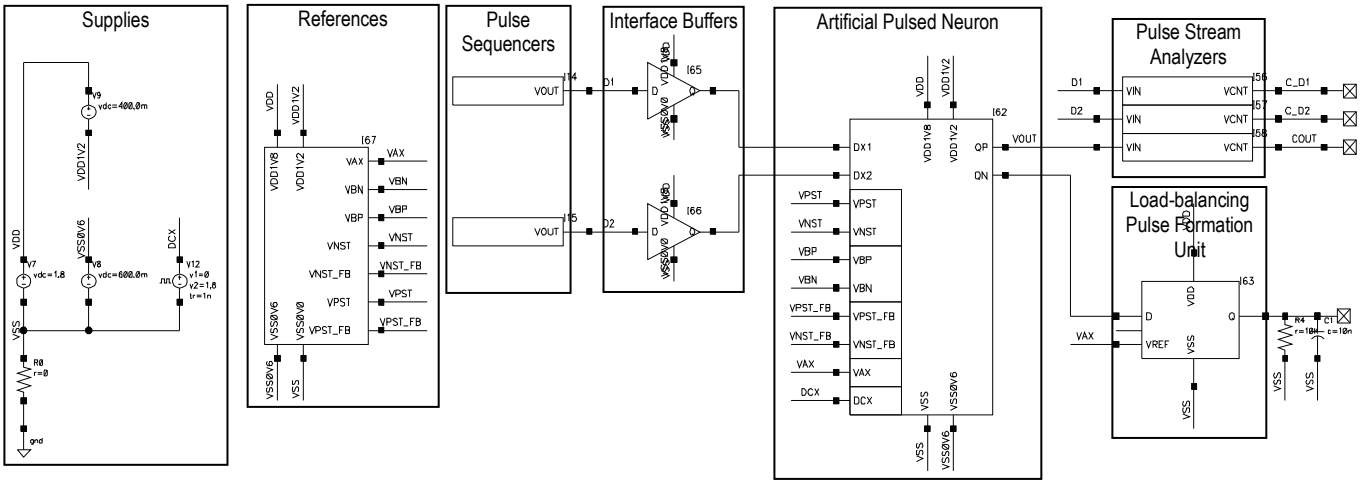

Fig. 3. Test bench schematics

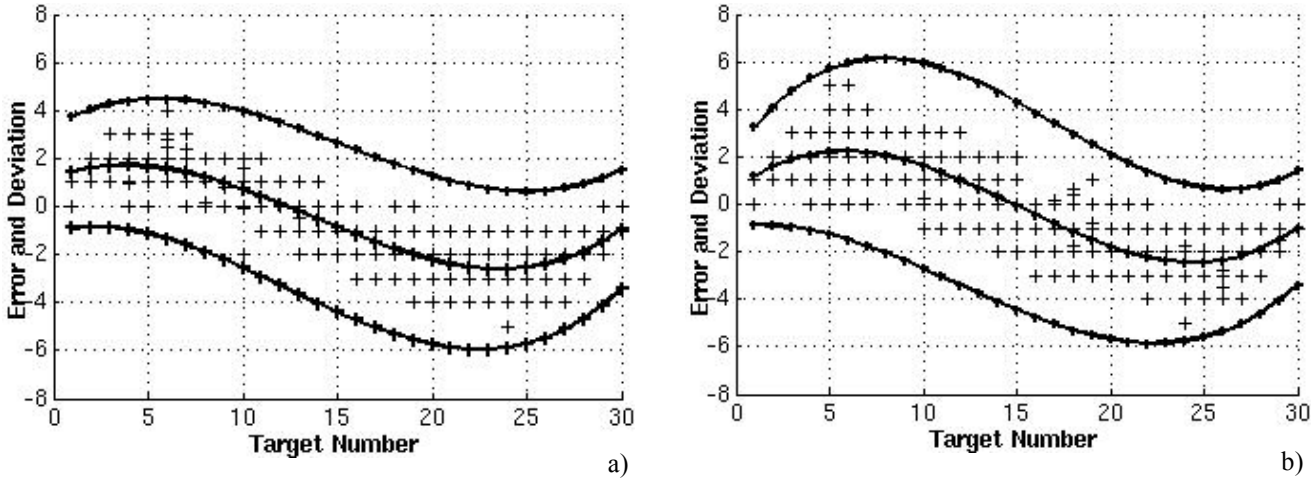

Fig. 4. The error and maximum deviation of the IC model of the artificial pulsed neuron: a) T=0ºC, b) T=125ºC

The input signals are generated by the pulse sequencers. This *Verilog-A* device is capable to generate pulse streams coding numbers ranging from 0 to 31. 30 randomly distributed layouts of pulse sequences for every number were generated outside the Verilog-A module to simulate the random fashion of input signal

The first measured characteristic is the difference between linear transfer response and the transfer response of the artificial pulsed neural network. This difference is called the error of the neural system and obtained at the temperatures 0ºC and 125ºC (Fig. 4). The small crosses depict actual simulation points.

Let us define that the transfer response error of the artificial neural cell is $r_m(n)$ where n represents the target value of this cell and m the selected experiment number. The mean transfer response error with M simulation or experimental data is

$$R(n) = M^{-1} \sum_{m=1}^{M} r_m(n). \quad (6)$$

Such mean transfer response error was approximated by the third order polynomial and considered as the standard transfer response error

$$S(n) = a_0 + a_1 n + a_2 n^2 + a_3 n^3 . \quad (7)$$

The factors of such polynomial approximating the error function of the EM comparing to the values obtained by simulating ISM [1] are presented in Table 1. The a_0 factor shows the systematic offset and it should be practically lower than the smallest number equal 1 as for ISM. The EM shows the a_0 value around the smallest number. The a_1 factor reports about the tendency of the average error to go down or up according to the target number. The a_1 is larger for both ISMs than EM for both models. In fact, it can be stated that factors are compensating each other while the actual average curves produced by both models are very similar in shape. The optimization accuracy of the EM could be one of the reasons of such situation.

Table 1. The polynomial factors of the transfer response error for ISM and EM

	a_0	a_1	a_2	a_3
Averages				
EM 0ºC	1.18	3e-1	-4.58e-2	1.1e-3
EM 125ºC	6.49e-1	6e-1	6.47e-2	1.4e-3
ISM Lin	9.17e-3	1.445	-1.38e-1	3.0e-3
ISM Exp	-3.31e-3	1.269	-1.16e-1	2.4e-3
Standard Deviation				
EM 0ºC	5.53e-1	9.38e-3	2.82e-3	8.5e-5
EM 125ºC	4.26e-1	8.5e-2	-3.97e-3	4.2e-5
ISM Lin	7.18e-2	2.04e-2	1.74e-3	4.5e-5
ISM Exp	5.80e-2	2.36e-2	1.99e-3	4.5e-5

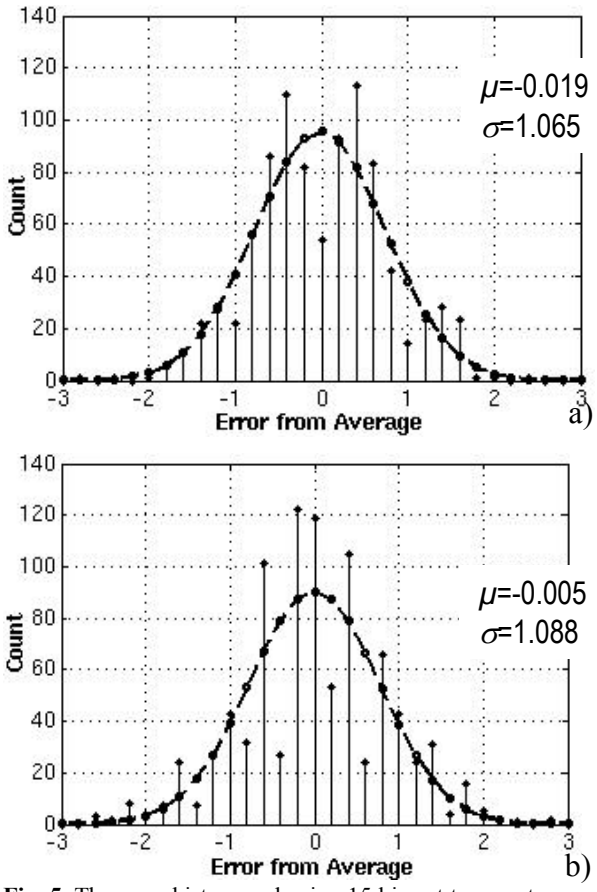

Fig. 5. The error histogram having 15 bins at temperatures: a) 0ºC, b) 125ºC

The third order polynomial is one type of error analysis. The other type is the histogram of the error and it is found when the average or standard third order polynomial values from actual experimental or simulation data are subtracted. Such histogram of the error obtained from 900 simulation data points is presented in Fig. 5. The profile was considered normal and, moreover, it is noticeable that the 95 percent of the output values have the error in the range from -2 to +2.

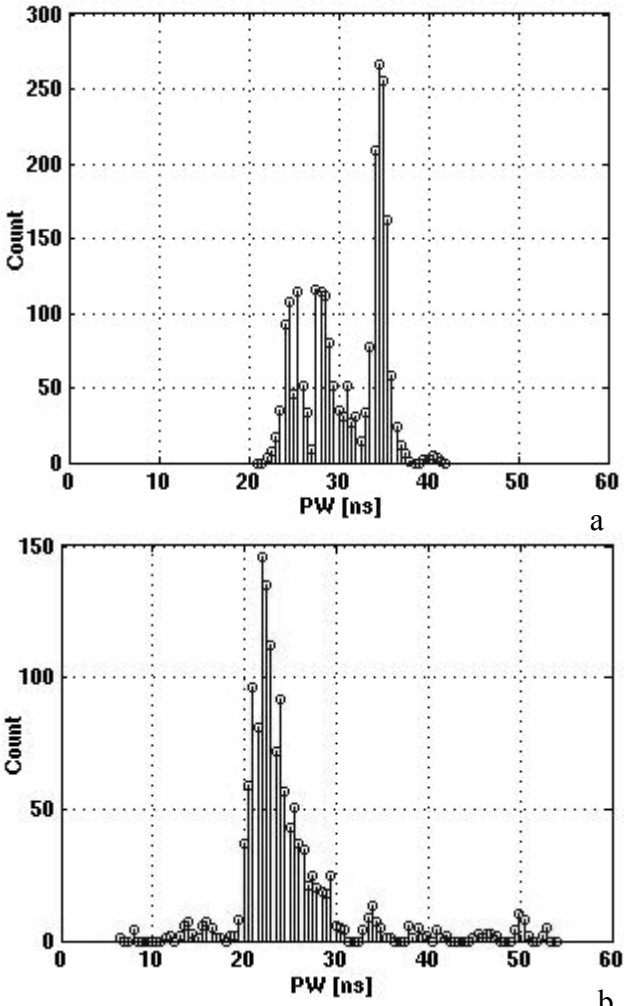

Fig. 6. Histograms of the output pulse width at temperatures: a) 0ºC, b) 125ºC

The difference between the values obtained by ISM and EM can be explained by the different models. Firstly, the ISM does not include the simulation of the pulse while the model in this paper does. The ISM treats multiplication and addition operations as the ideal ones while the EM lacks such ideal operations. The charging and discharging factors in the ISM can be defined very strictly while in the EM they are set by finite number of fingers on the both sides of the current mirror. The high-gain-amplifier is not an exception: amplification factor is finite, dependent on operation factors to some degree. In addition, such parameters as the charging-discharging current and pulse width also influence the final value of the output of the artificial neuron and even more, the whole artificial neural network.

The pulse formation unit, in ideal case, should form the output pulses with the same pulse width to pass the pulses to the next neural layer and to keep the power consumption constant due to operation conditions variation. The obtained pulse width histograms are presented in fig. 6 in nanoseconds for the temperatures 0ºC and 125ºC. Both histograms were obtained by measuring the pulse width of more than 1500 pulses. Looking at the histograms it can be concluded that pulse width varies with temperature which, in fact, is not welcomed, but the artificial neural layer still functions and with the presented dispersions of the pulse width. Additionally, the output pulses are directed only to feedback and the neural layer is not so sensitive to such pulse width fluctuations.

Figure 7 presents the conceptual layout for the neural system consisting of 3 charge pumps, 2 storage capacitors, high-gain-amplifier, output pulse conditioning circuitry. The cell's area is 50x40μm^2 and the schematics can be modified to even minimize the occupation area. For comparison, there can be approximately 450 such neurons in chip area of 1x1mm^2. Of course, it depends on the amount of charge pumps and output quantity. Therefore, it should be noticed that, there are no weight control circuitry which is necessary if the convolution kernel must be varied in time.

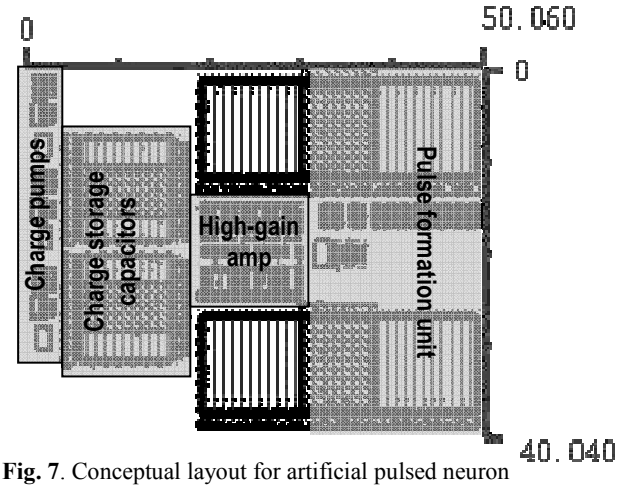

Fig. 7. Conceptual layout for artificial pulsed neuron

Synthesis Aspects

The presented artificial neuron can be easily composed from components (charge pumps, high-gain-amplifier, pulse-formation-unit). The homogenous artificial neural network is synthesizable easily then. To fulfill such task the following aspects should be taken into account:

•) the number of inputs determines the quantity of charge pumps which are wired together;

•) the charge storing capacitors always have the same quantity equal to 2;

•) the structures of high-gain-amplifier and pulse formation unit do not change;

•) pulse formation unit is loaded with the MOS transistors of the next neural layer. The pulse formation unit's output is digital and the load can be minimized by using addition CMOS buffers;

•) the weights are set and adjusted by current mirrors in the band-gap current reference;

•) the neural layer is composed from the equal neurons where inputs are wired in the desired manner and fed from the same current references to minimize power consumption.

Conclusions

1. The principle of operation of the artificial neural cell is based on the information streams model capable to be implemented on the semiconductor integrating circuit where the parameters are stabilized applying compensation techniques such as relative setting of the weights, proportional-to-absolute-temperature and band-gap current sources.

2. The simulation of the pulsed artificial neuron design shows that such a device is capable to operate in the temperature range from 0ºC to 125ºC giving the 95 percent of the output values in the range from -2 to +2.

3. The artificial pulsed neurons' transfer response error function qualitative difference between the information streams model and the electronic model is low and such an artificial neuron can be implemented in the IC.

4. The conceptual layout size is of 40x50μm^2 is two times less than the reported 200x300μm^2 for 0.5μm technology. The presented layout can be minimized by optimizing the schematics of the artificial pulsed neuron.

5. The large artificial pulsed network is synthesizable easily due to the artificial pulsed neuron composed from the fixed structure components (charge pump, high gain amplifier and pulse width formation unit).

References

1. **Paukštaitis V., Dosinas A.** Pulsed Neural Networks for Image Processing // Electronics and Electrical Engineering. – Kaunas: Technologija, 2009. – No. 7(95). – P. 15–20.
2. **Paukštaitis V., Dosinas A.** Multilayer Transformation of Different Resolution for Colour Image Analysis and Segmentation // Electronics and Electrical Engineering. – Kaunas: Technologija, 2006. – No. 8(72). – P. 49–54.
3. **Sadeghi-Emamchaie S., et al.** Digital Arithmetic Using Analog Arrays // Proceedings on the 8th Great Lakes Symposium on VLSI, 1998. – P. 202–205.
4. **Ibrahim Y., Jullien G. A., Miller W. C.** Ultra Low Noise Signed Digit Arithmetic Using Cellular Neural Networks // Proceedings of the 4th IEEE International Workshop on System-on-Chip for Real-Time Applications. – 2004. – P. 136-142.
5. **Taniguchi T., Horio Y., Aihara K.** Integrated Pulse Neuron Circuit for Asynchronous Pulse Neural Networks // Proceedings of the International Joint conference on Neural Networks, 2003. – Vol. 2. – P. 942-947.
6. **Liu B., Frenzel J. F.** A CMOS Neuron for VLSI Circuit Implementation of Pulsed Neural Networks // IEEE 28th Annual Conference of the IECON 02. – 2002. – Vol. 4. – P. 3182–3185.

Received 2009 09 28

V. Paukštaitis, A. Dosinas. Pulsed Neural Network Cell in Integrated Circuit // Electronics and Electrical Engineering. – Kaunas: Technologija, 2010. – No. 2(98). – P. 35–40.

The research of the created model of artificial pulsed neuron is presented. The peculiarities of building semiconductor integrated circuit with neuron cell and the artificial neural network constructed from such cells are discussed. The principle of operation of the artificial pulsed neural cell and the compensation methods of parameter variation due to temperature are unclosed for the reader. The model and the test bench of the artificial pulsed neuron cell are implemented in semiconductor factory calibrated design-kit. The obtained simulation results corroborated the research results of the earlier formed information streams model and showed that the designed artificial pulsed neuron cell is viable in the wide temperature range. The paper also presents the research of other electrical characteristics of the artificial pulsed neuron. The conceptual layout is presented and the synthesis aspects of the pulsed artificial networks are discussed. Ill. 7, bibl. 6 (in English; summaries in English, Russian and Lithuanian).

В. Паукштайтис, А. Досинас. Ячейка импульсной нейронной сети в микросхеме // Электроника и электротехника. – Каунас: Технология, 2010. – № 2(98). – С. 35–40.

Приводятся исследования модели исскуственного импульсного нейрона и рассматриваются особенности интегрирования ячеек такого нейрона и созданной на их основе нейронной сети в полупроводниковую интегральную схему. Представлен принцип действия ячейки импульсного нейрона и методы компенсирования изменения параметров ячейки под влиянием температуры. Модель ячейки такого нейрона и схема ее проверки была реализована на калиброванном производителем пакете проектирования. Полученные результаты моделирования подтвердили результаты, полученные во время исследования ранее разработанной модели информационных потоков, и показали работоспособность спроектированной ячейки импульсного нейрона в широком диапазоне изменения температур. В статье рассматриваются также и другие электрические характеристики ячейки импульсного нейрона. Приводится эскиз топографического чертежа ячейки, рассматриваются аспекты синтеза исскуственной импульсной нейронной сети, созданной из рассматриваемых в статье ячеек. Ил. 7, библ. 6 (на английском языке; рефераты на английском, русском и литовском яз.).

V. Paukštaitis, A. Dosinas. Impulsinio neuroninio tinklo ląstelė integriniame grandyne // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2010. – Nr. 2(98). – P. 35–40.

Pateikiamas sudaryto dirbtinio impulsinio neurono ląstelės modelio tyrimas, nagrinėjami tokio neurono ląstelės ir iš jų sudaryto impulsinio dirbtinio neuroninio tinklo integravimo į puslaidininkinį lustą ypatumai. Skaitytojui pristatomas neuroninio impulsinio tinklo ląstelės veikimo principas ir parametrų kitimo dėl temperatūros įtakos kompensavimo metodai. Neuroninio impulsinio tinklo ląstelės modelis ir jo tikrinimo schema sumodeliuoti gamintojo kalibruotu projektavimo paketu. Gautieji modeliavimo rezultatai patvirtino anksčiau sudaryto informacinių srautų modelio tyrimo rezultatus ir parodė, kad suprojektuota dirbtinio impulsinio neurono ląstelė yra gyvybinga plačiame temperatūros kitimo diapazone. Straipsnyje taip pat nagrinėjamos ir kitos impulsinio dirbtinio neurono ląstelės elektrinės charakteristikos. Pateikiamas ląstelės topografinio brėžinio eskizas, aptariami impulsinio dirbtinio neuronų tinklo, sudaryto iš nagrinėtų ląstelių, sintezės aspektai. Il. 7, bibl. 6 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).