Analysis of Analog Neural Network Model with CMOS Multipliers

Liliana DOCHEVA, Alexander BEKIARSKI, Ivo DOCHEV

Faculty of Communications Technic and Technologies, Technical University, Sofia, Bulgaria

docheva@tu-sofia.bg, aabbv@tu-sofia.bg, idochev@tu-sofia.bg

Abstract. The analog neural networks have some very useful advantages in comparison with digital neural network, but recent implementation of discrete elements gives not the possibility for realizing completely these advantages. The reason of this is the great variations of discrete semiconductors characteristics.

The VLSI implementation of neural network algorithm is a new direction of analog neural network developments and applications. Analog design can be very difficult because of need to compensate the variations in manufacturing, in temperature, etc. It is necessary to study the characteristics and effectiveness of this implementation. In this article the parameter variation influence over analog neural network behavior has been investigated.

Keywords
Analog neural networks, VLSI, robotic.

1. Introduction

Artificial Neural Networks is a new method wide used for information processing. Analog neural networks have the following advantages: high speed, low power consumption and compact implementation in comparison with competing digital signal processing approaches. With the help of the analog neural networks certain computations that are difficult or time-consuming for digital neural network can be done. A disadvantage of analog neural networks is their limited accuracy and nonlinear behavior. Variation in the size of discrete transistors and the local mobility will cause random parameter variation. Moreover an increase in the precision of any component has as a consequence an increase of its area. The aim of this paper is to investigate influence of some analog neural network parameters onto its recognition ability.

2. Implementation of Analog Neural Networks

Many investigations in the implementation of analog neural networks field are known [2, 3, 4, 5]. They are wide used because of their advantages, high speed, low power consumption and compact implementation. Nevertheless the analog computational hardware is typically limited to a relative precision of about 1%. For this reason it is preferable a simple ANN model to be used. The model must be compatible with the restrictions imposed by the analog VLSI technology. Otherwise the advantages of using the technology would be lost.

Fig. 1 depicts expandable neural network, which topology is though simple, a very capable one. In such a way we can implement systems of arbitrary size, fully connected between the layers.

Fig. 1. Expandable neural network.

There are many ways for analog neural network implementation, but it is necessary to study the characteristics and effectiveness of their implementation.

3. Parameter Variation Influence over Analog Neural Network Behavior

Variation in the size of individual transistors and the local mobility will cause random parameter variation. In this article analog parameter variation over analog neural network behavior is investigated.

In [1], [7], VLSI implementation of an analog neural network is depicted. The synapse chip consists of a number of inner product multipliers. The authors have chosen to use the MOS resistive circuit multiplier. Weight values are
stored by simple capacitive storage method with RAM Backup. The schematic of a single synapse is shown in Fig. 2.

![Schematic of a single synapse.](image)

**Fig. 2.** Schematic of a single synapse.

The transfer function of the synapse is the following

\[ i_{sk} = \frac{g_{mk}}{W_0/L_0V_c} \sum W_j/L_j V_{wj} v_{ij} \]  

(1)

where \( g_{mk} \) is the transconductance parameter, \( W/L \) are the MOS resistive circuit multiplier width/length ratios, \( V_c \) controls the total transconductance, \( V_{wj} \) are the weight values voltages, \( v_{ij} \) are the input voltages.

![Hyperbolic tangent neuron.](image)

**Fig. 3.** Hyperbolic tangent neuron.

The transfer function of the synapse is the following

\[ i_{sk} = \frac{g_{mk}}{W_0/L_0V_c} \sum W_j/L_j V_{wj} v_{ij} \]  

(1)

where \( g_{mk} \) is the transconductance parameter, \( W/L \) are the MOS resistive circuit multiplier width/length ratios, \( V_c \) controls the total transconductance, \( V_{wj} \) are the weight values voltages and \( v_{ij} \) are the input voltages.

On the base of the neuron and synapse equations of this implementation so as equations about analog parameter variation over analog neural network behavior have been worked out [6]. On this way we have investigated parameter variation influence over neural network behavior.

After parameter variation the synapse equation becomes

\[ i_{sk} = \frac{g_{mk} + \Delta g_{mk}}{W_0/L_0V_c} \sum W_j/L_j V_{wj} v_{ij} \]  

(3)

where \( l \) is the layer number, \( g_{mk} \) is the transconductance parameter, \( W/L \) are the MOS resistive circuit multiplier width/length ratios, \( V_c \) controls the total transconductance, \( V_{wj} \) are the weight values voltages, \( v_{ij} \) are the input voltages and \( \Delta g_{mk} \) is the step of transconductance parameter change.

Equation (4) is the neuron equation including parameter variation.

\[ v'_{sk} = \alpha_{FC} (I_b + \Delta I_b) \frac{\beta_{ORB} + \Delta \beta_{ORB} V_{OR}}{2(\beta_{ORB} + \Delta \beta_{ORB} V_{OR})} \]  

(4)

where \( l \) is the layer number, \( \alpha_{FC} \) is the emitter-collector current gain, \( I_b \) is the bias current, \( \beta_{ORB} \) and \( \beta \) are the MOSFET transconductance parameters, \( V_t \) is the thermal voltage, \( V_{IS} \) and \( V_{OR} \) are the control voltages, \( \Delta \alpha_{FC} \), \( \Delta \beta_{ORB} \) and \( \Delta \beta \) corresponds to the steps of \( \alpha_{FC} \), \( \beta_{ORB} \) and \( \beta \) parameters change.

On the base of the equations (3) and (4) investigation for analog parameter variation over analog neural network behavior has been done. A popular method for study of neural networks is network simulation using computers. In the analysis of any circuit it has been assumed that all the components were ideal. In this article a simulation using Matlab is presented, but in the real neural networks equations the parameters of real components take part. These parameters are: the MOS resistive circuit multiplier width/length ratios \( W/L \), the emitter-collector current gain \( \alpha_{FC} \), the bias current \( I_b \), MOSFET transconductance parameter \( \beta \) and thermal voltage \( V_t \).

The value of the emitter-collector current gain \( \alpha_{FC} \) is about \( \alpha_{FC} \approx 0.5 \). It’s variation because of parasitic processes is in a small range: \( \alpha_{FC} \approx (0.4÷0.55) \) \[1\]. In Tab. 1 for each \( \alpha_{FC} \) value error values and boundaries of weight values are given. It can be seen from the table that \( \alpha_{FC} \) variation is very important because this variation have effect on both: the error value and the range of weight values variation. Some of the weights can reach values of \( w = -93 \) and \( w = 64 \). With \( \alpha_{FC} \) increasing, the range of weight values variation decreases.

<table>
<thead>
<tr>
<th>( \alpha_{FC} )</th>
<th>( W^1 )</th>
<th>( W^2 )</th>
<th>( E )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>-93+64</td>
<td>-20+16</td>
<td>0.04</td>
</tr>
<tr>
<td>0.45</td>
<td>-10+5</td>
<td>-8+4</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>-3+2</td>
<td>-3+0.3</td>
<td>4.9.10^-4</td>
</tr>
<tr>
<td>0.55</td>
<td>-1.2+1.7</td>
<td>-1.8+1.4</td>
<td>1.9.10^-1</td>
</tr>
</tbody>
</table>

**Tab. 1.** Influence of \( \alpha_{FC} \) over error values and range of weight values variation.
The influence of $\alpha_{FC}$ over error values is depicted in Fig. 4. It can be seen that for values $\alpha_{FC}=0.4$ and $\alpha_{FC}=0.45$ the error increases compared to that achieved at the value $\alpha_{FC}=0.5$.

Because of their position in the equation (eq. 4) variations of parameters $\alpha_{FC}$, $I_B$ and $\beta_{OR}$ directly reflect on the output voltage range. This is the reason for error increasing for values of $\alpha_{FC}$ shown at the top of the Tab. 1. The influence of $\alpha_{FC}$ over forward mode neuron characteristics is depicted in Fig. 5.

The parameter $\alpha_{FC}$ variance strongly influences over analog neural network behavior. Regulating circuit for this parameter is depicted in [1].

The bias current $I_B$ is investigated in the range $I_B=(50 \text{ mA}; 70 \text{ mA})$. This parameter has effect on both: the error value and the range of weight values variation, too. For values $60 \text{ mA}>I_B>65 \text{ mA}$, some of the weights can reach values of $w=-47$ and $w=29$ (Tab. 2).

The influence of $I_B$ over error values is depicted in Fig. 6. The smallest error is obtained for the values $I_B=60 \mu\text{A}$ and $I_B=65 \mu\text{A}$, but for the second one the error increases after 300 epochs.

Variation of parameter $I_B$ directly reflects over output voltage range just like $\alpha_{FC}$. Graphics of the effect of $I_B$ on the forward mode neuron characteristics are quite similar to those of Fig. 5 and it is no reason to give them here.

The MOSFET transconductance parameter $\beta_{OR}$ is investigated in the range $\beta_{OR}=(25 \text{ }\mu\text{A}/\text{V}^2; 40 \text{ }\mu\text{A}/\text{V}^2)$. One is part of group of parameters that directly reflects over output voltage. Therefore $\beta_{OR}$ variation leads to rapid increasing of the output error and the range of weight values variation.

It can be seen from Tab. 3 that for values $\beta_{OR}≥30 \text{ }\mu\text{A}/\text{V}^2$ range of weight values strongly increases. For the value $\beta_{OR}=40 \text{ }\mu\text{A}/\text{V}^2$ one is $(-229+170)$. When analog neural network implementation has been discussed these are impermissible weight values.

From Fig. 7 it can be seen that the output error is maximum for the value $\beta_{OR}=25 \text{ }\mu\text{A}/\text{V}^2$. For the value $\beta_{OR}=20 \text{ }\mu\text{A}/\text{V}^2$ neural network cannot be learned.
The thermal voltage $V_t$ is investigated in the range $V_t=(23 \text{ mV}-37 \text{ mV})$. This corresponds to the thermal range $(-10^\circ\text{C} -150^\circ\text{C})$. The thermal voltage variation doesn’t lead to essential increasing of weight values range (Tab. 4). Output neuron error modifies slightly, too.

The error is comparatively low for the entire range of $V_t$ investigation (Fig. 8).

For values $\beta_s < 50 \mu\text{A}/\text{V}^2$ output neural network error decreases, but error decreasing holds on in the beginning of the learning process (see Fig. 10). For the value $\beta_s = 20 \mu\text{A}/\text{V}^2$ error decreasing begins after 300 epochs. Hence the leaning process becomes slower. This process doesn’t observe for values $\beta_s < 50 \mu\text{A}/\text{V}^2$. In this case, error increases slightly.

### Tab. 4. Influence of $V_t$ over error values and range of weight values variation.

<table>
<thead>
<tr>
<th>$V_t$, mV</th>
<th>$W^1$</th>
<th>$W^2$</th>
<th>$E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>-2.5+1.6</td>
<td>-2.8+0.3</td>
<td>4.2.10^-4</td>
</tr>
<tr>
<td>26</td>
<td>-2.7+1.6</td>
<td>-3+0.3</td>
<td>4.9.10^-4</td>
</tr>
<tr>
<td>29</td>
<td>-2.9+1.7</td>
<td>-2.9+0.3</td>
<td>5.5.10^-4</td>
</tr>
<tr>
<td>32</td>
<td>-3.2+1.7</td>
<td>-3.2+0.3</td>
<td>5.9.10^-4</td>
</tr>
<tr>
<td>37</td>
<td>-3.4+1.8</td>
<td>-3.4+0.3</td>
<td>6.2.10^-4</td>
</tr>
</tbody>
</table>

### Tab. 5. Influence of $\beta_s$ over error values and range of weight values variation.

<table>
<thead>
<tr>
<th>$\beta_s$, $\mu\text{A}/\text{V}^2$</th>
<th>$W^1$</th>
<th>$W^2$</th>
<th>$E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>-1.9+1.5</td>
<td>-1.8+0.4</td>
<td>4.2.10^-4</td>
</tr>
<tr>
<td>25</td>
<td>-1.97+1.55</td>
<td>-2+0.3</td>
<td>2.10^-4</td>
</tr>
<tr>
<td>30</td>
<td>-3.2+1.7</td>
<td>-3.2+0.3</td>
<td>2.4.10^-4</td>
</tr>
<tr>
<td>40</td>
<td>-3.4+1.8</td>
<td>-3.4+0.3</td>
<td>3.7.10^-4</td>
</tr>
<tr>
<td>50</td>
<td>-2.7+1.6</td>
<td>-3+0.3</td>
<td>4.9.10^-4</td>
</tr>
<tr>
<td>60</td>
<td>-2.7+1.6</td>
<td>-3+0.3</td>
<td>5.9.10^-4</td>
</tr>
<tr>
<td>70</td>
<td>-2.9+1.7</td>
<td>-2.9+0.3</td>
<td>6.2.10^-4</td>
</tr>
<tr>
<td>80</td>
<td>-3.2+1.7</td>
<td>-3.2+0.3</td>
<td>6.1.10^-4</td>
</tr>
</tbody>
</table>
The parameter $\beta_0$ variation reflects on the slope of the activation function. Therefore parameter $\beta_0$ variation slightly influences over analog neural network behavior. Parameter $\beta_0$ variation causes output neural network error variation that is admissible in bigger range than parameter $\beta_{OR}$ variation does.

The transconductance $g_{mk}$ variation reflects on the slope of the activation function and learning process. One is investigated in the range (1 mA/V ÷10 mA/V). It can be seen from Tab. 6 that $g_{mk}$ influence over low boundary of weight values. For the value $g_{mk}=1$ mA/V negative weights can reach the value $w=-11$. 

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
$g_{mk}$ & $W^1$ & $W^2$ & $E$
\hline
1 & -11+1.5 & -10+1 & $2.3 \times 10^{-3}$
2 & -5.1+1.9 & -5.4+0.6 & $9.9 \times 10^{-4}$
3 & -3.5+0.5 & -3.9+0.5 & $6.6 \times 10^{-4}$
4 & -2.7+1.6 & -3+0.3 & $4.9 \times 10^{-4}$
5 & -2.4+1.6 & -2.6+0.3 & $3.8 \times 10^{-4}$
6 & -2.1+1.5 & -2.3+0.3 & $3.0 \times 10^{-4}$
7 & -2+1.5 & -2+0.3 & $2.4 \times 10^{-4}$
8 & -1.9+1.5 & -1.9+0.3 & $1.9 \times 10^{-4}$
9 & -1.8+1.4 & -1.8+0.3 & $1.6 \times 10^{-4}$
10 & -1.8+1.4 & -1.7+0.3 & $1.5 \times 10^{-4}$
\hline
\end{tabular}
\caption{Influence of $g_{mk}$ over error values and range of weight values variation.}
\end{table}

Fig. 11 shows the error variation versus $g_{mk}$ variation. In the investigated range of variation error increasing is negligible. Although with $g_{mk}$ increasing learning speed decreases, because of error decreasing holds on in the beginning of the learning process.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig11.png}
\caption{Influence of $g_{mk}$ over error values.}
\end{figure}

4. Conclusion

The aim of this paper is to find the boundaries of parameters of analog neural network variations in which neural networks operate correctly. For that purpose a simulation using Matlab is presented. In the analysis of the circuits, components are not ideal. In neural networks equations the parameters of real components take part.

It is shown that variation of the first investigated group of parameters ($\alpha_{FC}, \beta_0$ and $\beta_{OR}$) parameters directly reflects on the output voltage range. Because of this, error increasing is fast for small variation parameters.

The parameters $V_o, \beta_0$ and $g_{mk}$ position is in the argument of $\tanh$ function. Therefore their variation doesn’t directly reflect on the output voltage range. Their variation reflects on the slope of the activation function. At significant variation of the second group of parameters learning speed decreases.

References


About Authors...

Liliana DOCHEVA (*1974 in Sofia, Bulgaria, M.S. degree in electronic and automatic engineering from the Technical Univ. (TU), Sofia, in 1996) is assistant professor at the TU, Sofia. Her research interests lie in the areas of neural networks, signal processing and computer vision.

Alexander BEKIARSKI (M.S. degrees in Communications in 1969, Ph. D in Television and Image Processing in 1975, both from the TU, Sofia) has been Assoc. Professor at the TU Sofia since 1987. He published over 120 research papers in Image Processing Systems, Pattern Recognitions, Neural Networks etc. His scientific ranges are in Image Processing Systems, Pattern Recognitions, Neural Networks, Digital Signal Processors for Image and Audio Processing, Polar Image Processing, Camera Eye Tracking.

Ivo DOCHEV (*1970 Sofia, Bulgaria, M.S. degree in electronic and automatic engineering from the TU Sofia in 1996) is assistant professor at the TU, Sofia. His research interests lie in the areas of neural networks, signal processing and measurement in communication.