Learning Accuracy Analysis of Memristor-based Nonlinear Computing Module on Long Short-term Memory

Hongyu An, Mohammad Shah Al-Mamun, Marius K. Orlowski, Yang Yi

Multifunctional Integrated Circuits and Systems (MICS) Group
The Bradley Department of Electrical and Computer Engineering
Virginia Tech, Blacksburg, VA, USA
July 25, 2018
Outline

• Backgrounds and Motivations
  • Von Neumann Computing Architecture Revisit
  • Emerging Neuromorphic Computing Architectures

• Memristor-based Nonlinear Computing Module

• Learning Accuracy Analysis of Memristor-based Nonlinear Computing Module on Long Short-term Memory

• Conclusions
Neuromorphic Computing

Explanation

Neuromorphic System

Reverse Engineering

Brain

Engineering Contributions:
• More power efficiency system
• Neuromorphic learning system;

Scientific Contributions:
• Optical illusion
• Mechanism of Memory
• Cognition

H. An.
Von Neumann Architecture: Design for Computing

Problems

Abstraction

Numbers

Equations

Computing

Results

Arithmetic/Logic Unit

Input Device

Encoding

Bus

Memory

Output Device

Decoding

H. An.
# Neuromorphic System: Design for Learning

<table>
<thead>
<tr>
<th>Digital Computer</th>
<th>Neuromorphic System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programs/Logic</td>
<td>Algorithm</td>
</tr>
<tr>
<td>Binary signals</td>
<td>Encoding Scheme</td>
</tr>
<tr>
<td>Von Neumann Architecture</td>
<td>Architecture</td>
</tr>
<tr>
<td>CPUs(Logic Gates, etc.), Memory(SRAM, etc.)</td>
<td>Devices</td>
</tr>
<tr>
<td></td>
<td>Neural Networks</td>
</tr>
<tr>
<td></td>
<td>Neurons and Synapses</td>
</tr>
</tbody>
</table>

- **Digital Computer**
  - Programs/Logic
  - Binary signals
  - Von Neumann Architecture
  - CPUs(Logic Gates, etc.), Memory(SRAM, etc.)

- **Neuromorphic System**
  - Algorithm
  - Encoding Scheme
  - Architecture
  - Devices
  - Neural Networks
  - Neurons and Synapses
Emerging Neuromorphic Computing Architectures: Distributed Neuromorphic Computing Architecture

Basic Building Modules

- Analog Signals
- Nonlinear Computing Module
- Matrix Computing Module
- Training Module

Distributed Neuromorphic Computing Architecture

[An, H., et al. (2017)]
Emerging Neuromorphic Computing Architectures: Cluster Neuromorphic Computing Architecture

Different sensory signals are processed in different regions;


Emerging Neuromorphic Computing Architectures: Associative Neuromorphic Computing Architecture

- Touch sensation
- Auditory sensation
- Vision
- Olfaction
- Gustatory sensation

Hardware Implementation: Nonlinear Computing Module

Distributed Neuromorphic Computing Architecture

Cluster Neuromorphic Computing Architecture

Associative Neuromorphic Computing Architecture

Physical Implementations

Analog Signals + Nonlinear Computing Module + Matrix Computing Module + Training Module

Memristor-based Nonlinear Computing Module

- Prerequisites:
  - The cascaded memristors can switch
  - The controllable set voltage
  - The controllable high resistance state/low resistance state (HRS/LRS)

\[
R^{(1)} = R_1^H + R_2^H + R_3^H + ... + R_i^H + ... R_n^H \\
R^{(2)} = R_1^L + R_2^H + R_3^H + ... + R_i^H + ... R_n^H \\
R^{(3)} = R_1^L + R_2^L + R_3^H + ... + R_i^H + ... R_n^H \\
R^{(i+1)} = R_1^L + R_2^L + ... + R_i^L + ... R_n^L
\]

- \(R_H\) is the high resistance value of each memristor; 
- \(R_L\) is the low resistance value of each memristor;
The switching Behavior Investigation of the Cascaded Memristors

Keithley 4200-SCS Semiconductor Parameter Analyzer

Probe

Microscope view

Memristor cell

Cu (150 nm)/TaOx(25 nm)/Rh(50 nm)
The set/reset voltage and HRS/LRS are determined by
• Materials;
• Physical geometry;
• Temperature;

https://nano.stanford.edu/stanford-memory-trends
The Mathematical Model of the Cascade Memristor-based Nonlinear Computing Module

Where

- $R$ is the total resistance of nonlinear module;
- $R_H$ is the high resistance value of each memristor;
- $R_L$ is the low resistance value of each memristor;
- $n$ is the total number of memristors in the module;
- $k$ is the step index whose value is from 0 to $n$;
- $\Delta I_{th}$ is an interval of threshold current values between two consecutive memristor ($I_{th_k} - I_{th_{k-1}} = \Delta I_{th}$), where $I_{th_k}$ is the threshold value of $k$th memristor.

\[
\begin{align*}
R &= (n - k)R_H + kR_L \\
I &= k \times \Delta I_{th} \\
R &= \left(n - \frac{I}{\Delta I_{th}}\right)R_H + \frac{I}{\Delta I_{th}}R_L \\
V_{out} &= -\left(\frac{R}{R_1}\right)V_{in}
\end{align*}
\]
Application to Digit Recognition with Long-short Term Memory

$$h_{t-1}, x_t \xrightarrow{\sum} \sigma$$

Output gate

$$h_t \xrightarrow{\tanh}$$

$$M_{t-1} \xrightarrow{\odot} M_t$$

$$M_t \xrightarrow{\odot} M_{t+1}$$

Forget gate

$$h_{t-1}, x_t \xrightarrow{\sum} \sigma$$

Input gate

$$h_{t-1}, x_t \xrightarrow{\sum} \sigma$$

Input

A Piecewise Approximation to Sigmoid Function
The training accuracies do not degrade by replacing the nonlinear function with piecewise approximation.
The testing accuracies decrease by the impact of the large resistance switching variation of memristor.
The testing accuracies decrease almost proportional with the increase of resistance variation of the memristor.
Conclusions

• Introduce three emerging neuromorphic architectures: Distributed Neuromorphic computing architecture; Cluster Neuromorphic Neuromorphic Computing Architecture; Associative Neuromorphic Computing Architecture

• We designed and evaluated a memristor-based nonlinear computing module in the Long Short-term Memory with the application on digit number recognition

• The training accuracy would not be degraded by using the proposed nonlinear computing module with ideal memristor

• The large resistance switching variation of memristor would significantly reduce the learning and testing accuracy, and the accuracies decrease is almost proportional to the increase of resistance variation of the memristor
Acknowledgement

Dr. Yang (Cindy) Yi  
Assistant Professor  
Virginia Tech  
cindy_yangyi@vt.edu

Mohammad Shah Al-Mamun  
Ph.D. Candidate  
Virginia Tech  
samamun@vt.edu

Dr. Marius K. Orlowski  
Professor  
Virginia Tech  
marius@vt.edu
References


Q & A

Hongyu An
Ph.D. Candidate
Virginia Tech
hongyu51@vt.edu

Thank you