Tunneling-Based Cellular Nonlinear Network Architectures for Image Processing

Pinaki Mazumder, *Fellow, IEEE*, Sing-Rong Li, and Idongesit E. Ebong

Abstract—The resonant tunneling diode (RTD) has found numerous applications in high-speed digital and analog circuits due to the key advantages associated with its folded back negative differential resistance (NDR) current-voltage (I-V) characteristics as well as its extremely small switching capacitance. Recently, the RTD has also been employed to implement high-speed and compact cellular neural/nonlinear networks (CNNs) by exploiting its quantum tunneling induced nonlinearity and symmetrical I-V characteristics for both positive and negative voltages applied across the anode and cathode terminals of the RTD. This paper proposes an RTD-based CNN architecture and investigates its operation through driving-point-plot analysis, stability and settling time study, and circuit simulation. Full-array simulation of a 128×128 RTD-based CNN for several image processing functions is performed using the Quantum Spice simulator designed at the University of Michigan, where the RTD is represented in SPICE simulator by a physics based model derived by solving Schrödinger's and Poisson's equations self-consistently. A comparative study between different CNN implementations reveals that the RTD-based CNN can be designed superior to conventional CMOS technologies in terms of integration density, operating speed, and functionality.

Index Terms—Resonant tunneling diode (RTD), cellular neural/nonlinear network (CNN), full array simulation, settling time analysis.

I. INTRODUCTION

SINCE its invention by Chua and Yang in 1988 [1], [2], the cellular neural/nonlinear network (CNN) has been much acclaimed as a powerful back-end analog array processor, capable of accelerating various computation-intensive tasks in image processing, pattern formation and recognition, motion detection, robotics, and various other real-time problem solving that requires complex computation [3]. In such real-world applications, massively parallel computation of spatial data over a 2-D surface is needed to process data in real-time, albeit computational functions are rather simple algebraic operations and each array element concurrently performs identical operation. The features of CNN that make it an easily implementable parallel computing architecture relative to fully connected neural networks are mentioned in [4] but reiterated here:

1) local interconnection between nodal components, with each node component called a cell or processing element; 2) regular placement of all cells in space; 3) identical cell configuration as well as space invariant interconnection; 4) real-time signal processing capability due to continuous-time dynamics and concurrent operation of cells; and 5) huge amount of templates to be exploited for various image processing algorithms. Numerous CNN implementations with versatile embedded applications have been developed using CMOS technology due to its low cost and high integration capability [5]–[8]. However, during the last two decades, CMOS technology has advanced by leaps and bounds so much so that it will encounter physical and manufacturing limitations, thereby ending the era of scaling down the transistor. To sustain the exponential growth of the integrated circuits as espoused by Moore's Law, several meso- and nano-scale technologies have been investigated to overcome the limitations of CMOS technology. Among several proposed nano-electronics devices, the resonant tunneling diode (RTD) has been explored for sometime due to its relatively easy fabrication process along with its unique folded-back negative differential resistance (NDR) current-voltage (I-V) characteristics [9]–[11]. The RTD has found several applications in both digital and analog circuits [12], [13].

Another advantage of RTD's include the possibility of cointegration with InP [14] or GaAs [15] three terminal devices such as HFETs, HEMTs, and HBTs that have one or two orders of higher electron mobility than CMOS. Previous work has shown that circuits implemented with RTDs have very fast operating speed, compact integration density, and rich functionality [12], [16]. These promising advantages the RTD possesses have spurred interest in its development for massively parallel architectures such as the CNN.

A compact bistable CNN architecture comprising the well-known monostable-bistable logic element (MOBILE) circuit models, first proposed by Maezawa *et al.* [16], were studied by Hanggi *et al.* [17]. Their preliminary study of RTD-based CNN array with 10^5 or more cells can be integrated in a standard CMOS chip. Dogaru *et al.* further extended the previous work by proposing an RTD based CNN cell configuration capable of performing various types of Boolean functions [18]–[20]. The operation principle of the RTD-based cell was clearly explained by them; however, no full-array simulation of a 2-D CNN array was presented in that work, and no image processing algorithms were simulated on a full array. Although subsequently Itoh *et al.* reported a full-array simulation for RTD-based CNNs [21], a simplistic piecewise linear model was used to represent the tunneling I-V characteristic of the RTD, thereby failing to precisely estimate the dynamics of the CNN in the real case.

P. Mazumder and I. E. Ebong are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: mazum@eecs.umich.edu).

S. R. Li was with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA.
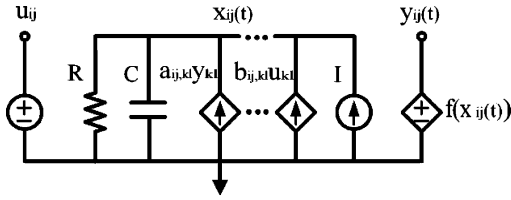
Fig. 1. Circuit model for the conventional CNN.

This paper extends previous works by presenting a comprehensive study of an RTD-based CNN architecture's operation. Section II deals with the study of the operation principle of a practical CNN. Section III discusses the stability and settling time issues of this architecture, and Section IV demonstrates full array simulation results utilizing a physics-based RTD model. Several image processing functions are demonstrated on both 12×12 and 128×128 CNN arrays by using an augmented SPICE simulator, Q-SPICE, developed at the University of Michigan.

Finally, it must be noted that though in this paper the NDR devices are realized with the intraband double barrier RTD, the CNN architecture proposed can be constructed with a host of other NDR devices [22], [23].

II. CNN OPERATION PRINCIPLE

A. CNN Based on Chua and Yang's Model

The cell model for the conventional CNN proposed by Chua and Yang (see Fig. 1) consists of one linear resistor, one linear capacitor, several linear voltage controlled current sources representing the feedback and feed-forward currents supplied by the neighboring cells, one independent current source, and one nonlinear voltage controlled voltage source.

According to this model, every cell acts as a nonlinear dynamic system, with its transient behavior governed by the following nonlinear ordinary differential equation:

$$C\frac{dx_{ij}(t)}{dt} = -\frac{x_{ij}(t)}{R}$$
$$+ \sum_{C(k,l)\in N_r(i,j)} (a_{ij,kl}f(x_{kl}(t)) + b_{ij,kl}u_{kl}) + I$$
$$1 < i < M; \quad 1 < j < N \quad (1)$$
$$f(x_{ij}) = 0.5 \times (|x_{ij} + 1| - |x_{ij} - 1|) \quad (2)$$

where $N_r(i,j)$ represents the neighborhood of cell (i,j) in an $M \times N$ CNN array; $x_{ij}(t)$, u_{ij}, and $f(x_{ij})$ are state, input, and output variables of cell (i,j), respectively. $a_{ij,kl}$ and $b_{ij,kl}$ are the space invariant feedback and feed-forward parameters, providing the weighting for the feedback and feed-forward currents from cell (k,l) to cell (i,j). The number of elements in the set formed by feedback parameters $a_{ij,kl}$ and in the set formed by feed-forward parameters $b_{ij,kl}$ depends on how far a central cell, cell (i,j), is connected to its adjacent cells, which is usually formulated as $(2r+1)^2$, where r is an integer starting from 1. For example, if r is equal to 1, there are 9 elements in each of the feedback and feed-forward parameter set: each cell communicates with its nearest 8 neighboring cells through feedback and feed-forward branches whereas it also contains a self-feedback loop and a self-feed-forward branch. Due to the space invariant
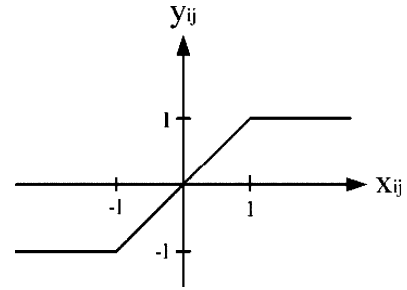
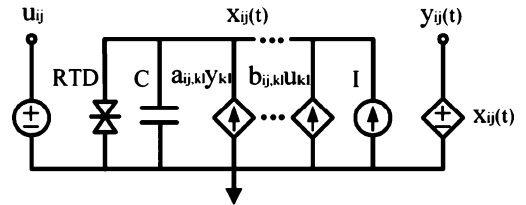

Fig. 2. Sigmoid state-to-output transfer function.

Fig. 3. Circuit model for the RTD-based CNN.

connection between cells in the CNNs, $a_{ij,kl}$ and $b_{ij,kl}$ are usually denoted by a pair of matrices of synaptic elements (e.g., 3×3 matrix for $r = 1$; 5×5 matrix for $r = 2$), called feedback template, $\mathbf{A}$, and feed-forward template, $\mathbf{B}$, respectively. In this work, we consider only the case where $r = 1$ since it generates the simplest structure. When the CNN is employed in image processing applications, its templates $\mathbf{A}$ and $\mathbf{B}$ merely act as image filters in order to map the input image into the desired output image. As a result, by designing templates of different matrix coefficients, the wide gamut of image processing algorithms can be executed by running an ordered sequence of different $\mathbf{A}$ and $\mathbf{B}$ matrices.

In conventional CNNs, the input and state variables are continuous values (analog signal) whereas the output variable is a binary value at steady state, either $+1$ or -1, according to the state-to-output transfer function $f(x_{ij})$. Fig. 2 graphically depicts this function as a piecewise linear function that saturates at either $+1$ or -1 when the state variable is greater than $+1$ or less than -1. The last component in the model presented in Fig. 1 is the independent current bias I; it is associated with each node as externally injected current which provides design flexibility for the CNN design.

B. CNN Equations Based on RTD's Model

The RTD is a symmetric two-terminal meso-scale device with non-monotonic I-V characteristics and extremely small intrinsic capacitance, which render its suitability in compact and high-speed CNN implementations. The RTD-based CNN cell model described in this paper is illustrated in Fig. 3, where one RTD replaces the linear resistor in the conventional CNN cell model.

Contrasting Figs. 1 and 3, the RTD introduces a nonlinearity that the resistor did not have. With the CNN in Fig. 1, the sigmoid representation in Fig. 2 is necessary for correct operation in representing the relationship between the input and output states, while with the RTD representation, the Fig. 2 transfer function is no longer necessary. The RTD would guarantee saturation due to its nonlinearity. This simplifies the relationship
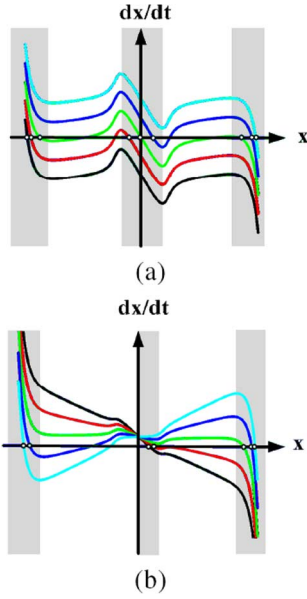
Fig. 4. Driving-point plots for (a) $a_{ij,ij} = 20\ \mu$A/V, $w = 100\ \mu$A, 50 μA, 0 μA, $-50\ \mu$A $-100\ \mu$A; (b) $w = 100\ \mu$A, $a_{ij,ij} = 400\ \mu$A/V, 200 μA/V, 0 μA/V, 200 μA/V, 400 μA/V.
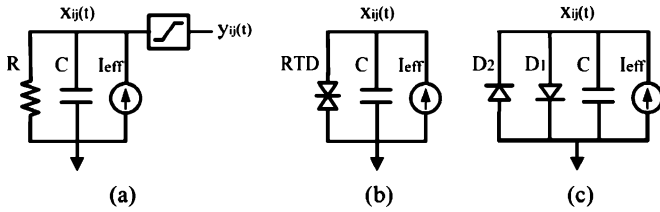

Fig. 5. Simplified circuit model for: (a) the conventional CNN; (b) the RTD-based CNN; and (c) the two-diode-based CNN.

between the state variable and output variable, making them equivalent to each other. Equation (1) is then modified to

$$C\frac{dx_{ij}(t)}{dt} = -h(x_{ij}(t))$$
$$+ \sum_{C(k,l)\in N_r(i,j)} (a_{ij,kl}x_{kl}(t) + b_{ij,kl}u_{kl}) + I. \quad (3)$$

$h(x)$ represents the RTD's I-V characteristics. An accurate physics model developed by Schulman *et al.* [24] is used to model $h(x)$:

$$h(x_{ij}(t)) = A \cdot \ln\left[\frac{1 + e^{(B-C+n_1 x_{ij}(t))q/kT}}{1 + e^{(B-C-n_1 x_{ij}(t))q/kT}}\right]$$
$$\cdot \left[\frac{\pi}{2} + \tan^{-1}\left(\frac{C - n_1 x_{ij}}{D}\right)\right]$$
$$+ H \cdot (e^{n_2 x_{ij}q/kT} - 1). \quad (4)$$

Equation (4) is current per unit area, and the parameters A, B, C, D, H, n_1, and n_2 depend on the physical model of the RTD. Equation (3) states that the current flowing into the capacitor is equal to the sum of the currents flowing out of the RTD, the feedback branches, the feed-forward branches, and the constant current bias. Referring to Fig. 3, this would be a Kirchhoff's Current Law (KCL) at the node denoted by the current state.

The RTD-based CNN's ability to perform the same functions as the conventional cell without the sigmoid function is explained through driving point plots in Fig. 4. The plots in Fig. 4 were generated for uncoupled CNNs—$a_{ij,kl} = 0$ for $kl \neq ij$. The x-axis indicates the state variable presented as a voltage while the y-axis indicates the derivative of the state variable with respect to time, which can be regarded as the normalized current flowing through the capacitor C in Fig. 3.

A parameter introduced in Fig. 4, w, is defined as

$$w = \sum_{C(k,l)\in N_r(i,j)} b_{ij,kl}u_{kl} + I. \quad (5)$$

This definition along with the uncoupling assumption thereby lets us write (3) as

$$C\frac{dx_{ij}(t)}{dt} = -h(x_{ij}(t)) + a_{ij,ij}x_{ij}(t) + w. \quad (6)$$

Fig. 4(a) shows the plots for different w—multiple combinations of input variables u_{kl}, feed-forward parameter $b_{ij,kl}$, and constant bias I—with a fixed self-feedback parameter $a_{ij,ij}$. Fig. 4(b) shows how the I-V relationship changes with different $a_{ij,ij}$ while other parameters stay the same.

The intersection points of the driving-point curves and the x-axis are the equilibrium states of the cell. The stable equilibrium states—marked with open circles—are distributed in three domains separated by the NDR regions of the RTD. Therefore the steady state of each cell can be assigned to three levels—$-1, 0$, or $+1$—by an analog-to-digital converter. Such conversion can be done using RTD-based quantizers whose threshold voltages are designed in the NDR regions [25].

With an RTD-based CNN, a need for a quantizer may not be necessary. From Fig. 4(a), when $w = 50\ \mu$A, we only have two equilibrium points. With proper design, a binary output (two stable points) can be obtained without the need of a quantizer.

C. Comparison Between Different CNN Models

To demonstrate the superior advantages of the RTD-based CNN cell, a comparison is made with a two diode model as well as conventional cell based on CMOS technology. Fig. 5 shows the three circuit models used in the comparison.

The two-diode model is introduced because the antiparallel configuration of the diodes—D_1 and D_2 in Fig. 5(c)—provides a similar saturating effect on the state variable as the RTD does. The assumptions for comparison are as follows.

- The circuit implementation for the RTD and the two diodes is the same for these three models in terms of device counts, area, power consumption, and the capacitance being driven.
- The total injected current from the feedback, the feed-forward, and the constant current bias is the same—represented as I_{eff} in Fig. 5.

Table I lists the comparison of circuit performance based on the simulation results, where the resistor, sigmoid circuits, and diodes are designed using 250-nm CMOS technology. In the conventional CNN cell, the sigmoid circuit would occupy a large amount of area to compensate for device mismatch issues. With continual CMOS scaling, the device mismatch issue is worsened since small atomic displacements render the operation of identical devices different. On the other hand, although the two-diode model holds similar nonlinearity as the

TABLE I
COMPARISON BETWEEN THE THREE CNN MODELS

	Conventional	RTD based	Two diode based
C		1pF	
R	5kΩ-10kΩ	Negligible	Negligible
Intrinsic Capacitance	Negligible	$0.75aF - 0.75fF$	$10pF - 110pF$
Estimated additional area	$>10\mu m^2$	$0.01 - 1\mu m^2$	$0.18 - 2\mu m^2$
Settling time (90%) zero initial condition with I_{eff}=200mA	20ns	$6.9ns - 12.5$ ns	$77.3ns - 716ns$
Saturated output voltage	-2V 2V	-1.8V to -1.4V 1.4V to 1.8V	-0.8V to -0.7V 0.7V to0.8V
# of steady states (Design flexibility)	2	3	2

Assumption: Same design, occupied area, power consumption for I_{eff} based on 0.25μm CMOS technology

RTD-based model, the intrinsic capacitance of the diodes is much larger than that of the RTD, resulting in slower operating speed. Moreover, the RTD-based model is the only one to have three output states due to its NDR property. This characteristic provides richer design flexibility than the other two methods. Table I summarizes the RTD-based model possesses the smallest area, smallest settling time, and gives the most design flexibility.

III. CIRCUIT ANALYSIS

The stability criteria and settling time analysis are two important factors regarding the design of a CNN architecture. Due to the complex interaction existing between cells, CNNs can be classified according to the feedback connection conditions: coupled CNNs and uncoupled CNNs ($a_{ij,kl} = 0$ for $kl \neq ij$). For uncoupled CNNs, the dynamic equation of one cell is represented as (6); this will enable our analysis henceforth.

A. Stability

Circuit stability analyses tend to use two methods, namely, Lyapunov theorem and graphing methods. According to the Lyapunov theorem, the energy function of the RTD-based CNN can be defined as [1]

$$E(t) = -\frac{1}{2}\sum_{(i,j)}\sum_{(k,l)} a_{ij,kl}x_{kl}(t)x_{ij}(t)$$
$$- \sum_{(i,j)}\sum_{(k,l)} b_{ij,kl}u_{kl}(t)x_{ij}(t)$$
$$- \sum_{(i,j)} Ix_{ij}(t) + \sum_{(i,j)} \int_0^{x_{ij}} h(s)ds. \qquad (7)$$

If a feedback template is symmetric $a_{ij,kl} = a_{kl,ij}$, the derivative of the energy function with respect to time t is

$$\frac{dE(t)}{dt} = -\sum_{(i,j)}\sum_{(k,l)} a_{ij,kl}x_{kl}(t)\frac{dx_{ij}(t)}{dt}$$
$$- \sum_{(i,j)}\sum_{(k,l)} b_{ij,kl}u_{kl}(t)\frac{dx_{ij}(t)}{dt}$$
$$- \sum_{(i,j)} I\frac{dx_{ij}(t)}{dt}$$
$$+ \sum_{(i,j)} \frac{dx_{ij}(t)}{dt}\frac{d}{dx_{ij}(t)}\int_0^{x_{ij}} h(s)ds \qquad (8)$$

$$\frac{dE(t)}{dt} = -\sum_{(i,j)} \frac{dx_{ij}(t)}{dt}$$
$$\times \left\{\sum_{(k,l)} a_{ij,kl}x_{kl}(t) + \sum_{(k,l)} b_{ij,kl}u_{kl}(t)\right.$$
$$\left. + I - \frac{d}{dx_{ij}(t)}\int_0^{x_{ij}} h(s)ds\right\} \qquad (9)$$

$$\frac{dE(t)}{dt} = -\sum_{(i,j)} C\left(\frac{dx_{ij}(t)}{dt}\right)^2 \leq 0. \qquad (10)$$

Equation (10) shows that the energy function is monotonically decreasing $|u_{ij}| < c_1, |x_{ij}(t=0)| < c_2, |x_{ij}(t)| < c$. Also, the energy function is bounded with certain constraints—$c_1, c_2,$ and c are constants. As a result, the state variable will also be bounded; dc output is always generated in the RTD-based CNN.

The graphical representation in Fig. 4 of the driving plots supports this conclusion. In Fig. 4(a), when w was changed, the plots shifted vertically, but each plot crossed the x-axis at least once. In Fig. 4(b), when $a_{ij,ij}$ was changed, the shape of the plots changed, but there is always an intersection with the x-axis. There exists at least one stable equilibrium state for every cell in the RTD-based CNN. Design flexibility comes into play because based on the number of equilibriums to design for, the method prescribed in [26] can be used to obtain a unary, binary, or ternary output.

B. Settling Time

CNNs prevail over sequential signal processors due to their real-time functional characteristics. Consequently, the operating speed is considered a critical performance index in CNN system design. To determine the operational speed of a dynamic system, the concept of settling time (t_s) is introduced. t_s of a cell is defined as the time required to reach a steady state (e.g., a stable equilibrium point) from the initial condition. The settling time of a CNN system is determined by the settling time of the slowest cell. The value of the settling time for a cell depends on several factors, such as initial condition, input, output, feedback, and feed-forward templates, constant bias, capacitance, and the size of the RTD (i.e., the current magnitude). Hence, analyzing the dependence of t_s on circuit parameters is the first step to design a high-speed CNN processor.
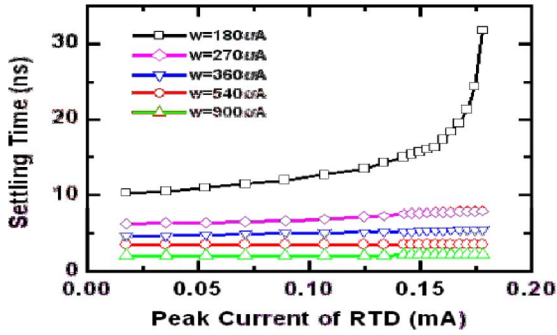
Fig. 6. Simulated settling time for various I_{peak} of RTD with $a_{ij,ij}$ for five different w.

Fig. 7. Simulated settling time for various $a_{ij,ij}$ with I_{peak} of RTD = 178 μA for five different w.

In this section, the settling time of one RTD-based CNN cell is investigated through SPICE simulation and mathematical modeling. Since the dynamics is very difficult to predict for propagated type CNN, this work focuses on the uncoupled RTD-based CNN. Without loss of generality, zero initial condition is assumed in our analysis.

1) Simulation Results: From (6), the settling time is dependent on w, $a_{ij,ij}$, and the area of the RTD, which govern the peak current of the RTD. Three cases are simulated to explore the relationship between the settling time and these three parameters. For the first case, we will investigate the dependence of settling time on the size of the RTD for five w values, 180, 270, 360, 540, and 900 μA, with $a_{ij,ij} = 0$. From the conclusion drawn in [26], we want the absolute value of w to be greater than the RTDs peak current (e.g., $I_{\text{peak}} = 178\ \mu$A for area $= 1\ \mu$m^2) for binary output. The simulation results illustrated in Fig. 6 show the settling time results for the first case.

In Fig. 6, the settling time has a weak dependence on peak current for the other w cases except for when $w = 180\ \mu$A. For this case, there is a drastic increase in settling time as peak current increases. This phenomenon may result from a very small current experienced by the capacitor during the trek from the initial state to the stable equilibrium state due to the small difference between the values of w and I_{peak}. Here we will define γ as the ratio of w/I_{peak}. When $1 < \gamma < 2$, the settling time greatly depends on I_{peak}. On the other hand, the settling time has a weak dependence on I_{peak} when $\gamma > 2$. From Fig. 6, cells with larger w take shorter time to achieve steady state.

In the second case, the settling time is simulated for various $a_{ij,ij}$ for the same five w values with an RTD area of 1 μm^2 ($I_{\text{peak}} = 178\ \mu$A). The results are illustrated in Fig. 7.

From Fig. 7, when $1 < \gamma < 2$, settling time decreases greatly as $a_{ij,ij}$ increases from 0 to 0.8 ms whereas the decreasing rate is not drastic when $\gamma > 2$. Note that as $a_{ij,ij}$ increases, the shape of the driving point plot changes, which increases the minimum current experienced by the capacitor, thus reducing the settling time.

The last case investigates the dependence of the settling time on the values of w. Different RTD sizes are used to modulate the peak current; the results are shown in Fig. 9.

From Fig. 8, the settling time seems to saturate when $w > 540\ \mu$A no matter what I_{peak} and $a_{ij,ij}$ are. Moreover, the set-
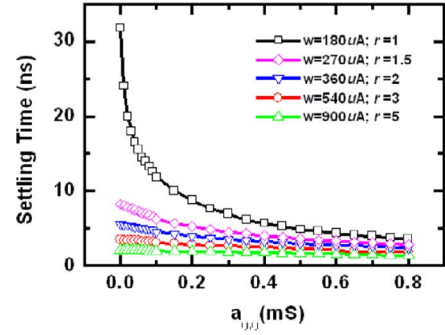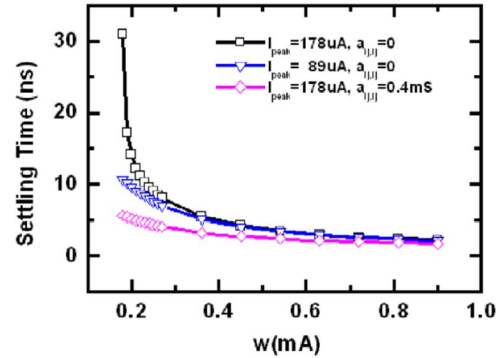

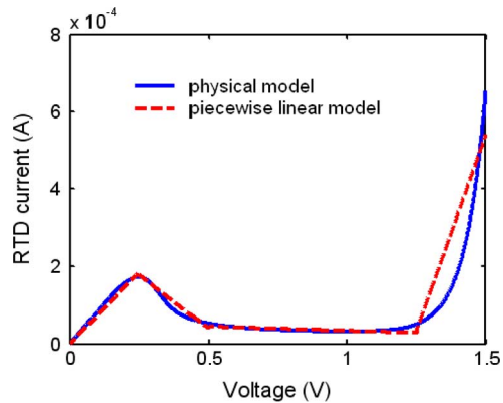Fig. 8. Simulated settling time for various w with different combinations of $a_{ij,ij}$ and I_{peak}.

Fig. 9. Physical model and piecewise linear model of the I-V curve of the RTD.

TABLE II
PARAMETERS FOR THE PWL MODELING

	A		A/V
a_0	0	a_1	6.96e-4
b_0	2.95e-4	b_1	-4.84e-4
c_0	5.41e-4	c_1	-2.27e-6
d_0	-2.93e-3	d_1	2.41e-3

tling time is reduced if gamma increases (from the black curve to the blue one) or $a_{ij,ij}$ increases (from the black curve to the violet one). This observation agrees with the previous experiments due to the same reason that w and $a_{ij,ij}$ have better control on the settling time as the effect of the RTD decreases. $a_{ij,ij}$ might
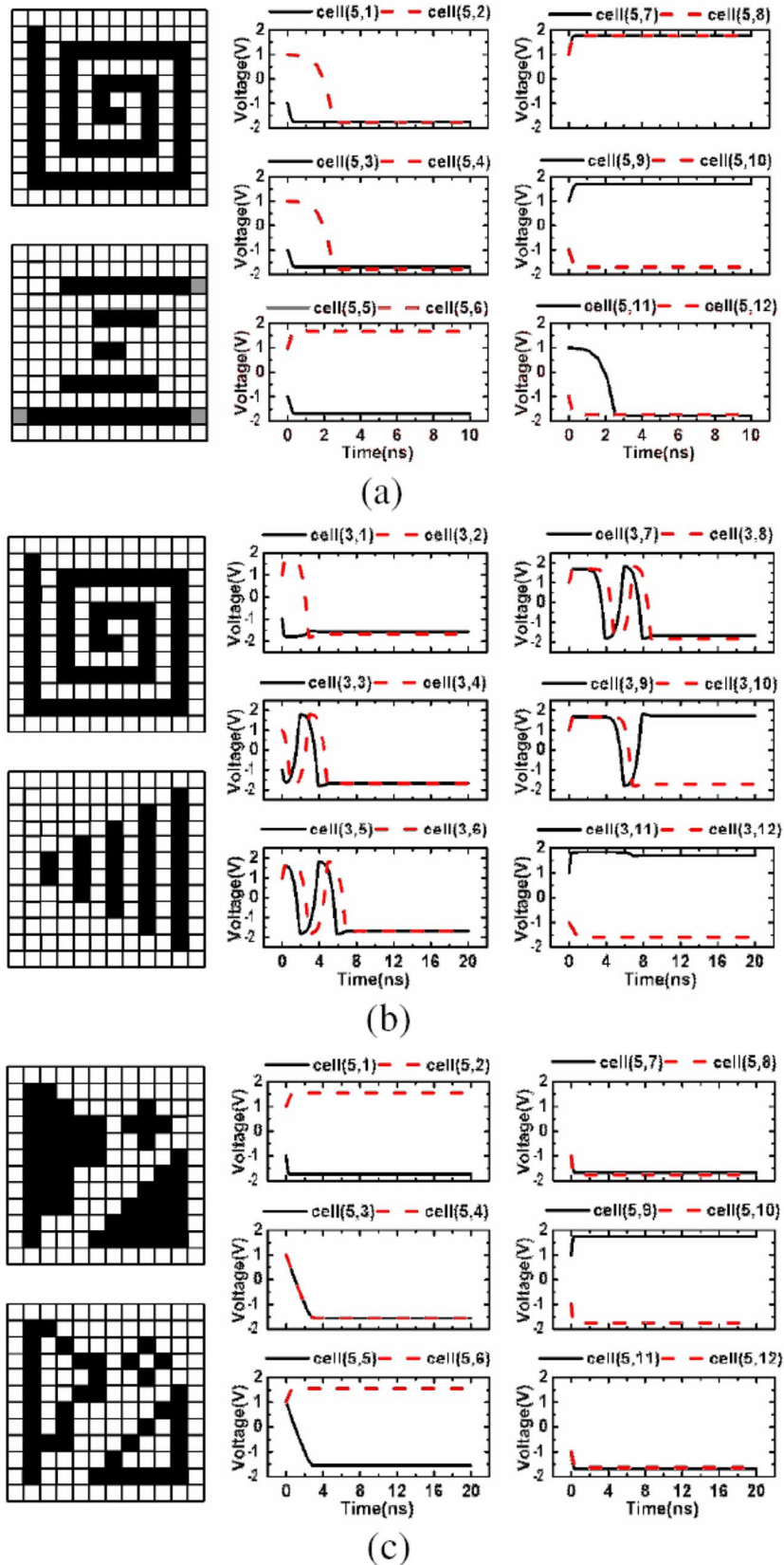
Fig. 10.   Simulation results of 12×12 RTD based CNN arra with the input and output image patterns for: (a) horizontal line detection; (b) horizontal physical model and piecewise linear model of the I-V curve of the RTD; (c) edge extraction.

not be controllable but since w depends on the constant current source I, we have more design control over this than feedback/ feedforward parameters.

2) Mathematical Modeling: Since the dynamic behavior of a single RTD-based CNN is determined by a nonlinear differential equation, it is very difficult to find an analytic solution

for the transient response of the state variable. Even though a complex analytic solution can be determined, it may not make any sense to us since we may not be familiar with the complex function. Hence, this work simplifies the settling time analysis by using a piecewise linear model for the I-V curve of the RTD as illustrated in Fig. 9. This method can provide an intuitive understanding for the transient response of a cell from the initial condition to the steady state as well as determine which parameters significantly affect the settling time.

From Fig. 9, the physical model of the RTD used in the simulation has a peak around 0.25 V and a valley current starting at 0.5 V followed by a wide valley region. In addition, the diode-like current activates at around 1.25 V. Therefore, the I-V curve of the RTD is modeled with a piecewise linear (PWL) function, with each section following:

$$h(x) = \begin{cases} a_0 + a_1 x & 0 \text{ V} \leq x < 0.25 \text{ V} & a_0 = 0, \quad a_1 > 0 \\ b_0 + b_1 x & 0.25 \text{ V} \leq x < 0.5 \text{ V} & b_0 > 0, \quad b_1 < 0 \\ c_0 + c_1 x & 0.5 \text{ V} \leq x < 1.25 \text{ V} & c_0 > 0, \quad c_1 < 0 \\ d_0 + d_1 x & 1.25 \text{ V} \leq x & d_0 > 0, \quad d_1 > 0. \end{cases}$$
(11)

The fitting parameters for (11) are summarized in Table II. The nonlinear differential equation describing the dynamics of one cell, thus, is decomposed into four linear ordinary differential equations (ODE), with each representing some period of the trek from the zero initial state to the steady state. Therefore, the settling time can be solved sequentially from one ODE to another ODE to obtain four solutions to (6) depending on initial and final conditions. The results of each of the four sections were combined linearly to get an overall approximation of settling time as depicted in (12)

$$t_s = \left(\frac{C}{a_{ij,ij} - d_1} \right) \cdot \ln \left[\frac{x_{ij,\text{steady}} - \left(\frac{d_0 - w}{a_{ij,ij} - d_1} \right)}{1.25 - \left(\frac{d_0 - w}{a_{ij,ij} - d_1} \right)} \right]$$

$$+ \left(\frac{C}{a_{ij,ij} - c_1} \right) \cdot \ln \left[\frac{1.25 - \left(\frac{c_0 - w}{a_{ij,ij} - c_1} \right)}{0.5 - \left(\frac{c_0 - w}{a_{ij,ij} - c_1} \right)} \right]$$

$$+ \left(\frac{C}{a_{ij,ij} - b_1} \right) \cdot \ln \left[\frac{0.5 - \left(\frac{b_0 - w}{a_{ij,ij} - b_1} \right)}{0.25 - \left(\frac{b_0 - w}{a_{ij,ij} - b_1} \right)} \right]$$

$$+ \left(\frac{C}{a_{ij,ij} - a_1} \right) \cdot \ln \left[\frac{0.25 - \left(\frac{-w}{a_{ij,ij} - d_1} \right)}{\frac{w}{a_{ij,ij} - a_1}} \right].$$
(12)

As can be seen from (12), the settling time depends on the self-feedback parameter, steady state target, $a_{ij,ij}$, w, and the size of the RTD (related to the parameters $(a_1, b_0, b_1, c_0, c_1, d_0, d_1)$. The settling time is proportional to the capacitance but is inversely proportional to $a_{ij,ij} - a_1, a_{ij,ij} - b_1, a_{ij,ij} - c_1$, and $a_{ij,ij} - d_1$. Another interesting phenomenon is that the impact of w is reduced due to the likely cancellation (if w dominates) of the numerator and denominator together with the log function. This result in conjunction with the results from Figs. 7 and 8 show that when w is large enough (greater than the peak current), additional increases in w does not affect the settling time tremendously.
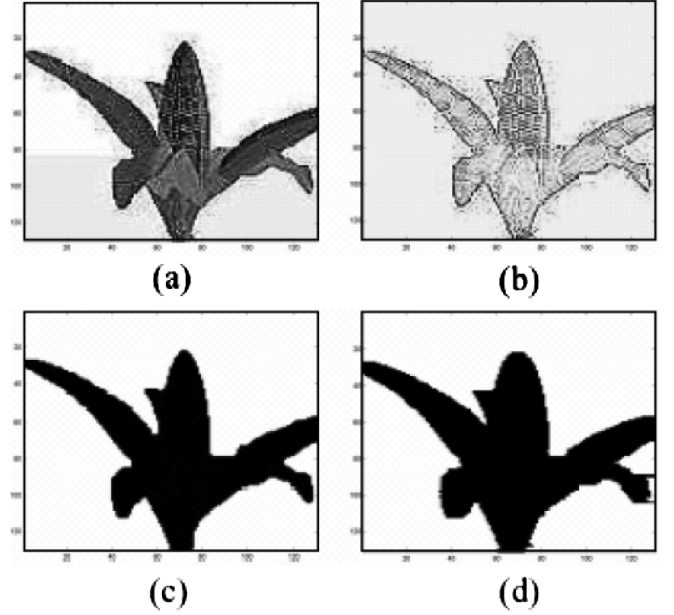

Fig. 11. Simulation results for 128×128 RTD-based CNN array. (a) Input image. (b) Output image for EE. (c) Output image for HF. (d) Output image for shadowing.

IV. SIMULATION RESULTS

In this work, a method of quantifying the effects of parameters that influence stability and settling time were presented. A simplification to uncouple the effects of neighboring cells were made with knowledge this method will introduce a percentage error between the values obtained during design and those acquired from simulation. A small 12×12 RTD-based CNN was designed in order to verify and quantify the errors associated with the assumptions made during analysis. The expression in (12) was used to appropriately design settling times in the nanosecond range. The calculated results were on the same order as what was observed during simulation. The results are shown in Fig. 10.

Fig. 10 was obtained through simulation results performed through SPICE simulation tools, where RTD's are represented by current sources describing the physics-based models and the feedback and feed-forward branches are modeled as ideal voltage controlled current sources. The boundary condition for the cells at the edge is zero, i.e., no feedback and feed-forward loops from outside of the array. In Fig. 11, the input voltages of black and white cells are $+1$ and -1 V, respectively, where as the output of each cell converges to values between $+1.5$ and $+1.8$ V or -1.5 and -1.8 V, representing $+1$ (black) or -1 (white) after A/D conversion.

The dynamics of the state variable of each cell in a full row [e.g., the fifth row for Fig. 10(a); the third row for Fig. 10(b); the fifth row for Fig. 10(c)] is demonstrated in the waveform. As can be seen, for the uncoupled image processing functions (e.g., the horizontal line detection and the edge extraction), the settling time is around a few nanoseconds with 1 pF capacitance. On the other hand, for the propagated type algorithm such as the horizontal connected component detection, the interaction between
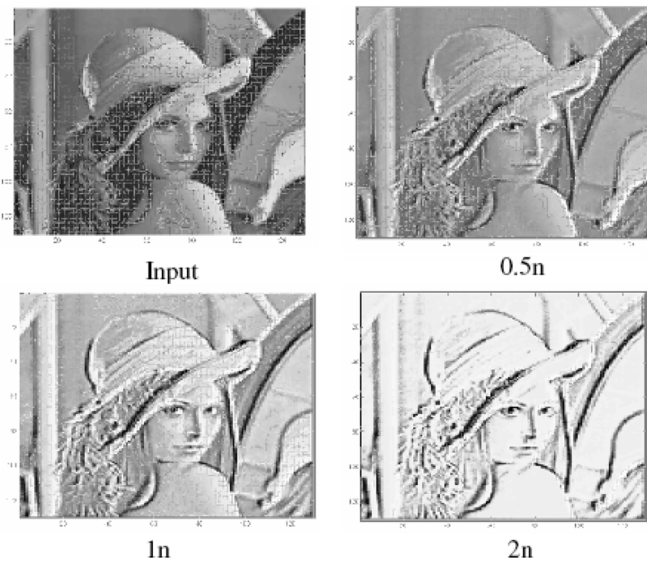
Fig. 12. Simulation results for 128×128 RTD-based CNN array for EE.
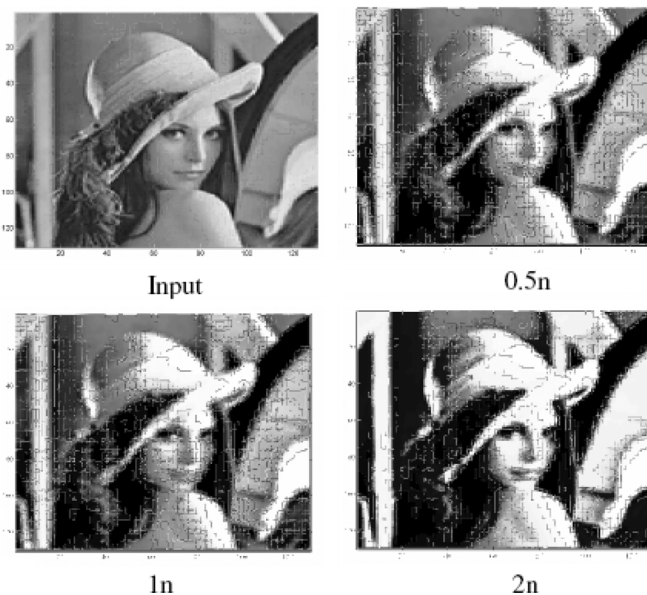

Fig. 13. Simulation results for 128×128 RTD-based CNN array for averaging.

neighboring cells can be obviously observed from the transient responses of state variables, which contributes to a much longer settling time.

For image processing on a grander scale, a full array for a 128×128 RTD-based CNN with grayscale input images was simulated. Fig. 11 demonstrates the results of different functions—edge extraction (EE), hole filling (HF), and shadowing—on the same input image.

In addition to the input image in Fig. 11, another image shown in Figs. 12 and 13 is used to show the intermediate process between the input and output images. These are different snapshot images at different timestamps of the transient response of the full array simulation.

V. CONCLUSION

Several image processing functions—EE, HF, shadowing, HLD, HCCD—are successfully demonstrated for a 12×12 and a 128×128 RTD-based CNN array using the QSPICE simulator developed at the University of Michigan. The RTD-based CNN is proven to always be stable if the feedback template is symmetric. A method and expression for approximating settling time of RTD-based CNN cells for binary outputs is presented. The settling time analysis shows that when the ratio of the total injected or extracted current in a cell to the RTD's peak current approaches 1, settling time increases exponentially. Simulation results on the comparison of the RTD-based CNN with the conventional CNN and the two-diode implementation show that the RTD presents unique design advantages in terms of compactness, speed, and design flexibility.

ACKNOWLEDGMENT

REFERENCES

[1] L. O. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.*, vol. CAS-35, no. 10, pp. 1257–1272, Oct. 1988.
[2] L. O. Chua and L. Yang, "Cellular neural networks: Applications," *IEEE Trans. Circuits Syst.*, vol. CAS-35, no. 10, pp. 1273–1290, Oct. 1988.
[3] L. O. Chua and T. Roska, "The CNN paradigm," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 40, no. 3, pp. 147–156, Mar. 1993.
[4] K. R. Crounse and L. O. Chua, "Methods for image processing and pattern formation in cellular neural networks: A tutorial," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 42, no. 10, pp. 583–601, Oct. 1995.
[5] T. Roska and A. Rodriguez-Vazquez, "Toward visual microprocessors," *Proc. IEEE*, vol. 90, no. 7, pp. 1244–1257, Jul. 2002.
[6] K. Karahaliloglu and S. Balkir, "Bio-inspired compact cell circuit for reaction-diffusion systems," *IEEE Trans. Circuits Syst. II, Express Briefs*, vol. 52, no. 9, pp. 558–562, Sep. 2005.
[7] J. Kowalski, "0.8 μm CMOS implementation of weighted-order statistic image filter based on cellular neural network architecture," *IEEE Trans. Neural Netw.*, vol. 14, no. 5, pp. 1366–1374, May 2003.
[8] P. Kinget and M. S. J. Steyaert, "A programmable analog cellular neural network CMOS chip for high speed image processing," *IEEE J. Solid-State Circuits*, vol. 30, no. 3, pp. 235–243, Mar. 1995.
[9] K. Karahaliloglu and S. Balkir, "Nanostructure array of coupled RTDs as cellular neural networks," *Int. J. Circuit Theory Appl.*, vol. 31, no. 6, pp. 571–589, 2003.
[10] P. Julian, R. Dogaru, M. Itoh, M. Hanggi, and L. O. Chua, "Simplicial RTD-based cellular nonlinear networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 4, pp. 500–509, Apr. 2003.
[11] M. Hanggi and L. O. Chua, "Cellular neural networks based on resonant tunnelling diodes," *Int. J. Circuit Theory Appl.*, vol. 29, no. 5, pp. 487–504, 2001.
[12] P. Mazumder, S. Kulkarni, M. Bhattacharya, J. P. Sun, and G. I. Haddad, "Digital circuit applications of resonant tunneling devices," *Proc. IEEE*, vol. 86, no. 4, pp. 664–686, Apr. 1998.
[13] Y. Tsuji and T. Waho, "Design of flash analog-to-digital converters using resonant-tunneling circuits," *IEICE Trans. Electron.*, vol. E87C, no. 11, pp. 1863–1868, 2004.

[14] J. I. Bergman, J. Chang, Y. Joo, B. Matinpour, J. Laskar, N. M. Jokerst, M. A. Brooke, B. Brar, and E. Beam III, "RTD/CMOS nanoelectronic circuits: Thin-film InP-based resonant tunneling diodes integrated with CMOS circuits," *IEEE Electron Device Lett.*, vol. 20, no. 3, pp. 119–122, Mar. 1999.

[15] Y. L. Huang, L. Ma, F. H. Yang, L. C. Wang, and Y. P. Zeng, "Resonant tunnelling diodes and high electron mobility transistors integrated on GaAs substrates," *Chinese Phys. Lett.*, vol. 23, no. 3, pp. 697–700, 2006.

[16] K. Maezawa, T. Akeyoshi, and T. Mizutani, "Functions and applications of monostable-bistable transition logic elements (MOBILE's) having multiple-input terminals," *IEEE Trans. Electron Devices*, vol. 41, no. 2, pp. 148–154, Feb. 1994.

[17] M. Hanggi and L. O. Chua, "Compact bistable CNNs based on resonant tunneling diodes," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2001, pp. 93–96.

[18] M. Hanggi, R. Dogaru, and L. O. Chua, "Physical modeling of RTD-based CNN cells," in *Proc. 6th IEEE Int. Workshop Cellular Neural Networks Their Appl. (CNNA)*, 2000, pp. 177–182.

[19] M. Hanggi, L. O. Chua, and R. Dogaru, "A simple RTD-based circuit for Boolean CNN cells," in *Proc. 6th IEEE Int. Workshop Cellular Neural Netw. Their Appl. (CNNA)*, 2000, pp. 189–194.

[20] R. Dogaru, M. Hanggi, and L. O. Chua, "A compact and universal cellular neural network cell based on resonant tunneling diodes: Circuit, model, and functional capabilities," in *Proc. 6th IEEE Int. Workshop Cellular Neural Netw. Their Appl. (CNNA)*, 2000, pp. 183–188.

[21] M. Itoh, P. Julián, and L. O. Chua, "RTD-based cellular neural networks with multiple steady states," *Int. J. Bifurcation Chaos*, vol. 11, no. 12, pp. 2913–2959, 2001.

[22] C. P. Gerousis and S. M. G. W. P. , "Nanoelectronic single-electron transistor circuits and architectures," *Int. J. Circuit Theory Appl.*, vol. 32, no. 5, pp. 323–338, 2004.

[23] S. Bandyopadhyay, K. Karahaliloglu, S. Balkir, and S. Pramanik, "Computational paradigm for nanoelectronics: Self-assembled quantum dot cellular neural networks," *IEE Proc. Circuits, Devices Syst.*, vol. 152, no. 2, pp. 85–92, 2005.

[24] J. N. Schulman, H. J. De Los Santos, and D. H. Chow, "Physics-based RTD current-voltage equation," *IEEE Electron Device Lett.*, vol. 17, no. 5, pp. 220–222, May 1996.

[25] S.-R. Li, P. Mazumder, and L. O. Chua, "On the implementation of RTD based CNNs," in *Proc. Int. Symp. Circuits Syst. (ISCAS)*, 2004, pp. III-25–III-28.

[26] S.-R. Li, P. Mazumder, and L. O. Chua, "Cellular neural/nonlinear networks using resonant tunneling diode," in *Proc. 4th IEEE Conf. Nanotechnol.*, 2004, pp. 164–167.

Pinaki Mazumder (F'99) received the Ph.D. degree from the University of Illinois at Urbana-Champaign, Urbana-Champaign, in 1988.

Currently, he is a Professor with the Department of Electrical Engineering and Computer Science, University of Michigan (UM), Ann Arbor. He is on leave for one year from the UM to serve as the lead Program Director of the Emerging Models and Technologies Program at the US National Science Foundation. He had worked for six years in industrial R&D centers that included AT&T Bell Laboratories, where in 1985 he started the CONES Project—the first C modeling-based VLSI synthesis tool, and India's premiere electronics company, Bharat Electronics Ltd., where he had developed several high-speed and high-voltage analog integrated circuits intended for consumer electronics products. He has published over 200 technical papers and 4 books on various aspects of VLSI research works. His research interests include current problems in Nanoscale CMOS VLSI design, CAD tools and circuit designs for emerging technologies including Quantum MOS and resonant tunneling devices, semiconductor memory systems, and physical synthesis of VLSI chips.

Dr. Mazumder was a recipient of Digital's Incentives for Excellence Award, BF Goodrich National Collegiate Invention Award, and DARPA Research Excellence Award. He is an AAAS Fellow (2008) and an IEEE Fellow for his contributions to the field of VLSI.

Sing-Rong Li was a graduate student at the University of Michigan, Ann Arbor. She left the University of Michigan in 2005 without completing the doctoral degree.

Idongesit E. Ebong received the B.S. and M.S. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, in 2006. He is currently working towards the Ph.D. degree in electrical engineering at the University of Michigan, Ann Arbor.

His research interests include digital/analog integrated circuit design, focused primarily on new devices and low power applications.