Implementation of PWL functions on the CNNUM

Víctor M. Preciado and Miguel A. Jaramillo

Departamento de Electrónica e Ingeniería Electromecánica,
Universidad de Extremadura,E-06071 Badajoz, Spain
{vpdiaz,miguel@unex.es}

Abstract. In this paper, the Cellular Neural Network Universal Machine (CNNUM) [7] is presented as a novel hardware architecture which make use of complex spatio-temporal dynamics [1] for solving real-time image processing tasks. Dealing with actual VLSI chip prototypes [6], we find the limitation of a fixed piecewise-linear (PWL) saturation output function. In this work, a novel algorithm for software emulation of any piecewise-linear (PWL) output functions on the CNNUM VLSI chip is presented.

1 Introduction

CNN topology is essentially characterized by a *local* interaction between non-linear dynamical cells distributed in a regular 2-D grid [1]. This fact makes the CNN an useful computation paradigm when the problem can be reformulated as a task where the signal values are placed on a regular 2-D grid, and interaction between signal values are limited within a finite local neighborhood [4]. Besides, local interaction facilitates the implemetation of this kind of networks as efficient and robust VLSI chips [5] ,[6]. The Cellular Neural Network Universal Machine (CNNUM) [7] is a programmable neuroprocessor based on CNN dynamics and implemented alongside photosensors which sense and process the image in a single VLSI chip. The main drawbacks encountered when using this chip in image processing application is the limitation in filtering capabilities to 3×3 dimension templates and the restricted possibilities provided by the fixed piecewise-linear (PWL) saturation output function in nonlinear filtering. Thus, we find in this framework a rigid relationship between an extremely high computing speed and a limitation in the complexity of the image processing operators.

In this work, the CNN dynamical model and the architecture of the Cellular Neural Network Universal Machine (CNNUM) prototyping system are introduced. Then, we describe a novel and general algorithm for achieving any type of PWL approximation of an arbitrary output function on the framework of the CNNUM chipset. By means of this methodology, we break the rigid relationship speed/complexity and it is provided a flexible framework to the designer of image processing algorithms.

2 CNNUM: Dynamics and Architecture

The dynamic of the array can be described by the following set of differential equations

$$\frac{d}{dt}x_{i,j}(t) = -x_{i,j}(t) + \sum_{k,l \in N_r} A_{k,l}\, y_{i+k,j+k}(t) + \sum_{k,l \in N_r} B_{k,l}\, u_{i+k,j+k}(t) + I \qquad (1)$$

with output nonlinear function

$$y(x) = \frac{1}{2}\left[|x-1| - |x+1|\right] \qquad (2)$$

The input,state and output, represented by $u_{i,j}, x_{i,j}$ and $y_{i,j}$ are defined on $0 \leq i \leq N_1$ and $0 \leq i \leq N_2$ and N_r represents the neighborhood of the cell with a radius r as $N_r = \{(k,j) : \max\{|k-i|, |l-j|\} \leq r\}$. The B template and I coeficient form a simple feedforward filtered (FIR) version of the input. On the other hand, the temporal evolution of the *dynamics network* is mathematically modeled by the A template operating in a *feedback* loop along with the fixed saturation nonlinearity previously defined.

The CNNUM architecture is an *analogic* (*analog+logic*) spatio-temporal array computer wherein analog spatio-temporal phenomena provided by the CNN and logic operations are combined in a programmable framework to obtain more sophisticated operation mode [7]. Every implemented neuron includes circuitry for CNN processing, binary and gray scaled images storage supplied by

- *a local analog memories (LAM)* which allow to save intermediate analog results of the algorithms.
- *a local logic memory (LLM)* distributed alongside analog processors
- *a local logic unit (LLU)* which permit logic operation among binary images
- *a local comunication and control unit (LCCU)* is the necessary configuration circuitry for electrical I/O and control of the different operations.

In this work we present an algorithmic way to deform the saturation output function defined in (2) in order to obtain a general piecewise-linear continuous function. Thus, it is convenient to study the stability of the dynamical network under the reshaped output function $y(x)$.

2.1 Stability Criteria

All image processing applications are based on the assumption that neither the oscillation nor the chaotic phenomenon is exhibited. In this section, some mathematical criteria which guarantee complete stability are presented.

Theorem 1 (State-Boundedness Criterion). *If the function $y(x)$ defined in (2) is continuous and bounded, then the state $x_{i,j}(t)$ of each cell of a standard CNN is bounded for all bounded threshold and bounded inputs.*

Proof. Equation (1) can be recast into the form

$$\frac{d}{dt}x_{i,j}(t) = -x_{i,j}(t) + g(t)$$

where

$$g(t) = \sum_{k,l \in N_r} A_{k,l} y_{i+k,j+k}(t) + \sum_{k,l \in N_r} B_{k,l} u_{i+k,j+k}(t) + I$$

Since both I and $u_{i,j}$ are bounded by hypotheses, there exists finite constant K such that

$$\max_{0<t<\infty} |g(t)| < K$$

It follows via Gronwall's Lemma that

$$\begin{aligned} |x_{i,j}(t)| &\leq \left| x_{i,j}(0) e^{-t} \right| + \left| \int_0^t e^{-(t-\tau)} g(\tau) d\tau \right| \\ &\leq |x_{i,j}(0)| \, e^{-t} + \max_{0<t<\infty} |g(t)| \int_0^t e^{-(t-\tau)} d\tau \\ &< |x_{i,j}(0)| + K, \ \forall t > 0. \end{aligned}$$

3 PWL Approximation by the Infinity Norm Criterion

When considering the VLSI CNN chip model, we deal with a rigid PWL saturation output function due to difficulties in implementing flexible non-linearities on silicon. In this Section, we present a general method to approximate any nonlinear output function on current CNNUM chips by superposition of piecewise-linear (PWL) saturation blocks as defined in (2).

3.1 Previous Definitions

The following notation is used: δ_{ij} denotes Kronecker delta, $B_{zo,r}$ denotes de open ball $B_{zo,r} := \{z \in Z : \|z - z_o\| < r\}$, $\|\cdot\|$ is the weighted Euclidean norm defined as $\|z\| = \left(\sum_{i=1}^{n} \omega_i z_i^2\right)^{1/2}$, with $\omega_i > 0$, $\|\cdot\|_\infty$ the weighted infinity norm. Increment and sum of succesive function in an indexed list is denoted by $\Sigma h_i := h_{i+1} + h_i$ and $\Delta h_i := h_{i+1} - h_i$, and the simbol ' denotes differentiate on variable x.

3.2 PWL Approximation by the Chebyshev Criterion

In order to approximate a desider output function, a superposition of piecewise-linear saturation functions are considered in the following structure:

$$\overline{f}(x) = \sum_{i=1}^{\sigma} \left(\frac{1}{b_i} y(a_i x - c_i) + m_i \right) \tag{3}$$

where $y(x)$ is defined in (2). Each summing term consists in a linear transformation of the same PWL original saturation function; thus, $\overline{f}(x)$ yields an adjustable PWL function where $a_i, b_i, c_i, m_i \in \mathbb{R}$, are the parameters of the structure. Basically, (3) is a nonlinear combination of linear affine lines,

$$\Pi_i := \frac{1}{2}\left[|(a_i x - c_i) - 1| - |(a_i x - c_i) + 1|\right], i \in [1, \sigma] \tag{4}$$

The problem under study can be stated as follows: Given a smooth function $f : S \to \mathbb{R}$, where $S \subset \mathbb{R}$ is compact, we want to design a PWL function $\overline{f}$ that minimizes the error between f and $\overline{f}$ in some sense. Formally, given a fixed number ε^* we want to find the optimal parameter vector $\theta^* = [a_i^*, b_i^*, c_i^*, m_i^*]$ that makes the objective functional $\mathfrak{J} := \left\|f(x) - \overline{f}(x)\right\|_\infty = \varepsilon^* \quad \forall x \in S$, with the most efficient shape.

The functional based on the infinite norm $\|\cdot\|_\infty$ is supported by a physical effect observed in the implementation of CNNUM VLSI chips: *the analog signals involved in the processing task are computed with an analog accuracy of 7 bits of equivalent digital accuracy [5].* Thus, this fact gives us a fixed value for $\varepsilon \leq 2^{-7}$, which assure that the error introduced by the approximation procedure is less or equal that the error introduced by the performance of the implemented chip.The functional proposed in this paper is an alternative to the $\left\|f(x) - \overline{f}(x)\right\|$ functional studied in several papers [3], [2]. This cuadratic criterion yields a nonlinear optimization problem characterized by the existence of several local minima. One practical technique used to undertake this serious problem consist in the use of iterative algorithms which produce new random search direction when a local minimum in reached.

The point of departure used to obtain the approximating PWL function based on the infinity norm is the following

Theorem 2 (Minimax). *Let $f(x)$ be a function defined in the open subset (x_i, x_{i+1}), $x_i, x_{i+1} \in \mathbb{R}$ and $P_n(x)$ a polynomial with grade n. Then $P_n(x)$ minimizes $\|f(x) - P_n(x)\|_\infty$ if and only if $f(x) - P_n(x)$ takes the value $\varepsilon := \max(|f(x) - P_n(x)|)$ at least in n+2 points in the interval (x_i, x_{i+1}) with alternating sign.*

Theorem 3. *Let $f(x)$ be a function with $f'' > 0$ in the interval (x_1, x_2), $x_1, x_2 \in \mathbb{R}$ and $P_1(x) := Mx + B$. Then $P_1(x)$[1] minimizes $\|f(x) - P_1(x)\|_\infty$ if and only if*

$$M = \Delta f / \ \Delta x_1 ;\ B = \frac{1}{2}\left[f(x_2) + f(x_a) - \Delta f_i / \ \Delta x_i \,(x_a + x_2)\right] \tag{5}$$

where x_a is obtained by solving

$$f'(x_a) = \Delta f_i / \Delta x_i \tag{6}$$

[1] This straight line is called *Chebyshev line* in the literature.

Proof. It follows from minimax theorem that it must be three points x_l, x_c, x_r in (x_1, x_2) which maximize $E(x) := f(x) - P_1(x)$. This condition implies that x_c is an intermediate point in the interval (x_1, x_2) with $E'(x)|_{x_c} = 0$; this is the same that $f'(x)|_{x_c} = M$. Since $f''(x) > 0$, $f'(x)$ is a strictly growing function and can equate M only once, this means that x_c is the only one intermediate point which minimizes E in the interval; thus $x_l = x_1$ and $x_r = x_2$. Applying the minimax condition we obtain $E(x_l) = -E(x_c) = E(x_r)$ and by solving these equations we can conclude

$$M = \Delta f_i / \ \Delta x_i; \ B = \frac{1}{2}\left[f(x_{i+1}) + f(x_a) - \Delta f_i / \ \Delta x_i \ (x_a + x_{i+1})\right] \qquad (7)$$

Corollary 1. *Under the previous conditions,* $\varepsilon := \|f(x) - P_1(x)\|_\infty$ *is given by*

$$\varepsilon = f(x) - \left[\frac{\Delta f_i}{\Delta x_i}x + \frac{1}{2}\left(f(x_{i+1}) + f(x_{ai}) - \frac{\Delta f_i}{\Delta x_i}(x_{ai} - x_{i+1})\right)\right] \qquad (8)$$

Remark 1. From the proof of this theorem it can be advised that in the case of $f'' < 0$, $\varepsilon = -E(x_l) = E(x_c) = -E(x_r)$

Theorem 4. *Let* $f(x)$ *be a function with* $f'' > 0$ *in the interval* (x_a, x_b), $x_a, x_b \in \mathbb{R}$, ε^* *an arbitrary small real number and* $\overline{f}(x) = \sum_{i=1}^{\sigma} \Pi_i$, *where* $\Pi_i := \frac{1}{2}\left[|(a_i x - c_i) - 1| - |(a_i x - c_i) + 1|\right]$, $i \in [1, \sigma]$; $a_i, b_i, c_i, m_i \in \mathbb{R}$, $i \in [1, \sigma]$. *Then* $\overline{f}(x)$ *makes* $\left\|f(x) - \overline{f}(x)\right\|_\infty = \varepsilon^*$ *minimizing the number of summing terms* σ *if the parameters of* $\overline{f}(x)$ *fulfill the following conditions:*

$$a_i = 2/ \ \Delta x_i, \ b_i = 2/ \ \Delta f_i, \ c_i = \Sigma x_i / \ 2, \ i \in [1, \sigma]; \qquad (9)$$

$$m_1 = \Sigma f_i / \ 2 - \varepsilon^*, \ m_j = \Sigma f_j / \ 2 - f(x_j) - \varepsilon^*, \ j \in [2, \sigma] \qquad (10)$$

where x_i *is obtained from the following set of discrete equations:*

$$\varepsilon^* - \frac{1}{2}\left[x_i + \frac{\Delta f_i}{\Delta x_i}(x_{ai} - x_i) - f(x_{ai})\right] = 0 \ \textit{being} \ f'(x_{ai}) = \frac{\Delta f_i}{\Delta x_i}, \ i \in [1, \sigma] \quad (11)$$

Proof. In order to demonstrates this theorem we can express Π_i as

$$\Pi_i := \begin{cases} m_i - b_i^{-1}, \ \forall x \in \left[c_i + a_i^{-1}, \infty\right) \\ m_i + \frac{\Delta f_i}{\Delta x_i}(x - c_i), \ \forall x \in B_{c_i, a_i^{-1}} \\ m_i + b_i^{-1}, \ \forall x \in \left(-\infty, c_i - a_i^{-1}\right] \end{cases}$$

Replacing the values of the parameters given in the statement of the theorem

$$\Pi_i := \begin{cases} \delta_{1i}\left(f(x_i) - \varepsilon^*\right), \ \forall x \in \left[c_i + a_i^{-1}, \infty\right) \\ \delta_{1i}\left(f(x_i) - \varepsilon^*\right) + \frac{\Delta f_i}{\Delta x_i}(x - x_i), \ \forall x \in B_{c_i, a_i^{-1}} \\ \delta_{1i}\left(f(x_i) - \varepsilon^*\right) + \Delta f_i, \ \forall x \in \left(-\infty, c_i - a_i^{-1}\right] \end{cases}$$

If we consider $x_a \in (x_j, x_{j+1})$ and expand $\overline{f}(x_a)$ taking into account the value of ε^* given in Corollary 1, it is obtained

$$\begin{aligned}
\overline{f}(x_a) &:= \Pi_1 + \textstyle\sum_{i=2}^{j-1} \Pi_i + \Pi_j + \sum_{i=j+1}^{\sigma} \Pi_i \\
&= (f(x_1) - \varepsilon^*) + \textstyle\sum_{i=2}^{j-1} \Delta f_i + \left[\frac{\Delta f_j}{\Delta x_j}(x - x_j)\right] \\
&= f(x_j) - \varepsilon^* + \frac{\Delta f_j}{\Delta x_j}(x - x_j) \\
&= \frac{\Delta f_i}{\Delta x_i} x + \frac{1}{2}\left[f(x_{i+1}) + f(x_a) - \frac{\Delta f_i}{\Delta x_i}(x_a + x_{i+1})\right]
\end{aligned}$$

this is the equation of the *Chebyshev line* that approximated $f(x)$ in the interval (x_j, x_{j+1}) with $\|f(x) - P_1(x)\|_\infty = \varepsilon^*$ as it was expressed in Theorem 3.

Corollary 2. *Since the PWL function is continuous in the intervals* (x_i, x_{i+1}) *and the term* $\sum_{i=j+1}^{\sigma} \Pi_i$ *is null in the expansion of* $\overline{f}(x_a)$ *performed in the previous proof, it can be affirmed that* $\lim_{\Delta x \to 0} f(x_i + \Delta x) = \lim_{\Delta x \to 0} f(x_i - \Delta x)$, *and* $\overline{f}(x)$ *is a PWL continuous function.*

Remark 2. Theorem 4 gives us the possibility of approximating any contiuous function $f(x)$ with $f'' > 0$ by means of a piecewise-linear function with an arbitrarily small infinite norm ε^*. Besides, the intervals of the approximation function can be obtained in a forward way if we know the analytical expression of $f(x)$, by means of solving the uncoupled set of discrete equations stated in (11). This fact supplies a direct method to design the intervals of approximations in comparison with the annealing iterative method needed in the minimization of the cuadratic norm.

3.3 CNNUM-based Approximation

In order to implement the previous theoretical results, we are going to modificate the original PWL original saturation as defined in (2) to adopt it to every affine plane Π_i as stated in Theorem 4 and superpose these modified saturation functions to obtain $\overline{f}(x)$ as defined in (3).

The saturation function $y(x)$ can be modified by the affine transformation $\frac{1}{b_i} y(a_i x - c_i) + m_i$. This reshaping translates the corners locates at (-1,-1) and (1,1) in the original saturation (2) to $(c_i - \frac{1}{a_i},\, m_i - \frac{1}{b_i})$, $(c_i + \frac{1}{a_i},\, m_i + \frac{1}{b_i})$ in the modified one. This transformation is performed on the CNN by means of the following two templates run in a sequential way

$$T_{k,1}^{PWL} = \{A_{ij}^2 = 0; B_{ij}^1 = a_k \delta_{2j} \delta_{2i}); I^k = -a_k c_k \ , \forall i, j\} \qquad (12a)$$
$$T_{k,2}^{PWL} = \{A_{ij}^2 = 0; B_{ij}^1 = b_k^{-1} \delta_{2j} \delta_{2i}); I^k = m_k \ , \forall i, j\} \qquad (12b)$$

where δ_{ij} denotes Kronecker delta.

The optimal parameter vector $\theta^* = [a_i^*, b_i^*, c_i^*, m_i^*]$ that makes the objective functional $\mathfrak{J} = \varepsilon^*$ can be obtained by means of calculating the intervals obtained by recursively solving the uncoupled set of discrete equations (11), and applying (10). Thus, we figure the templates given by (12) taking into account the values of

θ^* in order to obtain each summing term Π_i needed to define the approximating function $\overline{f}(x)$.

The most efficient procedure to add each term Π_i in the framework of CNNUM computing is introducing the input image into the CNN and operating the image by means of the nonlinear dynamics resulting from the connectivity defined in $T_{1,1}^{PWL}$ and $T_{2,1}^{PWL}$. After this, we save the result of the analog computation into the LAM. The same operation must be accomplished with templates $T_{1,2}^{PWL}$ and $T_{2,2}^{PWL}$ on the original input image. The result of the second step is accumulated in the LAM with the previous result. Making this process through every stage we finally obtain the image processed by a point operator that performs desired approximationg function $\overline{f}(x)$.

Lastly, it will be used a value $\varepsilon^* = 2^{-7}$ in the analytical deduction of the parameter vector θ^* because of the physical implementation of the CNN-UM chip allows an analog accuracy of this magnitude. In the case of $f(x) = \ln(x)$, the discrete equation in Theorem 4 yields the following implicit discrete equation

$$\ln\left(\frac{\Delta x_i}{\Delta \ln_i}\right) + \left(\frac{\Delta \ln_i}{\Delta x_i}\right) x_i - \ln(x_i) - 1 = 2\varepsilon^* \tag{13}$$

and in the approximation of an exponential function it can be similarly deduced the following condition

$$\frac{\Delta \exp_i}{\Delta x_i}\left[\ln\left(\frac{\Delta \exp_i}{\Delta x_i}\right) + x_i + 1\right] - \exp(x_i) \tag{14}$$

where $\Delta \ln_i = \ln(x_{i+1}) - \ln(x_i)$, $\Delta \exp_i = \exp(x_{i+1}) \exp(x_i)$ and $\varepsilon^* = 2^{-7}$. Both equation can be easily solved by standard numerical methods in order to obtain the neighboring points of the intervals that construct the PWL approximating function $\overline{f}(x)$ in a recursive and forward way.

4 Conclusions

In this paper, it has been introduced the equations that govern the complex CNN spatio-temporal dynamics and the CNNUM computational infrastructure implemented on silicon. After this, we have presented a general technique that allows us to approximate any nonlinear output function on the CNNUM VLSI Chip. For this purpose, we have given a theoretical analysis of an approximation technique based on the infinity norm criterion. Also, it has been comment the advantages of this technique in comparison with the cuadratic error criterion. The main motivation of this work is to release CNNUM emphanalogic (analog+logic) architecture from using emphdigital computers when CNN image processing computing capabilities are unable to perform any required nonlinear filtering step.

References

1. L. O. Chua, L. Yang. "Cellular neural networks: Theory". *IEEE Trans. Circuits Syst.*,vol. 35, no. 10. pp. 1257-1272, 1988.

2. L.O. Chua, and R.L.P.Ying, "Canonical piecewise-linear analysis," *IEEE Trans. Circuits Syst.*, vol. CAS-30, pp. 125-140, 1983.
3. L.O. Chua, and A. Deng, "Canonical piecewise-linear modeling," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 511-525, 1986.
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.,* Vol. 42, no. 10, 1995.
5. S. Espejo, R. Domínguez-Castro, G. Liñan, and A. Rodríguez-Vázquez, "64 x 64 CNN universal chip with analog and digital I/O", *IEEE Int. Conf. on Electronic Circuits and Systems*, 1998.
6. G. Liñan, R. Dominguez-Castro, S. Espejo and A. Rodriguez-Vazquez, "CNNUC3 user guide," *Instituto de Microelectronica de Sevilla Technical Report*, 1999.
7. T. Roska and L.O. Chua, "The CNN universal machine: An analogic array computer," *IEEE Trans. Circuits Syst.-II*, vol. 40, pp. 163-173, 1993.