LINEAR CELLULAR NEURAL NETWORKS.

Federico Lobato-López
Mexico Center for Semiconductor Technology (MCST)
Motorola Semiconductor Products Sector
Av. Esteban de Antonio No. 2702
Puebla, Pue. 72180, México

José Silva-Martínez and Edgar Sánchez-Sinencio
Texas A&M University
Department of Electrical Engineering
Analog and Mixed Signal Center
College Station, Texas 77843-3128

ABSTRACT

In this paper we introduce a new way of image processing based on CNN but whose activation function is a linear function. This fact enable us to develop gray-scale image processing; here we introduce the mathematical basis and the electrical model. In addition, some examples are presented and the properties of these networks are discussed.

1. INTRODUCTION.

The linear CNN has its basis on bayesian estimation and regularization theory. In areas such as signal processing, computational vision, patterns analysis, etc. exist problems called ill-posed problems. An ill-posed problem is defined as [1]:

Definition 1 Find a smooth function f, defined at the nodes of a regular lattice L, where the observations g can be modeled by:

\[ g(x) = Af(x) + n(x), \quad x \in S, \quad S \subseteq L \]  \hspace{1cm} (1)

where A is a non invertible operator and n(x) a random field (for example a Gaussian field).

The regularization solution for this kind of problems is derived from the Bayesian estimation theory [1, 2]:

\[ P_{g|f} = P_n(Af - g) = \frac{1}{K} \exp \left\{ - \sum_{x \in S} \frac{((Af)(x) - g(x))^2}{2\sigma^2} \right\} \]  \hspace{1cm} (2)

where K is a constant.

From (1), the estimation of f can be carried out by a Markov random field [2], which is defined by a set of neighborhoods on L and a set of potentials functions \( V_c \), where each function depends on the variables associated to each node belonging to a "clique" C on the neighborhood. In this way f is given by a Gibbs distribution:

\[ P_f = \frac{1}{Z} \exp \left\{ - \sum_{C} V_c(f) \right\} \]  \hspace{1cm} (3)

where Z is a normalization constant and the sum is over all "cliques" on L. Taken (2) and (3) and the Bayes rule, we obtain the posteriori distribution for f:

\[ P_{f|g} = \frac{1}{Z} \exp \left\{ - \frac{1}{2\sigma^2} \sum_{x \in S} ((Af)(x) - g(x))^2 \right\} \]  \hspace{1cm} (4)

From this the maximum a posteriori estimate (MAP) is defined as the minimizer of the functional:

\[ U(f) = \sum_{x \in S} \Phi_{f,g}(x) + \lambda \sum_{C} V_c(f) \]  \hspace{1cm} (5)

where \( \Phi_{f,g}(x) \) is a function depending of noise or the variance of noise given by the parameter \( \lambda \).

U(f) is an energy function, which represents the total energy of the system; so, minimization of (5) to obtain the MAP is interpreted as the system energy minimization or to find its steady state. The Minimization of (5) means to obtain the regularized solution to problems given by (1). Let's consider the regularization general formulation [3, 4]:

\[ E(f) = \frac{1}{2} \int (f - g)^2 \, dx + \sum_{i=1}^{\infty} \lambda_i \left( \frac{\delta f}{\delta x_i} \right)^2 ; \quad \lambda_i > 0 \]  \hspace{1cm} (6)

for the first order regularization, we have:

\[ E(f) = \frac{1}{2} \int (f - g)^2 \, dx + \lambda_1 \left( \frac{\delta f}{\delta x} \right)^2 \]  \hspace{1cm} (7)

Comparing (5) and (7) we can see that (5) is the regularization solution to problems given by (1). With all this in mind we present in the next section a derivation of an electrical implementation consider as a physical solution to (1); in section 2 the main characteristics of eqn. (5) are denoted and from all that; in section 3, 1-D and 2-D CNN approximation are obtained as well as its main characteristics are denoted; in section 4 results of the CNN approximation in 2-D applied to image processing are shown. Finally conclusions about this work are presented in section 5. All this work is considers as a first step toward a VLSI implementation of these systems.

2. LOW-PASS FILTERS.

From eqn. (5) and considering f as a smooth global field, then potential functions are taken like quadratic finite differences; approximations to partials derivatives (eqn. (7)). In this paper we take

0-7803-6685-9/01/$10.00 ©2001 IEEE
the first order Markov random field (membrane model), whose "cliques" are taken by pairs of neighbors \(\{a, b\}\) and the potential functions have the form:

\[
V_{ab} = (f(a) - f(b))^2
\]  
(8)

The minimization of (5) smooth the field \(g\), interpolating over those sites where the information is missing; in this way, the regularization can be considered as a low-pass filtering operation on \(g\).

If the lattice is considered unidimensional and boundless, and the operator \(A\) is the unitary one, and the noise is gaussian and the model is given by (8) then the gradient of (5) is equal to zero, then we obtain a set of linear equations:

\[
f(x) + \lambda (-f(x) + 2f(x) - f(x + 1)) = g(x)
\]  
(9)

Taking the Fourier transform on both sides results in:

\[
\mathcal{F}(u)(1 + \lambda(-e^{-i\omega} + 2 - e^{i\omega})) = \mathcal{G}(u)
\]  
(10)

solving for \(\mathcal{F}(u)\) we can see that regularization operation is like a low-pass filter whose transfer function is:

\[
\mathcal{H}(u) = \frac{1}{1 + 2\lambda(1 - \cos(\omega))}
\]  
(11)

and the value of \(\lambda\) is the parameter that defines the bandwidth of this kind of filters.

3. CNN APPROXIMATION.

3.1. 1-D Low-Pass Regularized Filter.

For the implementation of this filter, let us consider equation (9) which can be rewritten as follows:

\[
\frac{1}{\lambda}f(x) = \frac{1}{\lambda}g(x) + \left\{ [f(x) - f(x - 1)] + [f(x) - f(x + 1)] \right\}
\]  
(12)

This expression can be considered as the Kirchhoff Current Law (KCL) applied to node \((x)\), where \((x - 1), (x + 1)\) represent neighbor information variables. The terms \(\frac{1}{\lambda}f(x)\) and \(\frac{1}{\lambda}g(x)\) corresponds to currents at node \((x)\). Under these conditions the eqn. (12) can be represented by the circuit shown in the figure 1. Solving this circuit, eqn. (12) results, with R=1.

\(f(x)\) can also be generated by using voltage controlled current sources, as shown in figure 2. By applying superposition, it can be easily shown that the grounded resistance is given by \(\lambda||R/2\). In addition, the VCCS represents the effects of \(f(x - 1)\) and \(f(x + 1)\). These currents are given by:

\[
I_{x-1} = \frac{1}{R}f(x - 1)
\]

\[
I_{x+1} = \frac{1}{R}f(x + 1)
\]  
(13)

In real circuits, several parasitic capacitors are present; those are accounted in figure 3 by the capacitor \(C\).

The dynamics of the network can be expressed by the following equation:

\[
C\frac{df(x)}{dt} = -\frac{f(x)}{\lambda} + \frac{1}{R}\left[ f(x - 1) - 2f(x) + f(x + 1) \right] + \frac{g(x)}{\lambda}
\]  
(14)

It can be note that in steady state \(\left(\frac{df(x)}{dt} = 0\right)\) the solution of eqn. (14) is the same as eqn. (12). If the capacitor \(C\) in figure 3 is replaced by an open circuit, then eqn. (14) is reduced to eqn. (12), and both equations represent the steady state of the network.

3.2. 2-D Low-Pass Regularized Filter.

From eqn. (5) we can generalize the theory to the 2-D case, as follows:

\[
U(f) = \sum_{r} \left[ f(r) - g(r) \right]^2 + \lambda \sum_{<r,s>} \left[ f(r) - g(s) \right]^2
\]  
(15)

the notation \(<r,s>\) indicates differences of pairs taken one by one; this is true for all positions including the border cells. In this way, the neighborhood has other structure and depending of the position on the border the structure is different. For the centered cell, the Markov random field is shown in figure 4(a), and for a

III-438
corner cell the structure is shown in figure 4(b) and for an edge cell the structure is shown in figure 4(c). Those variations in the structure represent different boundary conditions.

To minimize the 2-D energy function given by (15) we proceed as in the 1-D case, but with \( r = (x, y) \); hence the function is obtained as follow:

\[
\frac{1}{\lambda} f(x, y) = \frac{1}{\lambda} g(x, y) + \left\{ f(x - 1, y) - f(x, y) \right\} + \\
+ \left\{ f(x + 1, y) - f(x, y) \right\} + \\
+ \left\{ f(x, y - 1) - f(x, y) \right\} + \\
+ \left\{ f(x, y + 1) - f(x, y) \right\}
\] (16)

Figure 5: Equivalent network for the 2-D low-pass regularized filter.

As we mentioned above the neighborhoods are not the same for all positions in the field, this condition is reflected in the templates of the CNN. In this case we have three different templates as shown in (18); \( A(i, j, k, l) \) and \( B(i, j, k, l) \) are matrices with the following values depending on the position at the cell:

\[
A = \begin{bmatrix}
0 & 1 & 0 \\
1 & -4 & 1 \\
0 & 1 & 0
\end{bmatrix} ;
B = \begin{bmatrix}
0 & 0 & 0 \\
0 & \frac{1}{\lambda} & 0 \\
0 & 0 & 0
\end{bmatrix}
\] (18)

\[
Ae = \begin{bmatrix}
0 & 1 & 0 \\
1 & -2 & 0 \\
0 & 0 & 0
\end{bmatrix} ;
Bc = \begin{bmatrix}
0 & 0 & 0 \\
0 & \frac{1}{\lambda} & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

\[
Ac = \begin{bmatrix}
0 & 1 & 0 \\
1 & -3 & 1 \\
0 & 0 & 0
\end{bmatrix} ;
Be = \begin{bmatrix}
0 & 0 & 0 \\
0 & \frac{1}{\lambda} & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

Templates \( Ac \) and \( Ae \) corresponds to cases 4b and 4e in figure 4, respectively; and these structures are reflected by the value given to the centered elements of the templates \( Ac \) and \( Ae \).

As we can see the minimization of (15) can be implemented by a CNN with \( l = 0 \) and templates given in (18). In this implementation the activation function is a linear function which permits us to implement gray-scale image processing as we will see in the next section.

4. GRAY-SCALE IMAGE PROCESSING

In order to demonstrated the capabilities of the linear–CNN, we have macromodeling the circuit shown in figure 6 and processing the pattern image shown in figure 7(b). The processing has been
Figure 8: Resulting images of processing figure 7(b) by the linear CNN.

done for different values of $\lambda$; in figure 8 are shown different images corresponding to several $\lambda$ values. For figure 8(a), $\lambda = 0.2$; figure 8(b), $\lambda = 0.5$; figure 8(c), $\lambda = 1.0$ and figure 8(d), $\lambda = 2.0$.

As can be seen from figure 8 the $\lambda$ parameter plays the role of the bandwidth factor for this kind of filters; for a unique image the values of $\lambda$ should be varied in one decade but for a processing system where different images with different quantities of noise are presented, $\lambda$ may take values as large as nearly two decades. The gray-scale image processing is obtained thanks to the negative feedback ($R_{gnd}$) which allow stable states stay in the linear range. The desired bandwidth characteristics are difficult to reach with very simple circuits in a VLSI implementation, but if an equivalent resistor obtained from the parallel connection of the resistors $\lambda$ with $R_{gnd}$ is considered then the bandwidth requirements can be relaxed and an efficient VLSI implementation may be obtained.

5. CONCLUSIONS.

An architecture for gray-scale image processing based on probabilistic and regularization theory and on CNN architecture has been presented. On the other hand, results for a Markov random field of first order (real-valued) has been shown; however, if the Markov random field is modified (complex-valued) then we can obtain macromodels of higher order with other properties and applications. For example, linear and nonlinear robust quadrature filters for determination of binocular disparity, texture segmentation and fring pattern analysis. By the free-boundary condition all the images can be processed without dummy cells; besides, irregular regions in an image can be processed in an easy way. VLSI implementations can be obtained without additional difficulty that those presented in classic CNN's implementations.

6. REFERENCES


