Multprocessor-Based Video Motion Detection Using Adaptive Neural Systems

Ji-Chien Lee, Bing J. Sheu, Chia-Fen Chang, Rama Chellappa

Department of Electrical Engineering
Signal and Image Processing Institute and Center for Neural Engineering
University of Southern California, Los Angeles, CA 90089-0271, U.S.A.

Abstract

Fast motion detection is an important step in the high-speed video/vision processing systems. System-level design of a 2-dimensional mesh-connected competitive neural network for video motion detection is presented. The motion information from a sequence of images can be obtained through mixed-analog/digital signal processing. Massively parallel neurocomputing is performed by multiple copies of compact and efficient neuroprocessors. Interprocessor data transfer is carried out by dedicated point-to-point analog interconnections. Global data communication between the host computer and neuroprocessors is through the digital common bus to maintain strong signal strength.

An extendable analog winner-take-all circuit is used to implement the competition function with a minimal delay time. A 1.5 × 2.8-cm² chip in a 1.2-μm CMOS technology can accommodate 64 velocity-selective neuroprocessors. This chip can achieve 83.2 Giga-connections per second. By using 128 VLSI neural chips, the speed-up factor over the Sun-4/60 SPARC station-I is estimated to be 54,545.

I. Introduction

Motion estimation extracted from inter-frame images provides useful information in tracking moving objects. Because optical flow represents the apparent motion of the brightness patterns, it corresponds to the motion field [1]. Conventional approaches of optical flow computation are sensitive to either rotation or noise as in the intensity-based method, or suffer from sparsity as in the token-based method. Zhou et al [2] described the algorithm of a locally connected neural network for computing optical flow based on rotation-invariant primitives extracted from multiple successive image frames. This deterministic neural network can detect large displacements. In this paper, the VLSI architecture and circuit implementation for optical flow computing are presented. Mixed-signal design techniques are utilized for neuroprocessor design to fully exploit the massively parallel computational power of neural networks. A neural chip containing 64 neuroprocessors has been developed. By using a two-dimensional mesh configuration, a high-speed video motion detection can be achieved.

II. 2-D Competitive Network Architecture

A VLSI architecture has been developed. The locally connected competitive neural network used for optical flow computing is shown in Fig. 1. Let the velocity field consist of two components \( k \) and \( l \). A set of \((2D_k + 1)(2D_l + 1)\) modules of neurons are used to represent the optical flow field, where \( D_k \) and \( D_l \) are the maximum values of velocity components in \( k \) and \( l \) directions, respectively. In digital image processing, the velocity component range is sampled using bins of size \( B \). Each module corresponds to a velocity value and contains \( N_k \times N_l \) neurons if the images are of size \( N_k \times N_l \). All neurons in the same module are self-connected and locally interconnected with other neurons in a neighborhood of size \( r \times r \). Every pixel is represented by \((2D_k + 1)(2D_l + 1)\) mutually exclusive neurons which form a hypercolumn for velocity selection. When the neuron at the point \((i, j)\) in the \((k, l)\)-th module is 1, the actual velocities in \( k \) and \( l \) directions at the point \((i, j)\) are \( kB \) and \( lB \), respectively.

Let \( V = \{v_{i,j,k,l} | 1 \leq i \leq N_k, 1 \leq j \leq N_l, -D_k \leq k \leq D_k, -D_l \leq l \leq D_l \} \) be a binary set of the neural network with \( v_{i,j,k,l} \) denoting the state of the \((i,j,k,l)\)-th neuron which is located at point \((i, j)\) in the \((k, l)\)-th module. \( T_{i,j,k,l,m,n,k,l} \) be the synaptic interconnection strength from...
neuron \((i,j,k,l)\) to neuron \((m,n,k,l)\) and \(I_{i,j,k,l}\) be the bias input. The interconnection weights \(T_{i,j,k,l;m,n,k,l}\) and bias input \(I_{i,j,k,l}\), can be obtained by comparing the coefficients of the error function for computing the optical flow and the energy function of the neural network. Each neuron receives a bias input \(I_{i,j,k,l}\) and binary inputs \(v_{m,n,k,l}\) from itself and neighboring neurons. The optical flow computing is performed by neuron evaluation based on a parallel updating scheme and the minimal mapping theory \([3]\). When the network reaches a stable state, the optical flow field is available at the neuron outputs.

III. Neuroprocessor Design

As shown in Fig. 2, each small frame represents one velocity-selective hypercolumn which contains \((2D_k+1)(2D_l+1)\) velocity-sensitive components. Each hypercolumn is locally interconnected to its \(\Gamma \times \Gamma - 1\) neighboring hypercolumns. The hypercolumn is designed as a neuroprocessor within which the velocity selectivity of an image pixel can be conducted. Mixed analog-digital design technologies are used to achieve compact and programmable synapses and neurons for massively parallel neural computation \([4, 5]\).

To simplify the two-dimensional interconnection problem for optical flow computing, both point-to-point analog connection and digital common bus are used. Since each pixel is affected by its near neighbors, the neuroprocessor has to receive information from the neighboring neuroprocessors during the network operation. Thus the data communication scheme between these locally interconnected neuroprocessors becomes a key factor to determine the overall system performance. By using a bit-parallel point-to-point interconnection scheme combined with the digital-to-analog and analog-to-digital conversion techniques, the required number of ports and the silicon area for interconnection are dramatically reduced. The \(v_{i,j,k,l}\) data is converted to an analog value before transmission and is converted back to a digital format upon reception. Data communication between neuroprocessors and the host controller is carried out in digital common bus to preserve signal strength and to achieve network scalability.

A functional diagram of the velocity-selective neuroprocessor is shown in Fig. 3. It includes a velocity-sensitive component array, and a neighbor interconnection block. The array has \((2D_k+1)(2D_l+1)\) velocity-sensitive components which are laterally connected through the winner-take-all circuit. The velocity of the neuroprocessor is determined by competition which is implemented with the winner-take-all circuit. Only one velocity component which has the maximum excitation will become the winner to represent the velocity of that pixel.

As shown in Fig. 4, each velocity-sensitive component is constructed with one synapse array, one summing neuron, and one winner-take-all cell. The synapse array contains \(\Gamma \times \Gamma + 1\) programmable synapses. A transconductance amplifier consisting of transistors \(M_1,M_5\) produces synapse output current \(I_{i,j}\) according to mask voltage \(V_{\text{mask}}\) and weight voltage \(V'_{i,j}\). The bias voltage \(V_{\text{bias}}\) controls the bias current in the transconductance amplifier. When the \(V_{\text{mask}}\) is at logic-1, the \(V_{\text{bias}}\) is connected to \(V_{\text{on}}\) to provide the amplifier with a specific bias current \(I_{\text{bias}}\) which is used to determine the dynamic-range of the synapse cell. When the \(V_{\text{mask}}\) is at logic-0, the \(V_{\text{bias}}\) is connected to the negative power supply to produce no synapse output current. Therefore, \(V_{\text{mask}}\) performs a masking operation on the synapse weight voltage \(V'_{i,j}\). The mask voltage of each synapse cell is directly related to the value of the \(v_{m,n,k,l}\) which represents the velocity information of the neighboring pixels. The summing neuron is a two-stage amplifier with externally adjustable voltage gain. Each summing neuron and its associated programmable synapse array perform a parallel inner-product computation. The outputs of the winner-take-all circuit are binary values. Only one winner cell has the output voltage of \(V_{\text{on}}\). The other outputs are inhibited to \(V_{\text{on}}\). The binary output of winner-take-all cell represents the velocity status.

The synapse weights and bias inputs are calculated by the host computer and stored in a digital static-RAM. The 8-bit D/A converter converts the digital representation of the synapse weights into analog values to charge the capacitors of the synapse matrix. A two-port static-RAM and differential amplifier-based synapse design allow network retrieving and learning processes to occur concurrently.

IV. Experimental Results

In our study, \(D_k = D_l = 2\) and a size of \(5 \times 5\) smoothing window are used for the demonstration purposes. The layout of one velocity-selective neuroprocessor is shown in Fig. 5. It occupies \(2,482 \times 5,636 \lambda^2\) and contains 25 neurons, \(25 \times 27\) synapse cells and is able to detect the moving object with 25 different velocities. With
an advanced 1.2-μm CMOS technology, 64 neuroprocessors can be accommodated into one VLSI neural chip of 1.5 × 2.8 cm² in size. The system-level block diagram for high-performance motion detection using multiple VLSI neural chips is shown in Fig. 6.

Processing speed for one neural network iteration is around 522 nsec. Each iteration cycle includes synapse multiplication, neuron thresholding, winner-take-all competition, data storage on latches, digital-to-analog, local data communication, and analog-to-digital functions. SPICE simulation results are listed in Table 1. The major delay comes from the synapse multiplication due to the significant capacitance loading on the current-summation line. For the D/A converter design, 5 pF and 50 pF capacitance loadings are estimated for inter-chip data communication and off-chip data communication, respectively. A total of 8.32 × 10¹⁰ connections per second can be achieved by using one VLSI neural chip containing 1600 neurons, 41,600 synapses cells, and operated at 2 MHz for the iteration cycle. Based on the results in Table 1, the speed comparison of the system using 128 VLSI neural chips with a Sun-4/60 SPARC station-1 is listed in Table 2. The speedup factor is estimated to be 54,545. System-level simulation has been conducted to further understand the performance of the motion detection chip. A set of successive image frames directly produced by a Sony XC-77 CCD camera was used. Figure 7(a) shows four successive image frames with a sedan moving from left to right against a stationary background. The size of each image frame is 100 × 120 pixels. The maximum displacement of the sedan between the time-varying image frame is about 7 pixels. By setting A = 4, B = 250, C = 0, Dₖ = 7, and D₁ = 1, the velocity field, as shown in Fig. 7(b), was obtained after 36 iterations.

VI. Conclusion

A compact and efficient VLSI neuroprocessor which includes 25 neurons and 25 × 26 synapse cells is able to detect the moving object with 25 different velocities. Multiple neuroprocessors can be connected as a two-dimensional mesh to fully exploit the massively parallel computational power of neural network. This electronic neuroprocessing design can be applied for various real-time signal and image processing tasks such as HDTV, or mobile robots.

Acknowledgement

The authors would like to thank Professor Irving Reed, Dr. Yi-Tong Zhou, Wai-Chi Fang, and Joongho Choi for valuable discussions.

References


Table 1. Circuit Response Time

<table>
<thead>
<tr>
<th>Circuit Component</th>
<th>Measured Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analog-to-Digital Conversion</td>
<td>73 ns</td>
</tr>
<tr>
<td>Synapse Multiplication</td>
<td>120 ns</td>
</tr>
<tr>
<td>Neuron Thresholding</td>
<td>20 ns</td>
</tr>
<tr>
<td>Winner-Take-All Operation</td>
<td>38 ns</td>
</tr>
<tr>
<td>Encoder &amp; Data Latch</td>
<td>8 ns</td>
</tr>
<tr>
<td>Digital-to-Analog Conversion</td>
<td>84 ns + 263 ns</td>
</tr>
<tr>
<td>Total</td>
<td>343 ns + 522 ns</td>
</tr>
</tbody>
</table>

Note: Tₜᵢ₀ = 202 A in MOSIS 1.2-μm CMOS technology. + with an output loading of 5 pF * with an output loading of 50 pF.

Table 2. Performance of VLSI Motion Estimation System

| Synapse Weight Loading Time (into SRAM) | 2,080 us |
| Network Execution Time (for 36 iterations) | 18.792 us |
| Neuron State Read Out Time | 409.6 us |
| Total Processing Time | 2,508 ms |
| Speed-Up Factor | 54,545 |

Note: + The comparison reference is Sun-4/60 SPARC station 1.
Fig. 1. A competitive neural network for optical flow computing.

Fig. 2. A 2-D array of velocity-selective hypercolumns. Each hypercolumn is locally interconnected with the neighboring hypercolumns.

Fig. 3. Functional diagram of each velocity-selective hypercolumn.

Fig. 6. System-level diagram of a multiprocessor system for optical flow computing.
Fig. 4. Circuit schematic of the velocity-sensitive component.

Fig. 5. The layout of one velocity-selective hypercolumn for one pixel. The synapse cell is shown in the insert.

Fig. 7. The simulation result on the real images.
(a) A sequence of images with a sedan moves from left to right.
(b) The estimated optical flow field.