Parallel Architecture for Binary Images Recognition

Saul Martinez-Diaz

***Abstract*— In this paper a new implementation of nonlinear composite filters for image recognition in parallel hardware is proposed. The architecture is designed for a Field Programmable Gate Array (FPGA) device. The filter design is based on logical operations and the correlation is computed with a nonlinear operation called morphological correlation. The proposed architecture reduces the time required for the nonlinear operations in the spatial domain. Simulation results are provided and discussed.**

***Index Terms*—Parallel processing, FPGA, nonlinear filters, pattern recognition, morphological correlation**

I. Introduction

Pattern recognition, especially image recognition, has been an area of intensive research over the last decades. The main reason for this is the great amount of applications for recognition systems. Since VanderLugt introduced the Matched Spatial Filter (MSF) in 1964 [1], correlation methods have been used extensively the last years for image recognition [2]-[10]. Firstly, correlation methods exploit all information from images in the recognition process. Besides correlation is shift-invariant and has solid mathematical foundation. In this case the basic recognition procedure is:

- Design a template (filter) with one or several training images.
- Correlate the filter with an input test image.
- Establish a threshold at the correlation output.

A correlation value greater than threshold indicates that target is located at coordinates of the correlation peak.

Correlation filters can be designed by optimizing one performance criteria with linear techniques. In addition, information from several distorted training objects could be incorporated. For example, synthetic discriminant functions (SDF) [2], [3] and minimum average of correlation energy (MACE) [4] filters can be used for distortion-invariant multiclass pattern recognition. In addition, an adaptive approach has been proposed [5], in order to reject other objects from scenes. However, linear filters are sensitive to most kind of real noise. On the other hand, several approaches of nonlinear filter design have been proposed too [6]-[10]. Recently, nonlinear composite filters for distortion-invariant pattern recognition were introduced [9], [10]. The filters are designed as a logical combination of binary objects. Correlation is computed among the filter and a test scene with a nonlinear operation called Morphological Correlation (MC) [11]. These kinds of filters have demonstrated a good discrimination capability and noise tolerance. With the help of threshold decomposition [12], this technique can be applied to grayscale images as well. A drawback of this process is the high computational cost for large images. Nevertheless, the calculation of the nonlinear correlation can be parallelized by using specialized hardware.

In this paper we propose the implementation of the nonlinear filtering process in parallel hardware. The aim is to reduce the processing time. The architecture is intended for a Field Programmable Gate Array (FPGA). FPGA´s have been used in several applications of parallel processing executing the more time consuming tasks [13]-[16]. Simulation results of proposed system are provided and discussed. The paper is organized as follows: Section II describes composite nonlinear filters. Section III presents an FPGA-based architecture. In section IV computer simulations are provided and discussed. Section V summarizes our conclusions.

Manuscript received February 21, 2012; revised March 8, 2012. This work was supported by the Instituto Tecnológico de La Paz (ITLP), grant ITPAZ-MSC-2011-002.

Saul Martinez-Diaz is with the Instituto Tecnológico de La Paz, La Paz BCS, México (e-mail: saulmd@itlp.edu.mx).

II. NONLINEAR FILTERING

A. Basics

The proposed technique is a locally adaptive processing of the signal in a moving window. The moving window is a spatial neighborhood containing pixels surrounding geometrically the central window pixel. The neighborhood is referred to as the *W*-neighborhood. The shape of the *W*-neighborhood is similar to the region of support of the target. The size of the neighborhood is referred to as $|W|$, and it is approximately taken as the size of the target. In the case of non-stationary noise or cluttered background (space-varying data), it is assumed that the *W*-neighborhood is sufficiently small and the signal and noise can be considered stationary over the window area.

B. Threshold Decomposition

Suppose a gray-scale image *I(m,n)* with *Q* levels of quantization, where (*m*,*n*) are the pixel coordinates. According to the threshold decomposition concept [12], the image *I* can be represented as a sum of binary slices

$$I(m,n) = \sum_{q=1}^{Q-1} I^q(m,n) \quad (1)$$

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

here $\{I^q(m,n), q=1,...Q-1\}$ are binary images obtained by decomposition of the greyscale image with a threshold *q*, as follows

$$I^q(m,n)=\begin{cases}1, & if \quad I(m,n)\geq q\\ 0, & otherwise\end{cases} \tag{2}$$

C. Filter Design

Now, assume that there are *M* reference objects to be recognized (true class) and N objects to be rejected (false class). We construct a filter as a logical combination of the training images:

$$H(m,n)=\sum_{q=1}^{Q-1}\left[\left(\bigcap_{i=1}^{N}T_i^q(m,n)\right)\bigcap\left(\bigcup_{j=1}^{M}\bar{F}_i^q(m,n)\right)\right], \tag{3}$$

Where $\{T_i^q(m,n), q=1...Q-1, i=1...N\}$ are the binary slices obtained by threshold decomposition of the true class images. $\{\bar{F}_j^q(m,n), q=1...Q-1, j=1...M\}$ are the logical complement of binary images obtained by threshold decomposition of false class images training images. $\bigcap$ represents the logical intersection: the result at coordinates (*m,n*) is 1 if the corresponding pixels of both planes are equal to 1; otherwise, the result is 0. $\bigcup$ represents the logical union: the result at coordinates (*m,n*) is 0 if the corresponding pixels of both planes are equal to 0; otherwise, the result is 1. The neighborhood *W* is taken as the region of support of the composite filter.

D. Morphological Correlation

Let $\{H(m,n)\}$ and $\{S(k,l)\}$ be a template and a test scene respectively, both with *Q* levels of quantization. The local nonlinear correlation (morphological correlation) between a normalized input scene and a shifted version of the target at coordinates (*k,l*) can be defined as

$$c(k,l)=\sum_{m,n\in W}\min\left(a(k,l)S(m+k,n+l)+b(k,l),H(m,n)\right) \tag{4}$$

Where *c(k,l)* is the local nonlinear correlation at the coordinates *(k,l)*. min*(x,y)* is the minimal value among *x* and *y*. The sum is taken over the *W*-neighborhood. *a(k,l)* and *b(k,l)* are local normalizing coefficients, which take into account unknown illumination and bias of the target, respectively. The coefficients estimates are given by:

$$a(k,l)=\frac{\sum_{m,n\in W}T(m,n)\cdot S(m+k,n+l)-|W|\cdot\bar{T}\cdot\bar{S}(k,l)}{\sum_{m,n\in W}\left(S(m+k,n+l)\right)^2-|W|\cdot\left(\bar{S}(k,l)\right)^2} \tag{5}$$

$$b(k,l)=\bar{T}-a(k,l)\cdot\bar{S}(k,l) \tag{6}$$

It can be shown that the nonlinear correlation in (4) can be computed with the binary slices obtained from threshold decomposition of the input scene and the filter as

$$c(k,l)=\sum_{q=1}^{Q-1}\sum_{m,n\in W}\left(S^q(m+k,n+l)\cap H^q(m,n)\right) \tag{7}$$

where $\{S^q(k,l)\}$ and $\{H^q(m,n)\}$ are binary slices of the gray-scale images obtained by threshold decomposition of the normalized input scene $\{a(k.l)\cdot S(k,l)+b(k,l)\}$ and the template $\{H(m,n)\}$, respectively.

III. HARDWARE

As can be seen, the nonlinear correlation process is computationally expensive. Suppose a test scene of KxL pixels and a template of MxN pixels; then computation of correlation requires KxLxNxM operations for each binary slice. On the other hand, these operations can be computed in a parallel way.

A. Field programmable gate arrays

Morphological correlation involves a high computational cost for a sequential computer. However, a faster response can be achieved by using parallel processing. A suitable device for this purpose is a field programmable gate array (FPGA). Figure 1 shows this basic structure of an FPGA: the device consists of an array of programmable basic cells (shown in white), an interconnect matrix surrounding the basic cells (shown in gray) and a set of input/output programmable cells (shown in black) [12]. Modern devices could contain millions of logic cells for a low cost.

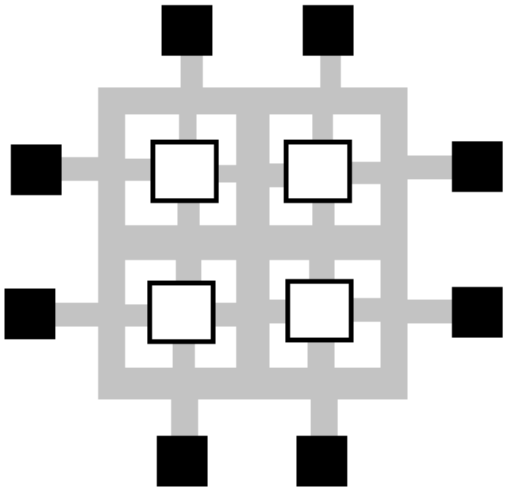

Fig. 1. Basic structure of a field programmable gate array (FPGA).

All kind of cells and the interconnections can be reprogrammed. Because of its capacity of reconfiguration, these kinds of device are a versatile choice for experimental designs. In addition, its high capacity and low cost make them suitable for implementation of specialized hardware. These kinds of devices have been used before in several applications [13]-[16].

B. Proposed architecture

Figure 2 is a block diagram of the proposed architecture. RAM1 memory stores the entire binary test scene, RAM2 stores the binary template and RAM3 stores the correlation output. The size of RAM1 and RAM3 are equal to the size

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

of the test scene. Dimensions of the template image are M rows and N columns. The size of RAM2 equals the size of the template. All data buses are M bits wide. The AND gates bank contains M units that execute the operation of intersection among an entire row of both, the local test scene and the template images. An adder unit calculates simultaneously the sum of correlation, which is entered to the accumulator unit. The final result is stored in the RAM3 memory. Besides, an address generator, which is clock synchronized, computes the memory allocation sequences to read data and store the results of the correlation process. Only a counter unit is needed to compute the address of RAM3 memory.

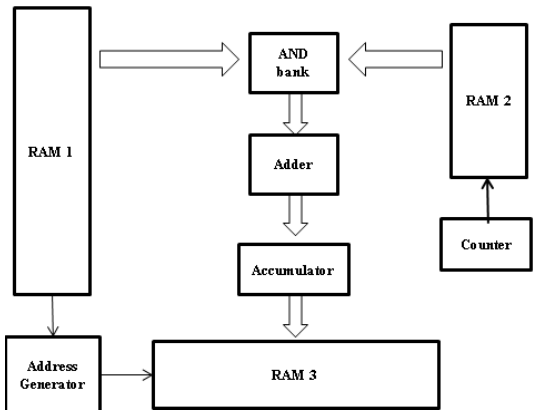

Fig. 2. Block diagram of the proposed architecture. All memory banks are clock synchronized.

C. Operation

The basic operation of proposed hardware is as follows:

The address of the first pixel of test scene is loaded and data is read from RAM1 memory. Since bus is M bits wide, an entire column is loaded in a single clock cycle from RAM1 and RAM2. In this way, the whole intersection among the local image and the template is processed in N clock cycles. The adder unit processes in parallel the M bits of each column. In consequence, not additional time is required for the sum. The partial result of each column sum is sequentially stored in the accumulator unit. Once processed a local image the address generator computes the address for storing the result in RAM3 memory and for reading the next local image from RAM1 memory. Besides, the counter unit is restarted to read the first column of template again. This procedure is repeated until all pixels of the test scene are processed.

Note that this procedure requires only O(KxLxN) operations. In addition, not additional time is required for executing arithmetic operations and loading data. The proposed architecture can be easily modified for larger images.

IV. COMPUTER SIMULATIONS

In this section computer simulations are provided. Figure 4 shows a test scene with two objects embedded. The template is 24x36 pixels and the test scene is 256x256 pixels. Both are grayscale images with 256 levels of quantization. The filter is designed including both butterflies. The target is the butterfly in the left side and the other butterfly is included to be rejected. First, threshold decomposition is executed by a computer. Next, each binary image is processed in the proposed architecture. Finally the correlation results are read from computer and normalized.

Figure 5 shows the final correlation output obtained. As can be seen, the correlation plane contains a sharp peak of magnitude equal to one at central coordinates of the target. On the other hand, at central coordinates of the false object the correlation value obtained is zero and no other sharp peak is distinguished. Note that the filter is able to detect the target embedded into the cluttered background and with a similar object in the same scene.

Fig. 3. Test scene including target (left butterfly) and an object to be rejected (right butterfly).

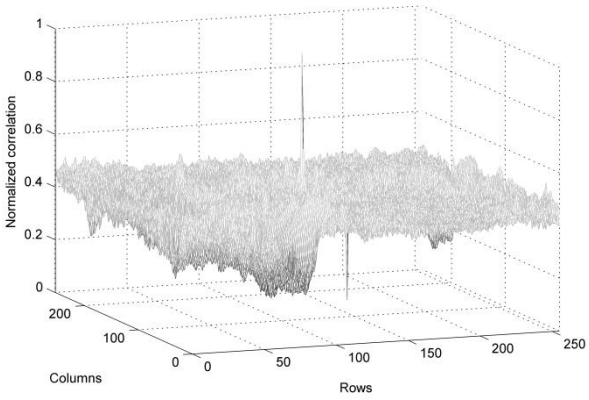

Fig. 4. Correlation plane obtained with proposed architecture for the test scene on figure 3.

With the proposed architecture the number of clock cycles required to process the entire image is only 3% of the cycles required with a sequential processor. In order to compare performance of proposed system with a sequential processor, several tests were executed in a personal computer. A 1.6 GHz sequential processor was utilized. Results of processing time in the personal computer were averaged. A 200 MHz FPGA was selected for the comparison. Because of the difference of speed among the processors, the interaction with the computer of FPGA and other factors, the time required to process an entire image with the proposed architecture was 50% of that required with the sequential processor. However, for largest images this ratio can be improved at the cost of more memory and gates on the FPGA device.

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

V. Conclusion

In this paper was presented a new architecture designed to perform nonlinear correlation for pattern recognition. The kind of filters employed has demonstrated robustness to non-Gaussian noise and good discrimination capability. Besides, the proposed architecture performs the most demanding time tasks of the correlation process. Moreover, the time consumption can be reduced by increasing the memory in such way that more correlations can be executed simultaneously. The proposed architecture is flexible and can be easily adapted to other sizes of images. Future work includes processing of the entire grayscale image into the FPGA device.

Acknowledgment

We want to thanks to the Programa de Mejoramiento del Profesorado (PROMEP) and the Instituto Tecnológico de La Paz (ITLP) for the facilities given to this project.

References

[1] A. B. VanderLugt, ”Signal detection by complex filtering,” in *IEEE Trans. Inf. Theory* vol. 10, 1964, pp. 139-145.

[2] C. F. Hester and D. Casasent, “Multivariant technique for multiclass pattern recognition,” *Appl. Opt.*, vol.19, pp. 1758-1761, Jun. 1980.

[3] D. C. W. Casasent, "Correlation synthetic discriminant functions," *Appl. Opt.*, vol. 25, pp. 2343-2350, Jul. 1986.

[4] A. Mahalanobis, B. V. K. Vijaya-Kumar and D. Casasent, “Minimum average correlation energy filters,” *Appl. Opt.*, vol. 31, pp. 1823-1833, Sep. 1987.

[5] J. A. González-Fraga, V. Kober, and J. Álvarez-Borrego, “Adaptive synthetic discriminant function filters for pattern recognition,” *Opt. Eng.*, vol. 45, pp. 057005-1-057005-10, May. 2006.

[6] Y. H. Doh, J. C. Kim, J. W. Kim, K. H. Choi, S. J. Kim and M. S. Alam, “Distortion-invariant pattern recognition based on a synthetic hit-miss transform,” *Opt. Eng.*, vol. 43, pp. 1798-1803, Aug. 2004.

[7] Q. Wang, Y. Deng and S. Liu, “Morphological fringe-adjusted joint transform correlation,” *Opt. Eng.*, vol. 45, pp. 087002-1-087002-9, Aug. 2006.

[8] P. Garcia-Martinez, M. Tejera, C. Ferreira, D. Lefebvre and H. H. Arsenault, “Optical implementation of the weighted sliced orthogonal nonlinear generalized correlation for nonuniform illumination conditions,” *Appl. Opt.*, vol. 41, pp. 6867-6874, Nov. 2002.

[9] S. Martínez-Díaz and V. Kober, “Nonlinear synthetic discriminant function filters for illumination-invariant pattern recognition,” *Opt. Eng,.*vol. 47, pp. 067201-1-067201-9, Jun. 2008.

[10] S. Martínez-Díaz and V. Kober, “Morphological correlation for robust image recognition,” in *Proc. IEEE 2011 International Conference on Computational Science and its Applications*, Santander, Spain, Jun. 20-23, 2011, pp. 263-266.

[11] P. Maragos, “Morphological correlation and mean absolute error criteria,” in *Proc. conf. IEEE Trans. Acoust. Speech Signal Process*, 1989, pp. 1568-1571.

[12] J. P. Fitch, E. J. Coyle and N. C. Jr. Gallagher, “Median filtering by threshold decomposition,” *IEEE Trans. Acoust. Speech Signal Process*, pp. 1183-1188, 1984.

[13] M. I. Vardavoulia, I. Andreadis and P. Tsalides, “Hardware implementation of soft color image morphological operations,” *Opt. Eng.*, vol. 41, pp. 1536-1545, Jul. 2002.

[14] S. Bouchoux and E. Bourennane, “Application based on dynamic reconfiguration of field programmable gate arrays: JPEG 2000 arithmetic decoder,” *Opt. Eng.*, vol. 44, pp. 107001-1-107001-6, Oct. 2005.

[15] E. Grivas, E. D. Kyriakis-Bitzaros, G. Halkias, S. G. Katsafouros, G. Morthier, P. Dumon and R. Baets, “Wavelength division multiplexing based optical backplane with arrayed waveguide grating passive router,” *Opt. Eng.*, vol. 47, 025401-1-025401-7, Feb. 2008.

[16] R. N. Rakvic, R. W. Ives, J. Lira and C. Molina, “Case for a field-programmable gate array multicore hybrid machine for an image-processing application,” *Journal of Electronic Imaging*, vol. 20, pp. 013015-1-013015-9, Jan-Mar 2011.

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)