AN OPTICAL CNN IMPLEMENTATION WITH STORED PROGRAMMABILITY

Research Report
DNS-5-2000

Sz. Tőkés, L. Orzó, Cs. Rekeczky, T. Roska, Á. Zarányd

Analogical & Neural Computing Systems Laboratory at the Computer and Automation Research Institute of the Hungarian Academy of Sciences,
Budapest, April 2000

ABSTRACT
The objective of this paper is to provide a framework for the implementation of Programmable Opto-Electronic Analogic CNN (POAC) Computers embedding CNN Universal Chips. Specifically, a new method for optical CNN implementation is provided and some details are experimentally studied. The POAC architecture includes the integration of an optical processing system, such as a joint transform correlator, with the fast spatio-temporal processing capabilities of a CNN-UM chip. We have built and tested an optical sub-unit of this experimental opto-electronic architecture to examine their processing capabilities for complex target recognition tasks. Preliminary result of these measurements will also be presented. The main idea is to introduce stored programmability into optical computing.

1 INTRODUCTION
The demand for fast identification and tracking of targets e.g. in surveillance systems has been increased dramatically during the last few years[1][2]. In several other image-processing tasks the quick recognition of particular structures is also important. The very fast, online pre- and post-processing of the flow of image data is inescapable. Optical information processing systems can provide appropriate speed to solve these demands. However, the so far published optical processing system architectures do not seem to be flexible enough to be applicable in different computational tasks.
In recent years, several studies have demonstrated that a cellular neural/nonlinear network (CNN) type architecture[3][4], provides exhaustive programming frame for several complex, image-processing tasks[7][8]. The CNN Universal Machine (CNN-UM) is a massively parallel nonlinear array processor[5][6]. While several emulated digital and mixed-signal analog implementations are emerging, there have been only a few attempts to build optical or opto-electronic implementation [10]-[18] of the CNN-UM. Optical correlators, however, can implement one of the basic operations of CNN computation: the convolution.
The POAC computer framework consisting of an optically implemented CNN combined with fast and re-programmable optical input VLSI CNN-UM chips present an ideal solution for the
above outlined problems. In the optical implementation a large number of templates can be stored and retrieved. In this framework a considerable, in some case the dominant part of the processing can be done at the speed of light, and the rest of the processing on a fast parallel opto-electric device.

In this paper, we propose a special hybrid opto-electronic CNN computer architecture: an implementation of Programmable Opto-Electronic Analogic CNN (POAC) Computers embedding CNN Universal Chips. We shall introduce and analyze this new type of feedforward only optical CNN implementation. We will demonstrate its flexibility in some image processing tasks. First we give the theoretical and structural description of the new optical CNN implementation (section 2) and this architecture's extension and combination with the existing VLSI CNN technology (section 3). Later we describe the measurements made on this architecture (section 4). Finally we demonstrate the architecture's image processing capabilities by some simple image processing tasks (section 5) and give the conclusion and discuss our results (section 6).

2 JTC FOR OPTICAL CNN

2.1 Optical correlators

The joint transform correlator (JTC)[19][20] is a powerful optical information-processing unit for pattern identification. It shows increased robustness comparing to a matched filter VanderLugt correlator (VLC)[21].

Advantages of JTC realization[22]:

- Use of a spatial-domain (impulse response) filtering (no previous calculations and computer processing is needed to synthesize the Fourier–domain filters which are necessary for VLC).
- The JTC has a higher space-bandwidth product and a lower carrier frequency.
- It has a higher index modulation.
- It is suitable for real time applications.
- It is much more robust against vibrations and misalignment compared with the VLC and its robustness is comparable with that of the incoherent correlators.

Drawbacks of JTC:

- It suffers from moderate (lower) detection efficiency when applied to multiple target recognition or targets embedded in intense background noise.
- Furthermore, high spatial coherence is required, but it does not need large coherence length.
- The JTC method is less efficient from energetic point of view as the first order diffractive beam, which is providing the desired convolution, carries only about 1/8th of the incident energy.

The following equations (1,2,3) describe mathematically the JTC's operations.

\[ \text{Input} = s(x + x_p, y) + t(x - x_p, y) \]  (1)

Where s and t corresponds to the input image and to the template.
Joint Power Spectrum = $S^2(\alpha, \beta) + T^2(\alpha, \beta) + 
+ S(\alpha, \beta)e^{-i\varphi_2(\alpha, \beta)}T(\alpha, \beta)e^{i\varphi_2(\alpha, \beta)}e^{-i2x_0a} + 
+ S(\alpha, \beta)e^{i\varphi_2(\alpha, \beta)}T(\alpha, \beta)e^{-i\varphi_2(\alpha, \beta)}e^{i2x_0a}$

where $S(\alpha, \beta)e^{i\varphi_2(\alpha, \beta)}$ and $T(\alpha, \beta)e^{i\varphi_2(\alpha, \beta)}$ corresponds to the Fourier transforms of $s(x, y)$ and $t(x, y)$.

\[ \text{Output} = s(x, y) \ast \tilde{s}(x, y) + t(x, y) \ast \tilde{t}(x, y) + 
+ \tilde{s}(x - 2x_0, y) \ast t(x, y) + s(x + 2x_0, y) \ast \tilde{t}(x, y) \]

The previous optical CNN implementations, however, had used mainly the VLC type correlator [16]. The main problem with optical CNN VLC implementation is the slow, offline construction of the appropriate complex, computer designed holographic filter, corresponding to the necessary template. This design is sensitive for the precise positioning of the elements.

2.2 Programmable Optical elements

In the last few years a lot of high-resolution programmable optical device have emerged. These are the electronically or optically addressable spatial light modulators. These new elements provide a new framework for optical information processing as well. These spatial light modulators can work parallel at a very high speed. The resolution of a typical Electronically Addressable Spatial Light Modulator (ESLM): 1024x1204 pixels (ferroelectric liquid crystal: binary or nematic LC: gray-scale).

The Optically Addressable Spatial Light Modulator (OASLM): The QuantaImage (CoreTek) OASLM is capable of resolving 220 lines/mm on a 1cm² surface, this means a resolution of about 2000x2000.

2.3 Basic structure and function of this design.

In our approach, the joint Fourier transform correlator (JTC) will be used in a novel way for preprocessing, since it accomplish the basic feedforward-only CNN operations. For example, this makes it possible to realize mathematical morphology (MM) processing in the CNN framework (for MM tasks the JTC seems to be more suitable than for general-purpose pattern recognition). The basic plan of the JTC based POAC is shown in Figure 1.

In the proposed architecture the unknown input image from an electronically addressed spatial light modulator (ESLM) is correlated with the template(s) considered as reference image(s). Here we have to mention, that in the current setup we use photographic images as input, but in the near future we will change it to a programmable (ESLM) device. Laser 1 is the coherent light source (red, He-Ne). Lens FL1 Fourier transforms the images on the OASLM in which the interference fringes are recorded.
Presently, we use a moderate resolution OASLM from Jenoptik GmbH. To increase the processing capability, a higher resolution OASLM - e.g. from CoreTek Inc. - will be used. The interferogram between the images is read-out by the light of laser 2 (green, He-Ne) and Fourier transformed by lens FL2 and projects the correlation peaks onto CCD camera. A personal computer (PC) is used to drive the image SLM.

So a classical JTC will be installed with OASLM as the holographic material. The template will also be, in the initial stages, a few photographic images (binary and gray-scale) and the output, the cross-correlation (convolution) terms, is recorded on a CCD camera.

3 POAC COMPUTER

3.1 Architecture

More complicated processing can be carried out by the composition of our optical CNN implementation with CNN-UM chips. The basic structure of the proposed architecture can be seen in Figure 2.

Possible advantages of POAC realization:

- *The large neighborhood templates can find complex images on the input image.*
- *A VLSI implemented CNN-UM can perform the necessary further computations, considering the feedback and complex algorithms.*
- *The CNN-UM can solve the adaptive scaling and thresholding of the optical feedforward only CNN (JTC) input.*

So the optical feedforward only CNN implementation output can be the optical input of the VLSI CNN-UM chip.
By the modification of the amplitude distribution of the reading light beam we can implement an additional template operation. The scheme of this operation is also denoted in Figure 2. This architecture provides the possibility to ensure adaptively the balance between input image and primary template’s illumination. If the primary template is only a dot (Dirac delta), the reconstructed correlation image should be the same as the input. By adaptive changes of the input image’s illumination we can achieve optimal reconstruction. Assurance of this balance is essential and usually unavoidable for any types of further processing. In the followings we can use the additional template (t2 in Figure 2) for programming the POAC system. In this case even speed of light computation is accessible.

![Figure 2. The proposed Programmable Opto-Electronic Analog CNN (POAC) computer architecture.](image)

Mathematically the next equation can describe the operations.

\[
Output = F^{-1}(F(s(x-x_0)) + F(t_2f(x+x_0)))^2F(t_2f(x)) =
\]

\[
= F^{-1}(T_2(\alpha)(F^2(\alpha)+T_2^*(\alpha))) + 
\]

\[
+ s(x-2x_0)*t_2(x)*t_2(x) + 
\]

\[
+ s(x+2x_0)*t_2(x)*t_2(x) 
\]

(4)

where \(F\) and \(F^{-1}\) are the Fourier transform and its inverse.

3.2 Optical disk, ESLM as programming elements of the templates (Library)

For a proficient applicability of the POAC it is necessary to ensure programmability, the fast alteration of templates. It can be done by an ESLM or by optical discs [23][24], where the library of templates is stored previously.

3.3 Large neighborhood templates

Main advantage of the optical implementation, that there is no harsh limitation on the size of the applied templates as in the case of VLSI implementations. So by the POAC computer it is easy to implement multi-scale image processing tasks by applying a set of scaled templates of different size.
4 MEASUREMENTS

4.1 Sensitivity analysis

For the hologram's appropriate recording we have to consider the OASLM's finite resolution. Template and input image have to have balanced intensities as it is mentioned above. The spatial frequency bandwidth product (SFBW) of the OASLM is a key parameter, and we have to determine the appropriate scaling of the input images. It is unavoidable even in the case, when the input and the template are on a photographic film.

4.2 Noise analysis

To increase the signal to noise ratio of the output image we have to identify the different noise sources and measure them. Such as the detector noise: the optical to electrical signal conversions, inherent coherent noise of the optical system (speckles). Object scene noise: background clutter, target noise. A theoretical study of the JTC with phase-only images will also be done. If phase images can be correlated in the JTC, an increase in the photon efficiency can be made, enabling low power diode lasers to be used. Several discrimination-improving methods (nonlinear recording of the JTC power spectrum, Difference Of Gaussians (DOG) [26] and wavelet filtering, phase-encoding, position encoding etc.), will be tested and composite filter synthesis will be applied to overcome the disadvantages of JTC but harness its advantages for a POAC. In the current setup the dynamic range of the OASLM is limited and the zero order terms seems to be saturated. We try to overcome this problem by the application of another high-resolution photosensitive material like bacteriorhodopsin [29] or using appropriate phase modulation, which distributes evenly the spectrum [30]. The results of the measurements will be presented.

5 MATHEMATICAL MORPHOLOGY OPERATIONS

For the demonstration of image processing capacity of this new optical CNN implementation we show its performance for some simple mathematical morphology operations [12]. Although we know, that these operations can be easy to implement on conventional VLSI chips but optical preprocessing can extend both the image and template sizes and can provide their cross-correlation nearly at the speed of light. In the next figure (Figure 3) we demonstrate the capabilities of the existing JTC system. The system can find the reference object within the input image. After suitable thresholding we can determine its appropriate positions in the output image.
Figure 3. By the optical correlator we can manage letter identification. The orientation of the image and the reference object, template (letter 'K' and 'A' in these cases) shows which letter was to be recognized. This is the first step of further CNN processing. 3a and 3a' are the inputs, 3b and 3b' are the output, 3c and 3c' are the thresholded correlogramm.

5.1 Simple CNN-POAC computation.

If we combine the JTC system with the capabilities of the CNN universal machine we can improve the capabilities both paradigms. To demonstrate it we made a simple example. The JTC system can solve relatively easily the position determination of different letters, while the CNN-UM system can solve the further necessary algorithmic steps. These later ones can be mathematical morphology operations (the JTC system will be able to perform these, but due to the limitations of the current setup these were implemented by the CNN simulator) and the recall procedure etc. In the next figure (Figure 4) we show the results of the combined system operations.
Figure 4. The POAC system can solve even complicated pattern recognition tasks. The JTC can determine the position of the C and N letters. The JTC input screens can be seen on a, and b. Only the template has been changed. The correlation peeks can be seen on c, d, and their thresholded version e, f, respectively. The recall of these letters (h,) from the original input (g,) can be solved by the CNN-UM by simple mathematical morphology and recall operations.

Our result tries to demonstrate the capabilities of the POAC system even in this state of the work. In a programmable POAC system, we will be able to implement more complex algorithms.
6 CONCLUSIONS AND DISCUSSION

The POAC computer architecture has been proposed that combines optical preprocessing opto-electronic CNN post-processing in the CNN-UM framework. The JTC based optical implementation has considerable advantages (programmability) and some disadvantages comparing to the so far published optical CNN implementations. Special efforts have to be made for noise reduction for discrimination enhancement. It seems to be possible by the introduction of optical and detection nonlinearities and the application of suitable adaptive thresholds. There are further conceptual competing architectural approaches. We shall later examine other architectures such as the dual axis JTC [27][28] and VLC with online, spatial domain template inputs. Further improvements and application of these techniques can ensure the realization of feedback, which is essential part of the CNN paradigm.

7 ACKNOWLEDGEMENTS

This research was supported by Computer and Automation Research Institute of the Hungarian Academy of Sciences (SZTAKI).

8 REFERENCES