HW/SW Implementation of Hyperspectral Target Detection Algorithm

Đordije Bošković, Milica Orlandić
Department of Electronic Systems
NTNU
Trondheim, Norway
Email: dordijeb@stud.ntnu.no, milica.orlandic@ntnu.no

Sivert Bakken, Tor Arne Johansen
Centre for Autonomous Marine Operations and Systems (NTNU-AMOS), Department of Engineering Cybernetics
NTNU
Trondheim, Norway
Email: sivertba@stud.ntnu.no, tor.arne.johansen@ntnu.no

Abstract—Hyperspectral images obtained by imaging spectrometer contain a vast amount of data which require techniques such as target detection to extract useful information. This article presents an implementation of the target detection method Adaptive Cosine Estimator (ACE) for hyperspectral images. The algorithm is implemented as hardware-software partitioned system on Zynq-7000 development platform. The computationally intensive operations are accelerated on FPGA with the speed-up factor of 28.54. The timing analysis presents results for the partitioned system as well as for the software implementation on Zynq processing system used for comparison. The detection performance of the implemented algorithm is tested and verified using publicly available hyperspectral scenes with ground truth data.

Index Terms—hyperspectral imaging, target detection, Adaptive Cosine Estimator (ACE), FPGA

I. INTRODUCTION

Hyperspectral imaging combines digital imaging and spectrometry; its main goal is capturing and processing of images consisting of many spectral components. As each object has its own spectral signature defined by varying reflectance as a function of wavelength, it is essential to measure the distribution of electromagnetic radiation in certain spectral bands. Opposed to typical true color imaging which is adjusted to human spectral sensitivity, hyperspectral imaging can also include abundance of wavelengths outside the visible spectrum. In satellite remote sensing applications, the data obtained from the imaging spectrometer provides significant information about the spectral characteristics of surfaces and materials of the Earth [1].

The data from the hyperspectral camera can be represented as a spectral cube, consisting of two spatial and one spectral axis. Therefore, a pixel in a hyperspectral cube is an array of captured spectral intensities for certain spatial coordinates. Depending on the underlying architecture, the cube is usually stored and accessed in BSQ, BIP or BIL format [2]. It is also useful to plot spectra in the spectral space, where each of the spectral bands can be considered as one dimension in N-dimensional space. Therefore, the point (or vector) in spectral space actually represents one spectrum. Since all materials exhibit some variability in their reflectance spectra, it is expected that a particular material will be empirically characterized as a cloud of points in spectral space.

The increasing amount of spatio-spectral data obtained by modern hyperspectral imagers has contributed to creating new challenges in hyperspectral data processing, especially in scenarios which require real-time operation [3]. Considering satellite hyperspectral missions, another important issue is the constrained down-link bandwidth from a satellite to the ground stations [4]. To fulfill this, intensive on-board processing performed on powerful platforms can be used. The systems with FPGAs have become an integral part of satellite remote sensing missions [5], characterized by low power consumption, inherent reconfigurability, high parallelization capabilities and commercial availability. The on-board hyperspectral data processing tasks are usually pipelined, consisting mainly of the following stages: binning, optical and sensor corrections, radiometric corrections, geo-referencing and registration, motion blur correction, super-resolution, atmospheric correction and dimensionality reduction, etc. The reduced data cube can then be processed by a target detection system.

A number of target detection algorithms, such as constrained energy minimization (CEM) [6], adaptive cosine estimator (ACE) [7] and spectral angle mapper (SAM) [1] have been widely used for hyperspectral images. Target detection is involved in many civilian and military applications [8], such as detection of vehicles, infrastructure, vegetation, pollution or harmful algae species in oceans. As such, the objects of interest can be scarcely populated in the scene or constitute a significant portion of the image [1]. The spectral signature of the target is required for mentioned algorithms, and it can be obtained from a spectral library or extracted from the scene for testing purposes.

The remainder of the paper is organized as follows. Section II describes target detection algorithms as well as metrics used to estimate their detection performance. In section III, analysis of detection performance of considered algorithms for FPGA implementation is briefly explained. Section IV describes hardware-software codesigned implementation of modified Adaptive Cosine Estimator (ACE) on ZedBoard Zynq-7000 System-on-Chip. Section V provides analysis of the proposed FPGA-based implementation using hyperspectral scenes with provided ground truth data. Finally, section VI concludes with guidelines for future development.
II. TARGET DETECTION ALGORITHMS

The objective of target detection algorithms is to find an object of interest in the hyperspectral image. The algorithms inspected in this paper are based on the statistical approach, where spectral reflectance features are exploited to identify the target. A typical target detection system consists of target detection algorithm and threshold selection system, as shown in Fig. 1.

![Hyperspectral target detection system](image)

Fig. 1. Hyperspectral target detection system, adapted from [8].

The target detection algorithm maps input pixel vector \( \mathbf{x} \) onto a scalar value \( y = D(\mathbf{x}) \), where \( y \) is called detection statistic. In other words, the target detection algorithm provides the system with a numerical value which is related to the probability of the inspected pixel to be a designated target. Afterwards, the detection statistic \( y \) is compared to a threshold value \( \eta \) in order to determine if the input pixel contains target signature. The threshold provided by threshold selection system is based on the estimated background image data. The optimum threshold is set so that a significant amount of present targets are detected and the false alarm rate is kept below a certain value. Threshold selection system is out of scope of this paper.

In statistical signal processing, target detection is regarded as binary hypothesis testing between a null and alternative hypothesis. Null hypothesis \( H_0 \) asserts that the pixel being tested is not a target, whereas the alternative hypothesis \( H_1 \) asserts the observed pixel as target. Modelling the signals under both hypothesis is characteristic for each target detection algorithm. A number of target detection algorithms are selected for further analysis.

A. Spectral Angle Mapper

The detection problem can be presented as follows:

\[
H_0 : \mathbf{x} = \mathbf{b} \\
H_1 : \mathbf{x} = \alpha \mathbf{s} + \mathbf{b}
\]  

(1)

where \( \mathbf{b} \) is background clutter and noise, \( \mathbf{s} \) is the known target vector, and \( \alpha \) represents a parameter influenced by illumination and sub-pixel mixing. In the case of the random, zero-mean and normally distributed background \( \mathbf{b} \), the spectral angle mapper [1] is defined as:

\[
D_{SAM}(\mathbf{x}) = \frac{(\mathbf{s}^T \mathbf{x})^2}{(\mathbf{s}^T \mathbf{s})(\mathbf{x}^T \mathbf{x})}.
\]  

(2)

An equivalent detection statistic is given as:

\[
D_{SAM}(\mathbf{x}) = -\cos^{-1} \frac{\mathbf{s}^T \mathbf{x}}{\sqrt{(\mathbf{s}^T \mathbf{s})(\mathbf{x}^T \mathbf{x})}}.
\]  

(3)

which represents an angle between a reference spectrum and a pixel under test [1].

B. Constrained Energy Minimization

The target detection algorithm can also be designed as a FIR linear filter \( \mathbf{h} = [h_1, h_2, ..., h_K]^T \) [6] with detection statistic given as:

\[
D(\mathbf{x}) = \mathbf{h}^T \mathbf{x}.
\]  

(4)

The vector \( \mathbf{h} \) is optimized so that the detection statistic better separates the background clutter and the target. The optimization is performed by minimizing the background energy under the following constraint:

\[
\min(\mathbf{h}^T \mathbf{R}) \text{ subject to } \mathbf{h}^T \mathbf{s} = 1
\]  

(5)

where \( \mathbf{R} \) is the sample correlation matrix. The resulting detection statistic of constrained energy minimization is then given as:

\[
D_{CEM}(\mathbf{x}) = \frac{\mathbf{s}^T \mathbf{R}^{-1} \mathbf{x}}{\mathbf{s}^T \mathbf{R}^{-1} \mathbf{s}}.
\]  

(6)

C. Adaptive Cosine Estimator

A model of the detection hypotheses given as:

\[
H_0 : \mathbf{x} = \beta \mathbf{b} \\
H_1 : \mathbf{x} = \alpha \mathbf{s} + \beta \mathbf{b}
\]  

(7)

leads to the Adaptive Cosine Estimator (ACE) [7], where the parameter \( \beta \) is the newly introduced scaling factor of the combination of noise and background clutter \( \mathbf{b} \sim N(\mu, \sigma) \). The detector is characterized by the following equation:

\[
D_{ACE}(\mathbf{x}) = \frac{(\mathbf{s}^T \Sigma^{-1} \mathbf{x})^2}{(\mathbf{s}^T \Sigma^{-1} \mathbf{s})(\mathbf{x}^T \Sigma^{-1} \mathbf{x})}
\]  

(8)

where all factors are mean centered and the covariance matrix \( \Sigma \) can be estimated from the sampled image data.

To satisfy real-time performance requirements and to obviate the need for mean centering of the hyperspectral data, an adaptation of ACE algorithm is proposed. The adaptation consists of replacement of the covariance matrix with the correlation matrix as follows:

\[
D_{ACE-R}(\mathbf{x}) = \frac{(\mathbf{s}^T \mathbf{R}^{-1} \mathbf{x})^2}{(\mathbf{s}^T \mathbf{R}^{-1} \mathbf{s})(\mathbf{x}^T \mathbf{R}^{-1} \mathbf{x})},
\]  

(9)

where the correlation matrix \( \mathbf{R} \) is estimated from the given dataset or its subset.
D. Target Detection Algorithm Performance Metrics

In order to choose the algorithm to implement and later verify the implementation, performance of the target detection algorithms is evaluated using the Matthews correlation coefficient (MCC) and visibility metric [9]. The performance of the algorithm is usually visualized using a confusion matrix that contains the number of true positives, true negatives, false positives and false negatives. True positives represent correctly detected targets, while false negatives are present targets which are not detected by the algorithm. On the other side, some pixels might be regarded as targets even if they are part of the background which belongs to false positives count (Fig. 1). In contrast with that, true negatives are correctly classified background pixels.

1) MCC metric: MCC metric is defined as:

\[
MCC = \frac{tp \cdot tn - fp \cdot fn}{\sqrt{(tp + fp)(tp + fn)(tn + fp)(tn + fn)}}
\]  

where \( tp \) are true positive, \( tn \) are true negative, \( fp \) are false positive and \( fn \) are false negative counts. MCC score is a value in range from \(-1\) to \(1\), where \( MCC = 1 \) means successful detection of all hyperspectral targets without false positives or negatives by target detection algorithm for a given threshold. On the other side, \( MCC = -1 \) indicates that the algorithm always gives the opposite class in case of binary classification. This metric involves all four quadrants of the confusion matrix.

2) Visibility metric: The robustness of an algorithm is evaluated using the visibility metric, which is a measure of the algorithm’s ability to separate background clutter and target. It is given as:

\[
Visibility = \frac{|T_t - T_b|}{T_{max} - T_{min}}
\]  

where \( T_t \) is the average detection statistic for target pixels, and \( T_b \) is the average detection statistic for non-target pixels based on the ground truth data. Factors \( T_{max} \) and \( T_{min} \) are the maximum and minimum evaluated detection statistics in the scene for a given algorithm. The best and the maximum score of visibility is \(1\), and the lowest score is \(0\).

III. ANALYSIS AND ADAPTATION OF ALGORITHMS

The algorithms are tested on two hyperspectral datasets, namely, Pavia University scene and Salinas scene, which are publicly available on [10]. The testing of each algorithm is performed on full image dimensionality (all spectral bands), as well as in pipeline with dimensionality reduction technique - Principal component analysis (PCA). For Pavia scene, Painted Metal sheets and Meadows signatures are used, while tests on Salinas scene are performed using Lettuce romaine 4th and 5th week signatures. For testing purposes, the thresholds are generated from a linearly spaced array with values between maximum and minimum detection statistic for a given scene and corresponding spectral signature. Thus, the MCC scores are obtained as the maximum achievable value over a range of thresholds.

The results for ACE, ACE-R, CEM and SAM algorithms for both scenes are shown in Fig. 2. Although the SAM algorithm shows good performance, it is characterized by low visibility. This makes the algorithm non-robust and shows its inability to separate background clutter from the possible targets. The CEM algorithm shows slightly lower MCC and visibility score over almost all scenes. Finally, it can be observed that the adapted ACE-R algorithm has the same performance or outperforms CEM, ACE and SAM algorithms. The targets are extracted from the provided scenes and especially ACE(-R) algorithm performs well in these conditions, which is consistent with the literature [1]. In particular, ACE-R proves to have high both MCC score and visibility on the used datasets.

Target detection with dimensionality reduction by PCA is also performed on Pavia and Salinas scenes. After the testing, it can be concluded that lowering the number of dimensions does not drastically degrade the performance metric values. In certain cases such as Salinas scene, MCC and visibility scores for ACE and ACE-R algorithms are improved. One of the reasons for the improvement can be found in elimination of the noise in discarded dimensions. Compared to full image dimensionality, dimensionality reduction shows advantages in both substantially shorter computing time as well as improvement of the performance scores for certain algorithms and scenes. Based on the analysis, ACE-R is chosen to be implemented on an FPGA-based HW/SW partitioned system.
IV. IMPLEMENTATION OF THE TARGET DETECTION CORE

In this paper, on-board target detection system for hyperspectral images is prototyped on ZedBoard. This board contains Zynq-7000 System-on-Chip consisting of processing system (PS) with ARM Cortex-A9 CPU and programmable logic (PL). The sensor data is stored in DDR3 memory, which can be accessed by both PS and PL of Zynq SoC. The programmable logic can communicate directly with DDR memory through AXI interface and AXI Direct Memory Access (DMA) [11]. As a step towards the full FPGA implementation, HW/SW codesign implementation of ACE-R target detection algorithm has been proposed. The overview of the system is shown in Fig. 3.

The operations of the algorithm are partitioned on the heterogeneous platform between processing system and programmable logic with a special consideration of background estimation. Since the correlation matrix can be estimated from the obtained image or already known from the previous runs, the correlation matrix and its inversion are computed in software and transferred to the hardware accelerator for the further algorithm steps. The computationally intensive matrix and vector operations such as dot product are then performed in hardware.

The block diagram of the implemented FPGA accelerator is shown in Fig. 4. The accelerator contains three pipelined HW stages: the inverted correlation matrix $R^{-1}$ is initially uploaded to the BRAM and serves during the whole algorithm execution.

In the first stage, each spectral component is streamed from DDR through DMA via AXI interfaces, and fed to the dot product (DP) modules which perform the following operation:

$$\begin{bmatrix} L_1(\lambda_1) & L_1(\lambda_2) & \ldots & L_1(\lambda_K) \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1K} \\ r_{21} & r_{22} & \ldots & r_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ r_{K1} & r_{K2} & \ldots & r_{KK} \end{bmatrix} = \begin{bmatrix} x^T \text{row}_1(R^{-1}) \\ x^T \text{row}_2(R^{-1}) \\ \ldots \\ x^T \text{row}_K(R^{-1}) \end{bmatrix},$$

where $r_{ik}$ is an element of the precomputed inverted correlation matrix $R^{-1}$. To compute this vector-matrix product and produce a vector of dot product elements:

$$(s^T R^{-1} \mathbf{x}, s^T R^{-1} \mathbf{x}, \ldots, s^T R^{-1} \mathbf{x}),$$

it takes $K + \text{delay}$ clock cycles where $K$ is number of spectral bands and $\text{delay}$ corresponds to the pipeline delay. The resulting vector is then stored in intermediate registers in stage 2. In this stage, incoming stream from DMA is stored in shift register and used for computation of dot product with computed vector from stage 1. This results in computation of $s^T R^{-1} \mathbf{x}$. Additionally, in stage 1 product $s^T R^{-1} \mathbf{x}$ is calculated, where $s^T R^{-1}$ vector is stored in BRAM and used for each incoming pixel. The result of this operation is squared in stage 2. The final stage 3 can either be implemented in dedicated HW or in SW, while for HW implementation, a fixed-point divider is required, such as AXI IP divider [12] provided by Xilinx. This substantially adds to the latency of the output, however it can achieve significant speedup. On the other side, when performed in SW, the pre-processed detection statistics coming from FPGA accelerator are multiplied and divided as being written to DDR memory. Finally, to stream the output data and communicate with DMA, Master Output module has been designed. It acts as an AXI stream master interface which are not memory-mapped and allow data-burst mode.

V. RESULTS

The FPGA accelerator shown in Fig. 4 has been implemented in VHDL targeting ZedBoard development platform. Performance analysis of the design and individual sub-modules has been performed with post-synthesis results. The results are shown in Table I with annotated bit widths used in each module.

The FPGA accelerator implementation is constrained by dot product datapath module speed. Although Stages 1 and 2 can operate at the same maximum frequency, other parts of the design are able to operate at significantly higher frequencies than the core datapath of the design. Thus, the dot product module has been deeply pipelined for the optimal performance.
Most importantly, the speed-up of the accelerator is 28.54 times when FPGA fabric is clocked with frequency of 100MHz. The speed-up is evaluated in comparison with ARM Cortex-A9 processor running on 666.67MHz, while assuming that the correlation matrix is uploaded once and reused for consequent algorithm runs to detect a target signature. Compared to full processing time of 4.00s in software for Salinas scene (reduced to 16 spectral bands using PCA), where ACE-R algorithm without correlation matrix calculation takes 0.59828s, the full computation time achieved with the use of the accelerator is 3.29s, while ACE-R takes 0.02096s.

Post-synthesis resource utilization has been presented in Tables II and III for 18-bit and 16-bit input hyperspectral data samples, respectively. Number of bands is set to 16 for both cases. It is important to note that the dedicated DSP block on ZYNQ PL has inputs that are 25 and 18 bits wide. However, since the design must accommodate different bit widths, those parameters are generic and affect the performance of the accelerator. In that sense, resource utilization for two sets of bit width parameters is presented, with the resource-optimized parameter set (25,18), and non-optimal set (32,16) providing higher precision. Therefore, the BRAM elements are of 25 and 32 bits, respectively. It should be noted that Top Level includes stage 1 and 2 of the FPGA accelerator as well as corresponding Master Output modules. All modules were synthesized out-of-context with default Vivado settings. The bit widths of the BRAM elements and input samples in Table II correspond to recommended input bit width for DSP blocks on ZYNQ PL. It can be observed that the DP datapath uses exactly one block for this setup and no other programmable logic to create a multiplier or an adder.

Table III shows utilization results for bit width values which are not optimal for this architecture, resulting in usage of 2 DSP blocks to create a multiplier and additional logic elements to synthesize the wider accumulator. Overall, the resource utilization is higher in this case, emphasizing the importance of correct use of dedicated FPGA functional units.

![Fig. 5. Part of FPGA accelerator pipeline with annotated bit widths](image)

Detection results on Pavia scene using ACE-R algorithm are shown in Fig. 6 where ground truth map is given in Fig. 6(a). The detection results obtained using HW/SW codesigned implementation are presented in Fig. 6(b) and the error induced by fixed-point implementation is shown in Fig. 6(c). Table IV shows that the proposed fixed-point implementation does not significantly degrade the results in terms of MCC and visibility score.

### Table I
**Performance Analysis of HW modules**

<table>
<thead>
<tr>
<th>Module</th>
<th>Minimum period</th>
<th>Maximum frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS-PL system (32 x 16)</td>
<td>6.835ns</td>
<td>146.43MHz</td>
</tr>
<tr>
<td>FPGA accelerator (32 x 16)</td>
<td>5.745ns</td>
<td>174.06MHz</td>
</tr>
<tr>
<td>Stage 1/2 (32 x 16)</td>
<td>5.745ns</td>
<td>174.06MHz</td>
</tr>
<tr>
<td>DP controller (16 bands)</td>
<td>1.644ns</td>
<td>608.27MHz</td>
</tr>
<tr>
<td>DP datapath (32 x 16)</td>
<td>5.745ns</td>
<td>174.06MHz</td>
</tr>
<tr>
<td>Master Output (16 packets)</td>
<td>2.628ns</td>
<td>380.52MHz</td>
</tr>
<tr>
<td>BRAM module (32)</td>
<td>4.663ns</td>
<td>214.45MHz</td>
</tr>
</tbody>
</table>

### Table II
**Resource utilization report 25x18, 16 bands, 32 bit output**

<table>
<thead>
<tr>
<th>Module</th>
<th>Slice LUTs</th>
<th>Slice Registers</th>
<th>DSP blocks</th>
<th>BRAM tiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS-PL system</td>
<td>6371</td>
<td>9935</td>
<td>32</td>
<td>10.5</td>
</tr>
<tr>
<td>Top Level</td>
<td>537</td>
<td>1772</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>Stage 1</td>
<td>8</td>
<td>5</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Stage 2</td>
<td>8</td>
<td>22</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>DP controller</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DP datapath</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Master Output (16 pkt)</td>
<td>83</td>
<td>298</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BRAM wrapper (25 bit)</td>
<td>133</td>
<td>300</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>

### Table III
**Resource utilization report 32x16, 16 bands, 32 bit output**

<table>
<thead>
<tr>
<th>Module</th>
<th>Slice LUTs</th>
<th>Slice Registers</th>
<th>DSP blocks</th>
<th>BRAM tiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS-PL system</td>
<td>7567</td>
<td>11838</td>
<td>54</td>
<td>10.5</td>
</tr>
<tr>
<td>Top Level</td>
<td>1470</td>
<td>2937</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>Stage 1</td>
<td>905</td>
<td>1178</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>Stage 2</td>
<td>57</td>
<td>125</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>DP controller</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DP datapath</td>
<td>53</td>
<td>69</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Master Output (16 pkt)</td>
<td>83</td>
<td>298</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BRAM wrapper (32 bit)</td>
<td>148</td>
<td>302</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>
VI. CONCLUSION

In this paper, HW/SW codesigned implementation of ACE target detection algorithm for hyperspectral data has been presented. A heterogeneous platform with processing system and programmable logic has been used along with specific communication interfaces. A number of algorithm implementations have been profiled for performance and resource utilization for HW implementation has been reported. The implemented design serves as a solid ground for the full HW implementation where it is required to implement correlation and inverse matrix calculation in the programmable logic in order to accelerate the execution of the algorithm to fulfill near real time requirement.

ACKNOWLEDGEMENT

This work was supported by the Research Council of Norway (RCN) through MASSIVE project, grant number 270959, and AMOS project, grant number 223254.

REFERENCES