ADAPTIVE NOC-BASED MPSoC SYSTEM FOR SPECTRAL IMAGING ALGORITHM DEDICATED TO ART AUTHENTICATION.

Junyan Tan, Virginie Fresse, Frédéric Rousseau

Hubert Curien Laboratory UMR CNRS 5516, Jean Monnet University- University de Lyon, 18 Rue du Professeur Benoît Lauras 42000 Saint-Etienne, FRANCE e-mail: virginie.fresse@univ-st-etienne.fr

TIMA Laboratory, UJF/CNRS/Grenoble INP System-Level Synthesis Group, 46, Avenue Félix Viallet, Grenoble, FRANCE e-mail: frederic.rousseau@imag.fr

ABSTRACT
We propose a NoC-based MPSoC architecture dedicated to art authentication applications for FPGA. The design of the low power architecture is based on the reuse of IP blocks and the adaptation of the communication according to the application requirements. The designer can adapt the architecture and change the parameters of the algorithm without redesigning the complete system. The NoC-based MPSoC architecture fully exploits task parallelism and pipeline of the application. The performance evaluation is easy of use. The designer can explore and estimate the total execution time with a restricted number of implementations.

Index Terms— spectral imaging, adaptive architecture, real time system, MPSoC, NoC, performance exploration.

1. INTRODUCTION
Spectral imaging is a technology originally developed for space-based imaging. Over the last 15 years, attempts at multi-spectral imaging have been carried out. This technology captures lights from more frequencies than the human eye. Spectral imaging allows us to investigate the artwork in ways that were impossible before. In 2004, this technique has been applied for the first time with the Joconde painting. Relevant unknown regions and information were then extracted [1]. Such real-time applications require embedded low power systems with a high number of computing resources. Existing architectures such as CPU and GPU (Graphical Processing Unit) cannot cope with all constraints that are size, speed and power consumption. Previous experiments [12] showed that CPU cannot handle the real-time constraints for spectral image applications. It also has been shown that GPU cannot handle massive data transfers from PC and GPU. Moreover GPUs are not low power architectures.

Multiprocessor architecture using Networks on Chip (NoC) structure is considered as one of the most appropriate solution for embedded low power real-time spectral imaging. The Field Programmable Gate Array (FPGA) devices are widely used for multi-processor architectures. FPGA can achieve high-speed performances in a relatively small footprint. Modern FPGAs integrate many heterogeneous resources on one single chip. One FPGA is capable to handle all processing operations, including the acquisition part. It means that incoming data from the camera or any other acquisition devices are directly processed by the FPGA. Using a NoC-based MPSoC architecture requires explorations and evaluations to tune this architecture (number of processor and size/parameters of the NoC) according to the algorithm requirements. In this paper, we propose an embedded and adaptive architecture dedicated to art authentication for configurable chip FPGA. The design of the system is based on an existing algorithm keeping in mind that hardware and software modifications should be considered. Components selected and inserted are low power, with few resources and with optimized execution times to meet the real-time requirements. The architecture is based on a Network On Chip (NoC) communication to exploit all parallelism sources of the application.

This paper is organized into 6 sections. Section 2 presents the context and the art authentication algorithm in spectral. Section 3 introduces the NoC-based MPSoC architecture and the design flow is detailed in section 4. Experiments are in section 5. Section 6 contains the conclusion.

2. CONTEXT
Several projects related to art authentication in spectral have been investigated from several decades. In the 90s, the VASARI camera project started. This camera produced 7 channels of data between 400 and 700 nm with the final image stitched together from a mosaic of low resolution shots [2]. The MARC project was to replace photograph used in the VASARI project with scanning and printing solutions. It uses three channels (R, G, B) to capture high
resolution images from paintings [4] at high speed. The NARCISSE European project also started in 1990 [3]. The objective was to scan larger images in photography and radiography. The objective of the CRISATEL project (2001) [7] was also to develop a system for capture, processing and storage paintings. The CRISATEL camera can scan in one shot a much larger paintings. It generates 13 very high definition images with a size of up to 12,000x20,000 pixels. There are many other projects related in literature. These projects are used for art authentication, storage and processing for paintings, sculpture and many other artworks [5][6]. For all these projects, the size of images becomes bigger and bigger with an increasing number of wavelengths. Many applications also require the integration of several spectral cameras (for 3D, reconstruction, tomography….). The development of more and more sophisticated spectral camera leads to an increasing number of data to process. The systems developed should scan larger and larger images with the increasing resolutions and with different size of channels.

First several significant regions of images are extracted from the original image and the compared image. Average color values from these regions are computed. According to the experiment, the size of windows is parameterized. For the following experimentations, the window is 3x3 pixels. Then some color projections transform the average color values to color space values. Color projections concern the original image and the compared image. These color images are used for the comparison process of the authentication. Color distance is just the basic neutral geometry distance between the compared images and the original images. All existing color distances can be implemented. Several multispectral algorithms are used to calculate the multispectral distance with the original multispectral image data. For this process, some algorithms require high precision operations which imply large amount of floating-point data and complex functions.

The communication architecture dedicated to this application is fully parameterized. The designer adapts:
- The size of original and compared images (4x4, 8x8, 16x16, 32x32)
- The size of data (8, 16, 32 or 64- bit data)
- The size of windows for the average
- The number and type of functions in the algorithm.
- The size and number of wavelengths.

3. NOC-BASED MPSOC ARCHITECTURE

The embedded system is designed considering the power, the performance (in size and execution time), the interfaces and others. The design of the system is based on:
- The adaptation to any type and number of spectral cameras.
- The use of an efficient low power scalable communication structure (NoC).
- The integration of Processing Elements (PE). PE is either a dedicated IP or embedded processors connected to the NoC.
- The optimization of the NoC and the exploitation of the parallelism in the application.

Several image processing system have been proposed in the field of art authentication in spectral imaging. The systems are based on PC or/and graphical card. The IIPImage system [8] is used to visualize high resolution multi-spectral 16- bit images, to view image details in color or for each spectral channel and to super-impose and compare different wavelengths. A 3D graphic card is inserted to the PC to reconstruct the resulting color dynamically while interactively changing the light spectrum [9].

In this paper, we propose to use programmable device (FPGA) for art authentication application. The algorithm used for the design and validation of the adaptive system is dedicated to art authentication for paintings. The algorithm is based on comparison between the relevant images extracted respectively from the original artwork and the artwork to authenticate. For the art authentication process, OI is the information of the true picture (original images), and CI is the information of the other “similar” pictures (compared images). The algorithm used for the authentication process presented in Fig 1.

Fig 1. The art authentication algorithm.

Several image processing system have been proposed in the field of art authentication in spectral imaging. The systems are based on PC or/and graphical card. The IIPImage system [8] is used to visualize high resolution multi-spectral 16- bit images, to view image details in color or for each spectral channel and to super-impose and compare different wavelengths. A 3D graphic card is inserted to the PC to reconstruct the resulting color dynamically while interactively changing the light spectrum [9].
1. Spectral camera

Many spectral cameras are proposed and can be integrated in the embedded system [1][2][4]. Spectral Camera is used to acquire the hyperspectral target image at tens or hundreds of wavelengths simultaneously. Spectral cameras are available with different interfaces to match closely to the different application requirements. Each type of camera has different specifications in type of sensor, range of wavelengths, number and size of pixels and other specifications as depicted in Fig 2. The user has to define the most appropriate camera according to the experiment.

2. The NoC communication structure

NoC architectures have been new communication architectures developed since ten years [10]. These communication architectures improve the flexibility of communications subsystem of SoC, with high scalability, high performance and energy efficient customized solution. Today, it is the most viable solution for on chip communication that it enables Giga-scale integration in the system on chip. Generally, NoC architecture is composed of several basic elements depicted in Fig 3:

- Network Adaptor (NA): it enables PE to communicate with routing node they are connected to.
- Routing node: according to routing type, the switch sends packets to the appropriate link in the network.
- Links: connect the routing nodes together or switches to NI.
- Processing Element: these units correspond to various modules of SoC, such as IP blocks, memories, processors.

![Fig 3. Structure of the NoC. PEs are connected to the NoC using network adapter. Routing nodes route packets inside the NoC.](image)

3. Design Flow

The designer implements the art authentication algorithm on the NoC-based MPSoC architecture using the following steps:

- Defining the set of parameters associated to the algorithm (this is done by changing constants in h file where all parameters are listed).
- Defining the parameters of the NoC.
- Mapping all tasks of algorithm on each processor according to the data flow graph of the application.
- Mapping the master processor on one of the idle node. The master processor is used for the global supervision of the system.
- Evaluating the execution time for each function and for each communication.
- Changing the set of parameters to evaluate the impact on the performances.

![Fig 4. The data flow graph of the art authentication algorithm.](image)

All exploration and evaluation of the architecture are done inside the FPGA (emulation) with automatic tools and associated IP blocks. In [12], we presented the design flow and associated IP (emulation block) to evaluate the NoC structure only according to the communication requirements. The advantages of emulation compared to simulations are: results are accurate and exploration and evaluation times are faster. The designer can easily and quickly tune the entire architecture.

The data flow graph of the algorithm is depicted in Fig 4. For the application, 14 tasks constitute the algorithm. We select the 4x4 Mesh NoC to be able to deploy each function on one node (including the master processor). Each Processing Element connected to routers is the Plasma processor. The data dependencies between functions help the designer to map tasks onto the NoC with optimized solutions.

![Fig 5. The task mapping of the algorithm onto the NoC.](image)
The mapping of the tasks onto the 4x4 NoC is depicted in Fig 5. For constraint reasons (input and output), the average functions are spread over the NoC (at the opposite corner of the structure). The designer first places the input function (average_1 and average_2). Then each function is sequentially placed inward the NoC by going over the data flow graph. The number of hops between a source and a destination is little using this method. In this paper, the task mapping is a manual task. The designer can use any existing mapping algorithm to optimize the solution.

5. EXPERIMENTS

The embedded MPSoC architecture is based on the 4x4 Hermes NoC with 16-bit flits. The PE is the 32-bit PLASMA processor (MIPS-like). Hermes is a NoC created by the Catholic University of Rio Grande do Sul (Porto Alegre, Brazil) [11]. Previous automatic experiments explored the appropriate parameters used to tune the NoC structure according to the transfers required [12]. Implementations are done on the ML506 platform with a virtex 5 FPGA. The experiments are carried out within the Atlas tool, ModelSim, Xilinx ISE with Synthesis tool. In this section, the experiments are:
- Performance evaluation: extracting the number of clock cycles for each function according to several parameters of the algorithm.
- Precision analysis: we see the impact on the accuracy of results with the use of the 32-bit fixed point representation compared to traditional 64-bit floating point architecture used by CPU.
- Design space exploration: we change the number of Processing Elements and we see the impact in the performance of the system with/without considering pipeline.

Performance evaluations are based on the number of clock cycles for each function and for the total application. We also define the execution time for processor with frequencies of 25 MHz (the PLASMA) and 100 MHz (as most of embedded processors run with this frequency).

Table 1. Execution time for each function on the PE.

<table>
<thead>
<tr>
<th>Functions</th>
<th>Nb of clock cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>30 287</td>
</tr>
<tr>
<td>XYZ</td>
<td>24 523</td>
</tr>
<tr>
<td>RGB</td>
<td>5 838</td>
</tr>
<tr>
<td>Lab</td>
<td>94 005</td>
</tr>
<tr>
<td>Dist_XYZ</td>
<td>29 595</td>
</tr>
<tr>
<td>Dist_RGB</td>
<td>29 518</td>
</tr>
<tr>
<td>Dist_WRMS</td>
<td>1 117 462</td>
</tr>
<tr>
<td>Dist_GFC</td>
<td>118 022</td>
</tr>
<tr>
<td>Dist_RMS</td>
<td>89 743</td>
</tr>
<tr>
<td>Total execution time</td>
<td>1 325 966</td>
</tr>
</tbody>
</table>

The execution times for each function of the application are evaluated on the Plasma PE (Table 1). Parameters are 16 wavelengths, 1 region and the 8x8 window. The timing does not consider the communication between PEs. The execution time for the WRMS function is the longest function compared to the other ones. It also represents 84% of the total execution time (Total E.T. in figures). The total latency of the application without considering the pipeline is 1.325,966 clock cycles. The time is 13.2 ms for the 100MHz processors and is 52.8 ms for 25MHz processors. Previous experiments in [12] showed that the execution time is 11.5 ms with GPU card (with a frequency of 576 MHz). Therefore a NoC-based MPSoC can provide identical performances than GPU with low power and without integrating PC.

Fig 6. The execution time for each function and the total execution time according to the number of wavelengths (16 and 64).

In-depth timing experiments are done with varying parameters of the application. The execution times for each function and for the complete algorithm are given according to the number of wavelengths (Fig 6 and Fig 7). The execution time of the dist_wrms functions significantly and linearly depends on the number of wavelengths. The number of clock cycles for dist_wrms is 1,107,329 for 16 wavelengths and is 4,278,147 for 64 wavelengths. In Fig 7, we see that the total execution time linearly depends on the number of wavelength used.

Fig 7. The total execution time according to the number of wavelengths.

We now consider the communication between blocks previously extracted in [12] and the functions executed. Parameters are 992 wavelengths, 1 region and the 8x8 window. In this case, task parallelism is exploited (functions are executed on the original data and compared data in the same time). The total execution time of the application is respectively 1.57s and 394.5ms for one region with the processor frequencies of 25 MHz and 100 MHz). We previously obtained 150 ms for a 576 MHz GPU
We can have identical execution time with 200 MHZ processor compared to GPU. It means that we can have identical timing results with low power processor.

We now analyze the NoC-based MPSoC architecture in the precision of results. GPU architectures are 64-bit floating point. The MPSoC architecture is a 32-bit fixed precision (16 bits are used for the decimal part). Table 2 shows the comparisons of results using both architectures. The analysis is based on existing spectral database of paintings (with 64 wavelengths, 8x8 pixels for one region). The error of precision is below 1.1% with an average precision of 0.21% of the global application (from the input data to the output data). According to the user, this precision is acceptable for art authentication application.

We decide to explore the design space of the architecture. Previous analysis did not consider the pipeline of the application that can be fully exploited by the MPSoC architecture. The size of the NoC is first changed from the size 1x2 to 4x4. The tasks are mapped according to the size of the NoC. For a small size of the NoC, more than one function is running per processor.

Table 2. Precision results for each function on 32-bit architecture compared to 64-bit precision simulation.

<table>
<thead>
<tr>
<th>Functions</th>
<th>Simulation (Matlab)</th>
<th>MPSoC architecture</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td>48.9847</td>
<td>48.984</td>
<td>0.0014</td>
</tr>
<tr>
<td>X</td>
<td>465.713</td>
<td>464.8841</td>
<td>0.1780</td>
</tr>
<tr>
<td>Y</td>
<td>60.38</td>
<td>60.3451</td>
<td>0.0578</td>
</tr>
<tr>
<td>Z</td>
<td>2291.4663</td>
<td>2290.498</td>
<td>0.0423</td>
</tr>
<tr>
<td>R</td>
<td>1085.1636</td>
<td>1082.7204</td>
<td>0.2666</td>
</tr>
<tr>
<td>G</td>
<td>-366.7739</td>
<td>-364.4709</td>
<td>0.6279</td>
</tr>
<tr>
<td>B</td>
<td>2099.1743</td>
<td>2077.3411</td>
<td>1.0401</td>
</tr>
<tr>
<td>L</td>
<td>82.0601</td>
<td>82.02</td>
<td>0.0489</td>
</tr>
<tr>
<td>A</td>
<td>426.1137</td>
<td>426</td>
<td>0.0267</td>
</tr>
<tr>
<td>B</td>
<td>-382.5499</td>
<td>-383</td>
<td>0.1177</td>
</tr>
<tr>
<td>Dist_XYZ</td>
<td>252.9862</td>
<td>252.85</td>
<td>0.0538</td>
</tr>
<tr>
<td>Dist_RGB</td>
<td>258.64</td>
<td>256.3455</td>
<td>0.8871</td>
</tr>
<tr>
<td>Dist_RMS</td>
<td>0.6622</td>
<td>0.662</td>
<td>0.0302</td>
</tr>
<tr>
<td>Dist_WRMS</td>
<td>0.7378</td>
<td>0.7375</td>
<td>0.0407</td>
</tr>
<tr>
<td>Dist_LAB</td>
<td>20.2119</td>
<td>20.2623</td>
<td>0.2494</td>
</tr>
<tr>
<td>Dist_GFC</td>
<td>1</td>
<td>1</td>
<td>0.0100</td>
</tr>
</tbody>
</table>

Then we extract the levels of pipeline of the algorithm for 2000 regions. The execution time considering or not the pipeline and the levels of pipeline are depicted in Fig 8. The application is data-flow oriented and the number of pipeline can be high compared to GPU architecture. GPU architecture does not support pipeline because of the memory and the communication used [12]. The architecture can reach a maximum of 10 levels of pipeline for the 4x4 architecture (pipelining data through the average, YYZ, RGB, lab and the distance functions associated to their communications). It indicates that 10 regions can be processed in the same time (4 regions for the GPU architecture only). The longest time in the application is the LAB function and the associated communication time which is 1,851,950 clock cycles. Parameters are 992 wavelengths and 8x8 windows. The architecture can reach 18s for the 25MHz frequency. The total execution time is 101s for the GPU architecture. Such application has a high level of pipeline that is fully exploited by the proposed MPSoC architecture.

![Fig 8. The total execution time according to the size of the NoC and the pipeline of the application.](image.png)

6. CONCLUSION

In this paper, we design an efficient low power MPSoC architecture that integrated embedded processors and the NoC structure dedicated to art authentication application. This architecture fully exploit pipeline and task parallelism on one single FPGA. The proposed architecture gives better performance than GPU with low power in an embedded system. The proposed architecture is generic enough to change parameters of the algorithm without a complete redesign compared to GPU design.

7. REFERENCES