EVPS: An Automotive Video Acquisition and Processing Platform

Christophe Flouzat, Erwan Piriou, Mickaël Guibert, Bojan Jovanović and Mohamad Oussayran
CEA-LIST, Centre CEA Saclay, 91191 Gif-sur-Yvette CEDEX, FRANCE - Email: surname.name@cea.fr

***Abstract*—This paper describes a versatile and flexible video acquisition and processing platform for automotive. It is designed to meet aggressive requirements in terms of bandwidth and latency when implementing ADAS functions. Based on a Xilinx Ultrascale+ FPGA device, a vision processing pipeline mixing software and hardware tasks is implemented on this platform. This setup is able to collect four automotive camera streams (MIPI CSI2) and process them in the loop before transmitting a more intelligible pre-processed/enhanced data.**

***Index Terms*—Automotive, Video processing, MPSOC, FPGA, PCIe, Ethernet TSN, ADAS**

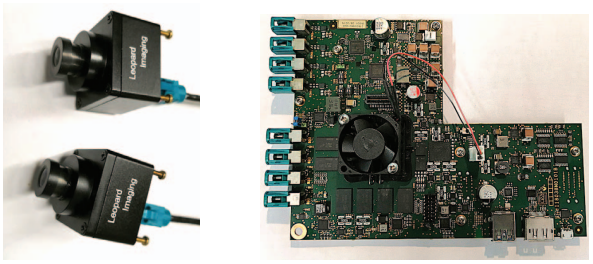

Fig. 1: The cameras and the *EVPS* prototype board

I. Introduction

Nowadays with the advent of the autonomous cars, automotive companies (from manufacturers to different tier levels) are involved in the design of more complex systems that depend on electrical and electronic (E/E) architecture. The transition from simple driving to aided driver assistance and finally to complete autonomous driving (AD) is based on a complete overhaul of existing platforms. All automotive actors have to face new challenges to integrate more and more advanced services while respecting drastic constraints in terms of safety, reliability, and security in operation [2]. Considering car perception, heterogeneous and flexible platforms are required to both perform near sensor processing (e.g. cameras, LIDAR) while more costly tasks such as fusion and decision making, run in a centralized manner. In particular, the aggregation of video streams and the processing of data streams is one of the major challenges. Collecting raw data pixels from typically 8 cameras stream mounted on a vehicle requires a large amount of data bandwidth (e.g. up to 6.4 Gbps for eight 1.3 MP cameras @60 fps). Furthermore, considering limitations of automotive TSN Ethernet in terms of bandwidth and latency (up to 1 Gbps in 802.1TSN) [1], an interfacing system able to handle both heterogeneous video streams and various automotive and others I/O interfaces is needed.

II. The *EVPS* hardware platform

The Embedded Vision Perception System *EVPS* board (Fig. 1) is designed to acquire simultaneously 4 camera streams through dedicated coaxial transmissions, along with on-the-fly data pixel processing onboard the Xilinx ZU5EV device.

The heart of the system is a *Xilinx* FPGA/MPSoC processing unit, the Ultrascale+ ZU5EV, which is the most powerful heterogeneous processing module available in the current and next-future roadmaps of Xilinx for the automotive industry. In a single chip, the device encloses a FPGA-like programmable logic (PL) part with many integrated IP cores, and a processing system (PS) made of a 64-bit quad-core ARM Cortex-A53 application processing unit, a 32-bit dual ARM Cortex-R5 real-time processing unit, and high-speed peripherals. To evaluate both *Maxim Integrated* Gigabit Multimedia Serial Link 2 (GMSL 2) and *Texas Instrument* Flat Panel Display Link III (FPDLink III) SerDes, the board is equipped with 2 sets of 4 FAKRA-Mini connectors, connected to the ZU5EV chip. Dedicated MAX96712 and DS90UB960-Q1 deserializers can be individually selected to convert serialized video data streams into (MIPI/CSI-2 format) compliant with FPGA inputs (Fig. 2). For standalone mode, the board embeds JTAG, ARM Trace, USB, Display Port interfaces, and circuitry for temperature, voltage as well as current monitoring. When plugged into a dedicated host platform, a Xilinx *TSN Endpoint Ethernet MAC* IP core implemented in the PL part manages data transfers at 100 Mbps and 1 Gbps. For higher data throughput, a 4-lane Generation 3 PCIe interface also in PL reaches up to 32 Gbps. Additionally, dedicated PCIe/TSN connectors are designed for automotive whereas a maximum of components onboard comply with the AEC-Q standard.

III. Video Processing chain for embedded systems

The video acquisition and processing pipeline is fully implemented in the PL part, while all tasks of configuration and control are left to the PS part running Petalinux (see Fig.2).

In the PL, two Xilinx *MIPI/CSI-2 Rx* IP cores extract data pixels into two AXI4-s data streams that are more convenient for further processing. These front-end cores include data bufferization essentially needed to alleviate the concurrency between the two MIPI paths toward the memory. By design, all nodes from camera sensors to the *MIPI/CSI-2 Rx* cores have a coherent throughput limitation of 6 Gbps per MIPI interface.

978-3-9819263-4-7/DATE20/©2020 EDAA

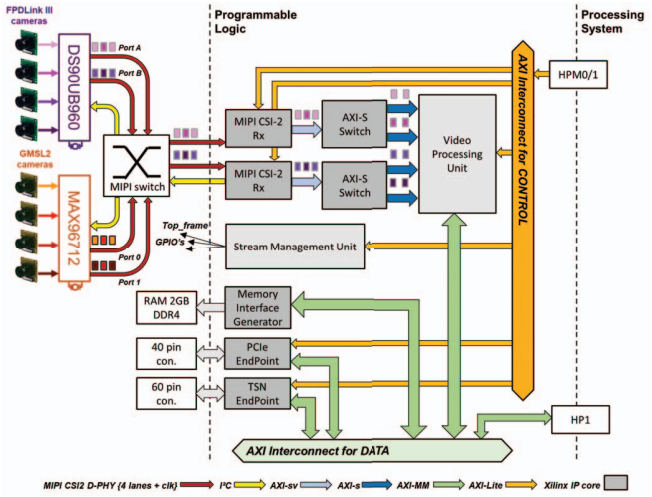

Fig. 2: Data Paths within the *EVPS* platform

The *Stream Management Unit* manages the synchronization of cameras for automotive stereovision use cases. Two Xilinx *AXI-s Switch* IPs sort camera data streams individually. The video processing unit includes a Xilinx *VDMA* core to transfer the pixels into the RAM dedicated to PL using a *"Genlock"* synchronization concept. PL internal memory consists of 5.1 Mb in BRAM and 18 Mb in URAM. In a standalone application, the PS can access the DDR4 PL-RAM (2 GB) for display purposes or even do more video processing using its own DDR4 PS-RAM (2 GB). When plugged to a motherboard, data can be accessed remotely through PCIe and TSN.

IV. Implementation

As an input, we consider two *FPDLink III* LI-AR0220 cameras developed by Leopard Imaging feeding the EVPS platform with a typical 1820x940/12-bit @ 60fps stream. The implemented VPU includes only 4 Xilinx demosaic IP cores followed by the corresponding VDMA's. The design works in a free running mode or as a triggered shooting of the cameras required by complex stereo vision algorithms like 3D stereo disparity [3] or point of interest tracking. The design also owns a 4-lane PCIe Gen.3 Xilinx IPs.

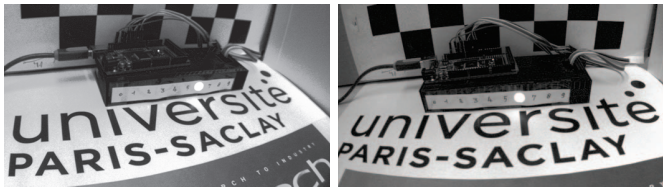

Fig. 3: Synchronous stereo acquisition of a LED chaser

A Petalinux OS running on the PS part includes all required drivers and a X server for monitoring through Display Port. A runtime task also handles pixels transfer from rolling buffer in the PL-RAM to the display. The validation of camera synchronization is shown in Fig. 3 with the tracking of a configurable LED chaser system (10 ms shift). Images taken in a single shot show systematically the same LED on.

TABLE I: Detail of the Resource relative to the full design

Entity	CLB LUTs	CLB Registers	BRAM	URAM
Top (units)	87726	109889	109.5	40
Available (units)	117120	234240	144	64
Utilization	74.90 %	46.91 %	76.04 %	62.50 %
Video Pipeline 0	22,40 %	20,58 %	11,87 %	50,00 %
Video Pipeline 1	22,40 %	20,60 %	11,87 %	50,00 %
PCIe	35,41 %	31,64 %	52,97 %	0,00 %
RAM interface	13,00 %	13,00 %	23,29 %	0,00 %
AXI interconnect	7,85 %	10,36 %	0,00 %	0,00 %

Table I illustrates resources utilization of our design in the ZU5EV device. The full design occupies 75% of the total Configurable Logic Blocks (CLB). Although, the main limitation relies on the PL internal memory utilization (BRAM/URAM). In our case-study, this limitation on distributed buffering requires a specific consideration when implementing complex vision IP's. According to Table I, PCIe IP and Memory Interface Generator (RAM DDR) need more than 50% and 20% of BRAM. Therefore, a balance must be reached between efficiency and internal buffers depth. Dealing with video pipeline, Xilinx demosaic IP cores is greedy in terms of internal memory resources. A custom implementation allows us to decrease the number of BRAM/URAM resp. CLB used by a factor of 3.5 resp. 5 while keeping same performances.

V. Conclusions

This paper describes an automotive prototyping platform EVPS capable of collecting four video streams with up to 3.2 Gbps throughput, processing and sending enhanced data through high speed interfaces such as Ethernet TSN/PCIe. In addition, most components also comply with the AEC-Q standard for reliability reasons. A flexible video processing pipeline is implemented on PL part while synchronization and software tasks are achieved on PS part. As nearly 75% of internal memory is used, future works will focus on dynamic reconfiguration and optimization of visual tracking IPs in order to reduce resource usage.

VI. Acknowledgments

This work was done within the FACE project with the *Renault-Nissan-Mitsubishi Alliance*. We wish to thank all the contributors of this industrial project.

References

[1] J. Migge, J. Villanueva, N. Navet, and M. Boyer, "Insights on the Performance and Configuration of AVB and TSN in Automotive Ethernet Networks," in *9th European Congress on Embedded Real Time Software and Systems (ERTS 2018)*, Toulouse, France, Jan. 2018. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01746132

[2] L. L. Bello, R. Mariani, S. Mubeen, and S. Saponara, "Recent advances and trends in on-board embedded and networked automotive systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 1038–1051, Feb 2019.

[3] M. Darouich, S. Guyetant, and D. Lavenier, "A reconfigurable disparity engine for stereovision in advanced driver assistance systems," in *Reconfigurable Computing: Architectures, Tools and Applications: 6th International Symposium, ARC 2010*, vol. 5992, 03 2010, pp. 306–317.