FPGA-based Image Analyzer
for Calibration of Stereo Vision Rigs

Aleksander Mielczarek, Dariusz Makowski, Piotr Perek, Andrzej Napieralski, Przemysław Sztoch

Abstract—**The paper presents a versatile solution facilitating calibration of stereoscopic camera rigs for 3D cinematography. Manual calibration of the rig can easily take several hours. The proposed device eases this process by providing the operator with several predefined layouts of the images from the cameras. The Image Analyzer is a compact stand-alone device, designed for the portable 19" racks. Almost all of the video processing is performed on a modern Xilinx FPGA. It is supported by ARM computer to provide control and video streaming over the Ethernet. The article presents its hardware design, as well as FPGA firmware and software architectures.**

Index Terms—**Computer aided analysis, Stereo image processing, Image fusion, Motion pictures**

I. Introduction

THE stereoscopic image recording is relatively new ground of the modern cinematography. The market for the stereoscopic motion pictures, often referred to as 3D films, is now in its rapid growth. This is caused by recent popularization of stereoscopic displays and development of related standards of video data transmission and storage. Although displaying the stereoscopic video material became quite easy, recording of such is still a complex task. Consumer-grade stereoscopic cameras are not suitable for professional productions hence the video acquisition is usually done by a set-up of two conventional cameras mounted on the camera rig.

The rig calibration is usually done by aligning image from cameras filming a board with sophisticated pattern of lines and other alignment markers [1]. Most rigs do not have any sensors nor actuators and are hence operated manually. Using the legacy equipment, the initial rig calibration can easily consume several hours.

The calibration time can be significantly reduced by providing the operator with several predefined layouts of the images from the cameras, facilitating calibration of the mechanical offsets as well as differences in camera zoom and focus. The preferred solution is a single integrated device, which connects to both cameras and provides analyzed image as well as pass-through signals.

II. Device Requirements

The Image Analyzer has to accept the video arriving from the cameras by means of the industry-standard Serial Digital Interface (SDI). The device should provide the composed output signal using HDMI and SDI interfaces. It should be also possible to provide the resulting video stream through a web service for preview e.g. on a mobile device.
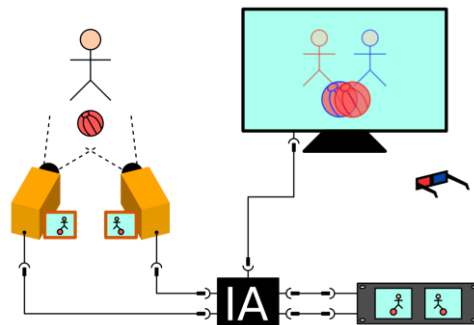

Fig. 1. Fundamental use case of the Image Analyzer

The Image Analyzer has to be a compact stand-alone device, that can be easily integrated with infrastructure already used on the set. Its place among other components (cameras, preview displays and main display) is illustrated in Fig. 1. Since there is often a small 19" rack present on the set, it was agreed to design the device as a regular 1 U rack module. The module has to accommodate mains or battery power supply.

III. Hardware Design

The Image Analyzer is composed of an FPGA board, several I/O modules, ARM-based Single Board Computer (SBC) and power supply. The device can be operated locally, by means of hardware keyboard, LCD and video OSD as well as remotely using a web-service.

The selected FPGA is modern Xilinx Kintex-7 integrated circuit: XC7K355T. It is responsible for almost all of the video processing. It is capable of receiving and sending 1080p60 (1920×1080 pixels, progressive − not interleaved, 60 FPS) stream over HDMI and 1080p30 stream over the SDI interfaces, by utilizing additional dedicated physical layer circuits. In case of SDI interfaces, only a proper signal equalization is needed.

P. Sztoch is with FINN Ltd., Lodz, Poland
Corresponding author: A. Mielczarek, e-mail: amielczarek@dmcs.pl.

The serialization and deserialization takes place directly in the FPGA. The HDMI interface was conveniently implemented with external highly-configurable serializers and deserializers. These devices communicate with FPGA by means of parallel 16-bit buses operating at the frequency of about 150 MHz.

The FPGA processing module cooperates with GateWorks Ventana GW5400 SBC. The module contains quad-core ARM Cortex-A9 processor running at 1 GHz. It is used for generation of overlay images as well as for streaming the analyzer output video through the wired Ethernet. This is possible, due to one uncommon board feature: it contains both the HDMI output and input interfaces. The board can generate and capture the 1080p30 video. The HDMI output is used for generation of the overlay, that is later composed in the FPGA to an OSD. The HDMI input enables streaming the image being the result of the analysis over the Ethernet.

More information on the design can be found in [1].

IV. FIRMWARE DESIGN

The FPGA firmware of the Image Analyzer is composed of two coupled systems: the control system and the video processing path.

The control system is shown in Fig. 2. It is governed by a Microblaze processor core, coupled with 128 KiB of the BlockRAM memory. This memory stores the processor executable code as well as its variables. The memory is preloaded with machine code during the FPGA boot process. The processor only schedules the transfers. It does not take part in the image processing.
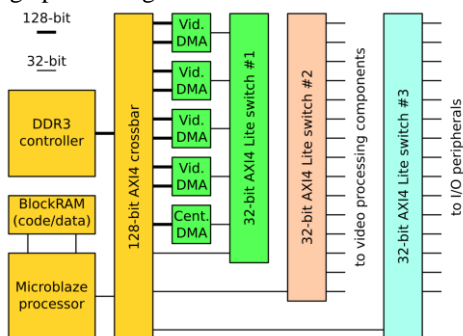
Fig. 2. Block diagram of the Image Analyzer's control system

The control system contains only the Commercial Off-The-Shelf (COTS) IP-cores. The situation is much different for the video processing path. There, the essential video processing components were developed from scratch. The structure of this system is shown in Fig. 3. For the sake of simplicity, the figure does not present the 14 AXI control interfaces connected to these components.

The signal from cameras is supplied through dual SDI interface. The deserialized data streams enter the SDI receiver blocks, where the video data is extracted and presented through an XSVI interface. The XSVI signals are provided to the video cross-switch, which has two purposes: it allows tests with identical video signal in both channels and it converts the video from XSVI to AXIS protocol. Usage of the depreciated XSVI interface comes from the Xilinx SDI input reference design.
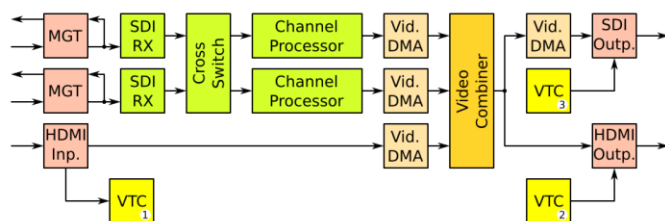
Fig. 3. Block diagram of the Image Analyzer's video processing system

The video from the cross-switch enters the channel processor, performing several monitoring and editing tasks. Next the video stream is consumed by the Xilinx Video DMA (VDMA). This component implements image buffer of three frames with dynamic GenLock synchronization. It offers seamless adaptation between input and output frame rate.

The HDMI input signal, from the SBC, has much simpler path. Firstly, the embedded synchronization data are extracted and then the signal is re-clocked and converted to AXIS standard. At the same time, the synchronization signals are provided to Xilinx Video Timing Controller (VTC) for detection of the resolution and frame rate. Finally, the video stream is provided to the Video DMA.

Signals from all the three Video DMA circuits are provided to a custom video combiner block operating at the frequency of 150 MHz. It is the solution dedicated for performing almost all the analyses that the Image Analyzer is required to provide. One of it functions is to calculate linear combination of the components of pixels from input streams. Moreover, it implements masking for inclusion of the On-Screen Display.

The video from the combiner is directly provided to the HDMI output path and follows its timing, which in turn is provided by the second VTC module. The HDMI stream is then enriched with embedded synchronization information and provided to a dedicated transmitter outside the FPGA. The video from the combiner block is also provided to the VDMA circuit isolating the combiner clock and timing domain from the SDI. The last, third, VTC generates timing for the SDI output.

V. CONCLUSIONS AND PLANS FOR THE FUTURE

The currently implemented set of features is focused mainly on alignment of the optical tract of the rig. Each functionality is tested with the FINN company from Łódź, which ordered the solution. The analyzer is hence routinely used for the calibration of the rig.

The solution is still under development. It is planned that in the future the device should perform scene depth analysis. The Image Analyzer was first publicly demonstrated during the MIXDES 2015 international conference held in Toruń, Poland.

REFERENCES

[1] P. Perek, D. Makowski, A. Mielczarek, A. Napieralski, and P. Sztoch, "Towards Automatic Calibration of Stereoscopic Video Systems", in MIXDES, Toruń, Poland, Jun. 2015, pp. 134–137

[2] A. Mielczarek, P. Perek, D. Makowski, A. Napieralski, and P. Sztoch, "Calibration of Stereoscopic Camera Rigs Using Dedicated Real-Time SDI Video Processor", in MIXDES, Toruń, Poland, Jun. 2015, pp. 138–140.