

VIDEO SIGNAL PROCESSING THROUGH LOW COST FPGA

Dr. R. Prakash Rao

Associate Professor, Electronics and Communication Engineering,
Matrusri Engineering College, #16-1-486, Saidabad, Hyderabad-500059, India.

ABSTRACT: This work depicts the utilization of present day low price Field Programmable Gate Arrays (FPGAs) for real time broadcast video processing. Capabilities of selected device family (Altera Cyclone IV) are discussed with regard to video processing. Example IP core deinterlacer is designed in Verilog HDL and the design flow is described. The IP core is implemented in real hardware system. The overall hardware system is described together with individual FPGA components providing video input/output and other I/O functions.

Keywords: FPGA, video processing, deinterlacing, Verilog HDL, hardware design flow

INTRODUCTION

Before designing the final hardware board to be used in the device, we designed an evaluation platform to test the video processing functions inside the FPGA and the interface chips used to convert the different video transmission standards to and from the FPGA input/output format. The evaluation board included an Cyclone III FPGA with 40k logic elements with the fastest speed grade available (EP3C40F484C6N). The FPGA had two DDR2 memory channels available, each consisting of two 16-bit DDR2 memory chips. The FPGA was connected to SDI input interface chips, DVI (TMDS) receiver, output DVI transmitter, USB Communication Bridge to allow for PC connection and other support ICs. On this board we evaluated the relevant IP cores to be used throughout the project and later developed our own intellectual property components [1].

1.1 Video and Image Processing Suite (VIP)

The first to evaluate were the components from the Altera Video and Image Processing Suite. We were mainly interested in the Deinterlacer and Switch IP cores. The VIP cores can be instantiated either standalone or inside an SOPC system. With both approaches, the user is offered a MegaWizard configuration interface to select the required core functionality. We used the core in the "Bob - Scanline Interpolation" mode. This deinterlacing method adds lines to each half field by calculating the missing odd/even lines of the field. We selected this mode for that this interpolation method produces relatively clean image with no visible artifacts from merging two fields (such as the Weave method does) and since this algorithm uses only a few lines of image to buffer data so it produces very little delay. The visual quality of the processed video feed was found acceptable for the project, but the stability of the IP generation was found unsatisfactory. For visual quality testing, we used the IP core version integrated in the Quartus II package version 9.0. This version performed with no problems. When we switched to the Quartus II version 9.1, we could not compile any design containing the core. When the core was configured and the system was being generated (the parametrization of the core was under way by the configurator), the Quartus IDE crashed and was not able to recover. Since we had to use the 9.1 version (because it included a Switch component which we needed and which was not included in the 9.0 version IP library), we had to abandon using the provided deinterlacer component from Altera and had to develop our own solution. At the time of writing this thesis, when testing the deinterlacer core the compilation runs without any problems [2]. This incident illustrates that the FPGA development toolchain is a rather complex software and should be thoroughly evaluated before considering using it in a design.

1.2 DDR2 High Performance Controller II

We needed some form of large temporary storage memory to store the video data when synchronizing two video streams (frame buffer) and to store the captured image for the still image capture function of the system. We decided to use DDR2 memory because it is the newest DDRx standard electrically supported on the Cyclone II/IV device family architecture. On the evaluation board we had integrated two channels of DDR2 memory channels, each consisting of two 16-bit DDR2 chips. This meant 64 bits effective transfer

size per clock and 128 bits smallest transfer size when considering DDR2 minimal memory side burst size of 4 beats according to the JEDEC specification. Regarding the memory access pattern, we needed to read and write sequential areas of memory and therefore did not need the short memory side burst lengths of DDR I memory, which could be more appropriate for other algorithms such as realtime image rotation. We tested the memory controller core by running the "memtest" example included in the Nios II Embedded Design Suite. The tests passed with no problems and we therefore decided to use this core [3].

1.3 NIOS II soft processor

To control the FPGA hardware an softcore processor was needed. Altera provides the Nios II 32-bit embedded processor for use on its devices. The processor core can be configured into three versions, Economy, Standard and Fast. Since we did not need any video processing functions done on the CPU, we could use the Economy version on the core. The JTAG debugging and communication feature of the Nios II EDS development software proved very handy when debugging the system later in the project.

II. Selected system structure

This section describes the resulting internal FPGA video processor structure, This setup emerged after several design iterations. The structure took shape after considering the project requirements described above. It was necessary to display both the video feed from endoscopic camera and the administrative GUI application running on the internal x86 system. Therefore the FPGA has two video inputs. It is necessary to display the output video, so the FPGA has a video output. We needed to somehow communicate with the PC for system control and captured image transfer. For this reason the FPGA is connected to and USB FIFO bridge. To provide storage space for triple buffering and captured image storage, the FPGA has an DDR2 memory attached. The design was created using the Quartus II IDE. As a top level entity was selected a schematic file to provide a clear way of showing the system structure inside the Quartus project[4]. Compared to a HDL top entity such as Verilog or VHDL file, the schematic quickly shows how the individual blocks are connected and communicates the information to the hardware designer. The block diagram and individual components of the system are discussed below. The components are covered only in brief detail, the three components comprising the core of this work are described in detail in a separate chapter.

2.1 Block diagram

The system block diagram on figure 7.1 shows only the components relevant to the video processing paths. Supporting components like clock domain crossing, external support signals for the PCIe grabber, video resolution detectors, color space conversion cores etc. are not shown to maintain clarity. The system takes two video streams as input, processes them and outputs a single video feed as output. The video inputs are the camera input and the PC video input. The timing of the PC video input is taken as a reference timing, onto which the camera video feed is synchronized and blended using the frame buffer component. The frame buffer is also connected to the USB link providing a way to "dump" the contents of a memory location containing a captured image to a PC.

One video processing path is the camera feed, another is the PC video feed. The camera video has to be synchronized to the PC video signal timing and if in an interlaced format it also must be deinterlaced [5]. Placing the deinterlacer after the frame buffer component saves memory bandwidth, since it allows for buffering of the half fields only and the final full frame is calculated using the deinterlacer after the synchronization phase. This also means that the images transferred to the host x86 system are half fields (for the 1080i interlaced input video format) and have to be stretched to original aspect ratio. It was found that this solution is perfectly acceptable since there is no visible reduction of quality of the captured image.

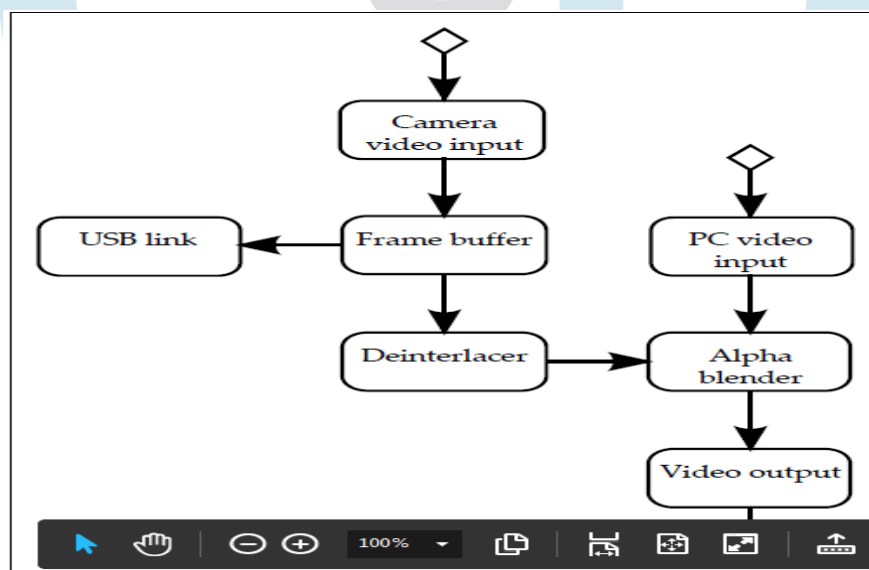


Figure 7.1: Final video processing system block diagram

2.2 Camera video input

The component providing the video input to the system is designed as an Avalon Interface Fabric compatible component. The input parallel video data from the external SDI receiver chip are converted from YCbCr color space into RGB color space using a simple pipelined calculation and then the data are fed into a dualclock FIFO (standard component provided by Altera). This effectively transitions the data from exact-time-formatted data for display into a data stream suitable for internal processing. The remainder of the component is an Avalon Memory Mapped Master externally controller by the Nios II CPU. The master component can be thought of as a DMA engine, which converts a video frame into the DDR2 storage memory using long Avalon side bursts [6]. To relax the frequency requirements for the core logic fabric, the width of the bus from the FIFO to the memory controller IP is set to 64 bits. This effectively halves the frequency at which the bus must run to transfer the data and therefore eases the fitter effort to reach timing closure. This component is displayed as standalone in the block diagram but is effectively part of the frame buffer subsystem.

2.3 Frame buffer

The frame buffer subsystem provides the means to synchronize the camera video feed to that of the PC. Although this introduces delay (of at most one half field), which could be avoided by synchronizing the PC video feed to that of the camera instead, it was supposed that since the PC video feed comes from inside of the device from the host x86 system, it is more reliable and "under control" than the unknown camera signal from outside of the system and is therefore more usable as a reference timing signal. The frame buffer uses a standard triple buffering scheduling algorithm, where one buffer is always available to save an incoming video frame. This provides the means to synchronize the two video streams, since the frame buffer scheduler can either drop or repeat a field to match the required timing [7].

Together with writing the raw image data to the DDR2 memory, the scheduler also registers whether the currently transferred field is odd or even. The scheduler has the field signal available from the external SDI receiver chip. This information is later used to properly configure the deinterlacer block at the output of the frame buffer. The frame buffer then includes an output component which reads a stored field from memory and outputs the data into an input dual clock FIFO of the deinterlacer component. The frame buffer component is described in detail in a separate chapter.

2.4 USB link

The frame buffer subsystem also contains a USB link component on the Avalon Interconnect Fabric. This provides the capability of the system to transfer the stored image data to the PC using a USB FIFO bridge (FT2232H from FTDI[21]). The size of a single half field in 1080i video format is about 4 megabytes and is transferred in under two seconds.

The USB interface IC has two channels, one is configured for the RS232 standard and is used for FPGA system control, the other channel is a one way communication link to the PC for captured still image transfers. The control channel is connected to an UART component of a controlling SOPC system with the Nios II soft core processor.

2.5 Deinterlacer

The deinterlacer component is fed the video data by the frame buffer component, this video data is deinterlaced (if requested) into a full frame and output to the alpha blender component. The deinterlacer core is described in a separate chapter.

2.6 PC video input

The video data from the host x86 system is fed into the FPGA using an external DVI receiver chip. The data passes into the alpha blender, where it is mixed with the video feed from the camera and output from the FPGA into an external DVI transmitter IC.

2.7 Alpha blender

The alpha blender IP core takes the two video streams and mixes them together using an alpha value provided by the Nios II controller system. The alpha value is controllable from the PC over the USB link and therefore allows for video stream switching. The alpha blender includes a transparent color definition. When the alpha blender encounters a pixel with this color in the video data of the host x86 system, the alpha blender displays the original pixel color from the camera feed regardless of the alpha setting. This basically provides the overlay function known from the PC world. This was implemented to allow the system not only to blend the two streams together, but to also enable the on screen display (OSD) generation. The transparent color definition allows for displaying original camera video data with non-transparent OSD mixed on top of this feed. The alpha blender component is also described in more detail in a separate chapter.

III Example video processing core

This section describes the IP cores developed to provide the video processing functions of the system, as required by the project requirements. This core was written using the Verilog HDL language [8]. Compared to VHDL, the Verilog hardware description language was perceived as more readable and "developer friendly". The cores process the video stream data in a stream format - the input components of the processing chain convert the video data from the exact timing format to a stream format, stripping the video data of the synchronization information and forwarding only the active picture data.

3.1 Deinterlacer

Deinterlacing is used to convert from interlaced video format to a progressive one. In interlaced video stream, each complete frame is transferred as two half fields, odd and even. Odd field contains odd picture lines and even field contains even picture lines. By splitting the complete frame into two half fields, the temporal resolution of the video feed is doubled and the motion appears more smooth. Progressive video format transfers frames as complete units, each frame containing all (odd and even) lines of the picture. Progressive video does not have the same temporal resolution as an interlaced video with the same bandwidth, on the other hand it offers better vertical resolution and therefore more detailed image. Interlaced format is commonly used in broadcast applications and TV industry, whereas progressive format is more common in the PC industry [9].

3.1.1 Algorithm overview

A system converting an interlaced video signal to progressive is called a deinterlacer. There are several methods on how to accomplish the conversion:

- Bob - line duplication
- Bob - line interpolation
- Weave - merging of lines of odd and even field
- Motion adaptive algorithms

"Bob" is a name given to algorithms needing only one half field to produce a complete progressive frame. The individual approaches are described below. Line duplication Line duplication algorithm simply takes the input line and produces two lines on output, each same as the image line on deinterlacer input. This is the simplest deinterlacing algorithm, however also the one with the lowest output image quality. Since the half field lines are duplicated, the output progressive image appears pixelated in the vertical direction. This is especially visible on sharp, highly contrasting edges in the image.

Line interpolation:

Line interpolation algorithm does not replicate the missing lines, but instead it calculates the missing line from the line above and below the missing one. This produces a complete frame from a single half field, with the quality of the output image better than the line duplication algorithm. The most visible improvement is that the sharp contrasting edges appear more smooth thanks to the interpolated lines.

Weave algorithm:

The weave algorithm uses two half fields to produce a progressive output frame. The method works by merging the odd and even fields directly into one frame. Compared to the Bob algorithms this method needs a storage memory to temporarily store the half field data. This also introduces a half field delay to the processing chain since the deinterlacer must wait for complete field to produce an output progressive frame. The output quality of this algorithm is compromised by artifacts on edges in the resulting progressive image; since the fields used to produce the output originate in different points in time, when the video feed contains scenes with fast movement, the edges appear distorted since each field captures the moving object in a different position [10]. Motion adaptive algorithm Motion adaptive algorithms try to predict the areas of the image with movement and try to compensate for the motion by calculating the final progressive frame from several preceding half fields of the interlaced input. Additionally to the requirements for storage of the preceding half fields, this algorithm also introduces delay to the video processing chain. This delay depends on the specific motion adaptive algorithm used.

3.1.2 Algorithm selection

After testing the above mentioned algorithms on a development board, we selected the line interpolation algorithm. The quality of the output image was found acceptable, since the edges appear smooth and there are no "saw tooth" artifacts as is the case with the weave algorithm. Also, since this method does not need any preceding half fields to produce an output frame, the latency introduced into the processing chain is very small - typically the duration of a single image line.

3.1.3 Principle of operation

The deinterlacer core uses the line interpolation algorithm to convert the input interlaced video to the progressive output format. The input data in stream format are fed to the core by the frame buffer component. The output of the core is connected to the alpha blender core, where it is mixed with a second video feed and output to the external DVI transmitter chip. The core is reset with the beginning of each input field. After the reset signal is deasserted, the core detects whether the current field is odd or even and also registers the video format resolution as detected by the preceding components of the video processing chain. The core uses two counters, x and y to store the actual position within the video frame. The core has three options as to what to do with each processed line:

- A - store the incoming line into the temporary buffer, and, at the same time, output the line
 - B - store the incoming line into the temporary buffer, and, at the same time, output the average (interpolation) of the line being currently stored in the temporary buffer and the line already saved in the temporary buffer
 - C - do not store anything, just output the line already stored in the temporary buffer
- The decision between performing the action A, B or C is made by the core scheduler. This part of the deinterlacer keeps the current position in the video image and performs configuration of the remaining parts of the component at the beginning of each image line.

3.1.4 Implementation

The core is implemented as a schematic file instantiating the subentities designed in Verilog hardware description language. The core also uses Altera specific components included in the Quartus IP library. The core has three main groups of virtual I/O pins exported to the higher level design file - video stream input, video stream output and control signals. Video stream input pin group is composed of signals fifo data input[63..0], fifo rdempty input and fifo rdreq output. These signals form an interface to the FIFO data buffer of the frame buffer component. Video stream output pin group is composed of pins out rdreq, out rdclk and out data[23..0]. These signals provide the interface to the alpha blender component mixing the two streams to form the output video signal. The remaining signals form the deinterlacer core control signals. The main signals in this group are clock, reset, video resolution and the deinterlacing enable signal to optionally disable the deinterlacing to let progressive video format pass through unchanged for the 720p progressive input camera video format. The core scheduler is located in the deinterlacer controller submodule. This module controls the state transitions at the beginning of each image line as described in the previous chapter. The core scheduler algorithm is the following:

```

if ( can advance == 1 )
begin
x = x + 1 ;
if ( x == x size )
begin
x = 0 ;
y = y + 1 ;
if ( field == 0 )
begin
if ( y == 0 ) master state = 2 ;
if ( ( y >= 1 ) & ( y < ( y size minus one ) ) )
master state [ 1 : 0 ] = fl ' b0 , y [ 0 ] g ;
if ( y == ( y size minus one ) ) master state = 0 ;

```



```

end
else
begin
if (y == 0) masterstate = 2;
if (y == 1) masterstate = 0;
if ((y >= 2) & (y < ysize))
masterstate[1:0] = fl'b0, ~y[0]g;
end
end
end

```

The variables x , y contain the actual position within the video image data, field is the even/odd field indicator, master state[1:0] is a variable indicating which of the actions A, B or C should the deinterlacer perform on the actual line and can advance is a signal indicating that the remaining core components are ready for next data item. The deinterlacer ram buffer module is Altera-specific instantiation of an embedded memory block forming a RAM memory to store the image line. The address to this embedded RAM memory block is controlled by the scheduler, the deinterlacer mem addr delay module delays the address signals for the line operation B. The operation B means that the deinterlacer must store the incoming line to the RAM buffer and at the same time load the data from the very same memory buffer. Therefore, it is necessary that the data from the buffer can be read out before the new image line data are saved to the buffer. The deinterlacer line switch module provides the switching between operations A, B and C as requested by the scheduler module. Operation A (master state = 2) means that the data received from the frame buffer component is stored to the RAM buffer and at the same time the data is routed through the deinterlacer line switch to the output FIFO. Operation B (master state = 1) means that the incoming data is stored to the RAM buffer and at the same time the previous line data stored in the RAM buffer are read out, sent to the deinterlacer line switch where the pixel data is averaged (interpolated) with the actual line data and sent to the output FIFO. Operation C (master state = 0) does not read the incoming pixel data but instead simply outputs the stored line from the RAM buffer to the output FIFO.

The remaining components of the deinterlacer core are mainly support functions to properly align the individual data and control signals to compensate for the latency of the respective communicating components. To relax the requirements for the maximum frequency of the device logic fabric, the deinterlacer core processes two pixels at a time. This doubles the used data bus width, but at the same time allows to halve the operating frequency while maintaining the required bus bandwidth. The deinterlacer core expects the field data in a standard RGB color space with every color component having 8 bit value range (0.. 255).

The interpolation (vertical averaging) of the neighboring half field image lines is done by adding the individual red, green and blue components of the pixel color (the two pixels in the RAM buffer from the previous image line and the two pixels currently being received and stored to the RAM buffer) together and then doing an one bit position shift right, thereby calculating an arithmetic average of the two values.

IV FPGA design flow

This section does not describe the FPGA design flow in the traditional sense- design entry, synthesis, fitting phase, timing analysis, programming - since there is plenty of information available on this topic elsewhere. Instead, it tries to provide a real world experiences and hopefully useful tips for designing an embedded system containing an FPGA. This section focuses mainly on the Altera Quartus II development software, but the principles can be used for design flow with any FPGA device vendor.

4.1 Separate projects for custom components

When developing a FPGA system with custom designed IP components, develop the custom cores as a standalone project (where possible). This allows the designer to focus on the component functionality rather than on the interactions with the rest of the system, reduces the design iteration compilation time and allows for component reusability.

4.2 Use standard interfaces

The system design can be greatly simplified by using standard interfaces wherever possible. A good example is the Avalon Interface used in SoPC designs using Altera devices. By designing the component interface according to a standard, the designer can then use the component together with IP cores developed by third parties. This saves development time and provides greater flexibility than by developing a custom interface [11].

4.3 Optimize the memory access pattern

When the design uses some form of DDRx memory for data storage, think about and optimize the memory access pattern. The most common caveats are not using burst transfers for sequential data access and inefficiencies in bus width adapters. For example, the Nios II softcore processor (at least the current version) accesses memory in 32-bit words. When the memory controller core has a different local side width, this may result in an inefficient transfers since e.g. half of the data word is discarded. This may hold true for other components of the Avalon fabric as well.

The memory access pattern is an important design decision as well. As described in the previous chapter, the differences in memory bandwidth between bursting and non-bursting transfers can be huge. There is plenty of information available on how to optimize the access patterns for a given task.

4.4 SignalTap II logic analyzer

A very useful feature of the Quartus design environment is the SignalTap logic analyzer. When included into the project, the SignalTap component creates a logic analyzer from unused device resources and allows the designer to see the system behavior in real time. This is especially useful for debugging, where by setting the appropriate trigger conditions the designer can observe the real behavior of problematic parts of the design.

4.5 Horizontal and vertical device migration

When designing a particular FPGA system, it is useful to include the option to migrate to a different device or even to a different device family with a single PCB layout. This adds more robustness to the project, because the hardware board is not limited to a single device and thus the system can be assembled even if the original device is not available. This mostly concerns the physical I/O pin mappings, where the designer has to do an "intersect" of the I/O pins of the devices considered for migration. There is also an added flexibility should the design requirements change - device with more or less resources can be used instead of the originally planned one. For example, with the recent earthquake and subsequent tsunami in Japan, the Altera supply chain was damaged and the target device for the project (Cyclone IV, 15k LEs, FBGA256) was not available from any major electronic components distributor. By designing the project for two device families (Cyclone III and IV), the system was assembled using Cyclone III family device (Cyclone III, 16k LEs, FBGA256) and the project schedule was not affected.

4.6 Physical I/O pin mapping

When mapping the design I/O pins to a specific device, the designer should work closely with the layouter (the person designing the printed circuit board where the FPGA chip is placed). By placing the design pins on optimal I/O positions of the device, the FPGA designer can reduce the complexity of the printed circuit board and therefore decrease the layout time and the cost of the PCB by reducing the number of required layers.

V. Resulting hardware

After the evaluation of the design on an development board a final hardware board was developed for the FPGA video processor. Based on the design requirements a target device was selected to fit the design. The device selected from the Cyclone IV family was EP4CE15F17C7N with the option to migrate to an older Cyclone III generation with a device EP3C16F256C7N. These devices have very similar pin out and therefore can be interchanged with no major problems.

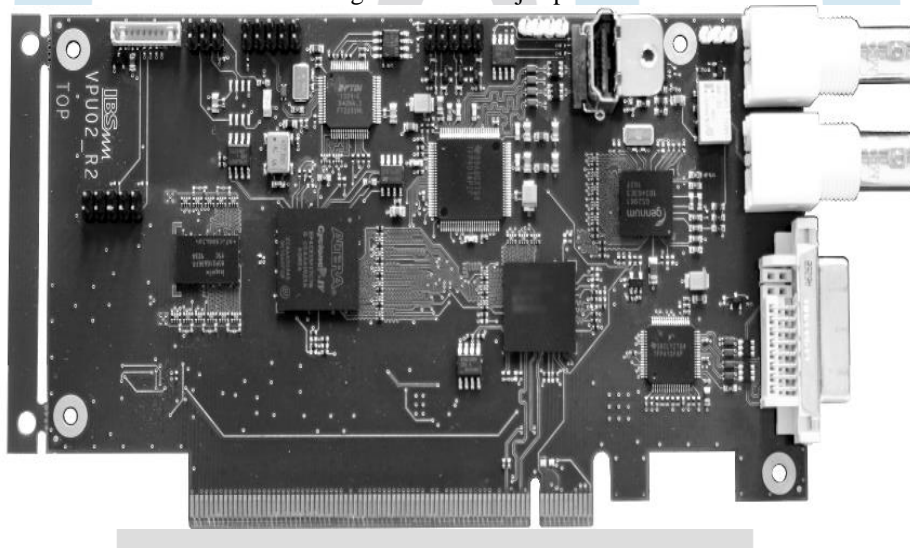


Figure 10.1: Final hardware board containing the FPGA video processor

The entire system resource usage (after optimizations done by the synthesis and fitter compilation phases) is 13565 logic elements (88% usage), 313072 memory bits (61% of the device resources) and 151 pins (91%). An illustration of the final pin usage can be seen in the appendix section to this work. Of this resource usage the deinterlacer block uses 501 LEs, alpha blender uses 228 and the frame buffer component uses 8269 LEs. The high resource usage of the frame buffer is mostly caused by the DDR2 memory controller core (6304 LEs) and Avalon bus arbitration (776 LEs). The reader and writer components use about 400 logic elements each. The DDR2 x16 memory is connected to banks 3 and 4, the other I/O pins are mostly used to connect to the parallel video data buses of the two video input ICs and one video output chip.

5.1 Verification of the hardware

The design was verified by several methods. The developed cores were simulated using (at that time) the Simulator utility integrated in the Quartus software. The designs were simulated using functional simulation to verify the correct functionality of the module. After successful simulation the cores were tested in a real system to validate correct function. The behavior of the core under test was again checked by the SignalTap II logic analyzer. After these stages the entire FPGA system was developed, placed on a real hardware board, assembled into a final prototype and this prototype was thoroughly tested. Visual image quality was checked,

communication with the PC and overall device stability were evaluated. The prototype also underwent thermal stress testing to validate the functionality in the entire operating temperature range as required by the design specification.

6. Conclusion

This work presented the overview of the development process of a hardware device using Field Programmable Gate Array and provided some example IP component for video processing to illustrate the possibilities and approaches used for said development. The capabilities of the selected low cost FPGA family from Altera was found to satisfy the requirements for processing power with no major problems. The flexibility of the FPGA architecture proved very useful since there were many minor modifications to the system behavior throughout the development of the project. Although the per-chip price of FPGA is still higher than that of the ASIC, for small production series this difference is balanced by the flexibility and the ability to customize the design to the project specific requirements.

REFERENCES

- [1] Ramachandran, S.: 'Digital VLSI System Design', New York: Springer, chapter 11, 2007.
- [2] Hammami, O., Wang, Z., Fresse, V., and Houzet, D.: 'A case study: Quantitative evaluation of Cbased high-level synthesis systems', EURASIP Journal on Embedded Systems, 2008.
- [3] Glasser, M.: 'Open Verification Methodology Cookbook', New York: Springer , chapter 1-3, 2009.
- [4] Man, K. L.: 'An overview of SystemCFL', Research in Microelectronics and Electronics, 2005, 1, pp. 145-148.
- [5] Hatami, N., Ghofrani, A., Prinetto, P., and Navabi, Z.: 'TLM 2.0 simple sockets synthesis to RTL', International Conference on Design & Technology of Integrated Systems in Nanoscale Era, 2000, 1, pp. 232-235.
- [6] Chen, W.: 'The VLSI Handbook'. 2nd edt., Boca Raton: CRC Press LCC, chapter 86, 2007.
- [7] Berkeley Design Technology, "An independent evaluation of high-level synthesis tools for Xilinx FPGAs", http://www.bdti.com/MyBDTI/pubs/Xilinx_hlstep.pdf.
- [8] Haastregt, S. V., and Kienhuis, B.: 'Automated synthesis of streaming C applications to process networks in hardware', Design Automation & Test in Europe, 2009, pp. 890 - 893.
- [9] Avss, P., Prasant, S., and Jain, R.: 'Virtual prototyping increases productivity - A case study', IEEE International Symposium on VLSI Design, Automation and Test, 2009, pp. 96-101.
- [10] He, W., and Yuan, K.: 'An improved canny edge detector and its realization on FPGA', in Proceedings of the 7th World Congress on Intelligent Control and Automation, 2008.
- [11] Gonzales, R. C., and Woods, R. E., 'Digital Image Processing', 3rd edt. New Jersey: Prentice Hall, 2007.