

12-1-2005

Real-Time 3D Image Visualization System for Digital Video on a Single Chip

Nader I. Rafla
Boise State University

Real-time 3D Image Visualization System for Digital Video on a Single Chip

Nader I. Rafla, Ph.D., PE,
Electrical and Computer Engineering Department
Boise state University
Boise, ID 83725

Abstract - Implementation of a real-time image visualization system on a reconfigurable chip (FPGA) is proposed. The system utilizes an innovative stereoscopic image capture, processing and visualization technique. Implementation is done as a two stage process. In the first stage, the stereo pair is captured using two image sensors. The captured images are then synchronized and sent to the second stage for fusion. A controller module is developed, designed, and placed on the FPGA for this purpose. The second stage is used for reconstruction and visualization of the 3D image. An innovative technique employing dual-processor architecture on the same single FPGA is developed for this purpose. The whole system is placed on a single PCB resulting in a fast processing time and the ability to view 3D video in real-time. The system is simulated, implemented, and tested on real images. Results show that this system is a low cost solution for efficient 3D video visualization using a single chip.

1. INTRODUCTION

Reconstruction of high quality real-time 3D video images is growing rapidly with the advanced development of digital image sensors and computer graphics. As a result, a great effort is placed into two aspects of this process: stereoscopic image sources used for image capturing [1]; image processing, fusion and presentation of 3D images [2]. Such systems are used in a wide variety of personal and social human activities such as: entertainment, education, sports, medicine and much more.

Several research groups developed systems for reconstruction of 3D video images [3-7]. All these systems use a group of distributed cameras to capture multi-view images of the same scene and process them for reconstruction of 3D images suitable for viewing. This process requires the use of a PC cluster or a highly expensive computational environment that makes them unusable in everyday life [8].

To reduce the cost and size of such systems and make them usable, the following problems must be addressed and taken into consideration:

- *Number of views:* The proposed system limits the number of views to only two. These views are captured using CMOS image sensors that are placed on a PCB. Additional logic placed on an FPGA is developed for routing and clocking. This setup is compact in size and

has the advantage of low power consumption. These sensor heads are used to capture the stereoscopic pair.

- *Computational speed:* Using an advanced FPGA architecture with two embedded processors allows both distributed and parallel processing. This rich computational setting eliminates the need of a PC cluster computing environment.
- *Image visualization:* 3D image visualization require the use of a 4D space, 3D geometric and 1D temporal [5]. Instead, 3D images can be viewed by proper alignment and displaying of red and blue/green image pairs.

2. Overall system design

The system described here uses an FPGA to synchronize stereo-pair two-dimensional digital images captured by two CMOS image heads. The image processing and fusion algorithm uses the red channel from one image and the blue/green channels from the other image. Data is then transferred to be displayed in the form of a three-dimensional image. The resulting image is visible by using anaglyphic glasses.

The system required a reconfigurable chip that is capable of interfacing and controlling the image sensors in addition to processing the captured images. A Xilinx Virtex-II Pro FPGA [9] is selected since it contains two build-in hardware multiprocessor cores (PowerPCs) among its resources.

The whole system is broken into several modules to facilitate the distributed and parallel processing nature of the image processing and fusion, as well as the interface and control required for the image sensor heads. Those modules that run on the PowerPCs are developed in a high level language, while the interface modules are written in VHDL and Verilog in the form of RTL level design. Inner communication among these modules is facilitated via the use of on-chip Wishbone-based bus architecture [10]. An interface to the image display is also included in the design. A block diagram of the system architecture is shown in Figure 1.

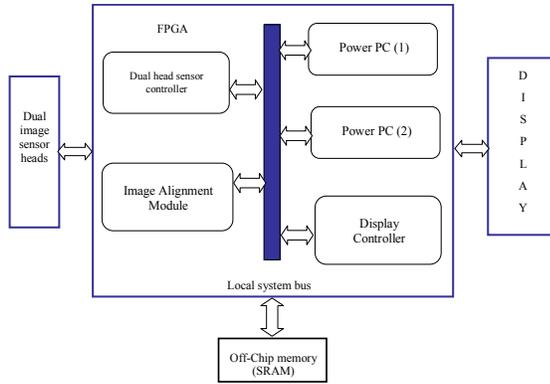


Fig. 1. Visualization System Architecture

3. IMAGE CAPTURE

Stereo 3D imaging systems use, at least, two CCD cameras to synchronously capture images from two perspectives. Objects in these two images are registered using computer algorithms to generate 3D image that contains depth information of the objects in the field of view of both cameras. The accuracy of stereo 3D imaging is usually limited by the rate at which the imaging system can capture images.

Currently, there exist several types of stereo-based cameras. These cameras are not suitable for our application in terms of size, cost, flexibility, or ease of control [11]. These image capture systems use two digital cameras mounted on a common physical support to form a 3D camera as shown in figure 2. The advantage of this configuration is that putting two cameras together on a rail is relatively simple. In the mean time, the relative position of the two cameras can be easily changed. The major problem of this set-up is the difficulty of obtaining simultaneous image capture from the two cameras to form a moving picture.



Fig. 2. Two digital cameras on a single mount

Therefore, a need arises to design a flexible image capture system that allows continuous capturing of images in real-time with an adequate storage and pre-processing capability.

The approach taken here is to integrate two image sensor heads onto a common board that is interfaced to an FPGA as shown in Figure 3.

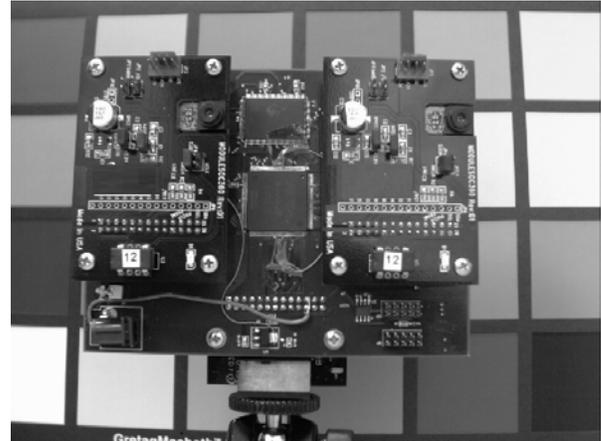


Fig. 3. Right/left image sensor heads

This sensor interface module communicates with the sensor head setup registers to control the image capture and data flow. This approach allows capturing the two images sequentially in time. The first captured image is stored temporary in an SRAM until the image from the second sensor head is captured. The two images are then delivered to the image synchronizer module simultaneously. Switching between the two sensors is accomplished via an I2C link. Hence, an I2C controller is implemented and included within this module. This allowed controlled switching between the left/right sensors for more accurate implementation of the stereoscopic image synchronizer. Figure 4 shows a block diagram of this part of the design

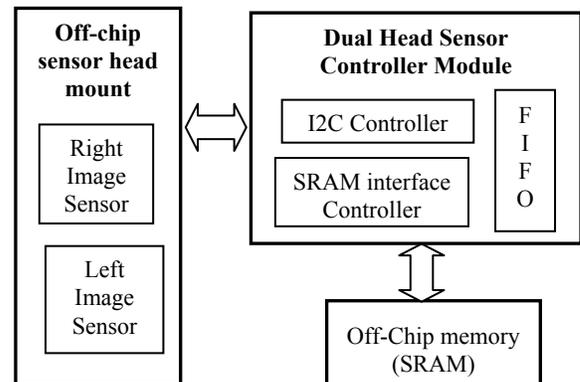


Fig. 4. Image capture block diagram

4. STEREOSCOPIC IMAGE SYNCHRONIZATION

Each related pair of images received from the two image sensors represents a 3D image, if synchronized in time. Practically, there exists a time delay between capturing image frames. In order to maintain image synchronization between the two related frames, another module is implemented. The finite state machine (FSM) designed for this module keeps track of the received frames and their sources, temporarily store them, and send them to the next processing stage. This state machine accounted for the fact that the frames are not synchronized and resulted in a stable mechanism for capturing the data. Table 1 summarizes the developed algorithm for synchronization.

This algorithm is implemented in hardware and placed on the same FPGA that contained the dual head sensor controller shown in figure 4. A FIFO is also designed and placed on the same chip. It is used as a temporary storage of the first frame of the stereo-pair. Due to the fast access time of the FIFO and its proximity to the image source, it was possible to achieve fast and accurate frame synchronization.

Table 1. Algorithm for image synchronization

1. Wait for a frame start bit from an image sensor
2. Tag the frame with a bit value that indicates which sensor is sending the frame
3. Temporary store the frame in a FIFO
4. Wait for another frame to start
5. If the frame is from the same sensor, store it in the SRAM. Otherwise tag it with a different bit value and store it in the FIFO
6. After the reception of two correlated frames, i.e. from the two sensors, is completed, send the two frames from the FIFO to the embedded processor for further processing

5. IMAGE RECONSTRUCTION AND VISUALIZATION

This part of the system is implemented using the on-chip PowerPC. A proprietary software program is written to extract the red channel from one image and the blue/green channels from the other image. This was based on the fact that the left image does not contain any blue or green data, and the right image does not contain red data. As soon as the three color channels are extracted, they are combined as an RGB signal and sent to the display for viewing. Another module is developed to dynamically adjust the alignment of the viewed images. The module was able to align 3D images for live viewing at a modest frame rate of 7.5 pairs per second.

6. INTEGRATED SYSTEM

The different components of the system are tested individually through simulation. After verifying the simulation results, the system components were integrated in a form of prototype and tested on real images. A PCB was then designed to accommodate the two image sensor heads as well as the FPGA. A special attention is paid to minimize the physical dimensions. The final PCB layout, all layer view, is shown in figure 5. Note the dimension information that is included

7. SYSTEM VERIFICATION

Although testing of image devices is more visual than technical, an innovative technique of verification of the integrated system was done to prove the concept and visualize the correct image out of the left/right pair after processing. That was achieved by simply covering one of the sensors and observing the resulted images. The following procedure is followed to visually verify the results:

- Cover the right image sensor leaving only the left sensor as the source of the image capture
- An all red image is observed, that indicated that the red channel was extracted correctly
- Cover the left image sensor. The image is only captured using the right sensor
- A blue/green image was viewed on the display indicating that the left sensor image is captured and processed correctly.

8. CONCLUSIONS

An integrated system to capture, process, and visualize live three-dimensional digital video images was designed and implemented on a single reconfigurable FPGA device. The system includes two image sensor heads mounted on the same PCB that contains the FPGA and a display to visualize the live 3D video. The design made full use of the FPGA resources and its embedded processors as a distributed and parallel processing system. Software was written and implement on the embedded PowerPC. Several other modules were designed and communication was achieved using Wishbone-based local bus architecture. The integrated system was built, verified and live video was observed on the display. This system is robust, compact in size and inexpensive. It has several applications in entertainment, education, sports, and medicine.

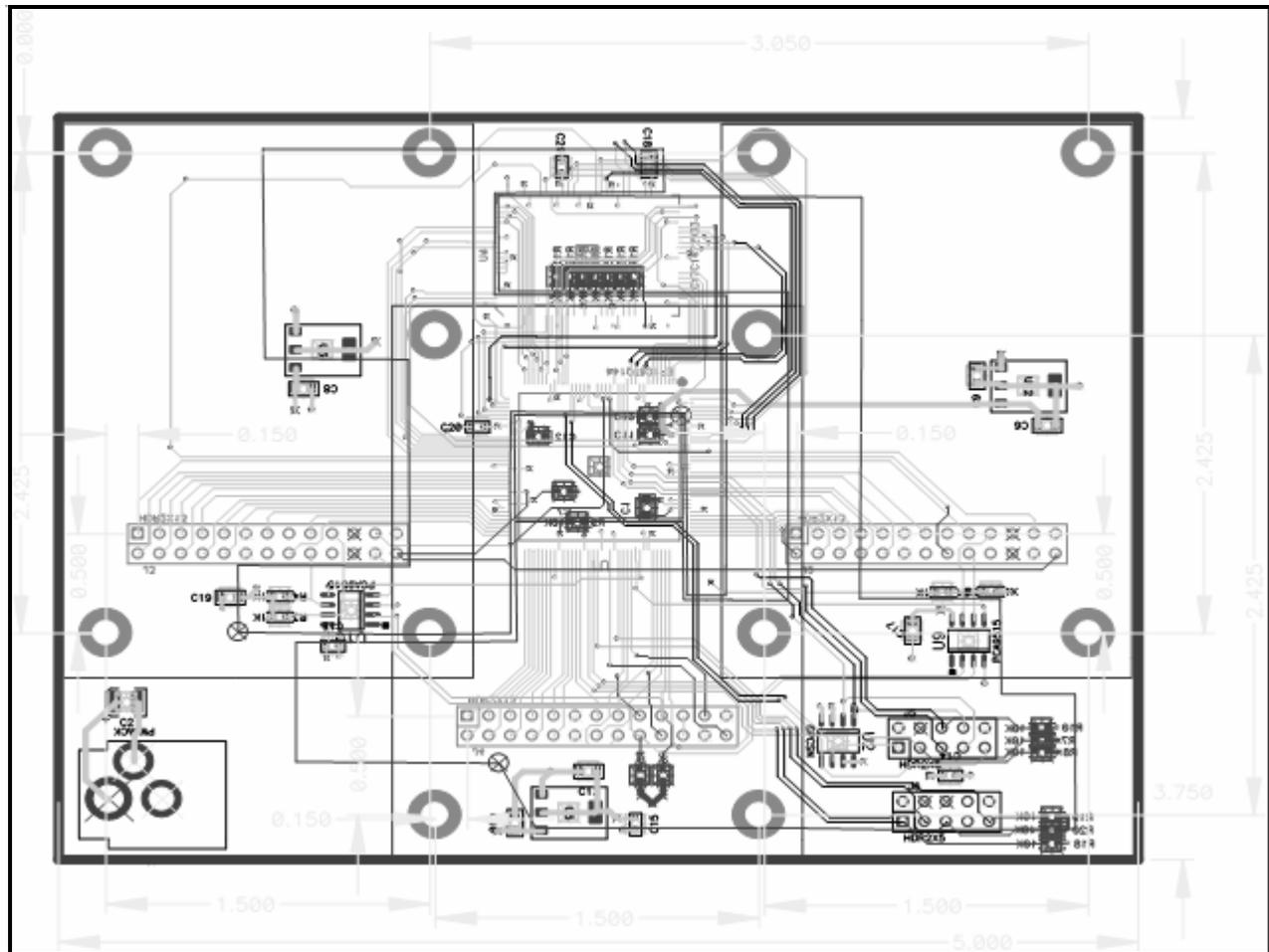


Fig. 5. PCB Layout

9. REFERENCES

- [1] *Image sensors market growth shapes to volume supply*, III-Vs Review, Volume 16, Issue 8, November 2003.
- [2] T. Matsuyama, X. Wu, T. Takai, S. Nobuhara, "Real-time generation and high quality visualization of 3D video", *Proc. Mirage, France, 2003 pp. 1–10*.
- [3] T. Kanade, P. Rander, P.J. Narayanan, "Virtualized reality: constructing virtual worlds from real scenes", *IEEE Multimedia*, 1997, pp. 34–47.
- [4] G. Cheung, T. Kanade, "A real time system for robust 3D voxel reconstruction of human motions", *Proc. of Computer Vision and Pattern Recognition*, South Carolina, USA, 2000, pp. 714–720.
- [5] T. Wada, X. Wu, S. Tokai, T. Matsuyama, "Homography based parallel volume intersection: toward real-time reconstruction using active camera", *Proc. of Internat. Workshop on Computer Architectures for Machine Perception*, Padova, Italy, 2003.
- [6] J. Carranza, C. Theobalt, M.A. Magnor, H. Seidel, "Free-viewpoint video of human actors", *ACM Trans on Computer Graphics Vol. 2, no.3*, pp. 569–577, July 2003.
- [7] Li Mingi, M. Marcus, H. Seidel. "Hardware-Accelerated Visual Hull Rendering. *Vision, Modeling, and Visualization Workshop*, Munich, Germany, 2003.
- [8] T. Matsuyama*, X. Wu, T. Takai, S. Nobuhara, "Real-time 3D shape reconstruction, dynamic 3D mesh deformation, and high fidelity visualization for 3D video", *Computer Vision and Image Understanding*, July 2004.
- [9] Vertex-IIPro Device Family Data Sheet, *ds083.pdf*, Xilinx Inc., <http://www.xilinx.com>
- [10] WISHBONE System-on-Chip (SoC) Interconnect Architecture, <http://www.opencores.org/projects.cgi/web/wishbone/>
- [11] PokeScope® Stereoscopic Resources, 3D cameras, <http://www.pokescope.co>