

Image Acquisition for VLSI Face Detection System

Lim Yu Ling¹, Siti Hawa Ruslan^{1*}

¹ Faculty of Electrical and Electronic Engineering
Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor,
MALAYSIA

DOI: <https://doi.org/10.30880/eeee.2020.01.01.020>

Received 5 August 2020; Accepted 6 September 2020; Available online 30 October 2020

Abstract: Face detection is the process to identify the face biometrics of people while image acquisition is the fundamental of face detection. Image acquisition system is a series process in capturing and loading image to Field-Programmable Gate Array (FPGA) board. This study implemented Viola-Jones face detection algorithm into DE2-115 board using TRDB-D5M as camera to acquire image. The captured image is segmented and stored to allocated memory address for image processing and external monitor is used to display the output. The first part of this study is using OV7670 camera and DE2-115 board to collect data for face detection in Very High Speed Integrated Circuit Hardware Description (VHDL) language while the second part is using TRDB-D5M and DE2-115 to test the function of TRDB-D5M camera in Verilog language. Then the last part of this study is the implementation of the TRDB-5M camera into the face detection system in VHDL language. The result obtained had shown that with OV7670 camera as the input, the acquisition of image is successfully done, stored and displayed on the monitor. However, the camera resolution is very low and the displayed images are very small in size. In enhancing the face detection system, TRDB-D5M camera is used to replace the OV7670 camera. The new setup is successful in displaying stream images showing that the implemented segmentation of the acquired image to the allocated memory is success, but for the still image only the default image is displayed.

Keywords: Face Detection, FPGA, Image Acquisition

1. Introduction

Face detection is a process to identify and locate the presence of faces from an image or video stream with the use of a computer by identifying the face biometrics. The face detection system is the pre-process of face recognition or verification of face. This embedded system majority is used for security purpose, in those areas that require verification to allow only appointed person to pass through the high security protected area and also the use of surveillance cameras. Furthermore, the face detection system is also commercialised for mobile phone usage such as to unlock the screen, applications of detect the face and process it to show the gender and age of the face, or it also can be used to swap the face of two-person. All of these applications are coming from the fundamental of the face detection system with the image acquisition that begins the processing.

An image acquisition system is a series of process in capturing and loading image before the image is used for face detection. It is a process of using electronic imaging sensor devices to capture a scene which is well known as a camera to convert analog image signal into digital image data for storing [1]. The image acquired is then used to detect a face. However, most of the existing face detection engines are run by software. The processing speed can be improved by implementing it using hardware such as using Field-Programmable Gate Array (FPGA) development board that has its own on-chip memory and Synchronous Dynamic Random Access Memory (SDRAM) to store the data for later process [2]. This implementation can speed up the image processing due to the delay of obtaining the image or video.

2. Methodology

The methods and procedures that are used to develop image acquisition hardware and software are described in this section. There will be two parts which are hardware implementation and software implementation respectively. Figure 1 depicts the whole system.

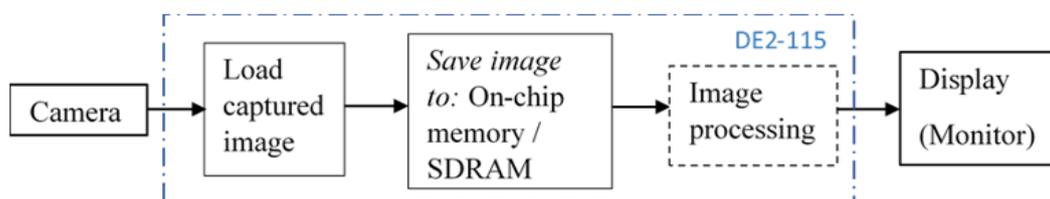


Figure 1: Block diagram of overall system

2.1 Hardware implementation

This section is divided into three parts which are input, FPGA and output. The input for this system is an image sensor which is the OV7670 and TRDB-D5M camera module. Then DE2-115 FPGA development kit is used to process the program and acts as a controller. After that, the image is displayed on a Video Graphics Array (VGA) monitor screen as output.

The input to the system being developed is coming from camera. A Complementary Metal Oxide Semiconductor (CMOS) image sensor implemented with advanced technology will provide high resolution and low noises images [3]. CMOS camera also provides high quality video with huge pixels (millions). OV7670 is a CMOS camera module that has a frame rate of 30 frame per second (fps) for a video stream with a frame size of 320*240 pixels, which is equivalent to 0.3 megapixels. OV7670 is suitable with 3.3 V power supply.

Another camera used is TRDB-D5M. It is a 5 megapixels CMOS sensor with a programmable frame rate that up to 70 fps for VGA display with a frame size of 640*480 pixels. The input voltage supply of TRDB-D5M is also 3.3 V. It has a colour filter array that filter the output image as Bayer pattern format.

The Terasic DE2-115 development and education board was chosen as the FPGA development kit in this study. This development kit consists of Altera Cyclone IV 4CE115 FPGA device which offer a low power consumption and high volume applications. DE2 board also consists of two 64MB SDRAM that is suitable to be used as memory architecture for the system being developed. The control switches, push-buttons and Light Emitting Diodes (LEDs) light that are built-in in the board are used for controlling purpose. This FPGA board also have a VGA digital analog converter (DAC) socket that enables the connection of the DE2 board with an external monitor to display the result with the aid of AV7123 (onboard digital to analog converter). Universal Serial Bus (USB) Blaster of the board provided a platform connection for programming. The general-purpose input/output (GPIO) of the

board is used to connect the camera module as image sensor input so that the image can be processed on the board before being displayed on the monitor.

For the displaying section, a VGA cable is used to connect an external monitor screen to the FPGA. ADV7123 converts the 24 bits Red Green Blue (RGB) (8 bits per colour) formatted data into analog RGB before displaying the outcome on the screen. The implementation of this connection provides 320*240 images on 640*480 VGA monitor.

2.2 Software implementation

In this study, Quartus Prime Lite version 15.1 software is used as the platform for Very High Speed Integrated Circuit Hardware Description Language (VHDL) code synthetization and programmed into the FPGA [5]. The Altera Quartus II application is the software which provides multiplatform design environment that can be adapted in various specific design needs. The code may be written in Verilog or VHDL depends on the designer requirement.

The first part of this study is using OV7670 as camera and DE2-115 board. The system architecture consist of a camera driver, VGA driver and image buffer. There is a reset button for the FPGA board which is switch 17 (SW17). The reset button is used to reset the whole board when needed.

The design of the camera driver consists of two components which are OV7670 controller and OV7670 capture. OV7670 controller is used to control the Serial Camera Control Bus (SCCB) interface input to the camera module with the interface to the development board. The OV7670 capture module received data from the camera module with a synchronous clock pulse of 30 Hz and the image size of 320*240. The output of the capture module is in a data format of 12-bit RGB colour (4:4:4) and then it is sent to image buffer with a 17 bits address.

The main function of the image buffer is to store the captured frames from the camera module and displayed the detected face on the VGA monitor. The image is then processed through Viola-Jones algorithm face detection and the output display will have a red box frame which is mentioned as a Facebox to indicate the detected face.

For the VGA driver, there are also two component parts included which are RGB module and VGA module. The VGA driver passes the signal through ADV7123 to convert the digital data into analog signal so that the image can be displayed through a VGA cable to an external monitor. RGB module converts the process images from grey level to 24 bits RGB (8:8:8) and drive into ADV7123. There are also another horizontal synchronization (Hsync) and vertical synchronization (Vsync) signals which are directly sent from the VGA module to VGA display.

The second part of this study is using TRDB-D5M camera that is connected to DE2-115 board to test the features of TRDB-D5M by referring to the provided TRDB-D5M system Compact disc (CD) [5]. This camera has higher resolution which is 5 megapixels, thus it is much more better than OV7670. It also having features of brightness adjustment, enlargement of capture image and snapshot mode.

The CCD_Capture and I2C_CCD_Config blocks are parts of the camera driver of TRDB-D5M. I2C_CCD_Config has a two-wire serial interface (serial data and serial clock) that are connected between the D5M camera and FPGA board. It is also responsible to respond to the specified switch and key of FPGA board for mode changes. The CCD_Capture module received 12-bits input pixel data in raw format from the CMOS sensor. Then the 12-bits raw data is sent to RAW2RGB module for Bayer to RGB conversion so that the output of image data is in RGB with 12-bits data for each colour.

SDRAM module in TRDB-D5M acts as a frame buffer that stores the data of input image that is converted by RAW2RGB module. There are a total of two SDRAM devices to store the three outputs from RAW2RGB. The [11:7] oGreen data is stored into the first SDRAM with [11:2] of oBlue while

[6:2] of oGreen and [11:2] of oRed are stored into the other SDRAM devices [6]. Therefore, the Bayer colour pattern data is saved to 30-bit RGB when delivered to VGA_Controller. Then the VGA driver passes the input signal getting from SDRAM through ADV7123 to convert the digital data into analog signal so that the image can be displayed through a VGA cable to an external monitor. There are also another horizontal synchronization (oVGA_H_SYNC) and vertical synchronization (oVGA_V_SYNC) signals which are directly sent from the VGA_Controller to VGA display.

The third part is the implementation of TRDB-D5M to the face detection system. It starts with a camera to capture image. The initial camera used for this study is OV7670 which has low resolution (0.3 megapixel) and the outcome is dependent on the brightness of surrounding. TRDB-D5M is a 5-megapixel camera with brightness mode and able to zoom in and out which should perform a better result.

The working principle of the implemented design is the D5M camera as the input data and the data will go through image processing by passing image buffer. After the image had been processed through the Viola-Jones algorithm, the system will detect the availability of face and frame the face with a red box after detecting the face. The monitor is displaying the input of camera in the real-time and red box is located as the face detected.

3. Results and Discussion

Both cameras (OV7670 and TRDB-D5M) had been used in this study. The software compilation are executed using Quartus 15.1 by using VHDL and Verilog language in respective cases. The study code is then programmed into FPGA board for image storing and the input images that are captured from camera module are then processed before displayed on the monitor screen.

3.1 Performance of OV7670

In terms of software implementation, the full design using VHDL code is compiled successfully. From the report of Quartus software, the flow status is successful which supported that the coding is running well and is ready to be programmed into FPGA board. The VHDL code is successfully run into FPGA board through USB-Blaster. The mode selected for this study is Joint Test Action Group (JTAG) which is for testing purpose and the program is directly downloaded into the Cyclone IV E FPGA.

As for hardware implementation, the overall connection of the system is also successfully developed as shown in Figure 2(a). The external monitor screen is connected with a VGA cable from FPGA to monitor while the OV7670 camera module is connected to GPIO of DE2-115 board. The monitor is displaying a video stream of what the camera module has captured. There is a red frame that represented that the face is detected. When the switch 15 (SW[15]) of DE2-115 is switched to high, the camera module is changed to capture mode and the screen display is displaying a static picture. SW[17] of DE2-115 board is assigned as the reset button. The two red LEDs are indicated as phase-locked loops (PLL) lock [LEDR0] and written of camera register [LEDR1] which represented that the camera module is functioning well.

The output is displayed on a monitor through a VGA cable. The display screen is 640*480 while the processed image is in the size of 320*240. Therefore, the display result is on the top left of the screen. The image acquisition from the camera module is successfully achieved and the output monitor is able to display the acquired image that had gone through segmentation at allocated memory. All of the samples are taken in the same location with the similar brightness of surrounding. However, the detection of the face is not stable and not 100% accurate. The sample images are displayed on a mobile phone and OV7670 camera module are used to acquire the image as input for face detection. Most of the outcome had a red frame for each. The low resolution of captured image and the skin tone of the sample face will affect the result.

OV7670 has 0.3 megapixels input data that is considered as low resolution and the captured object need to stay still for few seconds. In this situation, the detected faces are half of the samples being tested which is 5 out of 10. For the darker surrounding condition, the sample faces with darker skin tone are failed to be detected and vice versa. Table 1 shows the overall summary of the face detection result.



Figure 2: Overall connection of DE2-115 board with (a) OV7670 (b)TRDB-D5M

Table 1: Result of face detection

Sample face/s	Detected face/s
1	1
2	2
3	3
5	5
6	6
8	7
10	10
15	15
15	13
15	12
10	5

3.2 Performance of TRDB-D5M

The software implementation of TRDB-D5M camera and DE2-115 using Verilog code is successfully compiled and the programmer tools of Quartus is used to run the code into FPGA board through USB-Blaster.

The overall connection of the system is successfully developed as shown in Figure 2(b). This shows that the input data from TRDB-D5M is successfully interfaced with DE2-115 board for storage and process. Then the data is displayed on the monitor screen through a VGA cable in real-time image/video with size of 640*480. The TRDB-D5M is directly connected to the GPIO of FPGA board.

The system can capture real-time video and the system is reset when KEY[0] is pressed. In order to snapshot the moment of video into an image, KEY[2] is pressed and KEY[3] is pressed to resume it. The advantages of this camera module is that it can adjust the brightness of image captured, zoom in/out

and snapshot. SW[0] must be turned on and KEY[1] is pressed to adjust the image to be brighter. When KEY[1] is pressed without SW[0] activated, the output image is adjusted to become dimmer. When SW[16] is activated with the KEY[0] the size of captured image can be enlarged and vice versa.

The output of D5M is displayed on an external monitor through a VGA cable with resolution of 640*480. The default setting of D5M is dimmer and blur. The output are then tested with brightness adjustment, enlargement and snapshot modes successfully. The enlarged image is only a quarter (320*240) of the original image where it enlarged the right bottom size of the image. All of the result are taken in snapshot mode which is capturing the picture and this make the collection of data easier.

3.3 Implementation of TRDB-D5M

TRDB-D5M camera module is implemented into the reference face detection code by using VHDL language in Quartus 15.1. The replacement of camera module is to implement the advantages of TRDB-D5M to overcome the low resolution of OV7670 camera module to get a better result.

The implemented code is successfully compiled and the code is 100% successfully coded into FPGA board through USB-Blaster. The code is similar with the reference code of face detection designed in section 3.1 . The process of image storing and image processing is the same, only the camera module is replaced to TRDB-D5M.

The overall connection of the system is successfully developed as shown in Figure 3. The hardware connection is similar to the connection of section 3.2 but the key and switch buttons are as described in section 3.1. It is using the same face detection system architecture but the only difference is in the camera used.

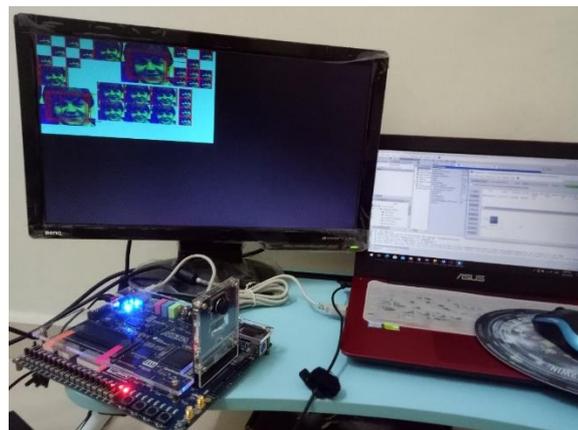


Figure 3: Overall connection of face detection using TRDB-D5M

The output is displayed on a monitor screen through a VGA cable. However, the output displayed on the monitor screen is not the image captured from the camera module. It is displaying a default picture for face detection. The main problem is that both VHDL language and Verilog language cannot be mix up in Quartus. When display the result in real-time, the red frame on the screen is blinking which means that the face detection process and boxing of faces are functioning well. The output connection is also not a problem because the default image is shown. Therefore, the interface of input module and FPGA board could be the cause of failure to display the captured image.

The OV7670 camera module has camera driver that consists of OV7670_controller and OV7670_capture before the data is sent into image buffer while the input module of TRDB-D5M camera consists of CCD_Capture, I2C_CCD_Config and RAW2RGB. The function of block OV7670_capture is similar to CCD_Capture while OV7670_controller is similar to I2C_CCD_Config. However, the overall module of D5M included RAW2RGB block that convert the raw data to RGB before sent to SDRAM for storage.

The difference between OV7670 and TRDB-D5M is OV7670 camera has 8 bits input data and the output data is RGB4:4:4 format which is 12 bits while TRDB-D5M has 12 bits input data and the 12 bits raw data is sent to RAW2RGB so that the input data is converted into RGB form. Therefore, the data sent into SDRAM is 12 bits for each RGB respectively.

Based on the information collected, the overall design met the main challenge at the input. The image capture failed to be loaded into FPGA board and this is the main challenge for this study. The TRDB-D5M camera has RGB Bayer pattern colour filter array and output of raw data is the cause of failure of the input code to communicate with FPGA board due to the code is based on the functionality of OV7670 camera.

4. Conclusion

In conclusion, the implementation of the face detection system being developed is partially successful. The overall software implementation has successfully designed by using Quartus 15.1. Using OV7670 camera as the input, the acquisition of image is successfully done, stored and displayed on the monitor. It has been tested using several samples and the result is very accurate. However, the camera resolution is very low and the displayed images are very small in size. In order to enhance the face detection system, TRDB-D5M camera is used to replace the OV7670 camera. With the new camera, the image acquisition part has failed to load the input data into the memory architecture. Nevertheless, the monitor successfully displayed the default image and the red frame is blinking during the video stream. This represented that the implemented segmentation of the acquired image to the allocated memory is success and the monitor is able to display the acquired image.

Acknowledgement

Thank you to the Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia for providing the platform and hardware for the study to be carried out successfully.

References

- [1] M. Kim, D. Lee, and K. Y. Kim, "System architecture for real-time face detection on analog video camera", *Int. J. Distrib. Sens. Networks*, vol. 11, pp. 1–11, 2015
- [2] A. R. P. Patil and M. R. Mulla, "A review : Design and implementation of image acquisition and voice based security system", *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, vol. 4, no. 3, pp. 1651–1656, 2015
- [3] J. Choi, J. Shin, D. Kang, and D. S. Park, 'Always-On CMOS image sensor for mobile and wearable devices', *IEEE J. Solid-State Circuits*, vol. 51, no. 1, pp. 130–140, 2016
- [4] "Altera: Introduction to The QUARTUS II Software | element14 | Technical Library". [Online]. Available: <https://www.element14.com/community/docs/DOC-40098/1/altera-introduction-to-the-quartus-ii-software>. [Accessed: 11-Dec-2019]
- [5] "Index of /downloads/cd-rom/d5m/". [Online]. Available: <http://download.terasic.com/downloads/cd-rom/d5m/>. [Accessed: 08-Jul-2020]
- [6] C. B. Cheng and A. B. Jambek, "Implementation of a camera system using Nios II on the Altera DE2-70 board", *Indones. J. Electr. Eng. Comput. Sci.*, vol. 14, no. 2, pp. 513–522, 2019