FPGA Based System Prototype for Face Location and Recording of People Getting on a Public Transport Bus.

Jorge Martínez-Carballido, Carlos García-Lucero, Juan Manuel Ramírez-Cortés, Rogerio Enríquez-Caldera

Electronics Department, National Institute of Astrophysics, Optics and Electronics
{jmc, cgl, jmram,rogerio}@inaoep.mx

Abstract

Assault and robbery on the public transport is a social problem. Almost every day, people lose possessions and, in some cases, thieves or citizens get hurt or even die, in the process. In Mexico, attempts to reduce or prevent crime include vigilance by armed agents [1] *and video cameras on public transport* [2].

This work contributes to crime reduction and/or prevention on public transport by taking photographic records of people as they get on the bus, using an FPGA based system with camera and Secure Digital (SD) storage.

An FPGA is used as the processing device which contains all the necessary logic to obtain the images from the camera, process them to locate the face, extract face region and store them into a SD media using an uncompressed BMP file.

For face area location single channel image was used to extract regional attributes and their inter-regional relationship, resulting in a face detection algorithm with tolerance to skin tones and light conditions.

1. Introduction

The use of technology as a tool for preventing and solving crime is a natural response to the decreasing security. Also, faster and smaller integrated circuits (ICs) which can be used to process ever increasing image resolution and higher processing capability of information are now available at lower cost. Increasing demand of high resolution cameras of cellular phones, personal devices and laptops, lead to economies of scale and decrease of prices.

Technology to prevent and decrease assault and robbery on public transport, and public places, is being used around the world. For instance, United Kingdom government has invested billions of pounds on video cameras to be used to monitor the security by police [3]. The initiative included:

Images database to track and identify offenders.

Publish images of robbery cases on internet.

National CCTV recordings and pictures database of convicted offenders as well as unidentified suspects.

Although a good strategy to reduce crime was planed, only 3% of assaults were solved by the use of CCTV technology [3]. Problem consists on people monitoring the cameras because they had little training, and the task of watching video for hours is not effective. Also when something occurs, search for the exact moment of occurrence is a hard work.

After training and new strategies, CCTV technology was helpful in solving 15-20% of street robberies [3].

In Mexico, the use of technology to prevent crime is very limited and little information can be found. In June, 2009, the State of Puebla Congress' security commission approved the use of camera installation on public transport [2]. This preventing technology is on its experimental phase, and Congress expressed its concern about the high cost of this technology for transport owners. The article didn't mention about the usage recordings and monitoring tasks that will be carried out.

2. Problem Description

Having real time video from transport to observe the events as they take place seems like a good strategy, because police could respond immediately if an assault is happening. But when talking about hundreds of transport units that could have the same technology, meaning that hundreds of videos are

being received for their inspection in real time, the monitoring task begins to be difficult, then perhaps their option is to record the video on some kind of storing media for later review/inspection which of course means police could not respond as the assault is taking place.

As other countries had experienced, people monitoring videos must be trained for this task, so assigning a police to monitor videos generally gives mixed results, but purposely trained police agents guarantee to have a stable performance.

This work approach is to use digital cameras as a media to obtain still images (frames) from people faces to store these into a storage media without human intervention during the process. When an assault takes place, the face image of the thief has already been stored on the media, which also prevents the thief from erasing its picture by destroying the camera, for its later identification. Also a massive publishing of this picture could prevent the thief from assaulting again, and give the police a good starting point to locate and process this delinquent. By no using any human intervention during the process, the system reduces its operating cost due to the use of less equipment and any kind of monitoring training is no longer needed.

3. Solution approach

The system has the following peripheral modules: camera [4], PSRAM [5], SD [6] [7], RTC [8], and people detection to the FPGA. A Block RAM, camera bus serial controller, I2C controller and image processing modules are implemented in VHDL and synthesized on a Xilinx XC3S1000-4FT256 FPGA as shown on Figure 1.
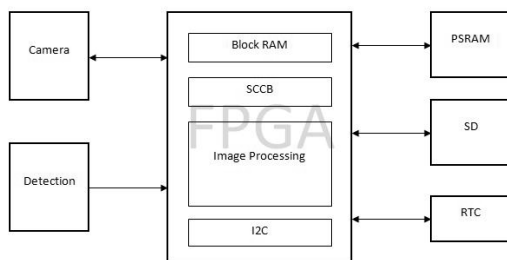

Figure 1. System's modules structure

4. Detection

By using a 3 axis accelerometer [9] positioned on the steps of the bus entrance as shown on Figure 2, we can obtain accelerometer signals to detect that

people has stepped on the bus after the bus has stopped.

Figure 2. Accelerometer position on bus entrance

Figure 3 shows the area of interest where one can detect from the three axis signals that somebody has stepped on to the bus after it has decelerated and stopped.
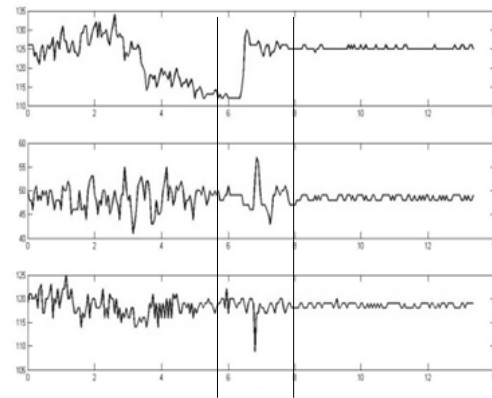

Figure 3. Accelerometer signals

5. Face location module

First we consider that from the field of view for the 1.3 Megapixel camera, only a sector of this field of view is where the face of people getting on the bus is possible to appear. Also it is possible that for a given bus, the area is different as well as whether the person is on the first or second step of the entrance. Tests with people of different heights were made so that the First Step Region (FSR) and Second Step Region (SSR) could be defined, along with a measuring stick. Figure 4 shows one of the tests.

Figure 4. FSR and SSR test images

Figure 4 shows First and Second Step Region with a rectangle of n by m pixels as obtained for an application. Next to locate a face within a region of interest we took the approach of selecting local properties of the sub-image of interest in a fraction of its possible dimensions and define their interrelationship to form the sub-image. Now when a person dresses with clothes of similar response to skin color on the selected plane, we have the potential of finding regions representing clothing. With this in mind a filter was used to improve face detection.

A binary matrix was used to represent local regions with properties that make them candidates to be a face region.
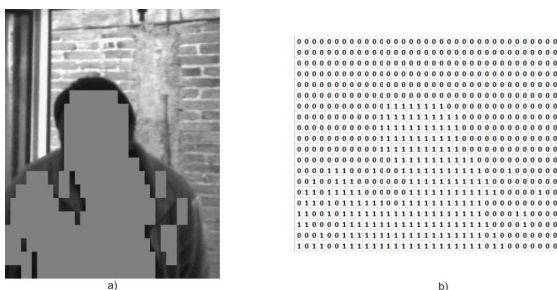

Figure 5. Binary matrix of local properties

In the binary matrix a 0 represents this local region *is not* a candidate, and a 1 represents this local region *is a candidate*. Figure 5 shows an image and its corresponding binary matrix.
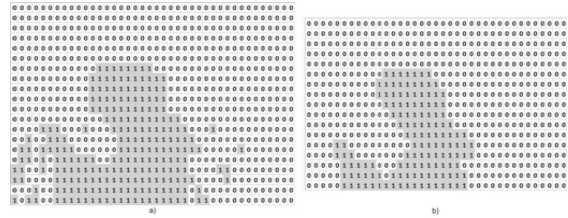

Figure 6. Binary matrix XOR filtered

After de-noising the binary matrix, another filtering is needed to exclude clothing, this is achieved by considering physical and dimensional properties of a face, as compared to the trunk section of a body. A sample is of this application is shown on Figure 7.
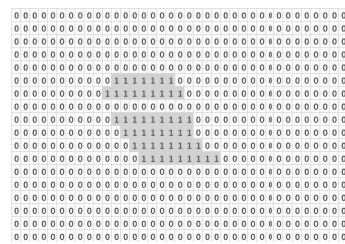

Figure 7. Isolated face region

6. FPGA implementation

The tool used to synthesize, place and route, and implement the VHDL code into the Spartan3 FPGA was the free software from Xilinx: ISE WebPACK 11.2. Figure 8 shows the RTL schematic from the VHDL TOP module. This figure shows the location of SD, PSRAM, Camera, XOR filtering and Coordinates modules generated by the ISE WebPACK.
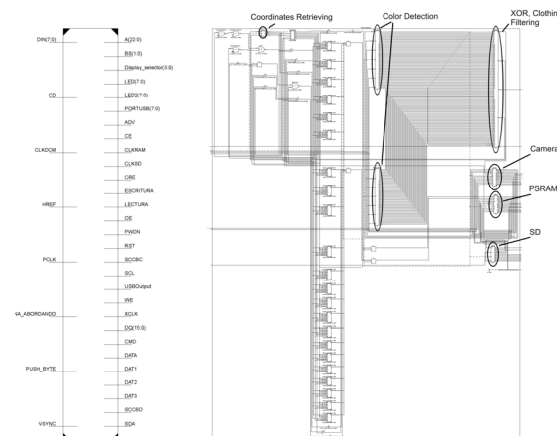

Figure 8. TOP module's RTL Schematic

7. Results and Tests

Several tests for good location of the system prototype within the bus where tried, Figure 9 shows the location used for the following tests.

Figure 9. Prototype location within the bus

Tests were taken for First and Second steps of the entrance ladder on the bus. Images on this paper are gray scale, but originals are in RGB555 BMP format.

Figure 10 show images stored on the SD of the system's prototype, once taken and extracted the located face on first and second step regions configured on the FPGA.
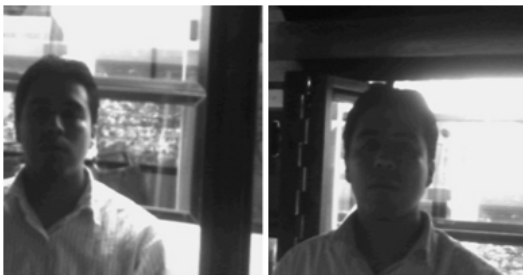

Figure 10 gray scale images for FSR and SSR

8. Conclusions

As final result we have a prototype based on a FPGA with a 1.3 pixel camera module, SD card module that is capable of taking images of people when they get on the bus on first and second steps of the entrance. From the whole 1.3 megapixel, the system on the FPGA locates the face and stores only the face region found on its corresponding first or second step region, this helps to identify people that have gotten on the bus, saves space on the SD card and in the event of assault or some incident, these images could be useful to locate the person or persons involved, thus helping to diminish events by helping authorities to identify and process legally any person breaking the law on a bus.

Use of local properties on a single channel image allowed a successful location of faces on all the tests and its FPGA system implementation simpler than algorithms on literature that require floating point and/or trigonometric operations [10] [11] [12] [13].

9. References

[1] Saúl Vela. *Milenio.com*. March 2009. [Online]. http://www.milenio.com/node/189736. [Accessed: September 03, 2009]

[2] Gerardo Orta. *periodicodigital.com.mx*. June 2009. [Online]. http://www.periodicodigital.com.mx/index.php?option=com_content&task=view&id=92893&Itemid=67. [Accessed: September 03, 2009]

[3] Owen Bowcott. *guardina.co.uk*. May 2008. [Accessed: December 12, 2008]

[4] Omnivision. (2005, August) OV9650FSL Color CMOS SXGA (1.3 MegaPixel). Concept Camera Module with OmniPixel® Technology.

[5] Micron Technology, Inc. (2004) Async/Page/Burst CellularRAM™ 1.5.

[6] SANDISK CORPORATION. (2003) Sand Disk Secure Digital (SD) Card Product Manual, Rev 1.9©.

[7] SD Group. (2000, February) SD Memory Card Specifications. Part 2 FILE SYSTEM SPECIFICATION.

[8] Philips Semiconductors. (1997, Jul) PCF8583 Clock/calendar with 240 x 8-bit.

[9] Pavel Lajšner and Radomír Kozub. (2007) Using the MMA7360L ZSTAR2 Demo Board.

[10] Shanq-Jang Ruan, Mon-Chau Shie, Yi-Chi Liu Yu-Ting Pai, "A SIMPLE AND ACCURATE COLOR FACE DETECTION ALGORITHM IN COMPLEX BACKGROUND," in *International Conference on Multimedia and Expo*, vol. 1-5, Taipei, Taiwan, 2006, pp.

1545-1548.

[11] Zhe-Ming Lu, Xiu-Na Xu Wei-Min Zheng, "A Novel Skin Clustering Method for Face Detection," in *First International Conference on Innovative Computing, Information and Control*, vol. 3, China, 2006, pp. 166-169.

[12] Chang-Tsun Li Wen-Hsiang Lai, "Skin Colour-based Face Detection in Colour Images," in *International Conference on Advanced Video and Signal Based Surveillance*, UK, 2006, pp. paper56:1-6.

[13] K.D. Adaos, G.P. Alexiou V. Mariatos, "Design and Implementation of a Reconfigurable, Embedded Real-Time Face Detection System," in *INTERNATIONAL WORKSHOP ON RAPID SYSTEM PROTOTYPING*, Greece, 2007, pp. 65-68.