A Real-time Object Detecting and Tracking Design for Railroad Crossing Application

Shih-Lun Chen[1], Chi-Hao Liao[2], Hao-Cheng Yu[3], Tang-Yu Lee[4], and Ting-Lan Lin[5]

Department of Electronic Engineer, Chung Yuan Christian University, Taiwan
200 Chung Pei Rd., Chung Li City, Taiwan, R.O.C.

[1]chrischen@cycu.edu.tw
[2] daughtry1745@gmail.com
[3] henry357951@hotmail.com
[4] u3050707@gmail.com
[5]tinglan@cycu.edu.tw

Abstract—**As the promotion of transportation construction, there are a lot of railroad crossings in Taiwan. However, the traffic accidents are always happening on the railroad crossings every year. In order to protect the drivers against the accidents on the railroad crossing, an object detecting and tracking design is proposed for an automatic warning system. Since it is necessary to identify a vehicle getting stuck on the railroad crossing or not in a very short time, the very large scaled integration (VLSI) technique was used to develop the image processor for this system. This design not only detects any vehicle getting stuck on the railroad crossing or not but also tracks all vehicles across the railroad crossing. This double checking is guaranteed to the driver's life. This design was realized by hardware description language (HDL) verilog and realized by an Altera DE2-115a field programmable gate array (FPGA) development board. By using the VLSI technique, this design has the benefits of high performance, low cost, and low power consumption. The accuracy of vehicle detection can be greatly improved with the combination of object detecting and tracking techniques.**

Keywords—**Feature extraction, FPGA; object detection; object tracking; railroad crossing; VLSI**

I. INTRODUCTION

In recent years, there are many traffic accidents happened on the railroad crossings. An accident completely destroys the happiness of the family. As the advantage of semiconductor and image processing, some techniques will provides an efficient way to avoid the accident to happen again on the railroad crossing. Hence, we hopes to develop an automatic warning system to avoid the accident happened on the railroad crossings. As shown in Fig. 1, this system has the ability to detect and track a vehicle when it crosses the railroad crossing. This system includes an image sensor device to capture the image on the railroad crossing and an image processor to detect and track the vehicles. So first of all we need to develop an efficient object detecting and tracking algorithms.

There are many efficient object detection algorithms proposed by some research groups. A basic efficient object detection algorithm was present by Horprasert [2], in which the screening differences were used to separate the object from the background of the images. Moreover, the selection of the appropriate threshold was used to determine the pixel if it

(a)

(b)

Fig 1. Vehicle detecting and tracking when it crosses the railroad crossing, the video source download from [1].

belongs to foreground, shadow or background. This algorithm successfully detected the object by a low complexity methodology. Stefano *et al.* [3] improved the method of [2] and proposed a high performance system. It divided the whole image into 8×8 matrix to calculate the gradient values. This technique greatly improved the performance and efficiently reduced the noise of the image. The difference base methods are widely used in many previous studies. It is also used in the proposed design to develop the object detection algorithm by the differences between the foreground and background.

As the development of technology, the integrity of monitor system increases day by day. The researching algorithms for the video surveillance systems are more and more important in now and future [4]. This study discussed the moving target detection and tracking algorithms and proposed a mixed solution of the multiple techniques. Javed *et al.* proposed a hierarchical background subtraction methodology by using the information of the color and gradient values [5]. It has the benefit to avoid the various interference of the light when it is used in outdoor environments. This technique is suitable for developing the proposed system since it can reduce the interference of the sunshine reflection and headlights from the vehicles.

Liao *et al.* [6] illustrated an experimental system to set up a CCD camera on self-propelled vehicle to real-time capture the images. The captured image was sent to the monitoring PC to real-time process it and then produced the control signals to feedback control the vehicle. This study proposed an efficient method to obtain the position and speed of the vehicle by calculating the color and the center of gravity. This study also presented an object detecting and tracing technique by comparing the shape, color, and surface of the vehicle. It provided an efficient method to track the vehicle entering and leaving a specific region.

An illegal entrant detection system in restricted area had been proposed by Yuan [7]. This system can detect and track any target when it enters the restricted area. Brémond [8] developed a method to be able to track multiple objects simultaneously. It used the amount of the moving information to estimate the location and size of the objects in the next screen. Zafar *et al.* [9] successfully developed a night vision detection Portrait method on FPGA development board. The parallel systolic array architecture and body contour capturing techniques had been used in this system. Finally, this system was successfully demonstrated by the real film. A dynamic object detecting algorithm by using Bayesian module was proposed by Sheikh *et al.* [10]. This technique successfully detected the object on a moving mirror according to the position and color information of the object.

The previous studies mentioned above had done great contribution in object detection and track regions. These studies also provided well base for us to develop a real time object detecting and tracking system for the railroad crossing application. In this paper, an object detecting and tracking algorithm will be proposed. It uses the travelling distance to track the vehicle due to the interference of environment light, reflection, and headlights from other vehicles. The block base technique, dividing whole frame into 8×8 matrix, is used to reduce the noise of the capture image. The object detecting and tracking algorithms are implemented by the HDL and realized by the FPGA development board to achieve the goal of real time vehicle detecting on the railroad crossings. With the development of this automatic warning system, we hope the accidents on the railroad crossings could be greatly reduced.

II. Object Detecting and Tracking Algorithms

The identification process of the proposed object detecting and tracking algorithms includes three steps. The first step is to choose the detect region. In this step, the image is captured by the camera first and then the system will select a range as the detection region. To set the detect region can avoid to detect a non-essential object or be interfered by external noise. The second step is to execute the object detecting and tracking processes. When the system finds any vehicle enters into the detection region by object detection algorithm, the system will start to track the vehicles. Simultaneously, the system will also calculate the moving distance of the vehicles as the determining information.

The last step is to determine whether the vehicle leaves the railroad crossing or not. In this step, the system calculates the moving distance of the vehicles first. If the moving distance of any vehicle is less than a threshold value during a time slice, the system would determine that the vehicle is in dangerous situation. The system will send a warning message automatically to the pilot of the train. By these three steps, this system can identify that the vehicles are staying on the railroad crossing or not. The details of object detecting and tracking algorithms will be introduced as follow.

A. Object Detecting Algorithm

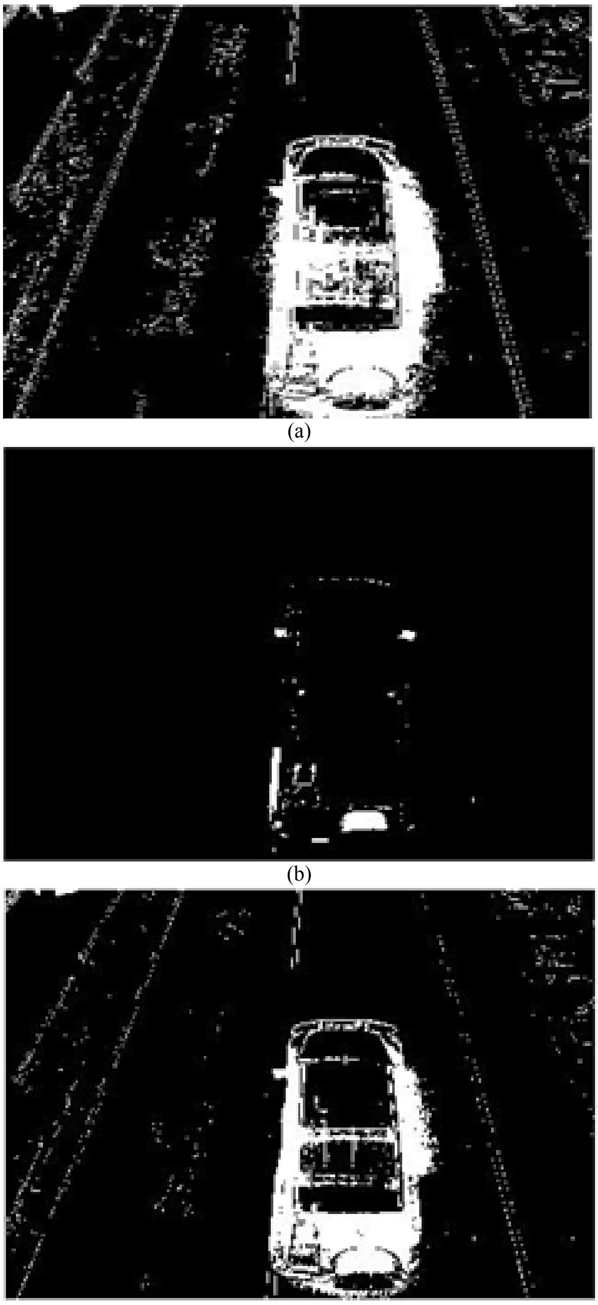
The first step of the object detecting algorithm is to find a reference image as the background image. The characteristic of the reference image is there is not any object in the detection region of this image. In order to test the proposed object detecting algorithm, a camera was used to capture a road continuously. First, we captured an image as the reference. Fig. 2 shows the captured reference image, in which there is not any object on the road. The camera of the system

Fig 2. The reference image

Fig 3. The current image in which a vehicle appears in the detection region.

captures the image. In the same time, the system executes the object detecting algorithm by comparing the current image with the reference image. If the detecting algorithm finds no object in the detecting region, the reference image would be replaced by the current image. The refresh of the reference image will reduce the distortion from changes of the weather and lights.

The reference image will be replaced continuously until the object appears in the detection region as shown in Fig. 3. In order to reduce the distortion and noise, the detected image is converted to a binary image. Fig. 4 shows three binary by the threshold values of 20, 80, and the threshold value obtained by OTSU's method respectively. The OTSU's method [11] can efficiently reduce the interference of the weather, light, and shadow from the outside world. The threshold value can be calculated by OTSU's method according to situation of the comparison of these three binary images, we can easily find that the OTSU's method used to find the appropriate threshold can get the highest performance of the object detection. By converting into the binary image, the distortion and noise can be greatly reduced, which will improve the accuracy of the object detection.

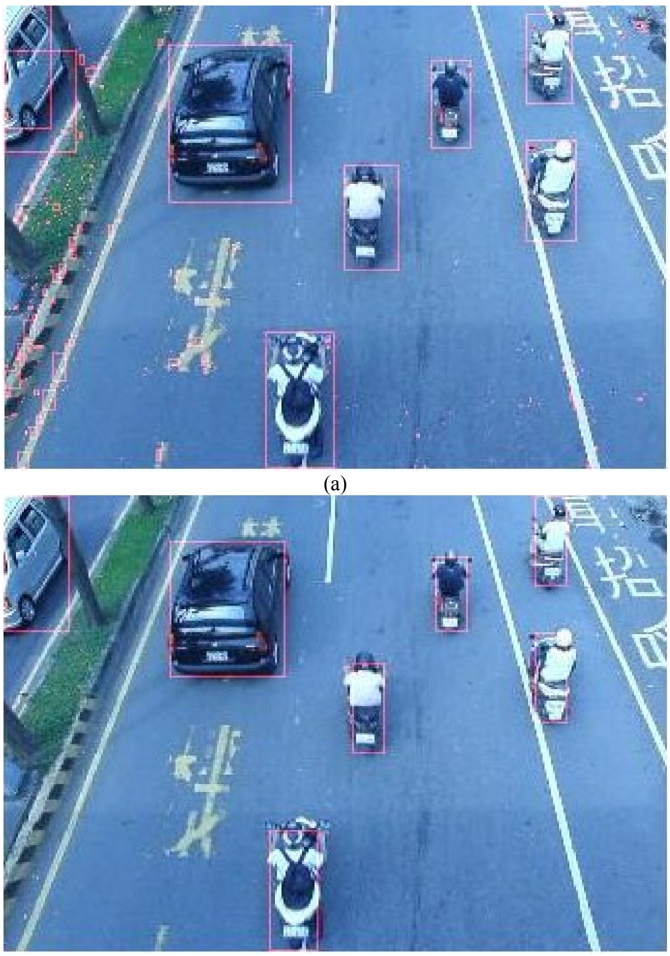


(a) (b) (c)

Fig 4. The binary images. (a) threshold value is 20. (b) threshold value is 80. (c) threshold value obtained by OTSU's method.

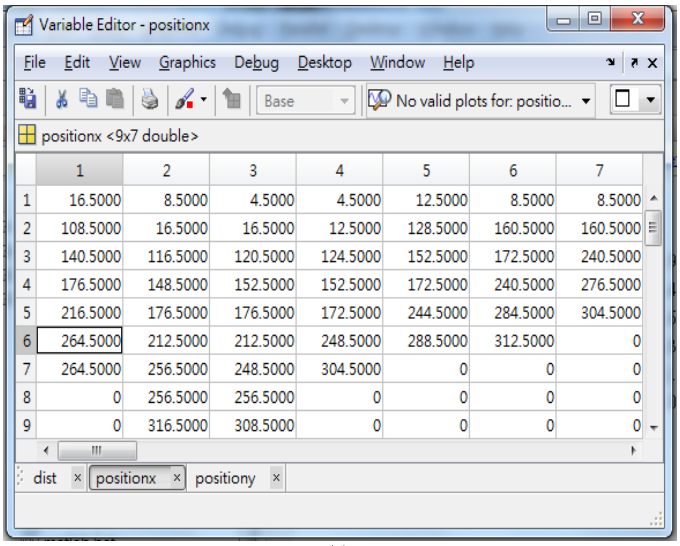
B. Object Tracking Algorithm

Although the binary image can reduce the interference from the outside world, it still has the some small regions of noise which causes the misjudgement of object tracking. In order to reduce more noise, the captured image was separated into the block of 8×8 matrix first and then each block was filtered by the image filter. After filtering, the small regions of noise can be eliminated completely, which makes the pixels of the vehicle more clearly. Fig. 5 shows the testing image before filtering and after filtering. We can easily find that the small regions of pixels are reduced by the filtering process.



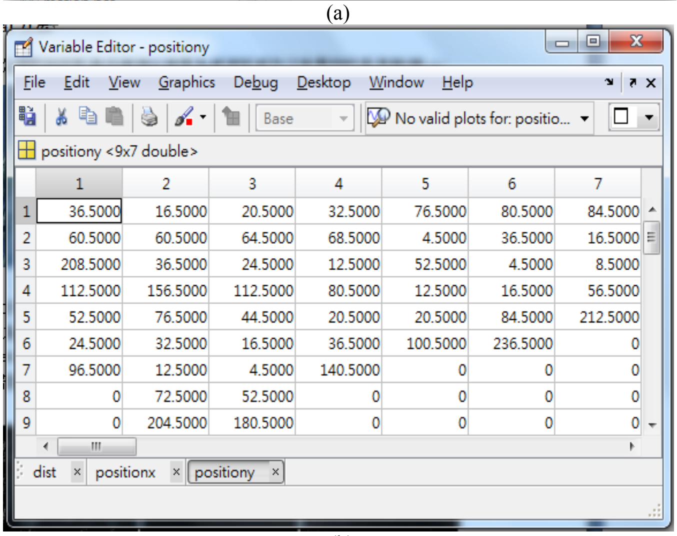
(a) (b)

Fig 5. (a) The image before filtering. (b) The image after filtering.



	1	2	3	4	5	6	7
1	16.5000	8.5000	4.5000	4.5000	12.5000	8.5000	8.5000
2	108.5000	16.5000	16.5000	12.5000	128.5000	160.5000	160.5000
3	140.5000	116.5000	120.5000	124.5000	152.5000	172.5000	240.5000
4	176.5000	148.5000	152.5000	152.5000	172.5000	240.5000	276.5000
5	216.5000	176.5000	176.5000	172.5000	244.5000	284.5000	304.5000
6	264.5000	212.5000	212.5000	248.5000	288.5000	312.5000	0
7	264.5000	256.5000	248.5000	304.5000	0	0	0
8	0	256.5000	256.5000	0	0	0	0
9	0	316.5000	308.5000	0	0	0	0

(a)



	1	2	3	4	5	6	7
1	36.5000	16.5000	20.5000	32.5000	76.5000	80.5000	84.5000
2	60.5000	60.5000	64.5000	68.5000	4.5000	36.5000	16.5000
3	208.5000	36.5000	24.5000	12.5000	52.5000	4.5000	8.5000
4	112.5000	156.5000	112.5000	80.5000	12.5000	16.5000	56.5000
5	52.5000	76.5000	44.5000	20.5000	20.5000	84.5000	212.5000
6	24.5000	32.5000	16.5000	36.5000	100.5000	236.5000	0
7	96.5000	12.5000	4.5000	140.5000	0	0	0
8	0	72.5000	52.5000	0	0	0	0
9	0	204.5000	180.5000	0	0	0	0

(b)

Fig 6. (a) The coordinate position of the vehicle in the X axle. (b) The coordinate position of the vehicle in the Y axle.

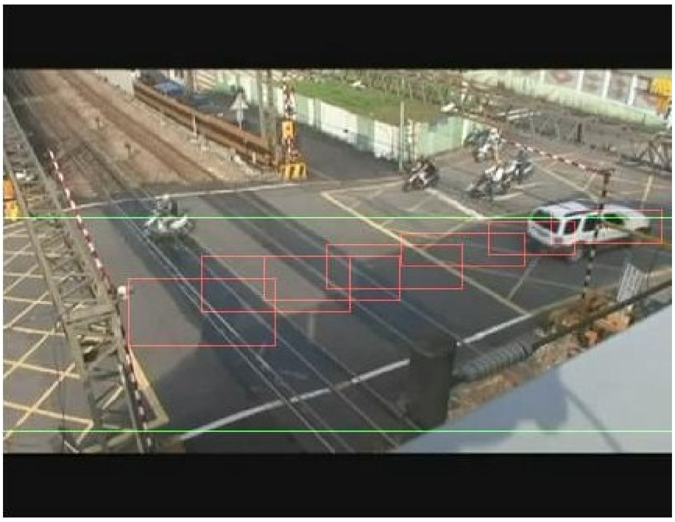
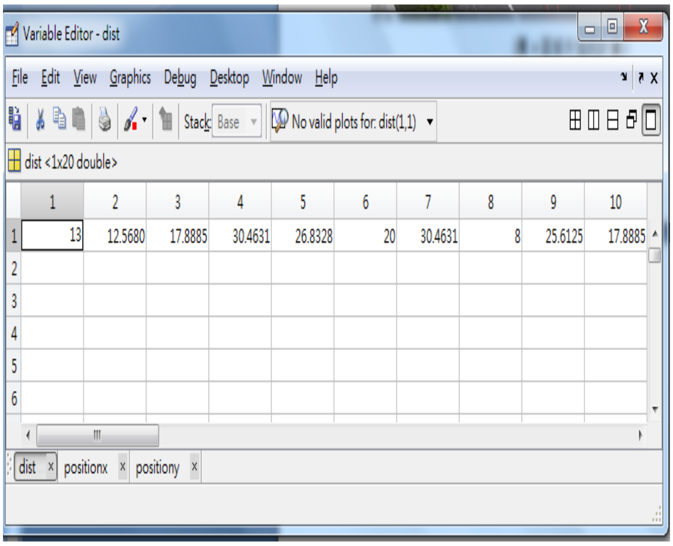


Fig 7. The vehicle is tracked in the specific region.

The 8-connedted components labelling technique was used to detect the center of position for each vehicle. By this technique, this can track multiple vehicles in the same time. Fig. 6 shows the experimental results of the coordinate position of each vehicle in the X and Y axles. The position of the center of each vehicle provides not only the position of the vehicle but also the speed of each vehicle by computing with the time information. If the system finds that the speed of the vehicle is zero, the system would change from the tracking mode to object detecting mode. The alternate tracking and detecting modes improve the accuracy of this system, which protect the drivers against accident more completely.

In order to verify this system, we used a real video captured from a railroad in Taiwan. Fig. 7 shows that the vehicle can be tracked successfully by this system. Fig. 8 shows the number of vehicles tracked by this system during a slice of time. Although the proposed two algorithms can detect and track vehicles successfully on the railroad crossing, the performance of the software system cannot satisfy the requirement of real time application. Hence, some partitions with high complexity are necessary to be implemented by the VLSI technique or FPGA. The hardware architecture will be introduced in the next section.



	1	2	3	4	5	6	7	8	9	10
1	13	12.5680	17.8885	30.4631	26.8328	20	30.4631	8	25.6125	17.8885

Fig 8. The number of vehicles tracked during a slice of time.

III. Hardware Architecture

In order to achieve the real time requirement for the railroad crossing application, it is necessary to realize the object detecting and tracking algorithms by VLSI technique. Fig. 9 shows the block diagram of the proposed design. First, the images were captured by the 5 Mega pixels digital camera on Altera DE2-115a field programmable gate array (FPGA) development board. Second, the captured images were stored into the DDR memory through a memory controller. Third, the captured images in red, green, and blue (RGB) format were read by the memory controller and then sent to a RGB to Y module. The RGB to Y transferred the captured images in RGB format to grayscale format.

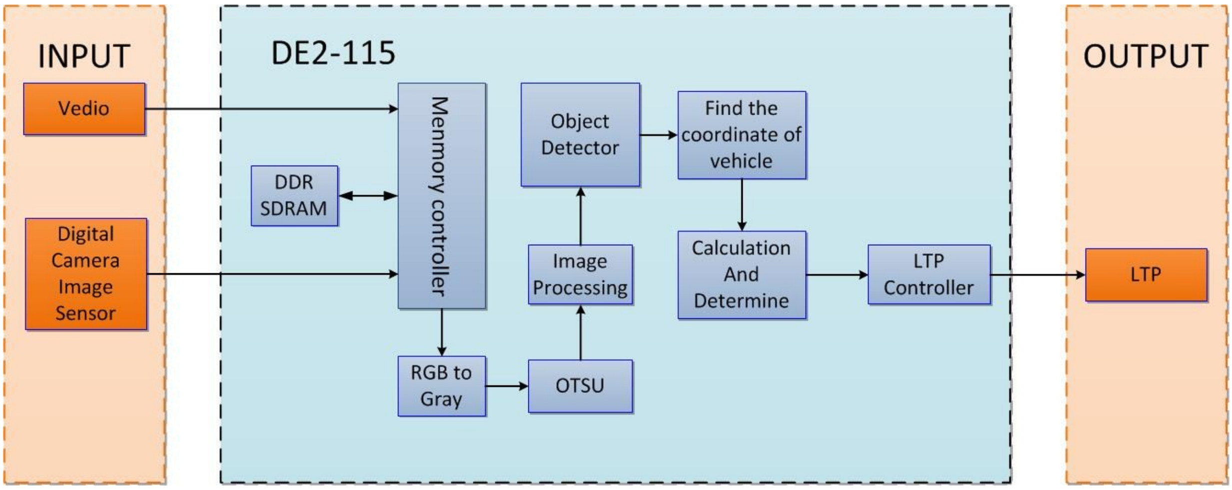


Fig 9. Block diagram of the hardware architecture.

Fourth, the grayscale images were processed into a binary image by an OTSU module. After converting into binary image, the object detector will find the coordinate positions of the vehicles. Finally, the result images were sent to touch panel on the FPGA development board for displaying. The warning signal is produced in the same time for the pilot if the train. Fig. 10 illustrates the data flow of the proposed design.

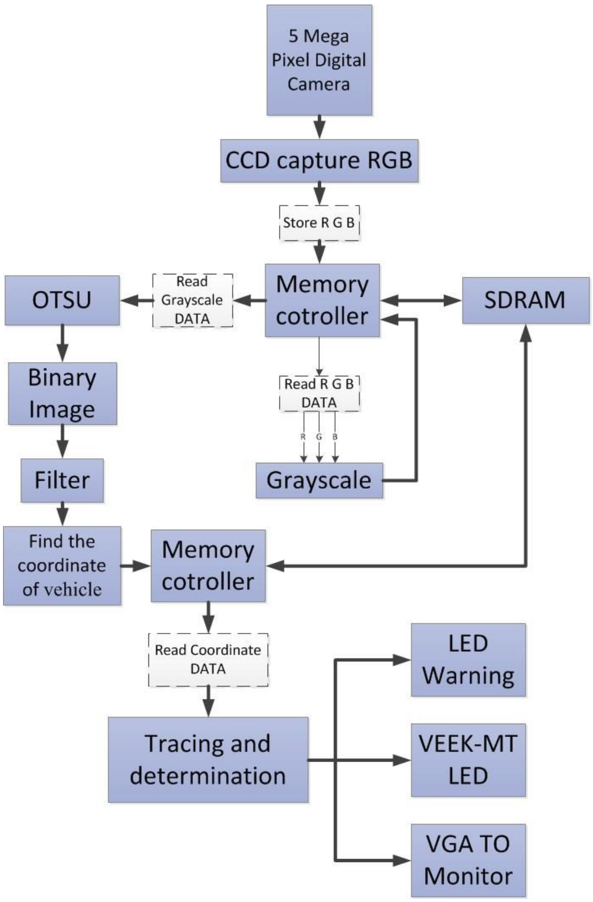


Fig 10. The data flow of the proposed design.

The design was also implemented by the Altera DE2-115a field programmable gate array (FPGA) development board. By this development board, the test images can be capture by the 5 Mega pixels digital camera, the HDL designs of the image processor can be download into the FPGA, the warning signal can be shown by the LEDs, and the result images can be displayed on the panel of the DE2-115 FPGA development board. The process of FPGA implementation can be separated into six steps. The details of each step are described as follow.

1. Capturing the testing images by the 5 Mega pixels digital camera with CCD and then writing the image data into the SDRAM.
2. Reading the image with the RGB format through the memory controller from the SDRAM and then writing the calculating grayscale image data back to the SDRAM.
3. Using OTSU's method to find the appropriate threshold for converting the grayscale image into binary image.
4. Using filter to remove noise and calculating the coordinate positions of the vehicles and then writing the image data into the SDRAM.
5. Determining whether the vehicles are leaving the railroad crossings or not by calculating the coordinate position of each vehicle.
6. If the system determines a vehicle stacking on the railroad crossing, the Red LEDs on the DE2-115 FPGA board are turned on and the panel is also displayed the result image by the image processor downloaded in the FPGA.

Fig. 11 shows the result of the warning signals turned on the red LEDS built in the FPGA development board. In the real application, the warning signal will send to the pilot of the train by wireless device. Fig. 12 shows the result image displayed on the touch panel built in the FPGA development board. In the real application, the result image will also send to the smart phone, tablet PC or LCD monitor on the train. The result image will help the pilot to judge whether stop the train emergently or not. Table I lists the performance of this design implemented by FPGA and TSMC 018-um CMOS process by operating at 50 MHz and 200 MHz respectively.

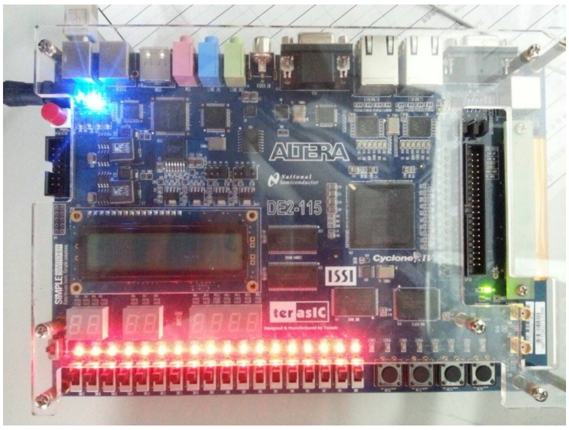


Fig 11. The warning signals turned on the red LEDS built in the FPGA development board.

Fig 12. The result image was displayed on the touch panel built in the FPGA development board.

TABLE I
THE PERFORMANCE OF THE PROPOSED DESIGN

Resolution (pixels) / Process & CLK	30 million	120 million	500 million
FPGA @50MHz	0.0985(s)	0.394(s)	1.64(s)
TSMC 018 @200MHz	0.0246(s)	0.098(s)	0.41(s)

The result shows this design can process an image with resolution of 30 million pixels within 0.1 second by FPGA device. Moreover, it consumed only 0.41 second to process an image with resolution of 500 million pixels by TSMC 018-um CMOS process. This design provides a high performance, low complexity, high accuracy, and high compatibility object detecting and tracking design for the automatic warning system for the railroad crossing application.

IV. CONCLUSIONS

This paper proposed an efficient design for object detecting and tracking. With the implementation of this design by VLSI technique and FPGA device, the design can detect and track a vehicle within one second. It provides a well base to develop an automatic warning system for the railroad crossings. We hope this system can protect people against the accidents on the railroad crossings. We will keep going to advance this design in future. We hope by our research, there are more and more life will be saved by this system.

ACKNOWLEDGMENT

This work was supported by the National Science Council, R.O.C, under Grant numbers of NSC-101-2221-E-033-036, NSC-101-2221-E-033-066, NSC-102-2221-E-033-018, NSC-102-2221-E-033-063 and NSC-102-2622-E-033-009-CC2 and the National Chip Implementation Center, Taiwan. The authors are also grateful to ALTERA and Terasic companies to hold such a good FPGA design competition for us.

REFERENCES

[1] (2013) The youtube website [online] Availble: http://www.youtube.com/watch?v=8fjlCxNDFQA&noredirect=1

[2] T. Horprasert, D. Harwood and L.S. Davis, "A statistical approachfor real-time robust background subtraction and shadow detection," i*n Proc. IEEE Frame-Rate Workshop*, 1999.

[3] L. D. Stefano, S. Mattoccia and M. Mola, "A change-detectionalgorithm based on structure and colour," *in Proc. IEEE Conf.Advanced Video and Signal Based Surveillance*, pp.252-259, 2003.

[4] Jiang-Li NA, "Video surveillance algorithm research and LabVIEW realization" Harbin University of Science and Technology, 2009.

[5] O. Javed, K. Shafiqu and M. Shah, "A hierarchical approach torobust background subtraction using color and gradientinformation," i*n Proc. IEEE Conf. Computer Vision and PatternRecognition*, pp. 22-27, 2001.

[6] Ming-Yang Liao, "Real-time imaging and fuzzy PID controlled robot-car applied in wireless intercepting mission" National Cheng Kung University Department of Aeronautics & Astronautics, 2007.

[7] Kai-Chun Yuan, "Illegal entrant detection in restricted area " National Central University Department of Computer Science & Information Engineering, 2005.

[8] F. Brémond and M. Thonnat, "Tracking multiple nonrigid objects in video sequences," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 8, no. 5, pp. 575-584, Sep. 1998.

[9] I. Zafar, U. Zakir, I. Romanenko, R. M. Jiang and E. Edirisinghe, "Human silhouette extraction on FPGAs for infrared night vision military surveillance", *Second Pacific-Asia Conference on Circuits, Communications and System*, 2011

[10] Y. Sheikh and M. Shah, "Bayesian modeling of dynamic scenes for object detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, PP 1778-1792, Nov. 2005.

[11] I. Sekita,"A thresholding method using the mixture of normal density functions" *International Symposium on Speech, Image Processing and Neural Networks*, Hong Kong vol.1 pp.304-307 ,1994

[12] Stan Z. Li "Recognizing multiple overlapping objects in image: an optimal formulation," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 9, no. 2, pp.273-277, Feb. 2000.

[13] Yan Nei Law, Hwee Kuan Lee, Chaoqiang Liu, and Andy M. Yip" A variational model for segmentation of overlapping objects with additive intensity value" *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 20, no. 6, pp.1495-1503, Jun. 2011

[14] Sahirzeeshan Ali* and Anant Madabhushi "An integrated region-, boundary, shape-based active contour for multiple object overlap resolution in histological imagery," *IEEE TRANSACTIONS ON MEDICAL IMAGING*, vol. 31, no. 7, pp.1448-1460, Jul. 2012.

[15] F. C. Cheng, S. C. Huang, and S. J. Ruan, "Illumination-sensitive background modeling approach for accurate moving object detection", *IEEE Transactions on Broadcasting*, vol. 57, no. 4, pp. 794-801, Dec. 2011.

[16] C. Kyrkou and T. Theocharides, "A flexible parallel hardware architecture for AdaBoost-based real-time object detection", *IEEE Transactions on Very Large Scale Integration (VLSI) System*, vol. 19, no. 6, pp. 1034-1047, Jun. 2011.