Pedestrian Detection System Using CENTRIST Algorithm on SIMT Based Image Signal Processor

Jun-mo Jeong¹, Hee-kyeong Jeon² and Jong-joon Park^{3*}

¹Electronic Engineering ²Computer Engineering ³Computer Science jjmo@skuniv.ac.kr¹, iq8580@skuniv.ac.kr², jong@skuniv.ac.kr³ *Corresponding Author

Abstract

This paper proposed a pedestrian recognition using CENTRIST algorithm on the basis of SIMT based Image signal processor. The conventional pedestrian recognition used algorithms such as Haar-like feature and HOG, but the huge amount of calculation makes it difficult to operate in real time in a restricted environment such as an embedded environment. This paper sought to improve the operation speed by processing the pedestrian recognition in parallel with the use of SIMT based Image signal processor and CENTRIST algorithm with less amount of calculation compared to other algorithms. The pedestrian recognition implemented in ISP exhibited the improved performance, which is about 2.7 times higher than that of the pedestrian recognition executed by Cortex-A9.

Keywords: Pedestrian detection, Image signal processor, CENTRIST, SIMT, processor

1. Introduction

With the recent development of image recognition technologies, various image recognition systems, including ADAS (Advanced driver assistant systems), security and intelligent transportation have spread. [1] In particular, the pedestrian recognition technology has widely been applied as an important field that can reduce casualties. However, the pedestrian recognition poses difficulties in maintaining high recognition rates due to the surrounding environment with various behaviors of persons and variables. In addition, the huge amount of calculation makes it difficult to operate the conventional algorithms used in the pedestrian recognition such as HOG and Haar-like feature in real time in a restricted environment such as an embedded environment. [2-3] To solve these problems, this paper attempted to implement the pedestrian recognition that operates in real time by finding out visual characteristics that can distinguish pedestrians using CENTRIST, improving target images with FFT operation through the use of SIMT based Image signal processor and processing the pedestrian algorithm in parallel.

2. Proposed Pedestrian Detection System

2.1. Pedestrian Detection Algorithm Using CENTRIST

CENTRIST (Census Transform) algorithm is a method to convert images by extracting feature points after comparing the intensity of surrounding pixels. The Census Transform makes the Census Transform Window with size of 3x3 in an image. And it arranges the neighboring pixels based on the center pixel in the Census Transform Window with size of 3x3 in a bit string and creates a histogram based on the bit string information. The

^{*} Corresponding Author

histogram is a type of bar graph used to show the characteristics of the image data at a time, the use of which can identify the distribution of contrast values configured in an image and analyze the image as shown in Figure 1. The CENTRIST-based pedestrian



Figure 1. Generating the Histogram for each Region

Recognition extracts the outline of the pedestrian by using Sobel and Census Transform algorithm with respect to the input image. And it generates the first score on the detect window with size of 36 x 108 through the use of integral image, learning data and image data in which the information of the outline is stored, and then finds out the detect window that has similar characteristics to the pedestrian. With the use of the histogram, it generates the second score and finds out the detect window that recognizes a pedestrian. For learning data, INRIA Person Dataset was used, and the learning data is generated using linear SVM. Since the pedestrian recognition algorithm extracts the pedestrian characteristics based on the outline, the recognition rate can be enhanced by using the image that performs Edge Enhancement process. However, an increase in the amount of calculation due to the Edge Enhancement process results in the increased amount of calculation in the entire process of image pre-processing and pedestrian recognition. In regards to this problem, the use of FFT with less amount of calculation compared to the conventional convolution operation in the Edge Enhancement process makes it possible to implement the pedestrian recognition with the improved recognition rate and faster processing speed compared to the conventional method.

2.2. SIMT Based Image Signal Processor

In this paper, SIMT-based Image signal processor (ISP) that supports the acceleration of operations for high-resolution images is utilized for parallel processing of the recognition algorithm and image enhancement process, which is the pre-processing process of the input image. The ISP is a command-based processor that can apply a variety of image processing algorithms, and it is also a structure that can freely change the size or type of filter depending on the image environment. In addition, the Image Enhancement process is performed in the frequency domain with less number of operations compared to the convolution operation in the spatial domain to support the image processing acceleration on the high resolution image.

Among various FFT algorithms, Cooley-Turkey algorithm suitable for a SIMT structure was used for a conversion from spatial domain to frequency domain. [4] Since the Cooley-Turkey algorithm has excellent regularity an- d modularity, it is suitable for the SIMT structure that has a plurality of execution flows with a single command.

The proposed ISP is composed of a total of 16 sub-ISPs and operates in 16 Warps and 16 Threads. One sub-ISP receives two commands from each of the odd Warp and even Warp and can process up to three ALU operations and one memory operation at the same

time. In addition, it has four OCs (Operand Collectors) that perform read/write operations on the Register to which the operand of the command refers from the Register Bank. Prior to processing OC, it has two types of execution flows such as odd Warp and even Warp, and therefore processes two execution flows separated from the sub-ISP. Figure 2 shows the parallel processing process of the Cooley-Turkey algorithm using this structure.



Figure 2. The Parallel Processing of the Cooley-Turkey Algorithm

Each sub-ISP reads 1 channel pixel value of the image from the memory and then performs FFT operations on the 2 point for each stage at the same time. Whenever the operation for one stage is completed, it performs a Dependency Check process to relocate the index of FFT value. By performing the Dependency Check process, it enables the parallel processing on the 2-point operation to be performed in the next stage. The proposed ISP operates in 16 Warps and 16 Threads and can process FFT operation of up to 512 points in parallel. In addition, the twiddle factor W is written out by LUT (Look Up Table) and used for the FFT acceleration. Since the Filter operation for Edge Enhancement after the FFT has no dependency, it performs the multiplication and addition operations of Filter and frequency image in all sub-ISPs. After the Filter operation, IFFT is processed in parallel in the same manner as FFT.

2.3. Proposed Pedestrian Detection System on ISP

The proposed ISP performs the parallelization of algorithms such as Sobel, Census Transform, Histogram and Integral Image and thus can obtain CENTRIST features for the pedestrian recognition. With the use of CENTRIST features, it performs the pedestrian recognition in parallel in each of the sub-ISP as shown in Figure 3.

Sobel and Census Transform algorithms perform the convolution operation by using size of 3×3 and process the input image in parallel simultaneously for each line after dividing it by the number of sub-ISPs. Histogram is also performed in a similar process,

but it can be processed by dividing it into size of more than 3 x 3. Unlike Sobel, Census Transform and Histogram with less dependency between data, Integral Image generation algorithm is not suitable for the parallel processing due to a high level of dependency. In order to solve this problem, Parallel pre- fix sum [5] algorithm that can process the Integral Image generation algorithm in parallel is used. Basically, the structure as shown in Figure 4 is used to obtain the ac- cumulated value of the pixel for the Integral Image generation of the input image.



Figure 3. The Parallel Processing of Sobel, Census Transform and Histogram



Figure 4. The Structure of Parallel Prefix Sum

The conventional method cannot perform the parallel processing in such a way that the sub-sequential operations are made to obtain the accumulated value, but the use of the Parallel prefix sum algorithm makes it possible to process the accumulated value ranging from X1 to X4 in parallel by dividing it into two operations. Figure 5 shows the parallel prefix sum operation in the sub-ISP.

When the Integral Image generation algorithm for eight data is performed, it is divided into a total of four stages, and the addition operation for two values is performed in one sub-ISP. Each stage is processed in parallel by a number of sub-ISPs, which enables the acceleration of the Integral Image.



Figure 5. The Parallel Prefix Sum Operation Using Sub-ISP

3. Experiment

For the proposed ISP, an experiment and verifications were conducted through the FPGA environment. In the FPGA environment, VC707 Board equipped with Xilinx Virtex-7 was used and Xilinx ISE 14.4 and Model-Sim SE 10.0b were used for the experiment.

The proposed ISP operates in 16 Warps and 16 Threads, and the operating frequency is 50MHz at the FPGA platform. For the performance and experimental results, FFT on the image, Edge Enhancement using FFT and CENTRIST-based pedestrian recognition algorithm were processed in parallel.

Table 1. The Comparison of the Processing Time between ConvolutionOperation and FFT

	Image size	Time(ms)
3x3 Filter(convolution)[6]	512x512	486
Butterworth Filter(FFT)	512x512	381

In order to identify the difference between the convolution operation of the Spatial domain and Butterworth Filter operation of the Frequency domain, the performance time of the 3x3 Filter operation(convolution) and Butterworth Filter (FFT) was compared as shown in Table 1. According to the comparison results, the Filter operation applied with the FFT showed the processing speed improved by about 30% compared to the convolution operation.

In the proposed ISP, the image improved by the Edge Enhancement process performs the pedestrian recognition algorithm additionally. Figure 6 shows the results of performing the CENTRIST-based pedestrian recognition algorithm on the original image, Blur-processed image and Edge Enhancement image.

The performance results showed that non-detection and false detection that determines the pedestrian with respect to the area where there are no pedestrians in the original image occurred, and non-detection that cannot recognize the pedestrian occurred in the Blur image.



Figure 6. The Result Image of Pedestrian Detection System for Each Enhanced Image (a) Original Image (b) Blur Image (c) Proposed Method

	Image size	Frame per second
ARM Cortex-A9	512x512	6.67
ARM Cortex-A15	512x512	12.01
Proposed ISP	512x512	18.16

Table 2. The Comparison of the Processing Time for Pedestrian Detection

On the other hand, it was confirmed that when the pedestrian recognition was performed with the image improved by the Edge Enhancement process, the pedestrian was accurately recognized without the false detection and non-detection that occurred in the original image as well as the non-detection that occurred in the Blur image. Table 2 shows a summary of the performance time on the pedestrian recognition in the proposed ISP and CPU of the embedded platform.

According to the experimental results, the proposed ISP performed an image improvement at a faster processing speed compared to the conventional convolution operation by using the FFT. In addition, it exhibited an enhanced recognition rate by performing the pedestrian recognition algorithm with the improved image, and an improved processing speed that is about 2.7 times higher than that of Cortex-A9, which is CPU of the embedded platform and about 1.5 times higher than that of Cortex-A15.

4. Conclusion

This paper proposed the SIMT-based ISP that is capable of processing a variety of image pre-processing and recognition algorithms unlike the conventional ISPs. The proposed ISP can be applied to the filter that has different types and sizes, which is used for the image pre-processing process based on the command, and is able to process the pedestrian recognition algorithm among recognition systems. With the use of Radix-2 FFT suitable for the SIMT structure, the Edge Enhancement process was performed, and the processing speed was further improved by about 30% compared to that of the image improvement algorithm using the convolution operation in the conventional ISP. A higher recognition rate than that of the original image or Blur-processed image was obtained by performing the pedestrian recognition algorithm based on CENTRIST using the improved image as the input image, and an improved processing speed that is about 2.7 times higher than that of Cortex-A9, which is CPU of the embedded platform and about 1.5 times higher than that of Cortex-A15 was achieved. With the development of the recognition system, various recognition algorithms through camera input are likely to be used in a variety of fields, and the ISP that can perform the improved recognition algorithm through the image enhancement of the input image in which several environments are present will be essentially required. For future work, it is expected that a study on the ISP that can perform various image pre-processing and recognition algorithms as well as the improved recognition system like the proposed ISP will be conducted.

Acknowledgments

This research was supported by Seokyeong University in 2014 and the IT R & D program of Ministry of Trade, Industry and Energy (10049192, Development of a Smart Automotive ADAS SW-Soc for a Self-Driving Car)

References

- [1] H. S. Yoon, "The Trend of Technology and Market of Image Recognition Service", Communications of the Korea Information Science Society, vol. 31, no. 2, (**2013**), pp. 23-31.
- [2] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, (2001), pp. 511-518.
- [3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, (2005), pp. 886-893.
- [4] H. Jeon, Y. Hwang and K. Lee, "An implementation of SIMT based Image Signal Processor for pedestrian detection", Advanced Science and Technology Letters, (Mechanical Engineering), vol. 129, (2016), pp. 25-31.
- [5] B. Bilgic, B. K. P. Horn and I. Masaki, "Efficient integral image computation on the GPU", Intelligent Vehicles Symposium (IV), (**2010**), pp. 528-533.
- [6] S. Agarwala, P. Koeppen, T. Anderson, A. Hill, M. Ales, R. Damodaran, L. Nardini, P. Wiley and S. Mullinnix, "A 600 MHz VLIW DSP", in Dig. Tech. Papers IEEE Int. Solid-State Circuits Conf. (ISSCC2002), vol. 2, (2002), pp. 38–390.

Authors



Jun-Mo Jeong, 1985: BS degree in Electronics Engineering, Hanyang University. 1987: MS degree in Electronics Engineering, Hanyang University. 1992: PhD degree in Electronics Engineering, Hanyang University. 1991~1995: Assistant Professor, Bucheon University. 1995: Professor, Seokyeong University Author's profile.



Hee-kyeong Jeon, 2015: BS degree in Computer Engineering, Seokyeong University. 2015: Studying for MS degree in Computer Engineering, Seokyeong University.



Jong-Joon Park, 1978 BS degree in Physics, Sogang University. 1980: MS degree in Electronics Engineering, Yonsei University. 1980~1990: Assistant Professor in the Dept. of Computer Science, College of Industry, Chosun University. 1994: Ph.D in Computer Science, College of Art and Science, Florida State University.