

P-5-13L

An Object Tracking Processor Core for Intelligent Surveillance System-on-Chip Applications

Yeong-Kang Lai and Yu-Chieh Chung

Department of Electrical Engineering, National Chung Hsing University
No.250, Kuo Kuang Road, Taichung, Taiwan
Phone: 886-4-22855268 E-mail: yklai@dragon.nchu.edu.tw

1. Introduction

The Intelligent Video-based Surveillance System has become a rather popular research topic recently. Its main purpose is to reach the target by detecting and tracking methods [1]-[5]. According to the different environments, it can be divided into two kinds of surveillance systems: the indoor system and the outdoor system. The indoor surveillance system emphasizes the purpose that recognizes the face of a person, and the outdoor one emphasizes the behavior analysis of the person. We develop the intelligent video-based surveillance system which is suitable for two kinds of environments using different algorithms and devices: A PTZ camera is for indoor environment, and another fixed wide-angle camera is for outdoor environment.

In this paper, an effective object tracking algorithm and its VLSI architecture is proposed for low cost multimedia surveillance applications. According to the proposed shape perceived architecture, it uses the building-block-based matching method. This method is fast, only one-pass scan and low-cost for objects classification, it does not need traditional four or eight adjacent connected component rule. We have implemented the design, and its maximum working frequency is 27 MHz, the gate count is 10848 (only the motion detector). The experimental results show that low-cost shape perceived object tracing architecture is feasible for intelligent surveillance systems.

2. Object Tracking Algorithm and Architecture

The feature extraction method using pelwise temporal filter employs the frame partition and background subtraction for a fixed wide-angle camera and a PTZ camera situation. The fixed wide-angle camera is used to monitor the whole surveillance environment. Using current image pixel to subtract with the same location of background image pixel, given in equation (1), to get the foreground target, it is suitable for stationary camera with wide-angle record.

$$Diff(i, j) = |Live_im(i, j) - Bg(i, j)| \quad (1)$$

Where $Live_im(i, j)$ is the current image, $Bg(i, j)$ is the background image. The difference image is simplified into a binary difference image after applying the current threshold. The threshold value is adaptively selected by the calculation of whole image pixels.

The presented motion detection algorithms are based on the detection of pixel change in the observed input

frame with respect to a recursively updated background. The overall performance depends on the type of the pixel temporal filter and the image features applied to it. In this paper, we employ the temporal low-pass filter defined by equation (2). The luminance extraction method is effective for wide-angle camera.

$$Bg_{K+1}(i, j) = Bg_K(i, j) + G * Diff_K(i, j), G = 0.125 \sim 0.25 \quad (2)$$

For object classification method, we use the building-block based connected component matching method. This method is fast, and there are only one-pass scan and low-cost for objects classification. It does not need four adjacent or eight adjacent traditional connected component rule. It just needs the building-block rule to connect components and classify the objects. The size of building block directly affects the target type which motion detection algorithm extracts. It is defined and decided by users according to detected target type such as cars, human beings or birds.

The building-block-based connected component matching algorithm is designed by the edge analysis of the proportional scale for binary image. It is suitable for the binary image which is captured from the fixed camera; the variation of the edge of binary image vertical project to proportional scale for each row is small, and the building block is easy to compose the binary image. Adjusting the edge variation degree of the building block, it will fit the detection of different target. Fig.1 shows the edge variation of building block. For example, the edge of car has bigger variation than human.

From 720x480 image segmentation, the block resolution is 24x16 by column data and row data accumulation, all these pixel luminance values would be extracted to a feature value. The block-level feature extraction architecture is shown in Fig.2. This architecture makes use of the right shift and rotation 30 of 8-bit registers to accumulate 24 pixels luminance values of each row. After the first row of luminance values accumulation, then we rotate the value to 8-bit addition for the luminance values of the second row of accumulation. After 16 rotations of accumulation, the feature value of the first line of blocks will be calculated.

The feature value of the first block in the frame will be input from the port "Frame1" and saved in 8-bit registers of the first row, then it is shifted right and the feature value of the second block will be input and then saved to 8-bit register again. After the feature values of block size 30x30 in Frame 1 are saved, the feature value of Frame 2 will start

inputting from the port “Frame2” and subtract with the feature value of block in Frame directly by 8-bit FS to get the difference value.

The difference value of the first block rotates to save the first 8-bit registers in the same buffer of the first row. We make use of this proposed reuseable 30×30 Feature Value buffer as circular queues. When all 720×480 pixels of Frame2 are processed, this content of 30×30 buffers is the difference value of two frames. This architecture of the reusable 30×30 8-bit buffers in Fig.3 saves 2/3 of the traditional buffer space.

3. Performance Analysis

In this paper, a fast and cost effective shape perceived motion detector is presented. From the viewpoint of VLSI implementation, the proposed architecture is simple, modular, regular, and cascade-able. The layout of the processor core is shown in Fig.4. The chip can process thirty frames with 720×480 resolution per sec. And the gate count is 10848 in UMC 0.18 technology. It needs 15 input pads, 38 output pads, and 18 power pads. The chip features are shown in Table 1. Its maximum working frequency is 27 MHz. These results show that low-cost motion detector is feasible.

4. Conclusions

This paper presents an effective object tracking algorithm and architecture for advanced intelligent surveillance system. According to the proposed shape perceived architecture, it uses the building-block-based matching method. This method is fast and cost effective for objects classification, only one-pass scan is employed, and it does not need four adjacent or eight adjacent traditional connected component rule. The results show that the low-cost shape perceived motion detector is feasible.

References

- [1] P.J. Burt, and J.R. Bergen, “Object tracking with a moving camera: An application of dynamic motion analysis” in *Proc. IEEE Intl. Conf. Visual Motion*, Page(s):2 – 12, 1989.
- [2] Cedars, C. and M. Shah, “Motion-based recognition: a survey,” *Image and Vision Computing*, vol.13, no.2, pp.129-155, March 1995.
- [3] Wren, C. R., A. Azarbayejani, T. Darrell, and A.P. Pentland, “Pfinder: real-time tracking of the human body,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.780-785, July 1997.
- [4] Makarov, A.; “Comparison of background extraction based intrusion detection algorithms”, in *Proc. Intl Conf Image Process.*, vol. 1, pp. 521 – 524, Sep. 1996.
- [5] G.L. Foresti, “A real-time system for video surveillance of unattended outdoor environments”, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 697–704, Oct. 1998.

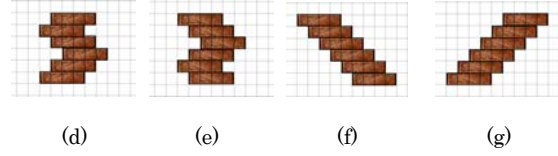
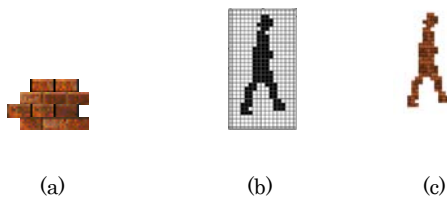


Fig.1 Matching method of building-block based connected components (a) Building blocks (b) proportional scale for binary image (c) connected component with building block. The edge variation of building block (d),(e): Max. Curves of edge variation. (f),(h): Max. Slopes of edge variation.

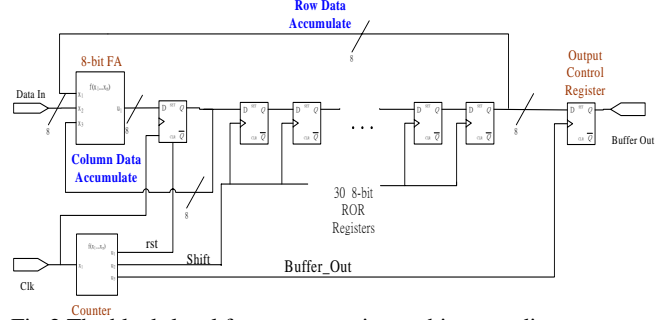


Fig.2 The block-level feature extraction architecture diagram

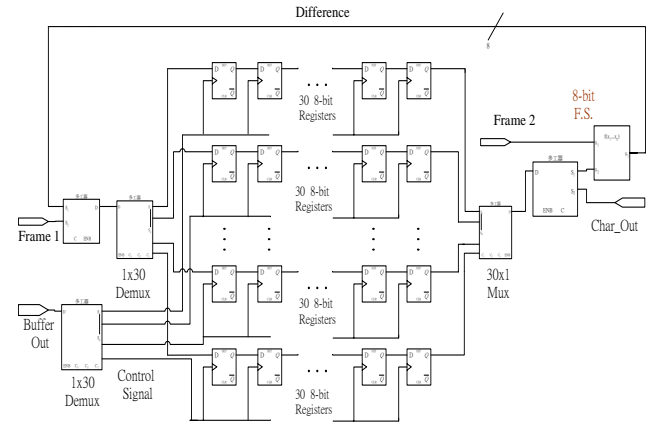


Fig.3 The re-useable feature value buffer architecture diagram

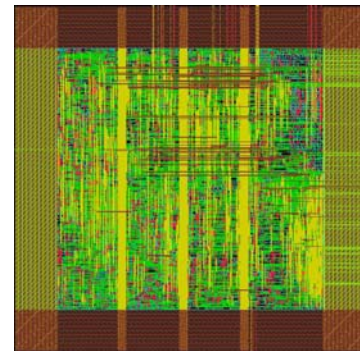


Fig. 4 Chip layout

Table 1. Chip features

Technology	UMC 0.18μm 1P6M
Clock Frequency	27MHZ
Video Process Ability	720x480 NTSC/PAL @ 30fps
Gate Count	10848 (NAND2)
Voltage	1.8V
Power	1.4456 mW