FPGA Implementation of Object-Based Real-Time Object Tracking Architecture

Kousuke Yamaoka, Takashi Morimoto, Hidekazu Adachi, Kazutoshi Awane, Tetsushi Koide and Hans Jürgen Mattausch

Research Center for Nanodevices and Systems, Hiroshima University, Higashi-Hiroshima, Japan
Phone: +81-82-424-6265   Fax: +81-82-424-3499   E-mail: yamaoka@sxsys.hiroshima-u.ac.jp

1. Introduction

Detecting moving objects in a video sequence and object tracking are indispensable technologies for surveillance and recognition systems. Various solutions have already been proposed for video object tracking (e.g. background difference based method [1], optical flow based method, and model-based method). However, each method has some short comings, for example, although the background subtraction method can be realized with relatively simple operations, it cannot be applied to the case of a moving camera. Also the optical flow based method needs complex calculations, so that real-time processing becomes difficult.

To overcome these problems, we have proposed a multi-object tracking algorithm based on image segmentation and simple object-feature matching [2]. We have confirmed the effectiveness of the proposed method which takes advantage of the spatial object information derived from image segmentation results for difficult cases such as rotating objects, partial occlusion between objects, or a moving camera.

In this paper, we introduce an FPGA implementation architecture for the proposed algorithm and report the verification of its effectiveness with our developed FPGA-based object tracking prototype system.

2. Proposed Object Tracking Architecture

Figure 1 shows the overall block diagram of the proposed hardware implementation architecture. This architecture roughly consists of 4 blocks. They are called "*Image Segmentation Block*", "*Object Feature Extraction Block*", "Object Matching Block" and "Estimated Position Calculation Block".

2.1 Image Segmentation Block

The first block is the image segmentation block in which all objects of the frame are extracted. The proposed implementation architecture uses a region-growing-type image segmentation algorithm, namely an image-scan based segmentation architecture which is suitable for VLSI implementation [3]. This architecture reduces hardware cost by dividing the input image into small pixel blocks and processing them sequentially. Figure 2 shows a conceptual diagram. The input image (e.g. 6×10 pixels) is divided into small pixel blocks (e.g. 6×2 pixels) from top to bottom. The pixels in each block are processed in parallel with a small image-segmentation processing array of the block size in sequential scan mode from top to bottom of the image. Between the processing steps of two blocks, the processing results of the finished block are stored in on-chip state memories and the processing status of the next block, obtained in
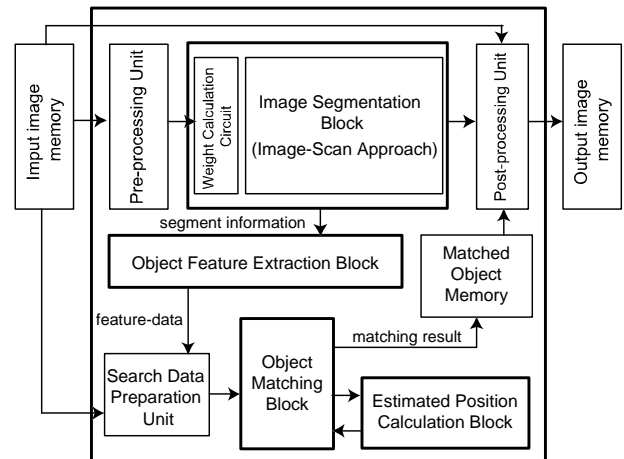

Fig. 1 Block diagram of the proposed hardware implementation architecture.
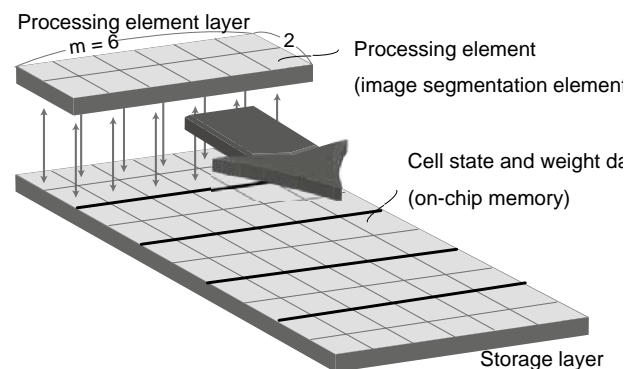

Fig. 2 Conceptual diagram of image-scan video segmentation. High access-bandwidth is needed between the processing element layer and the storage layer for real-time processing.

the previous scan of the image, is loaded into the processing elements of the array. These storing and loading steps require memories with very high access bandwidth. For realizing this requirement, we apply multi-bank embedded memory and give an efficient data mapping. Naturally, the block size is variable, and a trade-off can be exploited to optimize processing time versus hardware amount.

2.2 Object Feature Extraction Block

In this block, object features (position, size, color, area) for each segmented object are calculated using the image segmentation results.

Image Segmentation Block and Object Feature Extraction Block are connected with large access-bandwidth data busses so as to perform feature extraction using all pixel data of the segmented region. After feature calculation, pixel data are converted to a small amount of

feature data. This integration method enables seamless transmission of extracted object features for higher-level image processing such as object recognition.

2.3 Object Matching Block

The most similar object is searched among the reference object data from the previous frame in this block. The object matching circuit has two kinds of memories. One is for storing the object features in the current frame and the other is for storing the object features in the preceding frame as reference object data. Using these memories, all combinations of the currently segmented object and all reference objects in the preceding frame are compared for Manhattan-distance measure one by one. Manhattan-distance is calculated from normalized values of each object feature.

2.4 Estimated Position Calculation Block

For improving the precision of feature matching, we calculate the estimated position of each object in the next frame by using a motion vector.

As mentioned above, due to the sequential nature of the segmentation, we applied pipeline processing to interleave the processing steps of image segmentation and object feature matching for reducing processing time.

3. FPGA-Based Object Tracking Prototype System

For the verification of our proposed object tracking architecture, we have developed an FPGA-based real-time multi-object tracking prototype system and implemented the designed circuits with Verilog-HDL. We show the system overview and its configuration in Fig. 3. A VGA-sized input image from the video camera is resized to 80×60 pixels in the image resize block of the EPF10K250A device and stored in a clock-asynchronous external memory. Three external memories are used for storing 3 successive frame data. After segmentation and tracking processes are performed in the EP1S60 device, pixel and segmented label data are transmitted to the image restore block implemented in the EPF10K250A in combination with the tracking result. Finally, only the intended tracking target is displayed as a blue region. We can make a choice at random by using a push switch controller on EPF10K40.

Due to the limitation of the number of data pins among the FPGA boards, whole circuit blocks are partitioned to three FPGAs. The specification and the FPGA-resource usage of the prototype system are summarized in Table I. We can see that the usage of logic elements and on-chip memory of the main FPGA (EP1S60) is about 54% and 3%, respectively. Consequently, under the right conditions (e.g. pins construction), we can implement the whole circuit in a single FPGA chip. Figure 4 shows tracking results (30fps) with our prototype system. In the case of occurring object occlusion, correct tracking can be realized in real-time.

From the implementation result with 80×60 pixels image size, we can expect to realize real-time multi-object tracking of QVGA size images with a latest generation FPGA-chip (EP2S180 [4]).
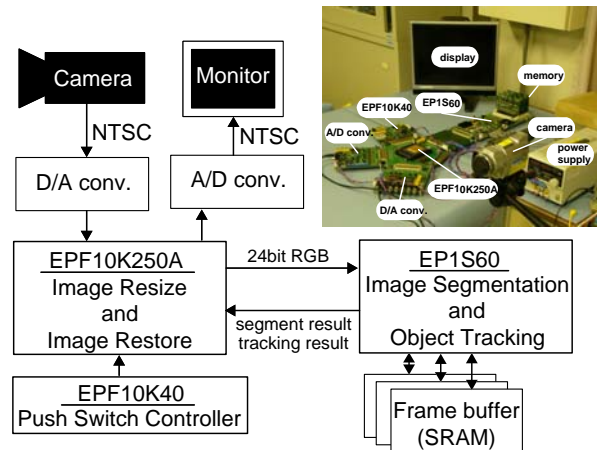

Fig. 3 Overview and structure of FPGA-based object tracking prototype system

Table I: Specification of the Prototype System

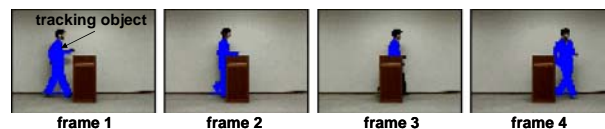| Input/Output Signal		NTSC Y/C Signal
Target Image Size		80×60 pixels
Processing Element Size		80×2 (2 line scan)
Clock Frequency		12.27 MHz
FPGA Devices	EP1S60	LE: 31,008/57,120 (54%) Mem: 0.17/5.2 Mbits (3%)
	EPF10K250A	LE: 157/12,160 (1%)
	EPF10K40	LE: 162/2,304 (7%)

Fig. 4 Example of object tracking results with the developed FPGA system. The tracking object is shown as a blue region

4. Conclusion

We have proposed a multi-object tracking architecture based on image segmentation and object feature matching. For both real-time processing and compact hardware implementation, we applied an image-scan-based segmentation architecture, which efficiently utilizes high access-bandwidth embedded memories. Furthermore, we have developed an FPGA-based object tracking system for verification of our proposed architecture. We could confirm a correct object-tracking result and real-time processing of 30fps for the system.

References

[1] S. Y. Chien et al., "Efficient moving object segmentation algorithm using background registration technique," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12 (7), pp. 577-586, 2002.
[2] T. Morimoto et al., "Object tracking in video pictures based on image segmentation and pattern matching," Proc. of the IEEE International Symposium on Circuits and Systems(ISCAS2005), pp.3215-3218, 2005.
[3] H. Adachi et al., "Image-scan architecture for efficient FPGA/ASIC implementation of video-segmentation by region growing," Proc. of the International SoC Design Conference(ISOCC2005), pp.301-304, 2005.
[4] StratixII, Altera Corporation, 2005, URL: http://www.altera.com/products/devices/stratix2