ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

Real Time Hardware Software Co-Simulation Edge Map Accumulation Based Feature Extraction

A.B. Wagh[1], M.R. Wankhade[2]

Assistant Professor, Dept. of ETC, SKN College of Engineering, Pune, India[1]

PG Student [VLSI& Embedded System], Dept. of ETC, SKN College of Engineering, Pune, India[2]

ABSTRACT: There are many applications like automotive vehicle control, video surveillance, remote gesture control etc. for which high speed image processing employing high frame rate image sensors provide details on motion which is useful for improving the speed and precision of the recognition. Therefore, the proposed system presents that the VLSI is capable for extracting motion features from moving images. The extraction of motion features from moving images is developed by using row parallel and pixel-parallel architectures. This system proposes a novel fast texture feature extraction method which takes advantage of the similarities between the neighbouring pixels to estimate texture values. Here, firstly by using FPGA and MATLAB edge detection and mapping take place.The feature extraction and the filtering process can be done through the MATLAB.

KEYWORDS: Motion-feature-extraction, parallel architecture, FPGA, MATLAB.

I.INTRODUCTION

The motivation behind the system of motion feature extraction is driven because motion recognition is becoming increasingly important in various applications such as video surveillance, automotive vehicle control, efficient computer-human interaction, remote gesture control and so on. High speed image processing having high frame rate image sensors can significantly provide much details of the motion in such applications, which is very useful for improving the speed and precision of the recognition. Motion Feature-Extraction (MFE), is an important and computationally heavy process in motion recognition is required to be processed for building such high speed systems. For performance enhancement in image processing algorithms a number of Very Large Scale Integration (VLSI) processors with parallel architectures have been developed.Image data processing usually includes some computationally intensive low-level data processing such as image filtering, which needs be performed repeatedly at least in the region of interest or in every pixel site in the entire image. Processors in improved the performance and power efficiency by developing parallel processing circuitries mainly for such low-level image processing tasks. Processors in implemented fine-grain parallel processing architectures by investigating image processing tasks at all levels (low, middle, and high levels) in some specific applications. Such processors have made real-time image processing systems feasible Therefore, the parallel architecture is important for improving the performance of image data processing [1].

The traditional software-based MFE algorithms are computationally very expensive, and building high speed systems is quiet difficult even when state-of-the-art multicore general purpose processors are utilized. By developing fine-tuned software exploring the Single Instruction Multiple-Data (SIMD) operations on modern processors and/or Graphics Processing Units (GPUs) for particular applications, high hardware costs and large power consumptions are severely limiting such approaches to be used in portable devices, as well as in large scale systems. As a result, high speed and low power architectures for MFE are now highly demanded [1].

II.LITERATURE SURVEY

Number of Very Large Scale Integration (VLSI) processors with parallel architectures has been developed for performance enhancement in image processing algorithms. Image data processing usually includes some computationally intensive low-level data processing such as image filtering, which needs be performed repeatedly in every pixel site in the entire image, or at least in the region of interest. Nowadays, motion recognition is becoming increasingly important in various applications such as automotive vehicle control, efficient human-computer interaction, video surveillance, remote gesture control and so on. For such applications, high speed image processing employing high frame rate image sensors can significantly provide much details of the motion, which is very useful for improving the speed and precision of the recognition. For building such high speed systems, Motion Feature-Extraction (MFE), as an important and computationally heavy process in motion recognition is required to be processed. [1]

T. Shanableh, K. Assaleh and M. Al-Rousan described Spatio-Temporal Feature-Extraction Techniques in which there are two types of feature extraction. First is Temporal Feature Extractions and second is Spatial-Domain Feature Extractions. In Temporal Feature Extractions the motion of the temporal sequence can be captured by removing temporal domain redundancies. The motion can be accumulated into one image that represents the activity of the whole temporal sequence. Temporal-domain redundancy reduction techniques are well established in the video compression literature. Video coders employ motion estimation and motion compensation prediction on blocks of pixels referred to as macroblocks to reduce the energy of prediction error. The outcome of the motion estimation process is a 2-D motion vector which represents the relative displacement of a macroblock relative to anchor picture or a reference. The motion compensation prediction subtracts the macroblocks of the current picture from the best matched location of the anchor picture as indicated by the relevant motion vector. Content-based querying of video databases utilizes such motion vectors for video indexing and retrieval. [2]

The extraction of such descriptors is facilitated by the syntax of the coded video stream, and it includes the needed motion information. Thus, neither further motion estimation nor compensation is needed in this case. However, in the proposed temporal motion activity extraction for SLR, computation of motion vectors is computationally prohibitive. Thus, it propose to capture the motion activity by examining the forward prediction error of successive pictures without motion compensation **.**Spatial-Domain Feature Extractions having reduced the temporal sequence of a given gesture into an image of accumulated differences, we now turn our attention to spatial-domain feature extractions. This section proposes two different approaches namely: Radon transformation followed by low-pass filtering and 2-D transformation followed by Zonal coding [2]

H. Yamasaki and T. Shibata developed an image-feature-extraction and vector-generation VLSI aiming at building real-time recognition systems. By employing arrayed-shift-register architecture, this system achieved scanning of the recognition window. It described that Directional-edge-based feature vector generation algorithm is used in which the feature vector generation is done by the Projected Principal-Edge Distribution (PPED) algorithm. Here, Arrayed-Shift-Register Architecture is used for Seamless Vector Formation. Seamless feature vector generation coherent to the continuous scanning movement of the recognition window is essential in recognition in a large scenery image and carrying out objects search In the edge-based image vector generation, like PPED, determining the threshold value for edge detection adaptive to local luminance variation is of paramount importance to perform robust image recognition. The median filter is known as a very powerful rank-order filter but is computationally very expensive filter and making threshold determination the bottleneck of this system. The binary search algorithm is a bit-comparison-based technique. This is quite compatible to direct hardware implementation. Since median filters are easily built using majority voting circuits (MVCs), less chip area is consumed as compared to word-comparison-based sorting circuits. The compact implementation is important when a large number of inputs are handled. [3]

Sigal Berman and Helman Stern developed A gesture recognition system (GRS) which is comprised of a gesture, gesture-capture device (sensor), feature extraction, tracking algorithm (for motion capture), and classification algorithm. It is important to examine the first apparatus that separates the human communicator and the device being controlled with the impending movement toward natural communication with mechanical and software systems although there are, a comprehensive analysis of the integration of sensors into GRSs and their impact on system performance is lacking in the professional literature and numerous reviews of GRSs. Determination of the sensor stimulus, sensor platform and context of use are major preliminary design issues in GRSs. Therefore, these three components form the basic structure of our taxonomy. This system emphasizes the relationship between these critical components and the design of the GRS in terms of its architectural functions and computational requirements. In this treatise, consider sensors that are capable of capturing dynamic and static arm and hand gestures. Although here discuss various sensor types, our main focus is on visual sensors as expected these to become the sensor of choice in the foreseeable future. [4]

W. Zhang, Q. Fu, and N. J. Wu proposes a novel programmable vision chip based on multiple levels of parallel processors. This chip integrates an embedded microprocessor unit (MPU) and, multiple-levels of SIMD parallel processors and CMOS image sensor. The multiple-levels of SIMD parallel processors consist of an array processor of SIMD processing elements (PEs) and a column of SIMD row processors (RPs). The RPs and PE array can be reconfigured to handle algorithms with different processing speeds and complexities. The PE array, RPs and MPU can execute low-level, mid-level and high-level image processing algorithms, respectively which efficiently increase the performance of the chip. This chip can satisfy flexibly the needs of different vision applications such as complicated feature extraction, image pre-processing, and over 1000 fps high-speed image capture. These applications including image recognition, pattern extraction and target tracking are demonstrated. It proposes a digital general-purpose programmable vision chip based on multiple levels of Parallel Processors. The proposed vision chip has better flexibility, scalability and performance compared with the previous vision chips, scalability and performance. The operation performance increases linearly with the number of PEs. One weakness of the chip is its inefficiency in performing multiplications. [5]

III. PROPOSED ARCHITECTURE

The above block diagram is the proposed system for edge mapping and edge detection which explains the hardware co simulation for edge mapping and edge detection. By using FPGA and MATLAB edge detection and mapping take place. In this process firstly video frames are captured through the MATLAB and this video frames are captured via stored video or webcam. And after that this video frames are passed to share memory. During this, Image will be stored to FPGA memory and MATLAB Tool will manage automatically. Then input image data is read by using data read controller block and after that this image data will be passed to edge detection block. In edge detection filter block four different filters will work simultaneously. These four filters are the 0/45/-45/90 degree filters. Then output of these filters will be given to the next block that is Gradient Selection and Thresholding block. Thus all the output of filters will be threshold. Now all edge map images will be combined as shown in above fig. by using edge map merging block and store in to memory. Now Image will be read from FPGA memory and MATLAB Tool will manage automatically. And finally both input and output that is edge map frames will be displayed in real time in MATLAB.

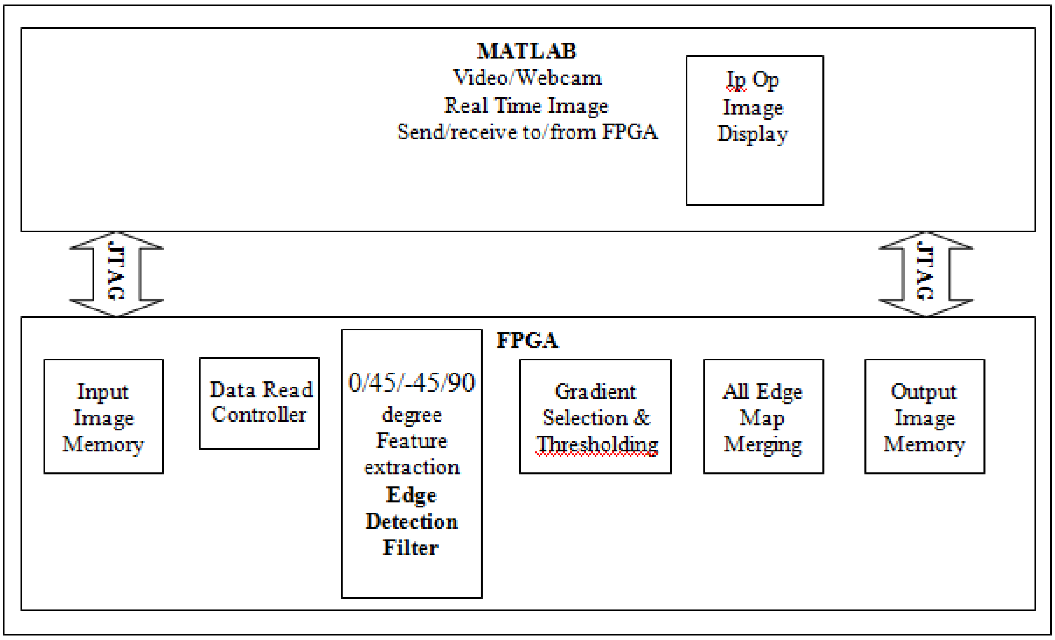

Fig. 1: Proposed Block Diagram for edge detection

The complete flow of the given system is summarising as shown in figure below.

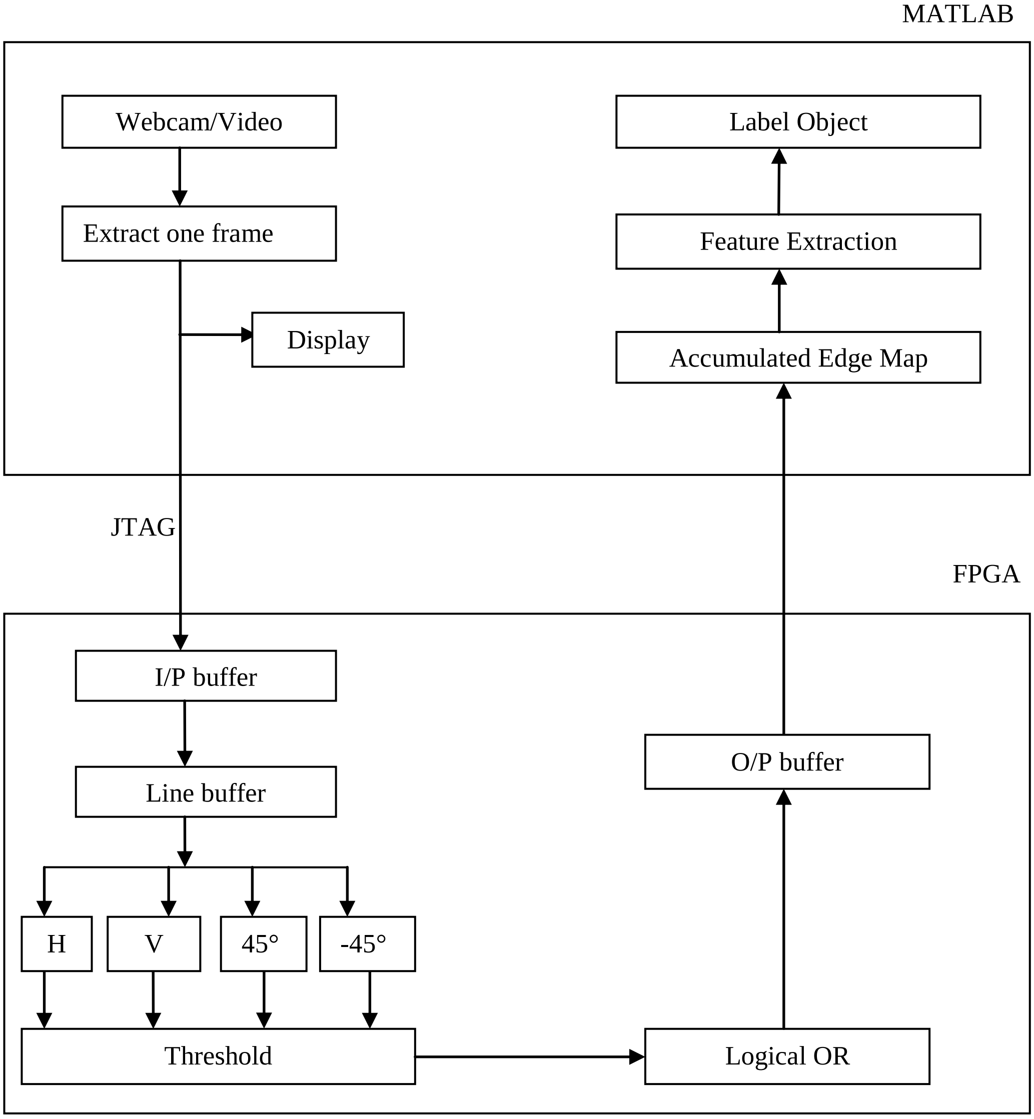

Fig.2: Proposed Architecture of the system

The Merged Significant Edge Map generation (MSEM) is used for the feature extraction.. Firstly from an input image four Significant Edge Maps are generated which is based on the four directional edge filtering (horizontal, vertical ,+45◦ and −45◦). Then by taking logical OR, the four SEMs are merged into a single edge map, which is called as MSEM. There are two main processes for four directional edge filtering: Global Feature Extraction (GFE) and the Local Feature Extraction (LFE) .These two processes where GFE is shown at the top and LFE is shown in the middle. In LFE firstly, convolutions between a 5×5 pixel local image centered at each pixel site and four 5×5-pixel filtering kernels (one for each direction) are calculated. Then based on the four convolution results, the maximum gradient value is selected to determine the edge direction at the pixel site and this convolution value is preserved. According to the edge direction at each pixel site, an edge flag bit “1” is set at the corresponding location in the respective binary edge map. For example, if the direction of a pixel is horizontal, then, this pixel has an edge flag in the horizontal edge map and no edge flag in the +45◦, vertical or −45◦ edge map. Such a process is repeated for every pixel, making the entire image seamlessly scanned by the four filtering kernels. Four directional kernels produce four binary edge maps which are called as local features. Since each pixel has an edge flag in one of the four binary edge maps after LFE, local features contain a lot of redundant information.

Therefore, after LFE, GFE is performed to retain only salient features. This is completed by selecting only a pre-determined percentage which is also called as the edge detecting threshold of significant edge flags out of all pixels that have larger gradient values than the rest. This type of selection is possible because convolution values are all preserved which obtained from every pixel site. The four edge maps having only the significant flags after the selection very well represent global features and we call them the "Significant Edge Maps (SEMs)".Only the salient features in the original image are highlighted. The SEMs are used for both static image recognition and motion analysis. For static image recognition, a feature vector representation algorithm called Projected Principal Edge Distribution (PPED) or Averaged Principal-Edge Distribution (APED) is employed to transform the four directional SEMs into a single 64-dimension feature vector. For motion analysis, the four directional SEMs are merged into one MSEM.

IV.CONCLUSION

The proposed system is the hardware solution to the edge detection and feature extraction of moving images. This system proposes a novel fast texture feature extraction method which takes advantage of the similarities between the neighbouring pixels to estimate texture values. Firstly system has demonstrated that edge detection and edge mapping is done by using hardware co simulation in between FPGA and MATLAB. Merged Significant Edge Map Generation is used for various motion feature extraction and then simulated by using MATLAB tool.

REFERENCES

[1] H. Zhu and T. Shibata, Member, IEEE "A Real-Time Motion-Feature-Extraction VLSI Employing Digital-Pixel-Sensor-Based Parallel Architecture" IEEE Transactions on Circuits and Systems for Video Technology, pp. 1-13 , 2014.

[2] T. Shanableh, K. Assaleh and M. Al-Rousan, "Spatio-Temporal Feature- Extraction Techniques for Isolated Gesture Recognition in Arabic Sign Language," IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 37, no. 3, pp. 641-650, Jun. 2007

[3] H. Yamasaki and T. Shibata, "A Real-Time Image-Feature-Extraction and Vector Generation VLSI Employing Arrayed-Shift-Register Architecture," IEEE J. Solid State Circuits, vol. 42, no. 9, pp. 2046-2053, Sep. 2007

[4] S. Berman and H Stern,"Sensors for Gesture Recognition Systems, IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 42, no. 3, pp. 277-290, May 2012.

[5] W. Zhang, Q. Fu, and N. -J. Wu, "A Programmable Vision Chip Based on Multiple Levels of Parallel Processors," IEEE J. Solid-State Circuits,vol. 46, no. 9, pp. 2132-2147, Sep. 2011

[6] TK. Ito, B. Tongprasit and T. Shibata, "A Computational Digital Pixel Sensor Featuring Block-Readout Architecture for On-Chip Image Processing," IEEE Trans. Circuits Syst.-I, Vol. 56, No. 1, pp. 114-123, 2009

[7] K. Fujita,K. Ito and T.Shibata"A single-motion-vectorcycle-generation optical flow processor employing directional-edge histogram matching," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2009), pp. 3022-3025, May. 2009.

[8] H. Zhu, P Zhao, and T. Shibata, "Directional-edge-based object tracking employing on-line learning and regeneration of multiple candidate locations," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2010), pp. 2630-2633, May. 2010

[9] Xilinx, System Generator for DSP Getting Started Guide, UG639 (v 14.2) July 25, 2012.