FPGA IMPLEMENTATION OF SPATIO TEMPORAL MARKOV RANDOM FIELD MODEL FOR REAL TIME HIGH DEFINITION VIDEO SEGMENTATION

M.Durka Devi
PG SCHOLAR
J.P COLLEGE OF ENGG.

Mrs.M.Shenbagavalli.M.TECH
DEPT.OF HOD/ECE
J.P COLLEGE OF ENGG.

Abstract:

Background identification is a common feature in many video processing systems. This paper proposes two hardware implementations of the OpenCV version of the Spatio temporal markov random field model (STMRF), a background identification algorithm. The implemented version of the algorithm allows a fast initialization of the background model while an innovative, hardware-oriented, formulation of the STMRF equations makes the proposed circuits able to perform real-time background identification on high definition (HD) video sequences with frame size 1920 × 1080. The first of the two circuits is designed with commercial field-programmable gate-array (FPGA) devices as target. When implemented on Virtex6 vlx75t, the proposed circuit process 91 HD fps (frames per second) and uses 3% of FPGA logic resources. The second circuit is oriented to the implementation in UMC-90 nm CMOS standard cell technology, and is proposed in two versions. Both versions can process at a frame rate higher than 60 HD fps. The first version uses the constant voltage scaling technique to provide a low power implementation. It provides silicon area occupation of 28847 μm2 and energy dissipation per pixel of 15.3 pJ/pixel. The second version is designed to reduce silicon area utilization and occupie 21847 μm2 with an energy dissipation of 49.4 pJ/pixel.

Index Terms—Application-specific integrated circuits (ASICs), computer vision, field programmable gate arrays (FPGAs), image motion analysis, object detection, subtraction techniques.

I. INTRODUCTION

The real-time detection of moving objects in a video sequence has many important applications, such as video surveillance and traffic monitoring, [1]–[4]. Various algorithms for the detection of moving objects have been developed during the years, [5]– [18]. A recent review on the topic is presented in [19]
.

Temporal difference techniques, [5]–[7], allow the detection of the dynamic parts of a scene(foreground, Fg) by comparing two consecutive frames.Background subtraction algorithms, [8]–[18], detect moving objects by analyzing the difference between a frame and a reference model that includes the static parts of the scene (background, Bg). Target of these algorithms is the determination of the Bg model that also needs to be updated to follow the changes of the scene.
As an example, in order to adapt the Bg model, [9] and [10] used a Kalmann filter, [11] used a Wiener filter, while [12] proposed an algorithm based on median filtering. These algorithms take into account changes in the lighting of the Bg and the eventuality of moving objects which become statics.

References [13]–[18] proposed Bg identification algorithms based on statistical models. In [14], an algorithm based on a static computational color model, which separates the brightness from the chromaticity is proposed.
The algorithm of [14] is extended in [15] where the shadow detection is included and the implementation on an FPGA platform is also proposed. In [16], the Bg is represented with a statistical model based on a single stmrf model per pixel. This allows a pixel by- pixel model of the Bg with larger variance for pixels that experience wide changes of lighting.

The algorithm proposed in [17] uses a statistical model similar to the algorithm of [16] but models each pixel with STMRF. The algorithm proposed in [17] is known as Spatio Temporal Markov Random Field model (STMRF), and provides good erformances in both presence of illumination changes and multimodal Bg. A multimodal Bg is characterized by objects showing repetitive motions, e.g., waves, moving leaves, or flickering light. When a pixel lies in a region where a repetitive motion occurs, its brightness oscillates between two or more values. This results in false Fg detection in most algorithms. On the contrary, when the STMRF algorithm is used, the intensity distribution of the pixel is modeled using two or more STMRF models and the problem of false Fg detection is solved.

Due to the good performances, the STMRF models has been selected as the Bg detection algorithm in the

OpenCV, Open source Computer Vision software library [20], developed by Intel. OpenCV provides a common base of computer vision instruments able to extract relevant details from the images and to process them in automatic way. The STMRF models proposed in the OpenCV library is an optimized version of the algorithm of [17]. It changes the initial learning phase for the Bg model, that is, significantly improved with respect to [17], and calculates the learning rate in a simplified yet efficient way.

In order to generate the updated Bg model, the STMRF model processes the video streams by computing a great number of parameters for each pixel of each frame, with a computational burden unreachable by computers in real-time applications. As example, in [17], only a frame rates of 11–13 fps is obtained for frame size 160 × 120 on an SGI O2 workstation. The authors of [25] conducted a test on a software implementation running on an AMD4400+processor and observed a frame rate of 4–6 fps for video sequences with 352 × 288 resolution.

Real-time video applications with larger frame size require dedicated hardware architectures. Hardware processors have been proposed in [21]–[28]. Papers [21] and [22] present GPU implementations based on GMM algorithm. Despite the fact that the approach described in [22] reaches 30 fps for HD videos, being a GPU implementation is an impediment for embedded systems and low power constraints.

Papers [23]– [28] propose FPGA implementations that are more suited to embedded and low power systems.

Reference [23] proposed a circuit able to process video sequences with frame size 1024 × 1024 at 38 fps when implemented on VirtexII FPGA platform. The processing capability of [23] is, therefore, 39.8 Mps (Mega pixels per second). The segmentation circuit of [23] is improved in [24] and [25]. The segmentation unit of the circuit proposed in [24] is able to run at 83 MHz on VirtexII xc2pro30. With respect to [23], the circuits of [24] and [25] improve the memory throughput employing a memory reduction scheme and propose the design of an automated digital surveillance system running in real-time on an embedded platform. However, the resulting processing capability of the overall system is reduced. In [25], the processing capability only reaches 7.68 Mps.

Reference [26] proposed a OpenCV STMRF model implementation able to process 22 HD (1920 × 1080 pixels) fps when implemented on Virtex5 xc5vlx50 FPGA. The circuit of [26] is improved in [27] where it is used in conjunction with a denoising block able to implement the operations of erosion, dilation, opening, and closing. The improvement in performance of the Bg segmentation unit of [27] with respect to [26] is mainly due to the substitution of the multipliers of the updating equations with shifters. The circuit of [27] is able to process 24 HD fps (50.5 Mps) on the same FPGA of [26].

This paper proposes two hardware implementations of the STMRF model that allow real-time processing of HD videos and comply with the OpenCV algorithm. The proposed Bg identification circuits process gray scale videos. When color videos need to be processed, the circuits, as shown in Section VII-B and reported in [25], can be fed with the luminance channel of the YCrCb color space.

The first circuit is optimized for an FPGA implementation while the second one is targeted to the standard cell implementation in UMC-90-nm technology. Both circuits exploit an innovative formulation of the algorithm that allows to obtain high performances with low area utilization.

The FPGA implementation, described in Section IV, processes one pixel per clock cycle and works at 189.3 MHz, when implemented on Virtex6.

The ASIC implementation is proposed in two versions, both able to process 60 HD fps. The first ASIC version uses the constant voltage scaling technique to reduce the energy dissipation per pixel and, with a supply voltage of 0.56 V, has a silicon area utilization of 28847 μm2 and energy dissipation of 15.3 pJ/pixel. The second ASIC version of the circuit uses folding technique and requires three clock cycles to process one pixel.

The resulting silicon area utilization of 21847 μm2 with an energy dissipation of 49.4 pJ/pixel.
The proposed circuit has been experimentally validated through experimental measurements of the working frequency and implementing two running on-line video systems equipped with monitor and video sensor.

Main contributions of this paper are the following.

1) An innovative, hardware-oriented, formulation of the STMRF equations that allows hardware saving and speed improvement without affecting the output of the STMRF model.

2) The implementation of the above cited STMRF equations in an FPGA-oriented circuit that outperforms previously proposed circuits.

3) The ASIC standard cell implementation of the proposed Bg identification circuit (actually in two versions that in turn optimize power or silicon area occupation), thus providing a performance reference

for ASIC designers that is still missing in the scientific literature.

4) The experimental demonstration of the proposed FPGA circuit in running on-line video systems.

II. STMRF MODEL

The GMM algorithm has been proposed by Stauffer and Grimson [17], with the target of efficiently dealing with multimodal Bg by using a statistical model composed by a mixture of Gaussian distributions. The GMM algorithm has been modified and included in the OpenCV libraries. A short description of the OpenCV GMM algorithm is given in the following. The differences with respect to the algorithm of [17] are indicated in the text

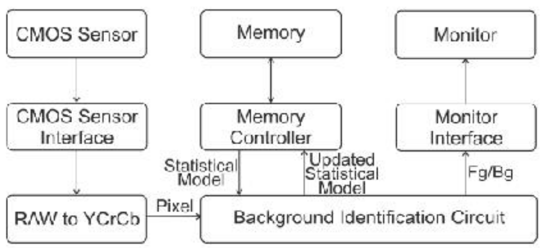


Fig. 1. Conceptual overview of a background identification system.

III. BACKGROUND IDENTIFICATION CIRCUIT

Fig. 1 shows a conceptual overview of a Bg identification system. A CMOS sensor captures each frame of the video sequence and a CMOS interface gives the pixel values to the Bg identification circuit. The Bg identification circuit processes the luminance value of the input “Pixel,” and the “Statistical Model” of the pixel for the given frame.

The output data are the “Updated Statistical Model” and the “Fg/Bg” tag. For each pixel, the Gaussian parameters are read from an external memory and the updated parameters are stored in the memory. For each frame of the input video sequence, the Fg/Bg tags produce a binary image that can be displayed on

a monitor.

In this paper, the “Background Identification Circuit” of Fig. 1 has been implemented on FPGA devices and then by using an UMC-90-nm standard cell library. Fig. 2 shows the block diagram of the FPGA implementation. The ASIC implementation is described in Section VI. Both FPGA and ASIC circuits are described in VHDL code and implement

the GMM algorithm by using, as suggested in [17], three Gaussian distributions for each pixel (k = [1, 2, 3]). As a consequence, the proposed circuits process 13 parameters per pixel (the luminance value and 12 Gaussians parameters for the statistical model of the pixel).

The software algorithm proposed in the OpenCV library represents the Gaussians parameters as double precision (64 bits) floating point numbers. Unfortunately, a hardware implementation that uses 64 bits floating point signals is not feasible.

IV.FPGA IMPLEMENTATION OF STMRF

BG/FG Identification:

- It’s common feature in many video processing systems
- Produce a binary image

If BG/FG have different pixel =>Lower m/y b.w and word length reduction

If BG/FG have same pixel=>higher m/y b.w and word length

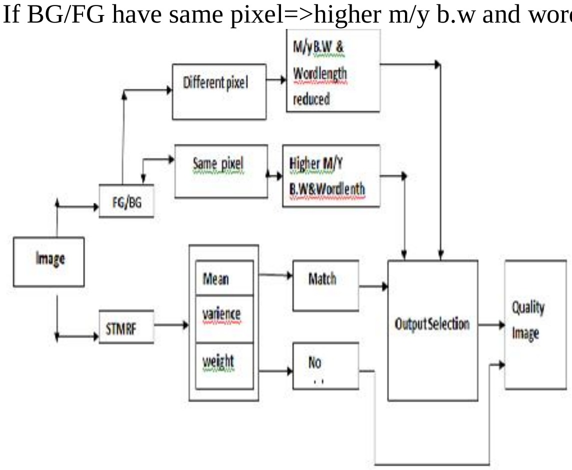


Fig:FPGA Implementation of STMRF

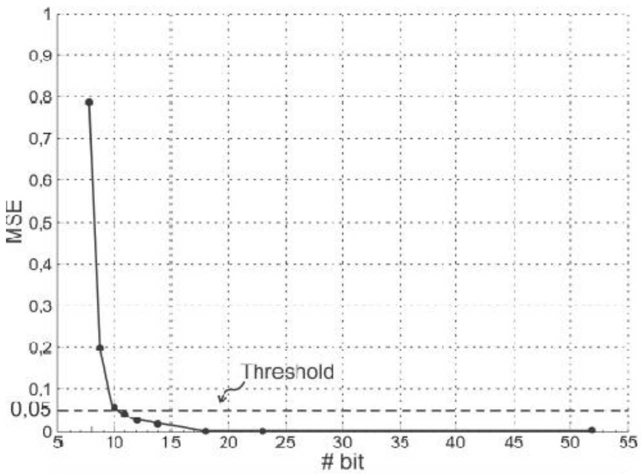
STMRF

The Performance sets Real videos compared with previous algorithm

- The STMRF they give rise to good, flexible image models

Two main issues

1. Estimation of noise model
2. Proper use of motion estimation



# bit	8	9	10	11	12	14	18	23	52
MSE	0.79	0.20	0.06	0.04	0.03	0.02	6E-4	1E-4	1E-4
% correctly	21.40	80.37	94.05	96.13	97.17	98.06	99.94	99.99	99.99

IV. BG IDENTIFICATION CIRCUIT—FPGA IMPLE MENTATION

The schematic diagram of the FPGA implementation of the GMM algorithm is shown in Fig. 2. A detailed explanation of the functionalities and of the hardware implementation of the circuital blocks of Fig. 2 is given in the following Output Selection.

This unit establishes the values of the updated parameters depending on whether the match condition is verified or not.

This paper proposed the FPGA implementation and the ASIC implementation of the OpenCV compatible GMMalgorithm for background identification able to process HD video in real time. Both implementations exploit an innovative, hardware-oriented, formulation of the equations of the GMM algorithm and take advantage of an optimization algorithm to minimize the word length of the signals.

The FPGA implementation when compared with previously proposed background identification circuits, provides improved speed and logic utilization.

CONCLUSION:

Both existing and proposed system are focusing on background identification using GMM. It differs on the resource utilization.GMM algorithm is a subtraction algorithm which compare each frame of video with reference frame. The comparison is mainly depending on the parameters like mean, variance and weight. For comparison and shifting, existing system is using separate comparator and shifter for each and every frame. So large area is occupied by existing system.But in Proposed system only one shifter and comparator is used for all frame. Though the proposed architecture in cooperate theses three processes, the ultimate goal that is background identification is obtained with less power consumption.

REFERENCES

[1] I. Haritaoglu, D. Harwood, and L. S. Davis, "A fast background scene modeling and maintenance for outdoor surveillance," in *Proc. 15th Int. Conf. Pattern Recognit.*, vol. 4. Jan. 2000, pp. 179–183.

[2] P. KaewTrakulPong and R. Bowden, "A real time adaptive visual surveillance system for tracking low-resolution colour targets in dynamically changing scenes," *Image Vis. Comput.*, vol. 21, no. 9, pp. 913–929, Sep. 2003.

[3] B. Gloyer, H. K. Aghajan, K. Y. Siu, and T. Kailath, "Video-based freeway monitoring system using recursive vehicle tracking," *Proc. SPIE*, vol. 2421, pp. 173–180, Mar. 1995.

[4] H. Jieménez and J. Salas, "Temporal templates for detecting the trajectories of moving vehicles," in *Advanced Concepts for Intelligent Vision Systems* (Lecture Notes in Computer Science), vol. 5807. New York, USA: Springer-Verlag, Oct. 2009, pp. 485–493.

[5] Z. Chaohui, D. Xiaohui, X. Shuoyu, S. Zheng, and L. Min, "An
improved moving object detection algorithm based on frame difference and edge detection," in *Proc. 4th Int. Conf. Image Graph.*, Aug. 2007, pp. 519–523.

[6] R. Mingwu and S. Han "A practical method for moving target detection under complex background," *Comput. Eng.*, vol. 31, no 20, pp. 33–34, Oct. 2005.

7] X. Benxian, L. Cheng, C. Hao, Y. Yanfeng, and C. Rongbao, "Moving object detection and recognition based on the frame difference algorithm and moment invariant features," in *Proc. 27th Control Conf.*, Jul. 2008, pp. 578–581.

[8] S. Y. Chien, S. Y. Ma, and L. G. Chen, "Efficient moving object segmentation algorithm using background registration technique," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 7, pp. 577–586, Jul. 2002.

[9] C. Ridder, O. Munkelt, and H. Kirchner, "Adaptive background estimation and foreground detection using Kalman filtering," in *Proc. Int. Conf. Recent Adv. Mech.*, 1995, pp. 193–199.

[10] K. P. Karmann and A. Brandt, "Moving object recognition using an adaptive background memory," in *Time-Varying Image Processing and Moving Object Recognition*, V. Capellini, Ed. Amsterdam, The Netherlands: Elsevier, 1990, pp. 289–307.