J Real-Time Image Proc (2013) 8:389–401
DOI 10.1007/s11554-011-0238-1

SPECIAL ISSUE

FPGA-based architecture for real time segmentation and denoising of HD video

M. Genovese · E. Napoli

Received: 1 February 2011 / Accepted: 14 December 2011 / Published online: 28 December 2011

Abstract The identification of moving objects is a basic step in computer vision. The identification begins with the segmentation and is followed by a denoising phase. This paper proposes the FPGA hardware implementation of segmentation and denoising unit. The segmentation is conducted using the Gaussian mixture model (GMM), a probabilistic method for the segmentation of the background. The denoising is conducted implementing the morphological operators of erosion, dilation, opening and closing. The proposed circuit is optimized to perform real time processing of HD video sequences (1,920 × 1,080 @ 20 fps) when implemented on FPGA devices. The circuit uses an optimized fixed width representation of the data and implements high performance arithmetic circuits. The circuit is implemented on Xilinx and Altera FPGA. Implemented on xc5vlx50 Virtex5 FPGA, it can process 24 fps of an HD video using 1,179 Slice LUTs and 291 Slice Registers; the dynamic power dissipation is 0.46 mW/MHz. Implemented on EP2S15F484C3 StratixII, it provides a maximum working frequency of 44.03 MHz employing 5038 Logic Elements and 7,957 flip flop with a dynamic power dissipation of 4.03 mW/MHz.

Keywords Image motion analysis · Image segmentation · Morphological operations · High definition video · Field programmable gate arrays

M. Genovese (✉) · E. Napoli
DIBET, University of Napoli Federico II, Napoli, Italy
e-mail: mariangela.genovese@unina.it

E. Napoli
e-mail: etnapoli@unina.it

1 Introduction

Motion detection is one of the most important tasks in computer vision. Segmentation algorithms able to find moving objects (Foreground, *Fg*) in video sequences have been developed in the past and electronic systems have been implemented and employed in applications like video surveillance [1–4], and traffic monitoring [5–8].

A segmentation algorithm begins with an identification phase that detects objects supposedly belonging to the *Fg* [9]. The input data of a segmentation algorithm are a video sequence that can be both a grayscale image or a RGB image. The output of the segmentation algorithm is a video sequence composed of binary images in which a pixel is represented with one bit that is equal to '0' if the pixel is classified as background or '1' if the pixel is classified as foreground (or viceversa). The identification is usually followed by a denoising phase that uses morphological operators to remove noise and enhance the appearance of the binary images [10].

The literature reports various contributions for the initial segmentation of the *Fg* pixels [11–22]. Some of the segmentation algorithms, named background subtraction methods, compare the frame with a reference model (background, *Bg*) [14–21]. Target of these algorithms is the determination of the *Bg* model. Further, this model has to be updated to track the background scene.

Chien et al. [14] builds a *Bg* model starting with the assumption that, if a pixel is stationary for a prefixed number of consecutive frames, the probability that it belongs to the *Bg* is high and, therefore, the pixel can be included in the *Bg* model. Ridder et al. [15] and Karmann and Brandt [16] use a Kalmann filter while Toyama et al. [17] employs a Wiener filter. Jacques et al. [18] proposes an algorithm based on median filtering to adapt the *Bg* model.