Optik 168 (2018) 605–618

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

Moving object detection using an adaptive background subtraction method based on block-based structure in dynamic scene

M.Fatih Savaş*, Hüseyin Demirel, Bilgehan Erkal

Karabük University, Faculty of Engineering, Baliklarkayasi Mevkii, 78050, Karabuk, Turkey

ARTICLE INFO

Article history:
Received 10 October 2016
Received in revised form 14 March 2018
Accepted 10 April 2018

Keywords:
Background subtraction
Block based
Adaptive update
Dynamic scene

ABSTRACT

Today, determining background which forms the basis of the video surveillance system whose use is increasing in parallel with developing technology, is a complicated process to implement for cases involving dynamic scene. In this study, we proposed a block-based adaptive method which can be adaptable to dynamic environmental conditions. By grouping the pixels in the picture frame as 2×2 non-overlapped blocks, we reduced the amount of noise and the time delay caused by processing of pixels. By a simple counter structure, we created an adaptive threshold parameter which can be adapted according to the case in N surrounding. We used this parameter to perform an update of the background model and reduced background model normalization process. So, we reduced the update time of background operably. Our proposed method achieved successful results in levels of grey levels.

1. Introduction

Determining moving objects in video sequences is an important process for video surveillance systems. A variety of approaches have been proposed in the literature to overcome the difficulties encountered particularly in dynamic scenes. Since the image in dynamic scenes contains a lot of noise, determining background according to a specific rule is difficult. Models determining background successfully need too much computational load and memory space requirement. These requirements pose problems in structures providing physical solution to the background model [1–3].

In video sequences, background subtraction methods, expressing pixels by various methods, have been proposed. There are three basic approaches for background subtraction [4]. These are pixel-based, block-based and texture-based methods. In video sequences, structures modeling each pixel individually are the pixel-based methods. Wren et al. [5] modeled each pixel as Gaussian distributions with a pixel-based method. However, these methods are inadequate in dynamic scenes. To eliminate the disadvantages of this method, Stauffer and Grimson [6] proposed Gaussian Mixture Model (GMM) modeling more than one Gaussian distribution for one pixel. Varadarajan et al. [7] have developed the GMM model, which is inadequate for dynamic backgrounds, using the spatial relationship of the pixels. However, this method has a problem due to the parameters set initially. For this problem, Chen et al. [8] tried to determine the appropriate parameter values by combining the codebook model with GMM. Kim et al. [9] offered a structure called the codebook for real-time applications in one of

* Corresponding author.
E-mail addresses: 2011538131007@ogrenci.karabuk.edu.tr (M.Fatih Savaş), hdemirel@karabuk.edu.tr (H. Demirel), berkal@karabuk.edu.tr (B. Erkal).

https://doi.org/10.1016/j.ijleo.2018.04.047

the studies aimed at coping with the multi modal backgrounds. In this structure, by quantizing sample background values of each pixel into the codebook, he updated the seconds in certain periods. In this structure, the codes which couldn't be reached for a long time are removed from the code table. To reduce the number of process, by using the color difference. Tu et al. [10] modeled the background with boxed based code book method. Shah et al. [11] designed a self-adaptive CodeBook (SACB) background model to remove the limitations of the CodeBook model. This model is a block-based structure using the close proximity of local neighborhoods. In this method, the exponential smoothing filter is adapted to keep the mean and variance values in order to automatically estimate the brightness boundary threshold and discolorations for each code word. They used YCbCr color space to calculate color distortions effectively and efficiently. For the codebook-based background subtraction algorithm, Zhi Zeng et al. [12] proposed a novel adaptive maintenance scheme. Authors concentrated on two sets of parameters on the model. These were the geometry of the newly added code word and the probability of learning. They set the geometry of the newly added code as the weighted averages of the code words in the code book. Learning probability is determined by the contrast between the misclassified foreground and background color model and the classification results and the spatial-temporal coherence. By testing this method for three test cases, they have improved on the classical codebook structure. Although codebook structure offers an effective solution for dynamic background modeling, it is not fast while modeling. To avoid the parameter determination problem in parametric structures, Elgammal et al. [13] proposed a nonparametric structure. In this method, the background was tried to be created according to the intensity values of pixels stored in the frame. To strengthen the background model in dynamic scenes, Sheikh and Shah [14] also used the spatial values of the pixels. However, in this method, video frames stored in memory to create an effective background model, posed significant problems for real time applications. Daniel Berjón et al. [15] proposed a nonparametric model for a GPU modeling both foreground and background temporally and spatially. The method was enhanced by a particle filter that updates the position of the reference data and reduced the processing time.

Zhu and Zeng [16] created the background $3 \times 3xk$ structure non-parametrically with the YCbCr color model, instead of storing each pixel in N size in memory, assigning each pixel a k-dimensional buffer. However, this method is not sensitive to situations similar to the foreground. Nonaka et al. [17] integrated background modeling based on spatial-temporal features by integrating pixel, region and frame-level background models and presented a powerful background model for changing environmental conditions. Jiang and Lu [18] proposed a variable weighted background model (WeSamBe) with very few examples and implemented an adaptive feedback technique to adapt the model to more demanding conditions.

Block-based methods are modeling structures which also consider the values of pixels' neighboring in video sequences [4]. These methods bring successful results especially in the dynamic scene. Ke et al. [19] used a block-based structure to determine the motion in scene by considering the computational redundancy. Experimental studies also showed that this method gives better results than pixel-based method. When creating the background model, not only did Jiang et al. [20] take the temporal dimension of the pixels into consideration but also they took the spatial neighborhood into consideration at the same time. They combined neighborhood values of the pixels by averaging them and used it for the model. This method can identify objects more precisely. Sajid and Cheung [21] proposed a universal background model for the background model at both pixel and mega pixel levels using multiple background models. In changing lighting conditions, they used RGB and YCbCr color spaces to identify the foreground better. They eliminated the necessity of adjusting the parameter with the updating mechanism dependent to the foreground.

Reddy et al. [22] proposed a block-based structure in order to cope with dynamic background, noise and light changes. They divided video frames into blocks that overlap and they formed a compact definition for each block with 2D Discrete Cosine Transform. They achieved good performance results in small block structures. Yang and Zou [23] used block-based RPCA (Robust Principal Component Analysis) in order to identify the foreground objects better. They defined algorithm in three parts to determine the foreground. These are first classification, detailed classification and the last processes step. In detailed classification, they solved the IALM (inexact augmented lagrangian multipliers) and RPCA problems by creating block based RPCA. In this study, they obtained better results when compared RPCA.

Heikkilä and Pietikäinen [24,25] proposed a texture-based background construction method using LBP (Local Binary Patterns).In this method, the video frame was divided into windows in certain sizes and a certain pattern was tried to be created by comparing the value of each pixel with neighborhood values. This method is much more successful in the non-dynamic background scene. A variety of hybrid methods have been proposed to resolve the weaknesses of this model. In LBP model, Yeh et al. [26] divided the frame into non-overlapped blocks to reduce the noise problem turned out due to texture structure based on center-pixel. They proposed a new hierarchical block-based coarse fine texture structure by taking the average of these blocks. Because the blocks have different complexity, they expressed blocks by k-bit mode. This proposed structure is tolerant to light changes and has less computational load. Kalpana and Jyoti [27] modeled the background model of a pixel with the histogram structure using the frame in LBP feature. They calculated the rate of learning of the model by evaluating how much the data point contributed to the model histogram. However, this model is not suitable for situations involving grey shadows.

2. Review of modify the original KDE method

KDE (Kernel Density Estimation) method, one of the most popular background subtraction methods, was proposed by Elgammal et al. [13] By using the past N pieces of the density value of a pixel, its density value at t time is estimated by nonparametric. In this method, controlled learning speed and model is calculated by updating sequentially, by using the