Signal Processing 96 (2014) 138–152

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Efficient modified directional lifting-based discrete wavelet transform for moving object detection

Chih-Hsien Hsia [a,*], Jing-Ming Guo [b]

[a] Department of Electrical Engineering, Chinese Culture University, Taipei City, Taiwan
[b] Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei City, Taiwan

ARTICLE INFO

Article history:
Received 3 December 2012
Received in revised form
17 May 2013
Accepted 9 September 2013
Available online 16 September 2013

Keywords:
Fine shape information
Modified directional lifting-based discrete wavelet transform
Moving object detection
Discrete wavelet transform

ABSTRACT

Moving object detection is a fundamental task on intelligent video surveillance systems, because it provides a focus of attention for further investigation. Thus, video object segmentation, which extracts the shape information of moving object from a video sequence, is a key operation for surveillance system. In this study, the current state-of-the-art in moving objects segmentation for intelligent video surveillance has been surveyed. An efficient modified directional lifting-based 9/7 discrete wavelet transform (MDLDWT) structure is proposed to further reduce the computational cost and preserve the fine shape information in low resolution image. Although perfect moving object detection in a practical environment is a challenging task due to the vague object shape issues in the low resolution configuration, the experimental results document that the proposed low-complexity MDLDWT scheme can provide more precise detection rate for multiple moving objects, and the fine shape information can be effectively preserved for the real-time video surveillance applications in both indoor and outdoor environments.

1. Introduction

A research issue in computer vision is how to efficiently detect moving objects in video sequences. More and more researches intent to develop intelligent video surveillance systems to replace the traditional passive video surveillance systems. The intelligent video surveillance system can detect moving objects in the early stage and subsequently deliver the corresponding possible value-aided functions such as object classification, object tracking, and object behavior descriptions. Thus, detecting moving object is a crucial step in computer vision.

To perfectly detect moving object in practical environment is a difficult task, since various issues are involved such as illumination changes, fake motion [1], crowded video scenes [2], and Gaussian noise in the background [3]. Due to the dynamic environmental condition such as the moving leaves of trees as part of the background. Detecting of moving objects in presence of litter background like leaves movement of trees and change of illumination in video sequences is a challenging issue. The unstationary background is often considered as a fake motion in contrast to the motion of the object of interest, which can cause the failure detection of the object. Because of these little movements within the background, it affects the performance of the automated detecting system. Thus, the fake motion issue is a critical task for a robust video surveillance system. Normally, there are three options for motion detection [4–6], namely background subtraction, temporal differencing, and optical flow. The background subtraction method detects moving regions between the current and the reference background frame. It provides the most complete motion mask data, yet it is vulnerable to handle dynamic scene changes due to lighting and extraneous events. Consequently, reference background

* Corresponding author. Tel.: +886 228610511x33234.
E-mail addresses: chhsia625@gmail.com (C.-H. Hsia), jmguo@seed.net.tw (J.-M. Guo).

0165-1684/$ - see front matter
http://dx.doi.org/10.1016/j.sigpro.2013.09.007

has to be updated frequently. The temporal differencing approach extracts the moving region using consecutive frames of the image sequences. It is suited for dynamic environment, yet it often extracts irrelevant motion object pixels. The optical flow method uses the characteristics of flow vectors of moving objects over time to detect moving regions. Yet, most of the optical flow methods are geared with higher complex computation.

To solve the aforementioned problems, several approaches for object detection were proposed. Although video detection systems can deal with input objects of various shape and size, the computational complexity is normally huge. For this, the size of the image can be reduced to ease the computational burden, and which can also remove part of the noises as well. Cheng et al. [1] used discrete wavelet transform (DWT) as a pre-processing to detect and track moving objects. The 2-D DWT was used to decompose an image into four subband images (LL, LH, HL, and HH). The LL_3-band is employed for further detection due to the consideration of low computing cost and noise reduction issues. Yet, the used Haar-based DWT leads to blur object shape. Bennett et al. [7] enhanced the underexposed low dynamic range videos using a dynamic function for the neighboring pixel values. This motion detection method can detect moving objects accurately, yet it suffers from high computing cost. Alsaqre et al. [8] used a local preprocessing method to reduce the noises and other small fluctuations from the image. However, the computational complexity induced from the post-processing is still an issue. Sugandi et al. [9,10] proposed a method for detecting and tracking objects using a low resolution image with the 5×5 Average Filter (AF), the image is generated by replacing each pixel value with the average value of its neighbors and itself. It was claimed that the low resolution image is insensitive to illumination changes and is able to reduce the small movement such as moving leaves of trees in the background. Yet, the low resolution images become more blurred than that of the LL-band image generated by the DWT. Huang et al. [11,12] enhanced the shape of the moving object from the frame difference of two consecutive frames. The change detection using inter-frame differences is a popular method for obtaining the shape information of a video object, since it is easy to implement in real-time systems. In [13], the background estimation system discriminates the foreground objects from the background by building the background model from images. It shows that a local contrast enhancement applied prior to Down-Sampling (DS) improves detection sensitivity. Tian et al. [14] presents a real-time algorithm to detect salient motion in complex environments by combining temporal difference and a temporal filtered motion field. To speed up the whole system using the DS instead of using the whole image (320×240), although less complex, resizing to a lower resolution can change the appearance of an image. In particular, the DS may cause an image blur. In addition, the edge information was employed for object tracking applications in recent years. The local edge binary pattern and the Garber filter are modified into local maxima edge binary pattern by [15] for image retrieval and object tracking. The directional line edges of neighbors for a particular center pixel are determined by eight directional windows. It extracts the information from images using maximum edges, calculated by the magnitude of local difference between the center pixel and its eight neighbors. Yet, if high resolution images for surveillance system will increase the computing complexity significantly and need larger size of memory in the future. It is a challenge for a software platform achieving real-time performance with high resolution images.

To detect the moving object more accurately, we propose a new approach, modified directional lifting-based 9/7 discrete wavelet transform (MDLDWT), which is based on the coefficient of lifting-based 9/7 discrete wavelet transform (LDWT). Clear shape information of moving objects may not be available from multiple-level decomposition image (such as LL_3). To overcome the aforementioned issues, the proposed method can preserve the shape of objects in the low resolution image. Thus, the MDLDWT is presented in this study for detecting foreground moving objects in spatial domain. It not only retains the features of the flexibilities for multiresolution, but also yields low computing cost when it is applied for LL-band images. The proposed method has the advantages of low critical path (the longest, time-weighted sequence of events from the start of the program to its termination), less multipliers/adder, and fast computational speed. Moreover, the LL_3-band of the MDLDWT is employed solely to reduce the image transform computing cost and remove noise. In addition, it preserves better image quality than that of the other approaches. As documented in the experimental results, the proposed method can better retain slow motion of objects than that of the former low resolution methods [1,9–14], and provide effective detection performance.

The rest of this paper is organized as follows. In Section 2, the general flow for detecting moving objects using the low resolution method is introduced. The proposed algorithm and the efficient model for moving object detection are presented in Section 3. Section 4 shows the experimental results, and the conclusions are drawn in Section 5.

2. Moving object detection techniques for sub-sampled images

Many practical video recording systems are with limited resources which may lead to poor image resolution. On the other hand, due to the imperfection of video acquisition systems and transmission channels, images are often corrupted by noises. Consequently, the issues lead to a significant reduction in image quality, which may jeopardize the further high-level computer vision applications, such as object tracking and recognition. Prior to the stage of motion object detection, several former methods [1,9–14] were presented in the past several years for removing noises and reducing computing cost for low resolution images. These are briefly introduced in the following sub-sections.

2.1. Down-sample (DS) scheme

In image processing, the DS is the process of reducing the sampling rate of a signal [3], and which is usually done to reduce the size of the signal. DS is the reduction in