Optik 157 (2018) 787–797

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

Contour-based object tracking in video scenes through optical flow and gabor features

S. Kanagamalliga*, S. Vasuki

Velammal College of Engineering and Technology, ECE Department, Viraganoor, Madurai, 625009, India

ARTICLE INFO

Article history:
Received 1 February 2017
Accepted 22 November 2017

Keywords:
Video surveillance
Background subtraction
Motion estimation
Feature extraction
Object classification
Occlusion detection
Object tracking

ABSTRACT

While many algorithms have been proposed for object tracking with demonstrated success, a crucial problem still persists which is to improve the performance of non-rigid object structures. This paper presents a new, efficient algorithm for Movement Estimation and object tracking in video scenes using Optical Flow and Gabor Features Based Contour Model. The target motion detection is done based on optical flow method to calculate the flow field, according to the optical flow distribution characteristics. Once the flow field has been determined it is used for motion analysis and the Expectation Maximization Based Effective Gaussian Mixture Model (EMEGMM) algorithm based background subtraction is performed to obtain the foreground pixels. With this method complete motion, shape and Gabor features are estimated. The extracted features are classified using Adaboost classifier for effectively handling the region of interest. Then contour based object tracking is carried out by locating the object region in every frame through the object model created by the previous frames. The object shapes are considered as boundary silhouettes and the tracking results obtained are updated dynamically in the video frames. Experimental outcomes validate that our proposed method runs faster and is more accurate, when compared to the several state-of-the-art tracking methods.

1. Introduction

Video processing is a particular case of signal processing where the input and output of video surveillance systems will be a video format or video streams [4]. The major three steps involved in video processing are detection, classification and tracking.

Most commonly used methods for detection of interested moving objects are background subtraction and optical flow method. The Optical flow method provides an apparent change of a moving object between the frames that determines the velocities and directions of each point of a video frame. Due to its higher detection accuracy, it is more suitable for non-rigid object analysis. Through optical flow estimation, motion information of moving objects can also be obtained for various video frames [1,19]. Gaussian Mixture Model (GMM) has been generally used for non-rigid object recognition due to its vast applicability. However, the GMM cannot appropriately model noisy or non-stationary background modes. The dense displacement fields, or optical flows, between consecutive video frames appear as the natural tool to build dense point trajectories: any optical flow technique, from the classic Horn and Schunk estimator and alternatives [10,11] to its most recent descendants [5], can be readily used to construct any number of point tracks over arbitrarily long video shots via numerical integration.

* Corresponding author.
E-mail addresses: malliga87@gmail.com (K. S.), sv@vcet.ac.in (V. S.).

https://doi.org/10.1016/j.ijleo.2017.11.181

Pattern classification methods have been presented to attain fruitful results in many areas of non-rigid object detection [6]. These methods can be decomposed into two significant components: feature extraction and classifier construction. In feature extraction, the dominant features are extracted from numerous training samples. These dominant features are used to train the classifier. During testing, the trained classifier cast an eye over the entire input to look for particular object patterns. The Support Vector Machine (SVM) classifier is broadly used for detection and recognition. The methods based on boosting [8] show impressive performance and attract much attention. Some existing approaches [12–14] have validated impressive recognition results. However, it is idealistic to expect flawless performance in non-rigid object recognition. Most non-rigid object recognition methods may miss a target or may also incorrectly classify a person in a stream of video scenes. Such inevitable errors would cause any object recognition method to misunderstand the object and background pixels. This method adds more difficulty to the non-rigid object recognition problem. As reported in a recent experimental study [8], the Adaboost-based approach has the fastest Detection Speed (DS) and comparable accuracy without the time constraint.

Tracking involves estimating the trajectory of an object moving around in a video scene [2,3]. Certain challenges in the literature survey have been made to use contour, segmenting method for dynamic target tracking [15,16]. Despite having the promising performance, the traditional trackers face a practical problem that they use the rectangular bounding box or oval to approximate the tracked target. However, non-rigid objects in practice may have complex shapes. Since the rectangle box used for presenting the tracked target directly determines the samples to be extracted in the subsequent target appearance modeling step, it is a critical factor in tracking performance. Inaccurate target presentation easily results in performance loss due to the pollution of non-object regions residing inside the rectangle box. Ideally, a better manner to describe the target is to use the accurate silhouette along the target's surface.

In this paper, the proposed optical flow is described for motion estimation, whose flow vectors are obtained by a combination of both Lucas-Kanade and Horn-Schunk method. The background subtraction method EM-EGMM is proposed to discard the noise and fill the holes for getting complete background region. Then, Adaboost classifier with the Gabor features is prescribed to guarantee both accuracy and speed necessities for real-world applications. Finally, Contour tracking has been chosen for tracking because silhouette based methods give a perfect shape sketch for the targets.

The remainder of this paper is structured as follows. In Section 2, the current state of the related work is reviewed. Then the proposed model is introduced in Section 3, and its generalized version is proposed in Section 4. Section 5 presents dense experiments conducted on a number of challenging video sequences. Section 6 concludes the paper.

2. Related work

2.1. Object detection

Frame differencing is a simple approach, which thresholds the difference between two image frames, and large changes are measured to be the foreground object. Another approach is to construct an illustration of the background that is used to evaluate against new images. Pixel wise median filter is a commonly used background modeling method, where the background is defined to be the median at each pixel location. Rather than using the median value of a group of pixels, a more reasonable way is that the pixel value follows a distribution of gaussian in the temporal route, and a model is used to calculate the likelihood of foreground and background for a particular pixel. When the single Gaussian is not able to satisfactorily account for the variance, a Mixture of Gaussians [9] is used to get better accuracy of the assessment. Another technique [18], estimates the probability of observing the pixel intensity values based on a sample of the intensity values of each pixel. The model is supposed to adapt quickly to changes in the video scene, thereby enabling responsive recognition of targets. An entirely dissimilar idea [17], investigates global information as an alternative of the local one. Alike to eigen faces, an eigen background is formed to confine the dominant variability of the background. In general, Markov Random Field (MRF) model representation takes a superior sense of balance between complexity and accuracy of the algorithm, in which the idea is to establish the mask from segmentation via a Maximum A-Posteriori (MAP) approximation.

2.2. Object classification

The performances of numerous object recognition approaches have been evaluated in [14]. With respect to the receiver operating characteristic performances and effectiveness manifold feature classifier combinations have been evaluated. Different features including PCA, local receptive fields (LRF) feature, and Haar wavelets are used to train neural networks, support vector machines (SVM) and Neural Networks (NN) classifiers. In addition, the Adaboost classifier based on features, has the maximum accuracy under a rigid constraint.

2.3. Moving object tracking

A real time tracking of non-rigid object shapes using mean-shift technique has been proposed [7]. The mean shift method is based on the iterations and the most probable target position in the current frame. The analysis of the method shows that it relates to the Bayesian structure at the same time as providing a resourceful solution. Thus, the tracking algorithm tries to identify the area of a video frame that is locally most related to a previously initialized representation and the object region to be tracked is carried out by a histogram. Here the target is represented by a rectangular or elliptical region. Some objects