Computers and Electrical Engineering 40 (2014) 375–389

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Fast pedestrian detection based on region of interest and multi-block local binary pattern descriptors ☆

Aminou Halidou [a,*], Xinge You [a], Mahamadou Hamidine [b], Roger Atsa Etoundi [c], Laye Hadji Diakite [a], Souleimanou [d]

[a] Huazhong University of Science and Technology, Hongshan, Wuhan, Hubei 430074, China
[b] University of Maradi, Faculty of Science and Technology, Department of Physics, Maradi 465, Niger
[c] University of Yaoundé I, Faculty of Science, Department of Computer Science, Yaoundé 4155, Cameroon
[d] University of Ngaoundéré, Faculty of Science, Department of Mathematics and Computer Science, Ngaoundéré 454, Cameroon

ARTICLE INFO

Article history:
Received 4 February 2014
Received in revised form 29 September 2014
Accepted 2 October 2014
Available online 30 October 2014

Keywords:
Pedestrian detection
Non-tensor product wavelet filter
ROI
Shapelet feature
MB-LBP
IFLDA

ABSTRACT

Nowadays pedestrian detection plays a crucial role in image or video retrieval, video monitoring systems and driving assistance systems. Detecting moving pedestrians is a challenging task, some of the detection methods are ineffective and slow. Occlusion, rotation, changes in object shapes, real time detection and illumination conditions are predominant obstacles. This paper is focus on the implementation of an efficient and speedy detector. A detection framework based on region of interest (ROI), full-body descriptor, body-part descriptors, and cascade classifier is proposed. ROI identifies, locates, and extracts candidate regions containing pedestrians, thus reducing the number of detection windows. In relation to human detection, independent information sources such as shapelet features and multi-block local binary pattern (MB-LBP) are used for features extraction. Experimental results showed that the proposed-model performs better than some state-of-the-art approaches, with suitable processing time for further operations such as tracking and imminent danger estimation.

1. Introduction

Pedestrian detection is a computer technology that provides the locations and sizes of persons in digital video images. It is the focus of recent research for its wide range of applications in our daily lives, including video surveillance, airport security, transportation system, image/video retrieval, and driving assistance.

Significant work has been devoted to detecting, locating, and tracking people in images and videos. Stereo-vision was the first approach based on hough transform algorithm, used for human detection in the late 1990s [1]. In 1998, an algorithm based on the movement of legs compared to the ground to detect and classify individual was developed by Heisele and Woehler [2]. Haar-like features and support vector machine (SVM) was proposed in the same year [3]. In 2000, researchers started taking advantage of the face detections algorithms. Following the face detection success, in 2001, the human detection work has witnessed rapid growth. Many works were developed based on face detection techniques such the Viola and Jones method which gave good result [4]. In 2005, histograms of oriented gradients (HOG) algorithm was introduced by the

☆ Reviews processed and approved for publication by the Editor-in-Chief.
* Corresponding author.
E-mail address: aminouhalidou@gmail.com (A. Halidou).

http://dx.doi.org/10.1016/j.compeleceng.2014.10.003
0045-7906/

researchers at the institut national de recherche en informatique et en automatique (INRIA), and proven to be efficient [5]. In 2006, credit was given to the covariance descriptor [6]. In 2007, the shape features are presented to describe local pieces (head, shoulder, and body) [7]. In 2008, researchers at Rutgers University improved the work by using the covariance matrix as descriptor [8]. In the same year Peking University employed the local binary patterns (LBP) used before in face detection area [9]. Significant performance was achieved by combining different pedestrian feature extraction approaches and classification algorithms [10].

Although the problem of pedestrian detection has been well studied in computer vision, detecting pedestrians remains a challenging task due to continuous changes in various factors like articulated pose, illumination, background, articulation, partial occlusions, style and colour of clothes. Some methods prefer to adopt existing methods for pedestrian detection, such as HOG detector [5]. Detections based on HOG are comparatively ineffective, due to their computational complexity. Few of these methods can work in realistic environment. On the other side, they used sliding-window, which scans the entire video frame without considering the potential pedestrian regions. The time taken by the sliding window to process video frames influence the detection speed. To address these problems, a novel pedestrian detection framework based on ROI, full-body detectors, body-part detectors, and improved fisher linear discriminant analysis (IFLDA) are proposed. The main contributions of this paper are summarized below:

1. New ROI method is introduced, decreasing the computational cost and improving the detection speed.
2. Image edges within the candidate region is improved using non-tensor product wavelet filter.
3. For features extraction, full-body descriptor based on shapelet and body-part descriptor based on MB-LBP are implemented.
4. Cascade ILFDA classifier is implemented to handle large data.

ROI extracts most informative regions containing pedestrians, it is used to avoid exhaustive search of the input image. In order to improve the efficiency and accuracy of pedestrian detection, we combine independent information sources, such as shapelet and MB-LBP for features extraction. Cascade IFLDA classifiers was used to improve the performance, and reliability of the detection system. The first stage uses full-body detector to generate pedestrians candidates, while the second stage uses body-part detector to confirm or reject the result of the first stage.

The rest of the paper is organized as follows: Section 2 reviews some of the prevalent available literature. Section 3 describes the methodology. Section 4 presents the experimental results. Section 5 summarizes the work and introduces the new directions.

2. Related work

Pedestrian detection has gained a vast amount of interest in the computer vision community over the past few years. The sliding window approaches are more promising when the image resolution is low. The first sliding window was proposed by Papa Georgiou, it was applied with the SVM to an over-complete dictionary of multi-scale Haar wavelets. Shape is represented using a pool of short line and curve segments (edgelet features) in [11]. The combination of many features increases the efficiency of the detector, LBP is a powerful discriminant feature, Wang combined it with HOG. We just mentioned few works, it is worthy to see [12] for more information. This work focuses on three basic problems: how to specify the ROI, how to combine efficient feature descriptor, and how to speed up the detection system.

For the full-body based pedestrian detection approach, Papageorgiou, Oren, and Poggio presented a pedestrian detector based on polynomial SVM using rectified Haar wavelets as input descriptors [3]. Gavrila and Philomin compared edge images to an example dataset using the chamfer distance [13]. Felzenszwalb, and Pedro presented the pedestrian as a shape model and used a generalization of the Hausdorff distance to compare edges with pedestrian templates [14]. Leibe, Seemann, and Schiele presented an algorithm that integrated evidence in multiple iterations and from different sources [15]. Zhu, Yeh, Cheng, Avidan, and Shai integrated the cascade rejectors approach with the HOG features to achieve a fast and accurate human detection system [16]. Munder and Gavrila studied the problem of pedestrian classification with different features and classifiers [17]. They found that local receptive fields do a better job of representing pedestrians and that both SVM and adaboost classifiers outperformed the other tested classifiers. Seemann, Fritz, Schiele presented a generative object model that could scale from a general object class model to a more specific object-instance model [18]. Maji, Berg, and Malik built histogram intersection kernel SVMs together with multi-level histograms of oriented edge energy features which performed much faster and better than existing systems [19]. Xu, Cao, and Qiao proposed an efficient tree classifier ensemble approach with a high detection and training speeds, which enables on-board detection in intelligent vehicles [20].

Most of pedestrian detection systems employ a fixed feature set. Gradient based features are among the most effective features for object detection. HOG has been widely used and generates many variants [5], some of which have achieved great success such as deformable part model [21]. But these approaches suffer from high computational expense, which makes it unsuitable for practical applications. The Haar-like wavelet features performed well for face detection but the performance was quite poor for pedestrian detection because of the large intra-class variability. Viola, Jones, and Snow improved their pedestrian detection system by integrating Haar-like wavelet features with motion features, but motion features are not available in a single image [22]. Wu and Nevatia [11] used a set of hand-coding mid-level features, called edgelet features