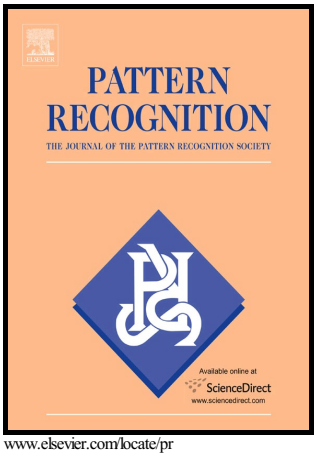
A Novel Pixel Neighborhood Differential Statistic Feature for Pedestrian and Face Detection

Jifeng Shen, Xin Zuo, Jun Li, Wankou Yang, Haibin Ling



www.elsevier.com/locate/pr

PII: S0031-3203(16)30271-0
DOI: http://dx.doi.org/10.1016/j.patcog.2016.09.010
Reference: PR5872

To appear in: *Pattern Recognition*

Received date: 14 January 2016
Revised date: 21 July 2016
Accepted date: 18 September 2016

Cite this article as: Jifeng Shen, Xin Zuo, Jun Li, Wankou Yang and Haibin Ling, A Novel Pixel Neighborhood Differential Statistic Feature for Pedestrian and Face Detection, *Pattern Recognition* http://dx.doi.org/10.1016/j.patcog.2016.09.010

A Novel Pixel Neighborhood Differential Statistic Feature for Pedestrian and Face Detection

Jifeng Shen[1], Xin Zuo[2], Jun Li[3], Wankou Yang[3], Haibin Ling[4]

[1]School of Electronic and Informatics Engineering, Jiangsu University, Zhenjiang, Jiangsu, 212013, China
[2]School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, 212003, China
[3]School of Automation, Southeast University, Nanjing, Jiangsu, 210096, China
[4]Dept. of Computer & Information Sciences, Temple University, Philadelphia, PA, 19122, USA

Abstract

Motivated by the successful application of Local Binary Pattern (LBP), in this paper we propose a novel pixel neighborhood differential statistic feature for pedestrian and face detection based on the multiple channel maps. The calculation of LBP comprises of two steps, Pixel Differential Feature (PDF) calculation and PDF sign encoding. The PDF distills discriminative information of local region that can improve the performance of the pedestrian detector, but the encoding step degrades the performance due to the quantization error. Although PDF is more discriminative than original channel maps, it has a much higher dimension than the original feature maps, and consequently requiring large computation cost. To address this issue, the pixel neighborhood differential pattern is learned with both supervised and unsupervised learning methods, which allow discovering discriminative pixel differential patterns in local area and achieving state-of-the-art results. Specifically, our method firstly aggregates the image channel maps into cell maps with max pooling. Then, pixel neighborhood differential feature based on each channel cell maps are calculated which contributes to encoding discriminative information in each local area and benefits the performance improvements. In addition, we attempt to learn discriminative differential statistic patterns by using linear discriminative analysis (LDA) and principle component analysis (PCA) for further performance improvement. Two sets of experiments are conducted on pedestrian detection and face detection respectively. The INRIA, Caltech, and ETH datasets are used for pedestrian detection, and the FDDB and AFW datasets for multi-view face detection. The experimental results show that our method achieves superior performance in comparison with the state-of-the-arts while running at 20 fps for 480×640 images.

Keywords

Pedestrian detection, Face detection, PDF, Neighborhood differential statistic patterns

1. Introduction

As a representative for the object detection problem, human detection is an active topic in computer vision. It poses great challenges due to large variability in deformation, illumination and occlusion present in images. In recent years, a great amount of attention has been paid to real-time human detection in applications such as vehicle autonomous driving, video surveillance and human activity understanding.

Despite a large body of work devoted to human detection in the last decades, it is still an open problem. Existing state-of-the-art methods [1-3] can be roughly categorized into two groups, those