Signal Processing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Signal Processing

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

Inferring occluded features for fast object detection

Shengye Yan*, Qingshan Liu

Jiangsu Key Laboratory of Big Data Analysis Technology, School of Information and Control, Nanjing University of Information Science and Technology, Nanjing 210044, China

ARTICLE INFO

Article history:
Received 31 July 2014
Received in revised form
23 October 2014
Accepted 26 October 2014

Keywords:
Object detection
Occlusion
Speed
Scanning window
Boosted cascade

ABSTRACT

In this paper, we study how to perform robust and fast detection of the partially occluded objects in natural scenes. Three techniques are proposed to achieve the objective. First, we propose an inference model based on linear reconstruction to reconstruct the unknown occluded features. The inference model is learned from the on-the-shelf training images without any occlusion. It takes advantage of the global structure constraint between the occluded part and the un-occluded part to recover the occluded features, which provides beneficial discriminative information during classification. Secondly, we design a new cascaded structure which is compatible with both high detection speed and high classification performance. In the proposed cascade structure, the earlier several stages are based on the more efficient un-occluded features to assure a high detection speed, while the later stages are based on both the un-occluded features and the reconstructed occluded features for better classification performance. Finally, we further propose an efficient occluded feature reconstruction method under the framework of Boosted cascade. In the reconstruction under the boosted cascade framework, only a few key features are selected by Boosting to conduct the reconstruction. And the reconstructed features along with the un-occluded features are used to learn the final classification model. Extensive experiments are conducted on car detection task in realistic environments and well recognized public pedestrian detection dataset (INRIA dataset). The experimental results demonstrate the effectiveness of the proposed method both in improving the classification performance and achieving a fairly high processing speed.

1. Introduction

Object detection is a fundamental task in computer vision and plays a very important role in a broad range of applications. Object detection itself can be used to indicate the appearance of an object. Based on this functionality, many applications can be developed such as the car\people counting system, the face-first autofocusing in camera industry, the automatic alarming system to avoid an unauthorized object class's intrusion. Moreover, object detection is the basic first step to conduct further object-related analysis. For example, face detection is the basic first step for face recognition, face expression recognition, and gender classification. Pedestrian detection is the basic first step for human pose estimation.

During the past fifteen years, object detection has attracted substantial attention and has made a big progress. Various features are designed or introduced for object detection such as Haar-like features [1,2], Histogram of Gradient Orientation [3,4,14], Local Binary Pattern [5], and Gabor [6].Various machine learning methods are investigated or proposed such as Neural Network [7,8], Bayesian classifier [9], SVM [10,11], and Boosting [1,12,13].

Though general object detection has been studied in many aspects, one issue is not solved well, i.e., the object detection with occlusions. In fact, occlusion not only exists in

* Corresponding author. Tel.: +86 15358126429.
E-mail address: shengye.yan@gmail.com (S. Yan).

http://dx.doi.org/10.1016/j.sigpro.2014.10.030
0165-1684/

Please cite this article as: S. Yan, Q. Liu, Inferring occluded features for fast object detection, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.10.030

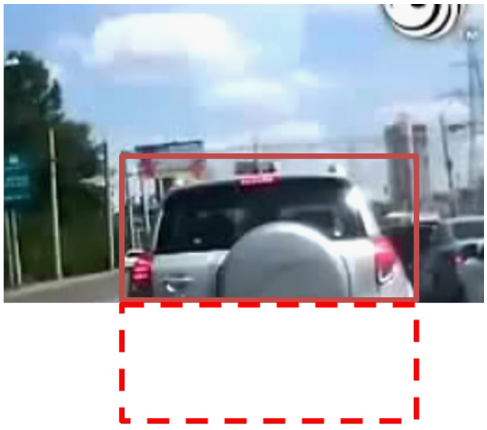

Fig. 1. A car image with only the upper part visible.

object detection, it is a common issue in computer vision tasks [22–27]. In this paper, we aim to solve a special kind of occlusion problem existing in object detection task. In this kind of problem, the occlusion mode is known. Here by mode we mean that the layout of the occluded area and the un-occluded part. As an example, under the most popular object detection framework of scanning-window, detecting the object falling in between the visible scope and the un-visible scope of an image is belonging to this case. In Fig. 1, an occluded car image with its scanning window is shown (the solid-line rectangle corresponds to the visible part. The dashed rectangle corresponds to the occluded part). Robust and fast detection of this kind of the partially visible object is very important in some real applications. As an instance, in a computer vision based rear-end collision avoiding system, a camera mounted in front of the car may miss in capturing the full car image in front of it as that in Fig. 1. But it is very important to detect the partial car robustly and fast, so that the car-safety system can provide an in-time reminder of the dangerous situation. In this kind of occluded object detection problem, traditional object detection methods have been observed to degrade in classification performance largely. In this paper, we preliminarily studied the problem. We want take advantage of the missing information to achieve a higher classification performance. The challenge of this problem comes from how to model the missing information in the occluded area which always largely deteriorates the object detection performance.

To deal with this kind of occlusion in object detection, the first optional solution coming into mind is to learn a classifier with only the un-occluded part of the object. Then in the testing phase, we directly apply the partial classifier to the un-occluded part. This may be the most popular manner in handling the kind of occluded object detection problem. For example, in pedestrian detection, instead of building a pedestrian classifier, the head-shoulder classifier is utilized to handle the occlusions [14–17]. Another stream of deformable component- or part- based model can also be considered as this kind of method to solve the partially occluded problem [18–22]. In this kind of method, each part has its own classification model. In the testing phase, each part or component is first justified to be occluded or not. And if the part is decided to be occluded, the part will not involve in further classification. This stream of solution is characterized with that it only takes advantage of the visible part of the object in performing the classification. And thus the specific global constraint between the occluded part and the un-occluded part in an object is neglected.

In contrast to the above-mentioned stream of solution, another solution is to recover the raw image of the occluded part with the aid of an on-the-shelf un-occluded object training set. In the object detection with this kind of solution, instead of learning a partial classifier, it learns a full classifier from the on-the-shelf un-occluded object training images. Then in the testing phase, it first recovers the raw image of the occluded part with the aid of the on-the-shelf un-occluded object images, and then applies the full classifier in the recovered testing image to perform a final classification. Although the reconstruction problem has been well-studied in computer vision [24–27] and the recovery effect is quite good, this kind of solution incurs computation issue when applying to the most popular fast and accurate object detection framework of scanning-window because there are such a great number of candidate windows needed to be recovered. If each of these occluded windows is applied with this kind of reconstruction, even with the simplest of reconstruction method, the computation cost increases largely. Besides, another issue in applying the reconstruction of raw occluded image is that the advanced technique of feature sharing [28–30] in object detection cannot be incorporated, which is always crucial to implement a fast object detection system.

In this paper, we propose a novel method for partially occluded object detection which can efficiently reconstruct the occluded part. Specifically, instead of recovering the raw image of the occluded input image, we propose to directly recover the occluded features. What is more, to simplify the reconstruction, we utilize a linear reconstruction formulation with a sparsity constraint. By enforcing the sparsity, only a small number of important un-occluded features are finally utilized to recover an occluded feature. And comparing to the first kind of solution for occlusion, the proposed method takes extra advantage of the global constraint in the object which we will demonstrate that can benefit the final classification.

Besides, we tailor a fast object detection framework with the proposed reconstruction method for real application. The traditional boosted cascade method is adopted in the classifier design. And considering the computation complexity of the reconstructed features, we have designed a new cascaded structure which is compatible with both the detection speed and the classification performance. In the proposed cascade structure, the earlier several stages are based on the more efficient un-occluded features to assure a high speed, while the later stages are based on both the un-occluded features and the reconstructed occluded features for better classification performance. Further, we propose an efficient occluded feature reconstruction method under the framework of boosted cascade. In the reconstruction under the boosted cascade framework, only a few key features selected by boosting are used to conduct the reconstruction. And the reconstructed features along with the un-occluded features are used to learn a better classification model.

The proposed method is evaluated on a real car detection dataset and a well recognized pedestrian benchmark

Please cite this article as: S. Yan, Q. Liu, Inferring occluded features for fast object detection, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.10.030