

摘要

物体检测近年来逐渐成为计算机视觉和模式识别领域中的一个研究热点。其之所以备受关注，主要出于以下两个原因：(1)物体检测领域技术的发展能够对其它相关领域问题提供借鉴，例如物体分类，图像标签等等；(2)应用上的需求进一步推动了物体检测方法的发展，例如在车载辅助驾驶系统、智能监控系统、人机交互系统以及基于内容的视频/图像管理系统中的应用。

物体检测中的难点问题可以归结为两个方面，类内散度大与信噪比低。类内散度大是指因物体表观形状上的差异而很难将物体的各个部分进行比较好的对齐。信噪比低是指物体数据中噪声所占比例较大而能够被用于对物体进行判别的信息相对较少。这两个难点问题综合作用的结果就是描述特定类物体异常困难。针对这两点，本文分别采用了提出更有判别力的局部特征和设计更学习能力更强的分类器的技术路线，并且在此技术路线的指导下提出了一系列物体检测特征提取与分类器学习方法。

论文的创新与主要贡献总结如下：

(1)从局部特征设计的角度，本文提出了一种边缘梯度方向直方图特征(Edged Histogram of Oriented Gradient, EGHOG)。EGHOG 使用区域的梯度方向直方图作为底层描述，并沿物体中常见的边缘将这些区域顺序串联起来。通过这种方式，局部边缘之间的关系可以被隐式的进行表达。在此基础上，我们提出一种基于局部距离关系的特征提取方法，以得到的距离向量反映区域的结构信息。实验结果表明，该特征相比于传统的轮廓特征与梯度方向直方图具有更好的判别力，同时可以很好的描述物体的结构特性。

(2)从局部特征设计的角度，本文提出了一种局部化的梯度方向共生直方图特征(Local Co-occurrence of Histogram of Oriented Gradient, LCoHOG)。传统的 CoHOG 特征提取整个图像内 31 种不同模式像素对的梯度方向共生直方图，然后使用线性支持向量机(Support Vector Machine, SVM)训练得到一个物体检测器。我们在此基础上将其局部化，采用 AdaBoost 方法选择最有判别力的区域与像素对模式构建层级物体检测器。实验结果表明，我们的方法与传统的 CoHOG 方法相比具有相同的检测率，在检测速度上有大约 30 倍的提升。

(3)从分类器设计的角度，本文提出了一种基于 AdaBoost(Adaptive Boosting)的样本空间预映射方法。在构建 AdaBoost 分类器的过程中，我们首先依据在原始空间中选择的“参考样本”(Reference Sample)对样本空间进行非线性映射，然后在映射之后的空间进行线性分类以学习弱分类器。“参考样本”对应核方法中的隐空间最优分类面，在映射之后的空间线性分类近似于在原始空间采用核方法的非线性分类，但计算量与线性方法相同。实验结果表明，我们的方法可以同时提升 AdaBoost 分类器的分类性能与计算效率。

(4)从分类器设计的角度，本文提出了一种全新的非线性弱分类器，分割向量弱分类

器(Partition Vector weak Classifier, PVC)。PVC 基于一个加性核函数与一系列预先定义的分割向量(Partition Vector, PV), 它的训练过程分为两步: (1)编码, 将样本空间显式映射到高维隐空间中训练线性分类面, 编码学习得到的线性分类面分类能力与原始空间核方法的非线性分类面完全相同。(2)解码, 将隐空间的线性分类转换为一系列原始空间基于 PV 的核函数运算。在检测过程中, 这些核函数运算可以进一步等价于分段线性函数, 计算的时间复杂度与线性方法完全相同。实验结果表明, 我们的方法相比于常用的线性弱分类器具有更优的检测性能与相同的检测速度, 相比于传统的非线性弱分类器具有相同的检测性能与更优的检测速度。同时, 该方法可以大大提升分类器的学习效率。

关键词: 物体检测, 局部特征, 分类器, 核函数, AdaBoost, 梯度方向直方图, 预映射, 弱分类器, 分割向量

The Research and Implementation of Object Detection Based on Local Features

Object detection is one of the most challenging and active research topics in computer vision and pattern recognition. This topic is attractive for the following two main reasons: (1) The development of the related technologies in object detection could provide valuable reference for other areas, such as image tagging and object classification; (2) The increasing demands in the practical applications, such as smart surveillance system, on-board driving assistance system and content based image/video management system.

The main challenges for object detection come from two aspects: large inner-class variation and low SNR. Large inner-class variation means that the shape of object can not be easily normalized due to different appearance and viewpoints. Low SNR means that the discriminative information is very limited in the object data comparing with the noise. From methodology point of view, we use utilizing more robust local features and designing more discriminative classifiers to deal with these two challenges. A series of local features and classifiers have been developed.

The main contributions of this paper can be summarized as follows:

(1) An Edged Histogram of Oriented Gradient feature (EGHOG) is proposed. EGHOG utilizes the Histogram of Oriented Gradient (HOG) of a series of regions as the base object description and concatenates these regions along an edge, so that the relationship of the local structure could be implicitly represented. We further propose a method for extracting the structure information of a local region, which is based on the distances between the gravity centers and the geometric centers. The experimental results show that the EGHOG is more discriminative than the Edgelet feature and traditional HOG feature. It could also reflect the structural information of local regions.

(2) A Local Co-occurrence of Histogram of Oriented Gradient (LCoHOG) feature is proposed. The tradition CoHOG feature extracts the co-occurrence histogram of the pixel pairs with 31 different patterns in a whole image. The linear Support Vector Machine (SVM) is further used to train a classifier. Our LCoHOG adopts the AdaBoost to select the most discriminative local regions and pixel pair patterns to construct a cascade classifier. The experimental results show that our method is at

least 30 times accelerated compared to the traditional CoHOG, while achieving the same accuracy.

(3) A sample pre-mapping method based on AdaBoost is proposed. In the training procedure, we apply a nonlinear mapping on original space by referring to the selected 'reference sample'. We train the weak classifier in the mapped space and further build the cascade classifier. The reference sample corresponds to an approximation of the optimal separating hyperplane in the implicit high dimensional space, so that the resulting classifier could achieve the performance similar to kernel method, while spending the computation cost as low as linear methods. Several experimental results on pedestrian and car show that our method could improve both the accuracy and the efficiency of the boosted classifiers.

(4) A novel nonlinear weak classifier named PVC (Partition Vector weak Classifier) is proposed. PVC is based on an additive kernel function and a series of pre-defined Partition Vectors (PV). The training of PVC consists of two steps: (1) an encoding step that uses a nonlinear function to map the feature vectors in the original space into a high dimensional implicit space, where a linear classification hyperplane is subsequently trained, so that the classification ability of the resulting hyperplane is the same as the kernel methods in the original space. (2) A decoding step which transforms the above classification function to a series of additive kernel functions referring to the PVs in the original space. The resulting classifier could be further accelerated as fast as the linear classifiers in detection, which is achieved by a linear piecewise function. Several experimental results show both the efficiency and effectiveness of the proposed method. Firstly, the boosted PVCs achieve significant improvement on the accuracy compared to linear weak classifiers. Secondly, it can obtain the same detection speed as linear weak classifiers, while the training is at least 4 times accelerated.

Keywords: Object Detection, Local Feature, Classifier, Kernel Function, AdaBoost, HOG, Sample Pre-mapping, Weak Classifier, Partition Vector Classifier