

基于 HOG 和 block 权重的快速人体检测方法

石志强, 赵向东, 李文军, 张挺

(航天科工集团第二研究院 207 所, 北京 100854)

摘要: 为解决 HOG 行人检测过程缓慢的问题, 在梯度向量直方图 HOG 中引入 block 权重的概念, 通过合理筛选 block, 组成行人的特征向量, 然后使用线性 SVM 作为分类器, 重新进行学习, 达到减少信息冗余、提高检测效率的目的。在 INRIA 库上进行实验, 结果表明, 通过合理选择 block, 能够在不影响检测效果的情况下, 显著减少 block 的数目, 达到提高检测速度的目的。

关键词: HOG; 行人检测; SVM 分类器; block 权重

中图分类号: TP391

文献标识码: A

文章编号: 1674-7720(2012)11-0044-03

Fast human detection method based on HOG and the weight of block

Shi Zhiqiang, Zhao Xiangdong, Li Wenjun, Zhang Ting

(Institute 207, the Second Academy of China Aerospace Science Corp, Beijing 100854, China)

Abstract: Based on the fact that HOG pedestrian detection process is slow, the concept of block weight is introduced into gradient vector histogram to reduce information redundancy and improve the detection efficiency. The theory lies in the base that the block weight can reduce information redundancy and improve the detection efficiency through the reasonable screening of the block to form pedestrian feature vector, and re-learning using the linear SVM classifier. The results of the experiments using INRIA library show that a reasonable choice of the block can reduce the number of the blocks and improve detection speed significantly while have little effect on the detection results.

Key words: HOG; pedestrian detection; SVM classifier; block weight

计算机视觉处理是计算机人工智能的一个重要研究方向, 它的最终目标是通过计算机模拟人类视觉, 实现对现实场景的理解和推理。其中, 对于输入场景通过计算机实现目标的自动识别分类, 即目标识别, 是实现计算机视觉的一个基础部分, 也是目前的一个研究热点。比较成功的商业化应用实例是人脸识别系统, 它是通过计算机对人脸的智能识别, 实现对不同人身份的确认。在人体识别方面, 由于人体自身姿态、衣着的多样性以及人体外观变化大, 目前尚没有较为成熟的算法。

目前, 人体识别主要有基于多模板匹配^[1]、基于边缘轮廓特征^[2]和基于运动特征^[3]等各种方法。然而, 从实际效果来看, 目前最好的是 DALAL N 和 TRIGGS B 提出的基于 HOG 特征的人体检测方法^[4]。DALAL N 等人通过使用 16×16 大小的块在检测窗口中滑动, 然后按一定方式统计块中的梯度向量直方图, 将各个块的梯度向量直方图串联, 组成特征向量, 最终利用 SVM 对获取的特

征向量进行训练, 获取分类器。按照该方法, 如果选择检测窗口大小是 128×64, block 大小为 16×16, cell 大小为 4×4, bin 分为 9 个方向, 每次 block 移动 8, 那么获得的向量大小为 3 780 维, 检测效果虽然理想, 但是耗时, 难以达到实时性的要求。

本文通过对基于 HOG 特征的人体检测算法进行分析, 提出了一种基于感兴趣区域进行 HOG 提取的算法。该算法在适当影响检测效果的条件下, 能够较好地减少计算量, 达到提高计算速度的目的。

1 行人检测算法

1.1 特征选取

HOG 是 DALAL N 等人在 2005 年提出的一种基于梯度的特征提取过程, 它的内容主要有 4 点: (1) 使用梯度作为特征提取对象, 将梯度向量划分为统计 bin; (2) 将梯度向量图划分为网格状的 cell, 以 cell 为单位统计梯度向量直方图; (3) 以 block 为单位, 对 cell 进行统

计,得到三维梯度向量直方图,并进行归一化,减少局部光照的影响;(4)收集所有的 block,合并成最终图像的 HOG 特征向量。

输入图像像素点 (x, y) 的梯度按如下方式来计算:

$$G_x(x, y) = H(x+1, y) - H(x-1, y) \quad (1)$$

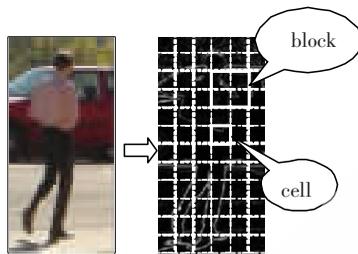
$$G_y(x, y) = H(x, y+1) - H(x, y-1) \quad (2)$$

其中, $G_x(x, y)$ 、 $G_y(x, y)$ 、 $H(x, y)$ 分别表示图像像素点 (x, y) 处 x 方向、 y 方向的梯度以及原始点像素值。像素点 (x, y) 的梯度和梯度方向定义为:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (3)$$

$$\alpha(x, y) = \tan^{-1} \left(\frac{G_y(x, y)}{G_x(x, y)} \right) \quad (4)$$

整个提取过程各概念之间的关系如图 1 所示。



(a)原图 (b)梯度图
图 1 block 和 cell 示意图

在 HOG 特征提取过程中, block 窗口按固定方式在检测窗口的梯度向量图上进行扫描获取 block 块内的局部特征,特征向量是通过每个 block 块的特征进行统计获得的,整个 HOG 计算过程以 block 为单位。由图 1 (b) 可以看到,检测窗口的大量 block 块并不包含目标信息,真正包含信息的 block 是包含行人边缘轮廓的 block。为了提高检测效率,可以采取降维的方式,通过选取包含人体信息、对分类有显著影响的 block 作为最终的特征,剔除掉不包含信息的 block,可以显著地减少行人的特征维数,达到快速检测的目的。

1.2 线性 SVM

SVM^[5] 是基于 Vapnik 等发展的统计学习理论的支撑向量机,具有相对优良的性能指标。通过算法的学习过程,SVM 可以自动寻找出那些对分类有较好区分能力的支持向量,构造出性能良好的分类器。

线性 SVM 假设样本是线性可分的情况下,寻找到一个可以使训练样本完全可分的超平面将样本分开,其本质上是一个凸二次规划问题。该超平面描述为:

$$w \cdot x + b = 0 \quad (5)$$

其中,“ \cdot ”是点积, w 是 n 维向量, b 为偏移量。

已知一组独立同分布的训练样本:

$$X = \{(x_i, y_i) | x_i \in R^d, y_i = \pm 1, i = 1, \dots, l\} \quad (6)$$

和假设函数集:

$$\{f(x, w, b) = \text{sgn}(w \cdot x + b)\} \quad (7)$$

其中, R^d 表示输入空间。

线性支撑向量机可以归纳为如下二次规划:

最小化:

$$\frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^l \xi_i \right)^p \quad (8)$$

约束条件:

$$y_i (w \cdot x_i + b) \geq 1 - \xi_i \quad (9)$$

$$\xi_i \geq 0 (i = 1, \dots, l) \quad (10)$$

其中, $C > 0$ 为惩罚因子,其取值越大,对经验误差的惩罚也越大。

取 $p = 1$, 通过 Lagrange 乘子方法,求原规划式的 Wolfe 对偶规划。

最大化:

$$L_D = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=1}^n a_i a_j y_i y_j (x_i \cdot x_j) \quad (11)$$

约束条件:

$$\sum_{i=1}^l a_i y_i = 0 \quad (12)$$

$$0 \leq a_i \leq C (i = 1, \dots, l) \quad (13)$$

求解上述二次规划问题,得到最优的 Lagrange 乘子 a_i ,那些对应于 Lagrange 乘子大于零的训练样本被称为支持向量。最终线性支持向量机的判决函数具有如下形式:

$$f(x) = \text{sgn}(w \cdot x + b) \\ = \text{sgn} \left(\sum_{i=1}^l a_i y_i x_i \cdot x + b \right) \quad (14)$$

将式(14)继续整理,得:

$$f(x) = \text{sgn} \left(\sum_{i=1}^l a_i y_i x_i \cdot x + b \right) \\ = \text{sgn} \left(\sum_{j=1}^d \sum_{i=1}^l a_i y_i x_{ij} \cdot x_j + b \right) \quad (15)$$

令

$$w_j = \sum_{i=1}^l a_i y_i x_{ij} \quad (16)$$

得到:

$$f(x) = \text{sgn} \left(\sum_{j=1}^d w_j x_j + b \right) \quad (17)$$

由式(17)可知, $|w_j|$ 越大,表示 x_j 是在行人识别中起较大作用的分量。因此,可以通过给出一个衡量 block 对分类影响大小的量 block 权重因子,通过对每个 block 权值的统计,寻找出在人体检测中有较大区分能力的 block,重新组成表征行人的特征向量。

首先对判决函数按 block 重新整理:

$$f(x) = \text{sgn} \left(\sum_{j=1}^d w_j x_j + b \right) \\ = \text{sgn} \left(\sum_k \sum_j w_{kj} x_{kj} + b \right) \quad (18)$$

定义第 k 个 block 权重因子:

$$w_k = \sum_j w_{kj}^2 \quad (19)$$

图 2 为改进的二次训练对 block 重新选择的过程。训练过程分为以下 4 步：

(1) 利用线性 SVM 分类器对依据传统 HOG 方法提取出来的特征进行训练, 得到支持向量的决策函数;

(2) 根据式 (19) 对步骤 (1) 得到的决策函数中的 w_{kj} 进行计算, 得到各个 block 的权重因子;

(3) 设定合适的阈值, 选取权重因子高于阈值的 block, 即对分类影响较大的 block 作为二次训练的 block;

(4) 以选择的 block 组成特征向量重新训练, 得到最终支持向量的决策函数。

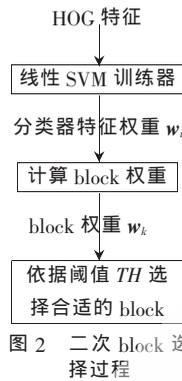


图 2 二次 block 选择过程

2 实验结果

为了证明该方法的有效性, 本文在 INRIA 样本库上进行验证。使用 LIBSVM 2.88 作为 SVM 工具, 训练样本选取正样本 3 542 个, 负样本 4 542 个, 测试样本选取正样本 1 000 个, 负样本 1 000 个。

原始 HOG 算法每个 block 由 36 维特征构成, 总共由 105 个 block 构成, 描述行人的特征向量长度为 3 780 维。

其中感兴趣 block 通过手动标定, 选取包含轮廓特征的 block 组成特征向量。实际操作过程中, 选取了 41 个包含轮廓特征的 block 形成 1 476 维特征向量, 经过训练获取 SVM 决策函数。

二次 block 选取过程基于 block 权重因子对 block 进行选择, 选取符合条件的 block 重新提取特征向量进行训练, 获取 SVM 决策函数。在实际实验中, 首先通过原始 HOG 学习过程获取 SVM 决策函数, 经过计算, 获得 block 权重因子, 选取阈值为 0.5, 对 block 权重因子进行筛选, 得到实际符合条件的 block 为 20 个, 如图 3 所示, 利用这 20 个 block 重新生成 720 维的特征向量, 经过训练获取 SVM 决策函数。

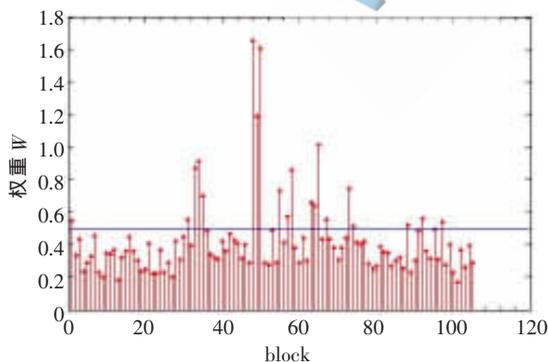


图 3 block 权重因子以及检测阈值

分别利用 3 种不同的方法通过改变决策函数中的阈值 b 对测试样本进行检测, 得到漏警率和虚警率的关系如图 4 所示。

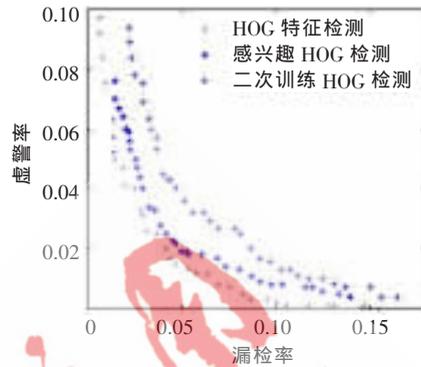


图 4 三种不同方法效果比较

由图 4 可以看出, 在 3 种检测方式中, 原始 HOG 检测方式检测效果最好, 通过人工选择感兴趣区域, 保留 41 个 block 后提取特征向量进行分类, 检测效果有一定的下降, 这主要是由于人类主观经验的不足, 不能完全合理确定起主要作用的特征, 导致特征选择不充分的原因。

二次 block 训练的方法通过引入 block 权重的概念, 利用原始训练得到的决策函数中的信息来确定感兴趣 block, 在选取的 block 数目减少到原始的 1/5 即 20 个时, 在显著提高计算速度的情况下仍然能够得到较精确的分类器。

实验结果证明, 本文提出的两种减少特征维数、提高分类效率的算法在实际的应用中较为有效。

HOG 特征是目前在行人检测过程中较为有效的一种方法, 在人体检测过程中有着很高的识别率, 然而 HOG 由于计算量较大、检测速度慢, 因此限制了 HOG 的应用。本文通过对 HOG 的分析, 发现可以通过减少 HOG 中冗余 block, 从而减少计算量, 提高检测速度。

参考文献

- [1] 吕治国, 徐昕, 贺汉根. 基于可变模板和支持向量机的人体检测[J]. 计算机应用, 2007(9): 2258-2261.
- [2] 陈实, 马天骏, 黄万红, 等. 基于形状上下文描述子的步态识别[J]. 模式识别与人工智能, 2007(6): 794-799.
- [3] 韩鸿哲, 王志良, 刘冀伟, 等. 基于线性判别分析和支持向量机的步态识别[J]. 模式识别与人工智能, 2005(2): 160-164.
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]. CVPR'05, 2005.
- [5] 焦李成, 周伟达, 张莉, 等. 智能目标识别与分类[M]. 北京: 科学出版社, 2010.

(收稿日期: 2012-03-13)

作者简介:

石志强, 男, 1986 年生, 硕士, 主要研究方向: 图像处理。

《微型机与应用》2012 年第 31 卷第 11 期