

采用在线高斯模型的行人检测候选框快速生成方法

覃 剑 王美华

重庆大学通信工程学院, 重庆 400030

摘要 行人检测是模式识别及机器学习领域的研究热点之一,广泛应用于智能监控、辅助驾驶等领域,而行人候选框的生成是识别及跟踪行人目标的一项重要的前期工作。针对静态监控场景以及特定情况下的车载监控场景,提出了一种基于在线高斯模型的行人检测候选框的快速生成方法(OL_GMPG)。该方法采用高斯模型拟合行人尺寸分布,可以通过生成较少数目的行人候选框达到较高的检测率;并可通过高斯模型的学习与更新过程,获取场景中行人频繁出现的位置以及对应的目标尺度信息,为后续的行人识别及跟踪过程提供辅助。

关键词 机器视觉; 行人检测; 高斯模型; 检测率; 尺度信息

中图分类号 TP391.41 **文献标识码** A

doi: 10.3788/AOS201636.1115001

Fast Pedestrian Proposal Generation Algorithm Using Online Gaussian Model

Qin Jian Wang Meihua

College of Telecommunication Engineering, Chongqing University, Chongqing 400030, China

Abstract Pedestrian detection is one of the most active research topics in the fields of pattern recognition and machine learning. It has been widely used in intelligent monitoring, auxiliary driving and so on. Generating pedestrian detection proposals is an important work in the early period of pedestrian recognition and pedestrian tracking. Based on the static monitoring scene as well as the on-board monitoring scene under specific circumstances, a novel method to generate pedestrian detection proposals quickly (OL_GMPG) is proposed by using online Gaussian model. High detection rate can be achieved by generating fewer pedestrian detection proposals through the Gaussian model fitting. Both the positions where people appear most frequently and the scale information of corresponding targets can be obtained through the learning and updating processes of the Gaussian model. The information is beneficial to subsequent pedestrian recognition or pedestrian tracking.

Key words machine vision; pedestrian detection; Gaussian model; detection rate; scale information

OCIS codes 150.1135; 100.3008; 100.5010; 110.2960

1 引 言

行人检测是模式识别及机器学习领域的研究热点之一,被广泛应用于智能监控、辅助驾驶等领域。现有方法一般采用生成检测候选区域和行人识别两个步骤来进行,即先在图像中选择可能为行人目标的区域,然后采用各种识别算法对该区域进行检测,以确定是否为行人目标。行人识别主要采用各种分类器实现,如基于特征的方法(方向梯度直方图^[1]等)和基于模型的方法(形变部件模型^[2]等)。生成检测候选区域则有基于运动检测、基于图像分布的方法等。基于运动检测的方法一般先假定行人目标是运动的,然后提取图像序列中的运动信息,得到运动目标区域,以此作为行人检测的候选区域。文献[3-5]采用了光流法,利用运动目标的光流随时间变化的特性,将光流计算得到的运动场作为一个重要的识别特性来生成运动目标的候选区域^[6]。文献[7-9]采用了背景差分法,先构建背景模型,然后将图像序列中的当前帧和背景模型进行比较来进行运动目标的检测。文献[10-12]采用帧间差分法,根据视频图像序列中连续两帧或者几帧图像的差异来

收稿日期: 2016-03-24; **收到修改稿日期:** 2016-07-01

基金项目: 国家自然科学基金(61108086)、重庆市自然科学基金(cstc2016shmszx0111)、中央高校基金(106112014CDJZR165503)

作者简介: 覃 剑(1977—),男,博士,副教授,主要从事图像处理、机器视觉方面的研究。E-mail: qinjian@cqu.edu.cn

进行运动目标检测,又称为时间差分法。基于运动检测的方法对持续明显运动的目标检测效果较好,但难以检测到慢速运动或静止的行人目标。基于图像分布的候选区域选择方法一般先根据某种规则对图像进行区域分割,从而得出有效的前景部分,文献[13-15]均属于此类方法。文献[16-17]利用深度信息进行目标检测,但前提是获得双目视频序列,难以应用到如今大量存在的单目监控或车载视频处理中。行人检测候选框的生成是识别行人的一项重要的前期工作,目前已有不少生成目标候选框的方法,比如早期的滑动窗搜索方式^[18]以及后来提出的选择性搜索^[13]与 Bing^[19]等方法。候选框的生成对行人检测的准确度和效率均有很大的影响。如果前期生成的候选框与真实目标的覆盖度较大,且候选框的数目不足以对整体检索效率造成明显的影响,就会对整个行人检测系统产生正面的作用,即提升行人检测的准确度和效率,并缩短搜索的时间。但是在很多情况下,这两个指标并不能同时满足。应用多尺度搜索时,滑动窗搜索方式往往会产生数以万计的搜索框,使目标检测的效率低下。因此学者们针对滑动窗搜索方式的弊端,提出了很多改进方法来增强检测的实时性,为使用更加复杂的识别算法留足空间。有一种简单的减少搜索框数目的方法,即先去掉图像上方1/3的区域,再进行滑动窗搜索^[18]。也有一些学者从其他角度生成候选框,如选择性搜索,综合了蛮力搜索和图像分割来获取可能出现目标的位置。

本文所述在线高斯模型的行人检测候选框快速生成(OL_GMPG)算法,针对静态监控视频以及车载监控视频等特殊场景,充分利用该类场景的固有特征,训练高斯模型,得出行人频繁出现的区域与检测点处的尺度信息。在之后的候选框生成过程中,在此区域进行滑动窗搜索,多尺度搜索则依据高斯模型提供的尺度信息实现。OL_GMPG算法采用在线学习机制,模型更新与检测同步进行。在加州理工学院行人检测数据库上应用OL_GMPG算法进行处理时,若在每幅图片上平均设置5387个候选框,可获得81%以上的平均覆盖率。

2 OL_GMPG 算法原理

静态监控场景的背景分布一般比较稳定。如位于小区或者路口的静态摄像头,主要拍摄对象是道路上的行人与车辆,监控图像的背景由若干固定方位的道路以及周围的绿化带等组成。根据日常生活经验,行人在道路上出现的概率比在周边区域(绿化带等)大得多。如果经过足够长时间的频次统计,该类场景下行人频繁出现的区域大致和道路区域相吻合。行驶中的车载摄像头所拍摄的图像大多是车前的路面以及远处的场景。车辆正常的行驶方向沿着道路方向,在路况没有发生明显改变时,车载监控图像的背景分布可以近似看作静态背景分布,从而可以与上述静态监控场景使用同样的方法来得出行人频繁出现的区域。但是,单纯使用长时间人数统计的方式虽然可以利用时间的累积效应得出人口密集区域,却无法得知其他信息。OL_GMPG算法由于采用了模型拟合的方式,可以巧妙地解决上述问题。

2.1 获取检测点

众所周知,高斯分布(正态分布)是一个在很多领域得到广泛应用的重要分布。现实中许多随机变量都服从或者近似服从高斯分布^[20]。

如果随机变量 x 的概率密度函数为

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], x \in (-\infty, +\infty), \quad (1)$$

式中 μ 和 σ 是常数, $\sigma > 0$, 则 x 服从高斯分布, 记为 $x \sim N(\mu, \sigma^2)$, 称 x 为正态随机变量。当 $\mu = 0, \sigma = 1$ 时, 该分布称为标准正态分布。

针对静态监控场景或特定情况下的车载监控场景,认为监控图像的某一确定位置处出现的行人目标的尺度大小近似服从正态分布,因此可以用数学期望 μ 和标准差 σ 两个参量来进行描述。在该假设下,某个像素点处高斯分布的期望即为此处出现概率最大的行人尺度大小,而标准差决定了高斯分布的幅度大小。设置各像素点处高斯模型的均值初始值为 900, 标准差初始值为 20。针对某一个位置,随着新尺度的引入,此处高斯模型的均值与方差被不断更新。模型均值、标准差和检测点坐标的单位均为像素。具体更新过程为^[21]

$$\mu_{T+1} = (1 - \vartheta) \times \mu_T + \vartheta \times S_T, \quad (2)$$

$$\sigma_{T+1} = \sqrt{(1 - \vartheta) \times \sigma_T^2 + \vartheta \times (S_T - \mu_T)^2}, \quad (3)$$

式中, μ_T 与 μ_{T+1} 分别为 T 时刻与 $T+1$ 时刻高斯模型的均值, σ_T 与 σ_{T+1} 分别为 T 时刻与 $T+1$ 时刻高斯模型的标准差。 ∂ 为更新率, 其值设置为 0.3, S_T 指当前输入的行人尺度大小, 即行人目标的外接矩形框的面积。

根据高斯模型的固有特性, 即模型的标准差大小反映高斯曲线概率峰值的大小, 设置一定的标准差阈值, 在学习过程中动态筛选出图像中模型标准差小于该阈值的若干检测点。标准差阈值设置为 18。根据高斯模型的物理意义, 上述符合要求的检测点即为行人频繁出现的位置。

衡量行人检测候选框生成算法的准确性与有效性的指标主要是针对每一帧图片的检测率以及所有待检测图像帧的平均检测率。当候选框与目标的重合度 (IOU) 大于设定的阈值时, 检测到判定目标。具体地, 设 D 表示该帧图片中检测出的标注行人目标数量, A 表示该帧图片中的所有标注行人目标数量, 则该帧图片对应的检测率为

$$r = D/A. \quad (4)$$

2.2 设置检测框

OL_GMPG 算法通过高斯模型的学习与更新过程, 得出场景中行人目标频繁出现的区域后, 再应用多尺度滑动窗搜索方式在检测区域均匀设置检测框。在检测点出现频率最高的行人尺寸由高斯模型的均值给出, 为此处多尺度检测框的设置提供基准尺寸。具体的实现细节将在第 3 节进行阐述。

3 OL_GMPG 算法的实现

3.1 初始化

开始进行模型的学习时, 由于输入的样本数目不足, 高斯模型暂未收敛。得出的行人频繁出现的位置较少, 形成的区域面积较小, 与实际情况不吻合。因此在开始的一段时间以内, 采用传统滑动窗搜索方式设置候选框, 防止出现检测率过低的情况。假设接下来的行人识别过程是无误差的, 通过传统滑动窗搜索方式检测出的标注目标将作为高斯模型更新的输入。在该阶段, 滑动窗搜索与高斯模型更新同步进行。在一段时间的学习过程后, 高斯模型趋于收敛状态, 得出的行人频繁出现的位置增多, 检测区域面积增大。当高斯方法在某一帧及其前 599 帧的平均检测率达到 0.8 时 (对应 20 s 视频), 由传统滑动窗搜索方式切换为 OL_GMPG 算法, 为后续的识别等模块提供候选框。实验结果显示, 在高斯模型学习过程进行到 2000 帧左右时满足上述条件。在进行传统滑动窗搜索时, 由于是全图密集搜索, 速度较慢。随着高斯模型的收敛, 会及时切换搜索方式, 不会对系统的实时性造成明显影响。使用 OL_GMPG 算法进行搜索时, 某一帧所检测出的标注目标会用于高斯模型的更新, 调整后的检测区域会用于下一帧的检测过程。

3.2 候选框的生成

OL_GMPG 算法以随时间变化的目标尺度为自变量, 建立图像各个位置处的高斯模型。根据 2.1 节所述的方法得出行人频繁出现的区域后, 利用滑动窗搜索方式在此区域均匀设置候选框。对于通过高斯模型的学习过程得出的检测点, 以均值尺寸为基准设置不同尺度的检测框。对于检测区域内部的其他位置, 要分几种情况进行讨论。首先计算距离这些位置最近的已知检测点 (由高斯模型直接得出的检测点) 以及最短距离。如果最短距离小于 15 pixel, 则用最近的已知检测点的基准尺寸作为该位置的基准尺寸; 如果最短距离为 15~37 pixel, 则使用所有已知检测点的中值尺寸作为该位置的基准尺寸; 如果最短距离大于 37 pixel, 则使用所有已知检测点的最大尺寸作为该位置的基准尺寸。各检测点设置的候选框尺度与基准尺度的比例范围为 0.6~2.6, 模板的宽高比锁定为 0.5。通过调整尺度变化的步长或者相邻模板质心之间的距离, 可以控制整幅图像设置的候选框数目。

3.3 算法步骤

OL_GMPG 算法具体步骤如图 1 所示。分为初始学习和在线处理两阶段, 由参数 flag 控制切换。记第 t 帧时的算法模型为 M_t , 根据 M_t 生成的搜索框集合为 φ_t^o , 对应检测率为 r_t^o ; 根据传统滑动窗方式生成的搜索框集合为 φ_t^s , 对应检测率为 r_t^s 。起始时令 flag 为 0, 算法处于初始学习阶段, 此时采用传统滑动窗方式生成候选框集合 φ_t^s 并检测, 再根据 (2)~(3) 式用得到的检测参数训练模型 M_t , 然后根据 (4) 式计算当前模型 M_t 对应的检测率 r_t^o 。若小于阈值, 则读入下一帧继续训练; 若达到阈值, 则将 flag 置为 1, 算法切换到在线

处理阶段。在在线处理阶段,对第 t 帧,先采用已有模型 M_{t-1} 生成候选框集合 φ_t^o 并检测,再根据(2)~(3)式用得到的检测参数更新参数 μ_t 及 σ_t ,得到新模型 M_t ,再进入下一帧循环往复。

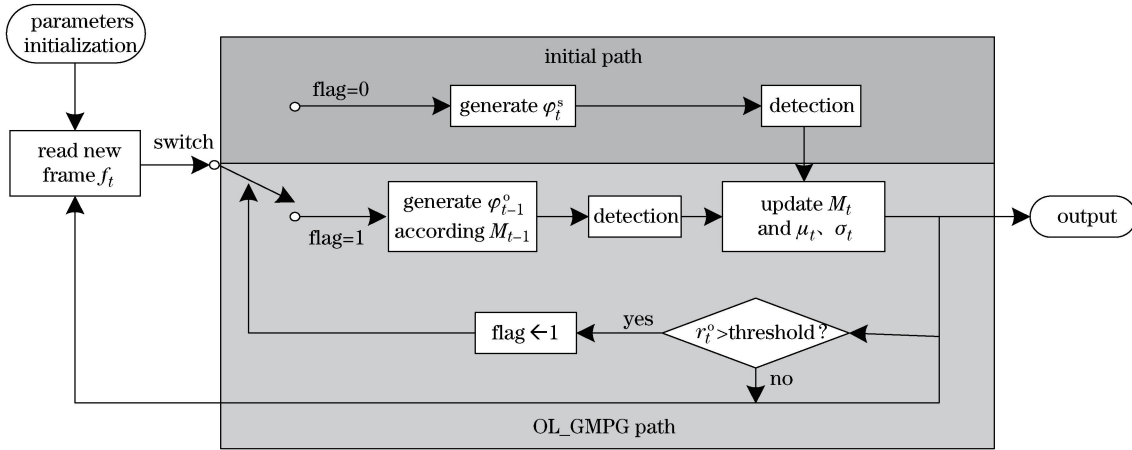


图 1 OL_GMPG 算法流程图
Fig. 1 OL_GMPG process flow

4 测试结果及分析

4.1 运用高斯模型生成检测点

所有实验均基于 Windows 7 系统下的软件平台 Matlab R2014a。测试数据集是加州理工学院行人检测数据库的一个分支, set00 数据集, 该数据集由一个行驶中的车载摄像头拍摄。在每帧图像都进行高斯模型的检测与更新, 并在此过程中动态生成模型标准差小于阈值 18 的位置。在第 3000 帧和第 8000 帧得出的检测点区域的热力图如图 1 所示。图中颜色不同于背景的点表示模型学习得到的行人频繁出现的位置, 由该图可以看出检测点区域随着高斯模型学习过程的变化情况。当搜索框间隔为 10 pixel 且每个检测点处设置 3 个搜索框时, 传统滑动窗搜索方式将在第 2207 帧时切换为 OL_GMPG 算法, 检测点在图像中的坐标以及相应的尺寸均值、标准差如表 1 所示。

表 1 2207 帧处检测区域各位置的坐标以及尺寸均值和标准差

Table 1 Coordinates, size averages and standard deviations of detection points in the 2207th frame

Coordinate		Size average	Standard deviation	Coordinates		Size averages	Standard deviations
X	Y			X	Y		
183	513	894.1	15.15638	188	597	899.1	16.77269
183	563	903.6	17.35419	188	599	899.1	16.77269
183	565	903.6	17.35419	183	503	897.9	16.94706
184	569	900.0	16.73320	177	552	900.0	16.73320
173	578	893.9	14.87246	177	554	900.0	16.73320
176	504	900.9	16.77269	190	527	900.9	16.77269
176	506	900.9	16.77269	47	317	896.4	17.35419
176	508	900.9	16.77269	53	80	896.4	17.35419
174	528	900.6	16.75076	46	316	903.6	17.35419
174	530	906.4	15.86621	44	315	896.4	17.35419
179	565	900.9	16.77269	44	55	899.1	16.77269
179	494	903.6	17.35419	176	544	900.9	16.77269
184	550	896.4	17.35419	179	102	897.9	16.94706
183	551	903.6	17.35419	179	109	897.9	16.94706
183	554	903.6	17.35419	179	115	903.6	17.35419
185	568	899.1	16.77269	178	121	903.6	17.35419
186	570	899.1	16.77269	185	510	836.9	13.91991
188	596	899.1	16.77269				

根据表 1 中的数据,不同位置处的平均尺度大小不同,模型标准差也不同。模型标准差越小,对应的高斯概率密度函数曲线越陡峭,说明此处行人目标尺度值大多集中在均值附近。一般情况下,均值尺寸越小的位置距离摄像头越远,反之则距离摄像头越近。这样的远近信息并不是通过人为观测获得,而是通过高斯模型的主动学习获得的。

4.2 OL_GMPG 算法测试结果

这些模型标准差小于阈值的位置共同组成了行人频繁出现的区域。图 2 中该区域位于图像中部偏上的位置,与实际情况比较吻合(行驶中的车辆近前方出现行人的概率较小,道路远处出现行人的概率较大)。为了增强算法的稳定性,将图 1 所示检测点集合的外接矩形框在图像宽度方向扩展 80 pixel 的长度(左右各扩展 40 pixel),在高度方向扩展 90 pixel 的长度(上下各 45 pixel)。在扩展后的矩形区域内进行多尺度滑动窗搜索,在数据集 set00 上的测试结果如图 3 所示。其中横坐标表示 IOU 阈值,纵坐标表示 IOU 超过某一阈值的比率。

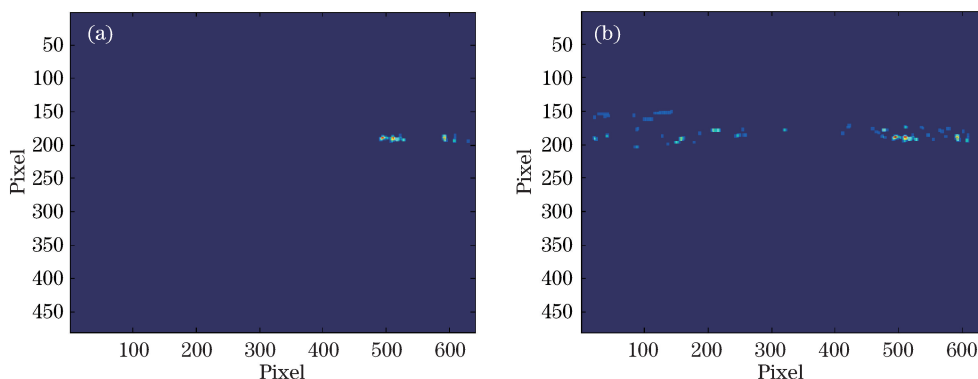


图 2 高斯模型生成的行人频繁出现的位置集合。(a) 第 3000 帧; (b) 第 8000 帧

Fig. 2 Pedestrians frequently appearing locations extracted by the Gaussian model. (a) The 3000th frame; (b) the 8000th frame

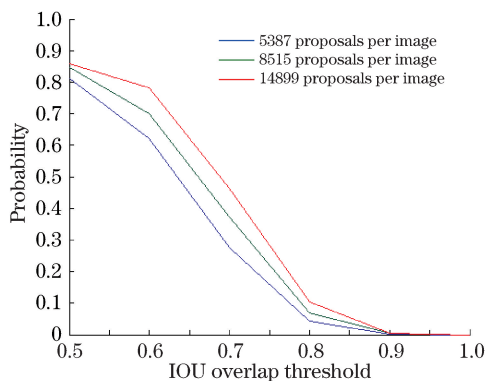


图 3 OL_GMPG 算法在不同候选框数量下的测试结果

Fig. 3 Results of the OL_GMPG algorithm for different numbers of detection proposals

通过控制相邻搜索窗质心的间隔,在测试图像上设置不同数目的候选框。如图 3 所示,当相邻搜索窗质心间隔为 10 pixel 时,整幅图像需要设置 5387 个候选框,此时 IOU 值大于 50% 的比率超过 81%。当在整幅图像设置 8515 个候选框时,IOU 值大于 50% 的比率接近 85%。进一步增加候选框的数量,整体检测率会继续提升。但是过多的候选框会导致系统检测效率低下,实时性变差。因此选择 5387 个检测框可以在不影响实时性的情况下获得较高的整体覆盖度。

图 2 显示了不同 IOU 阈值下的平均检测率。为了更加清楚地表示结果,画出了 IOU 阈值取 0.5 时每帧的检测率情况。如图 4 所示。图 4(a)、4(b)、4(c) 分别表示候选框数目为 5387、8515、14899 时每帧图片的检测率情况。如 3.3 节所述,在适当的帧数处会由传统滑动窗搜索方法切换为 OL_GMPG 算法,图 4(a)、4(b)、4(c) 对应的切换帧数分别为 2207、1927、2027。可以看出,在切换帧后算法的整体检测效果良好,且检测效果随候选框数目的增加有明显提升。

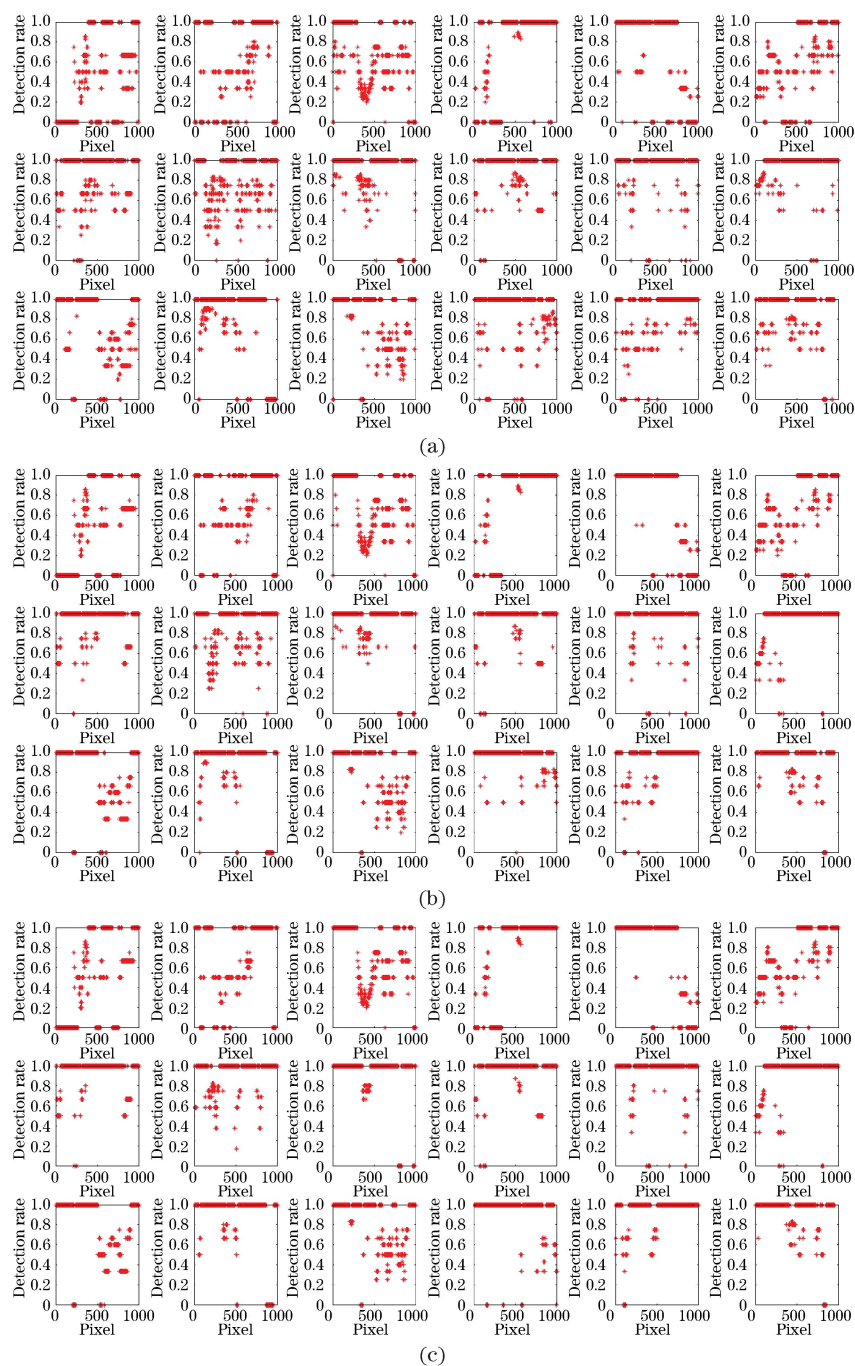


图4 每帧图像在不同候选框数目下的检测率。(a) 候选框数目为 5387；(b) 候选框数目为 8515；(c) 候选框数目为 14899

Fig. 4 Detection rates per image under different numbers of detection proposals.

(a) 5387 proposals; (b) 8515 proposals; (c) 14899 proposals

候选框的生成是后续目标识别、跟踪等操作的一项重要前期工作。利用 OL_GMPG 算法, 在每帧图片上设置 5387 个候选框时, 可以获得 81% 的检测率。在此参数设置下随机抽取测试集中的 4 张图片, 并画出采用 OL_GMPG 算法得出的 IOU 值最高的候选框, 标出重合度, 如图 5 所示。

同图 4 的参考图片和参数设置, 分别取三个 IOU 阈值 0.4, 0.5, 0.6, 得到不同 IOU 阈值下设置的候选框, 如图 6 所示:

4.3 OL_GMPG 方法与传统滑动窗方法的比较

4.3.1 传统滑动窗方法测试结果

使用类似的方法来度量传统滑动窗搜索方式的检测效果, 结果如图 7 所示。

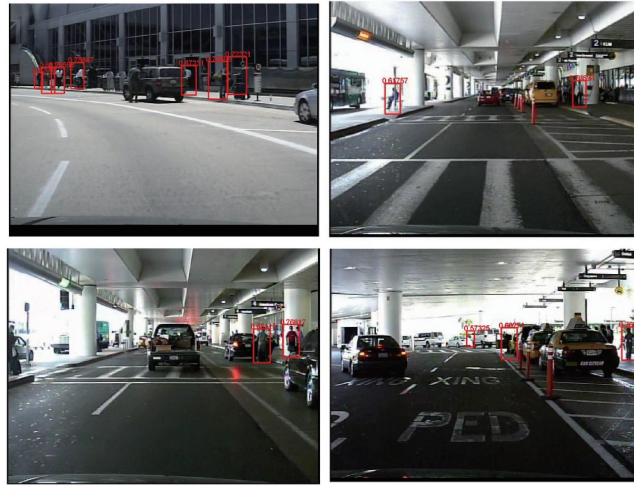


图 5 候选框的可视化表示(取最大 IOU 值)

Fig. 5 Visualization of detection proposals (take the largest value of IOU)

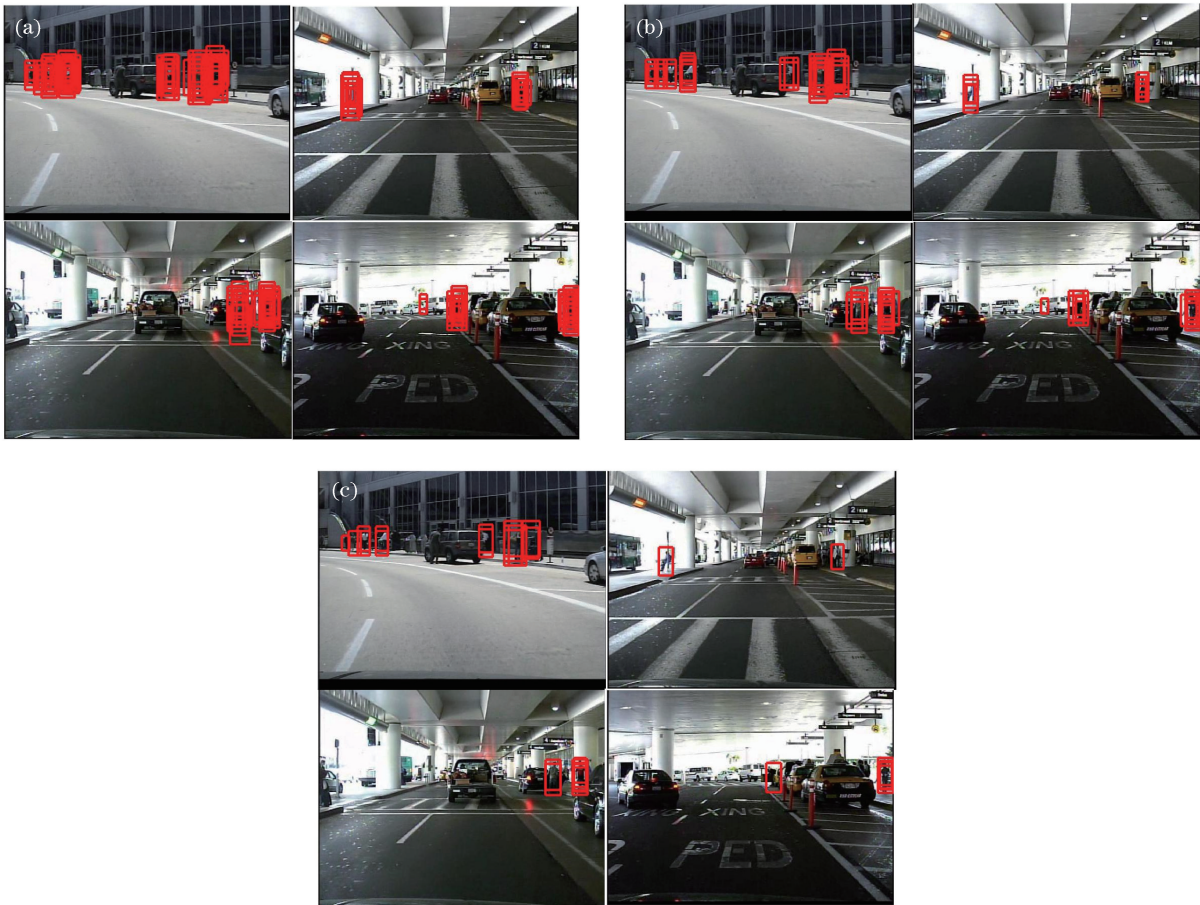


图 6 候选框的可视化表示。(a) IOU 阈值为 0.4；(b) IOU 阈值为 0.5；(c) IOU 阈值为 0.6

Fig. 6 Visualization of detection proposals. (a) IOU threshold is 0.4; (b) IOU threshold is 0.5; (c) IOU threshold is 0.6

参照文献[1],在尺寸为 480×640 的图像上均匀设置检测点。通过控制相邻搜索框质心之间的距离来改变候选框数目。为方便比较,基准尺寸与 OL_GMPG 算法设置相同的值,即 900,各检测点设置的检测框尺寸与基准尺寸的比例范围为 $0.6 \sim 2.6$ 。当每帧图片设置 14823 个候选框时,IOU 值大于 50% 的比率近似为 83%。准确度满足实际需求,但是超过 10^4 量级的检测框数目会降低检测系统的实时性。OL_GMPG 算法在设置 5387 个候选框时,便可获得 81% 的检测率。考虑到实时性和准确度之间的平衡,采用 OL_GMPG 算法生成行人候选框更加合理。

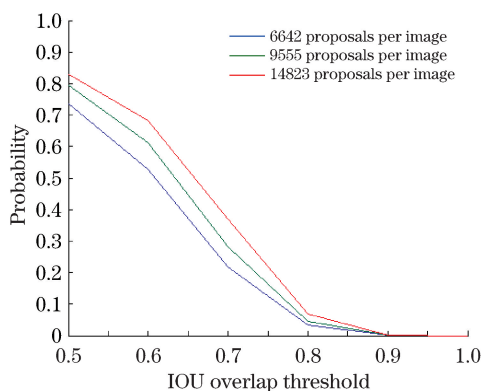


图 7 滑动窗搜索方法在不同数目的候选框下的测试结果

Fig. 7 Results of the sliding window method with different numbers of detection proposals

4.3.2 两种算法的比较结果

当 OL_GMPG 算法在每帧设置 5387 个候选框,传统滑动窗搜索方式在每帧设置 6642 个候选框时,二者的检测率随 IOU 阈值的变化情况如图 8 所示。OL_GMPG 算法可以在设置较少搜索框的同时达到更高的检测率。

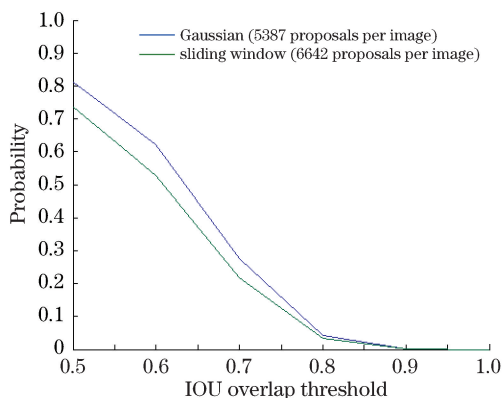


图 8 OL_GMPG 算法和滑动窗搜索方法在不同 IOU 阈值下的检测效果比较

Fig. 8 Comparison between the OL_GMPG algorithm and the sliding window method under different IOU thresholds

同时,用每帧图片的平均检测率随候选框数目的变化情况来度量算法的效果,并改变相邻搜索框间隔和改变尺度变化步长来控制候选框数量。使用改变相邻搜索框间隔的方法控制候选框数量,得出的效果比较如图 9 所示,相关数据如表 2 所示。使用改变尺度步长大小的方法控制候选框数量,得出的效果比较如图 10 所示,相关数据如表 3 所示。可见 OL_GMPG 算法与传统的滑动窗搜索方式相比较,在保证较高实时性的同时,可以获得理想的准确度。无论使用何种控制搜索框数目的方式,OL_GMPG 算法的检测效果均优于传统滑动窗搜索方式。

表 2 改变相邻搜索窗间隔时 OL_GMPG 算法与滑动窗搜索方法的检测数据比较

Table 2 Results of the OL_GMPG algorithm and the sliding window method by changing the interval of adjacent detection proposals

OL_GMPG		Sliding window method	
Number of proposals	Detection rate	Number of proposals	Detection rate
5387	0.8102726	6642	0.7339724
8515	0.8468724	9555	0.7931990
14899	0.8596323	14823	0.8301941
21286	0.8620764	26001	0.8412445

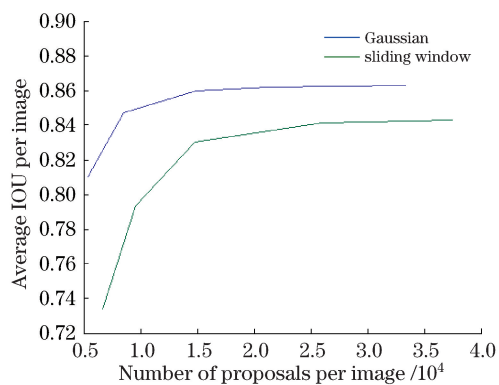


图 9 相邻搜索窗间隔时 OL_GMPG 算法与滑动窗搜索方法的检测效果比较

Fig. 9 Comparison between the OL_GMPG algorithm and the sliding window method under different intervals of adjacent detection proposals

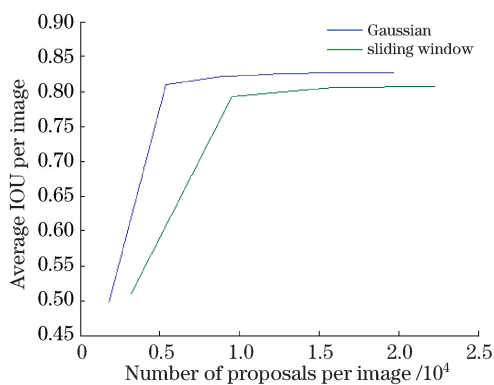


图 10 改变尺度变化步长时 OL_GMPG 算法与滑动窗搜索方法的检测效果比较

Fig. 10 Comparison between the OL_GMPG algorithm and the sliding window method by changing the scale step

表 3 改变尺度变化步长时 OL_GMPG 算法与滑动窗搜索方法的检测数据结果

Table 3 Results of the OL_GMPG algorithm and the sliding window method by changing the scale step

OL_GMPG		Sliding window method	
Number of proposals	Detection rate	Number of proposals	Detection rate
1818	0.4976346	3185	0.5099142
5387	0.8102726	9555	0.7931990
8971	0.8218085	15925	0.8058164
10766	0.8232002	22295	0.8076876

选择性搜索方法在 Pascal VOC 2007 数据集上获得了较好的测试效果。同样用选择性搜索方法来测试加州理工学院行人数据库分支 set00 数据集,具体测试结果如图 11 所示:

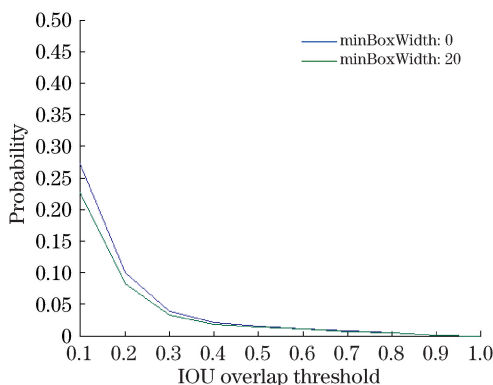


图 11 选择性搜索方法测试结果

Fig. 11 Results of the selective search method

选择性搜索方法的作者在测试 Pascal VOC 2007 数据集时,将参数 minBoxWidth 的值设为 20,以去除宽度或者高度小于该阈值的候选框。考虑到加州理工学院数据集是车载监控数据集,所拍摄行人的尺度较小,为了尽量避免漏检,还测试了阈值 minBoxWidth 取 0 的情况。采用的测试指标也是 IOU。该方法的测试结果如图 10 所示,可见选择性搜索方法在行人候选框生成的应用中效果并不理想,当 IOU 阈值取为 0.5 时,平均检测率只能达到 10^{-2} 数量级。当阈值 minBoxWidth 设置为 20 时,在所测试的 480×640 标准大小的图片上需要设置的候选框数目是 2.089345×10^4 。可见该方法在所用测试数据集上的计算复杂度较高,实时性不太理想。而 OL_GMPG 算法在设置 5387 个候选框时,便可获得 81% 的检测率。由于两种算法在应用于行人候选框的生成时,检测率和候选框数量的差异均较大,因此没有进行更多的比较分析。

5 结 论

OL_GMPG 算法充分利用静态监控场景与车载监控场景的固有特征,缩小滑动窗搜索的范围,该方法的优点主要如下。

1) OL_GMPG 算法是一种在线的行人候选框生成算法,当与后续的行人识别算法相结合时,可以做到行人候选框更新与行人识别同步进行,因此比较符合实际需求。

2) 在每帧图像设置较少数量的候选框,便可获得较高的检测率。OL_GMPG 算法在 480×640 大小的图片上设置 5387 个候选框时,获得了 81% 以上的检测率。一方面为后续的行人识别及跟踪过程提供了较为准确的候选框,另一方面不会影响整个系统的实时性。

3) OL_GMPG 算法通过训练图像各个像素点处基于尺度大小的高斯模型,不断更新模型的均值和标准差两个参量,使其更加符合该位置当前的尺度分布情况。在学习过程中,不仅得出若干行人频繁出现的位置,也得出了对应位置概率取值最大的尺度值,并根据这些尺度值调整候选框的大小。通过不同位置处尺度大小的比较,可以估计这些位置与车载摄像头之间的远近关系。

OL_GMPG 算法需要改进或者拓展之处包括:

1) 该方法针对静态监控场景与某些条件下的车载监控场景,即背景分布没有明显变动的情况,暂时不能处理背景剧烈变化的场景。

2) 没有加入后续的识别及跟踪算法,目前只是单纯地生成行人候选框。加入识别与跟踪算法可以看到 OL_GMPG 算法在综合应用中的效果。

3) OL_GMPG 算法可以得出相关位置的尺度信息,可以用来估计不同位置与摄像头之间的远近关系。

参 考 文 献

- Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 1: 886-893.
- Felzenszwalb P, Grishick R B, McAllister D, *et al.* Object detection with discriminatively trained part-based models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- Horn B K P, Schunck B G. Determining optical flow[J]. Artificial Intelligence, 1981, 17(1-3): 185-203.
- Terzopoulos D. Regularization of inverse visual problems involving discontinuities[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986(4): 413-424.
- Lucas B D, Kanade T. An iterative image registration technique with an application to stereo vision[C]. Proceedings of 7th International Joint Conference on Artificial Intelligence, 1981: 674-679.
- Deng Jinhao. Research of pedestrian detection algorithms based on video[D]. Guangzhou: South China University of Technology, 2011.
邓锦豪. 基于视频图像的行人检测算法研究[D]. 广州: 华南理工大学, 2011.
- Hsman H E. Hardware-based solutions utilizing random forests for object recognition[M]. // Köppen M, Kasabov N, Coinill G. Lecture notes in computer science description. Cham: Springer International Publishing, 2009: 760-767.
- Stauoer C, Grimson W E L. Adaptive background mixture models for real-time tracking[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1999, 2: 2246-2252.
- Arseneau S, Cooperstock J R. Real-time image segmentation for action recognition[C]. Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, 1999: 86-89.
- Collins R T, Lipton A J, Kanade T. Introduction to the special section on video surveillance[C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 745-746.
- Anderson C H, Bert P J, van der Wal G S. Change detection and tracking using pyramids transformation techniques[C]. SPIE, 1985(579): 72-78.
- Ai Haizhou, Lü Fengjun, Liu Wei, *et al.* Change detection and segmentation for visual surveillance[J]. Computer Engineering and Applications, 2001, 5: 75-77.
艾海舟, 吕风军, 刘 伟, 等. 面向视觉监视的变化检测与分割[J]. 计算机工程与应用, 2001(5): 75-77.
- Uijlings J R R, van de Sande K E A, Gevers T, *et al.* Selective search for object recognition[J]. International Journal of Computer Vision, 2013, 104 (2): 154-171.
- Tang Qing. Research of threshold segmentation algorithms and pedestrian detection on infrared image[D]. Guangzhou:

South China University of Technology, 2010.

唐 清. 阈值分割及红外图像行人检测研究[D]. 广州: 华南理工大学, 2010.

- 15 Liu Jian, Liu Yanan, Gao Enyang, *et al.* Human detection method based on foreground segmentation[J]. Journal of Chinese Computer Systems, 2014, 35(3): 654-658.
刘 剑, 刘亚楠, 高恩阳, 等. 基于前景分割的行人检测方法[J]. 小型微型计算机系统, 2014, 35(3): 654-658.
- 16 Liu Shumin, Huang Yingping, Zhang Renjie. Pedestrian contour extraction and its recognition using stereovision and snake models[J]. Acta Optica Sinica, 2014, 34(5): 0533001.
刘述民, 黄影平, 张仁杰. 基于立体视觉及蛇模型的行人轮廓提取及其识别[J]. 光学学报, 2014, 34(5): 0533001.
- 17 Gu Cheng, Qian Weixian, Chen Qian, *et al.* Rapid head detection method based on binocular stereo vision[J]. Chinese J Lasers, 2014, 41(1): 0108001.
顾 骋, 钱惟贤, 陈 钱, 等. 基于双目立体视觉的快速人头检测方法[J]. 中国激光, 2014, 41(1): 0108001.
- 18 Gerónimo D, López A M. Vision-based pedestrian protection systems for intelligent vehicles [M]. Cham: Springer International Publishing AG, 2014.
- 19 Cheng M M, Zhang Z M, Lin W L, *et al.* BING: Binarized normed gradients for objectness estimation at 300 fps[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 3286-3293.
- 20 Hogg R V, Craig A T. Introduction to Mathematical Statistics [M]. 5th edition. Upper Saddle River: Prentice Hall, 2004.
- 21 Chen Yin, Ren Kan, Gu Guohua, *et al.* Moving object detection based on improved single Gaussian background model [J]. Chinese J Lasers, 2014, 41(11): 1109002.
陈 银, 任 侃, 顾国华, 等. 基于改进的单高斯背景模型运动目标检测算法[J]. 中国激光, 2014, 41(11): 1109002.