

文章编号 1004-924X(2015)11-3227-11

视觉注意机制下的粒子窗快速目标检测

徐 超*, 高 敏, 杨锁昌, 方 丹, 卢志才
(军械工程学院 导弹工程系, 河北 石家庄 050003)

摘要:针对传统滑动窗目标检测方法需要在全图像范围内穷举搜索的缺点,提出了一种基于视觉注意机制的粒子窗检测方法,旨在保持较高检测精度的同时减少计算量。该方法将目标显著性作为先验知识引入搜索过程,采用“图像签名”方法生成显著图,然后通过阈值门限提取出包含目标真实位置的局部区域。利用蒙特卡洛采样在显著目标对应的图像范围内均匀生成粒子窗,并依据分类器的响应对粒子进行重采样,以凸显真实目标区域、避免滑动窗方法对搜索步长的依赖。建立了 Adaboost+类 Harr 特征(HLF)和支持向量机(SVM)+方向梯度直方图(HOG)的多级分类器结构,前级分类器用于大范围目标的快速筛选,后级分类器用于小范围目标的精确定位。将本文目标检测模型与传统滑动窗法和粒子窗法进行了比较,结果表明本文方法的受试者工作特征曲线(ROC)包含的面积更大,耗时仅为滑动窗法的 1/3 到 1/4,粒子窗法的 1/2,在保持较高检测精度的条件下显著提升了检测速度,实现了快速准确的目标检测。

关键词:目标检测;视觉注意机制;图像签名;粒子窗;多级分类器

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/OPE.20152311.3227

Visual attention mechanism-aided fast target detection by particle window

XU Chao*, GAO Min, YANG Suo-chang, FANG Dan, LU Zhi-cai

(Department of Missile Engineering, Ordnance Engineering College, Shijiazhuang 050003, China)

* Corresponding author, E-mail:475084845@qq.com

Abstract: As traditional sliding window detectors need to search the whole image by exhaustive method, a visual attention mechanism-aided target detection model by the particle window is proposed to reduce the calculational load while containing high detection accuracy. This model takes the target saliency as prior information of searching process, and then extracts the region of interest containing true target position by the “Image Signature” saliency map generator and entropy threshold. By uniformly drawing particle windows in an image range corresponding to the saliency targets with Monte Carlo sampling, the local region is treated as candidate detection points, thus resampling is carried out according to corresponding particle windows’ response. This strategy only focuses on the areas where the objects are potentially present and avoiding the tradeoff between accuracy and efficiency resulting from searching steps. A multi-stage classifier with Adaboost+HLF and SVM+HOG is established, the former is applied to once-over and the latter is used to locate precisely. The target detection model proposed is compared with the traditional sliding window method and particle window method, and the results show that the Receiver Operating Characteristic(ROC) curve by proposed method contains the area to be larger than that of the other methods and the time consuming is only 1/3 to 1/4 that of the sliding window method and 1/2 that of the particle window method. It increases significantly detection speeds at maintaining high precision detection speed and achieves fast

收稿日期:2015-08-10;修订日期:2015-10-10.

基金项目:军内科技创新项目(No. 2013562);军械工程学院科研基金资助项目(No. YJXM11018)

and accurate target detection.

Key words: target detection; visual attention mechanism; image signature; particle window; multi-stage classifier

1 引言

单帧图像的目标检测是计算机视觉中的基础问题之一^[1-2],其目的是将具有纹理和颜色特征的目标从杂乱无规律的背景中区分出来,但光照条件和背景等干扰因素使得该问题仍然面临着巨大挑战。当前的发展趋势是将目标检测转变为先验知识指导下的二分类问题^[3],利用预先训练的分类器将全图像区分为目标区域和非目标区域,并增强算法对复杂环境的适应能力。

二分类目标检测方法分为特征选择、分类学习和目标搜索定位等几个重要部分。其中,滑动窗是常用的目标搜索定位方法^[4],这类方法使用矩形窗口在位置和尺度空间内按步长依次滑动,对全图像进行遍历扫描,每一个矩形窗口对应一个候选目标,由分类器判断窗口所覆盖的区域是否通过检测。这种穷举搜索的方式虽可以保证不漏掉目标任何一个可能的位置,但是检测效率太低。文献[5]使用子窗口搜索策略对滑动窗算法进行了优化,其缺点是对目标函数有严格要求,并且每次搜索只能确定一个目标,需要重复多次才能完成多目标检测。文献[6-7]将多尺度策略应用到搜索过程,利用大步长在全图像寻找目标的大致位置,然后在分类器响应较强的候选目标附近以小步长对目标进行准确定位。这类方法的定位精度仍然依赖于最小步长,若最小步长设置过小,则检测效率提高不明显;若最小步长设置过大,则定位精度会下降。文献[8]引入统计理论,采用随机生成粒子的方式估计目标的后验概率密度函数,在理想情况下可以获得无限小步长才能达到的检测效果。但是它与上述几种滑动窗优化算法存在同样的问题,检测流程中缺少必要的目标位置提示环节,需要在全图像范围内完成搜索的初始化,使得大量窗口检测消耗在背景区域。文献[9]对位置提示环节进行了初步研究,尝试使用分割方法提取目标的大致轮廓,然后在候选点附近进行检测,但是复杂背景下的目标分割极易

出现错误,并且分割过程增加了整个算法的耗时,实用性不强。

本文通过对公共数据集的分析发现,由于目标的纹理和颜色特征明显,因此目标相对于背景都具有一定的显著性,以该特性作为先验知识对目标进行初始定位具有以下优点:1)视觉注意机制的相关研究提供了大量高效的显著性检测方法,尤其是基于频域的算法计算过程简单,能够以极快的速度提取显著目标区域;2)该区域只是搜索起始点,并不要求其必须覆盖目标整个轮廓,当背景复杂或目标不均匀时,能够避免分割方法无法获取完整目标。基于这种认识,本文将目标显著性和文献[8]的统计思想结合起来,建立了视觉注意机制下的粒子窗目标检测模型,在显著图等先验知识的指导下完成候选目标范围的初始化,后续的检测工作只需在目标真实位置附近展开;然后在离散的显著区域中优化粒子窗的生成和重采样方式,迭代实现快速准确的目标定位,大大降低了算法的计算量和对硬件设备的要求,满足了导弹、小型无人机等平台的快速检测需求。

目标检测的另一关键要素是选择合理的目标特征表达和分类器。Munder^[10]等对近年来应用较为广泛的主成分分析(PCA)、局部感受域(LRF)、局部轮廓特征(HOG)、HLF和Haar小波等特征,以及径向基神经网络(RBF)、支持向量机(SVM)、径向基-支持向量机(RBF-SVM)及Adaboost等分类器进行了整理分析,发现HOG是最有效的目标表达方式,而SVM和Adaboost是最准确的分类器。Enzweiler等^[11]通过对比Adaboost+HLF、RBF+LRF以及SVM+HOG^[12]等分类器组合进一步指出:在没有时间约束的前提下,SVM+HOG取得了最好的检测效果,而Adaboost+HLF虽然检测结果稍差,但是实时性超越了其他方法。因此,本文结合粒子窗的迭代搜索过程,建立了Adaboost+HLF和SVM+HOG多级分类器结构。考虑前级检测中的待检目标较多,使用弱分类器Adaboost和简单特征HLF快速去除大量非目标窗口来提高算

法的速度;而后级检测中候选目标窗口数量急剧减少,故使用强分类器 SVM^[13]和复杂特征 HOG^[14]进行精细筛选,保证了检测的精度。

本文的主要贡献可归纳为:(1)提出基于视觉显著性的位置提示方法,给定了目标搜索的起始点;(2)优化了粒子窗生成和重采样方式,使其适应离散的显著区域,提高了目标检测速度;(3)建立了 Adaboost + HLF 和 SVM + HOG 的多级分类器结构,应用由粗到精策略进一步增强了目标检测的时效性。

2 频域显著性模型分析及位置提示环节

引入位置提示环节的目的在于减少背景区域的窗口计算,这就要求提取显著图的耗时一定要少于背景区域中无效粒子窗的耗时。常用的 Itti、马尔科夫链、主动轮廓和图论等显著性模型计算方法复杂^[14-16],不适用于本文的位置提示环节。相比之下,运算速度具有优势的频域模型成为当前的研究热点之一。频谱残差(SR)^[17]模型将频域内的幅度谱残差作为显著性信息提取显著图。Guo 等^[18]认为显著区域取决于快速傅里叶变换(FFT)的相位谱,而不是幅度谱,并结合运动信息建立了四元数相位谱模型(PQFT)。Achanta 等^[19]提出了频率调谐方法(FT),通过求解差分高斯滤波器(DOG)滤波图像及其均值的欧氏距离得到显著图。Hou 等^[20]又利用“图像签名”算子(IS)对显著信息进行描述,在离散余弦变换(DCT)变换域中定义符号函数,通过对重构图像的平滑滤波提取显著区域。图 1(彩图见期刊电子版)给出了 4 种方法的运行时间比较结果。

2.1 不同显著性模型的对比分析

本节通过分析几种常用频域模型(SR、PQFT、FT、IS),选择合适的显著性模型作为位置提示环节,并对其进行了详细阐述。为了公平起见,本文从几种显著性方法的相关文献中各选一部分具有代表性的图片,建立公共的人眼跟踪测试数据库。数据库中包含 200 幅测试图像及其对应的人眼关注图,涵盖了不同场景和特点的目标。

在频域显著性模型的相关文献中,检测准确

度和速度是常用的评价标准,但作者均是在自己选定的图片库中进行算法测试,忽略了不同场景下算法的表现。因此,本文将算法的环境适应能力作为位置提示环节的重点关注指标,并分别从检测效果定量评价、运行时间对比和场景分析等 3 个方面综合衡量几种频域算法性能。

检测效果定量评价采用的是常见参数 AUC (Area Under the ROC Curve),公共数据库中 200 幅图像的 AUC 均值和标准差如表 1 所示。从 AUC 均值中可以看出,在提取显著目标时,IS 模型和 FT 模型明显优于 PQFT 模型和 SR 模型,并且 IS 模型的总体表现略好于 FT 模型。IS 模型、PQFT 模型和 SR 模型的 AUC 标准差明显小于 FT 模型,IS 模型最小,这表明了 IS 模型具有较好的一致性。4 种算法的运行时间对比如图 1 所示。由图 1 可知,FT 模型的平均运行时间为 19 ms,远远领先于 PQFT 模型和 SR 模型,而 IS 模型的平均运行时间只有 12 ms,比 FT 模型减少了 1/3。

表 1 不同频域模型的 AUC 均值和标准差对比

Tab. 1 Comparison of AUC mean and deviation among different spectral models

频域模型	AUC 均值	AUC 标准差
IS	0.750 2	0.091 7
FT	0.742 1	0.189 3
PQFT	0.675 4	0.108 3
SR	0.638 5	0.117 2

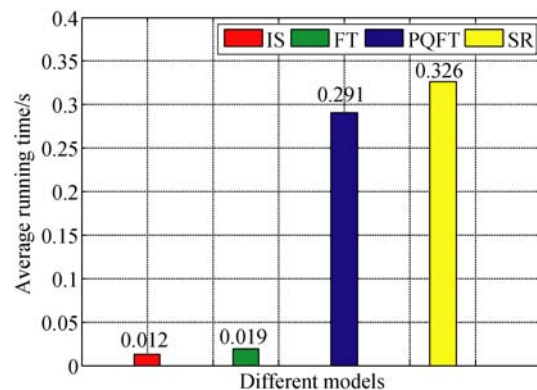


图 1 不同模型的平均运行时间

Fig. 1 Average running time of different models

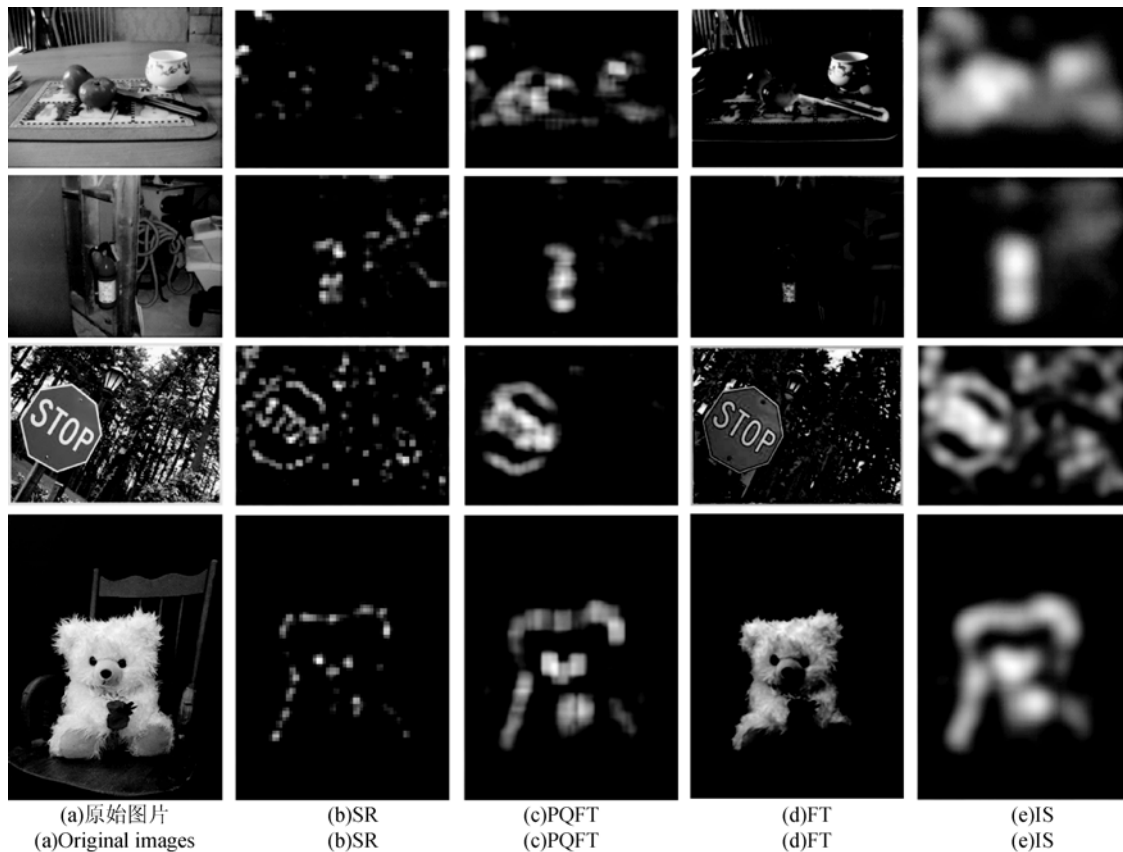


图 2 显著图视觉效果对比

Fig. 2 Visual effect comparison of saliency maps using different algorithms

FT 模型的 AUC 均值与 IS 模型非常接近, 但是其 AUC 标准差明显大于 IS 模型, 这表明在一部分场景中, FT 的检测准确度要优于 IS 模型。在公共数据库中选取几幅典型场景进行具体分析, 如图 2 所示。第一行和第二行的场景特点是小目标、复杂背景。对于这类目标, IS 模型可准确提取出显著区域, PQFT 模型和 SR 模型能检测出断续的目标, 而 FT 模型则很难区分出目标。第三行是具有细致纹理的目标, 第四行是背景简单的大目标, 此时, FT 模型能够提取出目标内部细节和完整的目标, 而 IS 模型、PQFT 模型和 SR 模型则只能检测出部分轮廓。

总的来说, IS 和 FT 模型相对于 PQFT 和 SR 模型具有明显优势, 其中 IS 模型在运行时间和总体检测准确度上稍微领先 FT 模型。在场景适应性方面, FT 模型不适于复杂背景检测, 其显著图中干扰太多, 很难在实际应用中准确地找到目标位置; IS 模型则不适于大目标及纹理检测, 但由于本文不关心目标内部纹理和完整轮廓, 只

要准确定位目标的一部分区域, 即可作为下一阶段的搜索起点, 以满足减少背景内的运算。因此, 综合衡量 4 种频域模型的性能, 本文选择 IS 模型作为位置提示环节的显著性方法。

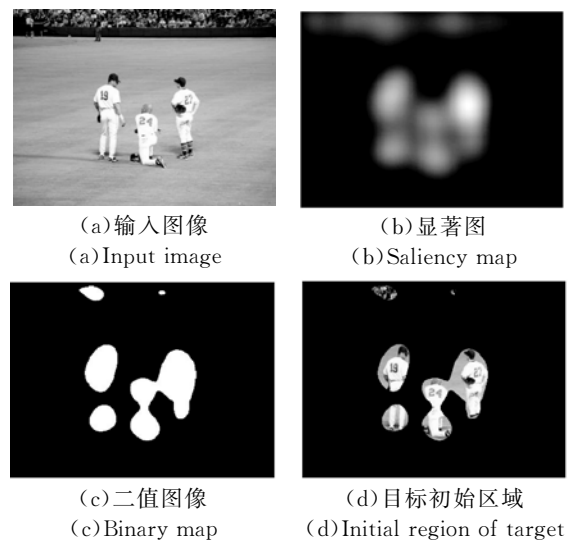


图 3 基于图像签名算法的初始提取

Fig. 3 Initial region extraction by IS model

2.2 基于图像签名算法的目标初始区域提取

利用图像签名算法提取显著图,然后使用自适应的阈值门限对其进行二值化处理,得到二值图像 b , 阈值 T_a 的选取原则参考文献[17]的前景检测方法,即:

$$T_a = \frac{\alpha}{W \times H} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} m(x, y), \quad (1)$$

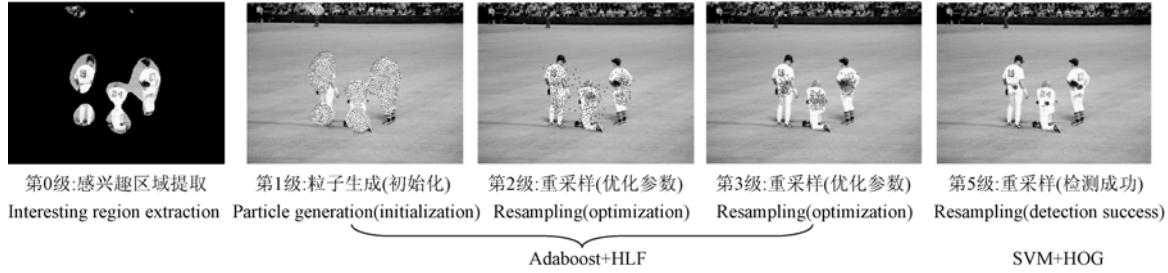


图 4 目标检测模型总体架构

Fig. 4 Whole framework of target detection model

3 视觉注意机制下的粒子窗目标检测模型

3.1 目标检测模型总体架构

本文利用位置提示环节将全图像随机搜索的目标检测问题转化为有监督的目标状态估计 $p(\mathbf{X}/\mathbf{Z})P(\mathbf{Z}/\mathbf{Y})$ 。其中 $\mathbf{X}=(\omega_x, \omega_y, \omega_s)$ 为粒子窗的 3D 向量,分别对应目标的位置坐标 x 、 y 和尺度 s , \mathbf{Y} 表示显著性先验知识, \mathbf{Z} 为对应的图像区域。目标检测模型的总体架构如图 4 所示,其中,在第 0 级中显著模型完成了目标区域初始化,后续的粒子窗检测可简化为 $p(\mathbf{X}/\mathbf{Z})$ 。粒子窗检测采用 4 级结构,第 1 级到第 3 级采用 Adaboost + HLF 分类器,第 4 级采用 SVM + HOG 分类器,各级递进的目标参数优化过程取决于分类器的响应和粒子的生成及重采样方法。

3.2 分类器响应

在 Adaboost 分类器 \mathbf{C} 中,窗口 ω 的检测结果是由其各层子分类器 C_i 的分类结果 $class(\omega, C_i) = \{\text{object}, \text{nonobject}\}$ 共同组成的,利用分类器响应 $R_A(\omega)$ 衡量最终的检测结果:

$$R_A(\omega) = \frac{j_\omega}{L}, \quad (2)$$

其中: L 表示级联分类器的层数, j_ω 表示最后一层检测结果为目标的子分类器编号。 $R_A(\omega)$ 的取

其中: W 、 H 是显著图的宽和高, $m(x, y)$ 为显著图在像素 (x, y) 上的值, α 为阈值系数。将二值图像 b 中的显著目标映射到原图像的对应位置,即可得到感兴趣区域 I 。基于图像签名算法的感兴趣区域提取结果如图 3 所示。

值范围为 $[0, 1]$, 其值越大,表明窗口与目标的相似度越高;当 $R_A(\omega) = 1$ 时,则该次检测成功,窗口覆盖的区域即为目标。

在 SVM 分类器中,窗口 ω 对应的输出定义为函数间隔 M :

$$M = W \cdot m + b, \quad (3)$$

其中: m 为窗口 ω 的特征,变量 W , b 用于定义 SVM 的超平面 $W \cdot m' + b = 0$, m' 为超平面上的数据特征。

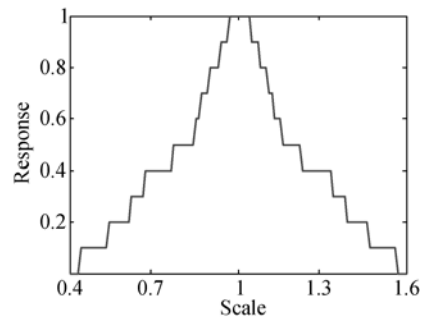
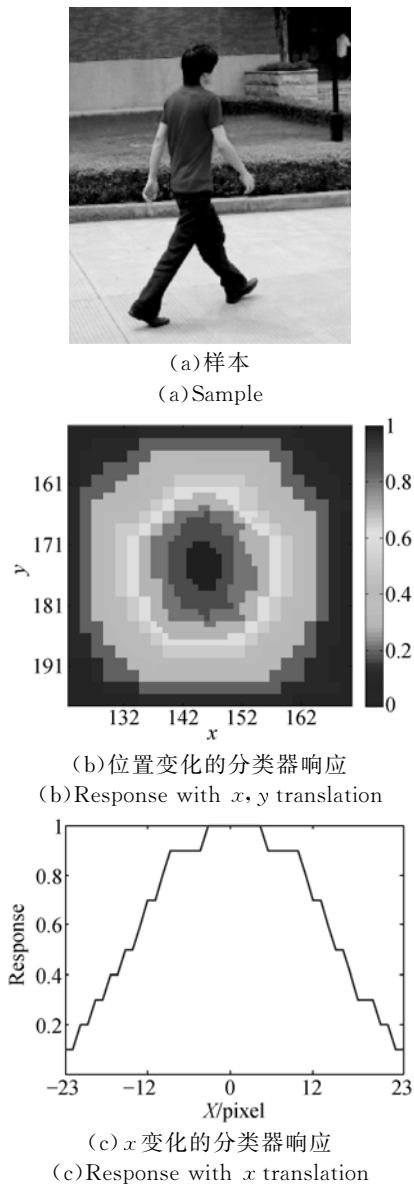
超平面将数据集分割为两类, $M > 0$ 的样本点为目标, $M < 0$ 的样本点为非目标,并且距离超平面越远其类别越明显,当 M 的取值为 $\pm \infty$ 时,可以完全确定样本为目标或者非目标。为了与 Adaboost 分类器响应一致,引入 sigmoid 函数将 M 转化成取值为 $[0, 1]$ 的分类器响应 $R_S(\omega)$:

$$R_S(\omega) = \frac{1}{1 + \exp(\alpha(M + c))}, M = [-\infty, +\infty], \quad (4)$$

其中,参数 α 和 c 可以通过训练数据集获得, α 的取值范围为 $(-\infty, 0)$, 以保证 sigmoid 函数单调递增。

分类器对位置和尺度具有良好的适应性,只要检测窗口出现在目标附近,分类器就会产生响应,并且窗口离目标越近,响应就越强烈,从而形成一个围绕着目标以 $R(\omega)$ 单调增加的“吸引域”。粒子窗模型中第 2 级到第 4 级的粒子重采

样都是在前级分类器指导下完成的,因此本文主要对 HLF+Adaboost 分类器的“吸引域”进行分析,获取该分类器的位置和尺度响应特性。以复旦大学行人检测数据库为例,选取目标位于中心的 50 幅图像,计算目标附近的 $R_A(\omega)$ 平均值,滑动窗口的位置间隔为 1 pixel,尺度变化为 0.01 倍,结果如图 5 所示。图 5(a)为数据库中的样本示例;图 5(b)为尺度固定不变时分类器响应随位置 x, y 的变化规律,“吸引域”的半径约为窗口大小的 20%;图 5(c)为位置“吸引域”在 x 方向的截面 $R_A(\omega) = f_1(x)$,用于近似两个位置方向独立的响应特性;图 5(d)为位置保持不变时分类器响应随尺度 s 的变化规律 $R_A(\omega) = f_2(s)$,”吸引域”的半径约为目标尺度的 50%。



(d) 尺度变化的分类器响应

(d) Response with scale translation

图 5 HLF+Adaboost 分类器的吸引域

Fig. 5 Attraction region of HLF+Adaboost classifier

3.3 粒子生成及重采样

位置提示环节在图像中标识出了可能出现目标的离散区域 $Area_1, Area_2, \dots, Area_N$, 则粒子的初始化建议分布 $q_0(\mathbf{X})$ 可以设定为各个离散区域对应的均匀分布 $U_1(\mathbf{X}), U_2(\mathbf{X}), \dots, U_N(\mathbf{X})$:

$$q_0(\mathbf{X}) = U_1(\mathbf{X}) + U_2(\mathbf{X}) + \dots + U_N(\mathbf{X}), \quad (5)$$

其中均匀分布中位置变量的取值范围即为离散区域本身,尺度的取值为 $[s/2, 2s]$, s 为各区域大小。

从 $q_0(\mathbf{X})$ 中采样 N_1 个粒子窗,每一个粒子窗代表一个候选目标。利用 HLF+Adaboost 分类器对各个粒子窗进行检测,并由检测结果得到近似的目标观测函数 $p_1(Z|\mathbf{X})$ 。其中,分类器响应为 1 的粒子确定为目标,将其记入集合 S ,不再作为下一级粒子重采样的起点。引入核密度估计理论将观测函数转化为 N_1' 个混合的高斯核函数, N_1' 为剩余粒子数量。对于第 j 个核函数来说,其均值 $\mu_i^{(j)}$ 设定为粒子 $p\omega_i^{(j)} = (\omega_{x_i}^{(j)}, \omega_{y_i}^{(j)}, \omega_{s_i}^{(j)})$ 。权值 $W_i^{(j)}$ 为 $R_A^\lambda(p\omega_i^{(j)})$,其中 $R_A(p\omega_i^{(j)})$ 为该粒子的分类器响应, λ_i 为第 i 级粒子窗的指数因子,在前级检测中 $\lambda_i \in (0, 1)$,使得响应较低的粒子不会直接被淘汰,在后级检测中 $\lambda_i \gg 1$,只保留响应较高的粒子,抑制非目标区域。协方差 $\Sigma_i^{(j)}$ 决定了粒子分布的范围,对粒子重采样至关重要,但文献[8]的协方差参数选取方式无法应对离散区域分布的随机性。本文通过分析分类器“吸引域”的发现,分类器响应与窗口偏离目标的程度成反比,因此将各个粒子对应的分类器响应 $R(\omega)$ 作为协方差矩阵选取的依据:

$$\Sigma_i^{(j)} = \text{diag}(|f_1^{-1}(R(\omega))|,$$

$$|f_1^{-1}(R(\omega))|, |\Delta f_2^{-1}(R(\omega))|). \quad (6)$$

第 i 级检测的目标观测函数 $p_i(Z|\mathbf{X})$ 可以由

上述核函数变量获得:

$$p_i(Z | \mathbf{X}) = \sum \pi_i^{(j)} \cdot N(\mu_i^{(j)}, \Sigma_i^{(j)}), \quad (7)$$

其中 $\pi_i^{(j)}$ 为归一化的权值。

$$\pi_i^{(j)} = \frac{W_i^{(j)}}{\sum W_i^{(j)}}. \quad (8)$$

似然函数 $p_i(\mathbf{X} | Z)$ 可以由目标观测函数 $p_i(Z | \mathbf{X})$ 获得,进而得到新的建议分布。为了控制粒子更新频率、防止漏检,在计算新的建议分布 $q_i(\mathbf{X})$ 时通常保留部分前一级的建议分布 $q_{i-1}(\mathbf{X})$:

$$q_i(\mathbf{X}) = (1 - \alpha_i) q_{i-1}(\mathbf{X}) + \alpha_i \frac{p_i(Z | \mathbf{X})}{\int p_i(Z | \mathbf{X}) d\mathbf{X}}, \quad (9)$$

其中 α_i 是自适应系数,随着检测的递进,其值逐渐减小。

下一级检测中对建议分布 $q_i(\mathbf{X})$ 进行重采样,并重复上述检测过程直至最后一级。在第 4 级中,利用 HOG+SVM 分类器对重采样粒子以及集合 S 中的粒子进行精确检测,并将分类器响应为 1 的粒子重新记入集合 S 。检测过程中经常出现多个窗口相互重叠的现象,此时,采用文献 [20] 中 SKDA (Sequential Kernel Density Approximation) 方法对窗口区域进行非极大值抑制,得到最终的目标似然函数 $p(\mathbf{X} | Z)$,函数的峰值对应目标检测结果。

3.4 算法流程

本文目标检测模型的算法流程如下:

(1)位置提示环节: $i=0$,利用“图像签名”方法生成显著图 m ,然后利用公式(1)得到阈值门限 T_a ,提取出包含目标真实位置的局部区域,并在该区域内采用公式(5)对建议分布 $q_0(\mathbf{X})$ 进行初始化。

(2)前级检测: $i=1:3$,从 $q_{i-1}(\mathbf{X})$ 中采样 N_i 个粒子窗,利用公式(2)中的 Adaboost+HLF 分类器对粒子窗进行检测,分类器响应 $R_A(\omega)$ 为 1 的粒子计入集合 S ,不参与下一级的粒子重采样。引入核密度估计理论将剩余粒子的观测函数转化为高斯核函数 $p_i(Z | \mathbf{X})$,然后将其用于公式(9)中,得到新的建议分布 $q_i(\mathbf{X})$,从而逐步缩小目标搜索范围。

(3)后级检测: $i=4$,从 $q_{i-1}(\mathbf{X})$ 中采样 N_i 个粒子窗,利用公式(4)中的 SVM+HOG 分类器

对新采样粒子和集合 S 中的粒子进行检测,将分类器响应 $R_S(\omega)$ 为 1 的粒子重新记入集合 S ,并利用 SKDA 非极大值抑制方法压缩重叠窗口的范围,从统计的角度实现最终的目标检测。

4 测量实验与结果

为了验证本文算法的性能,分别对行人目标和车辆目标进行检测和分析。其中行人目标来自 INRIA 数据库,车辆目标来自 INRIA 和 Caltech 数据库,图片大小按照 128 pixel×128 pixel 的基准缩放。引入检测率 (DR)、虚警率 (FPR) 和 ROC (Receiver Operating Characteristic) 曲线作为算法性能的评价指标。整个程序的硬件运行平台为 Intel 2.5GHz CPU+4G 内存,软件运行平台为 WindowsXP+Matlab 2011b。

4.1 综合性能测试

将本文算法 (V-PW) 和经典的滑动窗算法 (SW)、粒子窗算法 (PW) 进行比较,实验中选取了 3 种分类器结构 (Adaboost+HLF、SVM+HOG、多级分类器),检测结果如图 6 和表 2 所示。图 6(a) 和图 6(b) 分别为行人检测和车辆检测的 ROC 曲线,从两幅 ROC 曲线图可以看出, V-PW (Adaboost+HLF) 和 V-PW (SVM+HOG) 明显优于具有同样分类器的 SW 和 PW, V-PW (多级分类器) 比 V-PW (Adaboost+HLF) 和 V-PW (SVM+HOG) 的性能有更大提升。相比而言,车辆检测的整体性能要高于行人检测。这种现象是行人和车辆数据库的特征差异造成的,下面的场景分析部分将对其进行详细讨论。

表 2 是给定检测精度下的检测结果,第 1~8 行是行人检测,第 9~16 行是车辆检测。在行人检测中,第 1~3 行使用的是 Adaboost+HLF 分类器,第 4~6 行使用的是 SVM+HOG 分类器,同一分类器对应的各个检测算法具有相同的精度 (检测率)。以 Adaboost+HLF 分类器为例,首先从实用化角度选取 SW 的精度,以取得检测性能和运行时间的平衡,然后通过调整 PW 和 V-PW 的粒子窗数目达到与 SW 一致的检测精度。PW 和 V-PW 的粒子窗数目 N_w 与初始粒子数目 N_p 和级数 i 有关,PW 的级数为文献 [8] 中设定的 5 级,本文方法的级数为 4 级,每一级的粒子窗

数目设定为 $N_i = N_p / 2^{i-1}$, 则 $N_w = \sum N_i$ 。当检测率为 0.683 时, SW 使用的滑动窗口数量为 30 181, 而 PW 依靠粒子重采样策略将窗口数量减少至 8 000, 其耗时仅为前者的 48.61%, 虚警率与其相当。V-PW 利用显著性模型完成了目标区域初始化, 节省了背景内的计算时间, 但该区域通常只能涵盖目标的一部分, 因此其初始化粒子一般略多于 PW 第二级的粒子, 粒子窗总数约为 PW 的一半, 检测过程的耗时为位置提示环节和粒子窗模型之和, 约为 PW 的 54.4%。同时, 位置提示环节将搜索区域集中在真实目标周围, 大大降低了虚警的概率, 仅为 PW 的 70%。SVM+HOG 分类器的结果同样证实了 V-PW 的检测性能和效率优于 SW 和 PW。

为了验证本文多级分类器的性能, 在第 7 和第 8 行设置了 V-PW(多级分类器)的实验, 检测精度分别与第 3 行和第 6 行相同, 其粒子窗总数为 $N_w + N_s$, N_s 为集合 S 中的粒子数。与第 3 行相比, V-PW(多级分类器)增加了 N_s 个粒子, 且有 $N_4 + N_s$ 个粒子通过复杂分类器, 但后级分类器提高了检测精度, 降低了粒子窗总数, 耗时还略微下降, 其虚警率则仅为前者的 40%。与第 6 行相比, V-PW(多级分类器)的前级分类器大大降低了计算复杂度, 且该分类器比强分类器的漏检率要低, 同等精度下进一步减少了粒子数, 其耗时仅为第 6 行的 11.5%, 虚警率则为 58.8%。车辆检测的结论与行人检测相同, 这里不再赘述。

4.2 场景分析

本文方法对行人和车辆数据库中典型场景的检测结果如图 7 和图 8 所示, 第 1 行是待检图片, 第 2 行是显著图, 第 3 行是感兴趣区域, 第 4 行是标记的检测窗口。从第 2、3 行可以看出, 本文方法不以分割单个目标为目的, 而是利用显著性提取图像中的感兴趣区域, 大大降低了背景干扰的影响, 有效获取了部分或全部目标区域, 缩小了目标搜索的范围。但是不同场景中感兴趣区域和检测结果具有较大差异: 1) 图 7 第 1、2、4 和图 8 第 1、3 个场景中部分背景的显著性较强, 增加了无效显著区域的面积和检测时间; 2) 图 7 第 3、4 个场景较为复杂, 其中第 3 个场景中部分目标重叠严重, 无法检测, 而第 4 个场景中部分行人的尺度过小、显著性较低, 提取感兴趣区域时漏掉了该目标; 3) 图 8 第 2、3、4 个场景中单个目标的感兴趣区域相对于整个图像较大, 削弱了检测速度的提升。

由上述分析可知, 行人数据库中背景环境更为繁杂, 目标尺度和对比度变化范围更大, 比车辆数据库更容易出现漏检现象。车辆数据库中目标更大, 单个目标检测速度的提升小于行人数据库, 但行人数据库的行人数量更多, 单幅图像检测速度提升又小于车辆数据库。总的来说, 本文方法更适于显著性较强的中等大小目标检测, 但是对各种复杂环境同样有较强的适应性, 能够获得良好的检测效果。

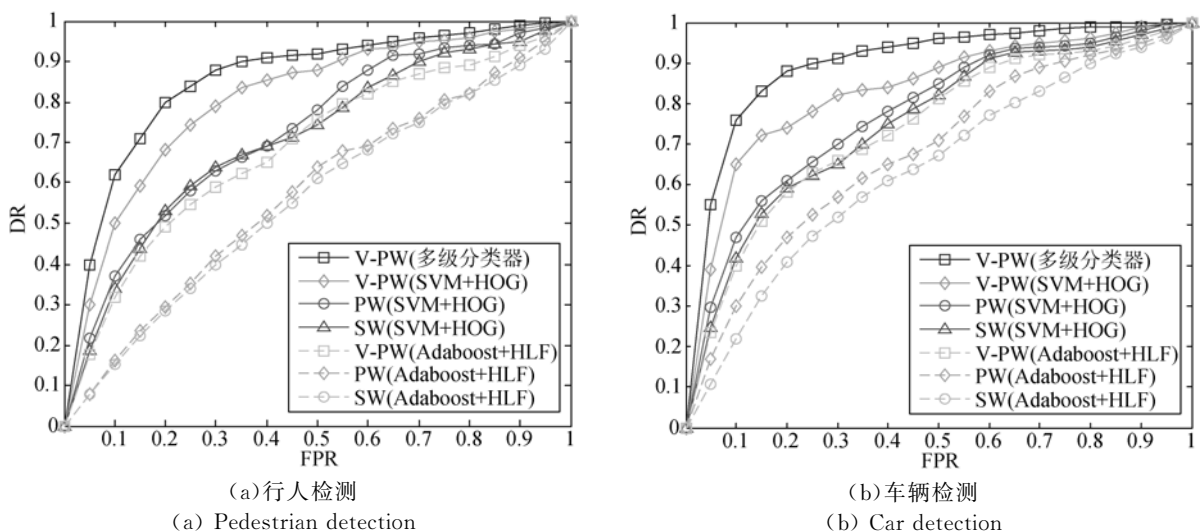


图 6 几种检测算法在行人和车辆检测中的 ROC 曲线

Fig. 6 ROC curves of V-PW, PW and SW in pedestrian and car data sets

表 2 行人和车辆检测结果

Tab. 2 Results of pedestrian and car detection

序号	数据库	分类器	检测模型	窗口数量	TP	TF	时间/s	
1			SW	30181	0.683	0.623	0.397	
2	行人检测数据库	Adaboost+HLF	PW	8000(4129/2064/1032/516)	0.683	0.617	0.193	
3			V-PW	3767(2153/1076/538)	0.683	0.432	0.105	
4		SVM+HOG	SW	21688	0.819	0.583	2.481	
5			PW	3415(1762/881/440/222/110)	0.819	0.515	1.832	
6		多级分类器	V-PW(A, a)	V-PW	1691(902/451/226/112)	0.819	0.354	0.985
7				V-PW(A, a)	1328(708/354/177/89)	0.683	0.173	0.093
8			V-PW(S, a)	V-PW(S, a)	1576(840/420/210/106)	0.819	0.208	0.113
9				SW	35272	0.768	0.604	0.421
10	车辆检测数据库	Adaboost+HLF	PW	5892(3041/1520/760/381/190)	0.768	0.571	0.175	
11			V-PW	3070(1638/819/409/204)	0.768	0.382	0.103	
12		SVM+HOG	SW	26679	0.843	0.494	2.757	
13			PW	4972(2566/1283/641/321/161)	0.843	0.435	1.564	
14		多级分类器	V-PW(A, a)	V-PW	2437(1300/650/325/162)	0.843	0.312	0.827
15				V-PW(A, a)	1597(851/425/214/107)	0.768	0.114	0.088
16			V-PW(S, a)	V-PW(S, a)	1965(1048/524/262/131)	0.843	0.165	0.097

注:A. a 表示与 Adaboost+HLF 分类器精度相同;S. a 表示与 SVM+HOG 分类器精度相同。



图 7 本文方法在行人数据库中的检测结果示例

Fig. 7 Detection results of proposed algorithm on pedestrian dataset

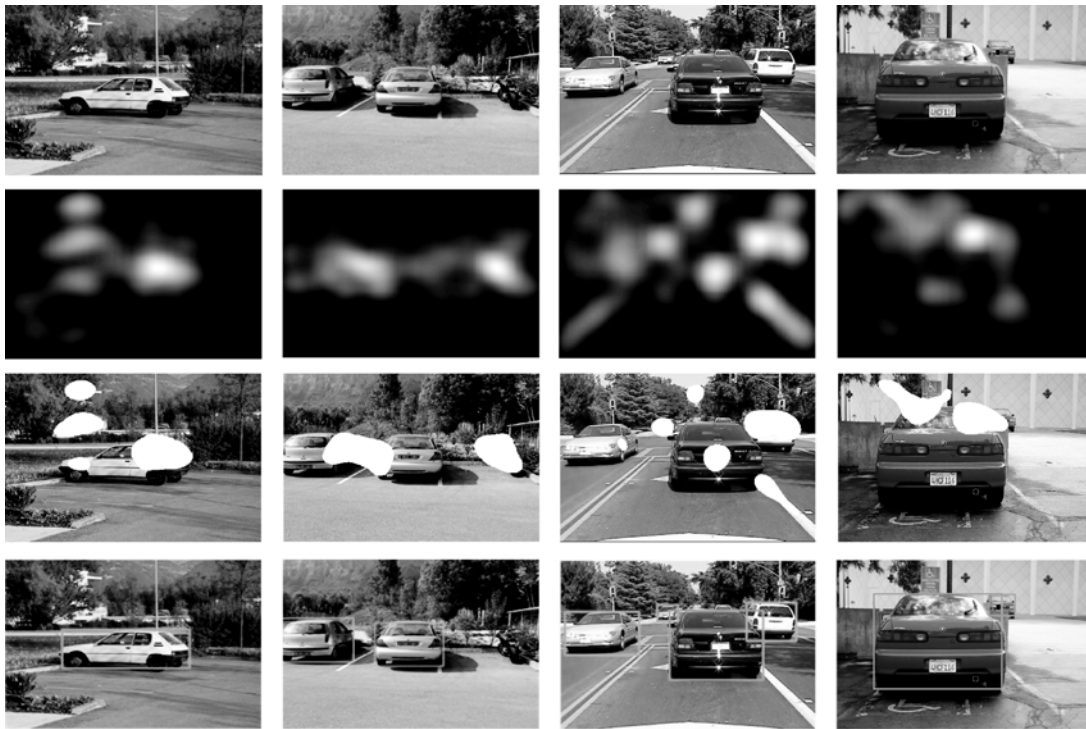


图 8 本文方法在车辆数据库中的检测结果示例

Fig. 8 Detection results of proposed algorithm on car dataset

5 结 论

本文将位置提示环节引入检测流程,建立了视觉注意机制下的粒子窗目标检测模型。通过提取目标感兴趣区域来快速定位目标范围,大大减少了背景中检测窗口的数量。显著性模型的选取直接影响位置提示环节的时效性,通过对不同显著性方法的分析发现,“图像签名”算法具有较强的环境适应能力,且平均耗时仅为 12 ms,完全满足位置提示环节的需求。故本文选择图像签名算法作为显著性模型在感兴趣区域中对粒子窗采样

方法进行了优化,根据分类器响应自主设定粒子散布的协方差矩阵,以适应位置和大小随机分布的离散区域,使粒子向目标真值逼近。将本文目标检测模型与传统滑动窗和粒子窗方法进行比较,结果表明本文方法 ROC 曲线包含的面积更大,耗时仅为滑动窗法的 1/3 到 1/4,大约为粒子窗的 1/2,结果证明,本文方法在保持较高检测精度的条件下显著提升了检测速度,实现了实时目标检测。同时,本文结合粒子窗迭代过程建立了多级分类器结构,与 Adaboost 和 SVM 分类器的对比实验证实了多级分类器能够进一步提高目标检测的时效性。

参考文献:

- [1] ZHANG Y X, DU B, ZHANG L P. A sparse representation-based binary hypothesis model for target detection in hyperspectral images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(3): 1346-1354.
- [2] VAZQUEZ D, LOPEZ A M, MARIN J, *et al.*. Virtual and real world adaption for pedestrian detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(4): 797-809.
- [3] 黄炎, 范赐恩, 朱秋平, 等. 联合梯度直方图和局部二值模式特征的人体检测[J]. *光学精密工程*, 2013, 21(4): 1047-1053.
- [4] HUANG Y, FAN C E, ZHU Q P, *et al.*. HOG-LBP pedestrian detection [J]. *Opt. Precision Eng.*, 2013, 21(4): 1047-1053. (in Chinese)
- [4] PEDERSOLI M, GONZALEZ J, HU X, *et al.*. Toward real-time pedestrian detection based on a deformable template model [J]. *IEEE Transactions on Intelligence Transportation Systems*, 2014, 15(1): 355-364.
- [5] LAMPERT C H, BLASCHKO M B, HOFMANN T. Efficient subwindow search; a branch and bound

- framework for object localization [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(12): 2129-2142.
- [6] YANG B, LEI Y Q. Vehicle detection and classification for low-speed congested traffic with anisotropic magnetoresistive sensor [J]. *IEEE Sensors Journal*, 2015, 15(2): 1132-1138.
- [7] PEDERSOLI M, GONZALEZ J, BAGDANOV A D, et al.. Recursive coarse-to-fine localization for fast object detection [C]. *Proceedings of 11th European Conference on Computer Vision*, 2010: 280-293.
- [8] GUALDI G, PRATI A, CUCCHIARA R. Multistage particle windows for fast and accurate object detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(6): 1589-1604.
- [9] 杨扬, 李善平. 分割位置提示的可变形部件模型快速目标检测 [J]. *自动化学报*, 2012, 38(4): 540-548.
- YANG Y, LI SH P. Fast object detection with deformable part models and segment locations' hint [J]. *Acta Automatic Sinica*, 2012, 38(4): 540-548. (in Chinese)
- [10] MUNDER S, GAVRILA D. An experimental study on pedestrian classification [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(11): 1863-1868.
- [11] ENZWEILER M, GAVRILA D. Monocular pedestrian detection: survey and experiments [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(12): 2179-2195.
- [12] 孙锐, 陈军, 高隽. 基于显著性检测与 HOG-NMF 特征的快速行人检测方法 [J]. *电子与信息学报*, 2013, 35(8): 1921-1926.
- SUN R, CHEN J, GAO J. Fast pedestrian detection based on saliency detection and HOG-NMF features [J]. *Journal of Electronics & Information Technology*, 2013, 35(8): 1921-1926. (in Chinese)
- [13] SUN W, ZHAO C, SUN M. Learning based particle filtering object tracking for visible-light systems [J]. *International Journal for Light and Electron Optics*, 2015, 126(19): 1830-1837.
- [14] 杨利平, 辜小花. 用于人脸识别的相对梯度直方图特征描述 [J]. *光学精密工程*, 2014, 22(1): 152-159.
- YANG L P, GU X H. Relative gradient histogram features for face recognition [J]. *Opt. Precision Eng.*, 2014, 22(1): 152-159. (in Chinese)
- [15] 钱生, 陈宗海, 林名强, 等. 基于条件随机场和图像分割的显著性检测 [J]. *自动化学报*, 2015, 41(4): 711-724.
- QIAN SH, CHEN Z H, LIN M Q, et al.. Saliency detection based on conditional random field and image segmentation [J]. *ACTA AUTOMATIC SINICA*, 2015, 41(4): 711-724. (in Chinese)
- [16] 修春波, 魏世安. 显著性直方图模型的 Camshift 跟踪方法 [J]. *光学精密工程*, 2015, 23(6): 1749-1757.
- XIU CH B, WEI SH A. Camshift tracking with saliency histogram [J]. *Opt. Precision Eng.*, 2015, 23(6): 1749-1757. (in Chinese)
- [17] 黎万义, 王鹏, 乔红. 引入视觉注意机制的目标跟踪方法综述 [J]. *自动化学报*, 2014, 40(4): 561-576.
- LI W Y, WANG P, QIAO H. A survey of visual attention based methods for object tracking [J]. *ACTA AUTOMATIC SINICA*, 2014, 40(4): 561-576.
- [18] GUO C, ZHANG L. A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression [J]. *IEEE Transactions on Image Processing*, 2010, 19(1): 185-198.
- [19] ACHANTA R, HEMAMI S, ESTRADA F, et al.. Frequency-tuned salient region detection [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami; FL*, 2009: 1597-1604.
- [20] HOU X D, HAREL J, KOCH C, et al.. Image signature: highlighting sparse salient regions [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(1): 194-201.

作者简介:



徐 超(1987—),男,安徽阜阳人,博士研究生,2009年、2011年于军械工程学院分别获得学士、硕士学位,主要研究方向为计算机视觉和图像末制导技术。E-mail: 475084845@qq.com

导师简介:



高 敏(1963—),男,山西临汾人,博士,教授,1983年于军械工程学院获得学士学位,1992年于北京理工大学获得博士学位,主要研究方向为计算机视觉和图像末制导技术。E-mail: gaomin1103@gmail.com

(版权所有 未经许可 不得转载)