

基于轮廓几何稀疏表示的刚性目标模型及其分级检测算法

林煜东¹ 和红杰¹ 陈帆¹ 尹忠科²

摘要 刚性目标轮廓具有明显几何特性且不易受光照、纹理和颜色等因素影响。结合上述特性和图像稀疏表示原理,提出一种适用于刚性目标的分级检测算法。在基于部件模型 (Part-based model, PBM) 的框架下,采用匹配追踪算法将目标轮廓自适应地稀疏表示为几何部件的组合,根据部件与目标轮廓的匹配度,构建描述部件空间关系的有序链式结构。利用该链式结构的有序特性逐级缩小待检测范围,以匹配度为权值对各级部件显著图进行加权融合生成目标显著图。PASCAL 图像库上的检测结果表明,该检测方法对具有显著轮廓特征的刚性目标有较好的检测结果,检测时耗较现有算法减少约 60%~90%。

关键词 目标检测, 刚性目标轮廓, 几何稀疏表示, 分级检测

引用格式 林煜东, 和红杰, 陈帆, 尹忠科. 基于轮廓几何稀疏表示的刚性目标模型及其分级检测算法. 自动化学报, 2015, 41(4): 843–853

DOI 10.16383/j.aas.2015.c130431

A Rigid Object Detection Model Based on Geometric Sparse Representation of Profile and Its Hierarchical Detection Algorithm

LIN Yu-Dong¹ HE Hong-Jie¹ CHEN Fan¹ YIN Zhong-Ke²

Abstract The profile of rigid objects has the geometrical characteristic and is insusceptible to illumination, texture or color. In this paper, a hierarchical detection algorithm for ridge objects based on geometric sparse representation of profile is presented. In the framework of part-based model (PBM), the object profile is automatically divided into geometrical parts by the sparse representation using the matching pursuit algorithm. To describe the spatial relationship of the geometrical parts, an ordered chain-like structure is constructed according to the order of the matching degree of the parts and the object profile. With the ordered chain-like structure, the detection range is gradually shrunk at each hierarchy. The final salient map of the object is the weighted summation of the parts' salient maps, and the weights are defined as the matching degrees. The simulation on the PASCAL datasets shows that the proposed method outperforms the existing models in rigid objects detection, and saves 60% to 90% detection time compared to the state-of-art methods.

Key words Object detection, profile of rigid object, geometric sparse representation, hierarchical detection

Citation Lin Yu-Dong, He Hong-Jie, Chen Fan, Yin Zhong-Ke. A rigid object detection model based on geometric sparse representation of profile and its hierarchical detection algorithm. *Acta Automatica Sinica*, 2015, 41(4): 843–853

目标检测是指利用图像处理和模式识别技术,从图像中找出已知目标对象的位置,其广泛应用于道路监控、军事侦察、生产线调控等领域。目标检测的方法种类较多,如基于视觉显著性^[1–3]、基于形状匹配^[4–5]、基于特征分类^[6–7]、基于投票^[8–9]等。Li 等^[10]对现有的检测模型进行综合分析,并总结为四类,分别为词袋模型 (Bag of word model)^[11]、基于部件模型 (Part-based model, PBM)^[12]、辨别模

型 (Discriminative model)^[13]和识别分割混合模型 (Combined recognition and segmentation)^[14]。本文主要研究基于部件的模型。

基于部件模型主要包括两部分:部件和部件间的空间关系。在部件选取方面,目前大部分的 PBM 以待检测目标的局部特征作为部件。这类部件通常易受光照、颜色、纹理和背景的影响,因此必须通过训练的方法^[15–16]或构建复杂特征集的方法^[17–18]来增强算子的描述能力,增加模型的复杂度。事实上,生活中大部分的待检测目标都是刚性目标。所谓刚性目标是指目标在一定的外力作用下,其形状和大小能保持不变,而且内部各部件相对位置保持恒定^[19]。刚性目标具有固定轮廓,且轮廓不易受光照、纹理和颜色等因素影响,目标的“轮廓特征”对查找目标具有重要作用。同时,刚性目标的轮廓具有明显的几何特性,采用稀疏表示方法能将其表示为少数几何图形的组合,简化目标模型。另外,在部

收稿日期 2013-05-22 录用日期 2014-11-15
Manuscript received May 22, 2013; accepted November 15, 2014
国家自然科学基金 (61461047) 资助
Supported by National Natural Science Foundation of China (61461047)

本文责任编辑 王飞跃
Recommended by Associate Editor WANG Fei-Yue
1. 四川信号与信息处理重点实验室, 西南交通大学 成都 610031 2. 北京遥感信息研究所 北京 100192
1. Sichuan Key Laboratory of Signal and Information Processing, Southwest Jiaotong University, Chengdu 610031 2. Beijing Institute of Remote Sensing Information, Beijing 100192

件间空间结构方面, 尽管目前已有较成熟的成果, 如 Fergus 等的星座结构^[20]、Crandall 等的 k -fans 结构^[21]、Felzenszwalb 等的树结构^[22] 和 Carneiro 等的稀疏柔性模型^[23] 等, 但它们都忽略了部件的差异性, 导致检测时需要对所有部件在所有位置上进行匹配检测, 增加检测时耗。

针对上述问题, 本文以部件模型为框架, 提出了一种适合于刚性目标的分级检测算法. 该算法利用稀疏表示理论将目标轮廓自适应地表示为少数几何图形(部件)的组合. 按部件与目标轮廓的匹配度构建部件的有序链式结构. 利用该结构的有序特性, 逐级缩小待检测范围, 并以匹配度为权值对各部件显著图进行加权融合, 生成目标显著图. 仿真结果表明, 本文提出的方法能根据目标轮廓复杂程度自适应选取部件, 对具有明显轮廓的刚性目标有较高的检测率和检测效率。

1 轮廓稀疏表示及有序链式结构

在传统的部件模型中, 目标各部件的局部特征和部件之间的关系是通过训练方法^[15-16] 得到. 以此建立的模型虽然能抵抗背景的干扰, 准确捕捉目标特征, 并适应一定的形变, 但训练增加建模时间. 刚性目标具有固定的轮廓, 部件间关系相对稳定; 轮廓具有明显的几何外观, 若用多个几何图形逼近表示, 即可获得其准确的边缘信息. 借助以上特点, 采用轮廓几何稀疏表示方法对刚性目标建模, 可避免复杂的训练过程, 提高建模速度。

轮廓稀疏表示方法首先需要对待检测目标样本图像(如图 1(a))进行二值化, 获得待检测目标的二值轮廓(如图 1(b)). 对背景简单的待检测目标样本图像(如图 1(a)中的瓶子和自行车), 二值化可采用 Otsu 阈值法实现. 而对背景复杂的待检测目标样本图像(如图 1(a)中的小轿车和帆船), 为给后面的几何表示提供准确的目标轮廓, 采用人工勾画的方法. 虽然需要人工参与, 但该勾画只对背景复杂的待检测目标样本图像执行, 且对这些目标的提取只需要进行一次, 与训练方法相比, 不会增加建模时间. 图 1 给出四种待检测目标的几何描述, 其中, 图 1(a)为待检测样本目标图像, 图 1(b)为待测样本目标图像二值化得到的目标轮廓, 图 1(c)为目标轮廓几何表示的近似逼近, 不同灰度表示不同的几何部件. 由图 1 可见, 轮廓的复杂度不同, 组成轮廓的几何形状个数也不同. 如瓶子只需要两个形状就能近似描述; 而结构复杂的自行车则需要七个形状才能近似描述. 同时, 目标的每个几何形状的面积各不相同, 图 1(c)中的数字表示几何形状的面积大小顺序. 几何形状的面积越大, 与目标轮廓越相似, 更能表明物体的存在。

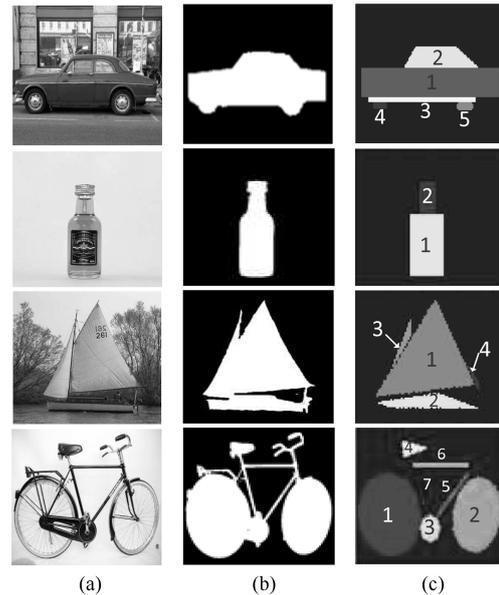


图 1 待检测目标样本图像、二值轮廓及其几何部件 ((a) 待检测目标样本图像; (b) 二值目标轮廓; (c) 几何部件)

Fig. 1 The sample image, binary object profile and the geometric parts of the detected rigid-object

((a) The sample image of the detected object; (b) Binary object profile; (c) Geometric parts)

基于这种思想, 本文提出了基于轮廓的几何稀疏表示的刚性目标检测模型. 该模型以 PBM 为框架, 结合稀疏表示原理, 将目标轮廓自适应地表示成若干个几何图形的组合. 每个几何图形定义为一个部件, 称为几何稀疏部件. 由于不同部件具有不同的重要性, 对目标轮廓的描述起不同的作用. 为反映部件间的差异, 本文进一步构造了有序链式结构。

1.1 轮廓几何稀疏表示

定义区域大小为 R 的二值几何原子 $g_{t,\theta,s,p}^R(x,y)$ 如下:

$$g_{t,\theta,s,p}^R(x,y) = \begin{cases} 1, & (x,y) \in \Gamma_{t,\theta,s,p}^R \\ 0, & \text{其他} \end{cases} \quad (1)$$

其中, $\Gamma_{t,\theta,s,p}^R$ 表示在区域 R 内的几何图形, 如图 2 所示. t 为类型参数, 用于描述几何图形形状, θ 为方向参数, s 为尺度参数, p 为位置参数. 图 2 给出了椭圆和梯形两种几何原子在不同参数情况下的空间表现。

图中四个原子大小均为 R , 其中, 两个椭圆原子具有相同的形状, 但方向、尺度和位置均不相同; 两个梯形原子具有相同的位置和尺度, 但形状和方向不同. 也就是说, 通过改变 t , θ , s 或 p 任意一个参数可产生不同的几何原子, 它们有不同的空间表现和描述作用. 对给定的目标轮廓 A , 建立几何原子库

如下:

$$D_A = \{g_{t,\theta,s,p}^{R_A}\} \quad (2)$$

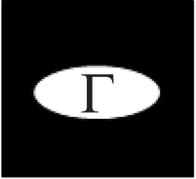
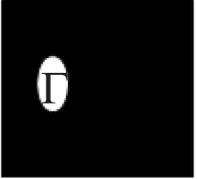
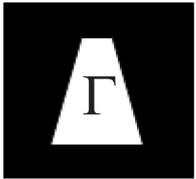
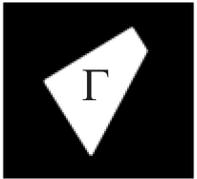
		$R = M \times N$	
椭圆			
	$t: a/b \quad \theta: 0$ $s: S \quad p: (M/2, N/2)$	$t: a/b \quad \theta: \pi/2$ $s: S/2 \quad p: (M/2, N/4)$	
	a 为椭圆长轴; b 为椭圆短轴; S 为尺度值		
梯形			
	$t: [a/h, b/h, 0] \quad \theta: 0$ $s: S \quad p: (M/2, N/2)$	$t: [a/h, b/h, (b-a)/2h] \quad \theta: \pi/4$ $s: S \quad p: (M/2, N/2)$	
	a 为上底; b 为下底; h 为高; S 为尺度值		

图2 几何原子

Fig.2 Geometric atoms

其中, R_A 为目标轮廓 A 的大小. 由式 (2) 的定义可知, 几何原子库由 t, θ, s 和 p 这 4 个参数通过取遍所有可能的取值所产生的几何原子组成, 包含不同形状、方向、尺度和位置, 能完全覆盖区域 R_A , 因此, 是一个过完备的原子库. 为阐述方便, 下文将库中原子统一表示为 $g_i^{R_A}$, i 为原子的索引编号. 利用该几何原子库可将 A 表示如下:

$$A = \sum_i c_i \cdot g_i^{R_A} \quad (3)$$

其中, $c_i \in \{0, 1\}$ 为二值系数, 表示是否选用 $g_i^{R_A}$ 来表示目标轮廓. 由式 (3) 可知, 对给定的目标轮廓, 只有系数 c_i 为 1 的原子才对表示起作用. 因此, 令 $\Omega_A = \{i | c_i = 1\}$ 为原子库的索引子空间, 则式 (3) 可等价表示如下:

$$A = \sum_{i \in \Omega_A} g_i^{R_A} \quad (4)$$

虽然 Ω_A 中的原子数目远小于原子库中的原子数目, 但该表示仍存在冗余. 事实上, Ω_A 中很多“小”原子在目标轮廓的逼近表示时几乎不起作用,

可以忽略. 因此, 将式 (4) 近似表示如下:

$$A \approx \sum_{i \in \Omega'_A} g_i^{R_A} \quad (5)$$

式中, $\Omega'_A = \{i | i \in \Omega_A, m_i \geq \varepsilon\}$, 表示 Ω_A 中作用较“大”的原子索引子空间. 原子的作用“大小”由匹配度 m_i 表示, 定义如下:

$$m_i = \frac{\langle g_i^{R_A}, A \rangle}{\langle A, A \rangle} \quad (6)$$

式中, $\langle \cdot \rangle$ 为内积算子, $m \in [0, 1]$.

为得到最优的几何原子组合, 建立最优化模型如下:

$$\begin{aligned} \min & \text{CARD}(\Omega'_A) \\ \text{s.t.} & \left\| A - \sum_{i \in \Omega'_A} g_i^{R_A} \right\|_2 < \sigma \\ & \langle g_i^{R_A}, g_j^{R_A} \rangle = 0, i \neq j \end{aligned} \quad (7)$$

式中, A 为目标轮廓 (如图 1 (b)), $\text{CARD}(\cdot)$ 为基数算子, 表示集合内元素个数, $\|\cdot\|_2$ 为 2 范数, 表示能量.

式 (7) 所描述的物理意义是, 在所有匹配度大于 ε 的几何原子中, 寻找一种正交组合, 使得该组合表示的轮廓与目标轮廓的能量差小于 σ . 由于该问题与信号稀疏分解问题^[24] 相似, 对式 (7) 可采用经典的稀疏分解优化方法求解, 本文采用了基于量子遗传优化的匹配追踪算法^[25]. 式 (7) 求解的结果如图 1 (c) 所示. 本文将这些几何原子定义为 PBM 中的部件, 即几何稀疏部件.

1.2 有序链式结构

PBM 最突出的优点在于构建了空间结构来描述零散部件间的关系, 使模型更准确地表示目标. 不同的结构有不同的特性, Fergus 等^[20] 的星座模型蕴含最多的部件空间关系, 但空间信息存在冗余, 建模及训练耗时; Crandall 等^[21] 将其简化为 k -fans 结构, 在保留空间信息的同时, 简化了建模复杂度. 除此之外, Felzenszwalb 等的树模型^[22], Carneiro 等的稀疏柔性模型^[23] 都从不同方面进行了改进. 然而, 传统的 PBM 模型缺少对部件差异性的描述. 由式 (6) 可知, 不同的部件具有不同的匹配度. 部件的匹配度越大, 其对目标轮廓的表示作用也越大. 这种差异性对后续的检测有导向作用, 可减少不必要的检索, 提高检测速度. 基于此, 本文构造了一种有序的链式结构. 该结构将部件按匹配度降序排序并链接. 根部件定义为目标轮廓, 即 $r = A$, 并置于链条的最后. 有序链式结构如图 3 所示.

图 3 中, P_i 表示第 i 个部件; r 表示根部件; $v_{i,j}$ 表示部件 P_i 和部件 P_j 的空间相对位置, 由下式求得:

$$v_{i,j} = P_j\{p\} - P_i\{p\} \quad (8)$$

其中, $P_i\{p\}$ 表示部件 P_i 中的位置参数 p .

由图 3 可见, 链式结构中相邻部件之间是相互关联的, 即根据前一个部件来确定后一个部件的锚点位置, 因此可避免空间结构过于松散. 有序的特点为后续的分级检测提供了基础.

2 分级检测算法

传统的 PBM 在检测时需要计算每个位置上根部件和所有子部件的匹配度, 尽管杨扬等^[26]提出借助预分割信息减少检测位置能一定程度上提高检测效率, 但检测的效果依赖于预分割的好坏. 事实上, 一些关键部件的存在与否能很大程度上决定目标是否存在, 因此, 本文在几何稀疏部件和有序链式结构的基础上, 提出了一种分级检测算法. 该算法根据每个部件的匹配度和检测结果对待检测位置进行筛选, 通过逐级缩小检测范围来避免不必要的匹配计算,

提高检测速度. 本文将该算法应用到猫头检测中, 其检测流程如图 4 所示, 算法伪代码如图 5.

分级检测算法以有序链式结构为序, 依次对每个部件和根部件进行检测. 图 4 的猫头检测例子中, 部件模型 (图 4(a)) 包含三个子部件和一个根部件 (图 4(b)), 共需要进行四级的检测.

对每一级的检测, 主要完成三个任务: 根据待检测点集扫描测试图像; 生成部件显著图和产生下一个部件的待检测点集 (如图 4 虚线中的步骤 1), 2) 和 3)). 生成部件显著图时, 首先根据式 (9) 对二值部件进行高斯模糊. 这是由于部件经过模糊后有效边界加粗. 用模糊后的部件检测不仅考虑到边界像素, 而且其相邻像素信息也用于部件检测, 从而使检测能适应少量的形变, 有利于防止部件漏检. 为了抵抗目标纹理、光照和颜色等因素的影响, 采用式 (10) 将部件和测试图像转换到 HOG 空间, 并定义部件与图像子块 HOG 特征的内积作为显著值, 如式 (11) 所示. 由式 (11) 可计算出测试图像中所有位置的部件显著值, 从而生成如图 4(e) 的部件显著图.

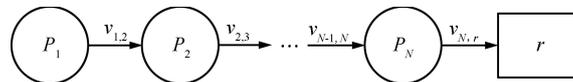


图 3 有序链式结构示意图

Fig. 3 Ordered chain-like structure

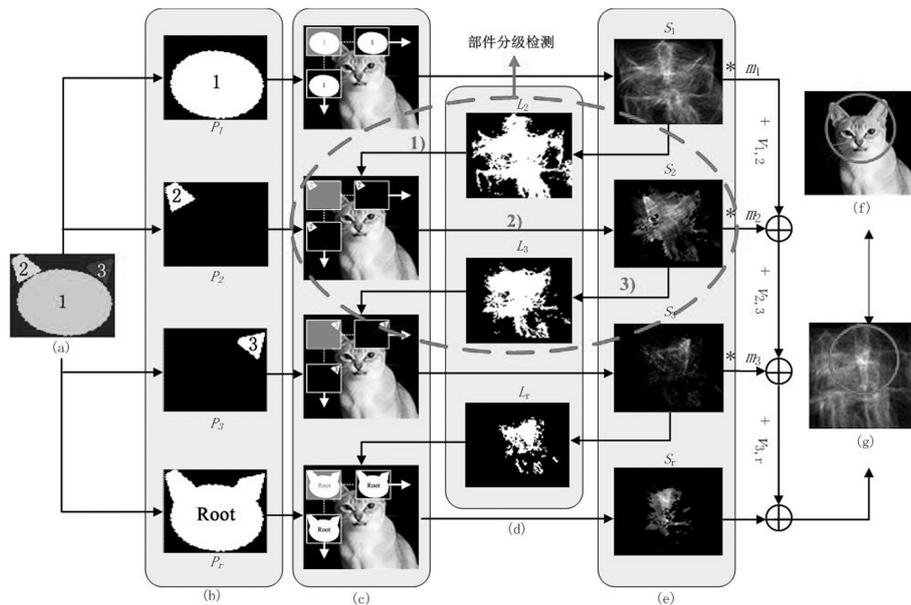


图 4 有猫头目标的分级检测过程 ((a) 部件模型; (b) 部件; (c) 部件扫描测试图像中的待检测位置; (d) 待检测点集; (e) 部件显著图; (f) 测试图像; (g) 目标显著图)

Fig. 4 Process of hierarchical detection of the cat's head ((a) Part-based model of the cat's head; (b) Model parts; (c) Each part scans the detected positions of the test image; (d) The detected position set of test image; (e) Salient maps of parts; (f) Test image; (g) Salient map of object)

输入	
测试图像: f 目标部件模型: 部件: $(P_1, P_2, \dots, P_N, P_r)$ 部件空间关系: $v_{1,2}, v_{2,3}, \dots, v_{N-1,N}, v_{N,r}$ 待检测点集: L_1 为测试图像中所有位置	
分级检测	
for $P_i, i := 1$ to $N + 1$ do	
对每一个部件执行如图 4 虚线框的部件分级检测，	1) 根据上一级得到的待检测点集 L_i ，每个部件扫描测试图像中的待检测位置
	2) 生成部件显著图: ● 部件 P_i 进行高斯模糊: $P_i^g = g(\mu, \sigma) * P_i \quad (9)$ 式中, g 为高斯函数; μ 和 σ 分别为高斯函数的均值和方差; $*$ 为卷积运算. ● 分别提取 P_i^g 及子图像 $f^{(x,y)}$ 的 HOG 特征图 ^[27] : $H_{P_i^g} = H(P_i^g), \quad H_{f_l} = H(f^{(x,y)}) \quad (10)$ 式中, $H(\cdot)$ 为 HOG 特征提取算子. ● 计算显著值: $S^{P_i}(x, y) = \langle H_{P_i^g}, H_{f_l} \rangle \quad (11)$
	3) 更新待检测点位置: ● 计算显著平均值: $\overline{S^{P_i}}$ ● 从待检测点集中去除不显著点: $L_{i+1} = \frac{L_i}{(x, y)}, \quad \text{if } S^{P_i}(x, y) < \overline{S^{P_i}} \quad (12)$
	end
显著图融合	
初始化显著图 S 为与测试图像相同大小的零矩阵; for $i := 1$ to $N + 1$ do	
融合得到目标显著图, 图 4 (e), (g)	● 校正显著图位置: $S^{P_i} = \text{shift}(S^{P_i}, v_{i,i+1}) \quad (13)$
	● 计算部件的匹配度: $m_i = \frac{\langle P_i, A \rangle}{\langle A, A \rangle} \quad (14)$
	● 显著图加权融合: $S = S + m_i S^{P_i} \quad (15)$
end	

图 5 分级检测算法伪代码

Fig. 5 Pseudo-code of hierarchical detection

为减少不必要的检测, 算法采用平均值阈值化的方法, 将显著值小于平均显著值的位置去除, 如图 5 中式 (12) 所示, 从而生成下一个部件的待检测点集. 比较图 4 (d) 各图可见, 待检测范围逐级缩小.

分级检测的最终结果是生成目标的显著图, 它由各部件 (包括根部件) 的加权融合得到, 如式 (15) 所示. 其中, 权值为各部件的匹配度, 由式 (14) 计算得到. 融合时, 由于部件间存在空间差异, 需要先进

行如式 (13) 的位置调整, 从而使显著点位于目标中心. 图 4(g) 为猫头检测的目标显著图.

对目标的定位, 本文采用非最大值抑制 (Non-maximum suppression, NMS)^[28] 算法. 该算法提取显著图的极值点位置, 并以其为中心, 获取一个与根部件大小相同的矩形区域. 若该区域覆盖真实目标超过 80%, 则认为检测成功. 对于多目标情况, 可采用 MAP (Maximum-a-posteriori)^[29] 算法在显著图上迭代检索极值点, 直到极值小于某个给定阈值, 检测结束.

3 实验仿真

3.1 评价指标

本文采用 PASCAL VOC2007^[30] 数据库作为测试图像库. 由于本文主要研究目标模型建立和分级检测算法, 为简化实验, 本文只考虑单目标检测的情况.

为了定性分析算法的有效性, 本文采用平均准确率 (Average precision, AP)^[31] 作为检测性能评价指标, 其定义如下:

$$AP = \int_R P(R) dR \quad (16)$$

式中, R 为召回率 (Recall), 表征检测系统的查全性; P 为准确率 (Precision), 表征检测系统的准确性.

平均准确率 AP 表征的是以 R 为横轴, P 为纵轴所构成的 PR 曲线所围成的面积. $AP \in [0, 1]$, AP 值越大, 说明系统性能越好.

3.2 模型建立

对每类待检测目标, 首先提取 (自动或手工) 目标样本图像的二值轮廓, 然后构建几何原子库, 并采用基于遗传优化的匹配追踪算法对目标轮廓进行几何稀疏表示, 生成每类目标的链式模型, 其中轮廓几何稀疏表示所用的参数在表 1 中给出.

表 1 轮廓几何稀疏表示参数表

Table 1 Table of parameters using in geometric sparse decomposition

	参数名称	取值
几何原子库	原子类型	矩形, 椭圆, 三角形, 梯形
	方向个数	16
量子遗传优化算法	种群数	50
	演化代数	100
	旋转角	0.01π

由第 2.1 节的式 (7) 可知, 阈值 σ 决定目标轮廓的部件的个数. σ 越小, 部件个数越多, 建模时所

耗费的时候也越多, 平均准确率越高. 由此可见, AP 值与部件个数、建模时间是相互矛盾的. 为了使这种矛盾最小, 建立指标如下:

$$Q = \frac{Ap}{Pn} + \frac{Ap}{Bt} \quad (17)$$

其中, Ap 表示归一化的平均准确率, Pn 表示归一化的部件个数, Bt 表示归一化的建模时间.

由式 (17) 可见, 指标 Q 与 Ap 成正比, 与 Pn 和 Bt 成反比. 我们希望在部件个数和建模时间较少的同时, 获得较高的准确率, 即寻找最优的 σ 使 Q 达到最大值. 本文通过数值统计的方法来实现对 σ 的最优选取: 分别计算帆船、自行车、瓶子、飞机、公共汽车、小汽车 6 种刚性目标及猫、马两种非刚性目标在 σ 取轮廓面积的 5% ~ 40% 内不同值时的 Q , 得到 $Q_i(\sigma)$, $i = 1, \dots, 8$. 构造综合指标曲线 $Q(\sigma)$ 如下:

$$Q(\sigma) = \sum_{i=1}^8 Q_i(\sigma) \quad (18)$$

通过 8 种目标的综合评价, $Q(\sigma)$ 在一定程度上反映了 σ 对 Q 的影响情况, $Q(\sigma)$ 曲线如图 6 所示.

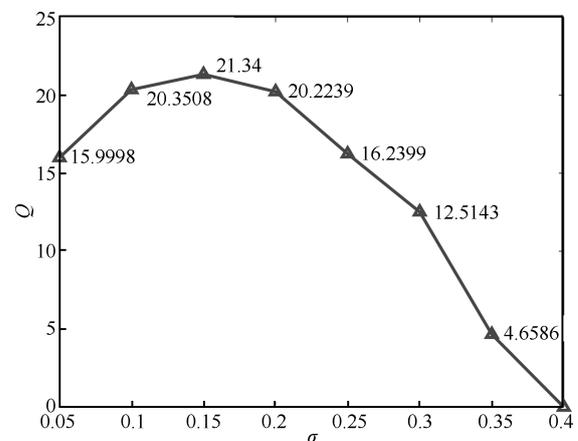


图 6 $Q(\sigma)$ 曲线

Fig. 6 Q curve

由图 6 可见, 当 σ 取轮廓面积的 15% 时, Q 值达到最大, 因此, 本文实验中的 σ 取目标轮廓面积的 15%.

3.3 仿真结果及分析

本文的实验仿真分为两部分, 首先验证部件的漏检, 然后给出本文与现有文献的对比检测结果.

3.3.1 部件漏检验证实验

以自行车目标为例, 通过研究各部件在对焦模糊、不均匀光照、复杂背景、不同纹理、噪声干扰及轻微遮挡等 6 种复杂情况下的检测情况, 验证部件的漏检. 图 7 给出了含自行车目标的 6 幅测试图

像, 其中图 7(a) 对焦不准导致自行车边缘模糊; 图 7(b) 由于光照角度导致自行车亮度不均匀, 且产生阴影干扰; 图 7(c) 的目标处于复杂背景中; 图 7(d) 为两辆纹理不同的自行车; 图 7(e) 为 PSNR 值仅有 9.08 的椒盐噪声干扰后的图像; 图 7(f) 为自行车被轻微遮挡的图像。

为研究以上 6 种复杂情况下各部件的漏检, 首先根据数据库标定的目标位置及目标的几何部件模型, 推算出各部件的期望位置; 然后通过图 5 求出各分级中部件在期望位置上的显著值 S^{P_i} 与平均显著值 $\overline{S^{P_i}}$; 最后计算各部件的显著值差值如式 (19):

$$\Delta S = S^{P_i} - \overline{S^{P_i}} \quad (19)$$

得到如图 8 所示的显著值差值图。

由图 8 可见, 在 6 种复杂情况下, 所有部件的显著值差值均大于 0, 说明部件没有被漏检; 另外, 从图 8 还可以发现, 显著值差值随部件编号增加呈减少趋势, 表明越靠前的部件越不容易被漏检。

上述实验结果表明, 本文的建模方法能减少部件检测时受外界的影响, 降低部件被漏检的概率; 由于靠前的部件不容易被漏检, 且具有较大的匹配度和权重, 即使在部分小部件被漏检的情况下, 经加权融合后, 能有效降低因部件漏检而导致整体目标漏检的可能性。

3.3.2 在 PASCAL VOC 2007 上的仿真实验

将本文方法在 PASCAL VOC 2007 数据库中的检测结果与文献 [15] 中的标准方法 (Base)、Bounding box 方法 (BB)、Context 方法和

文献 [16] 的 Latern hierartic sturcture (LHS) 方法进行对比, 说明本文方法对刚性目标的检测率和检测效率。图 9 中, 第 1 行和第 3 行分别为小轿车、瓶子、帆船和自行车的部分测试图像, 目标具有不同的纹理, 所处的背景差异也很大, 部分图像存在少量的视角偏移, 如第 1 行第 2 张和第 3 张测试图像; 第 2 行和第 4 行分别为与其相对应的显著图。高亮度 (图 9 中的矩形框区域) 标志着目标所在位置。从各检测结果可以看出, 测试图像中的目标无论具有怎样的纹理, 所处背景差异有多大, 只要与目标模型所描述的轮廓相似, 均能得到正确的检测结果。

表 2~表 4 分别给出了本文方法中各类别建模时所需部件数目、平均检测时间及 AP 值, 并与文献 [15] 的标准方法 (Base)、Bounding box 方法 (BB)、Context 方法和文献 [16] 的 Latern hierartic sturcture (LHS) 进行对比。

由表 2 可见, 在部件数目方面, 文献 [15] 各方法和文献 [16] 的部件数目是固定的, 而本文的部件数目根据目标轮廓的复杂程度自适应地变化。对于大部分具有刚性轮廓的人造目标, 如汽车、飞机等, 本文选取的部件数目均较少, 而对于结构复杂 (如自行车) 或者姿态多变 (如马、狗、人) 等目标, 本文的部件数目略多于文献 [15]。

表 3 反映的是各方法在平均检测时间方面的差异。由于本文需要的部件较少, 并采用了分级的检测方法, 检测时耗比文献 [15] 各方法和文献 [16] 都有不同程度的减少, 比文献 [15] 最多减少 64.88%, 比文献 [16] 最多减少 90.69%。由表 3 可见, 除了部件较多的自行车、马、狗和人这四种目标外, 检测时间均为最少。

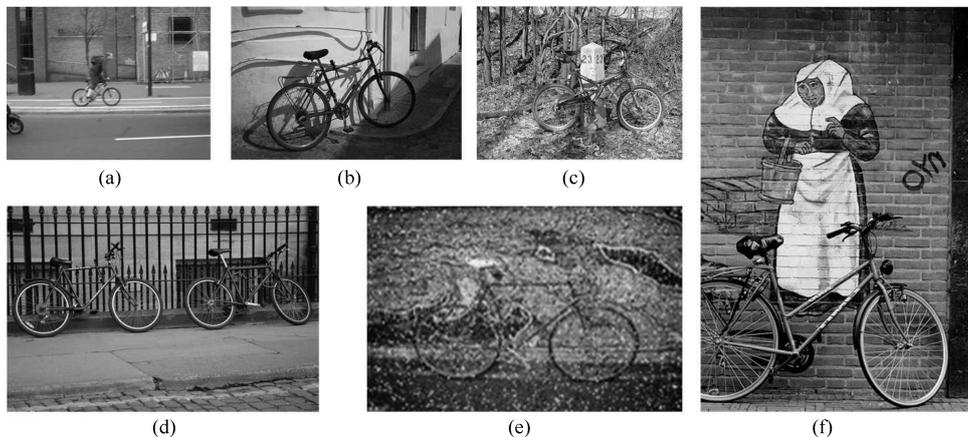


图 7 6 种复杂情况的测试图 ((a) 对焦模糊; (b) 不均匀光照; (c) 复杂背景; (d) 不同纹理; (e) PSNR 为 9.08 的噪声干扰; (f) 轻微遮挡)

Fig. 7 Test images with six different degeneration cases ((a) Blurred; (b) Illumination; (c) Complicated background; (d) Texture; (e) Noised; (f) Slight occlusion)

在 AP 指标方面, 本文方法对轮廓明显且具有刚性特征的目标, 如车类、帆船、飞机、瓶子等有较高的检测率; 而对姿态变化较多的动物类目标, 如鸟、猫、马等, 以及轮廓单一的目标, 如电视机, 其 AP 值均低于文献 [15] 和文献 [16]. 尤其对种类形状及视角变化较多的目标, 如椅子、桌子和沙发等, 其轮廓难以用几何部件模型描述, 因此本文方法对这类目标的检测性能较差.

总体来说, 本文方法能灵活自适应地选取部件, 有效减少部件数目和检测时间, 对具有刚性轮廓的目标有较好的检测结果.

由图 9 小轿车和瓶子的检测结果可见, 本文算法能适应因视角偏移而产生的少量形变, 但对姿态

变化或遮挡等问题暂时无法解决, 这将成为本文的未来工作继续展开.

4 结论

利用刚性目标具有固定轮廓且轮廓具有明显几何外观的特点, 提出了一种适用于刚性目标的分级检测算法. 该算法具有以下创新点:

- 1) 构建了二值几何原子库, 结合稀疏表示原理, 将刚性目标自适应地稀疏表示为几何部件的组合.
- 2) 根据部件与目标轮廓的匹配度, 构建了有序链式结构来描述部件间的空间结构.
- 3) 利用链式结构的有序特性, 逐级缩小待检测范围以提高检测速度.

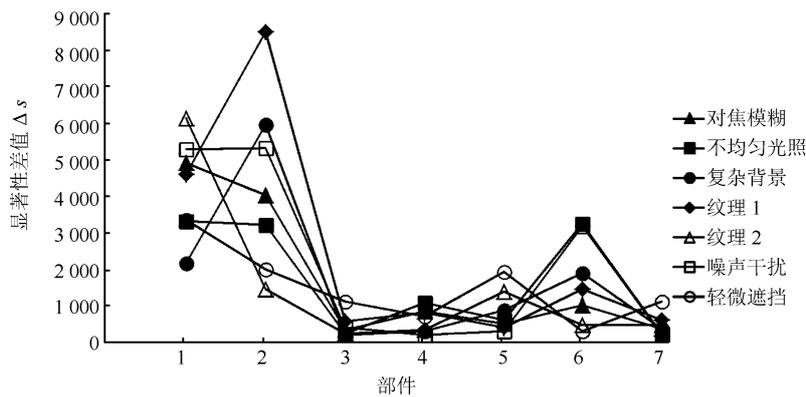


图 8 各部件在 6 种复杂情况下的显著值差值图

Fig. 8 Salient difference map of parts detection in six cases

表 2 本文方法与其他方法的部件数目比较

Table 2 Number of parts of our method and other methods

	Boat	Bike	Bus	Car	Aero	Bottle	Train	TV	Cat	House	Bird	Chair	Cow	Table	Dog	Pers	MBik	Sheep	Sofa
Base	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
BB	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
Context	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
LHS	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45
Ours	4	7	3	5	5	2	3	1	3	8	6	6	6	5	7	7	5	6	5

表 3 本文方法与其他方法的检测时间比较 (s)

Table 3 Detection cost of our method and other methods (s)

	Boat	Bike	Bus	Car	Aero	Bottle	Train	TV	Cat	House	Bird	Chair	Cow	Table	Dog	Pers	MBik	Sheep	Sofa
Base	2.11	2.34	1.98	2.01	2.21	2.05	2.04	1.99	2.18	2.13	2.07	2.11	2.23	1.89	2.02	1.97	2.13	2.08	1.94
BB	2.15	2.40	2.02	2.07	2.25	2.06	2.07	1.99	2.21	2.22	2.09	2.12	2.25	1.92	2.05	2.01	2.17	2.13	2.01
Context	20.8	21.7	20.4	20.4	21.6	20.1	20.8	21.1	21.3	20.5	21.4	20.5	21.4	20.6	21.6	20.9	21.2	21.3	20.7
LHS	8.21	8.04	7.83	7.79	7.93	7.73	7.94	7.53	8.12	8.17	8.18	7.78	8.02	7.82	8.13	8.09	8.11	8.08	7.98
Ours	1.47	2.49	1.01	1.73	1.83	0.72	0.82	0.33	1.07	2.72	1.68	1.56	1.61	1.52	2.11	2.24	1.47	1.87	1.50



图 9 VOC2007 测试库上的检测结果图

Fig. 9 Detection results on VOC 2007

表 4 本文方法与其他方法在 PASCAL VOC 2007 上的平均检测率比较

Table 4 Average precision of our method and other methods in PASCAL VOC 2007

	Boat	Bike	Bus	Car	Aero	Bottle	Train	TV	Cat	House	Bird	Chair	Cow	Table	Dog	Pers	MBik	Sheep	Sofa
Base	0.134	0.546	0.384	0.464	0.290	0.262	0.340	0.390	0.161	0.436	0.006	0.163	0.165	0.245	0.050	0.350	0.378	0.173	0.216
BB	0.143	0.552	0.394	0.501	0.289	0.266	0.340	0.381	0.165	0.451	0.006	0.165	0.166	0.245	0.050	0.362	0.383	0.174	0.228
Context	0.153	0.561	0.389	0.504	0.287	0.274	0.387	0.377	0.206	0.480	0.025	0.179	0.185	0.259	0.088	0.368	0.412	0.162	0.244
LHS	0.143	0.558	0.440	0.513	0.294	0.286	0.368	0.393	0.213	0.504	0.094	0.200	0.193	0.252	0.125	0.366	0.384	0.197	0.251
Ours	0.154	0.561	0.441	0.541	0.294	0.288	0.403	0.381	0.173	0.497	0.004	0.003	0.062	0.004	0.027	0.219	0.289	0.056	0.007

4) 以部件匹配度为权值加权融合各部件显著图来生成目标显著图, 突出了部件的差异性, 提高准确率.

在 PASCAL VOC 2007 图像库的检测测试中表明, 该算法对具有明显轮廓的刚性目标有较好的检测结果. 未来的研究方向主要集中在尺度变化、方向变化、姿态形变及遮挡等问题, 使模型更具鲁棒性.

References

- Li Wan-Yi, Wang Peng, Qiao Hong. A survey of visual attention based methods for object tracking. *Acta Automatica Sinica*, 2014, **40**(4): 561–576
(黎万义, 王鹏, 乔红. 引入视觉注意机制的目标跟踪方法综述. 自动化学报, 2014, **40**(4): 561–576)
- Jiang Xiao-Lian, Li Cui-Hua, Li Xiong-Zong. Saliency based tracking method for abrupt motions via two-stage sampling. *Acta Automatica Sinica*, 2014, **40**(6): 1098–1107
(江晓莲, 李翠华, 李雄宗. 基于视觉显著性的两阶段采样突变目标跟踪算法. 自动化学报, 2014, **40**(6): 1098–1107)
- Wang X, Lv Q, Wang B, Zhang L M. Airport detection in remote sensing images: a method based on saliency map. *Cognitive Neurodynamics*, 2013, **7**(2): 143–154
- Han Min, Zheng Dan-Chen. Shape recognition based on fuzzy shape context. *Acta Automatica Sinica*, 2012, **38**(1): 68–75
(韩敏, 郑丹晨. 基于模糊形状上下文特征的形状识别算法. 自动化学报, 2012, **38**(1): 68–75)
- Lin Y D, He H J, Yin Z K, Chen F. Rotation-invariant object detection in remote sensing images based on radial-gradient angle. *IEEE Geoscience and Remote Sensing Letters*, 2015, **12**(4): 746–750

- 6 Sun H, Sun X, Wang H Q, Li Y, Li X J. Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model. *IEEE Geoscience and Remote Sensing Letter*, 2012, **9**(1): 109–113
- 7 Liu L, Shi Z W. Airplane detection based on rotation invariant and sparse coding in remote sensing images. *Optik*, 2014, **125**(18): 5327–5333
- 8 Lei Z, Fang T, Huo H, Li D R. Rotation-invariant object detection of remotely sensed images based on texture forest and Hough voting. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, **50**(4): 1206–1217
- 9 Xu J, Sun X, Zhang D B, Fu K. Automatic detection of inshore ships in high-resolution remote sensing images using robust invariant generalized Hough transform. *IEEE Geoscience and Remote Sensing Letters*, 2014, **11**(12): 2070–2074
- 10 Li F F, Fergus R, Torralba A. Recognizing and learning object categories. [Online], available: <http://people.csail.mit.edu/torralba/shortCourseRLOC/>, March 15, 2013
- 11 Csurka G, Dance C R, Fan L X, Willamowski J, Bray C. Visual categorization with bags of keypoints. In: Proceedings of the 2004 European Conference on Computer Vision. Prague, Czech Republic: Springer, 2004. 1–22
- 12 Felzenszwalb P F, Huttenlocher D P. Pictorial structures for object recognition. *International Journal of Computer Vision*, 2005, **61**(1): 55–79
- 13 Song X, Luo P, Lin L, Jia Y D. A discriminative model for object representation and detection via sparse features. In: Proceedings of the 20th International Conference on Pattern Recognition. Istanbul, Turkey: IEEE, 2010. 3077–3080
- 14 Sun X, Wang H Q, Fu K. Automatic detection of geospatial objects using taxonomic semantics. *IEEE Geoscience and Remote Sensing Letters*, 2010, **7**(1): 23–27
- 15 Felzenszwalb P F, Girshick R B, McAllester D. Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(9): 1627–1645
- 16 Zhu L, Chen Y H, Yuille A, Freeman W. Latent hierarchical structural learning for object detection. In: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, CA, USA: IEEE, 2010. 1062–1069
- 17 Lu C E, Adluru N, Ling H B, Zhu G X, Latecki L J. Contour based object detection using part bundles. *Computer Vision and Image Understanding*, 2010, **114**(7): 827–834
- 18 Zhang H G, Wang J X, Bai X, Zhou J, Cheng J, Zhao H J. Object detection via foreground contour feature selection and part-based shape model. In: Proceedings of the 21st International Conference on Pattern Recognition. Tsukuba, Japan: IEEE, 2012. 2524–2527
- 19 Xi Hui-Ting. Study on Rigid Moving Target Tracking Algorithm [Master dissertation], East China Normal University, China, 2008.
(奚慧婷. 刚性运动目标的跟踪算法研究 [硕士学位论文], 华东师范大学, 中国, 2008.)
- 20 Fergus R, Perona P, Zisserman A. Object class recognition by unsupervised scale-invariant learning. In: Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Madison, Wisconsin, USA: IEEE, 2003. II-264– II-271
- 21 Crandall D, Felzenszwalb P F, Huttenlocher D P. Spatial priors for part-based recognition using statistical models. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE, 2005. 10–17
- 22 Felzenszwalb P F, Huttenlocher D P. Efficient matching of pictorial structures. In: Proceedings of the 2000 IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head Island, South Carolina, USA: IEEE, 2000. 66–73
- 23 Carneiro G, Lowe D. Sparse flexible models of local features. In: Proceedings of the 9th European Conference on Computer Vision. Graz, Austria: Springer, 2006. 29–43
- 24 Wang Jian-Ying, Yin Zhong-Ke, Zhang Chun-Mei. *The Sparse Decomposition and Application for Signals and Images*. Chengdu: Southwest Jiaotong University Press, 2006. 49–57
(王建英, 尹忠科, 张春梅. 信号与图像的稀疏分解及初步应用. 成都: 西南交通大学出版社, 2006. 49–57)
- 25 Li Heng-Jian, Yin Zhong-Ke, Wang Jian-Ying. Image sparse decomposition based on quantum genetic algorithm. *Journal of Southwest Jiaotong University*, 2007, **42**(1): 19–23
(李恒建, 尹忠科, 王建英. 基于量子遗传优化算法的图像稀疏分解. 西南交通大学学报, 2007, **42**(1): 19–23)
- 26 Yang Yang, Li Shan-Ping. Fast object detection with deformable part models and segment locations' hint. *Acta Automatica Sinica*, 2012, **38**(4): 540–548
(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. 自动化学报, 2012, **38**(4): 540–548)
- 27 Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE, 2005. 886–893
- 28 Neubeck A, van Gool L. Efficient non-maximum suppression. In: Proceedings of the 18th International Conference on Pattern Recognition. Hong Kong, China: IEEE, 2006. 850–855
- 29 Barinova O, Lempitsky V, Kohli P. On detection of multiple object instances using Hough transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(9): 1773–1784
- 30 Everingham M, van Gool L, Williams C K I, Winn J, Zisserman A. The PASCAL Visual Object Classes Challenge 2007 (VOC 2007) Results [Online], available: <http://www.pascal-network.org/challenges/VOC/voc2007/>, April 20, 2013
- 31 Lu Wen-Hao, Li Ya-Li, Wang Sheng-Jin, Ding Xiao-Qing. Improvements of 3D object detection with part-based models. *Acta Automatica Sinica*, 2012, **38**(4): 497–506
(鹿文浩, 李亚利, 王生进, 丁晓青. 基于部件的三维目标检测算法新进展. 自动化学报, 2012, **38**(4): 497–506)



林煜东 西南交通大学信息科学与技术学院博士研究生. 2006 年获得华南农业大学信息学院学士学位. 2009 年获得华南农业大学信息学院硕士学位. 主要研究方向为图像表示与图像目标检测.

E-mail: willianlam@126.com

(**LIN Yu-Dong** Ph.D. candidate at the School of Information Science and

Technology, Southwest Jiaotong University. He received his bachelor degree from South China University of Agriculture in 2006 and master degree from the same school in 2009. His research interest covers image representation and target detection in images.)



和红杰 博士, 西南交通大学信息科学与技术学院教授. 主要研究方向为图像处理, 图像取证.

E-mail: hjhe@home.swjtu.edu.cn

(**HE Hong-Jie** Ph.D. and professor at the School of Information Science and Technology, Southwest Jiaotong University. Her research interest

covers image processing and image forensics.)



陈帆 博士, 西南交通大学信息科学与技术学院副教授. 主要研究方向为多媒体数据安全, 图像处理, 计算机应用技术. 本文通信作者.

E-mail: fchen@home.swjtu.edu.cn

(**CHEN Fan** Ph.D. and associate professor at the School of Information Science and Technology, Southwest Jiao-

tong University. His research interest covers multimedia data security, image processing, and computer application technology. Corresponding author of this paper.)



尹忠科 博士, 北京遥感信息研究所教授. 主要研究方向为信号与信息处理, 图像传输与处理.

E-mail: yinzhongke@163.com

(**YIN Zhong-Ke** Ph.D. and professor at the Beijing Institute of Remote Sensing Information. His research interest covers signal and information

processing, and image transmission and processing.)