

基于级联 Adaboost 的目标检测融合算法

崔潇潇¹ 姚安邦¹ 王贵锦¹ 林行刚¹

摘要 单一特征的模型对于颜色纹理变化较大的目标的检测往往存在检测率不高或检测速度慢的缺点. 本文提出了一种基于级联 Adaboost 的“级联-加和”融合算法. 融合模型由两个独立训练得到的级联 Adaboost 分类器组成, 分别利用边界片段特征和矩形类 Haar 小波特征描述整个目标以及目标的一个稳定部件. 级联-加和的融合决策以样本在两个分类器中被拒绝或通过的级数信息为依据. 在多个数据库上的实验证明这种融合检测算法不仅综合了 Haar 小波特征检测速度快和边界片段特征鲁棒性好的优点, 而且与单一特征的分类器相比, 检测性能也有所提高.

关键词 目标检测, 融合模型, 边界片段特征, Haar 特征, 级联 Adaboost
中图分类号 TP391

Object Detection by Combined Model Based on Cascaded Adaboost

CUI Xiao-Xiao¹ YAO An-Bang¹ WANG Gui-Jin¹ LIN Xing-Gang¹

Abstract Single feature-based model always meets the difficulties of poor detection performance and slow detection speed for object with large variances in color, texture, and shape. A novel cascaded and additive model based on cascaded Adaboost classifier is proposed in this paper. This combined model consists of two cascaded Adaboost classifiers which are independently trained with edge-fragment feature and Haar feature to describe the whole object and one of its stable components, respectively. The final classification decision of the combined model is made according to the stage indexes by which a sample is rejected or accepted in the two cascaded classifiers. Experiments on several test databases show that the combined model can take advantages of the speed merit of Haar feature and the robustness of edge-fragment feature. Compared with single feature-based model, the detection performance of the combined model is greatly improved.

Key words Object detection, combined model, edge-fragment feature, Haar feature, cascaded Adaboost

特征选择与机器学习算法是目标检测中的两个关键问题. Viola^[1] 等提出的矩形类 Haar 小波特征以及级联的 Adaboost 学习算法已经在人脸检测上取得了巨大的成功. 矩形类 Haar 小波可以很好地描述块状分布的纹理, 并且具有计算简便的优点; 而级联的 Adaboost 算法可以在待检测图像中快速排除大量的负面样本. 但是, 矩形类 Haar 小波对于颜色、纹理以及形状变化较大的目标检测效果并不理想. 文献 [2] 指出对于这类变化较大的目标 (如马、狗等), 利用基于边缘的特征可以取得较好的检测效果. 文献 [2] 中利用离散抽取的边缘点所在邻域的对数极坐标直方图来描述目标, 文献 [3] 中也采用了类似的描述方法, 但是这种方法对边缘提取的要求较高. Shotton^[4] 以及 Opelt^[5] 等提出了一种基于边界片段的描述方法, 利用 Chamfer 匹配提取边界片段与待检测样本之间的相似度作为样本的特征, 再利

用 Adaboost 选取分辨能力强的边界片段组成强分类器. 由于这种算法只关注目标的边界信息, 因此对颜色纹理的变化有很强的鲁棒性, 采用边界片段而不是目标的整体轮廓作为特征, 也在一定程度上提高了检测算法在目标发生形变以及边界提取结果不理想等情况下的鲁棒性. 但是这种基于边界片段的特征也存在受复杂背景干扰较大、计算速度较慢的缺点.

由于单一种类的特征对目标的分辨力有限, 如何利用多种特征进行优势互补, 提高检测性能已经成为研究热点. 目前的融合检测算法主要分为两类: 一类是特征级融合, 另一类是分类器级融合. 前者将多种类型的特征混合在一起, 再由机器学习算法从中挑选出分辨力强的特征组成分类器. 文献 [3] 用主成分分析 (Principal component analysis, PCA) 特征和尺度不变特征变换 (Scale-invariant feature transform, SIFT) 特征分别描述目标的轮廓以及纹理, 其检测算法分为两步: 1) 从 PCA-SIFT 混合特征中选择合适的特征; 2) 训练出这些特征之间的空间关系. Opelt 等在文献 [6] 中提出了两种融合方法, 选择特征法 (Selected-features, SF) 以及联合模型法 (Combined-model, CM). 前者属于特征级融合, 利用 Boost 算法选择基于纹理的特征和基于轮廓的特征; 后者则属于分类器级的融合算法, 独立训练两

收稿日期 2008-01-21 收修稿日期 2008-07-16
Received January 21, 2008; in revised form July 16, 2008
国家自然科学基金 (60472028), 教育部博士点基金 (20040003015) 资助

Supported by National Natural Science Foundation of China (60472028) and Specialized Research Fund for the Doctoral Program of Ministry of Education of China (20040003015)

1. 清华大学电子工程系 北京 100084
1. Department of Electronic Engineering, Tsinghua University, Beijing 100084
DOI: 10.3724/SP.J.1004.2009.00417

个分别基于纹理特征和轮廓特征的分类器, 最终的分类决策由两个分类器的决策加权得到. 文献 [6] 的实验结果表明分类器级的融合算法要优于特征级的融合算法, 并指出这是因为特征级融合算法中的混合特征会在某些数据集上发生过学习现象, 而分类器级的融合算法由于对各分类器的输出进行了加权投影, 因此正确的检测结果受个别错误检测结果的影响较小^[6]. 文献 [7–10] 提出的分类器级融合模型是由多个基于部件的分类器组成的. 文献 [11–12] 提出将产生式模型和判别式模型相结合的方法.

从提高以边界片段作为特征的检测器的检测速度来看, 特征级的融合策略由于在每级分类器中都混有边界片段特征, 因此仍需要对待检测图像的大部分区域计算边界片段特征, 不能有效地提高检测速度. 但是文献中的分类器级的融合算法中也存在一些问题: 首先, 如果各分类器的决策都是在图像空间上独立做出再进行融合, 即并行融合策略, 那么各分类器都需要对图像上的所有检测窗口进行检测, 算法的时间复杂度会成倍增长; 其次, 虽然级联的融合形式具有较低的时间复杂度, 但是级联分类器的检测率不会超过各检测器所给出的最低检测率, 对各检测器的性能要求较高.

本文提出一种适用于检测颜色、纹理、形状变化较大的目标的融合检测算法, 它是基于级联 AdaBoost 算法的分类器级融合检测算法. 该“级联-加和”融合模型是通过一个多支路的融合框架将两个分类器连接在一起. 在每条检测支路上对两个分类器进行级联, 最终的融合检测结果是各检测支路结果的并集. 分类器分别由矩形类 Haar 小波特征和边界片段特征训练得到. 以样本在两个级联的 Adaboost 分类器中被排除或是被接受的级数信息作为最终融合决策的重要依据. 除了在融合算法框架上与现有的其他融合模型不同之外, 本文所提出的融合模型是由一个目标的稳定部件训练得到的局部模型和由目标整体训练得到的全局模型组成的, 前者利用 Haar 小波计算简便的优点, 而后者利用了边界片段特征对颜色纹理形状变化较大的目标的鲁棒性. 二者间的几何位置模型可以根据局部模型的检测结果对全局模型进行粗定位, 从而大大减少了全局模型的搜索时间, 同时也提高了融合模型对目标的不同姿态的适应性.

1 算法

1.1 全局模型

全局模型由边界片段特征训练得到.

1.1.1 生成边界片段模板

本文采用统计区域分割 (Statistical region

merging, SRM)^[13] 算法分割待检测样本, 然后再提取区域的边界. 这种基于区域生长的分割方法受噪声影响较小, 可以在一定程度上减小冗余边界片段对分类器的干扰.

从已分割的训练样本中选出一部分作为生成边界片段的模板. 采用文献 [5] 中的种子生长法生成边界片段. 每一个边界片段由它的起始点坐标和一段链码描述.

1.1.2 边界片段特征

根据每个边界片段模板, 可以提取它与样本的两个匹配代价特征, 一个是形状匹配代价, 一个是位置匹配代价. 前者可利用 Chamfer 匹配算法计算, 首先按下式计算待匹配样本边缘图像的距离变换

$$DT_E(\mathbf{x}) = \min_{\mathbf{e} \in E} d(\mathbf{x}, \mathbf{e}) \quad (1)$$

其中, $E = \{\mathbf{e}\}$ 表示待匹配样本的边缘点集合, d 表示欧氏距离.

然后将边界片段模板在距离变换后的图像上逐点移动, 它可以看作是一个二维滤波器, 计算边界片段模板在每个位置上与样本的形状匹配代价

$$Ch_{E,T}(\mathbf{x}) = \frac{1}{N_T} \sum_{\mathbf{t} \in T} DT_E(\mathbf{t} + \mathbf{x}) \quad (2)$$

其中, $T = \{\mathbf{t}\}$ 表示边界片段上边缘点的集合, N_T 表示边界片段的长度; \mathbf{x} 表示平移向量. 匹配代价越小说明该边界片段与样本的形状相似度越大. 为了减小复杂背景对边界片段模板匹配的干扰, 模板移动的范围被限制在边界片段的原始位置附近. 因此形状匹配代价定义为匹配代价的局部最小值

$$S_T(E) = \min_{|\mathbf{x}| < R} \{Ch_{E,T}(\mathbf{x})\} \quad (3)$$

位置匹配代价定义为边界片段的原始位置与最佳匹配位置之间的欧氏距离

$$D_T(E) = d(\mathbf{p}_{\text{opt}}, \mathbf{p}_T) \quad (4)$$

$$\mathbf{p}_{\text{opt}} = \mathbf{p}_T + \arg \min_{|\mathbf{x}| < R} \{Ch_{E,T}(\mathbf{x})\} \quad (5)$$

其中, \mathbf{p}_T 为边界片段的原始位置; \mathbf{p}_{opt} 为该边界片段在待匹配样本上的最佳匹配位置.

样本与该边界片段的匹配代价定义为形状匹配代价和位置匹配代价的加权平方和, 权系数 λ 由 Adaboost 训练算法得到

$$C_T(E) = S_T^2(E) + \lambda D_T^2(E) \quad (6)$$

1.1.3 Adaboost 分类器

Adaboost 弱分类器定义如下

$$h_T(E) = \begin{cases} a_T, & C_T(E) > th_T \\ b_T, & \text{其他} \end{cases}$$

$$a_T < b_T \quad (7)$$

由于匹配代价的值越小说明该样本与正面样本越相似, 所以上式中弱分类器的输出 a_T 必须小于 b_T . 离散化加权系数 λ , 并在训练过程中选择使弱分类器错误率最小的加权系数 λ .

通过级联 Adaboost 学习算法可得到一系列由边界片段组成的星状模型, 以耳朵直立的宠物狗检测器为例 (如图 1 所示), 每个边界片段都可在一定范围内移动, 提高了模型的灵活性.



图 1 全局模型

Fig. 1 The global model

1.2 局部模型

局部模型由目标的一个稳定部件训练得到. 它的颜色、纹理和形状的变化都较小, 适宜用矩形类 Haar 小波特征来描述. 但是局部模型只关注目标的一个部件, 丢失了其他信息, 限制了分类器的性能.

1.3 融合策略

1.3.1 级联 Adaboost 分类器性能分析

根据级联 Adaboost 分类器的训练原理, 在第 N 级被排除的样本为正面样本的概率为

$$p_N(o|\omega_b) = \frac{p_o(R_o)^{N-1}(1-R_o)}{p_o(R_o)^{N-1}(1-R_o) + p_b(R_b)^{N-1}(1-R_b)} = \left(1 + \frac{p_b(1-R_b)}{p_o(1-R_o)} \left(\frac{p_b}{p_o}\right)^{N-1}\right)^{-1} \quad (8)$$

其中, ω_b 表示被分类器排除的样本, o 表示真实目标, p_o 和 p_b 分别表示正负样本在图像中出现的概率, R_o 和 R_b 分别表示每一级 Adaboost 分类器的正负样本通过率. 图 2 给出了被排除样本是真实目标的概率随着级数增加的变化趋势的仿真结果 (参数设置为: $p_b/p_o = 10^4$, $R_b = 50\%$, $R_o = 99\%$).

从仿真结果可以看出在分类器的最后几级, 这个概率增加很快. 这是因为在分类器最后几级, 虽然分类器检测结果的准确率提高了, 但是漏检率提高的速度更快. 事实上, 由于单一特征的分辨能力有限, 级联 Adaboost 分类器的级数越高, 虚警率降低的趋势逐渐变缓, 与理论上的指数衰减偏离逐渐加大, 而正面样本随着分类器级数的增加被错误排除

的概率却在增大, 这就使得最后几级中被排出的样本是正面样本的概率增加得更快.

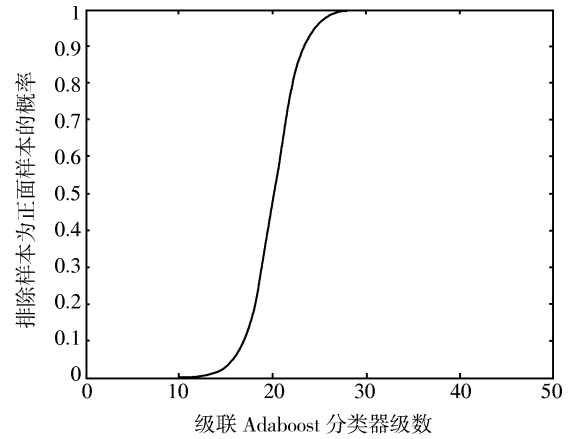


图 2 真实目标在被排除样本中的比例随分类器级数变化的仿真结果

Fig. 2 The simulation result of the rate of true object in the rejected samples vs. the stage index

1.3.2 融合策略

基于以上的分析, 本文提出了一种基于级联 Adaboost 分类器的“级联-加和”融合策略. 由于局部模型采用矩形类 Haar 小波作为特征, 检测速度较快, 所以在本文提出的融合模型中, 先采用局部模型快速排除待检测图片中的大部分背景区域, 然后再由信息量更丰富的全局模型验证局部模型的决策. 一般的级联策略是将待检测样本顺序通过各个分类器, 后面的分类器只关注那些通过前面分类器的样本. 而本文所提出的融合策略还关注那些在第一个分类器中被错误排除的正面样本, 因此融合模型中存在多条检测支路, 即第二个分类器不仅需要再次验证通过第一个分类器的样本, 还要对在第一个分类器的后几级中被排除的样本再次评估以给出综合决策. 对每一个样本来说, 检测过程是局部模型与全局模型的级联, 而对最后的检测结果集合来说, 它是各条检测支路的检测结果的并集. 融合策略的框架如图 3 (见下页) 所示.

图 3 中, R_n 表示在局部模型第 n 级中被排除样本是正面样本的置信度, P_n 表示通过局部模型第 n 级的样本是正面样本的置信度.

在全局模型中对再次评估的样本设置置信度阈值是设计融合模型的关键问题. 由于局部模型与全局模型是独立训练得到的, 各自所描述的目标也不完全相同, 所以可以假设两个分类器所产生的虚警是独立的. 基于这个假设, 如果在局部模型的第 N_c 级被排除的样本需要通过 N_g 级全局模型的验证, 那么融合模型所做出的最终决策的理论置信度为

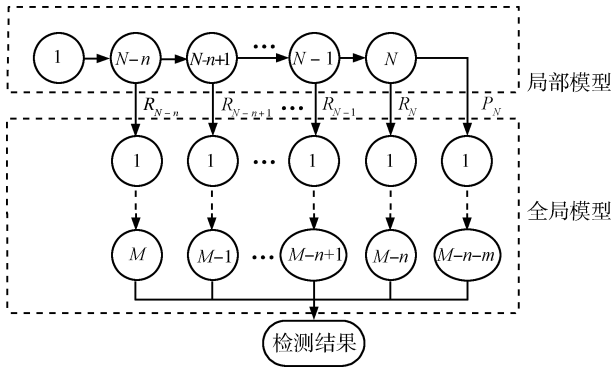


图3 级联-加和融合算法框架

Fig. 3 Framework of the "cascaded-additive" algorithm

$$p_N(o|\omega_o) = \left(1 + \frac{p_b(1-R_b^c)}{p_o(1-R_o^c)} \left(\frac{R_b^c}{R_o^c}\right)^{N_c-1} \left(\frac{R_b^g}{R_o^g}\right)^{N_g}\right)^{-1} \quad (9)$$

其中, ω_o 表示融合模型的检测结果, R_o^c 和 R_b^c 分别表示正面样本和负面样本在局部模型每一级中的通过率, R_o^g 和 R_b^g 分别表示正面样本和负面样本在全局模型每一级中的通过率. 从式 (9) 可以看出, 理论置信度由正负样本在各级中的通过率以及局部模型和全局模型中使用的分类器级数决定. 更进一步, 如果局部模型和全局模型的正负样本通过率一致, 那么最终的融合模型的决策置信度仅由两个级联的 Adaboost 分类器的级数决定. 为了保证融合模型给出的最终决策的置信度恒定, 在局部模型中越早被排除的样本在全局模型中需要通过越多级数的分类器验证, 具体分为以下两种情况:

1) 对于在局部模型的最后几级中被排除的样本, 全局模型中需要使用的级数与该样本在局部模型中被排除的级数之和为常数 N

$$N = N_c + N_g \quad (10)$$

其中, N_c 为局部模型的分器级数, N_g 为全局模型的分器级数.

2) 对于通过局部模型全部级联分类器的样本, 全局模型中需要使用的级数按下式计算

$$N_{\text{Alast}} = N_{\text{Rlast}} - \frac{\log\left(\left(\frac{R_b^c}{R_o^c}\right) \frac{(1-R_o^c)}{(1-R_b^c)}\right)}{\log\left(\frac{R_b^g}{R_o^g}\right)} \quad (11)$$

其中, N_{Alast} 表示对在局部模型的最后一级分类器被接受的样本在全局模型中需要的级数, N_{Rlast} 表示对在局部模型中被最后一级分类器排除的样本在全局模型中需要的级数.

式 (11) 的推导过程如下: 根据最终检测结果的置信度恒定的准则, 有

$$\left(1 + \frac{p_b(1-R_b^c)}{p_o(1-R_o^c)} \left(\frac{R_b^c}{R_o^c}\right)^{N_c-1} \left(\frac{R_b^g}{R_o^g}\right)^{N_{\text{Rlast}}}\right)^{-1} = \left(1 + \frac{p_b}{p_o} \left(\frac{R_b^c}{R_o^c}\right)^{N_c} \left(\frac{R_b^g}{R_o^g}\right)^{N_{\text{Alast}}}\right)^{-1}$$

$$\left(\frac{R_b^g}{R_o^g}\right)^{N_{\text{Rlast}}-N_{\text{Alast}}} = \left(\frac{R_b^c}{R_o^c}\right) \frac{(1-R_o^c)}{(1-R_b^c)}$$

由此即可得到

$$N_{\text{Alast}} = N_{\text{Rlast}} - \frac{\log\left(\left(\frac{R_b^c}{R_o^c}\right) \frac{(1-R_o^c)}{(1-R_b^c)}\right)}{\log\left(\frac{R_b^g}{R_o^g}\right)}$$

1.3.3 融合模型检测性能分析

假设全局模型与局部模型中正负样本在每一级的通过率都分别相等, 那么融合模型的理论检测率和虚警率为

$$DR_{\text{cm}} = n(R_o)^{N_c-1}(1-R_o)(R_o)^{N_g} + (R_o)^{N_c}(R_o)^{N_g-m} = nR_o^{N-1}(1-R_o) + R_o^{N-m} \quad (12)$$

$$FPR_{\text{cm}} = n(R_b)^{N_c-1}(1-R_b)(R_b)^{N_g} + (R_b)^{N_c}(R_b)^{N_g-m} = nR_b^{N-1}(1-R_b) + R_b^{N-m} \quad (13)$$

其中, R_o 为正面样本通过率, R_b 为负面样本通过率, n 为级联-加和融合模型中用来验证被局部模型排除的样本的支路的个数, 而单一特征的级联 Adaboost 分类器的理论检测率和虚警率为

$$DR_{\text{sm}} = R_o^N \quad (14)$$

$$FPR_{\text{sm}} = R_b^N \quad (15)$$

由式 (12)~(15), 图 4 (见下页) 对比了融合模型和单一特征模型的 ROC (Receiver operating characteristics) 曲线, 仿真参数为: $R_o = 99\%$, $R_b = 50\%$. 仿真结果表明, 融合模型的检测性能优于单一特征的 Adaboost 级联分类器.

图 5 (见下页) 给出了不同支路个数下融合算法的检测性能. 从图中可以看出随着 n 的增大, 融合模型的检测性能也逐渐提高. 但在实际中, 并不是选取的支路个数越多越好, 因为支路个数增加会导致运算量的增大, 所以在实际应用中支路个数的选择需要在检测性能与运算开销之间做出权衡. 当 $n = 0$ 时, 也就是融合模型中没有用来验证被局部模型排除的样本的支路时, 该融合模型退化为简单的级联模型.

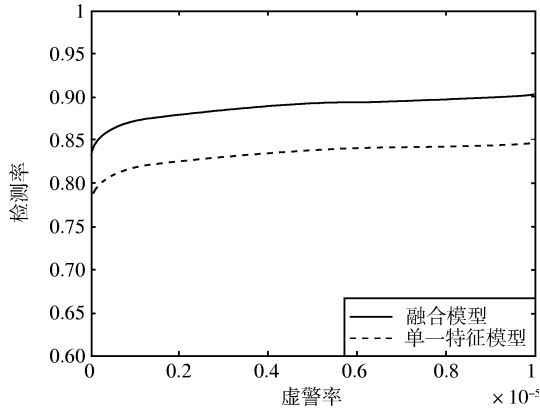


图 4 级联-加和融合算法与单一特征分类器性能比较的仿真结果

Fig. 4 Simulation result of comparison between cascaded-additive algorithm and classifier of single feature

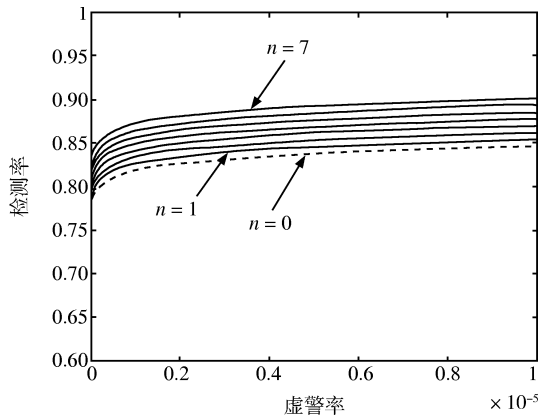


图 5 支路个数对检测器性能的影响

Fig. 5 Detection performance with different branch numbers

1.3.4 几何模型

由于局部模型与全局模型检测的目标不同, 在融合模型中, 需要建立二者之间的几何模型, 如图 6 所示. 根据局部模型的检测结果, 可以计算全局目标的大小并对全局目标进行粗定位. 由于局部模型的检测结果与真实目标在大小上存在差异, 为了提高估算的准确度, 对局部模型的检测结果进行一系列尺寸放缩. 另外, 由于对不同的样本来说, 局部模型和全局模型的相对位置可能存在一定变化, 为了提高融合检测模型的鲁棒性, 在全局目标可能出现的区域定义了搜索范围, 尺寸略大于全局目标的尺寸, 全局目标搜索窗口只需在这个搜索区域内滑动, 大大提高了基于边界片段的检测器的检测效率.

2 实验

为了验证算法的有效性, 我们在宠物狗和 UIUC 汽车数据库上进行了实验. 如无特别说明, 实验参数

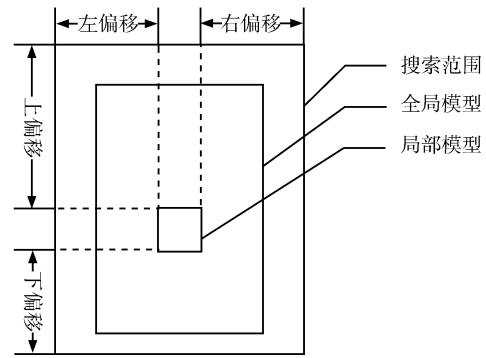


图 6 几何模型

Fig. 6 Geometric model

均设置如下: 正负训练样本在级联 Adaboost 分类器中的通过率分别设置为 100% 和 50%. 边界片段长度的取值范围为 10 到 100 个像素, 步长为 10 个像素. 检测器在尺度空间上的搜索步长为 1.05. 在全局模型的检测中, 对重叠面积超过检测目标大小 50% 的检测结果进行合并.

2.1 宠物狗数据库

该数据库从互联网上搜集得到. 旋转原始图像使目标的双眼保持水平. 全局模型针对宠物狗的整个头部训练, 各级正负训练样本各为 200 个, 归一化大小为 75×50 . 局部模型针对宠物狗的鼻子训练, 各级正负训练样本分别为 1564 个和 3000 个, 归一化大小为 24×24 . 全局模型由 50 级 Adaboost 分类器构成, 局部模型由 19 级 Adaboost 分类器构成. 几何模型中左偏移、右偏移、上偏移、下偏移与局部模型边长的比分别为 (2.5, 2.5, 4.5, 2). 全局模型与局部模型的宽度、高度之比分别为 4 和 6. 由于全局模型与局部模型的尺度比例变化较大, 为了提高融合模型的鲁棒性, 将局部模型的检测结果以 1.05 为步长进行 5 级放缩, 再分别根据放缩后的结果计算全局模型的大小以及搜索区域. 测试集共 292 幅, 320 个目标. 在尺度空间搜索时采用放缩待检测图像而保持检测窗口大小恒定的方法.

调整全局模型与局部模型的级数控制参数 N (取 50~61) 绘制 ROC 曲线. 级联-加和融合模型中支路的个数 n 取 7. 支路个数的选取原则是, 从局部模型检测率较高 (在本实验中为 95%) 的级数开始, 到局部模型虚警率较低 (本实验中取 0.33 个/图像) 的级数结束. 验证被排除样本的支路是从局部模型的 14 级开始到 19 级结束. 从局部模型的 14 级开始表明能够通过局部模型的前 13 级分类器的样本都可进入全局模型的再次验证环节.

2.1.1 与单一特征的模型比较

从图 7 的对比结果可以看出融合模型的检测性

能与单一特征的模型相比有较大改善. 实验结果也较好地符合图 4 的理论仿真结果.

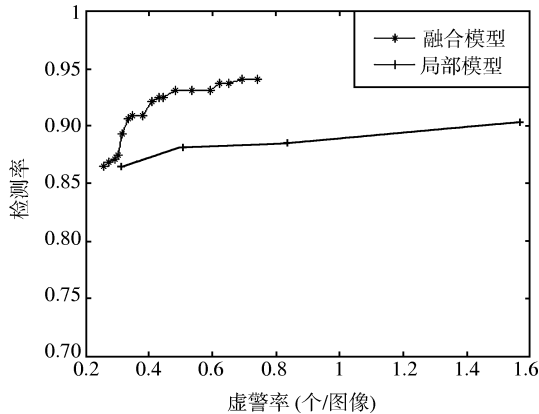


图 7 级联-加和融合算法与单一特征检测器性能比较的实验结果

Fig. 7 Experimental result of comparison between cascaded-additive algorithm and classifier of single feature

2.1.2 与简单级联模型比较

为了体现实验的公平性, 两种融合策略的检测率上限均为局部模型的前 13 级的检测率, 并采用完全相同的几何模型. 在简单级联的融合策略中局部模型的级数为 13 级, 通过调整全局模型的级数来绘制 ROC 曲线上的各点.

图 8 中的对比结果表明, 级联-加和的融合策略要优于简单级联的分类器. 正如第 1.3.3 节中所分析的, 简单级联的分类器是级联-加和融合框架在“验证被第一个分类器排除的样本的支路个数为零”的条件下的特例. 而图 8 出的实验结果也较好地符合了图 5 中的理论仿真结果.

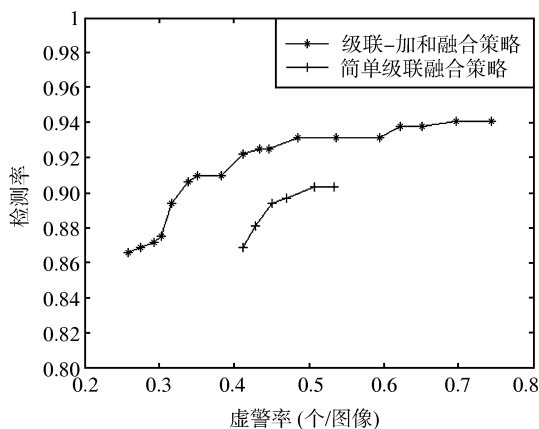


图 8 级联-加和融合策略与简单级联融合策略的检测性能对比

Fig. 8 Experimental result of comparison between cascaded-additive algorithm and simple cascaded algorithm

2.1.3 融合模型与全局模型的检测速度比较

以边界片段作为特征的检测算法的复杂度主要由两部分构成, 即检测算法的复杂度和图像分割算法的复杂度. 前者中边界片段特征的计算是影响检测速度的最主要原因. 在融合模型中, 由于基于 Haar 小波特征的局部模型可以快速排除待检测图像中的大部分区域, 从而极大地减小了全局模型在尺度上和空间上的搜索范围. 从图像分割算法的复杂度来看, 由于本文采用的分割算法涉及到对像素间差异大小的排序, 复杂度为 $O(n \log n)$, 因此受图像分辨率影响很大. 在融合模型中, 由于全局模型的待检测区域面积较小, 所以图像分割算法的效率也较高. 实验中对比了 39 级的全局模型以及控制参数 N 为 54 级的融合模型的检测时间. 在第 2.1.1 节的实验中, 由于全局模型需要对整幅图像在各尺度空间及空间位置上进行搜索, 对内存要求很高, 所以实验中将分辨率超过 950×850 的图片都归一化到 950×850 . 为了体现实验的公平性, 速度比较实验只在所有分辨率小于 950×850 的图片上进行, 共 222 幅. 在搜索步长, 分割参数与后处理方法完全相同的情况下, 融合模型所用的检测时间仅为全局模型的 17.8%.

在局部模型的检测结果中存在较多在尺度和空间位置上非常接近的检测结果, 融合实验中只合并了重叠部分占检测目标的比例超过 80% 的局部模型的检测结果, 这个参数并没有手工调整到最优, 因此全局模型实际上在一些位置上进行了重复检测, 耗费了较多的时间. 如果根据几何模型的设置参数对局部模型的后处理参数进行优化, 则可以进一步提高融合模型的检测速度.

2.1.4 检测结果示例

图 9 给出了一些成功的检测结果, 通过全局模型的再次验证, 成功地检测出了一些在局部模型的倒数几级中被排除的目标. 同时, 因为在几何模型中搜索范围要略大与全局目标的尺寸, 所以融合模型对目标姿势的变化有一定的鲁棒性, 并可以根据全局模型与局部模型检测结果之间的相对位置估计目标的姿势.



图 9 检测结果示例

Fig. 9 Samples of dog detection results

2.2 UIUC 汽车侧面图像数据库

2.2.1 与全局模型、局部模型的检测结果比较

对 UIUC 汽车侧面数据库, 局部模型针对车轮训练, 训练样本全部来自数据库的训练样本集. 各级正负训练样本分别为 1886 个和 2000 个, 归一化大小为 24×24 ; 全局模型针对整个轿车训练, 各级的正负训练样本均为 550 个, 归一化大小为 40×100 . 利用数据库中提供的训练样本集训练全局模型, 不区分轿车的朝向. 不同于文献 [4], 对测试样本也只进行一个方向的检测. 在融合模型中, 局部模型分类器共 16 级, 全局模型共 7 级. 融合模型的支路个数为 4. 采用数据库中提供的全部 170 张测试图片进行测试, 共包括 200 个目标 (而文献 [4] 只采用了其中的 164 幅测试图像, 共 193 个目标).

由于局部检测结果不能区分轿车的前后轮, 因此在实验中将每个局部检测结果分别作为前轮和后轮计算两个搜索范围. 搜索范围的左偏移, 右偏移, 上偏移, 下偏移与局部模型边长的比分别为 $(1, 3, 1, 0)$ 和 $(3, 1, 1, 0)$, 全局模型与局部模型的宽度、高度之比分别为 $25/6$ 和 $5/3$. 因为轿车车轮与车身的大小相对固定, 局部模型的检测结果仅以 1.05 的步长放缩 3 级. 如果检测器在两个搜索区域中都击中了目标, 那么选取分类器输出最大值较大的全局检测结果, 如果两个搜索区域中分类器输出的最大值相同, 则选取击中次数较多的搜索区域中的检测结果.

在表 1 的统计结果中, 局部模型只要检测到轿车的一个车轮就认为它可以正确检测到目标. 对比结果表明, 级联-加和融合算法由于综合了全局模型与局部模型的优势, 能够提高检测器的性能.

表 1 融合模型与全局模型和局部模型比较
Table 1 Comparison with global model and component model

算法	R-P EER
全局模型	93%
局部模型	88%
级联-加和融合算法	96%

2.2.2 与其他单一模型算法的比较

我们将本文提出的融合检测算法和文献 [4, 14–16] 中提出的单一模型的目标检测算法进行了比较. 文献 [14, 16] 中采用基于感兴趣点的图像块作为模板, 以模板与待检测图像的相关性作为特征. 文献 [15] 则进一步用 PCA 方法对图像块降维作为特征. 文献 [4] 以边界片段为特征, 与本文提出的算法相比, 文献 [4] 中采用了较多的后处理算法. 首先在图像

空间上对弱分类器的输出进行平滑以减少检测结果中的噪声, 然后比较检测结果的峰值输出位置与真实目标的中心位置之间的距离, 如果小于 25 像素, 则认为该目标被正确检测. 这相对于大小为 40×100 的目标来说是个较为宽松的阈值. 而本文提出的算法由于有局部模型以及全局模型的两次验证, 所以能够给出更加精确的目标位置 (见图 10).



图 10 UIUC 侧面轿车数据库检测结果

Fig. 10 Samples of UIUC side-view car detection results

表 2 中文献 [14–16] 的实验结果摘自文献 [4]. R-P EER 性能评价参数的对比结果表明, 本文提出的级联-加和融合检测算法的性能要明显优于其他单一模型的算法.

表 2 与其他单一模型的算法的比较
Table 2 Comparison with other state-of-the-art algorithms

算法	R-P EER
Agarwal ^[14]	79%
Fergus ^[15]	88.5%
Leibe ^[16]	91.0%
Shotton ^[4]	92.8%
级联-加和融合算法	96%

3 结束语

本文提出了“级联-加和”的融合检测算法, 它由独立训练得到的级联 Adaboost 全局模型和局部模型组成. 基于矩形类 Haar 小波特征的局部模型可以快速排除大量背景区域, 而对颜色纹理形状变化较为鲁棒的全局模型通过预先设定的几何模型对局部模型在最后几级的检测结果进行验证. 该融合框架不仅在特征层面综合了 Haar 小波特征和边界片段特征的优势, 在结构上也不同于简单的并行融合以及级联融合形式. 克服了并行融合形式时间复杂度以及级联融合形式检测率难以提高的缺点. 实验结果表明, 本文提出的融合检测算法的检测性能对于含有稳定部件的、在颜色、纹理以及形状上变化较大的目标有较好的检测效果.

事实上, 这种级联-加和的融合策略不仅适用于级联 Adaboost 分类器的融合. 只要级联分类器在每一级的正负样本通过率是可以估计的, 那么这种融合策略就可以推广到任何一种级联分类器的融合.

References

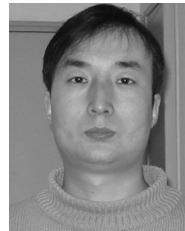
- 1 Viola P, Jones M J. Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE, 2001. 511–518
- 2 Leibe B, Schiele B. Analyzing appearance and contour based methods for object categorization. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Wisconsin, USA: IEEE, 2003. 409–415
- 3 Zhang W, Yu B, Zelinsky G J, Samarasinghe D. Object class recognition using multiple layer boosting with heterogeneous features. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 323–330
- 4 Shotton J, Blake A, Cipolla R. Contour-based learning for object detection. In: Proceedings of the 10th IEEE International Conference on Computer Vision. Beijing, China: IEEE, 2005. 503–510
- 5 Opelt A, Pinz A, Zisserman A. A boundary-fragment-model for object detection. In: Proceedings of the 9th European Conference on Computer Vision. Graz, Austria: Springer, 2006. 575–588
- 6 Opelt A, Pinz A, Zisserman A. Fusing shape and appearance information for object category detection. In: Proceedings of the 17th British Machine Vision Conference. Edinburgh, UK: Microsoft Lisa, 2006. 117–126
- 7 Mohan A, Papageorgiou C, Poggio T. Example-based object detection in images by components. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(4): 349–361
- 8 Mikolajczyk K, Schmid C, Zisserman A. Human detection based on a probabilistic assembly of robust part detectors. In: Proceedings of the 8th European Conference on Computer Vision. Prague, Czech Republic: Springer, 2004. 69–82
- 9 Heisele B, Serre T, Poggio T. A component-based framework for face detection and identification. *International Journal of Computer Vision*, 2007, **74**(2): 167–181
- 10 Ichikawa K, Mita T, Hori O, Kobayashi T. Component-based face detection method for various types of occluded faces. In: Proceedings of the 3rd International Symposium on Communications, Control and Signal Processing. Julians, Malta: IEEE, 2008. 538–543
- 11 Fritz M, Leibe B, Caputo B, Schiele B. Integrating representative and discriminant models for object category detection. In: Proceedings of the 10th IEEE International Conference on Computer Vision. Beijing, China: IEEE, 2005. 1363–1370
- 12 Yang J, Byun H. A combination of generative and discriminative approaches to object detection. In: Proceedings of the 18th International Conference on Pattern Recognition. Hong Kong, China: IEEE, 2006. 249–253
- 13 Statistical region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(11): 1452–1458
- 14 Agarwal S, Roth D. Learning a sparse representation for object detection. In: Proceedings of the 7th European Conference on Computer Vision. Copenhagen, Denmark: Springer, 2002. 97–101
- 15 Fergus R, Perona P, Zisserman A. Object class recognition by unsupervised scale-invariant learning. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Madison, USA: IEEE, 2003. 264–271
- 16 Leibe B, Leonardis A, Schiele B. Combined object categorization and segmentation with an implicit shape model. In: Proceedings of the 8th European Conference on Computer Vision, Workshop on Statistical Learning in Computer Vision. Prague, Czech Republic: Springer, 2004. 17–32



崔潇潇 清华大学博士研究生. 主要研究方向为计算机视觉和模式识别. 本文通信作者.

E-mail: cuixx03@mails.tsinghua.edu.cn
(**CUI Xiao-Xiao** Ph. D. candidate at Tsinghua University. Her research interest covers computer vision and video surveillance. Corresponding au-

thor of this paper.)



姚安邦 清华大学博士研究生. 主要研究方向为目标检测和识别, 目标追踪和特征选择.

E-mail: nianming@gmail.com
(**YAO An-Bang** Ph. D. candidate at Tsinghua University. His research interest covers object detection and recognition, object tracking, and fea-

ture selection.)



王贵锦 清华大学电子工程系副教授. 主要研究方向为视频处理与通信.

E-mail: wangguijin@tsinghua.org.cn
(**WANG Gui-Jin** Associate professor in the Department of Electronic Engineering, Tsinghua University. His research interest covers video processing and communication.)



林行刚 清华大学电子工程系教授. 主要研究方向为图像/视频处理与分析, 视觉信息压缩编码及应用和模式识别.

E-mail: xglin@tsinghua.edu.cn
(**LIN Xing-Gang** Professor in the Department of Electronic Engineering, Tsinghua University. His research interest covers image/video processing

and analysis, video coding and communications, and pattern recognition.)