

文章编号: 1001-0920(2011)11-1643-06

基于多尺度上下文信息的图像目标分类算法

杨蕾, 宋永红, 郑南宁

(西安交通大学 人工智能与机器人研究所, 西安 710049)

摘要: 针对真实场景图像的目标分类问题, 提出一种基于多尺度上下文信息的分类算法. 首先运用一种软判决采样机制对图像进行局部信息采样, 使场景内混合的各类信息以一种鲁棒的方式得到有效分离; 然后, 进一步基于软判决采样和统计特征表达机制, 计算各空间尺度下的目标上下文统计特征; 最后, 通过逻辑回归分类算法有效地融合多尺度的上下文信息, 并作出分类决策. 实验表明, 所提出的算法能更好地刻画真实场景下目标的特性, 明显提高图像目标分类性能.

关键词: 图像目标分类; 多尺度上下文信息; 软判决采样机制; 外观统计特征

中图分类号: TP391

文献标识码: A

Object categorization using context from multi-spatial levels

YANG Lei, SONG Yong-hong, ZHENG Nan-ning

(Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China. Correspondent: YANG Lei, E-mail: lyang@aiar.xjtu.edu.cn)

Abstract: To categorize objects in the real-world scene images, a method is proposed by exploiting multi-spatial extent context. Firstly, a soft decision-based sampling mechanism is utilized in the local image patch sampling process, by which, mixed information in the scene can be separated in an effective and robust way. Then, by using the soft decision-based sampling mechanism and the statistical representation methods, the statistical feature for each spatial extent context can be computed. Finally, a logistic regression classification method is adopted to integrate multiple spatial extent context information and make the final decisions. The experiments show that, the proposed method can better model the objects in the real world scenes, and thus apparently improves the object categorization performance.

Key words: object categorization; multiple extent context; soft decision-based sampling mechanism; statistical appearance representation

1 引言

目标分类技术是自动分析、理解图像与场景的关键技术之一, 旨在判断图像媒体中是否存在属于某类物体的实体, 可应用于图像自动标注、海量图像搜索、自动图片过滤、图像编码及压缩、智能监控等领域. 真实场景下, 如何体现形态各异同类个体之间的共性, 同时突出类间差异, 如何建立具有一般推广性的统一的分类模型以及如何消除复杂背景对感兴趣目标的干扰等, 都是目标分类技术当前所面临的主要挑战^[1].

目前, 非常流行的基于外观统计特征的 BoF (Bag-of-Features) 分类模型对目标类内差异、变形及遮挡表现出了良好的鲁棒性^[1-4], 但传统的 BoF 方法实质

上刻画的是图像内的整体场景信息, 并将该统计特征用于目标分类任务. 对于真实场景, 特别是复杂场景下的目标, 比如场景内出现的较小的目标, 传统 BoF 方法不能明确地刻画感兴趣目标的特性, 因而容易造成误判. 如何更有效地针对真实场景内的感兴趣目标进行建模, 同时保证算法具有较好的适应遮挡、变形以及类内差异的鲁棒性是本文研究的重点.

本文在 BoF 模型框架下, 提出了一种新的基于多尺度上下文信息的目标分类算法. 利用一种自顶向下的软判决机制指导局部信息采样过程, 以一种鲁棒的方式有效地提取并刻画目标自身外观上下文信息、各尺度局部上下文信息以及全局场景信息, 同时避免在目标分类过程中实现困难的目标定位或语义性分割.

收稿日期: 2010-07-02; 修回日期: 2010-08-31.

基金项目: 国家973计划项目(2007CB311005); 国家自然科学基金重大科学研究计划项目(90920008).

作者简介: 杨蕾(1983-), 女, 博士生, 从事目标分类与识别、机器视觉的研究; 郑南宁(1952-), 男, 中国工程院院士, 教授, 博士生导师, 从事自动控制、模式识别及计算机视觉等研究.

多尺度上下文外观统计信息通过逻辑回归分类算法有效地融合,并用于最终的分类决策.与传统的对全局场景建模的 BoF 方法相比,该方法首先分层次刻画了各尺度的目标上下文信息,然后通过决策层的信息融合更加有效地实现了对前景目标的建模,同时维持了统计表达方法的鲁棒性.

2 多尺度上下文信息

2.1 BoF 统计特征

本文借鉴 BoF 模型的统计特征计算方式来生成各个尺度的目标上下文特征.在 BoF 模型中,一幅图像被看成包含多个独立局部特征的特征包 (bag-of-features), 这些局部特征经量化后形成一组视觉关键词, 通过计算视觉关键词出现的频率, 每幅图像可以编码为一个用于分类的统计直方图特征向量. BoF 模型的统计特征计算主要包含 4 个模块: 局部区域采样、局部特征描述、局部特征量化和全局统计直方图特征计算模块, 具体流程如图 1 所示.

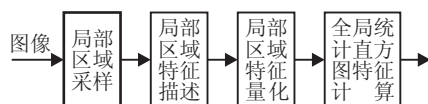


图 1 BoF 方法的统计特征计算流程

在最初的局部信息采样过程中,一幅图像依据某种采样准则抽取出 m 个局部小区域 $PT = \{pt_i, 1 \leq i \leq m\}$. 局部区域的采样结果将直接影响最终统计特征描述的内容.传统 BoF 使用的采样方法主要有: 兴趣点采样以及随机(密集)采样, 如使用 Harris-Laplace, HoG 和随机检测子等^[4]. 这些自底向上的采样机制对图像内的所有区域采用统一的度量检测标准, 前景和背景内的局部信息不加区分地混合体现在最终的用于分类的统计特征向量上. 因而, 基于这些方法的 BoF 模型实质上描述了全局场景信息. 然而, 在真实场景, 尤其在复杂场景中, 上述常用的采样方法容易导致前景目标的特性被背景噪声所掩盖, 当全局场景信息不具有刻画目标类的能力时, BoF 方法往往失效.

从另一个角度看, 与传统 BoF 模型中的全局场景信息一样, 单纯的目标前景外观信息不一定具有充分刻画感兴趣目标的能力. 图 2 中的 2 个例子反映了这种现象. 每一行从左到右依次是原图像, 感兴趣目标图像以及能够用于充分判定目标类属的相关区域范围. 对于第 1 幅车的图像, 局部的车图像便足以作为识别依据; 而对于第 2 幅图像中的沙发, 只有当观测到它的一定范围的局部上下文信息(人)时才能得出有效判断. 由此可见, 目标自身外观上下文信息、目标周围的局部上下文信息、全局场景上下文信息, 各自

在目标分类中的功能是随分类任务的不同而不同的. 对于不同的分类任务, 需要考虑不同空间尺度的目标上下文信息. 基于传统采样机制的 BoF 方法无法分层次刻画以感兴趣目标为中心的多尺度上下文信息. 为此, 本文在 BoF 框架下引入了一种软判决采样机制.



图 2 多尺度上下文信息在目标分类中的作用

2.2 软判决采样机制

实现精准并通用的感兴趣目标或目标部件检测本身是一个困难的开放式问题. 为了保证算法的鲁棒性, 结合 BoF 统计描述的特点, 本文采用了软判决采样机制. 所谓软判决采样策略, 是与基于目标(部件)检测的二值化判决策略相对应的, 即使用概率分布指导采样过程(而不是使用二值化的判定结果), 实现各种感兴趣信息的提取.

软判决采样机制的关键是用于指导采样的概率分布的生成. 通过计算图像场景中不同位置(坐标 (x, y)) 不同尺度 (s) 的局部小区域出现在感兴趣目标上的概率, 可以得到一个概率分布 $p(x, y, s)$, 而多尺度上下文信息的提取都是在计算这个分布的基础上进行的. 作者采用文献[5]提出的一种自顶向下的处理机制进行三维概率分布 $p(x, y, s)$ 的计算. 对于任一图像, 软判决采样机制的建立过程总结如下:

1) 图像过分割. 下述所有分析计算的基元是过分割后的图块 RG_i .

2) 结合区域颜色、纹理以及形状信息, 基于统计学习机制, 计算自顶向下的类别信息 S_O . 假设这几种视觉信息相互独立, 以获得松弛的类别约束. 令某视觉线索经量化后的特征空间内一特征为 F_i , $O(\bar{O})$ 表示存在(不存在)感兴趣目标, F_i 的分类能力 $R(F_i)$ 可以在训练过程通过学习获得, 有

$$R(F_i) = P(O|F_i) = \frac{P(F_i|O)}{P(F_i|O) + P(F_i|\bar{O})}. \quad (1)$$

对于任意分割区域 RG_i , 设其相应的各线索的特征为: 颜色 C_i , 纹理 TX_i 和形状 SH_i (量化后), 则由

$$S_O(x, y) = R(C_i)R(TX_i)R(SH_i), \quad \forall (x, y) \in RG_i, \quad (2)$$

可计算自顶向下的目标先验度量图 $S_O(x, y)$.

3) 归一化 S_O 到 $[0, 1]$ 区间. 为高效计算各位置各尺度局部区域的采样概率, 利用 S_O 的积分图像 $I(x,$

$y)^{[6]}$ 来快速计算 $p(x, y, s)$, 有

$$p(x, y, s) = I\left(x + \frac{s}{2}, y + \frac{s}{2}\right) + I\left(x - \frac{s}{2}, y - \frac{s}{2}\right) - I\left(x + \frac{s}{2}, y - \frac{s}{2}\right) - I\left(x - \frac{s}{2}, y + \frac{s}{2}\right). \quad (3)$$

4) 利用三维概率分布函数 $p(x, y, s)$ 指导采样过程, 重点提取感兴趣目标上的局部区域. 文献 [5] 中的实验表明, 结合统计特征表达模型 (BoF) 和松弛的类别约束信息可以有效鲁棒地描述前景目标外观, 抑制背景噪声.

5) 基于 $p(x, y, s)$, 计算各层次感兴趣信息采样使用的概率分布, 提取各类上下文信息.

2.3 多尺度上下文信息提取

传统 BoF 模型实际上是对全局场景上下文信息建模. 本文引入多空间尺度上下文信息进行目标描述. 这里多尺度上下文信息主要包括: 目标自身信息、局部上下文信息以及全局上下文信息. 2.2 节计算的 $p(x, y, s)$ 可以指导采样过程重点提取前景目标上的信息, 并对目标自身外观进行有效地刻画. 下面主要介绍 BoF 框架下, 基于 $p(x, y, s)$ 全局上下文及局部上下文信息的提取方法.

对于任一图像, 由式 (2) 计算的前景目标出现

的二维概率分布 $S_O(x, y)$, 可以计算背景出现的概率度量图 $S_{C_g}(x, y)$, 有

$$S_{C_g}(x, y) = 1 - S_O(x, y). \quad (4)$$

这个概率分布将用于描述全局上下文信息.

为计算目标周围的局部上下文信息, 首先通过二值化 $S_O(x, y)$ 来近似估算前景目标所占的空间区域, 如图 3(a) 所示. 令二值化阈值为 TH, δ 为二值化 (0 和 1) 判决函数, 设二值图 $\delta(S_O(x, y) > TH)$ 包含 n 个不同区域. 第 i 个区域的中心位置及内切圆半径分别为 (x_i, y_i) 与 r_i , 则用于指导该尺度下局部上下文信息提取的二维概率分布为

$$S_{C_l}(x, y) = (1 - S_O(x, y)) \times \text{sgn}\left(\sum_{i=1}^n \delta(\sqrt{(x - x_i)^2 + (y - y_i)^2} \leq 2r_i)\right), \quad (5)$$

式中 $\text{sgn}(\cdot)$ 是符号函数. 本文使用 2 个阈值 $\text{TH}_l = 0.5$ 和 $\text{TH}_h = 0.75$ 来分别计算 2 个不同尺度的局部上下文特征. 图 3(b) 中感兴趣的为目标为沙发, 从左到右分别显示了原图、针对于前景、全局上下文以及 2 个尺度的局部上下文描述所使用的二维采样概率. 使用式 (3) 中的方法, 上述二维概率分布最终均转化为关于位置和尺度的三维概率分布, 以指导采样.



(a) 目标所占空间区域的估算

(b) 提取各尺度上下文信息使用的采样概率分布示意

图 3 多尺度上下文信息提取示例

3 多尺度上下文信息融合及目标分类

3.1 整体分类算法描述

根据 2.3 节介绍的多空间尺度上下文信息提取方法, 对于每幅图像可以分别生成用于提取各种信息的概率分布, 进而基于软判决采样机制得到相应的多组采样结果集合. 对于每个集合内的各个局部区域, 再通过特征描述环节计算其特征, 则每组采样结果集合可转化为一个包含多个局部特征的特征包. 根据 BoF 统计特征计算流程, 各个特征包可进一步表示为一个统计直方图特征向量. 本文采用最基本的 BoF 系统框架^[3], 使用 SIFT 局部特征描述方法^[7], 特征空间的量化采用 K-Means 聚类. 对于任一图像, 针对某分类任务 (是否含有某类感兴趣目标) 可计算得到 4

个统计特征 $\{f_O, f_{C_g}, f_{C_{l1}}, f_{C_{l2}}\}$, 分别描述了目标外观信息、全局场景上下文信息、大尺度局部上下文信息以及小尺度局部上下文信息.

对于各尺度的目标上下文特征, 训练各自相应的分类器, 而对于一幅输入图像则可以获得一组目标出现的响应值 $\{t_O, t_{C_g}, t_{C_{l1}}, t_{C_{l2}}\}$. 基于这组前端分类器响应值, 通过训练一个逻辑回归分类器来实现多尺度上下文信息的决策层融合. 这里前端分类器使用了基于高斯核的 SVM 分类器^[8], 后端的决策层融合使用了 l_1 正则化逻辑回归分类器. 通过学习不同尺度上下文信息的预测结果在最终分类决策中的权重, 实现最优的目标上下文信息发掘及更为有效的目标建模. 本文实现的分类算法的基本思路如图 4 所示.

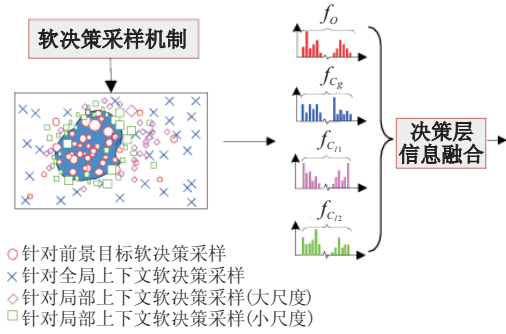


图 4 基于多尺度上下文信息的目标分类算法

3.2 决策层信息融合

逻辑回归分类器可以有效地实现决策层的信息融合,对于目标表征和分类问题,它能有效地学习和刻画各类信息在某个分类任务中的重要程度,即文中各尺度上下文信息在目标描述时的重要程度。 ℓ_1 正则化逻辑回归分类器是一种简单、行之有效的逻辑回归分类器^[9],这里利用它实现多尺度上下文信息融合。

令前端分类器响应构成一特征向量 $\mathbf{t} \in R^4$, $y \in \{-1, 1\}$ 为终端分类器的二值预测输出。在逻辑回归模型中,给定观测量 \mathbf{t} 事件 y 的发生概率为

$$P(y|\mathbf{t}) = \frac{1}{1 + \exp(-y(\boldsymbol{\omega}^T \mathbf{t} + \mathbf{v}))}, \quad (6)$$

其中 $\boldsymbol{\omega}$, \mathbf{v} 均为模型参数。当 $\boldsymbol{\omega}$ 不等于 0 时, $\boldsymbol{\omega}^T \mathbf{t} + \mathbf{v} = 0$ 表示特征空间内的分类超平面。

基于式(6), $\boldsymbol{\omega}$ 和 \mathbf{v} 可通过极大似然估计方法训练获得。设有 m 幅训练样本, $(\mathbf{t}^{(i)}, y^{(i)}) \in R^4 \times \{-1, 1\}$, $i = 1, 2, \dots, m$ 。定义逻辑损失函数

$$f(x) = \log(1 + \exp(-x)),$$

则 ℓ_1 正则化逻辑回归问题可归结为

$$\begin{aligned} & \text{minimize } l_{\text{avg}}(\boldsymbol{\omega}, \mathbf{v}) + \lambda \|\boldsymbol{\omega}\|_1 = \\ & \frac{1}{m} \sum_{i=1}^m f(\boldsymbol{\omega}^T \mathbf{t}^{(i)} y^{(i)} + \mathbf{v} y^{(i)}) + \lambda \sum_{i=1}^4 |\omega_i|. \end{aligned} \quad (7)$$

式中: $\|\cdot\|_1$ 为 1-范数, λ 为正则化参数。通过学习获得 $\boldsymbol{\omega}$, \mathbf{v} 的取值。当输入一个新的特征向量 \mathbf{t}^* 时,逻辑分类器最终输出二值化分类结果为

$$\phi(\mathbf{t}^*) = \text{sgn}(\boldsymbol{\omega}^T \mathbf{t}^* + \mathbf{v}). \quad (8)$$

在具体训练过程中,当计算出的某尺度上下文信息决策的权重为负时,将该特征对应前端分类器响应值置为 0,重新训练逻辑回归分类器参数。实际上,通过逻辑回归分类器参数的学习,实现了分类中自适应的目标空间尺度选择。

4 实验设计及结果分析

为了验证所提出的算法的有效性,本文采用了 PASCAL VOC 2008 目标分类数据集^[1]。该数据集收集了 20 类目标对象,所有的图像均取自真实场景。它一

共包含 2111 幅训练图像,2221 幅验证图像以及 4133 幅测试图像。本文使用全部训练图像训练自顶向下的类别约束机制以及基于各尺度上下文信息的前端分类器,使用全部验证图像训练后端的用于信息融合的逻辑回归分类器,并使用全部测试图像测试分类性能。

采用平均精确率 (AP) 评估每个二类分类任务 (是否为某特定目标类图像) 的性能,平均精确率可以通过计算分类的精确率/召回率曲线 (PRC) 下的面积获得。使用 VOC 2008 数据库内所有目标类的均值平均精确率 (MAP) 评价算法在该数据库上的整体性能。

下面主要通过两组实验来说明基于多尺度上下文统计特征的分类算法的性能。

实验 1 本文所提出的目标分类算法与传统的 BoF 方法性能比较。按文献[3]描述设置传统的 BoF 分类系统,图像局部信息采样使用常用的随机采样方式。为了保证对比实验的公正性,假设本文方法中各尺度上下文信息采样点数均为 n ,则 4 种尺度下累计总采样数为 $4n$,同样也设置传统 BoF 方法随机采样点数为 $4n$ 。图 5 表示了这两种方法在 VOC 2008 上分类的均值平均精确率随采样点数的增加而变化的曲线。

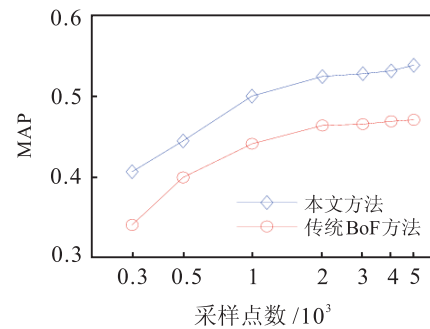


图 5 均值平均精确率随采样点数目增加而变化的曲线

从图 5 可以看到,本文提出的基于多尺度上下文统计特征的分类方法分类精确率明显好于传统的 BoF 方法。当总采样点数在 [300, 5000] 范围内变化时,分类算法在 VOC 2008 数据库上的均值平均精确率最大可以提升约 7.5 个百分点 (采样点数为 5000 时)。

实验 2 融合多尺度上下文信息的目标分类算法与基于各个单独尺度上下文信息的分类算法性能比较。在这组实验中,有 4 种基准算法与本文提出的基于多尺度上下文信息的目标分类算法进行比较,分别为:基于目标外观统计信息的分类算法,基于全局上下文统计信息的分类算法,基于大尺度局部上下文信息的分类算法,以及基于小尺度局部上下文信息的分类算法。基于单尺度上下文信息的分类算法采用与实验 1 中传统 BoF 方法相同的分类算法框架,只是将原有的随机采样方式替换为针对这种单

尺度信息的软判决采样方式. 为了更加直观地说明算法的有效性, 这里也详细地列出了传统 BoF 方法在 VOC 2008 数据库上针对各目标类具体的分类性能. 上述各算法的具体分类结果见表 1. 这里所有算法的控制采样点总数均为 4000.

表 1 各种算法在 VOC 2008 上平均分类精确率对比

目标	传统 BoF 方法	目标外观信息	全局上下文	大尺度局部上下文	小尺度局部上下文	融合多尺度信息
飞机	0.739	0.790	0.665	0.665	0.648	0.827
自行车	0.455	0.518	0.369	0.376	0.404	0.477
鸟	0.496	0.459	0.548	0.495	0.468	0.605
船	0.557	0.628	0.545	0.494	0.464	0.678
瓶子	0.200	0.195	0.185	0.227	0.239	0.253
巴士	0.447	0.517	0.412	0.412	0.451	0.554
轿车	0.549	0.577	0.498	0.504	0.528	0.578
猫	0.483	0.569	0.500	0.481	0.493	0.590
椅子	0.433	0.395	0.419	0.463	0.403	0.464
牛	0.288	0.278	0.305	0.298	0.258	0.350
餐桌	0.334	0.387	0.358	0.414	0.410	0.390
狗	0.445	0.473	0.454	0.466	0.432	0.475
马	0.626	0.595	0.632	0.606	0.598	0.634
摩托	0.579	0.630	0.563	0.571	0.569	0.632
人	0.811	0.851	0.574	0.702	0.756	0.838
盆栽	0.153	0.269	0.206	0.189	0.162	0.283
羊	0.321	0.303	0.329	0.354	0.288	0.358
沙发	0.343	0.275	0.375	0.403	0.402	0.447
火车	0.698	0.744	0.728	0.689	0.633	0.618
显示器	0.539	0.597	0.329	0.441	0.420	0.607
MAP	0.475	0.503	0.450	0.463	0.451	0.530

本文提出的算法在进行图像目标分类时综合考量了目标的多尺度上下文信息. 由表 1 可见, 该算法的分类性能整体上明显优于其他几种基准算法, 包括传统的 BoF 方法以及基于各种单尺度上下文信息的分类算法. 特别地, 通过仔细观察表 1 数据可以看出, 基于软判决采样机制获取感兴趣目标自身外观统计信息并实现分类的算法所取得的分类效果明显优于传统的 BoF 方法, 这从一定程度上说明了软判决采样机制的有效性. 这里, 软判决采样机制能够以一种鲁棒有效的方式提取图像内的前景目标信息, 减少了由于图像内各种信息相互混淆给分类带来的不确定性. 然而, 进一步对比上述基于目标外观信息的分类算法与本文方法, 有效融合多尺度上下文信息的分类算法比单纯的依赖目标外观信息的分类算法获得了更好的效果. 这说明针对目标分类任务, 场景上下文信息的引入以及适合的目标空间尺度的选择对增强感兴趣目标的刻画有着十分重要的作用. 本文提出的方法针对目标分类问题较好地实现了目标多尺度上下文信息的提取、融合与运用.

综上所述可以看出, 各尺度上下文信息对不同类的目标刻画及分类作用并不相同. 对于 VOC 2008 数据集的 20 个目标类, 根据 l_1 逻辑回归分类器中各尺度上下文信息预测值所获得的权重的大小, 在最终分类

决策时对于各个目标类起到最重要作用的目标上下文尺度如表 2 所示. 可以观察到, 单纯的目标外观信息对于分类任务并不一定始终能够起到最重要的作用, 这时全局上下文信息或局部上下文信息对于目标的认知反而具有至关重要的作用.

表 2 VOC 2008 数据集中各类目标的最有效的上下文尺度

目标类别	最有效上下文信息
{飞机、自行车、船、巴士、轿车、摩托、火车、显示器、猫、狗、人、盆栽}	目标外观
{鸟、马、牛}	全局上下文信息
{椅子、餐桌、沙发、羊}	大尺度局部上下文信息
{瓶子}	小尺度局部上下文信息

对于上下文信息作用机制, 可观察到以下几个现象: 1) 对于大部分目标类, 尤其是刚性目标(飞机、自行车、船、巴士、轿车、摩托、火车、显示器)以及具有充分学习样本的目标类(人), 目标自身外观对于分类时的有效目标刻画起到主要作用; 2) 对于某些具有规律性出现环境的目标类(鸟、马、牛), 特别是其中的某些非刚性目标类, 全局上下文是有效描述目标的重要信息; 3) 对于具有功能性定义并且类内个体外观差异巨大的目标类(椅子、餐桌、沙发), 或者弱小目标(瓶子), 局部上下文描述对于增强分类性能至关重要. 这些观察结果与前期研究者的观点也是基本一致的^[10]. 综上, 当计算机构建的模型不足以充分地描述目标自身结构及外观时, 或者目标外观由于遮挡变形等因素容易与其他事物混淆时, 上下文信息在分类识别中将会起到重要的辅助决策作用.

另外, 前面的讨论都是基于最基本的 BoF 系统框架的, 为进一步说明引入多尺度上下文信息对实现精确图像目标分类的促进作用, 这里也在更高级的 BoF 分类系统框架内进行了讨论. Soft 5 ColorSift 分类系统在 PASCAL VOC 目标分类竞赛中表现优异^[1], 它实质上是由基本的 BoF 方法衍生而来的. 基于该系统, 采用本文设计的多尺度上下文信息提取及融合算法, 在控制采样点总数为 4000 的前提下, 在 VOC 2008 数据库上获得了 0.560 的均值平均精确率. 改进后的 Soft 5 ColorSift, 原 Soft 5 ColorSift 以及 PASCAL VOC 2008 竞赛中的冠军系统的分类性能如表 3 所示. 显然, 经改进后的引入多尺度上下文信息的 Soft 5 ColorSift 系统分类性能得到了明显提高, 甚至取得了优于冠军系统的结果, 这表明本文设计的多尺度上下文信息提取与融合算法具有良好的应用潜质.

表 3 几种分类系统在 VOC 2008 上的均值平均精确率对比

	改进后的 Soft 5 ColorSift	原 Soft 5 ColorSift ^[1]	竞赛冠军系统 ^[1]
MAP	0.560	0.541	0.549

5 结 论

本文针对真实场景图像的目标分类问题, 在传统

的 BoF 统计外观模型基础上, 提出了一种基于多尺度上下文信息的目标分类算法. 与以往的分类方法相比, 本文的贡献主要在于以下 2 点: 1) 融合多尺度空间上下文信息, 包括目标自身外观信息, 有效增强了分类算法对真实场景图像中感兴趣目标的刻画能力, 提高了分类效果. 通过该融合机制解答了目标自身外观与各尺度场景信息对不同分类任务的作用机制问题, 一定程度上实现了分类的有效目标空间尺度的选择. 2) 基于软判决采样机制, 利用统计外观模型 (BoF) 框架, 实现了多尺度上下文统计特征的计算. 这种方法保留了传统 BoF 模型良好的鲁棒性, 克服了原模型混淆场景内各类信息以及无法计算局部上下文特征的弊端, 有效地实现了目标外观、各尺度上下文信息的分离与提取, 同时避免了实现困难的语义性分割或者基于二值化决策的目标定位过程. 在 VOC 2008 数据库上的实验表明, 对于真实场景图像, 本文提出的方法较传统的 BoF 以及基于单尺度上下文的分类方法可以显著提高目标分类的精度.

最后需要说明, 作为显著提高分类精确性的代价, 本文算法较之传统的 BoF 方法需要的训练时间有所增加, 这主要是由于训练时在局部信息采样过程中增加了弱监督学习机制以及相应的运算. 然而, 在测试中与传统的 BoF 方法相比, 本文算法对每幅测试图像的平均运算时间仅增加不到 3%, 在显著提高分类精确性的同时仍不失为一种高效的分类算法. 另外, 由于在目标分类中引入了场景上下文信息, 本文算法主要适用于一般性的真实场景图像, 对于某些人为设置的具有特殊场景的实验室采集图像并不能很好适用.

参考文献(References)

- [1] Everingham M, Van Gool L, Williams C K I, et al. The PASCAL visual object classes challenge 2008(VOC2008) Results[EB/OL]. (2009-03-17)[2010-06-27]. <http://www.pascalnetwork.org/challenges/VOC/voc2008/workshop/index.html>.
- [2] Csurka G, Dance C, Fan L, et al. Visual categorization with bags of keypoints[C]. Proc of ECCV Workshop on Statistical Learning in Computer Vision. Berlin: Springer Verlag, 2004: 1-22.
- [3] Zhang J, Marszalek M, Lazebnik S, et al. Local features and kernels for classification of texture and object categories: A comprehensive study[J]. Int J of Computer Vision, 2007, 73(2): 213-238.
- [4] Nowak E, Jurie F, Triggs B. Sampling strategies for bag-of-features images classification[C]. Proc of European Conf on Computer Vision(ECCV). Berlin: Springer Verlag, 2006: 490-503.
- [5] Yang L, Zheng N, Yang J, et al. A biased sampling strategy for object categorization[C]. Proc of Int Conf on Computer Vision(ICCV). Piscataway: IEEE, 2009: 1141-1148.
- [6] Viola P, Jones M. Robust real-time face detection[J]. Int J of Computer Vision, 2004, 57(2): 137-154.
- [7] Lowe D. Distinctive image features from scale-invariant keypoints[J]. Int J of Computer Vision, 2004, 60(2): 91-110.
- [8] Cortes C, Vapnik V. Support-vector networks[J]. Machine Learning, 1995, 20(3): 273-297.
- [9] Koh K, Kim S, Boyd S. An interior-point method for large-scale ℓ_1 -regularized logistic regression[J]. J of Machine Learning Research, 2007, 8: 1519-1555.
- [10] Uijlings J, Smeulders A, Scha R. What is the spatial extent of an object[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2009: 770-777.
- [11] Coleman T F, Li Y. An interior trust region approach for nonlinear minimization subject to bounds[J]. SIAM J on Optimization, 1996, 6(2): 418-445.
- [12] Drici Z, Lakshmikantham V, Walter W. Convex dependence of solution of differential equations in a Banach space relative to initial time[J]. Nonlinear Analysis, 1998, 34: 629-635.
- [13] Ralph D. Nonlinear programming advances in mathematical programming with complementarity constraints[R]. Article submitted to Royal Society, 2007.
- [14] 王健全, 陈晓峰, 杨梅强, 等. 考虑暂态稳定约束下的输电阻塞管理计算[J]. 电力系统及其自动化学报, 2007, 19(3): 66-71.
(Wang J Q, Chen X F, Yang M Q, et al. Calculation of transmission congestion management with transient stability constraints[J]. Proc of the CSU-EPSCA, 2007, 19(3): 66-71.)

(上接第1642页)