



本科毕业论文答辩：基于局部特征和上下文多类物体检测算法研究

作者：王利民 指导老师：路通 副教授

南京大学计算机科学与技术系

2011年5月30日

1 背景引言

2 算法框架

- 霍夫森林模型
- 上下文模型

3 实验结果

- 9类数据库实验结果
- LabelMe数据库实验结果



物体识别概述

- 物体识别是计算机视觉领域热点课题，现实中很多应用：Navigation for Robots, Content Based Image Retrieval etc.
- 物体识别大体分两类：Specific Object Recognition, Generic Object Recognition.
- 物体识别任务大体分三类：Image Classification, Object Detection, Image Segmentation.
- 物体识别存在很多挑战：Scale, Viewpoint, Illumination, Occlusion, Clutter etc.



Specific Object vs. Generic Object



Wild card



Carl Friedrich Gauss



Eiffel Tower



My neighbor's car

Specific objects



car



car



building



building

Generic object categories



面临挑战



occlusion



scale



deformation



clutter



illumination



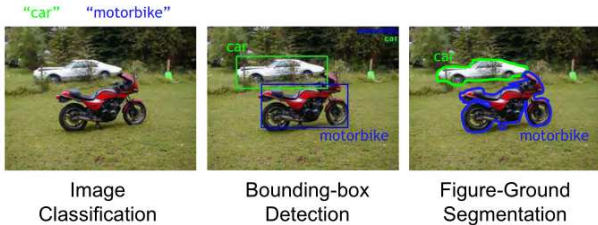
viewpoint



object pose



具体任务



- 本文工作主要是针对多类物体检测，目前主要存在两大类检测方法
 - 滑动窗口：SVM+HOG, Harris+Boosting etc.
 - 霍夫投票：ISM, Hough Forest etc.



1 背景引言

2 算法框架

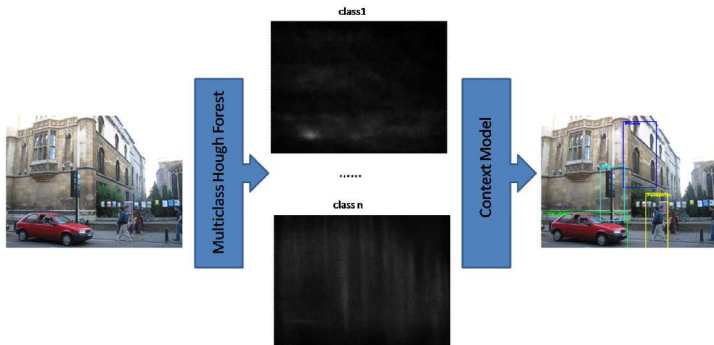
- 霍夫森林模型
- 上下文模型

3 实验结果

- 9类数据库实验结果
- LabelMe数据库实验结果



算法框架



1 背景引言

2 算法框架

- 霍夫森林模型
- 上下文模型

3 实验结果

- 9类数据库实验结果
- LabelMe数据库实验结果



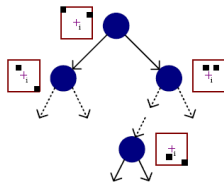
多类霍夫森林-图像Patch和叶节点

- 图像Patch: $\{\mathcal{P}_i = (\mathcal{I}_i, c_i, \mathbf{d}_i)\}$.
- 其中 \mathcal{I}_i 是图像Patch的特征, c_i 是图像Patch的类别标签, \mathbf{d}_i 是图像Patch相对物体中心的偏移量.
- 叶节点: Portion List $P_L = \{p_i\}$ and Offset Matrix $D_L = \{\mathbf{d}_{ij}\}$.
- p_i 表示第 i 类物体到达该叶子节点的图像Patch数目所占的比例.
- \mathbf{d}_{ij} 表示第 i 类物体到达该叶子节点的第 j 个Patch相对物体中心的偏移量.



多类霍夫森林-分类器

- 局部特征对应很多通道(Channel) $\mathcal{I}_i = (I_i^1, I_i^2, \dots, I_i^C)$
- 设计一个简单分类器:



$$t_{a,p,q,r,s,\tau}(\mathcal{I}) = \begin{cases} 0 & \text{if } l^a(p,q) < l^a(r,s) + \tau \\ 1 & \text{otherwise.} \end{cases}$$



多类霍夫森林-判断标准

- 首先是类别不确定性定义: $U_1(A) = |A| \times \sum_{i=0}^{|C|} p_i \log \frac{1}{p_i}$
- 其次偏移量不确定性定义: $U_2(A) = \sum_{i=1}^{|C|} \sum_{j: c_j=i} (\mathbf{d}_j - \mathbf{d}_{A_i})^2$
- 最后一个分类器性能标准:

$$Q_{multi}(t) = \alpha[U_1(A_0) + U_1(A_1)] + (1 - \alpha)[U_2(A_0) + U_2(A_1)]$$

- 其中, $A_k = \{P_i | t(I^i) = k\}$, $k \in \{0, 1\}$, α 是一个平衡两者之间的权重因子, 在实验中我们是随机选择0和1.



多类霍夫森林-物体检测

- 认为只有第 i 类物体部分的 Patch 才会对第 i 类物体中心的位置产生影响

$$\begin{aligned} p(E(\mathbf{x}), E(i)|\mathcal{I}(\mathbf{y})) &= p(\mathbf{d}(\mathbf{y}) = \mathbf{y} - \mathbf{x}, c(\mathbf{y}) = i|\mathcal{I}(\mathbf{y})) \\ &= p(\mathbf{d}(\mathbf{y}) = \mathbf{y} - \mathbf{x}|c(\mathbf{y}) = i, \mathcal{I}(\mathbf{y}))p(c(\mathbf{y}) = i|\mathcal{I}(\mathbf{y})) \end{aligned}$$

- 根据霍夫森林叶子节点记录的信息 P_L 和 D_L ，上述概率可以按如下展开：

$$p(E(\mathbf{x}), E(i)|\mathcal{I}(\mathbf{y}); \mathcal{T}) = \left[\sum_{j: \mathbf{d}_{ij} \in D_L} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|(\mathbf{y} - \mathbf{x}) - \mathbf{d}_{ij}\|^2}{2\sigma^2}\right) \right] \times \frac{p_i}{|D_L^i|}$$



1 背景引言

2 算法框架

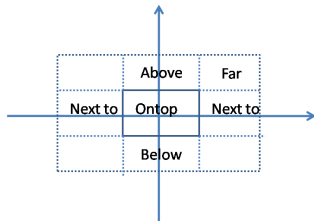
- 霍夫森林模型
- 上下文模型

3 实验结果

- 9类数据库实验结果
- LabelMe数据库实验结果



上下文关系建模



- 我们考虑上述五种位置关系，每种位置关系，我们建立如下模型：

$$p(< m, n >; \Phi^i) = \frac{1}{Z(\Phi^i)} \exp\{\phi_{mn}^i\}$$

- Φ^i 关系 R^i 的关系矩阵， ϕ_{mn}^i 为关系矩阵中的元素， $Z(\Phi^i) = \sum_{m,n} \exp\{\phi_{mn}^i\}$ ，是归一化常数.



Algorithm: The Learning Algorithm for Context.

Input: Frequency Matrix L^i .

Process:

1. Set Initial Value: $\Phi^i = \Phi_0, m_{t-1} = m_t = m_0$.
2. While ($m_{t-1} \leq m_t$) {
3. $m_{t-1} = m_t, m_t = 0$.
4. for $i = 1$ to 10 {
5. $\Phi^i = \Phi^i + \beta \nabla_{\Phi^i} \mathcal{L}(\Phi^i)$.
6. $Z(\Phi^i) = \text{Sampling Method}(\Phi^i)$.
7. $m_t = m_t + \mathcal{L}(\Phi^i)$.
8. }
9. $m_t = m_t/10$.
10. }

Output: Parameter: Φ^i .



Algorithm: Greedy search for multiclass object detection.

Input: Some candidates: $\{(\mathbf{x}_i, c_i)\}$ with appearance probability: $\{P_{app}(\mathbf{x}_i, c_i)\}$.

Process:

1. Set Initial Value: $R = \emptyset$, $P_{con}(\mathbf{x}_i, c_i) = 0$, $M = 0$.
2. $(\mathbf{x}_*, c_*) = \arg \max_{(\mathbf{x}_i, c_i) \notin R} P(\mathbf{x}_i, c_i)$.
3. While $(P(\mathbf{x}_*, c_*) > \theta)$ { % θ is the threshold
4. $R = R \cup \{(\mathbf{x}_*, c_*)\}$, $M = M + 1$.
5. For the remaining candidates {
6.
$$P_{con}(\mathbf{x}_j, c_j) = \frac{1}{M} \left[(M - 1)P_{con}(\mathbf{x}_j, c_j) + \frac{\exp(\phi_{c_j, c_*}^*)}{Z(\Phi^*)} \right]$$
7. }
8. $(\mathbf{x}_*, c_*) = \arg \max_{(\mathbf{x}_i, c_i) \notin R} P(\mathbf{x}_i, c_i)$.
9. }

Output: The detection result $R = \{(\mathbf{x}_i, c_i)\}$.



提纲

1 背景引言

2 算法框架

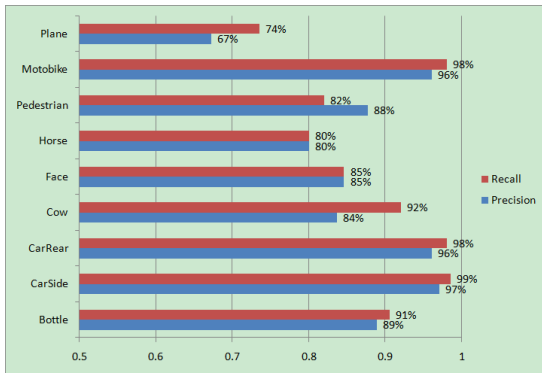
- 霍夫森林模型
- 上下文模型

3 实验结果

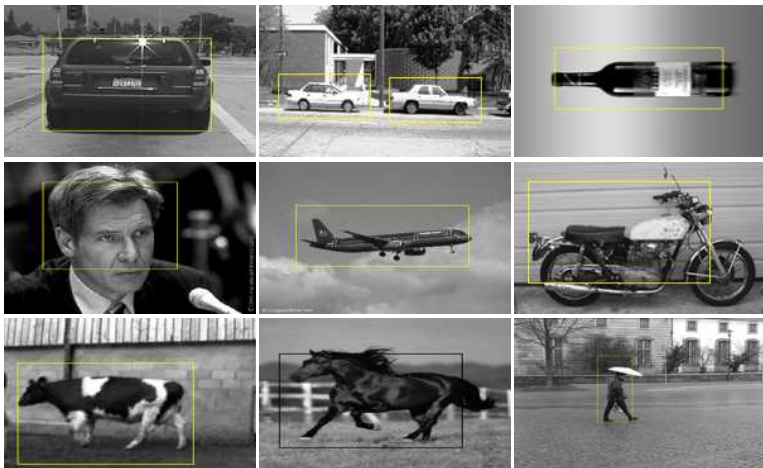
- 9类数据库实验结果
- LabelMe数据库实验结果



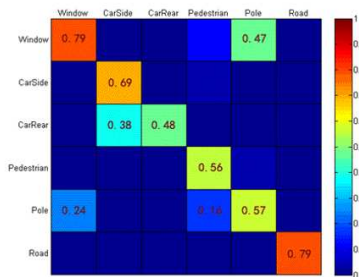
9类数据库的实验结果统计图



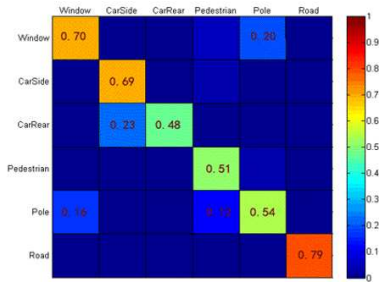
实验结果图展示



LebelMe数据库多类物体检测混合矩阵



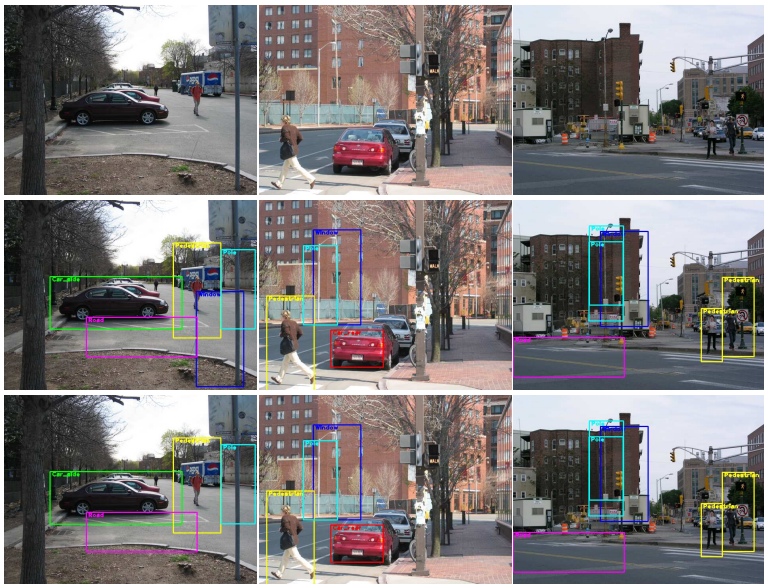
(a) 没有上下文模型



(b) 含有上下文模型







实验结果图展示1



实验结果图展示2



-  J.Gall, V.S.Lempitsky: *Class-specific hough forest for object detection*. In CVPR 2009, pp.1022-1029.
-  C.Galleguillos,A.Rabinovich,S.Belongie:*Object categorization using co-occurrence, location and appearance*. In CVPR 2008, pp.1-8.
-  C.Desai,D.Ramanan,C.Fowlkes:*Discriminative models for multiclass object layout*. in ICCV 2009, pp.229-236.
-  K.Grauman,B.Leibe:*Visual Object Recognition*. Morgan and Claypool Publishers, 2011.



Thank you!

Email: lmwang.nju@gmail.com

