

基于角度泛化的 SC 模型匹配的行人检测

胡春华¹, 钱堃²

(1. 江苏技术师范学院电气信息工程学院, 江苏 常州 213001; 2. 东南大学复杂工程系统测量与控制教育部重点实验室, 南京 210096)

摘要: 提出一种行人检测算法。该算法使用角度泛化的形状上下文描述子提取边缘采样点集的直方图分布, 以改进的 Hausdorff 距离作为模型与待测图像码本之间的匹配度量, 利用该相似度为人体中心位置投票, 并结合模板的二值掩码分割人体的前景与背景。采用颜色聚类与支持向量机形状分类器的两步验证法去除 2 类假阳结果。在自建行人图像库与 PASCAL 库上的实验结果表明, 角度泛化的形状上下文提高了对直立人体形状的局部描述性能, 两步验证明显降低了误检率。

关键词: 形状上下文; 行人检测; 码本; 支持向量机; 角度泛化

Pedestrian Detection Based on Angular-diffused Shape Context Model Matching

HU Chun-hua¹, QIAN Kun²

(1. School of Electrical and Information Engineering, Jiangsu Teacher University of Technology, Changzhou 213001, China;

2. Key Laboratory of Measurement and Control of CSE of Ministry of Education, Southeast University, Nanjing 210096, China)

【Abstract】 This paper proposes a novel pedestrian detection method. An angular-diffused Shape Context(SC) descriptor is proposed to obtain the histogram of the sampled edge map with considerations of different edge orientations. Modified Hausdorff Distance(MHD) is employed as the similarity likelihood between the codebook of the model and the test image. A voting map of the central hypotheses of pedestrians is generated, followed by the foreground and background segmentation which utilizes the binary masks of the templates. In the people candidate regions, a color-segmentation-based verification and a SVM-based shape classifier are subsequently performed to reduce two types of false positive results. Experiments on a pedestrian image database and PASCAL database illustrate the improved performance of the angular-diffused SC in representing the shape of up-right human body as well as the reduced false positive rate introduced by the two-step verification.

【Key words】 Shape Context(SC); pedestrian detection; codebook; SVM; angular-diffused

1 概述

在室外图像中检测直立行走的人体有着重要的应用价值。由于人体的非刚性结构、复杂的环境背景以及目标之间的遮挡等因素, 行人检测成为机器视觉领域中一项具有挑战性的研究课题。人体检测的主要方法大致分为自底向上与自顶向下 2 类。自底向上的方法^[1]从图像的底层特征(如边缘、颜色聚类)开始分析人的肢体, 并在人体结构约束下优化特定的代价函数以搜索人体存在区域, 该类方法不易出现误检, 但需要大量搜索计算。自顶向下的检测方法^[2]利用人体图像的基元特征训练分类器, 多采用尺度变换或者分类器级联^[2]技术实现任意尺度下的人体检测, 该类方法具有较高的计算效率, 但易出现误检。

人的轮廓形状描述是进行自顶向下检测的关键。传统的傅里叶描述子是对整体外形的全局参数化描述, 忽略了局部细节形状特征, 对于遮挡的鲁棒性较差。近年来, 形状上下文(Shape Context, SC)^[3-4]在图像检索与识别中得到了广泛的研究。它的突出优点在于: 是相对大尺度的局部描述, 参考点位置可以任意选择, 对虚假点和小范围的遮挡比全局法更为鲁棒。然而, 由于在构建形状上下文时一个像素点只能贡献于直方图中的一个扇区, 因此形状上下文对于物体的方向性扭曲等微小形变过于敏感^[4]。本文提出采用角度泛化的形状上下文人体检测法降低其对微小方向形变的敏感度而保持

对局部细节特征的描述能力。

2 行人检测算法原理

本文的行人检测算法主要包括构建人体轮廓的角度泛化形状上下文码本模型、基于该模型匹配的人体检测、快速验证去伪这三大步骤, 算法总体流程如图 1 所示。

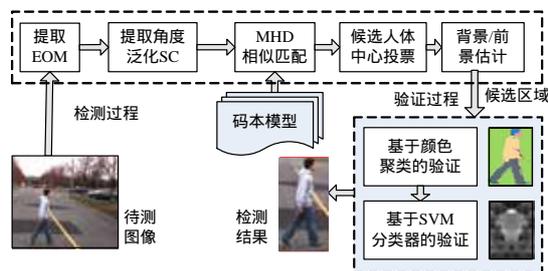


图 1 行人检测算法总体流程

首先, 在训练样本图像上同时提取形状上下文码本以及人体目标区域的二值掩码。形状上下文即在图像的边缘图(Edge Of Map, EOM)基础上, 以该边缘图上某点为中心的扇

基金项目: 江苏省高校自然科学基金资助项目(09KJD520003)

作者简介: 胡春华(1977 -), 女, 讲师、博士, 主研方向: 机器视觉, 模式识别; 钱堃, 博士研究生

收稿日期: 2010-03-22 **E-mail:** chunhh@163.com

形栅格内的统计直方图, 角度泛化即扇形区域内的任意点不仅对其所处的栅格投票, 还能以一定概率为其相邻角度的多个栅格投票。其次, 采用改进的 Hausdorff 距离 (Modified Hausdorff Distance, MHD) 为相似度量匹配模型与待检图像中边缘像素的码本, 利用匹配的边缘像素为人体中心区域投票, 并在覆盖了人体模型掩码的待检图像边缘图上估计前景与背景的概率, 阈值化后产生若干人体位置候选区域的二值掩码。最后, 对上述自顶向下的检测方法所造成的 2 类假阳错误, 依次采取基于聚类的颜色一致性验证与基于非线性支持向量机 (SVM) 的形状验证快速去除误检, 这 2 步分别利用了图像底层的颜色分割信息与连续边缘信息, 实际上是一种自底向上的验证过程。

3 偏角泛化的形状上下文码本

3.1 人体轮廓特征码本的构建

码本是一个由特征矢量构成的集合^[3]。模板图像的边缘像素数目为 L , 图像码本记为 $C = \{ce_1, ce_2, \dots, ce_L\}$, 一个像素对应一个码字。像素 i 属于某个物体, 其中心位于 (x, y) 处。每个码字 $ce_i (i=1, 2, \dots, L)$ 定义为 $ce_i = \langle u_i, r_i, m_i, w_i \rangle$, 记录了像素 i 的多个特征, 其中, u_i 是点 i 的形状上下文向量; r_i 是点 i 到中心 (x, y) 的距离; m_i 是点 i 所属的二值掩码数组; w_i 是像素 i 赋予 m_i 的权值。对各种姿态的人体样本模板图片提取的码本就作为人体模型。

3.2 边缘方向图上的偏角泛化形状上下文

某个像素点 d_i 的形状上下文定义在该像素附近的点在一个局部极坐标栅格中的统计直方图上。由于在具有相同梯度方向的边缘上进行匹配能有效减少误匹配, 因此首先从边缘图 E 中提取 EOM 并将方向离散化为 O 个方向, E 就由 O 个分量组成: $E = \{E_1, E_2, \dots, E_O\}$ 。直方图共有 $n_r \cdot n_\theta$ 个扇区 (bin)。点 d_i 的 SC 定义为 $u_{d_i} = \{h_{d_i}^1, h_{d_i}^2, \dots, h_{d_i}^O\}$, 即任意点 d_i 的形状上下文都包含在 O 个方向上的分量, 其中, h_o 定义在一个以 d_i 为圆心、 R 为半径、按对数距离等间隔建立的同心圆区域上:

$$h_{d_i}^o(k) = \#\{d_j \neq d_i : d_j \in E_o : (d_j - d_i) \in bin(k)\} \quad (1)$$

其中, $k=1, 2, \dots, n_r \cdot n_\theta$; 符号 $\#$ 表示统计数目。点 d_i 的 SC 实际上是一个二维分布, 第 k 个 bin 的径向序号与角度序号分别为 k_r 与 k_θ , 其中的值为该圆形范围内的像素统计数量, 是除点 d_i 以外该形状边缘上满足到 d_i 的径向距离为 k_r 、偏角为 k_θ 、梯度方向为 o 的点的数目总和。

在普通形状上下文的直方图计算中, 每个点都只能为一个 bin 投票, 因此, 某些形状相似而存在一定角度偏转的轮廓形状在直方图中的统计分布通常具有较大的差别。角度泛化的形状上下文使得每个点都能为多个 bin 投票, 相当于增加了一个针对角度的低通滤波器。首先设点 d_j 相对于参考点 d_i 所在的径向和偏角分别为 r_j 与 θ_j , 定义 d_j 的泛化 bin 域为 $S(d_j) = \{bin(l), l \in \{(\theta_{d_j} + i) \bmod n_\theta\}_{i=-b}^b\}$, 于是式(1)变为:

$$h_{d_i}^o(k) = \#\{d_j \neq d_i : d_j \in E_o : (d_j - d_i) \in S(d_j)\} \quad (2)$$

$$SC_{d_i}^o(k) = \sum_{d_j} g(\theta_{d_j}, k) h_{d_j}^o(k), k=1, 2, \dots, n_r \cdot n_\theta \quad (3)$$

其中, $g(\theta_{d_j}, k)$ 是以 θ_{d_j} 为对称中心、峰值为 1 的高斯密度函数。上述泛化考虑了以角度 k_θ 为中心的附近共 $2b+1$ 个 bin, 并以概率加权的方式将其融合。 $SC_{d_i}^o(k)$ 就是角度泛化后点 d_j

的形状上下文。

由于形状上下文描述中所有涉及的距离关系都是关于参考点的相对量的, 因此形状上下文具有平移不变性。尺度的不变性取决于模板的圆周半径 R , R 取为:

$$R = \frac{\alpha}{L(L-1)} \sum_{i \neq j, d_i, d_j \in E} \|d_i d_j\| \quad (4)$$

其中, α 是圆形半径的尺寸因子。参数 α 、 n_r 及 n_θ 决定了圆形区域的尺寸及扇形栅格的疏密, 其选取是一个折中问题: 圆形区域越小或扇形栅格越密, 则局部细节描述能力越强; 反之, 则全局形状描述能力越强。

4 人体检测算法

4.1 改进的 Hausdorff 距离相似匹配

对测试图像中每个边缘像素点的 SC 特征 $f = \langle x, y, v \rangle$ 与模板码本的每个码字 $ce = \langle u, r, m, w \rangle$ 进行匹配。其中, 向量 f 由点的位置 (x, y) 以及 SC 特征向量 v 组成。由于在待检图像的边缘像素上计算点 (x, y) 的 SC 容易受到背景边缘像素的干扰, 因此将附加在模板上的掩码 m 叠加于测试图像的相应位置上。叠加了掩码之后, v 的第 k 个 bin 被赋予一个权值, 其值正比于该 bin 与模板掩码重合的面积^[3]:

$$w(k) = \frac{\#\{d_i \in bin(k) \cap m\}}{\#\{d_i \in bin(k)\}} \quad (5)$$

其中, $\#\{\cdot\}$ 为统计符合条件的像素数目。用距离函数 $D_{MHD}(ce, f)$ 描述 ce 与 f 的 MHD:

$$D_{MHD}(ce, f) = \|u - w \cdot v\|_{MHD} \quad (6)$$

其中, u 是 ce 的 SC 向量; $w \cdot v$ 是 f 的加权 SC 向量。

设有待匹配点集 P 、 Q , 其 MHD 测度为:

$$\|P - Q\|_{MHD} = \max(D_{PQ}, D_{QP})$$

其中, D_{PQ} 、 D_{QP} 为对应轮廓点集间的有向匹配相似度:

$$D_{PQ} = \frac{1}{N_p} \sum_{p \in P} D(p, Q), D_{QP} = \frac{1}{N_q} \sum_{q \in Q} D(q, P) \quad (7)$$

$$D(p, Q) = \min_{q \in Q} \chi^2(\tilde{u}_p, \tilde{u}_q) = \min_{q \in Q} \frac{1}{2} \sum_{k=1}^{n_r \cdot n_\theta} \frac{(\tilde{u}_p(k) - \tilde{u}_q(k))^2}{\tilde{u}_p(k) + \tilde{u}_q(k)} \quad (8)$$

$$D(q, P) = \min_{p \in P} \chi^2(\tilde{u}_q, \tilde{u}_p) = \min_{p \in P} \frac{1}{2} \sum_{k=1}^{n_r \cdot n_\theta} \frac{(\tilde{u}_q(k) - \tilde{u}_p(k))^2}{\tilde{u}_q(k) + \tilde{u}_p(k)} \quad (9)$$

其中, $\tilde{u}_p(k)$ 是先将 $(n_r \cdot n_\theta)$ 行 O 列的矩阵 u_p 拉直、然后按照

$$\tilde{u}_p(k) = \frac{u_p(k)}{\sum_{k=1}^{n_r \cdot n_\theta} u_p(k)}$$

进行规一化之后的向量, $\tilde{u}_q(k)$ 的计算同 $\tilde{u}_p(k)$ 。将式(7)~式(9)中的 P 、 Q 分别替换为 u 和 $w \cdot v$, 即可计算出 $D_{MHD}(ce, f)$ 。

4.2 人体候选区域的投票产生

假设一个匹配上的点 d_i 的 SC 特征为 f_i , 则 f_i 与 ce_k 的相似概率为:

$$P(ce_k, d_i) \propto \exp(-\lambda_1 D_{MHD}(ce_k, f_i)) \quad (10)$$

其中, λ_1 是指数因子。将 f_i 与所有码本的码字匹配, 根据码本与模板人体中心的位置关系预测待检图像中可能的人体中心, 并对这个中心位置投票。若给定相匹配的 f_i 与 ce_k , 那么在 c 点为中心处存在一个人体 y 的概率为:

$$P(y, c | ce_k, d_i) \propto \exp(-\lambda_2 \|c + r_k - d_i\|^2) \quad (11)$$

式(10)与式(11)中的 λ_1 和 λ_2 都是距离变换为概率的系数。于是在 c 点为中心处存在一个假设人体 o 的概率按照全概率公式计算为:

$$P(y, c) = \sum_{i,k} p(y, c | ce_k, d_i) p(ce_k, d_i) p(d_i) \quad (12)$$

这个匹配过程是对 i 和 k 的双重遍历。 $P(y, c)$ 给出了一个投票结果图，具有最大概率值的点就认为是一个以此为中心的人体假设 $\{H_j\} = \{(y_j, c_j)\}$ 。

获得假设人体位置 H_j 的同时，为其投票的特征点 f_i 也被记录下来，接着验证该点是否属于人体 H_j 的前景区域。将与 f_i 匹配的码本 ce_k 的掩码 m_k 附加在测试图像以 H_j 为中心的相应位置处，作为一个假设的人体目标块。点 d_i 肯定属于前景，而图像中的任意点 p 属于前景的概率为^[3]：

$$P(p | H_j) \propto \sum_i \sum_k \exp(-\|p - l_i\|) m_k(p - l_i) p(f_i | ce_k) p(ce_k, l_i | H_j) \quad (13)$$

5 候选区域验证

上述检测仅利用了人体的轮廓特征，难以区分具有相似人形的干扰物。通过实验发现，容易出现如图 2 所示的 2 类假阳情况：(1) I 类假阳，假设区域位于真实人体区域之外，这里指重叠区域小于真实人体区域的 50%。(2) II 类假阳，假设区域包含了真实人体区域，但人体检测窗口存在尺度错误。

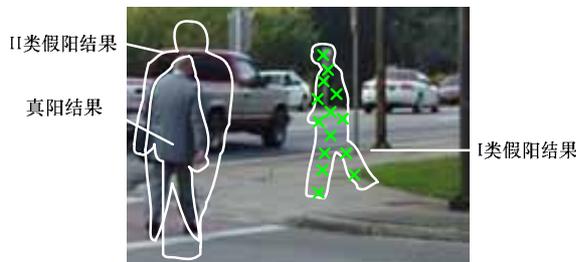


图 2 检测结果中的假阳现象

本文引入候选区域两步验证的过程以降低误检率，包括基于聚类的颜色一致性验证与基于分类器的形状验证：(1) 对于 I 类假阳，考虑到人的衣着颜色表现按照肢体结构一般具有一定连续性，因此，采用快速颜色聚类方法将图像分割为 $\{S_i\} (i=1, 2, \dots, N_{seg})$ 个聚类。设式(13)阈值化之后产生的 F 为人体候选区域(前景)， B 为非人体假设(背景)。对于聚类 S_i ，统计其中所有像素 $\{s_{i,j}\}$ 属于假设前景 F 与假设背景 B 的比

率 $\eta_i = \frac{\#\{s_{i,j} \in (S_i \cap F)\}}{\#\{s_{i,j} \in (S_i \cap B)\}}$ 对 η_i 进行阈值化就可以判断 S_i 是否属

于前景区域。(2) 对于 II 类假阳，导致结果尺度错误的原因主要是该窗口中存在一定数量的形状类似于头肩、手足等肢体轮廓的边缘像素，其并非来自真实且连续的人体轮廓。因此，在该窗口中采用自适应高斯核函数软间隔支持向量机^[5]作为二分类器，快速判决是否存在与窗口等比例尺度的人体。

6 实验结果

选用自建的行人图像数据库与通用的 PASCAL 2006 person 行人库测试本文方法的性能。本文方法采用角度泛化的形状上下文码本，其中，扇形栅格的同心圆数 $n_r = 5$ ；圆等分量 $n_\theta = 12$ ；泛化系数 $d = 1$ 。自建行人图像库采集于校园内，包括水平视角和监控视角(俯视)下的单人与多人(2 人~5 人)图像共 647 幅，尺寸不等。PASCAL 2006 person^[6]分别包含训练和测试图像 319 张和 1 153 张，分别含有人体目标 577 个和 1 153 个。实验中，正确检测率的评价标准为：检出窗口与人体目标的真实结果窗口存在 50% 以上的面积重叠。

首先在 PASCAL 2006 person 库中测试该方法，采用

PASCAL 挑战提出的测试准则，算法的性能用平均精度来评价^[6]，平均精度是精度召回曲线在 $\{0.0, 0.1, \dots, 1.0\}$ 处取值的平均，实验结果如表 1 所示。

表 1 平均精度比较

测试图像库	普通 SC	角度泛化 SC	角度泛化 SC+验证
Pascal-person 单人	0.296	0.323	0.398
Pascal-person 多人	0.160	0.169	0.174
自建库单人	0.538	0.621	0.702
自建库多人	0.356	0.463	0.598

自建行人图像库的检测结果表明，本文方法对于单人无遮挡图像一般具有较高检测率，而对于人体部分遮挡或者与环境背景融为一体难以区分时，出现了一些漏检现象。该库的整体检测结果用图 3 的 ROC 曲线表示，以描述误检率与正确检测率之间的变换关系。本文方法的曲线与横轴间所围成的面积较大，表示在所有代价比上的平均性能较优，也反映了角度泛化的 SC 描述正反 2 类样本的可分性好于普通的 SC。

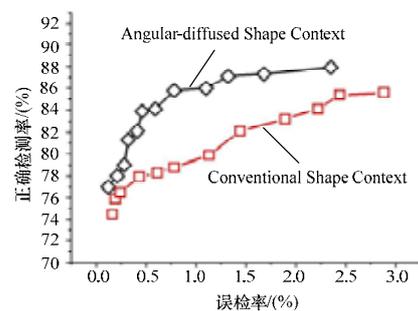


图 3 阈值为 0.6 时的检测结果 ROC 曲线

7 结束语

本文使用角度泛化的形状上下文构建人体轮廓形状特征向量，采用码本匹配法检测人体。针对 2 类假阳结果依次采用颜色聚类验证与形状分类器判决验证。对行人图像库的检测结果验证了该方法的有效性：角度泛化的形状上下文由于降低了直方图分布对形状角度微小扭曲的敏感度，因此能够得到更高的检测率。同时，两步验证的引入能够去除检测结果位置错误与尺度错误这 2 类常见的假阳，即使前景图像阈值设置较低，仍能够有效避免误检。

参考文献

- [1] 李江, 孙立军. 基于凸包裁剪的行人视频检测算法[J]. 计算机工程, 2010, 36(2): 173-175.
- [2] Viola P A, Jones M J. Rapid Object Detection Using a Boosted Cascade of Simple Features[C]//Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE Press, 2001: 511-518.
- [3] Wang Liming, Shi Jianbo, Song Gang, et al. Object Detection Combining Recognition and Segmentation[C]//Proc. of the 8th Asian Conference on Computer Vision. Tokyo, Japan: [s. n.], 2007.
- [4] Liu D, Chen T. Soft Shape Context for Iterative Closest Point Registration[C]//Proc. of 2004 International Conference on Image Processing. Singapore: [s. n.], 2004: 1081-1084.
- [5] 胡春华, 马旭东, 戴先中. 基于自适应高斯核函数支持向量机的室内人体存在检测[J]. 模式识别与人工智能, 2007, 20(4): 492-498.
- [6] PASCAL Object Recognition Database Collection[EB/OL]. (2008-11-05). <http://pascallin.ecs.soton.ac.uk/challenges/VOC/databases.html>.

编辑 张帆