

基于空间颜色特征的行人重识别方法

张 华

(武汉大学计算机学院, 湖北 武汉 430072)

摘要 行人重识别技术是嫌疑目标跨摄像头的完整活动路径分析系统的核心技术,但是在被用于实际应用时面临计算复杂度高和存储开销大等问题.针对这些问题,以视频侦查应用为背景,提出一种基于空间颜色特征的行人重识别方法,建立一种在计算复杂度和性能上较均衡的行人外貌特征描述符,对外貌中直观和重要的颜色特征采用对光照具有不变性的颜色描述符进行描述,一定程度上降低了光照变化对颜色特征稳定性的影响.为了弥补颜色直方图缺少空间信息的缺陷,并提高颜色直方图对光照和姿态变化的鲁棒性,把行人团块按语义分割成躯干和腿两部分,再把每一部分细分成若干子块,然后分别计算颜色描述符,并在计算相似度时引入位置信息.实验结果显示本方法在取得接近最好性能的同时具有低计算复杂度的优点.

关键词 视频检索; 视频监控系统; 行人重识别; 外貌描述; 相似性度量

中图分类号 TP37 **文献标志码** A **文章编号** 1671-4512(2013)S2-0209-04

Spatial color feature based pedestrian re-identification

Zhang Hua

(Computer School, Wuhan University, Wuhan 430072, China)

Abstract As one of the key technologies, pedestrian re-identification plays an important role in complete trajectory analysis of suspect objects cross multiple surveillance cameras, but it faces high computational complexity and large storage cost problems for practical application. Aiming at these problems, this paper provides a pedestrian re-identification method based on spatial color feature for video investigation application to construct appearance descriptor with a tradeoff of computational complexity and performance. The visual features of appearance are described by illumination invariant color descriptor, to a certain extent, which reduces the influence of illumination change on color stability. In order to compensate for the lack of spatial information of direct color histogram and improve appearance descriptor's robustness to illumination and pose changes, pedestrian blob is divided into two semantic parts: the torso and legs. Then, each part is subdivided into several particles and color descriptors are computed separately, and location information is introduced in similarity calculation. Experiment results show this method gets performance close to the best, and has advantages of low computational complexity.

Key words video retrieval; video surveillance system; pedestrian re-identification; appearance description; similarity measure

行人重识别是计算机视觉领域的一个较新的课题,目前还没有一个统一的、完整的理论框架,已提出的方法可以大致分为基于特征描述的方法和基于机器学习的方法.前者关注找到能够较好地描述外貌特征的方法,后者更关注通过学习的

方法找到有效的测量特征描述符间距离的模型.从实际应用的需求出发,基于特征描述的方法具有容易实现和无需训练过程的相对优势,因此,基于外貌的行人重识别技术已逐渐成为视频检索和分析方向的研究热点^[1-2].文献[3]提出的 SDALF

收稿日期 2013-07-25.

作者简介 张 华(1973-),男,讲师,E-mail: zhanghua@whu.edu.cn.

(symmetry-driven accumulation of local features)算法是基于特征描述的行人重识别方法中最具代表性的一个,也是性能最好的方法之一,被大量的研究者作为性能评价的基准.从应用角度来看,SDALF 算法存在以下问题:描述符计算复杂度高;描述符库存储开销大.

基于外貌的行人重识别技术主要关注的一个方面是如何建立有效的外貌描述,这包括从行人图像中提取特征,对特征进行描述,以及描述符间的相似性度量.基于外貌的行人重识别可以看成是行人团块图像间的匹配问题,因此可以用图像检索中的特征提取和描述技术来解决行人重识别问题.颜色是一种重要的图像特征,也是一种直观的有辨识性的外观属性,并且相对容易获取.因此,颜色是用来描述外貌特征的最常用的图像特征之一.关于这方面研究,文献[4-12]做了大量的工作,取得了一些研究成果.本文以视频侦查为背景,联合对光照具有不变性的颜色描述符和空间信息来描述行人外貌,建立一种在计算复杂度和性能上较均衡的行人外貌特征描述符.

1 基于空间颜色特征的行人重识别方法

1.1 外貌分解

首先,对行人的外貌进行分解,表示出人体的不同区域.本文假设监控视频中的行人团块图像已提取出来,团块包含处于直立行走或站立状态的行人的完整轮廓.为了把人体分成多个部件,考虑到效率因素,使用固定的参数去标识人体的三部分:头、躯干和腿,三部分的高度占团块图像高度的比率依次是:16%,29%和55%.本文方法忽略了头部信息和腿的下半部分信息,因为这两部分的有用信息量较少,却含有较多的背景信息.躯干和腿两部分也各自取中间区域提取特征,例如躯干部分的宽度为团块宽度的80%,腿部宽度为团块宽度的70%.

然后,对身体各部分进行细分来获得多个实例.本文方法没有采取随机密集提取子块的方法,因为这会带来高计算复杂度和大存储开销.如图1所示,本文方法进一步把躯干和腿两部分平均分成 n 个水平条带(子块).

设 $T = \{T_1, T_2, \dots, T_N\}$ 表示行人库,其中每一个对应一个行人.如图1所示,每一个行人 T_i 用一定数量的组的序列表示,与图像被细分的部件对应.即

$$T_i = \{T_{i,j}, T_{i,j+1}, \dots, T_{i,M}\}, \quad (1)$$

式中 M 为外貌分解后的部件数量,本文设置 $M=2$.

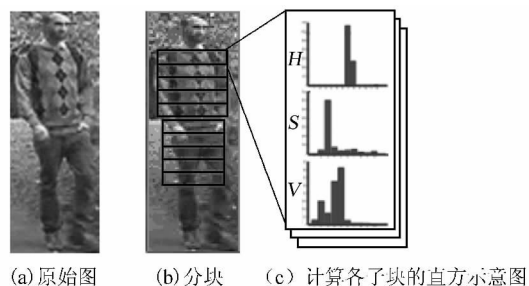


图1 基于空间颜色特征的行人外貌表示方法

1.2 特征提取和描述

对于人体的每一个子块,用 HSV 颜色直方图(分别计算 H , S 和 V 直方图)进行描述. HSV 颜色空间符合人对颜色的主观感觉,且对光照变化有一定的不变特性^[5-6].

因此,基于空间颜色特征的行人外貌描述符可以表示为

$$T_i = \{ \{ H_{\text{hsv } i, \text{torso}}^1, H_{\text{hsv } i, \text{torso}}^2, \dots, H_{\text{hsv } i, \text{torso}}^n \}, \{ H_{\text{hsv } i, \text{legs}}^1, H_{\text{hsv } i, \text{legs}}^2, \dots, H_{\text{hsv } i, \text{legs}}^n \} \}, \quad (2)$$

式中: H_{hsv} 是 HSV 颜色直方图; n 是身体各部分细分后的子块数量,本文设置 $n=5$.

1.3 相似度计算

直方图之间的相似度计算采用计算复杂度稍高,但较精确的 Bhattacharyya 距离(巴氏距离):

$$d_b(H_1, H_2) = \sqrt{1 - \sum_i \frac{\sqrt{H_1(i) \cdot H_2(i)}}{\sum_i H_1(i) \cdot \sum_i H_2(i)}}, \quad (3)$$

式中 H_i 为直方图.那么,子块间的相似度计算方法为

$$d_{\text{patch}}(H_{\text{hsv}1}, H_{\text{hsv}2}) = \alpha d_b(h_1, h_2) + \beta d_b(s_1, s_2) + \gamma d_b(v_1, v_2), \quad (4)$$

式中: H_{hsv} 为 HSV 颜色直方图; h , s 和 v 分别为 H , S 和 V 三个分量的直方图; α , β 和 γ 为权重值,本文设置三个系数的值分别为 0.5, 0.3 和 0.2.

为了联合空间信息,本文方法在计算行人各部分(躯干和腿)外貌描述符间的相似度时采取按位置计算对应子块间的相似度,并取平均值.

给定一个查询目标 Q ,计算行人整个外貌间的相似度时,分别计算身体各部分外貌间的相似度,然后取平均值,计算式为

$$D(T_i, Q) = f(d(T_{i,1}, Q_1), d(T_{i,2}, Q_2), \dots, d(T_{i,M}, Q_M)), \quad (5)$$

式中 $f(\cdot)$ 为求平均值的函数.

2 实验

本实验把 SDALF 方法作为参照方法,在相同的实验环境和材料下进行实验,通过对比实验结果来对本文提出方法的性能和计算复杂度进行分析,以检验该方法在性能和效率上的平衡性.为了检验本文方法对光照和视角变化的鲁棒性,实验使用光照和视角变化比较大的公开的标准行人数据集进行测试.另外,使用评价识别系统性能的标准方法对本文方法的性能进行评估.

2.1 数据集

本实验使用文献[13]介绍的一个用于评测视角不变行人识别方法性能的数据集 VIPeR(view-point invariant pedestrian recognition).该数据集包括 632 个不同行人的 1264 张图像,每人有两张采自不同光照条件下任意视角的图像,图像的大小都被规格化为 48×128 .因为图像中的光照和视角的变化很大(如图 2 所示,同一列为同一人的不同视角的图像),大多数都有 90° 的视角变化,所以分辨率相对较低,且用于建立行人外貌描述的信息有限,VIPeR 被认为是进行行人重识别最具挑战的数据集.



图 2 行人不同视角的示意图

2.2 实验方法

在一次测试中,首先从 VIPeR 中随机选取 316 个人,每人有两张图像,并把行人分成两组,每组都有 316 个人,每人一张图像,一组作为候选行人库,另一组作为要重识别的行人,即查询库.然后,把查询库中的每个人作为重识别目标在候选行人库中进行查找和匹配,并按相似度进行排序.最后根据相似度排序结果计算累加匹配特征(cumulative matching characteristic,CMC)曲线.把这样的测试做 10 次后计算 CMC 曲线的平均值.

本文把行人重识别看成是按相似度排序的问题,因此采用 CMC 曲线对行人重识别的性能进行评价.CMC 曲线是被广泛使用的重识别排名准确度评测指标,表示在前 K 个结果中找到正确匹配的期望,其计算方法可以用下式表示:

$$C_{CMC} =$$

排名结果的前 K 个行人中的正确数 / K . (6)

2.3 实验结果与分析

本文在一部 Lenovo ThinkPad 笔记本电脑(Intel Core 2 SP9400 2.4 GHz,4 GB,Intel X18-M 160 G2)上进行上述实验计算复杂度实验结果为:SDALF 方法计算单个行人外貌描述符所需时间为 13 s,本文方法(SC)只需 10.1 ms.

性能测试结果如图 3 所示,其中 RCH 是区域颜色直方图方法.RCH 方法不对人体的躯干和

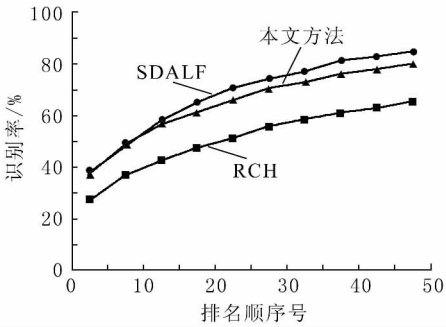


图 3 在 VIPeR 数据集上的 CMC 曲线

腿两个区域进行细分,而是直接从两部分提取颜色并用 HSV 颜色直方图进行描述.本文方法的识别率比 RCH 提高了约 25%,说明对躯干和腿两部分进行细分增强了描述符的特征描述能力和可分辨性,也增强了对视角变化的鲁棒性.在排名前 10 的结果中本文方法和 SDALF 的识别率相当,返回结果数增加后,本文方法的识别率略低于 SDALF,这是由于 SDALF 使用 MSCR 和 RHSP 特征提高了行人外貌描述符的可分辨力.另外,本文方法与目前有代表性的几种重识别方法在性能上的比较,如表 1 所示.其中,ELF 是文献[6]提出的一种视角不变条件下基于局部特征累积的行人识别方法,从行人的每一幅图像提取 8 种颜色特征和 19 种纹理特征.ERSVM 是文献[14]提出的方法,使用了与 ELF 相同的特征,通过 RankSVM 来训练模型参数,得到行人图像的距离度量模型.对于排名为 1 的正确识别率,本文方法要好于 ELF 和 ERSVM,说明联合空间信息和颜色直方图建立的行人外貌描述符的分辨力较好.可见,本文方法与几种主要的重识别方法的最好性能比

表 1 与几种主要的重识别方法在 VIPeR 数据集上的性能比较

排名	ELF	SDALF	ERSVM	本文方法
1	12	20	13	19
10	43	50	50	49
25	66	70	71	67
50	81	85	85	80

较接近。

实验显示本文提出的方法在识别准确度略低于 SDALF, 但计算复杂度明显降低, 在性能和效率间达到平衡。

在上述测试中, 本文方法的人体各部分的子块数量 $n=5$ 。为了检验子块数量对性能的影响程度, 把 n 设置为不同的值(2, 3, 5, 10, 15 和 20), 分别进行测试, 结果如图 4 所示。数据显示, n 值从 2 逐渐增加到 20, 性能逐渐由小幅提升转为小幅下降, 计算复杂度逐渐增加(完成一次完整测试的时间从 400 s 增加到 670 s)。 $n=5$ 时的总体性能最好, $n=20$ 时的性能最差, 显示过度细分会导致性能下降, 这是因为太小的子块包含的信息量不足。

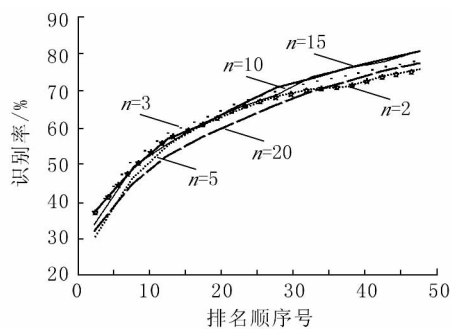


图 4 不同子块数量下的重识别性能

基于本文方法开发了实验系统——基于对象的检索系统 XObjectFinder, 图 5 是在 VIPeR 数据集中进行行人重识别实验的几次检索结果。如图 5 的(a), (b)和(c)所示, 同一行人的两张图像



图 5 行人重识别实验结果示例

中视角有变化, 光照也有变化, 正确结果的排名依然靠前。实验结果显示本文方法对光照和姿态变化具有鲁棒性。

为了建立兼顾性能和效率的对光照和姿态变化具有鲁棒性且有分辨力的基于颜色的行人外貌描述, 现有的方法一般采用颜色直方图作为外貌描述符, 但是由于在光照和姿态变化时颜色变化较大, 而且直方图没有保留空间信息, 从而导致基于颜色的行人外貌描述的性能较低。为此, 本文提出了基于空间颜色特征的行人外貌描述方法, 实验结果显示本方法在取得接近最好性能的同时具有低计算复杂度的优点。在未来的研究中将对如何从图像序列中提取有代表性的图像, 以及融合多张图像的空间颜色特征的方法进行研究。

参 考 文 献

[1] Doretto G, Sebastian T, Tu P, et al. Appearance-based person reidentification in camera networks: problem overview and current approaches[J]. Journal of Ambient Intelligence and Humanized Computing, 2011, 2(2): 127-151.
[2] Hu W, Xie N, Li L, et al. A survey o visual content-

based video indexing and retrieval[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2011, 41(6): 797-819.
[3] Farenzena M, Bazzani L, Perina A, et al. Person reidentification by symmetry-driven accumulation of local features[C]// Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco: IEEE, 2010: 2360-2367.
[4] Swain M J, Ballard D H. Indexing via colour histograms[C] // Active Perception and Robot Vision. Berlin: Springer, 1992: 261-273.
[5] van de Sande K E A, Gevers T, Snoek C G M. Evaluating color descriptors for object and scene recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1582-1596.
[6] Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features[C]// Computer Vision-ECCV 2008. Berlin: Springer Berlin Heidelberg, 2008: 262-275.
[7] Wang X, Doretto G, Sebastian T, et al. Shape and appearance context modeling[C]// IEEE 11th International Conference on Computer Vision. Rio de Janeiro: IEEE, 2007: 1-8. (下转第 217 页)

EMM 模型的 VANET 仿真体系架构,并给出了完整仿真流程和具体实现方法,实验结果验证了所提出的 VANET 仿真体系架构的可行性及 EMM 扩展移动模型的有效性.

参 考 文 献

- [1] Zhu J, Roy S. MAC for dedicated short range communications in intelligent transport system[J]. IEEE Commun, 2003, 41(12): 60-67.
- [2] Blum J J, Eskandarian A, Hoffman L J. Challenges of intervehicle ad hoc networks[J]. IEEE Trans on Intelligent Transportation System, 2004, 5 (4): 347-351.
- [3] Haerri J, Filali F, Bonnet C. Mobility models for vehicular ad hoc networks: a survey and taxonomy [J]. IEEE Communications Surveys and Tutorials, 2009, 11(4): 19-41.
- [4] Camp T, Boleng J, Davies V. A survey of mobility models for ad hoc network research [J]. Wireless Communications and Mobile Computing, 2002, 2(5): 483-502.
-
- (上接第 212 页)
- [8] Metternich M J, Worring M, Smeulders A. Color based tracing in real-life surveillance data[J]. Transaction on data hiding and multimedia security, 2010, 6010: 18-33.
- [9] Lin Z, Davis L S. Learning pairwise dissimilarity profiles for appearance recognition in visual surveillance [C]//Advances in Visual Computing. Berlin: Springer Berlin Heidelberg, 2008: 23-34.
- [10] Yu Y, Harwood K, Yoon K, et al. Human appearance modeling for matching across video sequences [J]. Machine Vision and Applications, 2007, 18(3): 139-149.
- [11] Gheissari N, Sebastian T B, Hartley R. Person re-identification using spatiotemporal appearance[C]// Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2006: 1528-1535.
- [12] Huang C, Wu Y, Shih M. Unsupervised pedestrian re-identification for loitering detection[C]//Proceedings of Advance in Image and Video. Berlin: Springer Berlin Heidelberg, 2008: 771-783.
- [13] Gray D, Brennan S, Tao H. Evaluating appearance models for recognition, reacquisition, and tracking [C]//Proceedings of 2007 IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. Rio de Janeiro: IEEE, 2007: 1-7.
- [14] Prosser B, Zheng W S, Gong S, et al. Person re-identification by support vector ranking[C]// British Machine Vision Conference. Aberystwyth: BMVC, 2010: 6-16.