

颜色特征和超像素特征融合的人体目标再识别

宋亚玲 张 良

(中国民航大学天津市智能信号与图像处理重点实验室, 天津 300300)

摘 要: 在非重叠视域的多摄像机监控系统中, 人体目标再识别有着重要的应用。针对再识别过程中面临的光照变化、视角变化、姿态变化、遮挡等问题, 提出了融合全局颜色特征和超像素特征的方法, 对颜色特征和超像素特征分配不同的权重, 进行人体目标间的相似性度量。超像素特征是将前景图像分割成多个超像素, 采用密集采样 SIFT 特征结合单词包(Bag-of-Words) 框架对每个超像素进行描述。将得到的超像素特征和全局颜色特征结合建立人体目标模型, 分别使用 EMD(Earth Mover's Distance) 距离和巴氏距离度量目标间的相似性。对多个数据库进行实验, 结果证明, 该算法具有较高的识别率。

关键词: 非重叠多摄像机; 人体目标再识别; 颜色特征; 超像素特征

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1003-0530(2015) 10-1378-05

Person Re-identification Based on Color Features and Superpixels Features

SONG Ya-ling ZHANG Liang

(Tianjin Key Lab for Advanced Signal and Image Processing of Civil Aviation University of China , Tianjin 300300 , China)

Abstract: Person re-identification is one of the key issues in a non-overlapping multi-camera surveillance system. The method of person re-identification must deal with several challenges such as variations of illumination conditions , poses and occlusions. To seek for more robust features , the unsupervised training method that combining the global color features with the superpixels features was proposed. Specifically , the color feature and the superpixels feature were assigned different weighted values. To obtain the superpixels feature of a human target , the foreground picture of a person should be divided into different patches using the superpixels segmentation's method. Then , dense SIFT features and Bag-of-Words model were applied to describe superpixels. At last , superpixels features and global color features were combined to represent a person , and EMD (Earth Mover's Distance) distance and Bhattacharyya distance were used to determine the similarity between the targets. Extensive experiments results show that the proposed method has a higher accuracy rate.

Key words: non-overlapping multi-camera; person re-identification; color features; superpixels features

1 引言

近年来 随着非重叠视域多摄像机监控系统的广泛应用 人体目标再识别得到了越来越多的重视。所谓人体目标再识别就是在分布式多摄像头监控系统中 识别出在不同的位置、不同时间拍摄的人体目标是否为同一个人^[1]。由于摄像机自身参数、光照、视角、姿态变化、遮挡等因素的影响 使得不同摄像机下

不同的人体目标看起来相似 而相同的目标看起来不同 这给再识别问题带来了严重的挑战。

目前关于再识别的研究方向主要分为两类, 一类是基于特征的方法, 一类是基于测度学习的方法^[2]。特征类方法的基本出发点是设计出具有较强的区分性和稳定性的特征, 主要使用的特征有颜色直方图、纹理、形状、梯度等。为了更好地应对再识别的诸多挑战 特征的空间信息也被作为再识别

的关键因素。SDALF(Symmetry-Driven Accumulation of Local Features)^[1]方法首先将人体分割成头部、躯干和腿部三部分,再利用人体的对称性分别将躯干和腿部分割,根据对称轴提取加权 HSV 颜色直方图、MSCR(Maximally Stable Color Regions)特征和 RHSP(Recurrent High-Structured Patches)特征对目标进行描述。文献[2]用水平条将人体分为 6 部分,并通过学习给每个部分的特征分配不同的权重。测度类方法的关注点是如何度量人体目标的相似性。通过建立正负样本训练,学习出不同摄像机下人体目标的距离函数,使同一个目标之间距离小、相似性高,不同目标间距离大、相似性小。用于再识别的距离度量学习方法有 LMNN(Large Margin Nearest Neighbour)^[3]、采用相对距离比较的 PRDC(Re-identification by Relative Distance Comparison)^[4]方法等。

通常来说,基于测度学习的方法对特征的设计要求较低,能取得较高的识别率。但是,测度学习方法需要对训练样本进行人工标注,时间花销大,复杂度高,在训练样本过少的情况下存在过度拟合现象。本文主要从特征的设计出发,构造鲁棒性高的特征描述,提出融合整体颜色特征和超像素特征的人体目标再识别方法。首先提取人体目标的颜色直方图特征,使用巴氏距离对目标进行相似性度量,之后使用超像素分割方法将人体分割成多个超像素,采用密集采样 SIFT 特征结合单词包框架对每个超像素进行描述,建立人体目标外观模型,使用 EMD 距离度量目标间的相似性,最后将两种特征融合确定目标间的相似性。

2 特征设计

本文人体目标再识别研究过程如下:

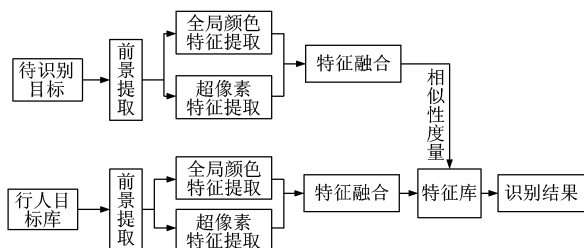


图 1 人体目标再识别流程图

Fig. 1 Flow char of person re-identification

2.1 全局颜色特征

在人体目标匹配过程中,背景的存在会严重影响

识别效果。因此,在提取特征之前,使用 STEL 成分分析方法^[5]将背景减除,在前景目标上提取特征。

颜色特征是图像的重要属性,而颜色直方图是用来描述目标外观的常用特征。在 HSV 空间中,色调(hue)代表颜色的种类,具有光照不变性和谱不变性。但是,饱和度(saturation)影响着色调,在饱和度较小时,色调值对灰度的表征不稳定。为了解决这一问题, Van de Weijer 在色调的稳定度与饱和度呈反比的前提下,引入误差分析,使用饱和度加权色调直方图^[6]。色调和饱和度从对立颜色空间得到:

$$\text{hue} = \arctan\left(\frac{O_1}{O_2}\right) = \arctan\left(\frac{\sqrt{3}(R-G)}{R+G-2B}\right) \quad (1)$$

$$\text{saturation} = \sqrt{O_1^2 + O_2^2} = \sqrt{\frac{2}{3}(R^2 + G^2 + B^2 - RG - RB - GB)} \quad (2)$$

O_1 和 O_2 是对立颜色空间的两个组成成分:

$$O_1 = \frac{1}{\sqrt{2}}(R - G) \quad (3)$$

$$O_2 = \frac{1}{\sqrt{6}}(R + G - 2B) \quad (4)$$

本文中加权后的色调直方图量化为 36 维的特征向量。

2.2 超像素特征

2.2.1 超像素分割

所谓超像素,是指具有相似纹理、颜色、亮度等特征的相邻像素构成的图像块,它利用像素之间特征的相似程度将像素分组,可以获取图像的冗余信息,在很大程度上降低了后续图像处理任务的复杂度^[7]。本文采用 SLIC^[8](simple liner iterative clustering)算法生成超像素,该方法在 Lab 颜色空间和 XY 空间操作,利用像素间距离大小来聚类相似的像素,可以生成如图 2 所示的紧凑、近似均匀的超像素,同时可以控制超像素个数和紧凑度。



图 2 不同超像素个数的分割图

Fig. 2 Segmentations with different number of superpixels

2.2.2 超像素描述

SIFT 特征描述图像的局部特征,对尺度变化、旋转保持不变性,对噪声、视角变化、仿射变换同样具有良好的稳定性^[9]。目前,SIFT 特征已经广泛应用于图像处理、目标识别等领域中。由于本文将图片分成了多个超像素,区域大小有限,为了保证每个区域都有 SIFT 特征点,我们放弃关键点检测,采用密集采样 SIFT(DSIFT) 特征,对前景中的每个像素点,按照统一方向、统一尺度进行 SIFT 描述,并与单词包(Bag-of-Words) 框架相结合,来表示超像素的特征。具体步骤如图 3 所示,首先从训练数据集中提取 DSIFT 特征,使用 K-means 聚类方法创建单词包特征所需要的词典;将每个像素对应的 SIFT 特征描述子映射到词典上,得到该描述子所对应的单词;统计超像素中每个单词出现的频率。词典的容量大小是影响识别率的关键因素,容量越大表征的细节越丰富,同时超像素直方图越稀疏,多个超像素直方图特征构成了一个人体目标的超像素特征描述。

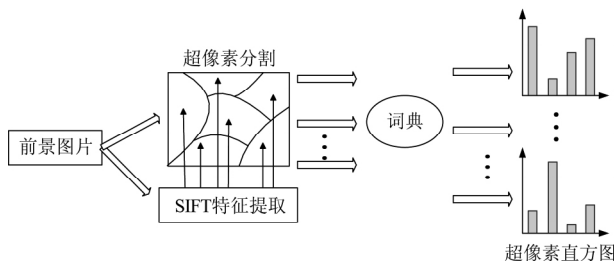


图 3 超像素特征描述过程

Fig. 3 The process of describing superpixels features

2.3 相似性度量

巴氏距离是测量颜色直方图相似性最常用的方法,运算速度较快。采用巴氏距离度量目标之间颜色特征的相似性,则目标 A 与目标 B 之间的颜色距离为:

$$d_{\text{color}}(A, B) = \sqrt{1 - \frac{\sum_i I_A(i) \cdot I_B(i)}{\sum_i I_A(i) \cdot \sum_i I_B(i)}} \quad (5)$$

其中 $I_A(i)$ 、 $I_B(i)$ 分别表示目标 A 与目标 B 的颜色特征向量,距离越小,相似性越高。

在不同时间、不同地点获取的同一个人目标往往会存在姿态变化,造成超像素直方图特征的漂移。如果采用简单的巴氏距离,仅将两个目标对应标号的超像素直方图相比较,必然会增加误配率,

而 EMD 交叉距离则会避免这种情况。EMD 距离用于度量集合的相似性,实现多对多的匹配。一个人体目标的超像素特征 I 表示为多个特征集的集合,即 $I = ((\alpha_1, \omega_{\alpha 1}), (\alpha_2, \omega_{\alpha 2}), \dots, (\alpha_m, \omega_{\alpha m}))$, 其中 α_i 表示一个超像素直方图向量, $\omega_{\alpha i}$ 表示向量 α_i 的权重。则目标 A 的超像素特征 $I_A = ((a_1, \omega_{a 1}), (a_2, \omega_{a 2}), \dots, (a_m, \omega_{a m}))$ 和目标 B 的超像素特征 $I_B = ((b_1, \omega_{b 1}), (b_2, \omega_{b 2}), \dots, (b_n, \omega_{b n}))$ 之间的 EMD 距离可以定义为:

$$d_{\text{EMD}}(A, B) = \min_{f_{ij}} \frac{\sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij}}{\sum_{i=1}^m \sum_{j=1}^n f_{ij}} \quad i = 1 \dots m; j = 1 \dots n \quad (6)$$

d_{ij} 为向量 a_i 与向量 b_j 的欧式距离, f_{ij} 为运输流,且满足以下约束条件:

$$f_{ij} \geq 0 \quad i = 1 \dots m; j = 1 \dots n$$

$$\sum_{j=1}^n f_{ij} \leq \omega_{a_i} \quad i = 1 \dots m$$

$$\sum_{i=1}^m f_{ij} \leq \omega_{b_j} \quad j = 1 \dots n$$

$$\sum_{i=1}^m \sum_{j=1}^n f_{ij} = \min(\sum_{i=1}^m \omega_{a_i}, \sum_{j=1}^n \omega_{b_j}) \quad (7)$$

将颜色特征和超像素特征匹配结果融合,度量目标间的相似性,则目标 A 与目标 B 之间的距离为:

$$d(A, B) = \gamma d_{\text{color}}(A, B) + \mu d_{\text{EMD}}(A, B) \quad (8)$$

γ 、 μ 分别为颜色特征和超像素特征的权值,经实验分析,将权重设定为固定值: $\gamma = 0.35$ $\mu = 0.65$ 。

3 实验结果及分析

一般来说,人体目标再识别会有两组数据集:查询集和候选目标集。本文在 VIPeR 数据库和 PRID450s 数据库进行实验,这两个数据库是目标再识别的常用数据库。将人体目标再识别问题看作是相似度排序问题,使用累积匹配特性(Cumulative Matching Characteristic, CMC) 曲线对再识别性能进行评估,CMC 曲线中的第 r 个排名代表在候选目标集中搜索待查询目标时,前 r 个搜索结果中能找到的待查询人的比率^[10]。

3.1 在 VIPeR 数据库的测试结果及分析

VIPeR 数据库由两组不同摄像机获取的 632 个人体目标组成。每个人体目标包含两张不同光照条

件下获取的图像,而且大部分图像有 90 度的视角变化,图像大小归一化为 48×128 。从 632 对人体目标中随机抽取 316 对作为训练集用于生成 Bag-of-Words 中的词典,剩下的 316 对作为测试集,测试集中的一组图像作为查询集,另一组图像作为候选目标集进行试验。在提取超像素特征时,将人体目标分成了 40 块。这个过程重复 10 次,得到平均 CMC 曲线。

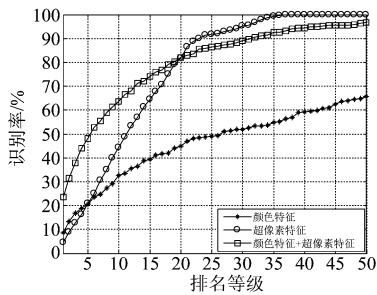


图 4 本文算法在 VIPeR 数据库上实验结果
Fig. 4 The experimental results of VIPeR dataset

表 1 本文算法在 VIPeR 数据库上与其他算法比较 (%)
Tab. 1 The compared results of VIPeR dataset between our method and other algorithms (%)

方法	第 1 匹 配率	第 10 匹 配率	第 20 匹 配率	第 30 匹 配率	第 50 匹 配率	训练 时间
SDALF ^[3]	20	50	65	75	85	-
PRDC ^[4]	16	54	70	78	87	15 min
ELF ^[11]	12	43	60	70	81	5 h
本文方法	23.4	64.5	82	88.9	94	6 min

由图 4 看出,将颜色特征与超像素特征结合,可以显著提高识别率。表 1 显示了本文算法与特征类方法 SDALF^[3]、测度类方法 PRDC^[4]和 ELF(Ensemble of Localized Features)^[11]测试结果对比。在第 1 匹配率上,特征类方法 SDALF 识别率为 20%,优于 PRDC 算法和 ELF 算法,而本文算法第一匹配率为 23.4%,第 30(候选目标集的前 10%)匹配率达到了 88.9%,远高于其他三类算法。

3.2 在 PRID450s 数据库的测试结果及分析

PRID450s 数据库由两组不同摄像机获取的 450 个人体目标组成。同一个人体目标包含两张光照强烈变化的图像,图像大小归一化为 64×128 。采用与 VIPeR 数据库相似的实验方法,从 450 对人体目标中随机选取 250 对用作训练生成词典,剩下的 200 对用作测试,测试集中的一组图像作为查询集,

另一组图像作为候选目标集进行试验。在提取超像素特征时,将人体目标分成了 36 块。这个过程重复 10 次,得到平均 CMC 曲线。

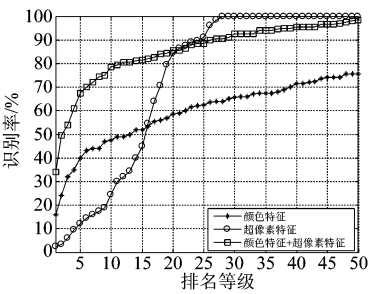


图 5 本文算法在 PRID450s 数据库上实验结果
Fig. 5 The experimental results of PRID450s dataset

图 5 显示了在 PRID450s 数据库上分别取颜色特征、超像素特征和特征结合三种情况下的识别结果。在仅使用颜色特征或者超像素特征的情况下,排名前 15 的识别率都比较低,但将两种特征融合后,识别率改善明显,第 1 匹配率达到了 34%、第 15 匹配率达到了 80%。

表 2 本文算法在 PRID450s 数据库上与其他算法比较 (%)
Tab. 2 The compared results of PRID450s dataset between our method and other algorithms (%)

方法	第 1 匹 配率	第 10 匹 配率	第 20 匹 配率	第 50 匹 配率
LMNN ^[3]	29	68	78	90
EIML ^[12]	35	68	77	87
KISSME ^[13]	33	71	79	90
本文方法	34	78.5	85.5	98.5

表 2 将本文算法与测度类方法 LMNN^[3]、EIML(Efficient Imposter-based Metric Learning)^[11]、KISSME(Keep It Simple and Straightforward MEtric)^[12]进行了比较。除了在第一匹配率上,本文算法比 EIML 方法稍低,在其他各个匹配率上,本文算法均取得了最高的识别率。

4 结论

本文提出了一种将颜色特征与超像素特征结合的人体目标再识别算法,经验证,将人体进行分块描述能增强对光照变化、视角变化、姿态变化的鲁棒性,取得较高的识别率,并且本文算法无需监督式训练,能取得与测度类方法近似的性能,大大

降低了复杂度。下一步研究主要是改善超像素特征 增强对遮挡问题的鲁棒性。

参考文献

- [1] Farenzena M , Bazzani L , Perina A , et al. Person re-identification by symmetry-driven accumulation of local features [C] // Computer Vision and Pattern Recognition. San Francisco: IEEE , 2010: 2360-2367.
- [2] Liu Chunxiao , Gong Shaogang , Loy Change Chen , et al. Person re-identification: what features are important? [C] // European Conference on Computer Vision. Heidelberg: Springer-Verlag Berlin , 2012: 391-401.
- [3] Weinberger K Q , John B , Saul L K. Distance metric learning for large margin nearest neighbor classification [J]. Journal of Machine Learning Research , 2009 , 10: 207-244.
- [4] Zheng Weishi , Gong Shaogang , Tao Xiang. Re-identification by relative distance comparison [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2013 , 35(3) : 653-668.
- [5] Jojic N , Perina A , Cristani M , et al. Stel component analysis: Modeling spatial correlations in image class structure [C] // Computer Vision and Pattern Recognition. Miami , FL: IEEE Conference , 2009: 2044-2051.
- [6] K. E. A. van de Sande , Gevers T , Cees G. M. Snoek. Evaluating color descriptors for object and scene recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2010 , 32(9) : 1582-1596.
- [7] 王春瑶 , 陈俊周 , 李伟. 超像素分割算法研究综述 [J]. 计算机应用研究 , 2014 , 31(1) : 6-12.
Wang Chunyao , Chen Junzhou , Li Wei. Review on superpixel segmentation algorithms [J]. Application Research of Computers 2014 31(1) : 6-12. (in Chinese)
- [8] Achata R , Shaji A , Smith K , et al. SLIC superpixels compared to state-of art superpixel methods [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2012 , 34(11) : 2274-2282.
- [9] 李丽丽 孙劲光. SIFT 特征区域 Contourlet 域遥感图像水印算法 [J]. 信号处理 2012 28(9) : 1254-1261.
Li Lili , Sun Jinguang. A Watermarking Algorithm for Remote Sensing Image Based on Contourlet Domain and SIFT Feature Region [J]. Signal Processing , 2012 , 28(9) : 1254-1261. (in Chinese)
- [10] 杜宇宁 , 艾海舟. 基于统计推断的行人再识别算法 [J]. 电子与信息学报 2014 , 36(7) : 1612-1618.
Du Yuning , Ai Haizhou. A Statistical Inference Approach for Person Re-identification [J]. Journal of Electronics & Information Technology , 2014 , 36(7) : 1612-1618. (in Chinese)
- [11] Gray D , Tao Hai. Viewpoint invariant pedestrian recognition with an ensemble of localized features [C] // Proceedings of the 10th European Conference on Computer Vision. Berlin Heidelberg: Springer , 2008: 262-275.
- [12] Hirzer M , Roth P M , Bischof H. Person re-identification by efficient imposter-based metric learning [C] // Proceedings of the IEEE Int'l Conference on Advanced Video and Signal-Based Surveillance. Beijing: IEEE , 2012: 203-208.
- [13] Köstinger M , Hirzer M , Wohlhart P , et al. Large scale metric learning from equivalence constraints [C] // Computer Vision and Pattern Recognition. Providence , RI: IEEE Conference , 2012: 2288-2295.

作者简介



宋亚玲 女 , 1990 年生 , 山东省滕州人 , 中国民航大学电子信息工程学院硕士生。研究方向为图像和视频处理。
E-mail: songyaling1990@163.com



张 良 男 , 1970 年生 , 山东人 , 中国民航大学 博士 教授 硕士生导师。研究方向为图像处理与模式识别、智能视频监控。
E-mail: l-zhang@cauc.edu.cn