

文章编号: 1006-4710(2013)02-0138-06

非重叠多摄像机目标识别方法研究

范彩霞, 朱虹

(西安理工大学 自动化与信息工程学院, 陕西 西安 710048)

摘要: 多摄像机间的目标识别是非重叠多摄像机监控系统中目标持续跟踪的关键。提出一种只依赖于目标外观模型, 不依赖于目标的捕获时间和路径限制的目标识别方法。通过提取三个互补的特征建立目标外观模型, 这些特征包括全局色彩特征、局部颜色和梯度特征以及全局纹理特征, 并通过多特征融合实现目标识别。所提出的方法不仅能用于单幅图像的目标识别, 也能用于视频或图像序列的目标识别。通过基于 VIPeR 数据库和 ETHZ 数据库实验得到的累计匹配特征 (CMC) 曲线表明所提出的算法对视角、姿态、遮挡和光照的变化具有一定的鲁棒性, 能够达到较高的识别率。

关键词: 非重叠多摄像机; 目标识别; 局部特征; 多特征融合

中图分类号: TP391.4

文献标志码: A

Object Recognition Algorithm of Non-Overlapping Cameras

FAN Caixia, ZHU Hong

(Faculty of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: The object recognition by the multi-cameras is the key issue in the non-overlapping surveillance system. A method is presented that depends only on the object appearance model, not on the capture time and path restrictions. The appearance model is established by extracting three complementary features, including global color features, local color and gradient features, and global texture features. Object recognition is realized by multiple cues fusion. The proposed method can be used to recognize not only the object of a single image, but also the object of video or image sequence. The experiments in VIPeR database and ETHZ database show that the algorithm proposed by the authors has a certain robustness of perspective, attitude, occlusion and lighting changes and can achieve high recognition rate.

Key words: non-overlapping cameras; object recognition; local features; multiple cues fusion

随着多摄像机的视频监控系统的广泛应用, 不同时刻、不同摄像机之间的目标匹配识别成为多摄像机跟踪的关键问题之一^[1]。

在非重叠的多摄像机监控系统中, 由于缺少可参考的时间和路径信息, 因此, 只能通过建立目标外观模型进行目标识别, 而且要求所建立的外观模型对光照条件、姿态、摄像机参数和遮挡具有不变性。

目标的外观模型可以通过一幅图像或多幅图像获取。建立外观模型的典型方法是首先计算图像的特征, 之后将它们用不同的策略结合在一起。图像的特征可以是颜色特征、边缘特征、梯度直方图特征

等, 此外, 也可以将图像特征与图像结构相结合。

目前, 建立目标外观模型的研究可分为两类, 一类是使用颜色或形状等目标外观的统计特征; 一类是通过训练获得具有最大区分度的特征建立目标外观模型。

文献[2]、[3]通过建立与空间相关的 HSV 颜色直方图进行目标识别。

文献[4]提出一种人体目标识别的匹配框架, 其将人体目标分为头部、躯干和腿部三部分, 分别在躯干和腿部提取重叠的多个区域的颜色特征进行目标匹配。

收稿日期: 2012-12-20

基金项目: 国家国际科技合作专项项目(2011DFR10480)。

作者简介: 范彩霞, 女, 讲师, 博士生, 研究方向为多摄像机联合监控。E-mail: fcx1981@xaut.edu.cn。

朱虹, 女, 博士, 教授, 博导, 研究方向为视频监控系统彩色图像处理。E-mail: zhuhong@xaut.edu.cn。

文献[5]~[7]提取人体目标的局部特征,并通过图、协方差矩阵等方法建立目标外观模型进行目标识别。

文献[8]通过 AdaBoost 算法选择局部颜色和纹理特征进行目标识别。

文献[9]通过局部最小平方方法将由颜色、纹理和梯度特征构成的高维信息降维后进行目标匹配。这两种方法在目标识别之前的训练阶段是必须的,如果有新的目标出现则必须重新进行训练。

针对人体目标识别问题,本文根据人类视觉特性提出一种多种特征融合的认识方法。首先依据人体比例模型将目标分为头部、躯干和腿部,头部特征参数不参与运算;之后分别在目标躯干和腿部提取外观特征:①基于空间直方图的全局色彩特征;②基于目标局部区域的色彩、梯度方向直方图等特征;③基于 LBP 的全局纹理特征;最后,通过固定权值进行多特征的融合实现人体目标识别。

本文提出的方法无需训练,对视角、姿态、光照和遮挡具有一定的鲁棒性,通过基于 VIPeR 数据库和 ETHZ 数据库的实验,表明本文提出的算法能够达到较高的识别率。

1 建立目标外观模型

为了提高特征的识别效率,本文通过人体比例将人体目标分为三部分:头部(16%);躯干(29%);腿部(55%)^[3]。由于头部信息量较小,不能为识别带来作用,忽略头部信息。因此,仅使用躯干和腿部的信息参与目标识别。

1.1 空间直方图

颜色直方图虽然对人体姿态和视角变化具有不变性,但是其完全忽略了颜色的空间结构信息,降低了模型的分辨能力。将直方图和更高阶的空间矩综合在一起的空间直方图可以保留直方图的空间信息。因此,本文使用二阶空间直方图来表示目标的整体色彩特征。

图像 I 的二阶空间直方图可以表示为^[10]:

$$S_I^{(2)}(b) = \langle n_b, \boldsymbol{\mu}_b, \boldsymbol{\varepsilon}_b \rangle, \quad b = 1, \dots, B \quad (1)$$

式中, B 为量化级数, n_b 为图像的量化直方图, $\boldsymbol{\mu}_b$ 和 $\boldsymbol{\varepsilon}_b$ 分别为均值矢量和协方差矩阵。

$$\begin{cases} n_b = \sum_{k=1}^N \delta_{kb} \\ \boldsymbol{\mu}_b = \frac{1}{n_b} \sum_{k=1}^N \mathbf{x}_k \delta_{kb} \\ \boldsymbol{\varepsilon}_b = \frac{1}{n_b} \sum_{k=1}^N (\mathbf{x}_k - \boldsymbol{\mu}_b) (\mathbf{x}_k - \boldsymbol{\mu}_b)^T \delta_{kb} \end{cases} \quad (2)$$

式中, N 为图像 I 的总像素数;如果第 k 个像素在量化级数 b 内,则 $\delta_{kb} = 1$, 否则 $\delta_{kb} = 0$; $\mathbf{x}_k = [x, y]$ 为像素 k 的空间坐标。

为了克服光照的影响,本文首先在 RGB 空间对图像进行直方图均衡,之后将图像转换到 HSV 空间,分别提取躯干和腿部的空间直方图,并将 HSV 直方图量化为 $16 \times 16 \times 4$ 。

1.2 局部特征

根据人类视觉特性,对局部特征的描述应该包括颜色信息和形状信息。

为了减少噪声的影响,本文首先对图像进行一定的平滑处理,之后通过 FCM 图像分割算法将目标分为若干个具有相似颜色的区域,目标分割会得到一些小区域,这些小区域可能为背景粘连或者由于遮挡粘连的其他目标信息,其在目标识别过程中通常是不可靠的。

为了消除这些小区域对目标识别的影响,得到能够表示目标的具有显著特征的局部区域,提高模型的稳定性,本文去除了区域总像素小于 30 像素的区域。

去除小区域后,对于其他的图像区域,本文采用颜色结合形状信息的描述方式进行局部特征的描述^[11],如公式(3)所示。

$$\mathbf{F} = (\hat{\mathbf{C}}, \lambda \hat{\mathbf{H}}) \quad (3)$$

式中, \mathbf{F} 为颜色和形状相结合的描述, λ 为权值参数,本文取 $\lambda = 0.6$, $\hat{\mathbf{C}}$ 为归一化 HSV 颜色空间的直方图描述, $\hat{\mathbf{H}}$ 为梯度方向直方图(HOG)的归一化描述,该局部特征可以抵抗遮挡、背景粘连以及光照的影响。

本文在特征空间进行主成分分析得到前 30 个特征值所对应的特征向量 \mathbf{V} 以对图像中的各个区域进行描述,结合各个区域的中心位置信息 $Cent$ 和区域大小信息 $Size$ 构成对图像的局部特征的描述。

1.3 LBP 纹理特征

本文采用图像的 LBP 纹理信息来弥补全局颜色特征和局部特征的不足^[12]。图像的 LBP 纹理特征计算公式为:

$$\begin{aligned} LBP_{P,R}(x_c) &= \sum_{p=0}^P s(g_p - g_c) 2^p \\ s(x) &= \begin{cases} 1 & x \geq T \\ 0 & x < T \end{cases} \end{aligned} \quad (4)$$

式中, R 为中心像素与相邻像素的距离,反映了纹理在空间上的分辨率, P 为邻域像素的数目, g_c 为中心点 x_c 处的灰度值, g_p 为以 x_c 为中心点、半径为

R 的圆环的第 p 个等分点的灰度值, T 是为了减轻噪声及局部灰度变化影响而设置的阈值。本文取 $R=1, P=8$, 纹理直方图被量化为 64 个等级。

2 特征匹配

2.1 空间直方图匹配

两个空间直方图 (S, S') 的相似性可以通过两个直方图的相似性的加权和来计算。

$$\rho(S, S') = \sum_{b=1}^B \psi_b \rho_n(n_b, n'_b) \quad (5)$$

式中, $\rho_n(n_b, n'_b)$ 为两个直方图间的相似性, 用 Bhattacharyya 距离计算, ψ_b 为空间相似性, 通过式 (6) 计算。

$$\psi_b = \eta \exp \left\{ -\frac{1}{2} (\boldsymbol{\mu}_b - \boldsymbol{\mu}'_b)^T \hat{\boldsymbol{\epsilon}}_b^{-1} (\boldsymbol{\mu}_b - \boldsymbol{\mu}'_b) \right\} \quad (6)$$

式中, η 为高斯归一化常数, $\hat{\boldsymbol{\epsilon}}_b^{-1} = \boldsymbol{\epsilon}_b^{-1} + (\boldsymbol{\epsilon}'_b)^{-1}$ 。

本文分别计算候选目标和待识别目标的躯干和腿部的空间直方图的相似性, 得到 ρ_T, ρ_L 。因为, 对于人体目标来讲, 一般上身穿着要比下身穿着的区分度大, 所以, 两目标之间的相似性通过式 (7) 计算得到。同理, 文中基于其他特征的目标之间的相似性采用相同的方法计算得到。

$$\rho = 0.7\rho_T + 0.3\rho_L \quad (7)$$

2.2 局部特征匹配

本文采用 EMD (Earth Movers Distance) 距离度量提取的目标局部特征间的相似性。具体描述如下。

设 $A = \{(a_1, w_{a1}), (a_2, w_{a2}), \dots, (a_m, w_{am})\}$ 为具有 m 个聚类的图像 A 的表述形式, 其中 a_i 为聚类的描述形式, w_{ai} 为聚类的权值。 $B = \{(b_1, w_{b1}), (b_2, w_{b2}), \dots, (b_n, w_{bn})\}$ 为具有 n 个聚类的图像 B 的表述形式。 $D = [d_{ij}]$ 为距离矩阵, 其中 d_{ij} 为聚类 a_i 和 b_j 之间的距离函数。则图像 A 和图像 B 的 EMD 距离经归一化处理为:

$$EMD(A, B) = \min_{f_{ij}} \sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij} \quad (8)$$

相应的约束条件为:

$$\begin{cases} f_{ij} \geq 0 & i = 1, \dots, m, j = 1, \dots, n \\ \sum_{j=1}^n f_{ij} = w_{ai} & i = 1, \dots, m \\ \sum_{i=1}^m f_{ij} = w_{bj} & j = 1, \dots, n \end{cases} \quad (9)$$

式 (8) 为线性规划问题, 求解方法参见文献 [13]。

本文中人体目标的局部特征采用的表示方

法为:

$$O = \{(o_1, w_{o1}), (o_2, w_{o2}), \dots, (o_m, w_{om})\}$$

式中 o_i 为由局部特征 V 和区域中心位置 $Cent$ 构成的向量表示, w_{oi} 由区域大小信息 $Size$ 表示。 d_{ij} 通过计算不同目标区域间的欧氏距离得到。

3 多特征融合

在目标识别中通常会有两组图像数据: 候选目标 P 和待识别目标 Q 。本文中, 两个目标间的相似性 S 通过目标的整体颜色特征、局部特征和纹理特征三种互补特征的线性融合得到。

$$S(P, Q) = \alpha \cdot S_{SH} + \beta \cdot S_{LR} + \gamma \cdot S_{LBP} \quad (10)$$

式中, S_{SH} 为目标 P 和目标 Q 的空间直方图的相似性, S_{LR} 为目标 P 和目标 Q 的局部特征的相似性, S_{LBP} 为目标 P 和目标 Q 的 LBP 纹理特征的相似性。本文中使用 Bhattacharyya 距离计算纹理特征的相似性。 α, β, γ 为各特征的权值。本文中, 经过实验设定各特征的权值为固定权值, 即 $\alpha = 0.4, \beta = 0.4, \gamma = 0.2$ 。

根据可获得的目标的帧数, 可以将目标识别分为三类:

- 1) 单对单 (Single vs Single, SvsS), 即候选目标和待识别的目标各只有一帧;
- 2) 多对单 (Multiple vs Single, MvsS), 即候选目标为多帧, 待识别的目标仅一帧;
- 3) 多对多 (Multiple vs Multiple, MvsM), 即候选目标和待识别的目标均为多帧。

因此, 本文通过两个不同的数据库来测试算法在这三种情况下的识别率。在基于单对单 (SvsS) 的目标识别中, 目标的特征仅有一组, 可直接计算出两个目标之间的相似性。

在基于多对单 (MvsS) 和多对多 (MvsM) 的目标识别中, 候选目标和待识别目标存在多组特征, 本文选择相似性最高的特征进行目标识别。

4 实验结果及分析

本文采用两个不同的数据库来验证算法的有效性, 每个数据库都包含有人体目标识别需解决的不同问题。

实验结果通过累计匹配特征 (Cumulative Matching Characteristic, CMC) 曲线进行评价 [14]。 CMC 曲线是指前 n 个匹配中包含正确匹配结果的比率。

4.1 VIPeR 数据库实验结果

VIPeR 数据库包含两组不同视角下的 632 个人体目标 [14]。数据库中的每一组图像都是在两个不

同摄像机(摄像机 A 和摄像机 B)、不同视角、不同姿态和不同光照条件下获取的。所有的图像都被归一化为 128×48 像素。

本文中摄像机 A 中的目标作为候选目标,摄像机 B 中的目标作为待识别目标。该数据库中候选目标和待识别的目标各只有一帧,属于 SvsS 的目标识别。

为了得到鲁棒的实验结果以及便于与文献[8]提出的 ELF 方法进行比较,本文分别从摄像机 A 和摄像机 B 的 632 个人体目标中随机抽取 316 个人体目标,共随机抽取 10 次,形成 10 组目标对,每组包含 316 对人体目标。

最终得到的 CMC 曲线为 10 组 CMC 曲线的平均值,如图 1 所示,其中“排名等级”是指相似度按降序排列的前 n 个目标,识别率是指前 n 个目标中包含正确匹配结果的比率。

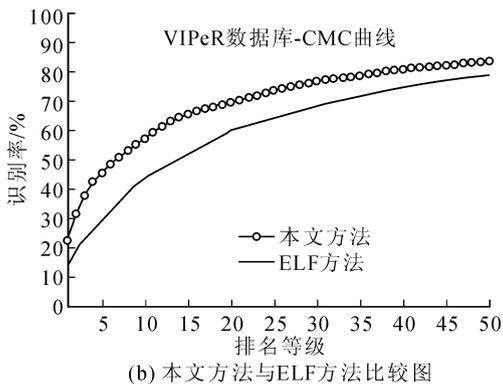
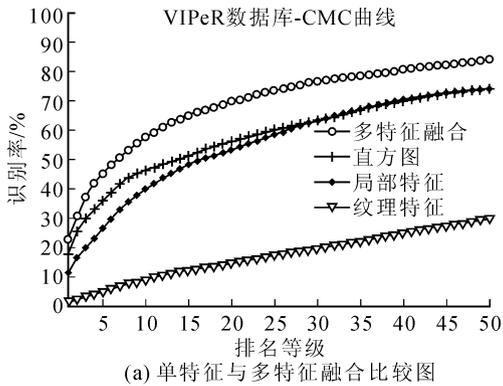


图1 VIPeR 数据库实验结果

Fig. 1 The experiment results of VIPeR database

图 1(a) 为各个单特征与多特征融合的结果比较,从图中可以看出经过多特征融合后,本文方法的识别率有了一定的提高,在第一列中单个特征的识别率不超过 20%,多特征融合后达到了 22%。同时可以看出空间直方图特征和局部特征对目标识别结果的贡献相当,都较高,纹理特征对匹配结果的贡献较小,说明本文在多特征融合中权值的设定合理。

图 1(b) 为本文方法与文献[8]的结果比较,文

献[8]建立了 VIPeR 数据库并针对该数据库提出目标识别方法,其将数据库分为训练数据和测试数据并采用 ELF 方法进行目标的识别。

从图 1(b) 可以看出,本文方法比 ELF 方法具有较高的识别率,在第一列 ELF 方法识别率不超过 20%,在第 15 列本文方法识别率为 66%,ELF 方法识别率为 52%,本文方法识别率高于 ELF 方法 14%,而且本文方法不用提前训练,当有新的目标加入时可以直接进行目标匹配。在 VIPeR 数据库中多数错误匹配的原因是因为相似穿着的目标的存在,此外,较大视角的改变,尤其是正面和背面,也是引起错误匹配的原因。

4.2 ETHZ 数据库实验结果

ETHZ 数据库是从运动的摄像机中获取的,其主要是用于人体目标检测的^[15]。

文献[9]将视频中不同的人体目标检测出来,组成一系列样本数据,用来检测其 PLS 算法。

在 ETHZ 数据库中光照条件的改变和遮挡较严重,人体姿态的变化较小。ETHZ 数据库的基本信息如下:序列 ETHZ1 中包含 83 个不同的人体目标共 4 857 帧图像;序列 ETHZ2 中包含 35 个不同的人体目标共 1 961 帧图像;序列 ETHZ3 中包含 28 个不同的人体目标,共 1 762 帧图像。本文中,ETHZ 数据库中所有的图像都被归一化为 128×64 像素。

在 SvsS 实验中,本文随机抽取 20 组不同的候选目标和待识别目标,并将得到的 CMC 曲线进行平均得到最终结果,为了方便与文献[9]的 PLS 方法进行比较,本文仅选择前 7 个具有较高匹配率的结果。

在 MvsS 和 MvsM 实验中,本文选择多幅关键帧图像分别为 $N = 3, 5, 10$ 三种情况进行实验,对于每种情况,本文随机选取图像并进行 20 次实验得到最终 CMC 曲线。

本文方法在 ETHZ 数据库中的 SvsS、MvsS 和 MvsM 的实验结果以及与文献[9]的对比结果如图 2、图 3 所示。

从图中可以看出,在 SvsS 中,本文算法的效果低于 PLS 算法。但是随着进行目标识别时使用的帧数的增加,即在 MvsS 和 MvsM 的情况下,本文算法的正确匹配率在逐步增加,在序列 ETHZ1 的 MvsS 实验中,当 $N = 3$ 时,在第一列的目标的正确识别率从 $N = 1$ 的 67% 提高到了 83%。在 MvsM 实验中,当 $N = 3$ 时,在第一列的目标的正确识别率从 $N = 1$ 的 67% 提高到了 80%。同样地,在序列 ETHZ2 和序列 ETHZ3 中目标识别率也有较大的提

高。此外,PLS 算法需要得到所有待识别目标后才能进行外观模型权值的估算,如果有新的目标进入则需要全部重新进行计算。而本文算法只需要计算新的目标的外观模型。

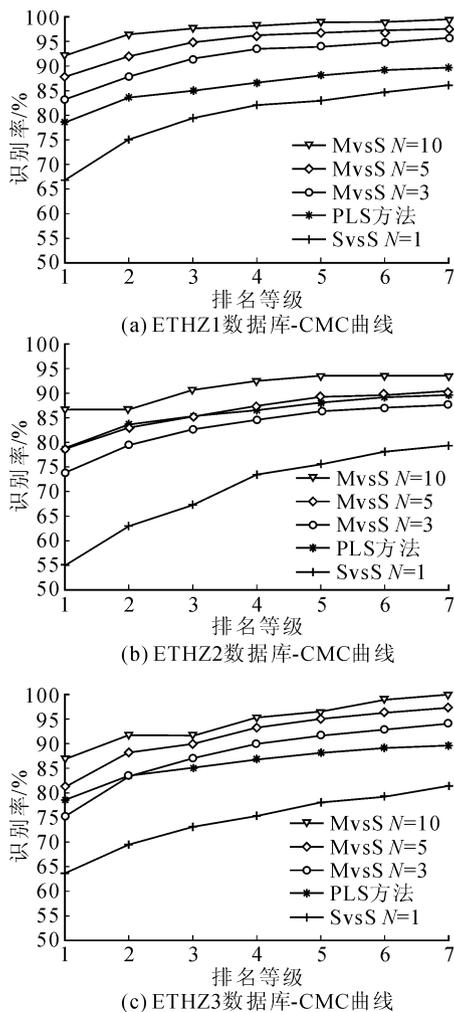


图2 ETHZ 数据库的 MvsS 实验结果
Fig. 2 The MvsS experiment results of ETHZ database

上述基于 VIPeR 数据库和 ETHZ 数据库的实验结果说明了本文所提出算法的有效性以及对姿态、视角、遮挡和光照变化的鲁棒性。

通过实验发现,当人体姿态变化不大时,人体目标的外观改变较小,基于外观模型的识别率较高。

但是,当人体姿态发生较大的改变时,如两个摄像机中目标的姿态分别为正面和背面,人体目标的外观改变较大,因为人的着装有时正背面差别较大,这种情况会导致基于外观模型的目标识别率的下降。

因此,在后续的研究中,有必要对人体的姿态进行简单的分类以提高目标识别率。

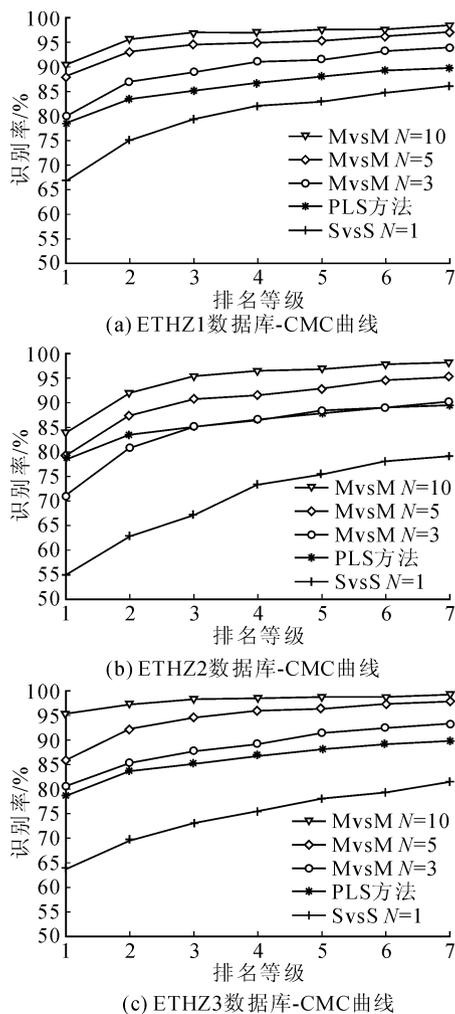


图3 ETHZ 数据库的 MvsM 实验结果
Fig. 3 The MvsM experiment results of ETHZ database

5 结语

本文提出一种基于多特征融合的基于目标外观模型的人体目标识别方法。

分别在人体目标的躯干和腿部提取基于全局色彩特征的空间直方图、基于颜色和梯度的局部特征和基于 LBP 的全局纹理特征,采用不同的相似性度量方法,并采用固定权值法进行多特征的融合得到最终的目标识别结果。

基于 VIPeR 数据库和 ETHZ 数据库的实验结果验证了本文算法的有效性。

本文所提出的算法无需训练,且可以用于 SvsS、MvsS 和 MvsM 任何类型的目标识别,对姿态、视角、遮挡和光照具有一定的鲁棒性。

参考文献:

- [1] 邓颖娜,朱虹. 构造多摄像机全景视图实现有重叠区域

- 的目标跟踪[J]. 西安理工大学学报, 2009, 25(2): 189-192.
- Deng Yingna, Zhu Hong. Object tracking across multiple cameras with overlapping view by panoramic surveillance view construction[J]. Journal of Xi'an University of Technology, 2009, 25(2): 189-192.
- [2] Cai Y, Pietikäinen M. Person re-identification based on global color context; Proceeding of ACCV, 2010 [C]. Queenstown, New Zealand, 2010.
- [3] Khan A, Zhang J, Wang Y. Appearance-based Re-identification of People in Video; Proceeding of DICTA, 2010 [C]. Sydney, Australia, 2010.
- [4] Satta R, Fumera G, Roli F, et al. A multiple component matching framework for person re-identification; Proceeding of ICIAP, 2011 [C]. Ravenna, Italy, 2011.
- [5] Brun L, Conte D, Foggia P, et al. People re-identification by graph kernels methods; Proceeding of GBRPR, 2011 [C]. Münster, Germany, 2011.
- [6] Bak S, Corvee E, Brémond F, et al. Person re-identification using spatial covariance regions of human body parts; Proceeding of AVSS, 2010 [C]. Boston, USA, 2010.
- [7] Hirzer M, Beleznai C, Roth P M, et al. Person re-identification by descriptive and discriminative classification; Proceeding of 17th Scandinavian Conference, SCIA, 2011 [C]. Ystad, Sweden, 2011.
- [8] Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features; Proceeding of ECCV, 2008 [C]. Marseille, France, 2008.
- [9] Schwartz W R, Davis L S. Learning discriminative appearance-based models using partial least squares; Proceeding of SIBGRAPI, 2009 [C]. Rio de Janeiro, Brazil, 2009.
- [10] Birchfield S T, Rangarajan S. Spatiograms versus histograms for region-based tracking; Proceeding of CVPR, 2005 [C]. San Diego, CA, USA, 2005.
- [11] Van De Weijer J, Schmid C. Coloring local feature extraction; Proceeding of ECCV, 2006 [C]. Graz, Austria, 2006.
- [12] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2002, 24(7): 971-987.
- [13] Rubner Y, Tomasi C, Guibas L J. The earth mover's distance as a metric for image retrieval [J]. International Journal of Computer Vision, 2000, 40(2): 99-121.
- [14] Gray D, Brennan S, Tao H. Evaluating appearance models for recognition, reacquisition, and tracking; Proceeding of PETS, 2007 [C]. Rio de Janeiro, Brazil, 2007.
- [15] Ess A, Leibe B, Van Gool L. Depth and appearance for mobile scene analysis; Proceeding of ICCV, 2007 [C]. Rio de Janeiro, Brazil, 2007.

(责任编辑 王卫勋)