



(12) 发明专利

(10) 授权公告号 CN 111783526 B

(45) 授权公告日 2022.08.05

(21) 申请号 202010434344.9

(22) 申请日 2020.05.21

(65) 同一申请的已公布的文献号  
申请公布号 CN 111783526 A

(43) 申请公布日 2020.10.16

(73) 专利权人 昆明理工大学  
地址 650093 云南省昆明市五华区学府路  
253号

(72) 发明人 李华锋 庞健 严双林 欧洋汛  
张亚飞 余正涛

(74) 专利代理机构 昆明人从众知识产权代理有  
限公司 53204  
专利代理师 沈艳尼

(51) Int. Cl.  
G06V 40/20 (2022.01)  
G06V 10/74 (2022.01)  
G06V 10/772 (2022.01)

(56) 对比文件

CN 107563328 A, 2018.01.09

CN 109214442 A, 2019.01.15

CN 104268583 A, 2015.01.07

CN 109190470 A, 2019.01.11

CN 110046870 A, 2019.07.23

CN 110349240 A, 2019.10.18

CN 105160312 A, 2015.12.16

CN 109101865 A, 2018.12.28

CN 106778464 A, 2017.05.31

CN 110826417 A, 2020.02.21

CN 111177447 A, 2020.05.19

US 2002068842 A1, 2002.06.06

US 2003019787 A1, 2003.01.30

Yang E.Shared predictive cross-modal  
deep quantization.《IEEE transactions on  
neural networks and learning systems》  
.2018,第5292-5303页.

(续)

审查员 周循

权利要求书3页 说明书8页 附图1页

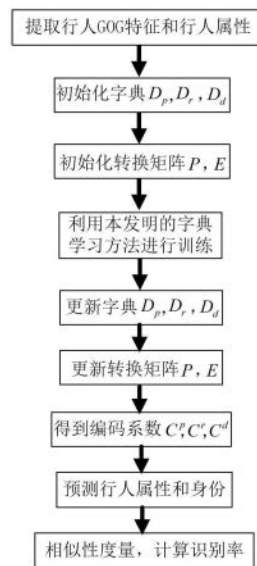
(54) 发明名称

一种利用姿势不变和图结构对齐的跨域行人重识别方法

(57) 摘要

本发明提出了一种利用姿势不变和图结构对齐的跨域行人重识别方法,属于计算机视觉领域。本发明提出了一个基于矩阵分解的字典学习算法来消除数据集间的域信息以及行人姿态信息对跨域行人重识别的影响。具体来说,将方法分为两部分:(1)基于矩阵分解的思想,将原始的视觉特征分解为姿态不变成分、域信息成分及干扰信息成分,目的是为了提取到不受域信息和行人姿态信息影响的视觉成分;(2)为了进一步提高模型泛化能力,通过引入超图结构对齐约束来建立姿态不变特征和语义属性的关系,以便后期能精确预测目标数据集的行人属性,最终能联合行人的姿态不变特征和语义属性来进行行人相似性度量,进一步提高识别性能。

CN 111783526 B



[接上页]

**(56) 对比文件**

张耿宁等. 基于特征融合与核局部Fisher判别分析的行人重识别.《计算机应用》.2016, (第09期),

郑伟诗等. 非对称行人重识别:跨摄像机持

续行人追踪.《中国科学:信息科学》.2018, (第05期),

詹敏等. 基于自适应度量学习的行人再识别.《电脑知识与技术》.2017, (第10期),

1. 一种利用姿势不变和图结构对齐的跨域行人重识别方法,其特征在于:包括如下步骤:

- 1) 定义数据集变量和所用行人特征及属性;
  - 2) 设计特征分解模块,确定含有姿态不变成分字典、域信息成分字典、干扰成分字典、转换矩阵的目标函数;
  - 3) 利用语义属性信息设计超图结构对齐模块;
  - 4) 设计能够减少域偏移的域自适应模块;
  - 5) 合并所提出的损失函数为最终优化函数;
  - 6) 利用交替优化算法得到字典和转换矩阵,从而进一步求得目标域数据编码系数;
  - 7) 通过目标域编码系数预测行人身份和属性;
  - 8) 利用余弦相似度结合预测的身份和属性计算行人间的相似性;
- 具体步骤如下:

步骤1:定义源数据集有K个行人,  $S = \{(x_i^s, a_i, y_i^s) | x_i^s \in X_s, a_i \in A_s, y_i^s \in Y_s\}_{i=1}^{N_s}$ , 其中  $x_i^s \in \mathbb{R}^d$  表示源域s第i个行人特征, d表示特征维度,  $a_i \in \mathbb{R}^c$  表示第i个行人属性, c表示属性维度,  $y_i^s$  表示第i个行人标签,  $N_s$  表示样本数量,  $X_s, A_s, Y_s$  分别表示源域特征集合, 源域属性集合, 源域标签集合, 定义目标数据集  $T = \{x_i^t | x_i^t \in X_t\}_{i=1}^{N_t}$ , 共包含  $N_t$  个样本,  $x_i^t \in \mathbb{R}^d$  表示目标域t第i个行人特征, 在特征层面使用GOG行人特征, 利用现有数据集的属性作为行人属性;

步骤2:设计如下损失函数特征分解项  $L_{FD}$  的目的是将源域特征集合  $X_s \in \mathbb{R}^{d \times N_s}$  分解为姿态不变成分, 域成分, 干扰成分:

$$L_{FD} = \min_{D_p, D_d, D_r, C_p^p, C_p^d, C_p^r} \sum_{i=1}^{N_s} \sum_{v=1}^{V_s} \left\{ \|X_{s,v,i} - D_p C_{s,v,i}^p - D_d C_{s,v,i}^d - D_r C_{s,v,i}^r\|_F^2 + \eta \|X_{s,v,i} - D_p C_{s,v,i}^p - D_r C_{s,v,i}^r\|_* \right. \\ \left. + \lambda_1 \|C_{s,v,i}^p\|_{2,1} + \lambda_2 \|C_{s,v,i}^d\|_{2,1} \right\} + \Phi(D_r, C^p, C^r) \quad (1)$$

$$s.t. \|d_{p,j}\| \leq 1, \|d_{d,j}\| \leq 1, \|d_{r,j}\| \leq 1, \forall j$$

其中,  $V_s$  表示源域视角总数,  $X_{s,v,i}$  表示训练集s中第v视角下第i个身份的特征,  $D_p, D_d, D_r$  分别表示姿态不变成分字典, 域信息成分字典以及干扰成分字典, 而  $C_{s,v,i}^p, C_{s,v,i}^d, C_{s,v,i}^r$  表示  $X_{s,v,i}$  分别对应三个成分字典的编码系数,  $\|\cdot\|_*$  表示矩阵的核范数,  $\|\cdot\|_{2,1}$  表示结构化稀疏范数,  $\eta, \lambda_1, \lambda_2$  表示正则参数, 其中  $\Phi(D_r, C^p, C^r)$  表示促进域分离的正则项, 具体如下:

$$\Phi(D_r, C^p, C^r) = \lambda_3 \left\| Q - (C^p)^T C^p \right\|_F^2 + \lambda_4 \left\| I - (D_r C^r)^T D_r C^r \right\|_F^2 \quad (2)$$

其中  $C^p, C^r$  表示数据集整体编码系数,  $\lambda_3$  和  $\lambda_4$  代表正则参数, I和Q分别表示单位矩阵和身份矩阵;

步骤3:为了增强语义属性的鲁棒性和域不变性,引入语义属性辅助跨域行人重识别, 损失函数超图结构对齐项  $L_{HSA}$  表示如下:

$$L_{HSA} = \min_{P, E} \alpha_1 \|H - PC^p\|_F^2 + \beta_1 \|H - EA_s\|_F^2 + \alpha_2 \text{tr}(C^p LC^{pT}) \quad (3)$$

$$s.t. \|p_j\| \leq 1, \|e_j\| \leq 1, \forall j$$

首先通过源域的图像样本和行人身份构造一个超图  $G(X, E)$ , 包括一组顶点

$J = [j_1, j_2, \dots, j_{|N_j|}]$  和一组超边  $R = [r_1, r_2, \dots, r_{|N_r|}]$ , 其中  $|N_j|$  和  $|N_r|$  分别表示顶点和超边的个数, 对于任意一个给定的超图, 能容易的转换它的超边为一个关联矩阵  $H \in \{0, 1\}^{|N_r| \times |N_j|}$ ,  $\alpha_1$ ,  $\alpha_2, \beta_1$  表示超参数,  $\text{tr}(C^p L C^{pT})$  表示两个超图拉普拉斯正则化,  $P$  和  $E$  表示线性转化系数矩阵,  $L = I - W$  表示超图拉普拉斯矩阵,  $W \in \mathbb{R}^{|N_j| \times |N_j|}$  表示超图的权重矩阵, 用来度量两个顶点之间相关的程度;

$$W = D_x^{-1/2} H W_e D_e^{-1} H^T D_x^{-1/2} \quad (4)$$

$D_x$  和  $D_e$  分别表示超边的degrees和顶点的degrees的对角矩阵,  $W_e$  表示超边权重的一个对角矩阵;

步骤4: 为了解决域偏移, 引入一个域适应项, 用一部分目标域的未标记数据参与特征分解模型的训练, 损失函数域适应项  $L_{DA}$  表示如下:

$$L_{DA} = \min_{D_p, D_d, D_r, C_{t,v,i}^p, C_{t,v,i}^d, C_{t,v,i}^r, Q} \sum_{i=1}^{N_t} \sum_{v=1}^{V_t} \left\{ \|X_{t,v,i} - D_p C_{t,v,i}^p - D_d C_{t,v,i}^d - D_r C_{t,v,i}^r\|_F^2 + \|X_{t,v,i} - D_p C_{t,v,i}^p - D_r C_{t,v,i}^r\|_* + \lambda_2 \|C_{t,v,i}^d\|_{2,1} \right\} \quad (5)$$

$$s.t. \|d_{p,j}\| \leq 1, \|d_{d,j}\| \leq 1, \|d_{r,j}\| \leq 1, \forall j$$

其中,  $V_t$  表示目标域视角总数,  $N_t$  表示目标域样本数量,  $X_{t,v,i}$  表示目标数据集  $t$  中第  $v$  视角下第  $i$  个身份的行人图像特征序列, 而  $C_{t,v,i}^p, C_{t,v,i}^d, C_{t,v,i}^r$  表示  $X_{t,v,i}$  分别对应三个成分字典  $D_p, D_d, D_r$  的编码系数,  $\lambda_2$  为正则化参数, 最后, 整个目标函数表示为:

$$L = L_{FD} + L_{HSA} + L_{DA} \quad (6)$$

步骤5: 接下来对所提函数进行整理合并, 步骤4中整体损失函数  $L$  可展开成如下形式:

$$L_{rec} = \min_{D_p, D_d, D_r, C_{s,v,i}^p, C_{s,v,i}^d, C_{s,v,i}^r, Q} \sum_{i=1}^{N_t} \sum_{v=1}^{V_t} \left\{ \|X_{s,v,i} - D_p C_{s,v,i}^p - D_d C_{s,v,i}^d - D_r C_{s,v,i}^r\|_F^2 + \eta \|X_{s,v,i} - D_p C_{s,v,i}^p - D_r C_{s,v,i}^r\|_* + \lambda_1 \|C_{s,i}^p\|_{2,1} \right. \\ \left. + \lambda_2 \|C_{s,v}^d\|_{2,1} \right\} + \lambda_3 \|Q - (C^p)^T C^p\|_F^2 + \lambda_4 \|I - (D_r C^r)^T D_r C^r\|_F^2 + \alpha_1 \|H - P C^p\|_F^2 \\ + \beta_1 \|H - E A_s\|_F^2 + \alpha_2 \text{tr}(C^p L C^{pT}) \\ + \sum_{i=1}^{N_t} \sum_{v=1}^{V_t} \left\{ \|X_{t,v,i} - D_p C_{t,v,i}^p - D_d C_{t,v,i}^d - D_r C_{t,v,i}^r\|_F^2 + \|X_{t,v,i} - D_p C_{t,v,i}^p - D_r C_{t,v,i}^r\|_* + \lambda_2 \|C_{t,v,i}^d\|_{2,1} \right\} \\ s.t. \|d_{p,j}\| \leq 1, \|d_{d,j}\| \leq 1, \|d_{r,j}\| \leq 1, \|p_j\| \leq 1, \|e_j\| \leq 1, \forall j \quad (7)$$

步骤6: 步骤5中有9个变量需要求解, 利用交替迭代优化算法对每个变量进行求解, 过程中求解其中一个变量需要固定其他变量, 通过求解得到姿态不变成分字典  $D_p$ , 域信息成分字典  $D_d$ , 干扰成分字典  $D_r$ , 以及转换矩阵  $P$  和  $E$ , 有了这些字典, 能利用通过下式计算其对应的编码系数  $c_{t,i}^p, c_{t,i}^d, c_{t,i}^r$ :

$$\{c_{t,i}^p, c_{t,i}^d, c_{t,i}^r\} = \arg \min_{c_{t,i}^p, c_{t,i}^d, c_{t,i}^r} \|x_{t,i} - D_p c_{t,i}^p - D_d c_{t,i}^d - D_r c_{t,i}^r\|_2^2 + \zeta \|c_{t,i}^p\|_2^2 + \zeta \|c_{t,i}^d\|_2^2 + \zeta \|c_{t,i}^r\|_2^2 \quad (8)$$

$\zeta$  表示正则化参数;

步骤7: 当计算得到  $c_{t,i}^p$  后, 利用步骤6求得的转换矩阵  $P, E$  能够通过等式 (9) 和 (10) 来得到  $h_{t,i}$  和  $a_{t,i}$ :

$$h_{t,i} = Pc_{t,i}^p \quad (9)$$

$$a_{t,i} = \arg \min_{a_{t,i}} \alpha_2 \left\| h_{t,i} - Ea_{t,i} \right\|_F^2 \quad (10)$$

上式中将 $h_{t,i}$ 和 $E$ 可看作定值,通过寻找最优的 $a_{t,i}$ ,使得右边项F范数开平方后取最小值即求得此时的 $a_{t,i}$ ,对于测试样本有了预测得到的身份表示 $h_{t,i}$ 和语义属性 $a_{t,i}$ , $\alpha_2$ 表示正则化参数;

步骤8:最后能够通过等式(11)的余弦距离计算公式分别计算行人图像对在身份空间和语义空间的相似性成绩 $\text{sim}_h$ 和 $\text{sim}_a$ ,

$$\text{sim}(z_a, z_b) = \frac{z_a^T \cdot z_b}{\|z_a\|_2 \cdot \|z_b\|_2 (\|z_a - z_b\|_2 + \varepsilon)} \quad (11)$$

其中 $z_a$ 和 $z_b$ 分别表示当前行人身份表示向量和语义属性向量与步骤7中所求 $h_{t,i}$ 和 $a_{t,i}$ 为同一表示,区别在于 $z_a$ 和 $z_b$ 泛指当前行人的身份表示和语义属性,而 $h_{t,i}$ , $a_{t,i}$ 表示第 $i$ 个行人的身份表示和语义属性, $\varepsilon$ 为常数0.0000001,将身份空间和语义属性空间分别得到的相似性成绩进行加权和,加权之后的相似性成绩作为最终的行人对相似性成绩进行相似性度量:

$$\text{sim}_{\text{final}} = \tau \text{sim}_a + (1 - \tau) \text{sim}_h \quad (12)$$

其中 $\tau > 0$ 表示每个空间所占的权重,设置 $\tau = 0.2$ ,通过这种方法最终就可以利用求解的变量衡量目标数据集中行人的相似性。

## 一种利用姿势不变和图结构对齐的跨域行人重识别方法

### 技术领域

[0001] 本发明涉及一种利用姿势不变和图结构对齐的跨域行人重识别方法,属于计算机视觉领域。

### 背景技术

[0002] 随着人工智能的飞速发展,将基于高维特征的行人重识别技术应用到现实生活中的成为一种需要。为此国内外学者在行人重识别方面取得了一系列重大研究进展,发展起了很多方法。一些方法针对目标数据集设计对光照、视角等变化鲁棒的判别性手工艺特征,或者对未标记目标数据进行聚类。但是这类方法的性能很差,主要原因是目标数据没有标签,模型对于挖掘判别性信息非常困难。一些比较先进的方法将行人重识别看作是无监督的域适应问题,它着重进行源域到目标域的知识迁移。与传统的无监督域适应方法相比,行人标签在源域和目标域完全不同,因此挑战更大。这类方法与有监督相比较,性能仍然较差。

### 发明内容

[0003] 本发明目的是提供一种利用姿势不变和图结构对齐的跨域行人重识别方法,用于解决现有行人重识别算法部署困难的问题,本方法能够进行跨域行人重识别,利用域信息的低秩特征和行人的姿态特性,提出一个无监督跨域行人重识别字典分解模型,通过这个分解模型,消除域信息和姿态信息对跨域行人重识别的影响,缩小不同域间的差异;引入一个有效的超图结构对齐约束,建立姿态不变特征和语义属性的转换关系,充分结合两者的优点进行联合度量,具体流程见图1。该方法与现有方法相比可以做跨域重识别任务,即将训练好的模型部署到一个全新的摄像头网络下进行行人的识别。

[0004] 一种利用姿势不变和图结构对齐的跨域行人重识别方法,包括如下步骤:

[0005] 1) 定义数据集变量和所用行人特征及属性;

[0006] 2) 设计特征分解模块,确定含有姿态不变成分字典、域信息成分字典、干扰成分字典、转换矩阵的目标函数;

[0007] 3) 利用语义属性信息设计超图结构对齐模块;

[0008] 4) 设计能够减少域偏移的域自适应模块;

[0009] 5) 合并所提出的损失函数为最终优化函数;

[0010] 6) 利用交替优化算法得到字典和转换矩阵,从而进一步求得目标域数据编码系数;

[0011] 7) 通过目标域编码系数预测行人身份和属性;

[0012] 8) 利用余弦相似度结合预测的身份和属性计算行人间的相似性。

[0013] 具体步骤如下:

[0014] 步骤1:定义源数据集有K个行人,

[0015]  $S = \{(x_i^s, a_i, y_i^s) | x_i^s \in X_s, a_i \in A_s, y_i^s \in Y_s\}_{i=1}^{N_s}$ , 其中  $x_i^s \in \mathbb{R}^d$  表示源域s第i个行人特征,d

表示特征维度,  $a_i \in \mathbb{R}^c$  表示第*i*个行人属性,  $c$  表示属性维度,  $y_i^s$  表示第*i*个行人标签,  $N_s$  表示样本数量。  $X_s, A_s, Y_s$  分别表示源域特征集合, 源域属性集合, 源域标签集合。定义目标数据集  $T = \{x_i^t \mid x_i^t \in X_t\}_{i=1}^{N_t}$ , 共包含  $N_t$  个样本,  $x_i^t \in \mathbb{R}^d$  表示目标域*t*第*i*个行人特征,  $d$  表示特征维度。

在特征层面使用GOG行人特征, 利用现有数据集的属性作为行人属性。

[0016] 步骤2: 设计如下损失函数特征分解项 (Feature Decomposition, FD)  $L_{FD}$  的目的是将源域特征集合分解为姿态不变成分, 域成分, 干扰成分:

$$L_{FD} = \min_{D_p, D_d, D_r, C_{s,v,i}^p, C_{s,v,i}^d, C_{s,v,i}^r} \sum_{i=1}^{N_s} \sum_{v=1}^{V_s} \left\{ \left\| X_{s,v,i} - D_p C_{s,v,i}^p - D_d C_{s,v,i}^d - D_r C_{s,v,i}^r \right\|_F^2 + \eta \left\| X_{s,v,i} - D_p C_{s,v,i}^p - D_r C_{s,v,i}^r \right\|_* \right. \\ \left. + \lambda_1 \left\| C_{s,v,i}^p \right\|_{2,1} + \lambda_2 \left\| C_{s,v,i}^d \right\|_{2,1} \right\} + \Phi(D_r, C^p, C^r) \quad (1)$$

$$s.t. \left\| d_{p,j} \right\| \leq 1, \left\| d_{d,j} \right\| \leq 1, \left\| d_{r,j} \right\| \leq 1, \forall j$$

[0018] 其中,  $V_s$  表示源域视角总数,  $X_{s,v,i}$  表示训练集*s*中第*v*视角下第*i*个身份的特征。  $D_p, D_d, D_r$  分别表示姿态不变成分字典, 域信息成分字典以及干扰成分字典。而  $C_{s,v,i}^p, C_{s,v,i}^d, C_{s,v,i}^r$  表示  $X_{s,v,i}$  分别对应三个成分字典的编码系数。  $\| \cdot \|_*$  表示矩阵的核范数,  $\| \cdot \|_{2,1}$  表示结构化稀疏范数。  $\eta, \lambda_1, \lambda_2$  表示正则参数。其中  $\Phi(D_r, C^p, C^r)$  表示促进域分离的正则项, 具体如下:

$$\Phi(D_r, C^p, C^r) = \lambda_3 \left\| Q - (C^p)^T C^p \right\|_F^2 + \lambda_4 \left\| I - (D_r C^r)^T D_r C^r \right\|_F^2 \quad (2)$$

[0020] 其中  $C^p, C^r$  表示数据集整体编码系数。  $\lambda_3$  和  $\lambda_4$  代表正则参数。  $I$  和  $Q$  分别表示单位矩阵和身份矩阵。

[0021] 步骤3: 为了增强语义属性的鲁棒性和域不变性, 引入语义属性辅助跨域行人重识别。损失函数超图结构对齐项 (Hypergraph Structure Alignment, HSA)  $L_{HSA}$  表示如下:

$$L_{HSA} = \min_{P,E} \alpha_1 \left\| H - PC^p \right\|_F^2 + \beta_1 \left\| H - EA_s \right\|_F^2 + \alpha_2 \text{tr}(C^p LC^{pT}) \quad (3)$$

$$s.t. \left\| p_j \right\| \leq 1, \left\| e_j \right\| \leq 1, \forall j$$

[0023] 首先通过源域的图像样本和行人身份构造一个超图  $G(X, E)$ , 包括一组顶点  $J = [j_1, j_2, \dots, j_{|N_j|}]$ , 和一组超边  $R = [r_1, r_2, \dots, r_{|N_r|}]$ , 其中  $|N_j|$  和  $|N_r|$  分别表示顶点和超边的个数。对于任意一个给定的超图, 能容易的转换它的超边为一个关联矩阵  $H \in \{0,1\}^{|N_r| \times |N_j|}$ ,  $\alpha_1, \alpha_2, \beta_1$  表示超参数,  $\text{tr}(C^p LC^{pT})$  表示两个超图拉普拉斯正则化,  $P$  和  $E$  表示线性转化系数矩阵,  $L = I - W$  表示超图拉普拉斯矩阵,  $W \in \mathbb{R}^{|N_j| \times |N_j|}$  表示超图的权重矩阵, 用来度量两个顶点之间相关的程度:

$$W = D_x^{-1/2} H W_e D_e^{-1} H^T D_x^{-1/2} \quad (4)$$

[0025]  $D_x$  和  $D_e$  分别表示超边的degrees和顶点的degrees的对角矩阵。  $W_e$  表示超边权重的一个对角矩阵。

[0026] 步骤4: 为了解决域偏移, 引入一个域适应项, 用一部分目标域的未标记数据参与特征分解模型的训练, 损失函数域适应项 (Domain Adaptation, DA)  $L_{DA}$  表示如下:

$$L_{DA} = \min_{D_p, D_d, D_r, C_{t,v,i}^p, C_{t,v,i}^d, C_{t,v,i}^r, P, Q} \sum_{i=1}^{N_t} \sum_{v=1}^{V_t} \left\{ \|X_{t,v,i} - D_p C_{t,v,i}^p - D_d C_{t,v,i}^d - D_r C_{t,v,i}^r\|_F^2 + \|X_{t,v,i} - D_p C_{t,v,i}^p - D_r C_{t,v,i}^r\|_* + \lambda_2 \|C_{t,v,i}^d\|_{2,1} \right\} \quad (5)$$

s.t.  $\|d_{p,j}\| \leq 1, \|d_{d,j}\| \leq 1, \|d_{r,j}\| \leq 1, \forall j$

[0028] 其中,  $V_t$  表示目标域视角总数,  $N_t$  表示目标域样本数量,  $X_{t,v,i}$  表示目标数据集  $t$  中第  $v$  视角下第  $i$  个身份的行人图像特征序列。而  $C_{t,v,i}^p, C_{t,v,i}^d, C_{t,v,i}^r$  表示  $X_{t,v,i}$  分别对应三个成分字典  $D_p, D_d, D_r$  的编码系数。 $\lambda_2$  为正则化参数。最后, 整个目标函数表示为:

$$L = L_{FD} + L_{HSA} + L_{DA} \quad (6)$$

[0030] 步骤5: 接下来对所提函数进行整理合并, 步骤4中整体损失函数  $L$  可展开成如下形式:

$$L_{rec} = \min_{D_p, D_d, D_r, C_{s,v,i}^p, C_{s,v,i}^d, C_{s,v,i}^r, P, Q} \sum_{i=1}^{N_s} \sum_{v=1}^{V_s} \left\{ \|X_{s,v,i} - D_p C_{s,v,i}^p - D_d C_{s,v,i}^d - D_r C_{s,v,i}^r\|_F^2 + \eta \|X_{s,v,i} - D_p C_{s,v,i}^p - D_r C_{s,v,i}^r\|_* + \lambda_1 \|C_{s,v,i}^p\|_{2,1} \right. \\ \left. + \lambda_2 \|C_{s,v,i}^d\|_{2,1} \right\} + \lambda_3 \|Q - (C^p)^T C^p\|_F^2 + \lambda_4 \|I - (D_r C^r)^T D_r C^r\|_F^2 + \alpha_1 \|H - PC^p\|_F^2 \\ + \beta_1 \|H - EA_s\|_F^2 + \alpha_2 \text{tr}(C^p L C^{pT}) \\ + \sum_{i=1}^{N_t} \sum_{v=1}^{V_t} \left\{ \|X_{t,v,i} - D_p C_{t,v,i}^p - D_d C_{t,v,i}^d - D_r C_{t,v,i}^r\|_F^2 + \|X_{t,v,i} - D_p C_{t,v,i}^p - D_r C_{t,v,i}^r\|_* + \lambda_2 \|C_{t,v,i}^d\|_{2,1} \right\} \\ \text{s.t. } \|d_{p,j}\| \leq 1, \|d_{d,j}\| \leq 1, \|d_{r,j}\| \leq 1, \|p_j\| \leq 1, \|e_j\| \leq 1, \forall j \quad (7)$$

[0032] 步骤6: 步骤5中有9个变量需要求解, 利用交替迭代优化算法对每个变量进行求解, 过程中求解其中一个变量需要固定其他变量。通过求解得到姿态不变成分字典  $D_p$ , 域信息成分字典  $D_d$ , 干扰成分字典  $D_r$ , 以及转换矩阵  $P$  和  $E$ 。有了这些字典, 能利用通过下式计算其对应的编码系数  $c_{t,i}^p, c_{t,i}^d, c_{t,i}^r$ :

$$\{c_{t,i}^p, c_{t,i}^d, c_{t,i}^r\} = \arg \min_{c_{t,i}^p, c_{t,i}^d, c_{t,i}^r} \|x_{t,i} - D_p c_{t,i}^p - D_d c_{t,i}^d - D_r c_{t,i}^r\|_2^2 + \zeta \|c_{t,i}^p\|_2^2 + \zeta \|c_{t,i}^d\|_2^2 + \zeta \|c_{t,i}^r\|_2^2 \quad (8)$$

[0034]  $\zeta$  表示正则化参数。

[0035] 步骤7: 当计算得到  $c_{t,i}^p$  后, 利用步骤6求得的转换矩阵  $P, E$  能够通过等式 (9) 和 (10) 来得到  $h_{t,i}$  和  $a_{t,i}$ :

$$h_{t,i} = P c_{t,i}^p \quad (9)$$

$$a_{t,i} = \arg \min_{a_{t,i}} \alpha_2 \|h_{t,i} - E a_{t,i}\|_F^2 \quad (10)$$

[0038] 上式中将  $h_{t,i}$  和  $E$  可看作定值, 通过寻找最优的  $a_{t,i}$  使得右边项  $F$  范数开平方后取最小值即求得此时的  $a_{t,i}$ 。对于测试样本有了预测得到的身份表示  $h_{t,i}$  和语义属性  $a_{t,i}$ 。 $\alpha_2$  表示正则化参数。

[0039] 步骤8: 最后能够通过等式 (11) 的余弦距离计算公式分别计算行人图像对在身份空间和语义空间的相似性成绩  $\text{sim}_h$  和  $\text{sim}_a$ 。

$$\text{sim}(z_a, z_b) = \frac{z_a^T \cdot z_b}{\|z_a\|_2 \cdot \|z_b\|_2 (\|z_a - z_b\|_2 + \varepsilon)} \quad (11)$$



[0041] 其中 $z_a$ 和 $z_b$ 分别表示当前行人身份表示向量和语义属性向量与步骤7中所求 $h_{t,i}$ 和 $a_{t,i}$ 为同一表示,区别在于 $z_a$ 和 $z_b$ 泛指当前行人的身份表示和语义属性,而 $h_{t,i}, a_{t,i}$ 表示第 $i$ 个行人的身份表示和语义属性。 $\epsilon$ 为常数0.0000001。将身份空间和语义属性空间分别得到的相似性成绩进行加权和,加权之后的相似性成绩作为最终的行人对相似性成绩进行相似性度量。

$$[0042] \quad \text{sim}_{\text{final}} = \tau \text{sim}_a + (1 - \tau) \text{sim}_h \quad (12)$$

[0043] 其中 $\tau > 0$ 表示每个空间所占的权重。在本发明中设置 $\tau = 0.2$ 。通过这种方法最终就可以利用求解的变量衡量目标数据集中行人的相似性。

[0044] 本发明的有益效果如下:

[0045] (1) 通过所提出的分解模型,消除了数据集间的域信息和行人姿态信息对跨域行人重识别的影响,缩小不同域间的差异。有利于模型在现实场景中提取行人更具鲁棒性的特征。

[0046] (2) 通过引入一个有效的超图结构对齐约束,建立了姿态不变特征和语义属性的转换关系,结合两者进行的相似性度量方法使得模型对于不同行人之间更具有鉴别性,例如两个行人外观非常相似,但是通过属性信息可以避免将两者识别为同一行人从而避免误判。

## 附图说明

[0047] 图1为本发明利用姿势不变和图结构对齐的跨域行人重识别方法的流程图。

## 具体实施方式

[0048] 下面结合附图和具体实施例,对本发明做进一步说明。

[0049] 实施例1:如图1所示,一种利用姿势不变和图结构对齐的跨域行人重识别方法,包括如下步骤:

[0050] 1) 定义数据集变量和所用行人特征及属性;

[0051] 2) 设计特征分解模块,确定含有姿态不变成分字典、域信息成分字典、干扰成分字典、转换矩阵的目标函数;

[0052] 3) 利用语义属性信息设计超图结构对齐模块;

[0053] 4) 设计能够减少域偏移的域自适应模块;

[0054] 5) 合并所提出的损失函数为最终优化函数;

[0055] 6) 利用交替优化算法得到字典和转换矩阵,从而进一步求得目标域数据编码系数;

[0056] 7) 通过目标域编码系数预测行人身份和属性;

[0057] 8) 利用余弦相似度结合预测的身份和属性计算行人间的相似性。

[0058] 进步一地,具体步骤如下:

[0059] 步骤1:定义源数据集有 $K$ 个行人,

[0060]  $S = \{(x_i^s, a_i, y_i^s) | x_i^s \in X_s, a_i \in A_s, y_i^s \in Y_s\}_{i=1}^{N_s}$ ,其中 $x_i^s \in \mathbb{R}^d$ 表示源域 $s$ 第 $i$ 个行人特征, $d$ 表示特征维度, $a_i \in \mathbb{R}^c$ 表示第 $i$ 个行人属性, $c$ 表示属性维度, $y_i^s$ 表示第 $i$ 个行人标签, $N_s$ 表

示样本数量。 $X_s, A_s, Y_s$ 分别表示源域特征集合,源域属性集合,源域标签集合。定义目标数据集  $T = \{x_t^i | x_t^i \in X_t\}_{i=1}^{N_t}$ , 共包含  $N_t$  个样本,  $x_t^i \in \mathbb{R}^d$  表示目标域  $t$  第  $i$  个行人特征,  $d$  表示特征维度。在特征层面使用GOG行人特征,利用现有数据集的属性作为行人属性。

[0061] 步骤2:设计如下损失函数特征分解项 (Feature Decomposition, FD)  $L_{FD}$  的目的是将源域特征集合  $X_s \in \mathbb{R}^{d \times N_s}$  分解为姿态不变成分,域成分,干扰成分:

$$L_{FD} = \min_{D_p, D_d, D_r, C_p^p, C_p^d, C_p^r} \sum_{i=1}^{N_s} \sum_{v=1}^{V_s} \left\{ \|X_{s,v,i} - D_p C_{s,v,i}^p - D_d C_{s,v,i}^d - D_r C_{s,v,i}^r\|_F^2 + \eta \|X_{s,v,i} - D_p C_{s,v,i}^p - D_r C_{s,v,i}^r\|_* \right. \\ \left. + \lambda_1 \|C_{s,v,i}^p\|_{2,1} + \lambda_2 \|C_{s,v,i}^d\|_{2,1} \right\} + \Phi(D_r, C^p, C^r) \quad (1)$$

$$s.t. \|d_{p,j}\| \leq 1, \|d_{d,j}\| \leq 1, \|d_{r,j}\| \leq 1, \forall j$$

[0063] 其中,  $V_s$  表示源域视角总数,  $X_{s,v,i}$  表示训练集  $s$  中第  $v$  视角下第  $i$  个身份的特征。  $D_p, D_d, D_r$  分别表示姿态不变成分字典,域信息成分字典以及干扰成分字典。而  $C_{s,v,i}^p, C_{s,v,i}^d, C_{s,v,i}^r$  表示  $X_{s,v,i}$  分别对应三个成分字典的编码系数。  $\|\cdot\|_*$  表示矩阵的核范数,  $\|\cdot\|_{2,1}$  表示结构化稀疏范数。  $\eta, \lambda_1, \lambda_2$  表示正则参数。其中  $\Phi(D_r, C^p, C^r)$  表示促进域分离的正则项,具体如下:

$$\Phi(D_r, C^p, C^r) = \lambda_3 \|Q - (C^p)^T C^p\|_F^2 + \lambda_4 \|I - (D_r C^r)^T D_r C^r\|_F^2 \quad (2)$$

[0065] 其中  $C^p, C^r$  表示数据集整体编码系数。  $\lambda_3$  和  $\lambda_4$  代表正则参数。  $I$  和  $Q$  分别表示单位矩阵和身份矩阵。

[0066] 步骤3:为了增强语义属性的鲁棒性和域不变性,引入语义属性辅助跨域行人重识别。损失函数超图结构对齐项 (Hypergraph Structure Alignment, HSA)  $L_{HSA}$  表示如下:

$$L_{HSA} = \min_{P,E} \alpha_1 \|H - PC^p\|_F^2 + \beta_1 \|H - EA_s\|_F^2 + \alpha_2 \text{tr}(C^p LC^{pT}) \quad (3)$$

$$s.t. \|p_j\| \leq 1, \|e_j\| \leq 1, \forall j$$

[0068] 首先通过源域的图像样本和行人身份构造一个超图  $G(X, E)$ , 包括一组顶点  $J = [j_1, j_2, \dots, j_{|N_j|}]$ , 和一组超边  $R = [r_1, r_2, \dots, r_{|N_r|}]$ , 其中  $|N_j|$  和  $|N_r|$  分别表示顶点和超边的个数。对于任意一个给定的超图,能容易的转换它的超边为一个关联矩阵  $H \in \{0,1\}^{|N_r| \times |N_j|}$ ,  $\alpha_1, \alpha_2, \beta_1$  表示超参数,  $\text{tr}(C^p LC^{pT})$  表示两个超图拉普拉斯正则化,  $P$  和  $E$  表示线性转化系数矩阵,  $L = I - W$  表示超图拉普拉斯矩阵,  $W \in \mathbb{R}^{|N_j| \times |N_j|}$  表示超图的权重矩阵,用来度量两个顶点之间相关的程度:

$$W = D_x^{-1/2} H W_e D_e^{-1} H^T D_x^{-1/2} \quad (4)$$

[0070]  $D_x$  和  $D_e$  分别表示超边的degrees和顶点的degrees的对角矩阵。  $W_e$  表示超边权重的一个对角矩阵。

[0071] 步骤4:为了解决域偏移,引入一个域适应项,用一部分目标域的未标记数据参与特征分解模型的训练,损失函数域适应项 (Domain Adaptation, DA)  $L_{DA}$  表示如下:

$$\begin{aligned}
L_{DA} = & \min_{D_p, D_d, D_r, C_{t,v,i}^p, C_{t,v,i}^d, C_{t,v,i}^r} \sum_{i=1}^{N_t} \sum_{v=1}^{V_t} \left\{ \|X_{t,v,i} - D_p C_{t,v,i}^p - D_d C_{t,v,i}^d - D_r C_{t,v,i}^r\|_F^2 + \|X_{t,v,i} - D_p C_{t,v,i}^p - D_r C_{t,v,i}^r\|_* + \lambda_2 \|C_{t,v,i}^d\|_{2,1} \right\} \\
[0072] \quad & \text{s.t. } \|d_{p,j}\| \leq 1, \|d_{d,j}\| \leq 1, \|d_{r,j}\| \leq 1, \forall j \\
& (5)
\end{aligned}$$

[0073] 其中,  $V_t$  表示目标域视角总数,  $N_t$  表示目标域样本数量,  $X_{t,v,i}$  表示目标数据集  $t$  中第  $v$  视角下第  $i$  个身份的行人图像特征序列。而  $C_{t,v,i}^p$ ,  $C_{t,v,i}^d$ ,  $C_{t,v,i}^r$  表示  $X_{t,v,i}$  分别对应三个成分字典  $D_p, D_d, D_r$  的编码系数。 $\lambda_2$  为正则化参数。最后, 整个目标函数表示为:

$$[0074] \quad L = L_{FD} + L_{HSA} + L_{DA} \quad (6)。$$

[0075] 步骤5: 接下来对所提函数进行整理合并, 步骤4中整体损失函数  $L$  可展开成如下形式:

$$\begin{aligned}
L_{rec} = & \min_{D_p, D_d, D_r, C_{s,v}^p, C_{s,v}^d, C_{s,v}^r, P, Q} \sum_{i=1}^{N_s} \sum_{v=1}^{V_s} \left\{ \|X_{s,v,i} - D_p C_{s,v,i}^p - D_d C_{s,v,i}^d - D_r C_{s,v,i}^r\|_F^2 + \eta \|X_{s,v,i} - D_p C_{s,v,i}^p - D_r C_{s,v,i}^r\|_* + \lambda_1 \|C_{s,v,i}^p\|_{2,1} \right. \\
& + \lambda_2 \|C_{s,v,i}^d\|_{2,1} \left. \right\} + \lambda_3 \|Q - (C^p)^T C^p\|_F^2 + \lambda_4 \|I - (D, C^r)^T D, C^r\|_F^2 + \alpha_1 \|H - PC^p\|_F^2 \\
& + \beta_1 \|H - EA_s\|_F^2 + \alpha_2 \text{tr}(C^p LC^{pT}) \\
[0076] \quad & + \sum_{i=1}^{N_t} \sum_{v=1}^{V_t} \left\{ \|X_{t,v,i} - D_p C_{t,v,i}^p - D_d C_{t,v,i}^d - D_r C_{t,v,i}^r\|_F^2 + \|X_{t,v,i} - D_p C_{t,v,i}^p - D_r C_{t,v,i}^r\|_* + \lambda_2 \|C_{t,v,i}^d\|_{2,1} \right\} \\
& \text{s.t. } \|d_{p,j}\| \leq 1, \|d_{d,j}\| \leq 1, \|d_{r,j}\| \leq 1, \|p_j\| \leq 1, \|e_j\| \leq 1, \forall j \\
& (7)
\end{aligned}$$

[0077] 步骤6: 步骤5中有9个变量需要求解, 利用交替迭代优化算法对每个变量进行求解, 过程中求解其中一个变量需要固定其他变量。通过求解得到姿态不变成分字典  $D_p$ , 域信息成分字典  $D_d$ , 干扰成分字典  $D_r$ , 以及转换矩阵  $P$  和  $E$ 。有了这些字典, 能利用通过下式计算其对应的编码系数  $c_{t,i}^p, c_{t,i}^d, c_{t,i}^r$ :

$$[0078] \quad \{c_{t,i}^p, c_{t,i}^d, c_{t,i}^r\} = \arg \min_{c_{t,i}^p, c_{t,i}^d, c_{t,i}^r} \|x_{t,i} - D_p c_{t,i}^p - D_d c_{t,i}^d - D_r c_{t,i}^r\|_2^2 + \zeta \|c_{t,i}^p\|_2^2 + \zeta \|c_{t,i}^d\|_2^2 + \zeta \|c_{t,i}^r\|_2^2 \quad (8)$$

[0079]  $\zeta$  表示正则化参数。

[0080] 步骤7: 当计算得到  $c_{t,i}^p$  后, 利用步骤6求得的转换矩阵  $P, E$  能够通过等式 (9) 和 (10) 来得到  $h_{t,i}$  和  $a_{t,i}$ :

$$[0081] \quad h_{t,i} = P c_{t,i}^p \quad (9)$$

$$[0082] \quad a_{t,i} = \arg \min_{a_{t,i}} \alpha_2 \|h_{t,i} - E a_{t,i}\|_F^2 \quad (10)$$

[0083] 上式中将  $h_{t,i}$  和  $E$  可看作定值, 通过寻找最优的  $a_{t,i}$  使得右边项  $F$  范数开平方后取最小值即求得此时的  $a_{t,i}$ 。对于测试样本有了预测得到的身份表示  $h_{t,i}$  和语义属性  $a_{t,i}$ 。 $\alpha_2$  表示正则化参数。

[0084] 步骤8: 最后能够通过等式 (11) 的余弦距离计算公式分别计算行人图像对在身份空间和语义空间的相似性成绩  $\text{sim}_h$  和  $\text{sim}_a$ 。

$$[0085] \quad \text{sim}(z_a, z_b) = \frac{z_a^T \cdot z_b}{\|z_a\|_2 \cdot \|z_b\|_2 (\|z_a - z_b\|_2 + \varepsilon)} \quad (11)$$

[0086] 其中 $z_a$ 和 $z_b$ 分别表示当前行人身份表示向量和语义属性向量与步骤7中所求 $h_{t,i}$ 和 $a_{t,i}$ 为同一表示,区别在于 $z_a$ 和 $z_b$ 泛指当前行人的身份表示和语义属性,而 $h_{t,i}, a_{t,i}$ 表示第 $i$ 个行人的身份表示和语义属性。 $\epsilon$ 为常数0.0000001。将身份空间和语义属性空间分别得到的相似性成绩进行加权和,加权之后的相似性成绩作为最终的行人对相似性成绩进行相似性度量。

$$[0087] \quad \text{sim}_{\text{final}} = \tau \text{sim}_a + (1 - \tau) \text{sim}_h \quad (12)$$

[0088] 其中 $\tau > 0$ 表示每个空间所占的权重。在本发明中设置 $\tau = 0.2$ 。通过这种方法最终就可以利用求解的变量衡量目标数据集中行人的相似性。

[0089] 在上述提出的模型中,有11个参数需要设置,包括字典 $D_p, D_d, D_r$ 的原子大小 $d_p, d_d, d_r$ ,以及正则项参数 $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \alpha_1, \alpha_2, \beta, \zeta$ 。在实验中,分别设置这些参数为 $d_p = 600, d_d = 180, d_r = 180, \lambda_1 = 0.0001, \lambda_2 = 0.0001, \lambda_3 = 0.01, \lambda_4 = 1, \alpha_1 = 0.1, \alpha_2 = 0.1, \beta = 0.1, \zeta = 0.1$ 。

[0090] 使用GOG特征作为行人视觉特征,用标准的已经表示的语义属性作为行人属性表示。为了体现算法可部署在现实生活中,在VIPeR数据集上进行实验。该数据集包含两个摄像头,每个摄像头每人捕获一个图像。数据集中有多种行人姿态变化,还有视角和光照变化。以prid2011和grid作为源数据集,将模型平均分为训练和测试。重复训练10次求得平均值作为最终性能。比较结果如表1所示。该实验证明了的方法可以直接将训练好的模型部署到VIPeR场景下进行识别,并且保持良好的识别率。

[0091] 表1 VIPeR数据集

Methods	rank1	rank5	rank10	rank20
Adversarial(2016)	22.80	38.60	50.30	<b>63.90</b>
UMDL(2016)	31.50	-	-	-
SDC(2017)	25.80	-	-	-
CAMEL(2017)	30.90	-	-	-
UJSDL(2018)	20.30	38.04	49.11	60.38
AIESL(2020)	28.92	40.41	46.58	52.44
DAPR-SSAE(2020)	26.84	39.72	49.27	60.38
Ours	<b>31.80</b>	<b>46.30</b>	<b>54.81</b>	63.26

[0094] 本发明同样在CUHK01数据集上进行了实验,该数据集采集自香港中文大学校园场景,摄像机分别部署在教学楼以及室外场景中,视角分步广泛。以VIPeR为源数据集,CUHK01为目标数据集进行测试。结果如表2所示,表中还展示了其他方法的性能,从结果可以看到,的方法达到了比较高的性能。

[0095] 表2 CUHK01数据集

Methods	Rank1	rank5	rank10	rank20
UDML(2016)	27.10	-	-	-
TSR(2017)	22.40	35.90	47.90	64.50
CAMEL(2017)	57.30	-	-	-
DAS(2018)	54.90	-	-	-
UJSDL(2018)	27.74	48.81	57.71	66.85
DECAMEL(2019)	65.81	-	-	-
AIESL(2020)	63.26	81.63	87.36	91.84
Ours	<b>64.76</b>	<b>82.68</b>	<b>89.20</b>	<b>93.63</b>

[0096]

[0097] 以上结合附图对本发明的具体实施方式作了详细说明,但是本发明并不限于上述实施方式,在本领域普通技术人员所具备的知识范围内,还可以在不脱离本发明宗旨的前提下作出各种变化。

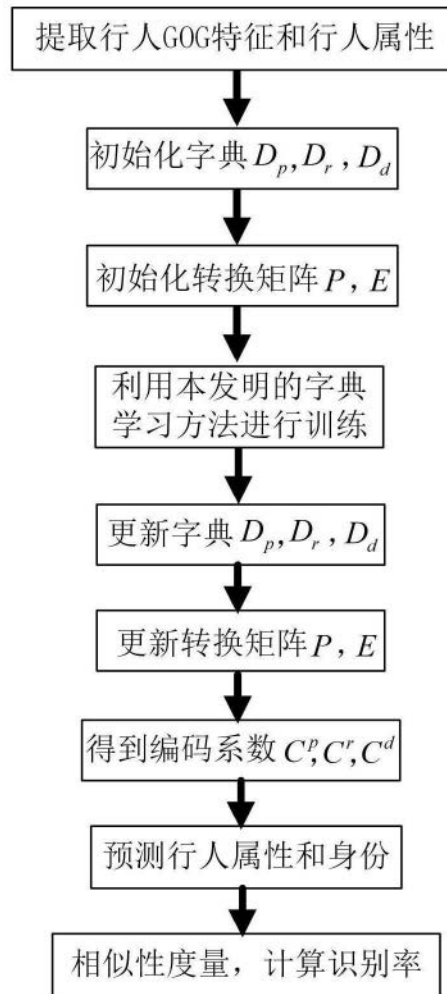


图1