

智能视觉监控中行人再识别技术研究

Person Re-Identification in Intelligent Visual Surveillance

答辩人：四建楼

导师：张洪刚老师

sijianlou@bupt.edu.cn

北京邮电大学信息与通信工程学院
模式识别实验室

April 19, 2018

目录

绪论

行人再识别中的相关算法

基于空间金字塔统计特征及多核学习的行人再识别算法

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于正则化度量学习的行人再识别算法

总结和展望

目录

绪论

行人再识别中的相关算法

基于空间金字塔统计特征及多核学习的行人再识别算法

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于正则化度量学习的行人再识别算法

总结和展望

绪论

研究背景

- 视频监控网络、互联网摄像头等越来越普及
- 随着监控数据量的不断增长，单纯依靠人力进行数据的分析和处理变得代价高昂，甚至不切实际
- 为了实现多摄像头协同监控，我们有必要探索高效的数据关联策略使不同摄像机下的行人监控视频分析结果得以结合

绪论

研究背景

- 视频监控网络、互联网摄像头等越来越普及
- 随着监控数据量的不断增长，单纯依靠人力进行数据的分析和处理变得代价高昂，甚至不切实际 – **智能视觉监控 (Intelligent Visual Surveillance, IVS) 技术**
- 为了实现多摄像头协同监控，我们有必要探索高效的数据关联策略使不同摄像机下的行人监控视频分析结果得以结合 – **行人再识别 (Person Re-Identification, ReID)**

绪论

研究背景—IVS 与 ReID

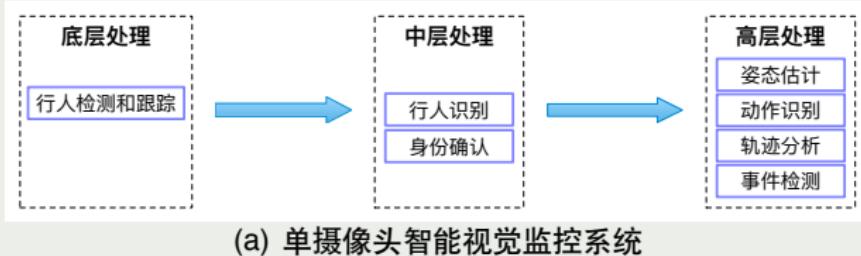


图 1: 典型的智能视觉监控系统所涉及的关键技术步骤。

绪论

研究进展—ReID 基本框架

ReID 的目的是，基于行人的视觉外观特征（如衣着、肤色、体态等），进行跨摄像头跨视域的行人身份一致性关联。

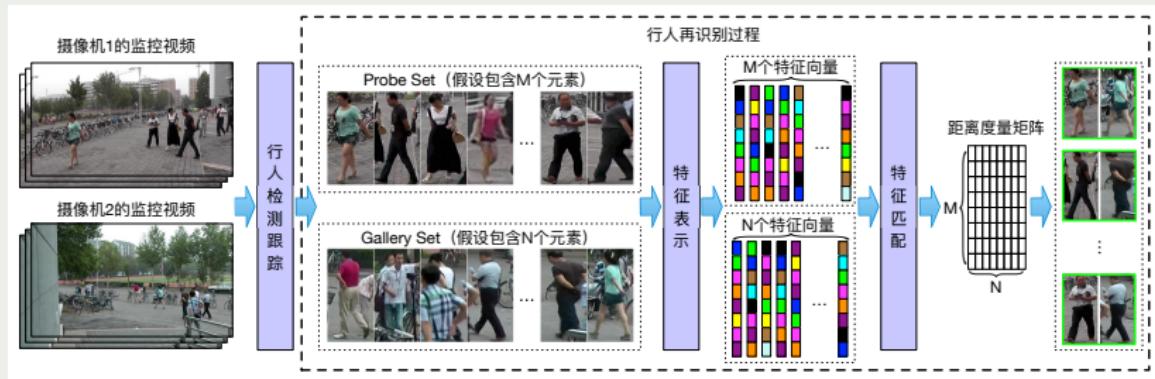


图 2: 行人再识别基本框架。图中所用到的图像样本均来自于数据集 PRW[64]。

绪论

研究进展—性能指标

- 首位准确度 (Rank-1 Accuracy, Rank-1)
- 累计匹配性能 (Cumulative Matching Characteristic, CMC),
 $cmc(N) = \sum_{n=1}^N r(n)$
- 均值平均精度 (Mean Average Precision, mAP)

研究进展—常用数据集

表 1: 常用行人再识别数据集。

数据集	# 摄像机	# 行人	# 图片	图片尺寸	单图	多图	序列	标记
VIPeR	2	632	1,264	128×48	✓	✗	✗	手工
CUHK01	2	971	3,884	160×60	✗	✓	✗	手工
3DPeS	8	192	1,011	不固定	✗	✓	✗	手工
DukeMTMC	8	1,812	36,441	不固定	✗	✓	✗	手工
PRID2011	2	934	24,541	128×64	✓	✓	✓	手工
Market1501	6	1,501	32,217	128×64	✗	✓	✗	自动
MARS	6	1,261	1,191,003	256×128	✗	✓	✓	自动

绪论

面临的挑战一特征表示

- 复杂的光照和成像条件
- 局部的背景干扰和遮挡
- 多变的拍摄角度和行人姿态
- 不固定的衣着外观

面临的挑战一特征匹配

- 类间混叠和类内偏移
- 局部不对齐
- 训练样本缺乏
- 模型泛化能力

绪论

面临的挑战—举例 (1)



(a) VIPeR



(b) CUHK01



(c) 3DPeS



(d) DukeMTMC-reID

图 3: 行人再识别常用数据集样本举例。绿色和红色矩形框代表图片来自不同摄像头, 黄色矩形框代表干扰项。

绪论

面临的挑战—举例 (2)



(a) Market1501



(b) MARS

图 4: 行人再识别常用数据集样本举例。绿色和红色矩形框代表图片来自不同摄像头, 黄色矩形框代表干扰项。

绪论

我们的主要工作

- 从手工特征设计和多特征融合分步模型的角度，我们提出了一种统一的局部统计特征提取框架，并结合多核学习实现 ReID 中的多特征融合。
- 从特征序列提取和序列匹配端到端模型的角度，我们提出了一个上下文敏感的特征序列提取以及基于双重注意力机制序列匹配的深度 ReID 模型。
- 从小数据集上模型泛化能力的角度，我们提出利用正则化的度量学习方法来增强小数据集上 ReID 模型的泛化能力。

目录

绪论

行人再识别中的相关算法

基于空间金字塔统计特征及多核学习的行人再识别算法

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于正则化度量学习的行人再识别算法

总结和展望

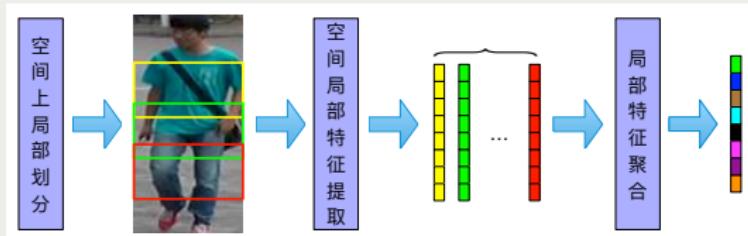
行人再识别中的相关算法

特征表示方法

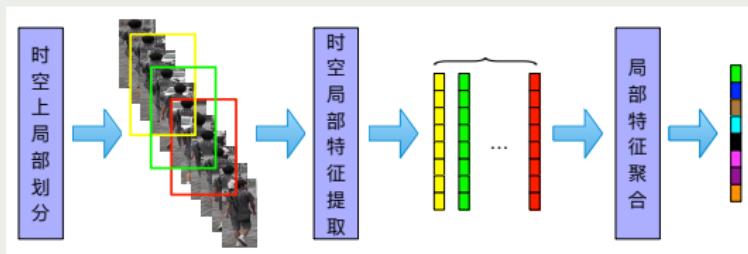
- 底层视觉特征提取
- 高层语义特征学习

行人再识别中的相关算法

特征表示方法—底层视觉特征提取



(a) 行人图像的底层视觉特征提取流程



(b) 行人跟踪序列的底层视觉特征提取流程

图 5: 底层视觉特征提取基本流程。

行人再识别中的相关算法

特征表示方法—底层视觉特征提取

- 局部划分：基于图像块、图像条、人体结构、光流能量分布图等
- 局部特征提取：颜色特征（直方图等）、纹理特征（边缘滤波、HOG/LBP/DSIFT 等）、组合特征（LOMO/GOG 等）、时空特征（HOG3D 等）等
- 局部特征聚合：不同类型特征融合、局部特征聚合成全局特征

行人再识别中的相关算法

特征表示方法—高层语义特征学习

图像的高层特征往往是对底层视觉特征的进一步抽象和提炼，产生语义相关的信息。

- 属性特征：“男性、白人、短发、贝雷帽、红色外套、灰色裤子、等等”
- 深度特征：CNN、RNN

行人再识别中的相关算法

特征匹配算法

ReID 系统对特征匹配算法的基本要求就是，保证具有同一身份的行人特征之间的距离（或相似度）小于（或大于）具有不同身份的行人特征之间的距离（或相似度）。

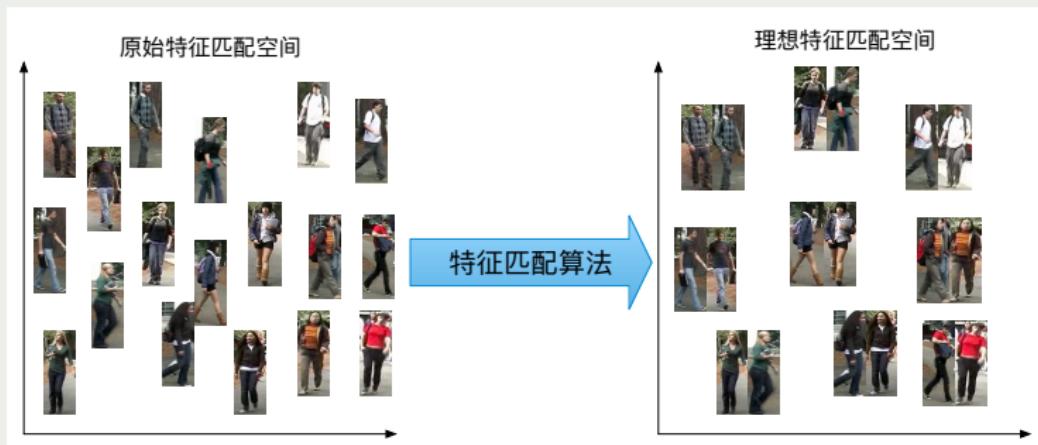


图 6: 特征匹配算法效果示意图。

行人再识别中的相关算法

特征匹配算法

- 度量学习
- 投影学习
- 局部对应关系学习

行人再识别中的相关算法

特征匹配算法—度量学习

- 度量学习是一类最直接的特征匹配算法，直接学习任务专用的距离度量方式。如果定义距离度量函数为 $d(\cdot, \cdot; \theta)$ ，度量学习的目标是，对于任意的输入样本对 (x, y) 或 (x, z) 都满足如下公式：

$$\begin{aligned}d(x, y; \theta) &< \varepsilon_1, \\d(x, z; \theta) &> \varepsilon_2,\end{aligned}$$

其中 x 和 y 为具有相同身份标签的样本对的特征表示， x 和 z 为具有不同身份标签的样本对的特征表示， ε_1 和 ε_2 分别为依据经验设定的阈值， θ 为距离度量函数的参数。

- 马氏距离： $d(x, y; M) = \sqrt{(x - y)^T M (x - y)}$

行人再识别中的相关算法

特征匹配算法—投影学习

- 对称投影 : $d(\mathbf{x}, \mathbf{y}; \mathbf{L}) = \sqrt{(\mathbf{L}\mathbf{x} - \mathbf{L}\mathbf{y})^T(\mathbf{L}\mathbf{x} - \mathbf{L}\mathbf{y})}$
- 非对称投影 : $d(\mathbf{x}, \mathbf{y}; \mathbf{L}, \mathbf{H}) = \sqrt{(\mathbf{L}\mathbf{x} - \mathbf{H}\mathbf{y})^T(\mathbf{L}\mathbf{x} - \mathbf{H}\mathbf{y})}$
- 稀疏表示 :

(1) $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_C]$, 其中 $\mathbf{D}_c = [\mathbf{x}_{c_1}, \mathbf{x}_{c_2}, \dots, \mathbf{x}_{c_{k_c}}]$;

(2) $\boldsymbol{\beta} = \arg \min_{\boldsymbol{\beta}} \|\mathbf{y} - \mathbf{D}\boldsymbol{\beta}\|_2^2 + \lambda_1 \|\boldsymbol{\beta}\|_1 + \lambda_2 r(\boldsymbol{\beta})$;

(3) \mathbf{y} 的身份标签 $= \arg \min_c \|\mathbf{y} - \mathbf{D}\delta_c(\boldsymbol{\beta})\|_2^2$;

- 耦合字典学习 :

$$\begin{aligned} \min_{\mathbf{D}_x, \mathbf{D}_y, \boldsymbol{\alpha}, \boldsymbol{\beta}} & \|\mathbf{x} - \mathbf{D}_x \boldsymbol{\alpha}\|_2^2 + \|\mathbf{y} - \mathbf{D}_y \boldsymbol{\beta}\|_2^2 + \lambda_x \|\boldsymbol{\alpha}\|_1 + \lambda_y \|\boldsymbol{\beta}\|_1 \\ & + \lambda r(\mathbf{D}_x, \mathbf{D}_y, \boldsymbol{\alpha}, \boldsymbol{\beta}), \end{aligned}$$

行人再识别中的相关算法

特征匹配算法—局部对应关系学习

- 基于位置
- 基于语义

行人再识别中的相关算法

深度神经网络算法

- 网络结构
- 损失函数

行人再识别中的相关算法

深度神经网络算法—网络结构

一般分别从特征表示和特征匹配角度去设计网络结构

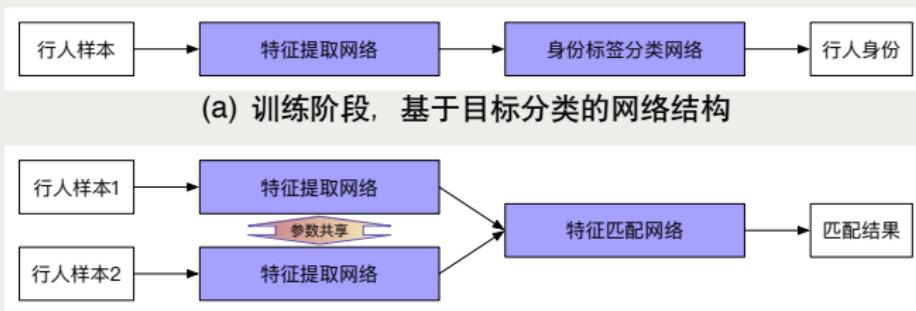
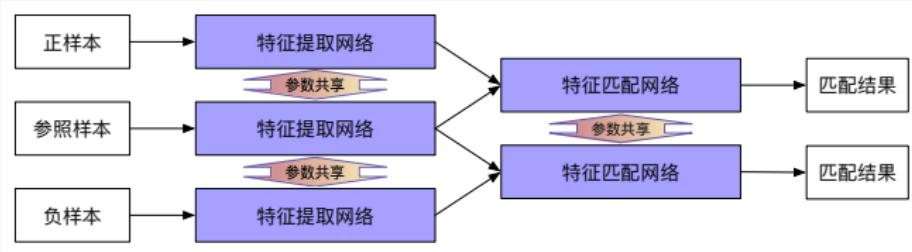


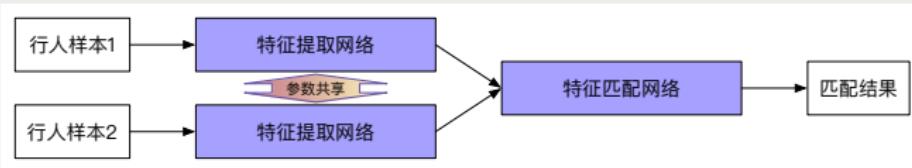
图 7: 行人再识别中常用网络结构。

行人再识别中的相关算法

深度神经网络算法—网络结构



(a) 训练阶段, 基于三元组匹配的网络结构



(b) 测试阶段网络结构

图 8: 行人再识别中常用网络结构。三元组中的正样本为与参考样本身份一致的行人, 负样本为与参考样本身份不同的行人。

行人再识别中的相关算法

深度神经网络算法—损失函数

- Cross Entropy Loss

$$f = - \sum_{c=1}^C l_c \ln (\hat{s}_c)$$

- Binary Classification Loss

$$f = -l \ln(\hat{s}_1) - (1-l) \ln(\hat{s}_2)$$

- Contrastive Loss

$$f = ld(\mathbf{x}, \mathbf{y}) + (1-l) \max(M - d(\mathbf{x}, \mathbf{y}), 0)$$

- Triplet Loss

$$f = \max(d(\mathbf{x}, \mathbf{y}) - d(\mathbf{x}, \mathbf{z}) + M, 0)$$

目录

绪论

行人再识别中的相关算法

基于空间金字塔统计特征及多核学习的行人再识别算法

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于正则化度量学习的行人再识别算法

总结和展望

基于空间金字塔统计特征及多核学习的行人再识别算法

研究动机

- 由于影响行人外观的因素众多，很难构造单一特征来全面、稳定地刻画行人外观。
- 不同特征的提取过程相对独立、缺乏系统而详细地评估分析，很难启发研究人员充分发掘特征的性能或设计其他有效的特征表示。

创新点

- 提出一种基于空间金字塔的统计特征提取框架，在此框架下实现了多种 ReID 中常用局部统计特征的提取和改进。
- 提出基于多核局部 Fisher 判别分析的进行多特征融合。
- 进行了大量的对比实验，来评估特征提取每个环节的作用以及不同的特征融合策略的有效性。

基于空间金字塔统计特征及多核学习的行人再识别算法

算法流程

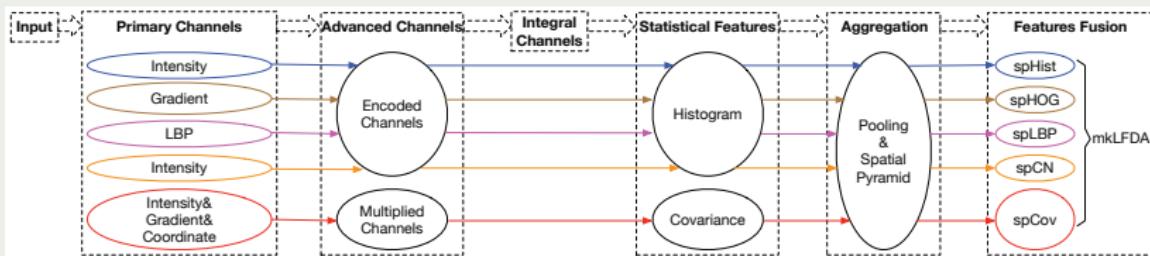


图 9: 基于空间金字塔统计特征及多核学习的行人再识别算法流程图。

完整的框架流程由生成初级特征通道 (Generating Primary Channels)、构造高级特征通道 (Constructing Advanced Channels)、提取局部区域统计特征 (Extracting Local Region Statistical Features)、局部特征聚合 (Features Aggregation,)、以及多特征融合 (Features Combination) 五个步骤组成。

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—生成初级特征通道

假设单一通道的原始图像可表示为 $\mathbf{I} \in \mathbb{R}^{H \times W}$, 其中 H 和 W 分别为图像的高度和宽度; 那么, 初级特征通道可以表示为 $\mathbf{C} = \Omega(\mathbf{I}) \in \mathbb{R}^{H \times W}$, 其中 Ω 表示某种简单的变换操作, 如论文 [11] 所提到的线性滤波、恒等变换、非线性变换、或者逐像素映射操作等。

- 亮度通道 (Intensity Channels)
- 坐标通道 (Coordinate Channels)
- 梯度通道 (Gradient Channels)
- LBP 通道 (LBP Channels)

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—构造高级特征通道

给定一系列初级特征通道 $\mathcal{C} = \{\mathbf{C}^{(m)} \in \mathbb{R}^{H \times W}, m = 1, \dots, M\}$, 则高级通道可表示为 $\mathcal{C}_A = \Omega(\mathcal{C}) \in \mathbb{R}^{H \times W \times N}$, 其中 Ω 代表某种映射操作, 比如编码或者乘积操作。

- 编码特征通道 (Encoded Image Channels)

给定一幅图像的 M 个初级特征通道组成的特征图谱以及一个定义好的码书 (Codebook)

$\mathbf{V} = \{\mathbf{v}_n \in \mathbb{R}^M, n = 1, \dots, N\}$, 图谱中每个位置的 M 维特征向量 \mathbf{f}_{ij} 都可以被编码成关于码书 \mathbf{V} 中元素的 N 个组合系数 $\{a_{ij}^n, n = 1, \dots, N\}$ 。那么, 第 n 个编码特征通道的每个像素可以表示为 $\mathbf{C}^{(n)}(i, j) = a_{ij}^n$ 。

- (1) 直方图编码 (Histogram Encoding, HE)
- (2) 核码书编码 (Kernel Codebook Encoding, KCE)
- (3) 线性插值编码 (Linear Interpolation Encoding, LIE)
- (4) 显著颜色编码 (Salient Color Encoding, SCE)

- 乘积特征通道 (Multiplied Image Channels)

给定一系列初级特征通道 $\mathcal{C} \in \mathbb{R}^{H \times W \times M}$, 乘积通道可以按照如下方式构造 :

$$\mathbf{C}^{m_1, m_2} = \mathbf{C}^{(m_1)} \odot \mathbf{C}^{(m_2)}.$$

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—特征图道可视化

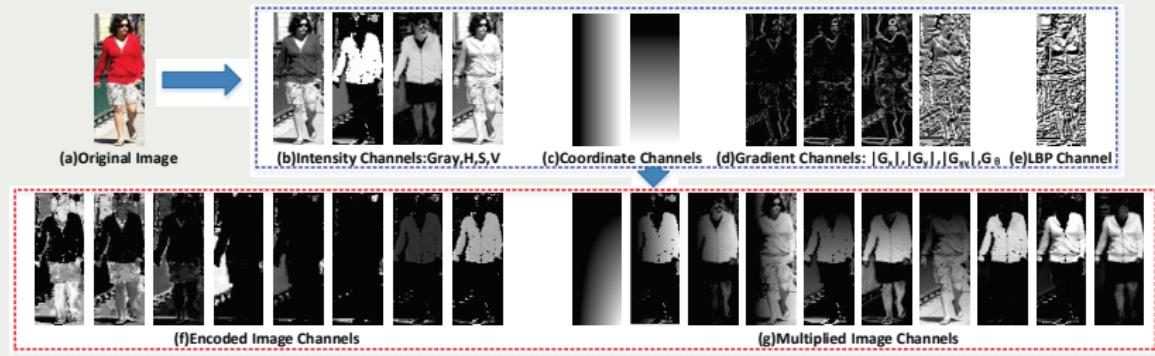


图 10: 特征通道举例。 (a) : 原始图像 ;(b) - (e) : 初级特征通道 ;(f) 和 (g) : 高级特征通道。

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—提取局部区域统计特征

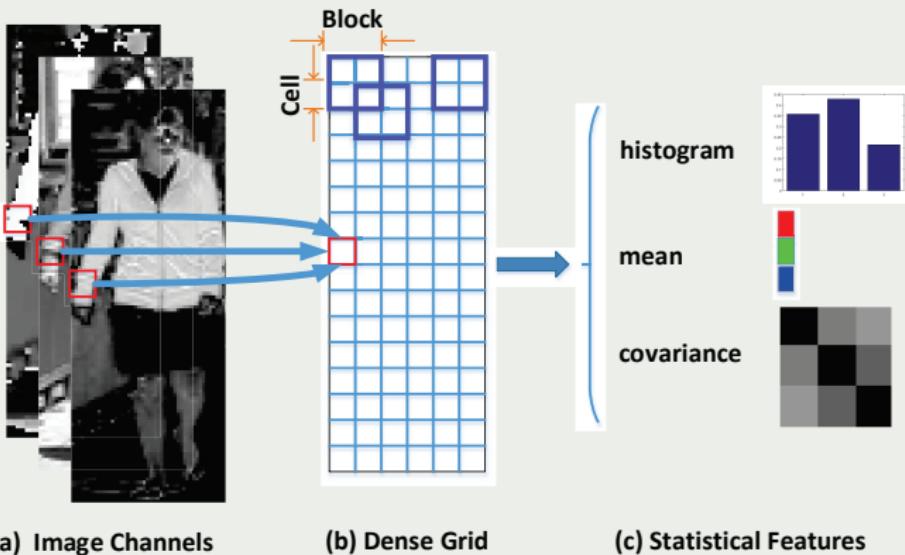


图 11: 局部统计特征提取过程示意图。通过计算不同特征通道局部区域的像素值之和, 可以得到三种不同类型的统计特征, 分别是类直方图特征、均值向量、和协方差矩阵。

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—提取局部区域统计特征（快速）

- 基于积分图通道快速提取局部区域统计特征

(0) 积分图通道为：

$$\mathbf{C}_{Intg}(i', j') = \sum_{i \leq i', j \leq j'} \mathbf{C}(i, j)$$

令 $\mathbf{p}_{i,j} = \mathcal{C}_{Intg}(i, j, :)$, $\mathbf{q}_{i,j} = \mathcal{C}_{Intg'}(i, j, :)$, 和 $\mathbf{E}_{i,j} = \mathcal{C}_{Intg''}(i, j, :, :)$ 。

(1) 均值向量：

$$\mathbf{u}_{(i', j'; i'', j'')} = \frac{1}{S} (\mathbf{p}_{i'', j''} + \mathbf{p}_{i'-1, j'-1} - \mathbf{p}_{i'-1, j''} - \mathbf{p}_{i'', j'-1})$$

(2) 直方图特征：

$$\mathbf{h}_{(i', j'; i'', j'')} = (\mathbf{q}_{i'', j''} + \mathbf{q}_{i'-1, j'-1} - \mathbf{q}_{i'-1, j''} - \mathbf{q}_{i'', j'-1})$$

(3) 协方差矩阵：

$$\mathbf{O}_{(i', j'; i'', j'')} = \frac{[\mathbf{E}_{i'', j''} + \mathbf{E}_{i'-1, j'-1} - \mathbf{E}_{i'-1, j''} - \mathbf{E}_{i'', j'-1} - S \mathbf{u}_{(i', j'; i'', j'')} \mathbf{u}_{(i', j'; i'', j'')}^T]}{S - 1}$$

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—提取局部区域统计特征（快速）

- 基于卷积积分图通道快速提取局部区域统计特征

为了克服小胞边界上像素引起的空间混叠问题，引入论文 [48] 中的卷积积分图。

基于卷积积分图通道：

$$\mathbf{C}_{Intg}(i', j') = \sum_{i \leq i', j \leq j'} \mathbf{C}_{Cov}(i, j), \text{ where } \mathbf{C}_{Cov} = \mathbf{C} \otimes \mathbf{K}$$

其中卷积核定义如下：

$$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \text{ or } \frac{1}{256} \begin{bmatrix} 1 & 2 & 3 & 4 & 3 & 2 & 1 \\ 2 & 4 & 6 & 8 & 6 & 4 & 2 \\ 3 & 6 & 9 & 12 & 9 & 6 & 3 \\ 4 & 8 & 12 & 16 & 12 & 8 & 4 \\ 3 & 6 & 9 & 12 & 9 & 6 & 3 \\ 2 & 4 & 6 & 8 & 6 & 4 & 2 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 \end{bmatrix}$$

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—提取局部区域统计特征（局部对比归一化）

- 将 2×2 个小胞 (Cell) 特征组成空间上存在一半重叠的区块 (Block) 特征：串联、均值
- 归一化

$$l_1\text{-norm} : \mathbf{x} \rightarrow \mathbf{x}/\|\mathbf{x}\|_1 + \epsilon$$

$$l_1\text{-sqrt} : \mathbf{x} \rightarrow \sqrt{\mathbf{x}/\|\mathbf{x}\|_1 + \epsilon}$$

$$l_2\text{-norm} : \mathbf{x} \rightarrow \mathbf{x}/\|\mathbf{x}\|_2 + \epsilon$$

$l_2\text{-clip}$: 在 $l_2\text{-norm}$ 基础上，限制最大值

$$l_1^2\text{-norm} : \mathbf{x} \rightarrow \mathbf{x}/\|\mathbf{x}\|_1^2 + \epsilon$$

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—局部特征多尺度聚合

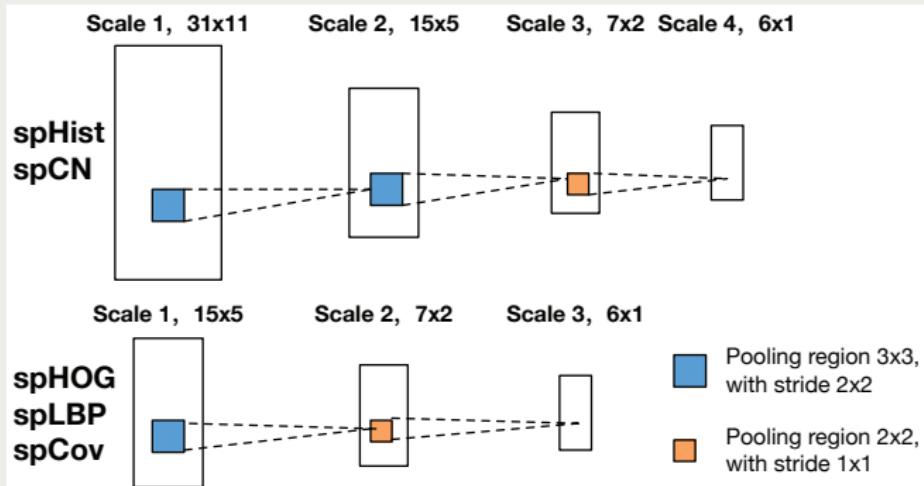


图 12: 多尺度池化操作示意图。在我们的方法中，输入图像被统一缩放到 128×48 大小，因此局部归一化以后每幅图像会产生 31×11 个基于区块的 spHist 和 spCN 特征，产生 15×5 个基于区块的 spHOG、spLBP、和 spCov 特征。

基于空间金字塔统计特征及多核学习的行人再识别算法

基于空间金字塔的统计特征提取框架—数据流可视化

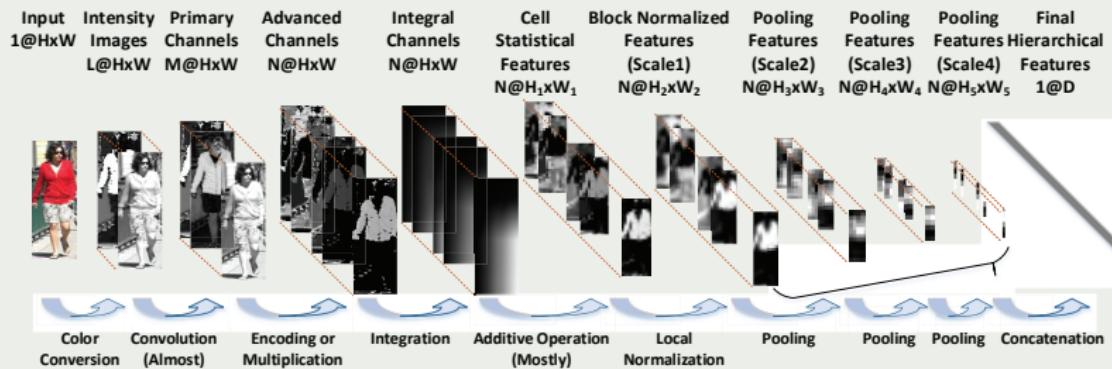


图 13: 特征提取流程中数据流的可视化。顶部的文字代表不同的数据类型, 底部的文字代表对应的操作。

基于空间金字塔统计特征及多核学习的行人再识别算法

基于多核局部 Fisher 判别分析的特征融合—核化局部 Fisher 判别分析

- 优化目标函数

$$\max_{\mathbf{A}} \frac{\text{tr}(\mathbf{A}^T \tilde{\mathbf{S}}^{(b)} \mathbf{A})}{\text{tr}(\mathbf{A}^T \tilde{\mathbf{S}}^{(w)} \mathbf{A})},$$

其中 $\text{tr}(\cdot)$ 代表矩阵的迹, $\mathbf{A} \in \mathbb{R}^{s \times d}$ 代表投影矩阵, s 表示数据集的大小。符号 $\tilde{\mathbf{S}}^{(b)}$ 和 $\tilde{\mathbf{S}}^{(w)}$ 分别表示类间和类内的局部散度矩阵, 表示为:

$$\tilde{\mathbf{S}}^{(b)} = \frac{1}{2} \sum_{i,j=1}^s \tilde{\mathbf{W}}_{ij}^{(b)} (\mathbf{k}_i - \mathbf{k}_j)(\mathbf{k}_i - \mathbf{k}_j)^T \in \mathbb{R}^{s \times s},$$

$$\tilde{\mathbf{S}}^{(w)} = \frac{1}{2} \sum_{i,j=1}^s \tilde{\mathbf{W}}_{ij}^{(w)} (\mathbf{k}_i - \mathbf{k}_j)(\mathbf{k}_i - \mathbf{k}_j)^T \in \mathbb{R}^{s \times s},$$

其中 $\mathbf{k}_i = [\kappa(\mathbf{x}_1, \mathbf{x}_i), \dots, \kappa(\mathbf{x}_s, \mathbf{x}_i)]^T \in \mathbb{R}^s$, $\tilde{\mathbf{W}}^{(b)}$ 和 $\tilde{\mathbf{W}}^{(w)}$ 分别表示样本类间和类内的局部邻接图的权重矩阵。

- 等价于广义特征值问题

$$\tilde{\mathbf{S}}^{(b)} \mathbf{A} = \lambda \tilde{\mathbf{S}}^{(w)} \mathbf{A}$$

基于空间金字塔统计特征及多核学习的行人再识别算法

基于多核局部 Fisher 判别分析的特征融合—多核局部 Fisher 判别分析

- 多核学习

$$\kappa(\mathbf{x}, \mathbf{x}') = \sum_{p=1}^P \beta_p \kappa^{(p)}(\mathbf{x}, \mathbf{x}')$$

- 多核局部 Fisher 判别分析优化目标函数

$$\min_{\mathbf{A}, \boldsymbol{\beta}} \text{tr}(\mathbf{A}^T \widetilde{\mathbf{S}}_{\boldsymbol{\beta}}^{(w)} \mathbf{A}), \text{ s.t. } \text{tr}(\mathbf{A}^T \widetilde{\mathbf{S}}_{\boldsymbol{\beta}}^{(b)} \mathbf{A}) = 1 \text{ and } \boldsymbol{\beta} \geq \mathbf{0}, \quad (1)$$

其中,

$$\boldsymbol{\beta} = [\beta_1, \dots, \beta_P]^T \in \mathbb{R}^P,$$

$$\mathbb{K}^{(i)} = [\mathbf{k}_i^{(1)}, \dots, \mathbf{k}_i^{(P)}] \in \mathbb{R}^{s \times P},$$

$$\widetilde{\mathbf{S}}_{\boldsymbol{\beta}}^{(w)} = \sum_{i,j=1}^s \frac{1}{2} \widetilde{\mathbf{W}}_{ij}^{(w)} (\mathbb{K}^{(i)} - \mathbb{K}^{(j)}) \boldsymbol{\beta} \boldsymbol{\beta}^T (\mathbb{K}^{(i)} - \mathbb{K}^{(j)})^T,$$

$$\widetilde{\mathbf{S}}_{\boldsymbol{\beta}}^{(b)} = \sum_{i,j=1}^s \frac{1}{2} \widetilde{\mathbf{W}}_{ij}^{(b)} (\mathbb{K}^{(i)} - \mathbb{K}^{(j)}) \boldsymbol{\beta} \boldsymbol{\beta}^T (\mathbb{K}^{(i)} - \mathbb{K}^{(j)})^T.$$

基于空间金字塔统计特征及多核学习的行人再识别算法

基于多核局部 Fisher 判别分析的特征融合—多核局部 Fisher 判别分析

- 优化算法：交替迭代优化

- (1) 当参数 β 固定，优化参数 A 时，mkLFDA 问题退化为 kLFDA 问题，可以直接利用广义特征值方法求解；
 (2) 当参数 A 固定，优化参数 β 时，可以将原始的 mkLFDA 问题重新构造成等价的非凸的 QCQP 问题，形式如下：

$$\begin{aligned} & \min_{\beta} \beta^T \tilde{\mathbf{S}}_A^{(w)} \beta \\ & \text{s.t. } \beta^T \tilde{\mathbf{S}}_A^{(b)} \beta = 1, \text{ and } \beta \geq 0, \end{aligned}$$

其中

$$\tilde{\mathbf{S}}_A^{(w)} = \sum_{i,j=1}^s \frac{1}{2} \widetilde{\mathbf{W}}_{ij}^{(w)} (\mathbb{K}^{(i)} - \mathbb{K}^{(j)})^T \mathbf{A} \mathbf{A}^T (\mathbb{K}^{(i)} - \mathbb{K}^{(j)}),$$

$$\tilde{\mathbf{S}}_A^{(b)} = \sum_{i,j=1}^s \frac{1}{2} \widetilde{\mathbf{W}}_{ij}^{(b)} (\mathbb{K}^{(i)} - \mathbb{K}^{(j)})^T \mathbf{A} \mathbf{A}^T (\mathbb{K}^{(i)} - \mathbb{K}^{(j)}).$$

该问题可以借助它的半正定规划松弛快速求解 [29]。

基于空间金字塔统计特征及多核学习的行人再识别算法

基于多核局部 Fisher 判别分析的特征融合—多核局部 Fisher 判别分析

- 优化算法：交替迭代优化

- (1) 当参数 β 固定，优化参数 A 时，mkLFDA 问题退化为 kLFDA 问题，可以直接利用广义特征值方法求解；
 (2) 当参数 A 固定，优化参数 β 时，可以将原始的 mkLFDA 问题重新构成等价的非凸的 QCQP 问题，形式如下：

$$\min_{\beta, B} \text{tr}(\tilde{S}_A^{(w)} B)$$

$$\text{s.t. } \text{tr}(\tilde{S}_A^{(b)} B) = 1, \beta \geq 0, \text{ and } \begin{bmatrix} 1 & \beta^T \\ \beta & B \end{bmatrix} \succeq 0,$$

其中

$$\tilde{S}_A^{(w)} = \sum_{i,j=1}^s \frac{1}{2} \widetilde{W}_{ij}^{(w)} (\mathbb{K}^{(i)} - \mathbb{K}^{(j)})^T A A^T (\mathbb{K}^{(i)} - \mathbb{K}^{(j)}),$$

$$\tilde{S}_A^{(b)} = \sum_{i,j=1}^s \frac{1}{2} \widetilde{W}_{ij}^{(b)} (\mathbb{K}^{(i)} - \mathbb{K}^{(j)})^T A A^T (\mathbb{K}^{(i)} - \mathbb{K}^{(j)}).$$

该问题可以借助它的半正定规划松弛快速求解 [29]。

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—空间金字塔统计特征的详细性能评估

- 输入图像颜色空间
- 编码方法
- 积分图通道类型
- 局部对比归一化方法
- 池化方法
- 特征尺度以及全局归一化方法

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—空间金字塔统计特征的详细性能评估

表 2: 利用不同参数提取的 spHist 特征在 VIPeR 数据集上的 Rank- r 准确率。

参数配置	Rank1	Rank5	Rank10	Rank20	Rank50
(a) spHist	35.19	66.30	80.38	91.23	98.20
(b) Hist	29.49	57.15	71.14	84.15	96.39
(c) RGB	15.66	41.27	57.41	73.70	90.82
(d) YUV	24.78	54.81	69.62	83.58	94.27
(e) LAB	26.30	55.16	70.22	83.83	95.35
(f) KCE, $\gamma = \frac{1}{24}$	34.49	62.66	75.63	87.66	96.84
(g) KCE, $\gamma = \frac{1}{6}$	33.89	65.51	79.05	89.81	98.10
(h) LIE	34.91	65.79	78.58	90.38	97.59
(i) HE	31.93	62.28	75.82	88.04	97.06
(j) Orig	35.70	65.92	79.02	89.59	98.07
(k) No normalization	31.52	63.29	77.15	89.18	98.01
(l) ℓ_1 -norm	34.62	67.18	80.51	90.76	98.23
(m) ℓ_1 -sqrt	36.74	68.20	80.63	90.66	97.94
(n) ℓ_1^2 -norm	35.66	67.85	80.92	90.47	98.13
(o) ℓ_2 -clip	34.65	67.25	80.38	90.38	98.10
(p) Avg	34.84	67.12	79.97	90.85	98.07
(q) Multi, ℓ_2	35.22	66.27	80.22	90.85	98.13

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—空间金字塔统计特征的详细性能评估

表 3: 利用不同参数提取的 spHOG 特征在 VIPeR 数据集上的 Rank- r 准确率。

参数配置	Rank1	Rank5	Rank10	Rank20	Rank50
(a) spHOG	26.58	56.96	71.39	83.67	95.66
(b) HOG	5.16	16.65	26.90	39.94	60.76
(c) Gray	5.92	20.32	30.76	46.08	68.39
(d) HSV	23.92	50.92	63.86	78.20	92.34
(e) RGB	9.49	27.75	40.38	56.96	80.82
(f) LAB	24.18	53.07	67.91	81.96	95.06
(g) KCE, $\gamma = \frac{\pi}{18}$	23.54	52.78	66.90	80.66	94.37
(h) KCE, $\gamma = \frac{\pi}{2}$	22.53	50.38	65.70	79.81	93.86
(i) LIE	23.04	52.12	66.55	80.54	94.40
(j) HE	16.52	40.89	55.60	71.87	88.45
(k) Orig	23.83	53.77	68.99	82.59	94.91
(l) No normalization	17.44	41.84	55.35	71.77	89.59
(m) ℓ_1 -norm	23.48	52.69	66.30	79.08	92.82
(n) ℓ_1 -sqrt	23.45	52.06	65.92	79.59	92.63
(o) ℓ_2 -norm	24.24	53.07	67.88	79.75	92.53
(p) ℓ_2 -clip	23.07	52.44	66.93	79.49	93.48
(q) Max	24.46	53.67	68.99	82.72	94.62

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—空间金字塔统计特征的详细性能评估

表 4: 利用不同参数提取的 spLBP 特征在 VIPeR 数据集上的 Rank- r 准确率。

参数配置	Rank1	Rank5	Rank10	Rank20	Rank50
(a) spLBP	20.32	49.97	64.08	78.99	94.08
(b) LBP	4.21	14.21	22.82	37.78	60.95
(c) Gray	4.49	17.37	29.21	45.13	69.46
(d) HSV	13.58	36.46	51.99	68.13	88.99
(e) RGB	9.72	29.62	43.83	60.98	83.77
(f) LAB	16.61	41.61	56.87	73.20	90.35
(g) Orig	19.18	47.37	61.80	76.93	93.58
(h) No normalization	14.18	38.77	54.78	72.47	92.15
(i) ℓ_1 -norm	18.32	45.19	60.32	77.15	92.63
(j) ℓ_1 -sqrt	18.01	44.91	59.62	75.44	91.77
(k) ℓ_1^2 -norm	19.18	46.74	62.50	78.35	93.54
(l) ℓ_2 -clip	18.45	45.22	61.74	76.84	93.04
(m) Avg	19.53	47.78	62.41	78.01	93.77
(n) Multi, ℓ_1 -norm	20.47	48.29	62.63	78.07	93.39
(o) Single, ℓ_2 -norm	19.11	45.47	60.92	76.27	92.47

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—空间金字塔统计特征的详细性能评估

表 5: 利用不同参数提取的 spCN 特征在 VIPeR 数据集上的 Rank- r 准确率。

参数配置	Rank1	Rank5	Rank10	Rank20	Rank50
(a) spSCN	20.13	48.96	63.99	78.42	90.92
(b) SCN	19.02	46.49	61.11	75.41	88.99
(c) HSV	21.17	49.34	63.83	77.53	92.63
(d) YUV	20.06	44.94	59.43	74.08	89.65
(e) LAB	14.56	37.25	52.88	68.73	86.74
(f) KCE, $\gamma = \frac{1}{10}$	12.25	31.30	43.83	57.85	75.47
(g) KCE, $\gamma = \frac{1}{2}$	15.60	40.70	54.84	71.90	89.59
(h) KCE, $\gamma = 1$	14.97	37.31	51.27	67.69	85.73
(i) HE	16.61	41.52	55.28	70.09	86.68
(j) Orig	20.13	48.51	63.39	77.91	90.79
(k) No normalization	16.27	43.54	58.80	74.94	88.96
(l) ℓ_1 -norm	18.89	47.25	61.20	76.74	89.94
(m) ℓ_1 -sqrt	20.70	48.64	63.20	77.75	91.46
(n) ℓ_1^2 -norm	19.97	47.94	63.07	77.97	90.16
(o) ℓ_2 -clip	19.56	47.59	61.96	76.23	89.49
(p) Avg	20.28	48.35	62.72	76.42	90.00
(q) Multi, ℓ_1 -norm	20.16	48.48	63.67	78.32	90.98

基于空间金字塔统计特征及多核学习的行人再识别算法

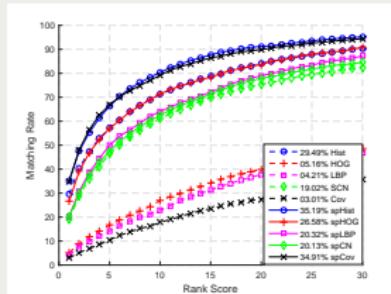
实验结果及分析—空间金字塔统计特征的详细性能评估

表 6: 利用不同参数提取的 spCov 特征在 VIPeR 数据集上的 Rank- r 准确率。

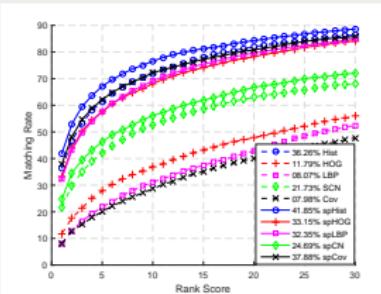
参数配置	Rank1	Rank5	Rank10	Rank20	Rank50
(a) spCov	34.91	66.80	79.15	89.87	98.20
(b) Cov	3.01	10.47	17.82	27.41	47.50
(c) Gray	4.34	14.94	23.89	37.34	61.17
(d) No normalization	18.77	47.47	63.23	80.06	94.78
(e) ℓ_1 -norm	33.10	66.23	79.18	89.87	98.29
(f) ℓ_1^2 -norm	33.42	64.27	77.91	88.45	97.34
(g) Max	34.46	65.63	79.08	88.99	97.85
(h) Without Mean	30.38	61.46	74.72	86.39	96.39
(i) Multi, ℓ_1 -norm	33.73	67.15	79.27	88.99	98.20
(j) Single, ℓ_2 -norm	31.80	63.04	77.56	88.54	97.50

基于空间金字塔统计特征及多核学习的行人再识别算法

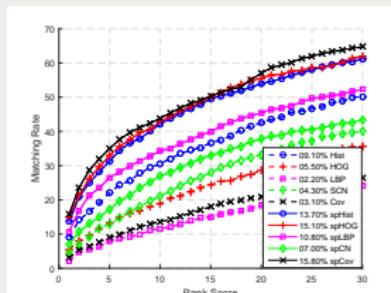
实验结果及分析—空间金字塔统计特征与原始特征的性能对比



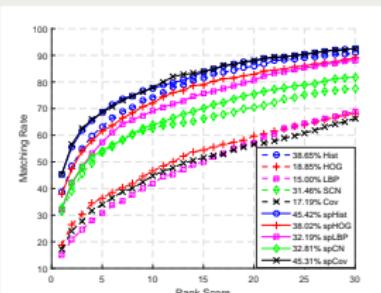
(a) VIPeR



(b) CUHK01



(c) PRID2011



(d) 3DPeS

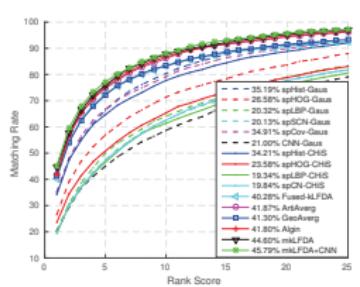
基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—空间金字塔统计特征与原始特征的性能对比

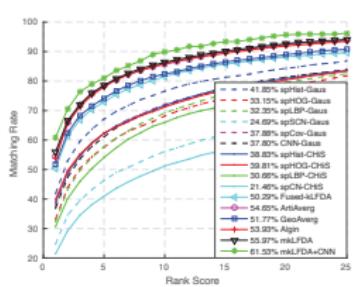
- 丰富的初级特征通道信息可以增强特征的表示能力和可辨识度
- 合适的编码方法可以减少量化误差带来的信息损失
- 恰当的局部对比归一化方法可以平滑掉特征的局部噪音变化
- 空间金字塔的层次化策略可以减轻特征的局部不对齐问题

基于空间金字塔统计特征及多核学习的行人再识别算法

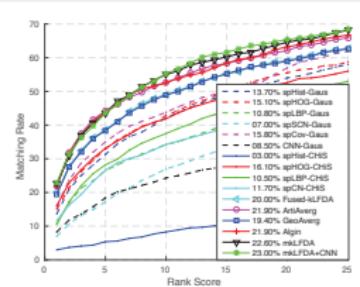
实验结果及分析—mkLFDA 与其他多核学习算法的性能对比



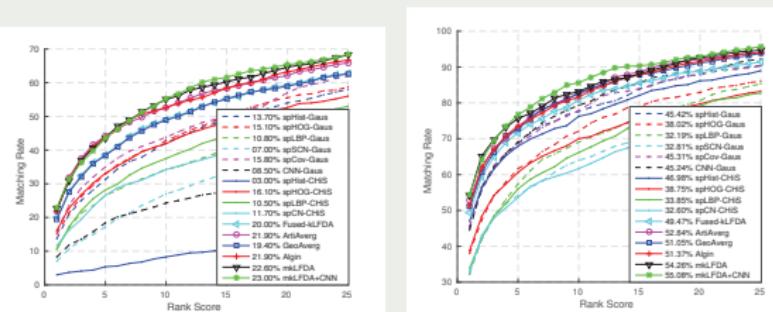
(a) VIPeR



(b) CUHK01



(c) PRID2011



(d) 3DPeS

Person Re-Identification

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—与同期其他先进 ReID 算法的性能对比

表 7: 在数据集 VIPeR 上, 基于空间金字塔统计特征和 mkLFDA 多特征融合的 ReID 算法与同期其他先进 ReID 算法的 Rank- r 准确率比较。

方法	Rank1	Rank5	Rank10	Rank20	Rank50
mkLFDA+CNN	45.8	77.3	88.4	95.7	99.9
mkLFDA	44.6	75.8	87.7	95.3	99.5
Ensemble[40]	45.9	77.5	88.9	95.8	99.5
LateFusion[63]	30.2	51.6	62.4	73.8	-
ELF[13]	12.0	31.0	41.0	58.0	-
MidFilter[61]	43.4	73.0	85.0	93.7	-
MTL[44]	42.3	72.2	81.6	89.6	-
Transfer[43]	41.6	71.9	86.2	95.1	-
SCNCD[53]	37.8	68.5	81.2	90.4	97.0
ExplicitPoly[5]	36.8	70.4	83.9	91.7	97.8
SalMatch[59]	30.2	52.0	65.5	79.2	-
Salience[60]	26.7	50.7	62.4	76.4	-
ColorInv[16]	24.2	-	57.1	69.7	87.0
ViewInv[49]	21.4	45.9	62.6	79.7	-
eBiCov[36]	20.7	42.0	56.2	68.0	-
SDALF[12]	19.9	38.9	49.4	65.7	92.2
Attribute[18]	17.4	39.0	50.8	-	86.4

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—与同期其他先进 ReID 算法的性能对比

表 8: 在数据集 VIPeR 上, 基于空间金字塔统计特征和 mkLFDA 多特征融合的 ReID 算法与同期其他先进 ReID 算法的 Rank- r 准确率比较。

方法	Rank1	Rank5	Rank10	Rank20	Rank50
mkLFDA+CNN	45.8	77.3	88.4	95.7	99.9
mkLFDA	44.6	75.8	87.7	95.3	99.5
MLAPG[26]	40.7	-	82.3	92.4	-
XQDA[25]	40.0	-	80.5	91.1	-
KernelML[51]	36.1	68.7	81.3	91.1	-
RMLLC[6]	31.3	62.1	75.3	86.7	-
LADF[24]	30.0	65.0	79.0	91.0	98.0
LAFT[22]	29.6	-	69.3	-	96.8
MtMCML[37]	28.8	59.3	75.8	88.5	-
RPLM[15]	27.0	-	69.0	83.0	95.0
FuncSpace[38]	25.8	-	69.6	83.7	95.1
LFDA[41]	24.2	-	67.1	-	94.1
KISSME[17]	22.0	-	68.0	-	93.0
PCCA[39]	19.6	48.9	64.9	80.3	-
RDC[66]	18.3	42.7	57.8	72.4	-

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—与同期其他先进 ReID 算法的性能对比

表 9: 在数据集 VIPeR 上, 基于空间金字塔统计特征和 mkLFDA 多特征融合的 ReID 算法与同期其他先进 ReID 算法的 Rank- r 准确率比较。

方法	Rank1	Rank5	Rank10	Rank20	Rank50
mkLFDA+CNN	45.8	77.3	88.4	95.7	99.9
mkLFDA	44.6	75.8	87.7	95.3	99.5
ImprDeep[2]	34.8	63.0	76.0	-	-
CPDL[20]	34.0	64.2	77.5	88.6	-
ISR[30]	27.0	-	61.0	72.0	94.1
SSCDL[32]	25.6	53.7	68.1	83.6	-
MirrorRep[9]	43.0	75.8	87.3	94.8	-
CSL[42]	34.8	68.7	82.3	91.8	96.2
CompTemp[52]	24.0	47.0	60.0	75.0	-

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—与同期其他先进 ReID 算法的性能对比

表 10: 在数据集 CUHK01 上, 基于空间金字塔统计特征和 mkLFDA 多特征融合的 ReID 算法与同期其他先进 ReID 算法的 Rank- r 准确率比较。

方法	Rank1	Rank5	Rank10	Rank20	Rank50
mkLFDA+CNN	61.5	81.7	90.0	96.3	99.1
mkLFDA	56.0	78.5	86.3	92.6	97.4
Ensemble[40]	53.4	76.4	84.4	90.7	96.4
MidFilter[61]	34.3	55.1	65.0	74.9	-
Transfer[43]	31.5	52.5	65.8	77.6	-
SalMatch[59]	28.5	46.0	56.0	-	-
ImprDeep[2]	47.5	72.0	80.0	-	-
DeepReid[23]	27.9	64.0	77.0	88.0	-
MirrorRep[9]	40.4	64.6	75.3	84.1	-

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—与同期其他先进 ReID 算法的性能对比

表 11: 在数据集 PRID2011 上, 基于空间金字塔统计特征和 mkLFDA 多特征融合的 ReID 算法与同期其他先进 ReID 算法的 Rank- r 准确率比较。

方法	Rank1	Rank5	Rank10	Rank20	Rank50
mkLFDA+CNN	23.0	43.7	55.4	65.4	78.9
mkLFDA	22.6	43.5	55.0	64.6	77.5
Ensemble[40]	17.90	39.0	49.0	62.0	-
MidFilter[61]	12.5	23.9	30.7	36.5	51.6
MTL[44]	18.0	37.4	50.1	66.6	82.3
SalMatch[59]	4.9	17.5	26.1	33.9	47.8
RPLM[15]	15.0	-	42.0	54.0	70.0
PCCA[39]	3.5	10.9	17.9	27.1	45.0
PRDC[65]	4.5	12.6	19.7	29.5	46.0

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—与同期其他先进 ReID 算法的性能对比

表 12: 在数据集 3DPeS 上, 基于空间金字塔统计特征和 mkLFDA 多特征融合的 ReID 算法与同期其他先进 ReID 算法的 Rank- r 准确率比较。

方法	Rank1	Rank5	Rank10	Rank20	Rank50
mkLFDA+CNN	55.1	76.9	86.2	93.0	99.7
mkLFDA	54.3	75.6	83.1	92.6	99.3
Ensemble[40]	53.3	77.0	85.0	92.0	-
kernelML[51]	54.0	77.7	86.0	92.4	-
LFDA[41]	33.4	-	70.0	-	95.1
PCCA[39]	41.6	70.5	81.3	90.4	-
rPCCA[51]	47.3	75.0	84.5	91.9	-
CSL[42]	57.9	81.1	89.5	93.7	-

基于空间金字塔统计特征及多核学习的行人再识别算法

实验结果及分析—算法时间复杂度分析

表 13: 在数据集 VIPeR 上, 算法运行时间复杂度分析。

Feature Extraction Phase					Matching Phase	
spHist	spHOG	spLBP	spCN	spCov	training	testing
~90ms	~80ms	~260ms	~60ms	~150ms	~20min	~260ms

目录

绪论

行人再识别中的相关算法

基于空间金字塔统计特征及多核学习的行人再识别算法

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于正则化度量学习的行人再识别算法

总结和展望

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

研究动机

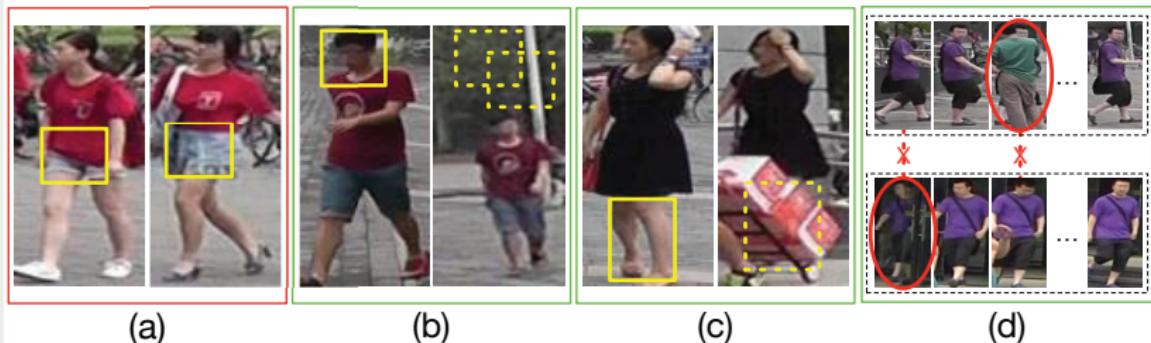


图 16: 行人再识别中的错误样例。(a) : 具有相似衣着的负样本对 ; (b) : 具有较大空间位置不对齐的正样本对 ; (c) : 具有严重身体部分遮挡的正样本对 ; (d) : 含有干扰帧 (图中的椭圆标记位置) 以及具有时间戳不对齐 (图中的“ \times ”符号标记的位置) 的正视频样本对。

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

研究动机

- 单一特征向量无法很好地保留视觉细节信息
- 特征集合或者特征序列语义不对齐

创新点

- 提出了一种新颖的端到端可训练的双重注意力匹配网络 (Dual Attention Matching Network, DuATM), 该模型可以联合学习上下文敏感的行人特征序列和实现基于双重注意机制的特征序列匹配。
- 提出利用双重注意力机制同时进行特征序列内的特征精炼和特征序列间的特征对齐。
- 将 DuATM 构造成含有三分支的孪生网络进行训练，并联合使用了 Triplet Loss、Cross Entropy Loss、和 De-Correlation Loss 进行模型优化。

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

DuATM 框架示意图

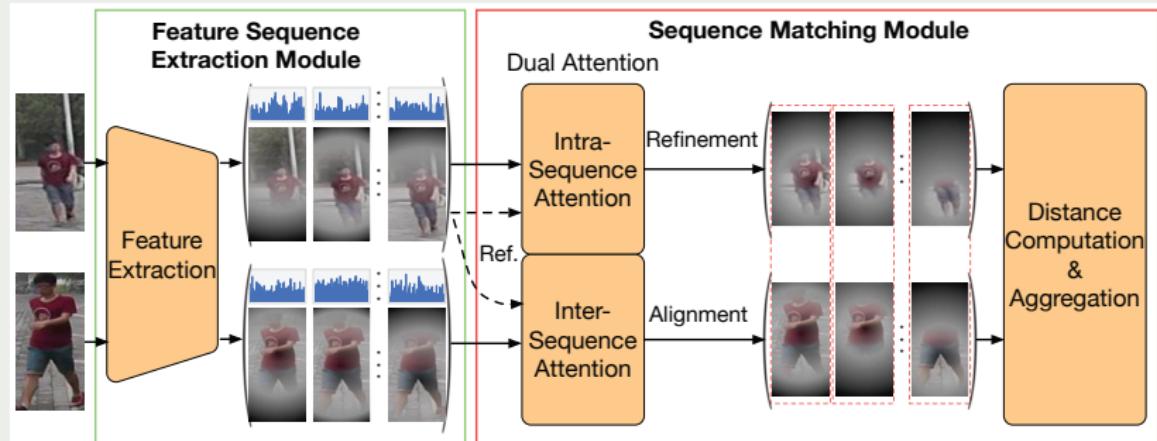
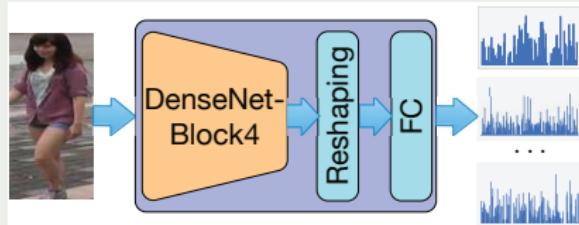


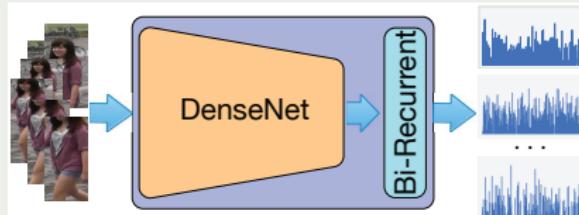
图 17: DuATM 框架示意图。

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

上下文敏感的特征序列提取模块



(a) 图像特征序列提取模块



(b) 视频特征序列提取模块

图 18: 特征序列提取模块。

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于双重注意力机制的特征序列匹配模块

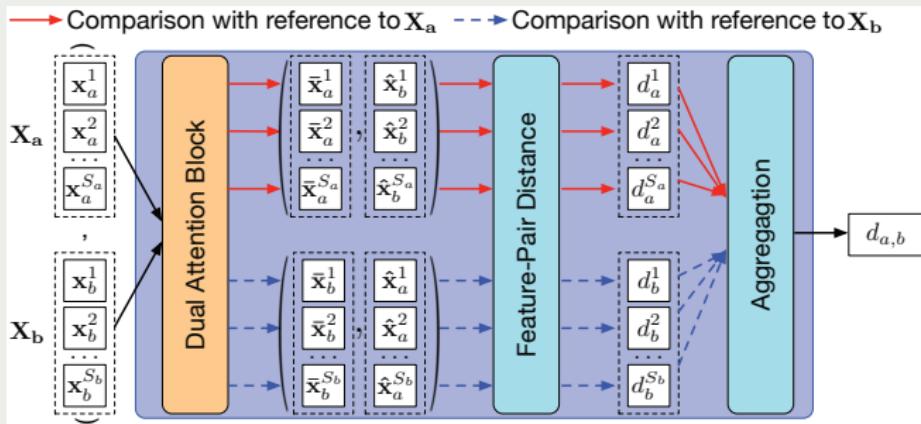


图 19: 特征序列匹配模块。

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于双重注意力机制的特征序列匹配模块—双重注意力区块

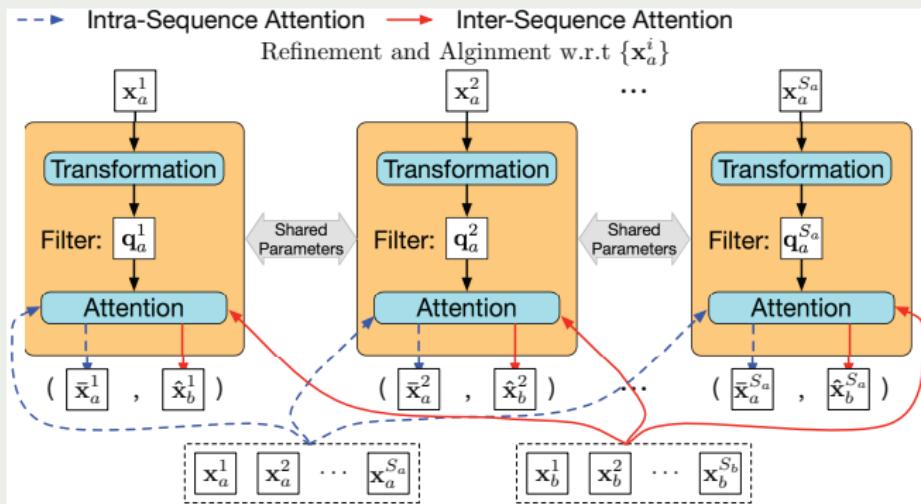


图 20: 双重注意力区块。

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于双重注意力机制的特征序列匹配模块—双重注意力区块

- 对于输入的参考特征向量，利用转换层生成对应的注意力滤波器

$$\mathbf{q}_a^i = \text{ReLU}(BN(\mathbf{W}\mathbf{x}_a^i + \mathbf{b})),$$

- 生成用于特征序列内部的去噪精炼的注意力权重、以及用于特征序列间的语义对齐的注意力权重

$$\bar{e}_a^{i,m} = \langle \mathbf{q}_a^i, \mathbf{x}_a^m \rangle, \quad \hat{e}_b^{i,n} = \langle \mathbf{q}_a^i, \mathbf{x}_b^n \rangle,$$

- 生成去噪精炼和语义对齐的特征对 $(\mathbf{x}_a^i, \hat{\mathbf{x}}_b^i)$

$$\mathbf{x}_a^i = \sum_{m=1}^{S_a} \sigma(\bar{e}_a^{i,m}) \mathbf{x}_a^m, \quad \hat{\mathbf{x}}_b^i = \sum_{n=1}^{S_b} \sigma(\hat{e}_b^{i,n}) \mathbf{x}_b^n,$$

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于双重注意力机制的特征序列匹配模块—距离计算和汇总

- 先利用欧氏距离计算每一对具有可比性的特征对的距离

$$d_a^i = \|\bar{\mathbf{x}}_a^i - \hat{\mathbf{x}}_b^i\|_2, \quad i = 1, \dots, S_a,$$

$$d_b^j = \|\bar{\mathbf{x}}_b^j - \hat{\mathbf{x}}_a^j\|_2, \quad j = 1, \dots, S_b.$$

- 利用计算均值的方式将这些局部距离汇总成序列对 $(\mathbf{X}_a, \mathbf{X}_b)$ 的整体距离

$$\|\mathbf{X}_a - \mathbf{X}_b\|_{\mathcal{M}} = \frac{1}{2S_a} \sum_{i=1}^{S_a} d_a^i + \frac{1}{2S_b} \sum_{j=1}^{S_b} d_b^j,$$

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

损失函数

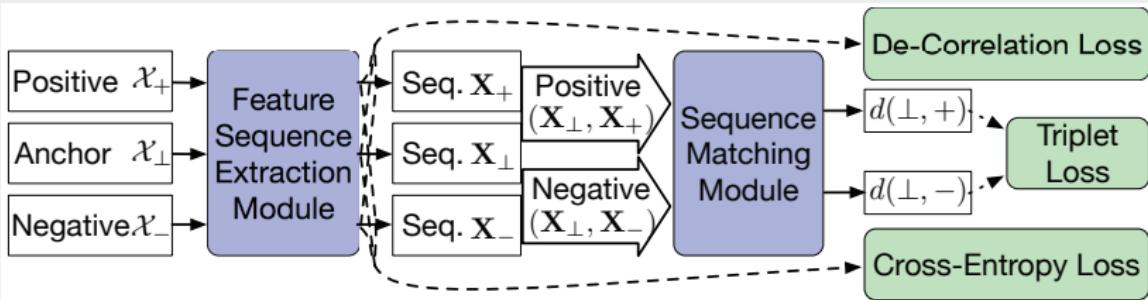


图 21: 训练阶段模型结构示意图。

$$\ell = \ell^{(0)}(\mathcal{X}, \Theta_{\mathcal{F}}, \Theta_{\mathcal{M}}) + \lambda_1 \ell^{(1)}(\mathcal{X}, \Theta_{\mathcal{F}}) + \lambda_2 \ell^{(2)}(\mathcal{X}, \Theta_{\mathcal{F}}, \boldsymbol{\theta}),$$

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

损失函数—Triplet Loss

$$\ell^{(0)}(\mathcal{X}, \Theta_{\mathcal{F}}, \Theta_{\mathcal{M}}) = \max\{0, \gamma + \|\mathcal{F}(\mathcal{X}_{\perp}) - \mathcal{F}(\mathcal{X}_+)\|_{\mathcal{M}} - \|\mathcal{F}(\mathcal{X}_{\perp}), \mathcal{F}(\mathcal{X}_{-})\|_{\mathcal{M}}\},$$

损失函数—De-Correlation Loss

$$\ell^{(1)}(\mathcal{X}, \Theta_{\mathcal{F}}) = \frac{1}{N^2} \|\mathbf{I} - \mathcal{F}(\mathcal{X})^T \mathcal{F}(\mathcal{X})\|_F^2,$$

损失函数—Cross Entropy Loss

- 将特征序列内元素组合成单一特征向量： $\mathbf{z} = \sum_{i=1}^S \omega_i \mathbf{x}_i$
- Cross Entropy Loss： $\ell^{(2)}(\mathcal{X}, \Theta_{\mathcal{F}}, \theta) = -\ln \sigma(\mathbf{w}_c \mathbf{z} + b_c)$

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 中的损失函数评估

表 14: 在数据集 Market1501 上，DuATM 与基准模型及不同损失函数下模型性能对比结果。^{*} 在这组织实验中，我们将损失函数中的超参按照参数分析中的结果调节到最优配置。^{**} 在这组实验中，我们在性能评估阶段采取了数据扩充的策略。

方法 & 损失函数	R1	R5	R20	mAP
AvePool+ $\ell^{(0)}$	74.20	89.67	95.58	56.88
DuATM+ $\ell^{(0)}$	79.66	91.15	96.73	63.46
DuATM+ $\ell^{(0)} + \ell^{(1)}$	81.83	92.46	97.33	65.21
DuATM+ $\ell^{(0)} + \ell^{(2)}$	87.50	95.37	98.01	70.02
DuATM+ $\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	88.75	95.78	98.46	70.46
DuATM*+ $\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	89.96	96.53	98.72	75.22
DuATM**+ $\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	91.42	97.09	98.96	76.62

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 中的损失函数评估

表 15: 在数据集 DukeMTMC-reID 上，DuATM 与基准模型及不同损失函数下模型性能对比结果。^{*} 在这组实验中，我们将损失函数中的超参按照参数分析中的结果调节到最优配置。^{**} 在这组实验中，我们在性能评估阶段采取了数据扩充的策略。

方法 & 损失函数	R1	R5	R20	mAP
AvePool+ $\ell^{(0)}$	64.05	79.44	87.52	43.79
DuATM+ $\ell^{(0)}$	68.40	81.73	89.77	48.65
DuATM+ $\ell^{(0)} + \ell^{(1)}$	69.17	82.23	89.36	49.48
DuATM+ $\ell^{(0)} + \ell^{(2)}$	79.40	90.04	94.25	61.55
DuATM+ $\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	81.06	91.11	95.02	62.27
DuATM*+ $\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	81.46	90.75	95.11	63.14
DuATM**+ $\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	81.82	90.17	95.38	64.58

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 中的损失函数评估

表 16: 在数据集 MARS 上, DuATM 与基准模型及不同损失函数下模型性能对比结果。* 在这组实验中, 我们将损失函数中的超参按照参数分析中的结果调节到最优配置。** 在这组实验中, 我们在性能评估阶段采取了数据扩充的策略。

方法 & 损失函数	R1	R5	R20	mAP
AvePool+ $\ell^{(0)}$	65.45	81.92	90.10	47.26
DuATM+ $\ell^{(0)}$	66.36	83.13	90.40	48.44
DuATM+ $\ell^{(0)} + \ell^{(1)}$	66.52	83.78	91.21	49.07
DuATM+ $\ell^{(0)} + \ell^{(2)}$	73.74	87.73	93.84	56.36
DuATM+ $\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	74.43	89.08	94.13	58.19
DuATM*+ $\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	76.36	90.10	95.30	58.96
DuATM**+$\ell^{(0)} + \ell^{(1)} + \ell^{(2)}$	78.74	90.86	95.76	62.26

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 的模型简化测试

表 17: 在数据集 Market1501 上, DuATM 的模型简化测试结果。

方法 & 损失函数	R1	R5	R20	mAP
AvePool+ $\ell^{(0)}$	74.20	89.67	95.58	56.88
Intra+ $\ell^{(0)}$	78.78	90.69	96.73	61.76
Inter+ $\ell^{(0)}$	72.36	87.74	95.19	53.91
DuATM+ $\ell^{(0)}$	79.66	91.15	96.73	63.46

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 中的参数评估

- 参数 λ_1 、 λ_2 、 p 评估

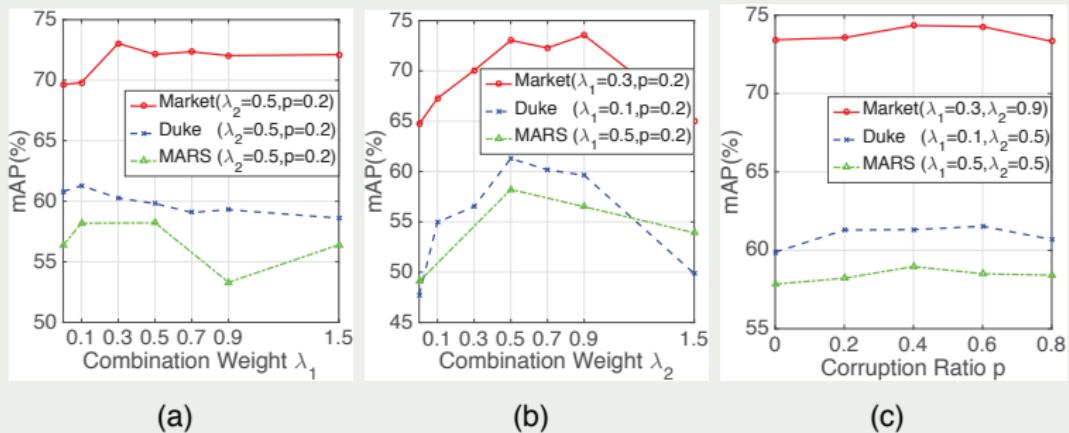


图 22: 损失函数中参数设置对模型性能影响的评估实验。

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 中的参数评估

- 特征向量维度 D 和视频序列长度 T

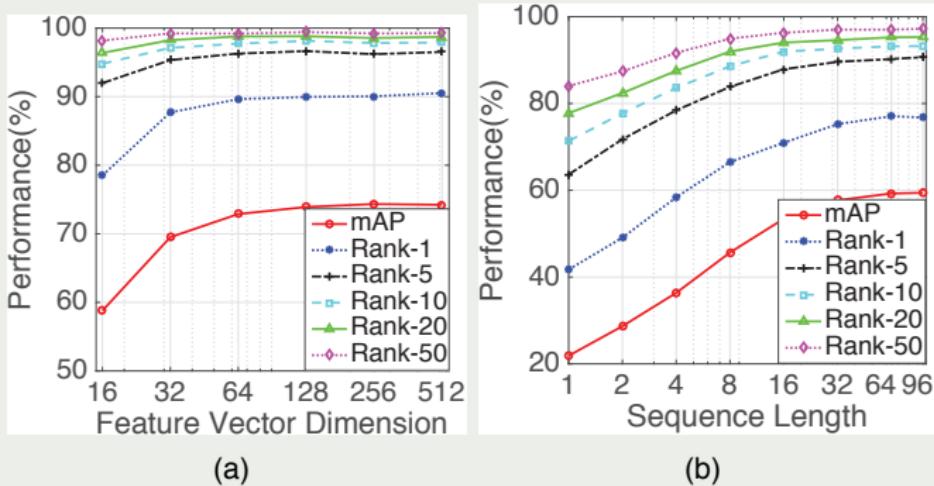


图 23: 特征向量维数和视频段产长度对模型性能影响的评估实验。

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 与其他注意力模型的性能对比

表 18: DuATM 与其他注意力模型的性能对比结果。

数据集	方法	R1	mAP	文献
Market	CAN	48.24	24.43	2017 TIP[31]
	HP-Net	76.90	-	2017 ICCV[33]
	DuATM	91.42	76.62	This paper
MARS	ST-RNN	70.60	50.70	2017 CVPR[71]
	QAN	73.74	51.70	2017 CVPR[34]
	DuATM	78.74	62.26	This paper

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 与其他特征集合或特征序列匹配模型的性能对比

表 19: DuATM 与其他特征集合或特征序列匹配模型的性能对比结果。

数据集	方法	R1	mAP	文献
Market	SCSP	51.90	26.35	2016 CVPR[4]
	SpindleNet	76.90	-	2017 CVPR[57]
	DLPAR	81.00	63.40	2017 ICCV[58]
	DRL-PL	88.20	69.30	2017 Arxiv[54]
	DuATM	91.42	76.62	This paper

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 与同期最先进算法的性能对比

表 20: 在数据集 Market1501 上, DuATM 与同期其他先进 ReID 算法的性能对比。

方法	R1	R5	mAP	文献
BOW	44.42	63.90	20.76	2015 ICCV[62]
LDNS	61.02	-	35.68	2016 CVPR[56]
Re-Rank	77.11	-	63.63	2017 CVPR[68]
SSM	82.21	-	68.80	2017 CVPR[3]
S-LSTM	61.60	-	35.30	2016 ECCV[47]
G-CNN	65.88	-	39.55	2016 ECCV[46]
CRAFT	68.70	-	42.30	2017 TPAMI[10]
P2S	70.72	-	44.27	2017 CVPR[70]
CADL	73.84	-	47.11	2017 CVPR[27]
USG-GAN	78.06	-	56.23	2017 ICCV[67]
LDCAF	80.31	-	57.53	2017 CVPR[19]
SVDNet	82.30	92.30	62.10	2017 ICCV[45]
TriNet	84.92	94.21	69.14	2017 Arxiv[14]
JLML	85.10	-	65.50	2017 IJCAI[21]
DarkRank	89.80	-	74.30	2017 Arxiv[7]
DuATM	91.42	97.09	76.62	This paper

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—DuATM 与同期最先进算法的性能对比

表 21: 在数据集 DukeMTMC-reID 上, DuATM 与同期其他先进 ReID 算法的性能对比。

方法	R1	R5	mAP	文献
BOW	25.13	-	12.17	2015 ICCV[62]
LOMO	30.75	-	17.04	2015 CVPR[25]
USG-GAN	67.68	-	47.13	2017 ICCV[67]
OIM	68.10	-	-	2017 CVPR[50]
APR	70.69	-	51.88	2017 Arxiv[28]
SVDNet	76.70	86.40	56.80	2017 ICCV[45]
DPFL	79.20	-	60.60	2017 ICCVW[8]
REDA	79.31	-	62.44	2017 Arxiv[69]
DuATM	81.82	90.17	64.58	This paper

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

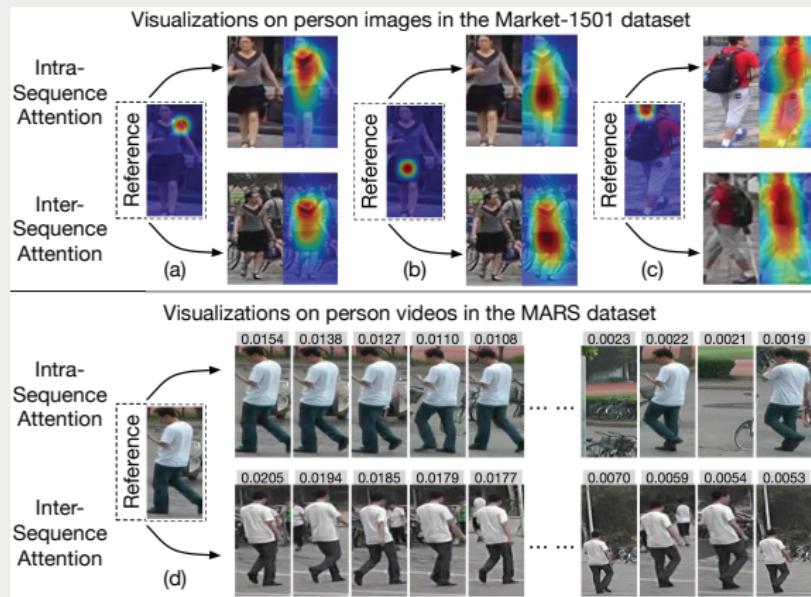
实验结果及分析—DuATM 与同期最先进算法的性能对比

表 22: 在数据集 MARS 上, DuATM 与同期其他先进 ReID 算法的性能对比。DuATM*: 采用与论文 [14] 一样的图像尺寸 256×128 重新训练模型。

方法	R1	R5	mAP	文献
SMP	23.59	35.81	10.54	2017 ICCV[35]
BOW	30.60	46.20	15.50	2015 ICCV[62]
DGM	36.80	54.00	21.30	2017 ICCV[55]
Re-Rank	73.93	-	68.45	2017 CVPR[68]
IDE	65.10	81.10	45.60	2016 ECCV[1]
LDCAF	71.77	86.57	56.50	2017 CVPR[19]
TriNet	79.80	91.36	67.70	2017 Arxiv[14]
DuATM	78.74	90.86	62.26	This paper
DuATM*	81.16	92.47	67.73	This paper

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

实验结果及分析—双重注意力机制的可视化



目录

绪论

行人再识别中的相关算法

基于空间金字塔统计特征及多核学习的行人再识别算法

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于正则化度量学习的行人再识别算法

总结和展望

基于正则化度量学习的行人再识别算法

研究动机

在实际应用中，用于度量学习训练的数据集往往比较小，因此很容易使所学模型趋于过拟合的状态，因此降低了实际应用中 ReID 算法的稳定性和准确性。

创新点

我们打算通过向度量学习模型中引入正则化约束条件的方式，来限制模型的复杂度，从而提升模型在小数据集上的泛化能力。具体来说，我们将从马氏距离学习、对称投影学习、以及非对称投影学习三个角度理解度量学习，并结合对具体学习算法实例的正则化，来证明正则化度量学习的有效性。

基于正则化度量学习的行人再识别算法

不同形式的度量函数—马氏距离

$$d_{\mathbf{M}}^2(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^T \mathbf{M} (\mathbf{x} - \mathbf{y}),$$

不同形式的度量函数—对称投影

$$d_{\mathbf{M}}^2(\mathbf{x}, \mathbf{y}) = (\mathbf{Lx} - \mathbf{Ly})^T (\mathbf{Lx} - \mathbf{Ly}).$$

不同形式的度量函数—非对称投影

$$d_{\mathbf{M}}^2(\mathbf{x}, \mathbf{y}) = (\mathbf{Lx} - \mathbf{Hy})^T (\mathbf{Lx} - \mathbf{Hy}),$$

基于正则化度量学习的行人再识别算法

正则化度量学习算法—正则化的最大间隔最近邻算法

- 最大间隔最近邻算法 (LMNN)

$$\begin{aligned} & \min_{\mathbf{M} \in \mathbb{S}_+^d} (1 - \mu) \sum_{i, j \rightsquigarrow i} d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_j) + \mu \sum_{i, j \rightsquigarrow i, l} (1 - g_{il}) \xi_{ijl}, \\ & \text{s.t. } d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_l) - d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_j) \geq 1 - \xi_{ijl}, \quad \xi_{ijl} \geq 0, \end{aligned}$$

基于正则化度量学习的行人再识别算法

正则化度量学习算法—正则化的最大间隔最近邻算法

- 核范数正则化的最大间隔最近邻算法 (nuLMNN): 促使 LMNN 学习到一个低秩的 PSD 矩阵 \mathbf{M} , 进行特征选择

$$\begin{aligned} & \min_{\mathbf{M} \in \mathbb{S}_{+i, j \rightsquigarrow i}^d} (1 - \mu) d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_j) + \sum_{i, j \rightsquigarrow i, l} \mu(1 - g_{il}) \xi_{ijl} + \lambda \|\mathbf{M}\|_* \\ & \text{s.t. } d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_l) - d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_j) \geq 1 - \xi_{ijl}, \quad \xi_{ijl} \geq 0, \end{aligned}$$

基于正则化度量学习的行人再识别算法

正则化度量学习算法—正则化的最大间隔最近邻算法

- LogDet 散度正则化的最大间隔最近邻算法 (trLMNN): 促使模型学习到的距离度量更简单

$$\begin{aligned} & \min_{\mathbf{M} \in \mathbb{S}_{+i,j \rightsquigarrow i}^d} (1 - \mu) d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_j) + \sum_{i, j \rightsquigarrow i, l} \mu(1 - g_{il}) \xi_{ijl} + \lambda D_{ld}(\mathbf{M}, \mathbf{M}_0) \\ & \text{s.t. } d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_l) - d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_j) \geq 1 - \xi_{ijl}, \quad \xi_{ijl} \geq 0. \end{aligned}$$

其中 $D_{ld}(\mathbf{M}, \mathbf{M}_0) = \text{tr}(\mathbf{M}\mathbf{M}_0^{-1}) - \log \det(\mathbf{M}\mathbf{M}_0^{-1}) - d$ 。

基于正则化度量学习的行人再识别算法

正则化度量学习算法—正则化的最大间隔最近邻算法

- nuLMNN 和 trLMNN 的优化算法

$$\mathbf{G}_t = (1 - \mu) \sum_{i, j \sim i} \mathbf{C}_{ij} + \mu \sum_{i, j \sim i, l} (\mathbf{C}_{ij} - \mathbf{C}_{il}) + \lambda \mathbf{I}$$

$$\mathbf{G}_t = (1 - \mu) \sum_{i, j \sim i} \mathbf{C}_{ij} + \mu \sum_{i, j \sim i, l} (\mathbf{C}_{ij} - \mathbf{C}_{il}) + \lambda (\mathbf{I} - \det(\mathbf{M}_t) \mathbf{M}_t^{-1})$$

其中 $\mathbf{C}_{ij} = (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T$ 。由于矩阵 \mathbf{M}_t 有时可能是不可逆的，因此我们常常用 $\widehat{\mathbf{M}}_t = (1 - \alpha)\mathbf{M}_t + \frac{\alpha}{N}\text{tr}(\mathbf{M}_t)\mathbf{I}$ 代替原始 \mathbf{M}_t ，其中 $0 \leq \alpha \leq 1$ ， N 为样本数。

基于正则化度量学习的行人再识别算法

正则化度量学习算法—正则化的 Fisher 判别分析算法

- Fisher 判别分析 (FDA)

$$\max_{\mathbf{L} \in \mathbb{R}^{k \times d}} \frac{\text{tr}(\mathbf{L}\mathbf{S}^b\mathbf{L}^T)}{\text{tr}(\mathbf{L}\mathbf{S}^w\mathbf{L}^T)},$$

- 稳定 Fisher 判别分析 (sFDA)

$$\widehat{\mathbf{S}}^w = (1 - \alpha)\mathbf{S}^w + \frac{\alpha}{N}\text{tr}(\mathbf{S}^w)\mathbf{I},$$

- sFDA 优化算法：广义特征值问题

基于正则化度量学习的行人再识别算法

正则化度量学习算法—正则化的决策函数学习算法

- 决策函数

$$f(\mathbf{x}_i, \mathbf{x}_j) = d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_j) - t,$$

- 改进的决策函数

(1) $d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{x}_j)$ 选择非对称投影形式

(2) t 定义为二次型函数

$$t(\mathbf{x}_i, \mathbf{y}_j) = \frac{1}{2} \mathbf{x}_i^T \tilde{\mathbf{A}} \mathbf{x}_i + \frac{1}{2} \mathbf{y}_j^T \tilde{\mathbf{B}} \mathbf{y}_j + \mathbf{x}_i^T \tilde{\mathbf{C}} \mathbf{y}_j + \mathbf{w}^T (\mathbf{x}_i + \mathbf{y}_j) + \mathbf{w}_0,$$

(3) 改进的决策函数

$$\begin{aligned} f(\mathbf{x}_i, \mathbf{y}_j) &= d_{\mathbf{M}}^2(\mathbf{x}_i, \mathbf{y}_j) - t(\mathbf{x}_i, \mathbf{y}_j) \\ &= \frac{1}{2} \mathbf{x}_i^T \mathbf{A} \mathbf{x}_i + \frac{1}{2} \mathbf{y}_j^T \mathbf{B} \mathbf{y}_j - \mathbf{x}_i^T \mathbf{C} \mathbf{y}_j - \mathbf{w}^T (\mathbf{x}_i + \mathbf{y}_j) - \mathbf{w}_0, \end{aligned}$$

基于正则化度量学习的行人再识别算法

正则化度量学习算法—正则化的决策函数学习算法

- 正则化的决策函数学习 (regularized Decision Function Learning, rDFL)

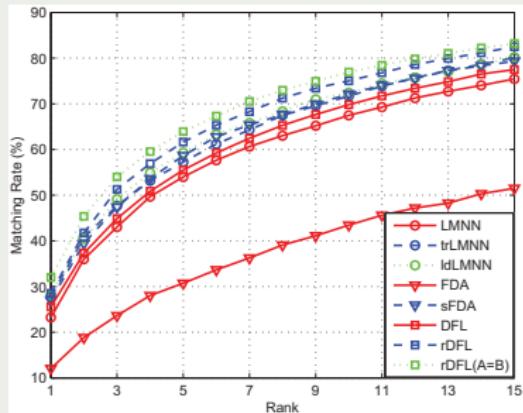
$$\min_{\Theta} E(\Theta) = \sum_i \sum_j h_{\beta} (g_{ij} f(\mathbf{x}_i, \mathbf{y}_j)) + r(\Theta),$$

其中, $\Theta = \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{w}\}$, $r(\Theta) = \lambda_1 \|\mathbf{A}\|_F^2 + \lambda_2 \|\mathbf{B}\|_F^2 + \lambda_3 \|\mathbf{C}\|_F^2$ 。

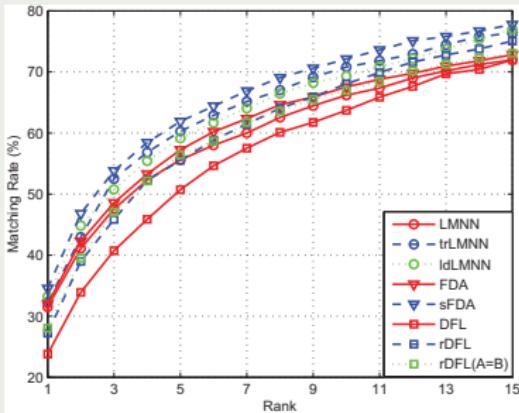
- 优化算法 : 利用梯度下降法, 交替迭代优化 \mathbf{A} 、 \mathbf{B} 、 \mathbf{C} 、以及 \mathbf{w} 。

基于正则化度量学习的行人再识别算法

正实验结果及分析—与基准模型的性能对比



(a) VIPeR



(b) CUHK01

图 25: 基于正则化度量学习的方法与基准模型的性能对比。

基于正则化度量学习的行人再识别算法

正实验结果及分析—与同期其他先进 ReID 模型的性能对比

表 23: 在数据集 VIPeR 上, 基于正则化度量学习方法与同期其他先进 ReID 模型的性能对比。

方法	Rank1	Rank5	Rank10	Rank20
SCNCD[53]	33.7%	62.7%	74.8%	85%
SDALF[12]	19.9%	38.9%	49.4%	65.7%
SDC[60]	26.7%	50.7%	62.4%	76.4%
kMFA[51]	31.1%	65.2%	79.6%	90.2%
LFDA[51]	21.5%	49.6%	64.6%	79.1%
LADF[51]	30.1%	63.2%	77.4%	88.1%
LMNN	23.2%	54.0%	67.5%	80.8%
trLMNN	27.8%	57.3%	71.7%	84.0%
IdLMNN	28.2%	59.3%	72.4%	84.6%
FDA	12.1%	30.7%	43.5%	57.7%
sFDA	27.0%	58.7%	72.2%	83.8%
DFL	25.6%	55.5%	69.8%	81.9%
rDFL	28.5%	61.7%	75.0%	87.0%
rDFL(A=B)	32.0%	63.9%	77.0%	88.6%

目录

绪论

行人再识别中的相关算法

基于空间金字塔统计特征及多核学习的行人再识别算法

基于上下文敏感特征序列及双重注意力匹配的行人再识别网络

基于正则化度量学习的行人再识别算法

总结和展望

总结和展望

总结

- 提出了一种统一的局部统计特征提取框架，并结合多核学习实现 ReID 中的多特征融合
- 提出了一个端到端的上下文信息敏感的特征序列提取以及基于双重注意力机制的序列匹配 ReID 模型
- 提出利用正则化的度量学习方法来增强小数据集上 ReID 模型的泛化能力

展望

- 利用丰富的无标签或者弱标签数据，进一步提升 ReID 算法的泛化能力
- 设计高效以及可以在超大规模数据集上执行的 ReID 算法
- 联合行人检测、行人跟踪、行人再识别等不同智能视觉监控阶段，实现整体监控系统的性能最优

感谢各位专家、老师、和同学的
聆听!

-  **MARS: A Video Benchmark for Large-Scale Person Re-identification, 2016.**
-  **E. Ahmed, M. Jones, and T. K. Marks.**
An improved deep learning architecture for person re-identification.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3908–3916, June 2015.
-  **S. Bai, X. Bai, and Q. Tian.**
Scalable person re-identification on supervised smoothed manifold.
pages 3356–3365, 2017.
-  **D. Chen, Z. Yuan, B. Chen, and N. Zheng.**
Similarity learning with spatial constraints for person re-identification.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1268–1277, June 2016.
-  **D. Chen, Z. Yuan, G. Hua, N. Zheng, and J. Wang.**
Similarity learning on an explicit polynomial kernel feature map for person re-identification.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1565–1573, June 2015.



J. Chen, Z. Zhang, and Y. Wang.

Relevance metric learning for person re-identification by exploiting listwise similarities.

IEEE Transactions on Image Processing (TIP), 24(12):4741–4755, Dec 2015.



Y. Chen, N. Wang, and Z. Zhang.

Darkrank: Accelerating deep metric learning via cross sample similarities transfer.

arXiv, 2017.



Y. Chen, X. Zhu, and S. Gong.

Person re-identification by deep learning multi-scale representations.

In *Proc. IEEE International Conference on Computer Vision Workshop (ICCVW)*, 2017.



Y.-C. Chen, W.-S. Zheng, and J. Lai.

Mirror representation for modeling view-specific transform in person re-identification.

In *Proc. International Joint Conference on Artificial Intelligence (IJCAI)*, pages 3402–3408, 2015.

-  Y. C. Chen, X. Zhu, W. S. Zheng, and J. H. Lai.
Person re-identification by camera correlation aware feature augmentation.
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 40(2):392–408, Feb 2018.
-  P. Dollar, Z. Tu, P. Perona, and S. Belongie.
Integral channel features.
In *Proc. British Machine Vision Conference (BMVC)*, pages 91.1–91.11, 2009.
-  M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani.
Person re-identification by symmetry-driven accumulation of local features.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2360–2367, 2010.
-  D. Gray and H. Tao.
Viewpoint invariant pedestrian recognition with an ensemble of localized features.
In *Proc. European Conference on Computer Vision (ECCV)*, pages 262–275, 2008.

-  A. Hermans, L. Beyer, and B. Leibe.
In defense of the triplet loss for person re-identification.
CoRR, abs/1703.07737, 2017.
-  M. Hirzer, P. M. Roth, M. Köstinger, and H. Bischof.
Relaxed pairwise learned metric for person re-identification.
In *Proc. European Conference on Computer Vision (ECCV)*, pages 780–793, 2012.
-  I. Kviatkovsky, A. Adam, and E. Rivlin.
Color invariants for person reidentification.
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 35(7):1622–1634, July 2013.
-  M. Köstinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof.
Large scale metric learning from equivalence constraints.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2288–2295, 2012.
-  R. Layne, T. M. Hospedales, and S. Gong.
Person re-identification by attributes.

In Proc. British Machine Vision Conference (BMVC), pages 24.1–24.11, 2012.

 D. Li, X. Chen, Z. Zhang, and K. Huang.

Learning deep context-aware features over body and latent parts for person re-identification.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 7398–7407, July 2017.

 S. Li, M. Shao, and Y. Fu.

Cross-view projective dictionary learning for person re-identification.

In Proc. International Joint Conference on Artificial Intelligence (IJCAI), pages 2155–2161, 2015.

 S. Li, T. Xiao, H. Li, B. Zhou, D. Yue, and X. Wang.

Person search with natural language description.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5187–5196, 2017.

 W. Li and X. Wang.

Locally aligned feature transforms across views.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3594–3601, June 2013.

-  W. Li, R. Zhao, T. Xiao, and X. Wang.
Deepreid: Deep filter pairing neural network for person re-identification.
In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 152–159, June 2014.
-  Z. Li, S. Chang, F. Liang, T. S. Huang, L. Cao, and J. R. Smith.
Learning locally-adaptive decision functions for person verification.
In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3610–3617, June 2013.
-  S. Liao, Y. Hu, X. Zhu, and S. Z. Li.
Person re-identification by local maximal occurrence representation and metric learning.
In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2197–2206, June 2015.
-  S. Liao and S. Z. Li.
Efficient psd constrained asymmetric metric learning for person re-identification.
In Proc. IEEE International Conference on Computer Vision (ICCV), pages 3685–3693, Dec 2015.



J. Lin, L. Ren, J. Lu, J. Feng, and J. Zhou.

Consistent-aware deep learning for person re-identification in a camera network.

In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5771–5780, 2017.



Y. Lin, L. Zheng, Z. Zheng, Y. Wu, and Y. Yang.

Improving person re-identification by attribute and identity learning.
arXiv, 2017.



Y. Y. Lin, T. L. Liu, and C. S. Fuh.

Multiple kernel learning for dimensionality reduction.

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 33(6):1147–1160, June 2011.



G. Lisanti, I. Masi, A. D. Bagdanov, and A. D. Bimbo.

Person re-identification by iterative re-weighted sparse ranking.

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 37(8):1629–1642, Aug 2015.



H. Liu, J. Feng, M. Qi, J. Jiang, and S. Yan.

End-to-end comparative attention networks for person re-identification.

IEEE Transactions on Image Processing (TIP), 26(7):3492–3506, July 2017.

-  X. Liu, M. Song, D. Tao, X. Zhou, C. Chen, and J. Bu.
Semi-supervised coupled dictionary learning for person re-identification.
In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3550–3557, June 2014.
-  X. Liu, H. Zhao, M. Tian, L. Sheng, J. Shao, S. Yi, J. Yan, and X. Wang.
Hydraplus-net: Attentive deep features for pedestrian analysis.
In Proc. IEEE International Conference on Computer Vision (ICCV), pages 350–359, Oct 2017.
-  Y. Liu, J. Yan, and W. Ouyang.
Quality aware network for set to set recognition.
In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 4694–4703, July 2017.
-  Z. Liu, D. Wang, and H. Lu.
Stepwise metric promotion for unsupervised video person re-identification.
In Proc. IEEE International Conference on Computer Vision (ICCV), pages 2448–2457, Oct 2017.



B. Ma, Y. Su, and F. Jurie.

Bicov: a novel image representation for person re-identification and face verification.

In *Proc. British Machine Vision Conference (BMVC)*, page 11 pages, Guildford, United Kingdom, Sep 2012.



L. Ma, X. Yang, and D. Tao.

Person re-identification over camera networks using multi-task distance metric learning.

IEEE Transactions on Image Processing (TIP), 23(8):3656–3670, Aug 2014.



N. Martinel, A. Das, C. Micheloni, and A. K. Roy-Chowdhury.

Re-identification in the function space of feature warps.

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 37(8):1656–1669, Aug 2015.



A. Mignon and F. Jurie.

Pcca: A new approach for distance learning from sparse pairwise constraints.

In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2666–2672, June 2012.

-  S. Paisitkriangkrai, C. Shen, and A. van den Hengel.
Learning to rank in person re-identification with metric ensembles.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1846–1855, June 2015.
-  S. Pedagadi, J. Orwell, S. Velastin, and B. Boghossian.
Local fisher discriminant analysis for pedestrian re-identification.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3318–3325, June 2013.
-  Y. Shen, W. Lin, J. Yan, M. Xu, J. Wu, and J. Wang.
Person re-identification with correspondence structure learning.
In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 3200–3208, Dec 2015.
-  Z. Shi, T. M. Hospedales, and T. Xiang.
Transferring a semantic representation for person re-identification and search.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4184–4193, June 2015.
-  C. Su, F. Yang, S. Zhang, Q. Tian, L. S. Davis, and W. Gao.

Multi-task learning with low rank attribute embedding for person re-identification.

In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 3739–3747, Dec 2015.

 Y. Sun, L. Zheng, W. Deng, and S. Wang.

Svdnet for pedestrian retrieval.

In *Proc. IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

 R. R. Varior, M. Haloi, and G. Wang.

Gated siamese convolutional neural network architecture for human re-identification.

In *Proc. European Conference on Computer Vision (ECCV)*, 2016.

 R. R. Varior, B. Shuai, J. Lu, D. Xu, and G. Wang.

A siamese long short-term memory architecture for human re-identification.

In *Proc. European Conference on Computer Vision (ECCV)*, 2016.

 X. Wang, T. X. Han, and S. Yan.

An hog-lbp human detector with partial occlusion handling.

In Proc. IEEE International Conference on Computer Vision (ICCV), pages 32–39, Sept 2009.

 Z. Wu, Y. Li, and R. J. Radke.

Viewpoint invariant human re-identification in camera networks using pose priors and subject-discriminative features.

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 37(5):1095–1108, May 2015.

 T. Xiao, S. Li, B. Wang, L. Lin, and X. Wang.

Joint detection and identification feature learning for person search.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3376–3385, 2017.

 F. Xiong, M. Gou, O. Camps, and M. Sznajer.

Person re-identification using kernel-based metric learning methods.

In Proc. European Conference on Computer Vision (ECCV), pages 1–16, 2014.

 Y. Xu, L. Lin, W. S. Zheng, and X. Liu.

Human re-identification by matching compositional template with cluster sampling.

In Proc. IEEE International Conference on Computer Vision (ICCV),
pages 3152–3159, Dec 2013.

 Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Z. Li.

Salient color names for person re-identification.

In Proc. European Conference on Computer Vision (ECCV), pages
536–551, 2014.

 H. Yao, S. Zhang, Y. Zhang, J. Li, and Q. Tian.

Deep representation learning with part loss for person re-identification.

CoRR, abs/1707.00798, 2017.

 M. Ye, A. J. Ma, L. Zheng, J. Li, and P. C. Yuen.

Dynamic label graph matching for unsupervised video re-identification.

In Proc. IEEE International Conference on Computer Vision (ICCV),
pages 5152–5160, Oct 2017.

 L. Zhang, T. Xiang, and S. Gong.

Learning a discriminative null space for person re-identification.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition
(CVPR), 2016.

 H. Zhao, M. Tian, S. Sun, J. Shao, J. Yan, S. Yi, X. Wang, and X. Tang.

Spindle net: Person re-identification with human body region guided feature decomposition and fusion.

In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 907–915, July 2017.

-  L. Zhao, X. Li, Y. Zhuang, and J. Wang.

Deeply-learned part-aligned representations for person re-identification.

In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 3219–3228, 2017.

-  R. Zhao, W. Ouyang, and X. Wang.

Person re-identification by salience matching.

In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 2528–2535, Dec 2013.

-  R. Zhao, W. Ouyang, and X. Wang.

Unsupervised salience learning for person re-identification.

In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3586–3593, June 2013.

-  R. Zhao, W. Ouyang, and X. Wang.

Learning mid-level filters for person re-identification.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 144–151, June 2014.

-  L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian.
Scalable person re-identification: A benchmark.

In Proc. IEEE International Conference on Computer Vision (ICCV),
pages 1116–1124, Dec 2015.

-  L. Zheng, S. Wang, L. Tian, F. He, Z. Liu, and Q. Tian.
Query-adaptive late fusion for image search and person re-identification.
- In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1741–1750, June 2015.

-  L. Zheng, H. Zhang, S. Sun, M. Chandraker, Y. Yang, and Q. Tian.
Person re-identification in the wild.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3346–3355, July 2017.

-  W. S. Zheng, S. Gong, and T. Xiang.
Person re-identification by probabilistic relative distance comparison.
- In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 649–656, 2011.

-  W. S. Zheng, S. Gong, and T. Xiang.
Reidentification by relative distance comparison.
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 35(3):653–668, March 2013.
-  Z. Zheng, L. Zheng, and Y. Yang.
Unlabeled samples generated by gan improve the person re-identification baseline in vitro.
In *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2017.
-  Z. Zhong, L. Zheng, D. Cao, and S. Li.
Re-ranking person re-identification with k-reciprocal encoding.
In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
-  Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang.
Random erasing data augmentation.
arXiv, 2017.
-  S. Zhou, J. Wang, J. Wang, Y. Gong, and N. Zheng.
Point to set similarity based deep feature learning for person re-identification.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5028–5037, July 2017.



Z. Zhou, Y. Huang, W. Wang, L. Wang, and T. Tan.

See the forest for the trees: Joint spatial and temporal recurrent neural networks for video-based person re-identification.

In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 6776–6785, July 2017.