|  |  |
| --- | --- |
| 行人ReID调研报告 | 文件类型：技术文档 |
| 文件编号： |
| 面向的部门： 研发中心 |
| 保密等级：高 |
| 作者： 刘凯 |
| 日期：2018-08-28 |
| 版本：1.0V |

行人ReID调研报告



目 录

[1 MTCNN算法原理 3](#_Toc523408927)

[1.1 算法框架 3](#_Toc523408928)

[1.2 网络结构 4](#_Toc523408929)

[1.3 训练 6](#_Toc523408930)

[2 MTCNN算法实现 6](#_Toc523408931)

[参考文献 1](#_Toc523408932)

# 行人ReID简介

* 1. 简介

行人重识别(Person Re-Identification，ReID)技术是现在计算机视觉研究的热门方向，主要解决跨摄像头跨场景下行人的识别与检索。在监控视频中，由于相机分辨率和拍摄角度的缘故，通常无法得到质量非常高的人脸图片。该技术能够根据行人的穿着、体态、发型等信息认知行人，与人脸识别结合能够适用于更多新的应用场景，将人工智能的认知水平提高到一个新阶段。ReID已经在学术界研究多年，但直到最近几年随着深度学习的发展，才取得了非常巨大的突破。

行人重识别的难点在于实际场景中数据非常复杂，会受到各种因素的影响，这些因素是客观存在的，如低分辨率、姿态变化、配饰、遮挡等因素，如图所示。另外一个难点在于行人重识别的数据集非常难获得，现有的数据集的图片大概在几万张，ID数量基本小于2000，摄像头在10个以下，与人脸数据集动辄几百万张图片的数据集比，对算法要求更高。同时ReID数据必须跨摄像头采集，需要对同一个人采集多张全身照图片，因此对数据采集要求很高。

在行人重识别中，通常用query/probe表示查询图像，用gallery表示候选行人库。行人重识别会从gallery中选取与query最形似的几张图片，作为识别结果。通常用Rank-k作为识别性能的度量(见2.2节)，图1为数据库DukeMTMC-reID的示意图。

行人重识别可以大致分为如下几类：

**基于表征学习(Representation learning)**，这类算法将ReID当成一个分类问题(classification)或验证问题(verification)。分类问题是指利用行人的ID或者属性等作为训练标签来训练模型；验证问题是指输入一对（两张）行人图片，让网络来学习这两张图片是否属于同一个行人。

**基于度量学习(Metric learning)**，度量学习旨在通过网络学习出两张图片的相似度。在行人重识别问题上，具体为同一行人的不同图片相似度大于不同行人的不同图片。最后网络的损失函数使得相同行人图片（正样本对）的距离尽可能小，不同行人图片（负样本对）的距离尽可能大。常用的度量学习损失方法有对比损失(Contrastive loss)[5]、三元组损失(Triplet loss)[6-8]等。

**Partial Person ReID**，上面两类算法通常针对的是行人的全身照片，但很多时候遮挡会降低识别性能。Part-based方法通常提前行人图片的局部信息。一种局部信息是行人姿态估计和关键点检测，通过对其行人关键点来识别。另一种局部信息是将行人图片切分为几个部分，每部分分别提取局部信息。基于局部信息的行人重识别是解决遮挡问题的有效手段。



图1：行人重识别示意图

# 数据集和度量

* 1. 数据集

常用的行人重识别数据集包括[1-5]：Market-1501, DukeMTMC-reID, CUHK03以及部分行人ReID数据集Partial REID, Partial-iLIDS等。前三个为全身的行人图片集，后两个为部分(partial)的行人图片集，行人照片多被遮挡。

Market-1501有32668张行人图片，共1501个行人，由6个摄像头在一个超市前拍摄得到。其中训练集包含12936张图片共751个行人，测试集包含19732张图片共750个行人。

DukeMTMC-reID有36411张图片，共1812个行人，由8个摄像头拍摄得到。

CUHK03有14096张图片，共1467个行人，由10个摄像头拍摄得到。

Partial REID有600张图片，共60人，每个人有5张全身图片和5张半身(partial)图片。

Partial-iLIDS有476张人物图片，共119个人。部分图片被遮挡。

* 1. 度量

常用的行人重识别度量包括：Rank-*k*，mAP(Mean Average Percision)和CMC(Cumulative Matching Characteristics)。

Rank-*k*，其中k为正整数，如1,2,3,5,10等。将gallery中的图片按照与query的相似性进行排序，Rank-1为首位命中率，类似地，Rank-k表示前k张图中包含正确结果的概率。

mAP，P表示Precision，即识别的准确率。AP指一次识别的平均准确率。mAP表示多次识别的平均识别率。mAP是比Rank-k更准确的度量。mAP的计算示例如图所示。

CMC，一条性能曲线，横轴为Rank-k，纵轴为准确率，曲线递增，如图所示。

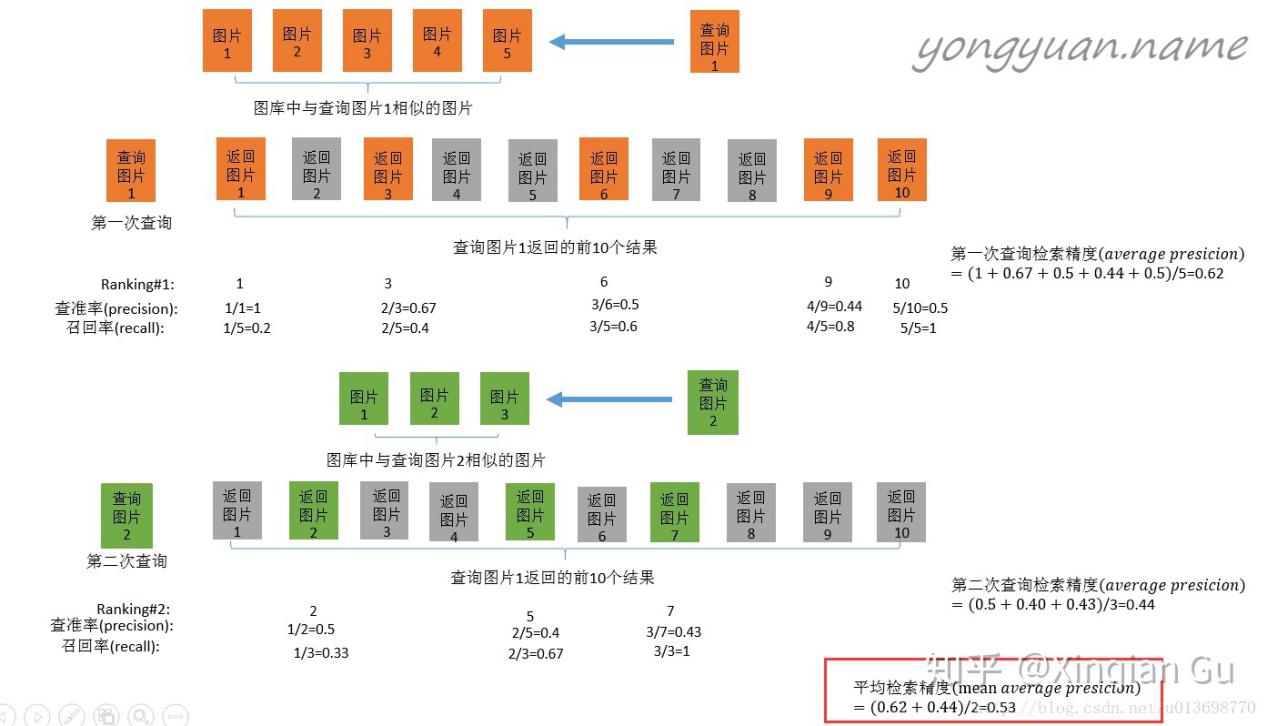


图2：mAP计算示意图

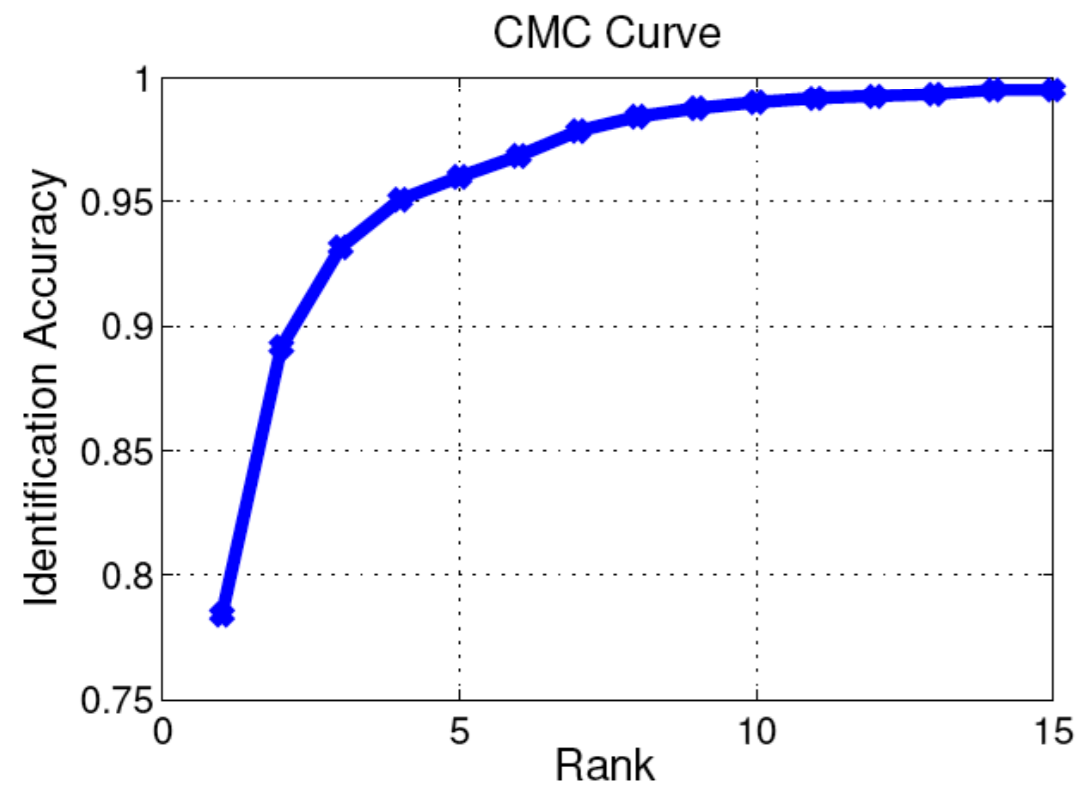


图3：CMC曲线示意图

参考文献

1. Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., Tian, Q.: Scalable person reidentification: A benchmark. In: ICCV (2015)
2. Ristani, E., Solera, F., Zou, R., Cucchiara, R., Tomasi, C.: Performance measures and a data set for multi-target, multi-camera tracking. In: European Conference on Computer Vision workshop on Benchmarking Multi-Target Tracking (2016)
3. Li, W., Zhao, R., Xiao, T., Wang, X.: Deepreid: Deep filter pairing neural network for person re-identification. In: CVPR (2014)
4. Zheng, W.S., Li, X., Xiang, T., Liao, S., Lai, J., Gong, S.: Partial person reidentification. In: ICCV (2015)
5. Zheng, W.S., Gong, S., Xiang, T.: Person re-identification by probabilistic relative distance comparison. In: CVPR (2011).