

基于统计推断的行人再识别算法

杜宇宁* 艾海舟

(清华大学计算机系 北京 100084)

摘要: 行人再识别是指给定一张行人图像, 在已有的可能来源于非交叠摄像机视场的行人图像库中, 识别出与此人相同的图像。研究该问题有着非常重要的现实意义, 同时也面临许多挑战。该文提出一种基于统计推断的行人再识别算法。该算法从统计推断的角度出发学习两幅行人图像的相似度度量函数, 利用此函数从行人图像库中搜索待查询的人。在公共数据集 VIPeR 上的实验表明, 该算法性能优于已有的行人再识别算法, 学习相似度度量函数的时间开销明显少于已有的基于学习的算法, 并且在只有少量训练样本时, 缓解了学习相似度度量函数的过拟合问题。

关键词: 计算机视觉; 行人再识别; 相似度度量函数; 统计推断

中图分类号: TP391.4

文献标识码: A

文章编号: 1009-5896(2014)07-1612-07

DOI: 10.3724/SP.J.1146.2013.01144

A Statistical Inference Approach for Person Re-identification

Du Yu-ning Ai Hai-zhou

(Computer Science and Technology Department, Tsinghua University, Beijing 100084, China)

Abstract: Person re-identification, identifying the same person's images in an existing database come from non-overlapping camera views, is a valuable but challenging task. This paper proposes a statistical inference approach for person re-identification. A similarity measure of two person images is learned from a statistical inference perspective. Then the similarity measure is utilized to query a person from a gallery set. The proposed approach is demonstrated on VIPeR dataset, and the experiment shows that it outperforms the state-of-the-art approaches. Besides, it costs less time than the existing learning-based ones in training, and alleviates the over-fitting problem when there are few training data.

Key words: Computer vision; Person re-identification; Similarity measure; Statistical inference

1 引言

随着行人检测和跟踪技术的发展^[1-3], 从视频中很容易获得行人图像库。行人再识别问题是指给定一张行人图像, 在已有的可能来源于非交叠摄像机视场的行人图像库中, 识别出与此人相同的图像。研究此问题有着非常重要的现实意义和广泛应用。行人再识别技术一个典型的应用场景是刑事侦查。给定嫌疑人照片, 收集与嫌疑人相关的视频监控录像, 进而搜集与嫌疑人有关的线索。在实际应用中, 一般通过有经验的公安人员观看相关的监控视频获取信息。但是, 由于相关的视频监控录像中涉及的行人图像的数量往往是海量的, 此方法侦查效率非常低下。如果利用行人检测、跟踪和行人再识别技术, 计算机能自动标定出嫌疑人在相关视频中的位

置, 无疑将会大大提高警察搜捕犯罪嫌疑人的效率。行人再识别技术除了应用在刑侦方面, 还可以用于在公共场所中寻找丢失的小孩、个人的相册管理、电子商务等领域。在自然场景中, 由于人的姿态、视角、光照以及人所处背景的变化都会导致同一个人不同的场景中表现变化很大, 因此行人再识别问题也是一个很有挑战性的问题。正因为行人再识别问题的重要现实意义和面临的困难, 在计算机视觉领域, 近几年越来越多的研究者关注此问题^[4-13]。为了叙述方便, 给定一对行人图像, 如果它们来源于同一个行人, 则称其为相关行人对, 否则称其为不相关行人对。

目前已有的行人再识别算法大概分为两类: 直接法^[4-6]和基于学习的算法^[7-13]。

直接法的基本出发点是手动设计具有区分性和稳定性的描述子, 利用这些描述子来度量两张行人图像的相似度。文献[4]将行人图像粗略分为头部、上半身和下半身 3 部分, 然后串联每一部分的颜色直方图对此人进行描述。文献[5]采用分割的技术提

2013-07-30 收到, 2014-02-04 改回

国家自然科学基金(61075026)和国家 973 计划项目(2011CB302203)资助课题

*通信作者: 杜宇宁 dyn10@mails.tsinghua.edu.cn

取人的前景, 并利用行人区域的对称性和非对称性将人的前景划分成不同的区域。对于每个区域, 提取带权重的颜色直方图特征、极大稳定颜色区域(maximally stable color regions)特征和重复度高的结构区域(recurrent highly structured patches)特征描述它们。文献[6]采用图案结构(pictorial structure)算法定位图像中人的各个部件所在的区域。对于每个部件的区域, 提取与文献[5]类似的颜色直方图特征、极大稳定颜色区域特征来描述它们。虽然直接法不需要训练, 但是在现实中, 由于同一个人的表现变化往往很大, 很难手动设计出具有区分性和稳定性的特征。此外, 由于提取这些特征的时间花销往往很大, 所以与基于学习的算法相比, 直接法的搜索效率往往不高。

与手动设计特征的直接法不同, 基于学习的算法则是利用机器学习的方法, 学习出两张行人图像的相似度量函数, 使相关的行人图像对的相似度尽可能高, 不相关的行人图像对的相似度尽可能低。探究一个相似度量函数的性能需要考虑4个方面的因素: (1)算法的识别准确率。(2)算法的查询效率。(3)学习相似度量函数的时间花销。利用行人跟踪和检测技术, 从视频中很容易获得行人的跟踪片段, 尤其在监控视频中。学习一个适用于所有场景通用的相似度量函数往往很困难。针对不同的场景, 学习适合相应场景的相似度量函数更有实际意义。但与此同时, 从大量的行人图像对中学习相似度量函数的时间花销也将成为一个重要问题。(4)只有少量样本时训练的过拟合问题。除了利用行人检测和跟踪技术获得用于学习相似度量函数的行人图像对, 也可以通过手动标定获得训练样本。但是, 因为标注工作枯燥易错, 所以训练样本的数目往往很少。因此在设计行人再识别算法时也需要关注在只有少量训练样本时训练的过拟合问题。

文献[7]将所有的水平条带的颜色特征和纹理特征定义为特征池, 利用 Adaboost 算法学习最优的特征组合来度量一对行人图像的相似度。但是在只有少量的训练样本情形下, 此算法往往过拟合严重。为了缓解少量样本过拟合问题, 文献[8]使用和文献[7]同样的特征, 利用随机森林算法学习相似度量函数。文献[9]则将人的图像平均分成6个水平条带。对于每个水平条带, 抽取相应的颜色特征和纹理特征。将所有特征串联, 利用支持向量机学习相似度量函数。值得关注的是, 近两年研究者非常亲睐基于度量学习的算法^[10-13]。这主要有两方面的原因: 一方面, 即使使用简单的颜色直方图作为特征,

这些算法的性能往往优于其它算法。另一方面, 因为这些算法只是将原始特征空间线性投影到另一个更具区分性的特征空间, 所以与其它的算法相比它们具有更高的查询效率。但是这些算法学习相似度量函数的时间消耗比较大, 在只有少量样本时训练的过拟合问题依然存在。

本文提出一种基于统计推断的行人再识别算法。该算法从统计推断的角度出发学习两幅行人图像的相似度量函数。除了具有基于度量学习算法的优点(特征简单和搜索效率高), 该算法学习的相似度量函数更易区分相关的和不相关的行人对, 识别性能优于已有的行人再识别算法。由于训练过程只利用了训练样本的子样本空间, 并且没有复杂的优化迭代过程, 所以该算法学习相似度量函数的时间花销明显少于已有的基于学习的算法。此外, 该算法在只有少量训练样本时缓解了训练的过拟合问题。

本文其余章节的组织安排如下。第2节介绍本文提出的行人再识别算法。第3节介绍本文算法在公共数据集上的实验。第4节总结全文。

2 基于统计推断的行人再识别算法

本节首先介绍基于统计推断的相似度量函数的学习方法; 其次介绍算法中所使用的行人图像的特征。

2.1 相似度量函数的学习

行人再识别的目的是给定一张行人图像, 在已有的可能来源于非交叠摄像机视场的行人图像库中, 搜索出此人的图像。为达到这个目的, 可以通过学习一个相似度量函数 $f(\bullet)$ 来估计一对行人图像的相似度, 从而使相关的行人图像对的相似度尽可能高, 不相关的行人图像对的相似度尽可能低。给定一对行人图像 z_i 和 z_j , 将产生两对有序行人图像对 (z_i, z_j) 和 (z_j, z_i) 。从统计学的角度考虑, 一对有序行人图像对 (z_i, z_j) 的相似度可以表示为

$$f(z_i, z_j) = \frac{P((z_i, z_j) \in S)}{P((z_i, z_j) \in D)} \quad (1)$$

其中 S 表示相关的行人图像对空间, D 表示不相关的行人图像对空间, $P((z_i, z_j) \in S)$ 表示 (z_i, z_j) 属于相关的行人图像对空间的概率, $P((z_i, z_j) \in D)$ 表示 (z_i, z_j) 属于不相关的行人图像对空间的概率。如果行人图像 z_i 和 z_j 越相似, 则 (z_i, z_j) 属于相关的行人图像对空间的概率越大, 即 $P((z_i, z_j) \in S)$ 越大, 属于不相关的行人图像对空间的概率越小, 即 $P((z_i,$

$z_j \in D$) 越小, 从而 z_i 和 z_j 的相似度量值 $f(z_i, z_j)$ 越大。反之亦然。无论 z_i 和 z_j 是属于相关的图像空间对, 还是不相关的图像空间对, 作为一种集合上的“关系”, (z_i, z_j) 和 (z_j, z_i) 都是对称的, 即具有相同的属性, 所以 $f(z_i, z_j) = f(z_j, z_i)$ 。

假设 x 是行人图像 z 的特征向量, 为了估计一对有序行人图像对 (z_i, z_j) 属于相关行人图像对空间的概率和不相关行人图像对空间的概率, 将 (z_i, z_j) 投影到特征差分空间 $x_i - x_j$ 。虽然 (z_i, z_j) 和 (z_j, z_i) 在特征差分空间的投影不同, 但是因为它们是对称的, 所以它们在特征差分空间的投影也是对称的。与 KISSME 算法^[10]类似, 假设相关的图像空间对和不相关的图像空间对在特征差分空间的投影分别服从均值为 0 , 协方差矩阵为 Σ_S 和 Σ_D 的多维正态分布, 则相似度量函数 $f(\bullet)$ 表示为

$$f(z_i, z_j) = \frac{\left(1/\sqrt{2\pi|\Sigma_S|}\right)e^{-1/2(x_i-x_j)^T \Sigma_S^{-1}(x_i-x_j)}}{\left(1/\sqrt{2\pi|\Sigma_D|}\right)e^{-1/2(x_i-x_j)^T \Sigma_D^{-1}(x_i-x_j)}} \quad (2)$$

为了计算方便, 式(2)两边分别采用对数运算, 则变换为

$$\begin{aligned} \lg f(z_i, z_j) = & \lg \frac{1}{\sqrt{2\pi|\Sigma_S|}} - \frac{1}{2}(x_i - x_j)^T \\ & \cdot \Sigma_S^{-1}(x_i - x_j) \\ & - \lg \frac{1}{\sqrt{2\pi|\Sigma_D|}} + \frac{1}{2}(x_i - x_j)^T \\ & \cdot \Sigma_D^{-1}(x_i - x_j) \end{aligned} \quad (3)$$

因为相似度量函数中的常数不影响相似度的度量, 所以可以忽略式(3)中右边的常数, 得到

$$\lg f(z_i, z_j) = \frac{1}{2}(x_i - x_j)^T (\Sigma_D^{-1} - \Sigma_S^{-1})(x_i - x_j) \quad (4)$$

假设 $M = \Sigma_D^{-1} - \Sigma_S^{-1}$, 根据式(4), 可以利用 M 来估计一对行人图像的相似度。与以往的度量学习

方法不同, 由于 M 可能不是半正定矩阵, 所以 M 可能不是度量矩阵(metric)。但是通过实验发现, 利用 M 和式(4)可以学习到一个好的相似度量函数, 其性能优于已有的采用度量学习算法学习的度量矩阵。

利用式(4)学习相似度量函数的时间花销主要在相关的行人图像对和不相关的行人图像对的协方差矩阵的估计, 即估计 Σ_S 和 Σ_D 。已有的度量学习算法^[10-13]使用训练集中所有的行人图像对去学习相似度量函数。当训练集中标定了 N 对相同行人的图像, 相关的行人图像对数目为 N_S , 不相关的行人图像对数目为 N_D , 则 $N_S = N, N_D = 4 \cdot [(N-1) + (N-2) + \dots + 1] = 2N(N-1)$ 。当 $N = 316$ 时, $N_S = 316, N_D = 199080$ 。很明显, 训练的时间花销主要在估计 Σ_D 。为了减少训练的时间花销, 可以选择不相关的行人图像对的子样本空间, 即随机选择 $\alpha N_D (0 < \alpha \leq 1)$ 对行人图像对去估计 Σ_D 。为了探究采用不同规模的子样本空间学习的相似度量函数的性能差异, 从 VIPeR 数据集中, 随机选择 316 对相同行人的图像用于学习相似度量函数, 其余的 316 对相同行人的图像形成相关的行人图像对空间和不相关的行人图像对空间。利用学到的相似度量函数, 可以统计出两个空间的行人图像对的相似度概率分布。图 1 显示了当 α 取 0.02, 0.10 和 1.00 时, 将行人图像对的相似度区间平均分成 32 个小区间, 统计出的两个空间的行人图像对的相似度概率分布。从图 1 中可以看出, α 取值较小时和 $\alpha = 1$ 时两个空间的行人图像对的相似度概率分布变化不是很大, 学习的相似度量函数的性能相当。但是 α 取值较小时学习相似度量函数的时间花销远远小于当 $\alpha = 1$ 时, 这也是本文算法训练时间少于其它的基于度量学习的算法的原因。算法 1 总结了本文算法的实现步骤(表 1)。

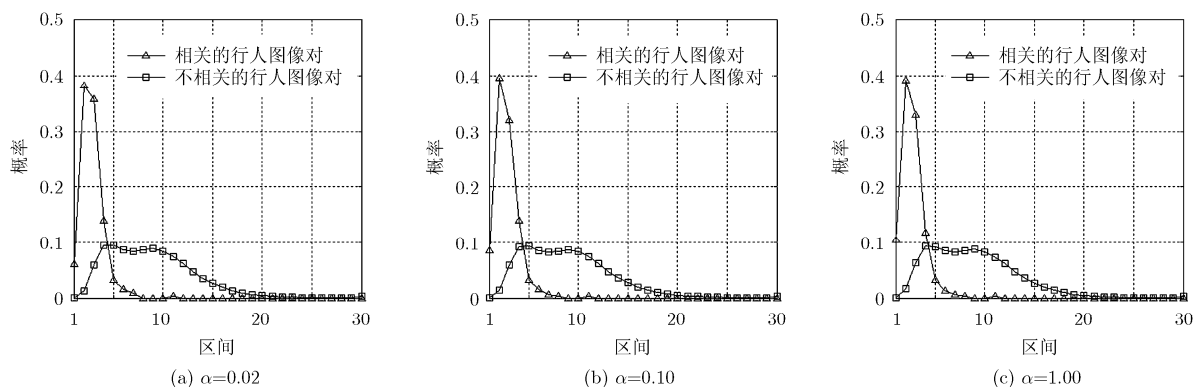


图 1 不同 α 取值下不同空间的行人图像对的相似度概率分布

表 1 基于统计推断的行人再识别算法步骤

| | |
|------------------------------|--|
| 算法 1 基于统计推断的相似度量函数的学习 | |
| 输入: | N 对相同行人的图像, 参与训练的样本空间规模参数 α |
| 输出: | 相似度矩阵 M |
| 步骤 1 | 从 N 对相同行人的图像产生 N 对相关行人图像对, 估计 Σ_S^{-1} 。 |
| 步骤 2 | 从 N 对相同行人的图像产生 $2\alpha N(N-1)$ 对不相关的行人图像对, 估计 Σ_D^{-1} 。 |
| 步骤 3 | $M = \Sigma_D^{-1} - \Sigma_S^{-1}$ |

2.2 行人图像的特征

在已有的行人再识别算法中, 颜色和纹理是常用的两种特征。本文后续实验表明, 不用纹理特征, 只用颜色特征也能学到很好的相似度量函数。在本文实验中, 我们从规范化的 RGB(NRGB), HSV, YCbCr 和 CIE Lab 4 种颜色空间提取特征。图 2(a) 给出了提取行人图像特征的示例。一张行人图像被平均分成 6 个水平条带, 对于每个水平条带的每种颜色空间的每个通道, 提取 16 维的颜色直方图特征, 将所有的颜色直方图串联到一块形成行人图像的特征。因此本文使用的行人图像的特征的维数为 $1152(6 \times 4 \times 3 \times 16)$ 。如果采用原始的 1152 维特征直接学习相似度量函数, 需要估计 $2 \times 1152 \times 1152$ 个参数, 参数量太多。因此, 实验中利用 PCA 降维, 将行人图像的特征维数降到 55, 保持大约 88.25% 的能量。

当行人目标未能正好填充 ROI 区域, 本文使用的特征也有一定的适用性。虽然行人的姿态变化很大, 但是变化主要在水平方向上, 在垂直方向上变化不是很明显。例如, 给定一对行人图像, 他们的头部、上半身和下半身在垂直方向上的范围基本一致。如果行人目标不能很好地填充 ROI 区域, 可以

在行人检测器后面加上一步后处理, 使行人目标较好填充 ROI 区域。

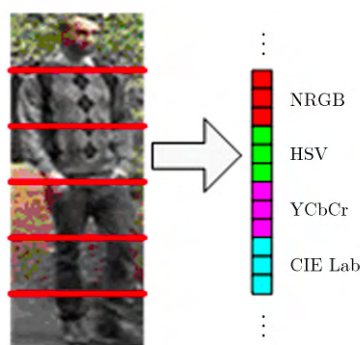
3 实验

本节首先介绍实验中所使用的测试数据和算法性能的评测准则。然后比较当 α 取值不同时学习的相似度量函数的性能差异, α 的含义见 2.1 节。最后介绍本文算法与已有的行人再识别算法的性能比较。

3.1 实验数据和算法性能评测准则

本文将在行人再识别广泛使用的公共数据集 VIPeR^[4]上评测本文算法。VIPeR 数据集是由 632 对行人图像组成, 如图 2(b)所示, 其中同一列中的行人图像为同一个人。每对行人图像来源于不同的摄像头场景。同一个人的表现由于姿态、视角、光照和背景的变化而变化很大。为了与已有算法公正比较, 在实验中, 采用先前工作普遍采用的评价框架。随机选择 p 对行人图像对作为测试集, 余下的行人图像对作为训练集。测试集由查询集和行人图像库两部分组成。每对行人图像, 随机选择一张图像并入查询集, 另一张则并入行人图像库。当给定一个行人再识别算法, 衡量该算法在行人图像库中搜索待查询行人的能力来评测此算法的性能。为了测试算法在只有少量的训练数据时的性能, p 分别取 316, 432, 532 进行实验。 p 取值越大, 则测试样本越多(匹配越困难), 训练样本越少(训练越困难)。对于每组实验, 以上产生测试集和训练集的过程重复 10 次, 10 次实验结果的平均值作为本组实验的结果。所有的实验是基于 MATLAB 实现的, 实验平台是 4 G 内存的 Intel(R) Core(TM)2 Quad 2.5 GHz PC 台式机。

已有的行人再识别算法大部分采用累积匹配特性(Cumulative Match Characteristic, CMC)曲线评价算法性能。给定一个查询集和行人图像库, 累积



(a) 行人图像特征的提取



(b) 公共数据集VIPeR中图像举例

图 2 行人图像及其特征

匹配特征曲线描述的是在行人图像库中搜索待查询的行人,前 r 个搜索结果中找到待查询人的比率。第 1 匹配率($r = 1$)很重要,因为它表示的是真正的识别能力。但是当 r 取值很小时,第 r 匹配率也很有现实意义。因为在实际应用中,反馈的前 r 个搜索结果中,可以通过人眼判断找到查询目标。从这个角度看,行人再识别问题也是一个信息检索问题。因此,除了累计匹配特性曲线,我们引入信息检索中常用的规范化折扣累计增益(Normalized Discounted Cumulative Gain, NDCG)评测算法性能。根据累积匹配特性曲线,可以利用式(5)计算规范化折扣累计增益。实验采用 NDCG@20 评测算法性能。

$$\text{NDCG}@k = \frac{\sum_{i=1}^k W(i) \cdot C(i)}{\sum_{i=1}^k W(i)} \tag{5}$$

其中 $C(i)$ 是 CMC 曲线中第 i 匹配率。

$$W(i) = \begin{cases} 1, & i = 1 \\ 1/\log_2(i), & 2 \leq i \leq k \end{cases}$$

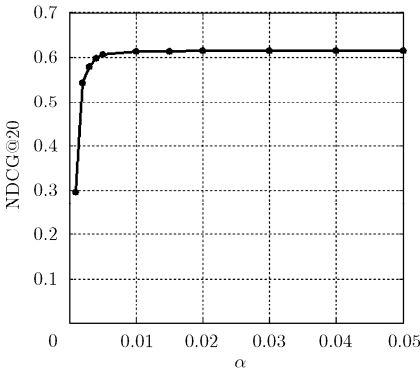
3.2 α 取值不同时算法的性能比较

图 3 显示了在测试集规模为 $p = 316$, α 取值为 $[0.001, 0.050]$ 时,学习的相似度量函数的性能 NDCG@20 和学习相似度量函数的时间花销随 α 取值不同的变化曲线。从图 3 中可以看出,当 α 达

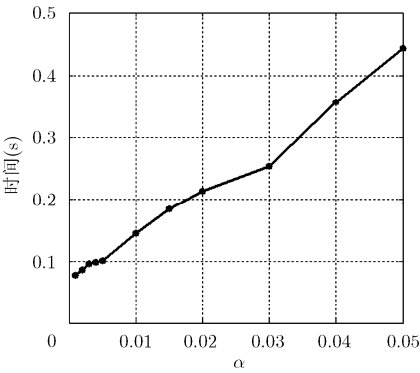
到一定值时,算法的识别性能不会随着 α 的增大而显著提高,但是学习相似度量函数的时间花销却随着 α 的增大而不断增加。表 2 显示了在不同的测试集规模上,当 $\alpha = 0.02$ 和 $\alpha = 1.00$ 时,学习的相似度量函数的性能比较。当测试集规模为 $p=316$ 时,从表 2 中的第 1,5,10,20 的匹配率和 NDCG@20 可以看出, $\alpha=0.02$ 和 $\alpha=1.00$ 学习的相似度量函数性能相当。但 $\alpha=0.02$ 的训练时间明显少于 $\alpha=1.00$ 。当测试集规模为 $p = 432$ 和 $p = 532$ 时,以上结论依然成立。一般情况下,当 $\alpha = 0.02$ 时,本文算法可以在较少的时间内学习出较好的相似度量函数。

3.3 与已有的行人再识别算法性能比较

表 3,表 4 对比了本文提出的行人再识别算法与已有算法的性能,其中 $\alpha = 0.02$ 。因为很难获得已有算法实验中累积匹配特性曲线每个匹配数的匹配率,所以本文只采用累积匹配特性曲线的第 1,5,10,20 匹配率进行性能对比。由于本文算法和 KISSME^[10] 算法想法类似,并且文献[10]表明, KISSME 算法性能与已有的基于度量学习的算法(除 RPML^[13])相当,但学习相似性度量函数的时间花销远远少于已有算法,所以在本文使用的特征基础上实现了 KISSME 算法,进而对比不同算法训练所需要的时间花销。此外,作为基准,我们也评测了本文使用的特征采用欧氏距离的性能。从表 3 可



(a) NDCG@20的变化曲线



(b)学习度量函数的时间花销变化曲线

图 3 不同 α 取值下本文算法学习的相似度量函数的评测

表 2 在不同的测试集规模上, $\alpha = 0.02$ 和 $\alpha = 1.00$ 的性能比较

| p | α | 第 1 匹配率 | 第 5 匹配率 | 第 10 匹配率 | 第 20 匹配率 | NDCG @ 20 | 时间花销(s) |
|-----|----------|---------|---------|----------|----------|-----------|---------|
| 316 | 0.02 | 0.297 | 0.598 | 0.730 | 0.841 | 0.616 | 0.217 |
| | 1.00 | 0.298 | 0.594 | 0.721 | 0.843 | 0.616 | 6.088 |
| 432 | 0.02 | 0.225 | 0.486 | 0.614 | 0.744 | 0.517 | 0.084 |
| | 1.00 | 0.228 | 0.487 | 0.618 | 0.746 | 0.520 | 2.471 |
| 532 | 0.02 | 0.124 | 0.311 | 0.430 | 0.567 | 0.353 | 0.035 |
| | 1.00 | 0.130 | 0.318 | 0.439 | 0.575 | 0.362 | 0.658 |

以看出，本文算法性能优于欧氏距离，行人再识别直接法中的 SDALF^[5]和 PS^[6]算法，基于度量学习的算法 KISSME^[10]，PRDC^[11]，RPML^[13]，ITML^[14]，LMNN^[15]和 MLCC^[16]，以及其它的基于学习的算法 ELF^[7]，RECF^[8]和 PRSVM^[9]。可见，本文提出的基于统计推断的行人再识别算法可以学到一个很好的相似度量函数。在测试集规模 $p = 316$ 时，虽然第 1 匹配率只有 29.7%，但是第 20 匹配率达到 84.1%，

因此本文算法有一定的实际应用价值。此外，从表 4 可以得出，在只有少量的训练样本情况下，该算法性能同样优于已有算法，可见该算法缓解了学习相似度量函数的过拟合问题。

表 5 对比了在不同的测试集规模上，本文算法和 KISSME 算法训练所需要的时间。由于本文算法只利用子样本空间进行训练，所以学习相似度量函数的时间花销明显少于 KISSME 算法。

表 3 当测试集规模为 316 时，本文算法与已有行人再识别算法性能比较

| 算法 | 第 1 匹配率 | 第 5 匹配率 | 第 10 匹配率 | 第 20 匹配率 |
|------------------------|---------|---------|----------|----------|
| 本文算法 | 0.297 | 0.598 | 0.730 | 0.841 |
| 欧氏距离 | 0.055 | 0.142 | 0.195 | 0.293 |
| KISSME ^[10] | 0.222 | 0.517 | 0.678 | 0.818 |
| SDALF ^[5] | 0.200 | 0.390 | 0.490 | 0.660 |
| PS ^[6] | 0.220 | 0.450 | 0.570 | 0.710 |
| ELF ^[7] | 0.120 | 0.320 | 0.445 | 0.615 |
| RECF ^[8] | 0.170 | 0.310 | 0.426 | 0.564 |
| PRSVM ^[9] | 0.163 | 0.382 | 0.537 | 0.699 |
| PRDC ^[11] | 0.157 | 0.384 | 0.539 | 0.701 |
| LMNN-R ^[12] | 0.200 | 0.490 | 0.680 | 0.800 |
| RPML ^[13] | 0.270 | 0.560 | 0.690 | 0.830 |
| ITML ^[14] | 0.116 | 0.314 | 0.458 | 0.639 |
| LMNN ^[15] | 0.062 | 0.197 | 0.326 | 0.523 |
| MLCC ^[16] | 0.152 | 0.418 | 0.576 | 0.734 |

表 4 当测试集规模为 432 和 532 时，本文算法与已有行人再识别算法的性能比较

| 算法 | $p = 432$ | | | | $p = 532$ | | | |
|------------------------|-----------|---------|----------|----------|-----------|---------|----------|----------|
| | 第 1 匹配率 | 第 5 匹配率 | 第 10 匹配率 | 第 20 匹配率 | 第 1 匹配率 | 第 5 匹配率 | 第 10 匹配率 | 第 20 匹配率 |
| 本文算法 | 0.225 | 0.486 | 0.614 | 0.744 | 0.124 | 0.311 | 0.430 | 0.567 |
| 欧氏距离 | 0.048 | 0.115 | 0.164 | 0.232 | 0.040 | 0.103 | 0.148 | 0.209 |
| KISSME ^[10] | 0.176 | 0.426 | 0.566 | 0.715 | 0.113 | 0.294 | 0.421 | 0.562 |
| RECF ^[8] | 0.149 | 0.289 | 0.383 | 0.502 | 0.127 | 0.249 | 0.332 | 0.440 |
| PRSVM ^[9] | 0.106 | 0.297 | 0.423 | 0.583 | 0.089 | 0.229 | 0.327 | 0.460 |
| PRDC ^[11] | 0.126 | 0.320 | 0.443 | 0.600 | 0.091 | 0.242 | 0.344 | 0.486 |
| RPML ^[13] | 0.200 | 0.420 | 0.560 | 0.710 | 0.110 | 0.280 | 0.380 | 0.520 |
| ITML ^[14] | 0.084 | 0.245 | 0.368 | 0.523 | 0.042 | 0.111 | 0.172 | 0.246 |
| LMNN ^[15] | 0.051 | 0.131 | 0.203 | 0.339 | 0.040 | 0.097 | 0.142 | 0.212 |
| MLCC ^[16] | 0.113 | 0.324 | 0.473 | 0.629 | 0.050 | 0.163 | 0.259 | 0.396 |

表 5 在不同测试集规模上，本文算法与 KISSME 算法训练所需要的时间比较(s)

| 算法 | $p = 316$ | $p = 432$ | $p = 532$ |
|------------------------|-----------|-----------|-----------|
| 本文算法 | 0.217 | 0.084 | 0.035 |
| KISSME ^[10] | 6.209 | 2.492 | 0.687 |

4 结论

行人再识别是一个具有实际应用价值同时也很有挑战性的问题。本文提出一种基于统计推断的行人再识别算法，从统计推断的角度出发学习相似度量函数，进而估计两幅行人图像的相似度。该算

法性能优于已有的行人再识别算法。由于训练过程只利用了训练样本的子样本空间, 并且没有复杂的优化迭代过程, 所以该算法学习相似度度量函数的时间花销明显少于已有的基于学习的算法。此外, 该算法缓解了只有在少量训练样本时训练的过拟合问题。

参 考 文 献

- [1] Duan Gen-quan, Ai Hai-zhou, Xing Jun-liang, *et al.* Scene aware detection and block assignment tracking in crowded scenes[J]. *Image and Vision Computing*, 2012, 30(4): 292-305.
- [2] Kalal Z, Mikolajczyk K, and Matas J. Tracking-learning-detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7): 1409-1422.
- [3] Dollar P, Wojek C, Schiele B, *et al.* Pedestrian detection: an evaluation of the state of the art[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(4): 743-761.
- [4] Gray D, Brennan S, and Tao Hai. Evaluating appearance models for recognition, reacquisition, and tracking[C]. *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, Rio de Janeiro, 2007: 41-47.
- [5] Farenzena M, Bazzani L, Perina A, *et al.* Person re-identification by symmetry-driven accumulation of local features[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, 2010: 2360-2367.
- [6] Cheng Dong-seon, Cristani M, Stoppa M, *et al.* Custom pictorial structures for re-identification[C]. *British Machine Vision Conference*, Dundee, 2011: 749-760.
- [7] Gray D and Tao Hai. Viewpoint invariant pedestrian recognition with an ensemble of localized features[C]. *European Conference on Computer Vision*, Marseille, 2008: 262-275.
- [8] Du Yu-ning, Ai Hai-zhou, and Lao Shi-hong. Evaluation of color spaces for person re-identification[C]. *IEEE Conference on Pattern Recognition*, Tsukuba, 2012: 1371-1374.
- [9] Prosser B, Zheng Wei-shi, Gong Shao-gang, *et al.* Person re-identification by support vector ranking[C]. *British Machine Vision Conference*, Aberystwyth, 2010: 231-242.
- [10] Kostinger M, Hirzer M, Wohlhart P, *et al.* Large scale metric learning from equivalence constraints[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, 2012: 2288-2295.
- [11] Zheng Wei-shi, Gong Shao-gang, and Xiang Tao. Person reidentification by probabilistic relative distance comparison [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, Colorado Springs, 2011: 649-656.
- [12] Dikmen M, Akbas E, Huang T S, *et al.* Pedestrian recognition with a learned metric[C]. *The Asian Conference on Computer Vision*, New Zealand, 2010: 501-512.
- [13] Hirzer M, Roth P M, Kostinger M, *et al.* Relaxed pairwise learned metric for person reidentification[C]. *European Conference on Computer Vision*, Florence, 2012: 780-793.
- [14] Davis J V, Kulis B, Jain P, *et al.* Information-theoretic metric learning[C]. *International Conference on Machine Learning*, Corvallis, 2007: 209-216.
- [15] Weinberger K Q and Saul L K. Distance metric learning for large margin nearest neighbor classification[J]. *The Journal of Machine Learning Research*, 2009, 10(2): 207-244.
- [16] Globerson A and Roweis S. Metric learning by collapsing classes[C]. *Advances in Neural Information Processing Systems*, Vancouver, 2006: 451-458.

杜宇宁: 男, 1987 年生, 博士生, 研究方向为计算机视觉。

艾海舟: 男, 1964 年生, 教授, 博士生导师, 研究方向为计算机视觉。