



(12) 发明专利

(10) 授权公告号 CN 110969112 B

(45) 授权公告日 2022. 08. 16

(21) 申请号 201911189515.X

CN 108257158 A, 2018.07.06

(22) 申请日 2019.11.28

CN 110428448 A, 2019.11.08

(65) 同一申请的已公布的文献号

CN 106056581 A, 2016.10.26

申请公布号 CN 110969112 A

US 2013287248 A1, 2013.10.31

US 2017245092 A1, 2017.08.24

(43) 申请公布日 2020.04.07

(73) 专利权人 福州大学

余春艳 等. 鉴别性特征学习模型实现跨摄像头下行人即时对齐.《计算机辅助设计与图形学学报》.2019, 第31卷(第4期), 第602-611页.

地址 350108 福建省福州市闽侯县福州大学城乌龙江北大道2号福州大学

余春艳 等. 鉴别性特征学习模型实现跨摄像头下行人即时对齐.《计算机辅助设计与图形学学报》.2019, 第31卷(第4期), 第602-611页.

(72) 发明人 余春艳 钟诗俊 赖奇嵘

夏天 等. 基于深度学习和时空约束的跨摄像头行人跟踪.《计算机与数字工程》.2017, 第45卷(第11期), 第2270-2274页.

(74) 专利代理机构 福州元创专利商标代理有限公司 35100

专利代理师 丘鸿超 蔡学俊

孙锐 等. 联合多级深度特征表示和有序加权距离融合的视频行人再识别方法.《光学学报》.2019, 第39卷(第9期),

(51) Int. Cl.

G06V 40/10 (2022.01)

G06V 10/762 (2022.01)

G06V 10/74 (2022.01)

G06V 10/80 (2022.01)

G06V 10/82 (2022.01)

G06K 9/62 (2022.01)

G06N 3/08 (2006.01)

W.Chen 等. An Equalized Global Graph Model-Based Approach for Multicamera Object Tracking.《in IEEE Transactions on Circuits and Systems for Video Technology》.2016,

(56) 对比文件

审查员 孔月瑶

CN 108198200 A, 2018.06.22

权利要求书3页 说明书7页 附图3页

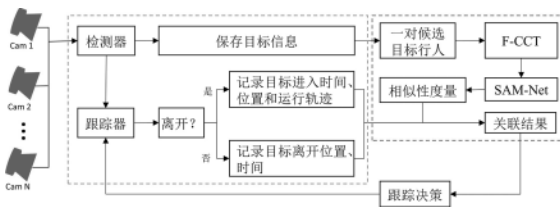
(54) 发明名称

跨摄像头场景下行人身份对齐方法

(57) 摘要

本发明提出一种跨摄像头场景下行人身份对齐方法, 其在于解决跨摄像头下的行人关联问题, 以对单摄像头下的目标行人进行持续正确的跟踪为基础, 在目标离开摄像头视野区域后, 进入盲区, 跟踪断续, 但当其再次出现在摄像头下时, 能够重新识别出该行人并维持其身份标识不变延续跟踪。当检测器检测到新的行人时, 将其加入待关联的候选池中, 之后从候选池中选择一对行人使用F-CCT模型完成图像预处理, 处理后的图像将作为SAM-Dets模型的输入数据并求得两个行人的外观适配度。候选池中的行人两两配

对计算外观适配度后, 根据每对行人适配度并结合时空关系建立最小费用流图模型求解最优的行人关联解; 最后根据关联结果对行人保持原有标识或者赋予新的身份标识。



CN 110969112 B

1. 一种跨摄像头场景下行人身份对齐方法,其特征在于,包括以下步骤:

步骤S1:多个摄像头各自将其通过检测器检测到的行人图像加入待关联的候选池中;

步骤S2:对待关联的候选池中属于不同摄像头的两个行人图像计算外观适配度;

步骤S3:将待关联的候选池中的行人图像两两配对计算完外观适配度后,根据每对行人外观适配度,结合时空关系,建立最小费用流图模型,求解最优的行人关联解;

步骤S4:根据步骤S3的关联结果,对行人进行保持原有标识或者赋予新的身份标识的操作;

在步骤S1中,所述检测器为Faster R-CNN;

步骤S2中,计算外观适配度的具体步骤采用步骤A21-步骤A23,或采用步骤B21-步骤B24;

其中,步骤A21-步骤A23具体为:

步骤A21:使用模糊C均值聚类F-CCT模型完成图像预处理,设行人图像A的整体性特征为 $X = \{x_1, x_2, \dots, x_N\}$,行人图像B的整体性特征为 $Y = \{y_1, y_2, \dots, y_N\}$;

步骤A22:将步骤S21处理后的图像作为融合细粒度表征的行人关联模型SAM-Dets的输入数据:以X为输入向量,Y为权值向量,通过所述融合细粒度表征的行人关联模型SAM-Dets编码行人A具有的局部细粒度特征 f_1 ;以Y为输入向量,X为权值向量,通过所述融合细粒度表征的行人关联模型SAM-Dets编码行人B具有的局部细粒度特征 f_2 ;

步骤A23:将 $f_s = (f_1 - f_2)^2$ 作为两个核大小为 $1 \times 1 \times 4096$ 的卷积层C的输入值,将softmax作为输出函数,输出一个二维向量 (q_1, q_2) ,表示输入两个对象属于现实世界中同一个人的概率值,作为外观适配度;

在步骤A22中,所述融合细粒度表征的行人关联模型SAM-Dets的结构包括:K个注意力分支和拼接层;每一所述注意力分支均包括以下六层,其中:

第一层为卷积层A,用于提取输入的行人整体特征的高层特征;

第二层为激活层,激活函数为softmax;

第三层为维度扩大层;

第四层为求和层,将行人整体特征与第三层获得的结果相加;

第五层为全局平均池化层,用于降低特征维度;

第六层为全连接层,用于完成输入向量与权值矩阵中的权值向量的内积计算;

所述拼接层将K个注意力分支得到的结果按通道拼接,输出行人局部细粒度特征;

所述卷积层A的卷积核尺寸为 1×1 ,步长为1;所述维度扩大层将通道维度扩大为512维;所述全局平均池化层的尺寸为 1×1 ,步长为1;

步骤B21-步骤B24具体为:

步骤B21:使用模糊C均值聚类F-CCT模型完成图像预处理,设行人图像A的整体性特征为 $X = \{x_1, x_2, \dots, x_N\}$,行人图像B的整体性特征为 $Y = \{y_1, y_2, \dots, y_N\}$;

步骤B22:将行人图像A和行人图像B经过DR-ResNet基础网络提取行人抽象特征;

步骤B23:使用卷积层B进一步提取目标行人高层特征,作为分类模型和融合细粒度表征的行人关联模型SAM-Dets的输入数据;所述分类模型分别输出行人图像A和行人图像B各自的身份识别表示号,所述融合细粒度表征的行人关联模型SAM-Dets输出行人A具有的局部细粒度特征 f_1 ,行人B具有的局部细粒度特征 f_2 ;

步骤B24:将 $f_s = (f_1 - f_2)^2$ 作为两个核大小为 $1 \times 1 \times 4096$ 的卷积层C的输入值,将softmax作为输出函数,输出一个二维向量 (q_1, q_2) ,表示输入两个对象属于现实世界中同一个人的概率值,作为外观适配度;

在步骤B22中,所述DR-ResNet基础网络包括两个权重共享的完全相同的深度卷积孪生神经基础网络模块R-ResNet;所述深度卷积孪生神经基础网络模块R-ResNet的结构包括四十九层卷积层、三个并行的卷积层、以及末端卷积层:

其中,第一卷积层的卷积核大小为 $(7, 7, 64)$,max-pooling为 $(3, 3)$,滑动步长为2;

第二卷积层到第四卷积层的卷积核大小分别为 $(1, 1, 64)$ 、 $(3, 3, 64)$ 、 $(1, 1, 256)$,激活函数都采用ReLU函数;三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第二卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第五卷积层到第七卷积层,以及第八卷积层到第十卷积层都采用与第二卷积层到第四卷积层相同的结构;

第十一卷积层到第十三卷积层的卷积核大小分别为 $(1, 1, 128)$ 、 $(3, 3, 128)$ 、 $(1, 1, 512)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第十一层卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第十四卷积层到第十六卷积层、第十七卷积层到第十九卷积层以及第二十卷积层到第二十二卷积层都采用了与第十一卷积层到第十三卷积层相同的结构;

第二十三卷积层到第二十五卷积层的卷积核大小分别为 $(1, 1, 256)$ 、 $(3, 3, 256)$ 、 $(1, 1, 1024)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第二十三卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第二十六卷积层到第二十八卷积层、第二十九卷积层到三十一卷积层、第三十二卷积层到第三十四卷积层、第三十五卷积层到三十七卷积层、第三十八卷积层到第四十卷积层均采用与第二十三卷积层到第二十五卷积层相同的结构;

第四十一卷积层到第四十三卷积层的卷积核大小分别为 $(1, 1, 512)$ 、 $(3, 3, 512)$ 、 $(1, 1, 2048)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第四十一卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第四十四卷积层到第四十六卷积层、第四十七卷积层到四十九卷积层采用与第四十一卷积层到第四十二卷积层相同的结构;

在第四十九卷积层后为三个并行的卷积层,每个卷积层,使用2048个卷积核,第一并行卷积层到第三并行卷积层的尺寸大小分别为 $(3, 3, 1024)$ 、 $(5, 5, 1024)$ 和 $(7, 7, 1024)$,通过一个连接层将三个并行的卷积层的通道进行合并,其后的max-pooling为 $(4, 4)$;

最后一层是使用1024个卷积核,且尺寸大小为 $(2, 2, 2048)$ 的末端卷积层;

所述卷积层B使用2个卷积核,尺寸大小为 $(1, 1, 4096)$;

所述融合细粒度表征的行人关联模型SAM-Dets的结构包括:K个注意力分支和拼接层;每一所述注意力分支均包括以下六层,其中:

第一层为卷积层A,卷积核尺寸为 1×1 ,步长为1;用于提取输入的行人整体特征的高层特征;

第二层为激活层,激活函数为softmax;

第三层为维度扩大层,将通道维度扩大为512维;

第四层为求和层,将行人整体特征与第三层获得的结果相加;

第五层为全局平均池化层,尺寸为 1×1 ,步长为1,用于降低特征维度;

第六层为全连接层,用于完成输入向量与权值矩阵中的权值向量的内积计算;

所述拼接层将K个注意力分支得到的结果按通道拼接,输出行人局部细粒度特征;

在步骤S3中,根据每对行人外观适配度,结合时空关系,建立最小费用流图模型,求解最优的行人关联解的具体过程,包括以下步骤:

步骤S31:设给定 t_p-1 时刻下的完成即时对齐后的费用流图为 G_{t_p-1} ,则当 t_p 时刻,为视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 和出视野行人集 $\bigcup_{i=1}^n OUT_i^{t_p}$ 中每一个目标新增进出两个节点,更新新增节点与源点、汇点间有向边连接;

步骤S32:根据视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 和出视野行人集 $\bigcup_{i=1}^n OUT_i^{t_p}$ 中两两目标间的行人外观适配度,更新相应节点间有向边,得到 t_p 时刻新的费用流图 G_{t_p} ;

步骤S33:删除所有对齐目标节点以及视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 剩余未对齐的目标节点,并将视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 剩余未对齐的目标节点为新进入的目标行人,得到费用流图 G_{t_p} ,等待下一时刻更新对齐;

在步骤S4中,对行人进行保持原有标识或者赋予新的身份标识的操作之后,交由跟踪器延续追踪;所述跟踪器采用KCF算法进行跟踪;所述KCF算法为每个行人分配一个跟踪器;

将摄像头的跟踪视野区域划分为核心区域和临界区域,在步骤S1中,通过所述检测器只检测临界区域的行人。

跨摄像头场景下行人身份对齐方法

技术领域

[0001] 本发明属于机器视觉、智能安防领域,尤其涉及一种跨摄像头场景下行人身份对齐方法。

背景技术

[0002] 随着经济社会的不断发展,人们对于安全的需求越来越大。因此,智能安防领域不断的被讨论和发展,使用范围也越来越大,类似于行人检测和跟踪等技术开始成为当前研究的热点问题。对于行人追踪问题,自然是跨摄像头下的研究才具有实际意义,且要求要对多个行人进行处理。尽管现在单摄像头场景下的行人跟踪技术已经相对成熟,但是多摄像头,尤其是非重叠视野区域情形下由于盲区的存在使得目标的时空信息变得不再可靠,从而给不同时刻不同空间下不同摄像头中同一目标的识别跟踪和检索造成了很大的困扰。所以,跨摄像头场景下的行人追踪技术这一研究正在被兴起。其中,最重要的部分便是在不同的摄像头下怎么让行人身份匹配。

[0003] 跨摄像头行人身份对齐主要是以行人为研究对象,关注非重叠视野区域的多摄像头多目标跟踪问题。这一问题,目前常见的解决机制是分为两个步骤:首先,使用检测和跟踪算法获得目标在单摄像头下得运行轨迹。其次,使用关联算法将摄像头间独立的行人运行轨迹进行关联整合,从而获得每个目标完整的运动轨迹。上述机制的局限在于只能处理离线的数据,本质上适用于检索场景,无法支持在线跟踪。究其原因在于,目标行人离开当前摄像头视域后,由于盲区的存在目标在进入下一摄像头视域时,时空信息均缺失,增加了将目标行人从上一个摄像头正确移交给下一个摄像头的难度。这种机制还产生了副效应:使得跨摄像头行人跟踪的结果严重依赖单摄像头下得行人跟踪效果。

[0004] 实现行人跨摄像头身份对齐的关键在于将不同视域内的同一目标行人正确关联。针对目前大部分跨摄像头行人跟踪算法中对于行人特征学习能力有限,并不能学习到较为鲁棒的行人特征。因此,最终影响到了后面的行人相似性度量的精度,最终造成了并不理想的数据关联结果。所以很难适应于跨摄像头行人跟踪的复杂环境。

[0005] 尽管现有的跨摄像头身份对齐相关的研究可以比较有效地解决一些对于离线数据的行人跟踪,但是并不能满足对于需要即时在线跟踪的要求,并且也不能在未知行人进出区域的时候进行有效的跟踪。

发明内容

[0006] 为了克服现有技术存在的空白和不足,本发明的方案在于解决跨摄像头下的行人关联问题,主要功能是给新进入目标分配新的身份表示或者与之前离开的目标关联成功,则将之前的目标赋予新的身份标识。其以对单摄像头下的目标行人进行持续正确的跟踪为基础,在目标离开摄像头视野区域后,进入盲区,跟踪断续,但当其再次出现在摄像头下时,能够重新识别出该行人并维持其身份标识不变延续跟踪。当检测器检测到新的行人时,将其加入待关联的候选池中,之后从候选池中选择一对行人使用F-CCT模型完成图像预处理,

处理后的图像将作为SAM-Dets模型的输入数据并求得两个行人的外观适配度。候选池中的行人两两配对计算完外观适配度后,根据每对行人适配度并结合时空关系建立最小费用流图模型求解最优的行人关联解;最后根据关联结果对行人保持原有标识或者赋予新的身份标识,交由跟踪器延续跟踪。

[0007] 本发明具体采用以下技术方案:

[0008] 一种跨摄像头场景下行人身份对齐方法,其特征在于,包括以下步骤:

[0009] 步骤S1:多个摄像头各自将其通过检测器检测到的行人图像加入待关联的候选池中;步骤S2:对待关联的候选池中属于不同摄像头的两个行人图像计算外观适配度;

[0010] 步骤S3:将待关联的候选池中的行人图像两两配对计算完外观适配度后,根据每对行人外观适配度,结合时空关系,建立最小费用流图模型,求解最优的行人关联解;

[0011] 步骤S4:根据步骤S3的关联结果,对行人进行保持原有标识或者赋予新的身份标识的操作。

[0012] 优选地,在步骤S1中,所述检测器为Faster R-CNN。

[0013] 优选地,步骤S2中,计算外观适配度具体包括以下步骤:

[0014] 步骤A21:使用模糊C均值聚类F-CCT模型完成图像预处理,设行人图像A的整体性特征为 $X = \{x_1, x_2, \dots, x_N\}$,行人图像B的整体性特征为 $Y = \{y_1, y_2, \dots, y_N\}$;

[0015] 步骤A22:将步骤S21处理后的图像作为融合细粒度表征的行人关联模型SAM-Dets的输入数据:以X为输入向量,Y为权值向量,通过所述融合细粒度表征的行人关联模型SAM-Dets编码行人A具有的局部细粒度特征 f_1 ;以Y为输入向量,X为权值向量,通过所述融合细粒度表征的行人关联模型SAM-Dets编码行人B具有的局部细粒度特征 f_2 ;

[0016] 步骤A23:将 $f_s = (f_1 - f_2)^2$ 作为两个核大小为 $1 \times 1 \times 4096$ 的卷积层C的输入值,将softmax作为输出函数,输出一个二维向量 (q_1, q_2) ,表示输入两个对象属于现实世界中同一个人的概率值,作为外观适配度。

[0017] 优选地,在步骤A22中,所述融合细粒度表征的行人关联模型SAM-Dets的结构包括:K个注意力分支和拼接层;每一所述注意力分支均包括以下六层,其中:

[0018] 第一层为卷积层A,用于提取输入的行人整体特征的高层特征;

[0019] 第二层为激活层,激活函数为softmax;

[0020] 第三层为维度扩大层;

[0021] 第四层为求和层,将行人整体特征与第三层获得的结果相加;

[0022] 第五层为全局平均池化层,用于降低特征维度;

[0023] 第六层为全连接层,用于完成输入向量与权值矩阵中的权值向量的内积计算;

[0024] 所述拼接层将K个注意力分支得到的结果按通道拼接,输出行人局部细粒度特征。

[0025] 优选地,所述卷积层A的卷积核尺寸为 1×1 ,步长为1;所述维度扩大层将通道维度扩大为512维;所述全局平均池化层的尺寸为 1×1 ,步长为1。

[0026] 优选地,步骤S2中,计算外观适配度具体包括以下步骤:

[0027] 步骤B21:使用模糊C均值聚类F-CCT模型完成图像预处理,设行人图像A的整体性特征为 $X = \{x_1, x_2, \dots, x_N\}$,行人图像B的整体性特征为 $Y = \{y_1, y_2, \dots, y_N\}$;

[0028] 步骤B22:将行人图像A和行人图像B经过DR-ResNet基础网络提取行人抽象特征;

[0029] 步骤B23:使用卷积层B进一步提取目标行人高层特征,作为分类模型和融合细粒

度表征的行人关联模型SAM-Dets的输入数据;所述分类模型分别输出行人图像A和行人图像B各自的身份识别表示号,所述融合细粒度表征的行人关联模型SAM-Dets输出行人A具有的局部细粒度特征 f_1 ,行人B具有的局部细粒度特征 f_2 。

[0030] 步骤B24:将 $f_s = (f_1 - f_2)^2$ 作为两个核大小为 $1 \times 1 \times 4096$ 的卷积层C的输入值,将softmax作为输出函数,输出一个二维向量 (q_1, q_2) ,表示输入两个对象属于现实世界中同一个人的概率值,作为外观适配度。

[0031] 优选地,所述DR-ResNet基础网络包括两个权重共享的完全相同的深度卷积孪生神经基础网络模块R-ResNet;所述深度卷积孪生神经基础网络模块R-ResNet的结构包括四十九层卷积层、三个并行的卷积层、以及末端卷积层:

[0032] 其中,第一卷积层的卷积核大小为 $(7, 7, 64)$,max-pooling为 $(3, 3)$,滑动步长为2;

[0033] 第二卷积层到第四卷积层的卷积核大小分别为 $(1, 1, 64)$ 、 $(3, 3, 64)$ 、 $(1, 1, 256)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第二卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第五卷积层到第七卷积层,以及第八卷积层到第十卷积层都采用与第二卷积层到第四卷积层相同的结构;

[0034] 第十一卷积层到第十三卷积层的卷积核大小分别为 $(1, 1, 128)$ 、 $(3, 3, 128)$ 、 $(1, 1, 512)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第十一层卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第十四卷积层到第十六卷积层、第十七卷积层到第十九卷积层以及第二十卷积层到第二十二卷积层都采用了与第十一卷积层到第十三卷积层相同的结构;

[0035] 第二十三卷积层到第二十五卷积层的卷积核大小分别为 $(1, 1, 256)$ 、 $(3, 3, 256)$ 、 $(1, 1, 1024)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第二十三卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第二十六卷积层到第二十八卷积层、第二十九卷积层到三十一卷积层、第三十二卷积层到第三十四卷积层、第三十五卷积层到三十七卷积层、第三十八卷积层到第四十卷积层均采用与第二十三卷积层到第二十五卷积层相同的结构;

[0036] 第四十一卷积层到第四十三卷积层的卷积核大小分别为 $(1, 1, 512)$ 、 $(3, 3, 512)$ 、 $(1, 1, 2048)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第四十一卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第四十四卷积层到第四十六卷积层、第四十七卷积层到四十九卷积层采用与第四十一卷积层到第四十二卷积层相同的结构;

[0037] 在第四十九卷积层后为三个并行的卷积层,每个卷积层,使用2048个卷积核,第一并行卷积层到第三并行卷积层的尺寸大小分别为 $(3, 3, 1024)$ 、 $(5, 5, 1024)$ 和 $(7, 7, 1024)$,通过一个连接层将三个并行的卷积层的通道进行合并,其后的max-pooling为 $(4, 4)$;

[0038] 最后一层是使用1024个卷积核,且尺寸大小为 $(2, 2, 2048)$ 的末端卷积层;

[0039] 所述卷积层B使用2个卷积核,尺寸大小为 $(1, 1, 4096)$;

[0040] 所述所述融合细粒度表征的行人关联模型SAM-Dets的结构包括:K个注意力分支和拼接层;每一所述注意力分支均包括以下六层,其中:

[0041] 第一层为卷积层A,卷积核尺寸为 1×1 ,步长为1;用于提取输入的行人整体特征的

高层特征；

[0042] 第二层为激活层，激活函数为softmax；

[0043] 第三层为维度扩大层，将通道维度扩大为512维；

[0044] 第四层为求和层，将行人整体特征与第三层获得的结果相加；

[0045] 第五层为全局平均池化层，尺寸为 1×1 ，步长为1，用于降低特征维度；

[0046] 第六层为全连接层，用于完成输入向量与权值矩阵中的权值向量的内积计算；

[0047] 所述拼接层将K个注意力分支得到的结果按通道拼接，输出行人局部细粒度特征。

[0048] 优选地，在步骤S3中，根据每对行人外观适配度，结合时空关系，建立最小费用流图模型，求解最优的行人关联解的具体过程，包括以下步骤：

[0049] 步骤S31：设给定 $t_p - 1$ 时刻下的完成即时对齐后的费用流图为 G_{t_p-1} ，则当 t_p 时刻，为视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 和出视野行人集 $\bigcup_{i=1}^n OUT_i^{t_p}$ 中每一个目标新增进出两个节点，更新新增节点与源点、汇点间有向边连接；

[0050] 步骤S32：根据视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 和出视野行人集 $\bigcup_{i=1}^n OUT_i^{t_p}$ 中两两目标间的行人外观适配度，更新相应节点间有向边，得到 t_p 时刻新的费用流图 G_{t_p} ；

[0051] 步骤S33：删除所有对齐目标节点以及视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 剩余未对齐的目标节点，并将视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 剩余未对齐的目标节点为新进入的目标行人，得到费用流图 G_{t_p} ，等待下一时刻更新对齐。

[0052] 优选地，在步骤S4中，对行人进行保持原有标识或者赋予新的身份标识的操作之后，交由跟踪器延续追踪；所述跟踪器采用采用KCF算法进行跟踪；所述KCF算法为每个行人分配一个跟踪器。

[0053] 优选地，将摄像头的跟踪视野区域划分为核心区域和临界区域，在步骤S1中，通过所述检测器只检测临界区域的行人。

[0054] 相较于现有技术，本发明及其优选方案实现了跨摄像头行人的在线跟踪，其识别准确、效率高，且不受目标离开摄像头视野区域后，进入盲区，跟踪断续的影响。

[0055] 其中，跟踪功能的实现由现有成熟FasterR-CNN实现行人检测，由KCF算法实现在线行人跟踪；而作为实现本发明及优选方案的核心在于根据时空信息和行人相似性值建立最小费用流图模型即时完成行人身份对齐任务，以及整合了F-CCT和SAM-Dets模型或SAM-Dets模型和DR-ResNet网络模型融合，解决了行人的外观适配度量度的问题。

附图说明

[0056] 下面结合附图和具体实施方式对本发明进一步详细的说明：

[0057] 图1是本发明实施例1整体流程示意图；

[0058] 图2是本发明实施例SAM-Dets模型结构示意图；

[0059] 图3是本发明实施例SAM-Dets模型网络结构示意图；

[0060] 图4是本发明实施例2行人细粒度关联模型示意图；

[0061] 图5是本发明实施例2在跨摄像头场景下行人身份对齐模型效果示意图1；

[0062] 图6是本发明实施例2在跨摄像头场景下行人身份对齐模型效果示意图2。

具体实施方式

[0063] 为使本专利的特征和优点能更明显易懂,下文特举2个实施例,作详细说明如下:

[0064] 如图1所示,在本发明的第一个实施例中,实现跨摄像头场景下行人身份对齐的整体方案包括以下步骤:

[0065] 步骤S1:当行人进入到摄像头视野,多个摄像头各自将其通过检测器检测到的行人图像加入待关联的候选池中。

[0066] 在本实施例中,检测器选用基于深度学习的目标检测代表性方法Faster R-CNN,其在选择候选框时使用RPN以检测网络共享全图的卷积特征,让分类和回归任务有相同的卷积特征。

[0067] 步骤S2:对待关联的候选池中属于不同摄像头的两个行人图像计算外观适配度。

[0068] 在实施例中,计算外观适配度具体包括以下步骤:

[0069] 步骤A21:使用模糊C均值聚类F-CCT模型完成图像预处理,设行人图像A的整体性特征为 $X = \{x_1, x_2, \dots, x_N\}$,行人图像B的整体性特征为 $Y = \{y_1, y_2, \dots, y_N\}$;使用模糊聚类算法对图像进行聚类域划分,通过源和目标图像的聚类域匹配,在聚类域间实现局部颜色亮度迁移,并引入隶属度因子提高颜色亮度迁移效果。

[0070] 步骤A22:将步骤S21处理后的图像作为融合细粒度表征的行人关联模型SAM-Dets的输入数据:以X为输入向量,Y为权值向量,通过融合细粒度表征的行人关联模型SAM-Dets编码行人A具有的局部细粒度特征 f_1 ;以Y为输入向量,X为权值向量,通过融合细粒度表征的行人关联模型SAM-Dets编码行人B具有的局部细粒度特征 f_2 。

[0071] 如图2所示,在本实施例中,融合细粒度表征的行人关联模型SAM-Dets,由多个注意力组成,模型中每个分支具有相同的功能模块,输入数据都为基础网络提取的行人全局特征。每一条注意力分支都一次经过局部检测器、全局池化和线性嵌入三个模块,最后将K条分支的结果按通道进行拼接获得注意力模型的完整输出结果。

[0072] 在局部检测器模块中,首先获取输入信息的高层特征,使用softmax函数对高层特征完成归一化得到符合概率分布取值区间的注意力权重,为下一步的求和操作需要先对高层特征的维度进行扩大处理,最后使用加权求和函数获得相应特征的注意力分配概率分布。全局池化模块和线性嵌入模块中,注意力分配概率分布筛选保留对应的行人特征,最后输出具有注意力分布的高层行人特征。

[0073] 如图3所示,具体地,融合细粒度表征的行人关联模型SAM-Dets的网络结构包括:K个注意力分支和拼接层;每一注意力分支均包括以下六层,其中:

[0074] 第一层为卷积层A,用于提取输入的行人整体特征的高层特征;

[0075] 第二层为激活层,激活函数为softmax;

[0076] 第三层为维度扩大层;

[0077] 第四层为求和层,将行人整体特征与第三层获得的结果相加;

[0078] 第五层为全局平均池化层,用于降低特征维度;

[0079] 第六层为全连接层,用于完成输入向量与权值矩阵中的权值向量的内积计算;

[0080] 拼接层将K个注意力分支得到的结果按通道拼接,输出行人局部细粒度特征。

[0081] 卷积层A的卷积核尺寸为 1×1 ,步长为1;维度扩大层将通道维度扩大为512维;全局平均池化层的尺寸为 1×1 ,步长为1。

[0082] 步骤A23:将要计算输入的一对行人的相似性值,转化成对 f_1 和 f_2 特征的相似性比较。引入一个无参数层Square层来对 f_1 和 f_2 特征求解平方差,作为 f_1 和 f_2 相似性比较层,并记该Square层为: $f_s = (f_1 - f_2)^2$;将 $f_s = (f_1 - f_2)^2$ 作为两个核大小为 $1 \times 1 \times 4096$ 的卷积层C的输入值,将softmax作为输出函数,输出一个二维向量 (q_1, q_2) ,表示输入两个对象属于现实世界中同一个人的概率值,作为外观适配度。

[0083] 进一步地,根据获得的一对行人间的相似性概率值作为图的权值,将新进入的行人和待关联的目标行人分别作为两个不同的顶点集合,建立带权值匹配图;通过求解最大权值匹配图问题的解,获得新进入的行人与等待关联的目标行人间的数据关联的解。

[0084] 步骤S3:将待关联的候选池中的行人图像两两配对计算完外观适配度后,根据每对行人外观适配度,结合时空关系,建立最小费用流图模型,求解最优的行人关联解。

[0085] 其具体包括以下步骤:

[0086] 步骤S31:设给定 $t_p - 1$ 时刻下的完成即时对齐后的费用流图为 G_{t_p-1} ,则当 t_p 时刻,为视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 和出视野行人集 $\bigcup_{i=1}^n OUT_i^{t_p}$ 中每一个目标新增进出两个节点,更新新增节点与源点、汇点间有向边连接;

[0087] 步骤S32:根据视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 和出视野行人集 $\bigcup_{i=1}^n OUT_i^{t_p}$ 中两两目标间的行人外观适配度,更新相应节点间有向边,得到 t_p 时刻新的费用流图 G_{t_p} ;

[0088] 步骤S33:删除所有对齐目标节点以及视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 剩余未对齐的目标节点,并将视野内行人集 $\bigcup_{i=1}^n IN_i^{t_p}$ 剩余未对齐的目标节点为新进入的目标行人,得到费用流图 G_{t_p} ,等待下一时刻更新对齐。

[0089] 步骤S4:根据步骤S3的关联结果,对行人进行保持原有标识或者赋予新的身份标识的操作,并交由跟踪器延续追踪;跟踪器采用采用KCF算法进行跟踪;KCF算法为每个行人分配一个跟踪器。该算法在目标区域形成循环矩阵,再利用循环矩阵在傅里叶空间可以对角化等一些性质,通过回归岭回归得到通用的预测公式。

[0090] 同时,本实施例中,将摄像头的跟踪视野区域划分为核心区域和临界区域,在步骤S1中,通过检测器只检测临界区域的行人。这是考虑到任何目标的进入和离开都必先经过临界区域,因此,只考察临界区域内的行人具有合理空间转移关系,以最大程度地保证后续行人对齐求解的普适性。

[0091] 在本发明的第二个实施例中,如图4所示,其提供了步骤S2:对待关联的候选池中属于不同摄像头的两个行人图像计算外观适配度的另一种优选的可实现方案,其具体包括以下步骤:

[0092] 步骤B21:使用模糊C均值聚类F-CCT模型完成图像预处理,设行人图像A的整体性特征为 $X = \{x_1, x_2, \dots, x_N\}$,行人图像B的整体性特征为 $Y = \{y_1, y_2, \dots, y_N\}$;

[0093] 步骤B22:将行人图像A和行人图像B经过DR-ResNet基础网络提取行人抽象特征;

[0094] 步骤B23:使用卷积层B进一步提取目标行人高层特征,作为分类模型和融合细粒度表征的行人关联模型SAM-Dets的输入数据;分类模型分别输出行人图像A和行人图像B各自的身份识别表示号,融合细粒度表征的行人关联模型SAM-Dets输出行人A具有的局部细

粒度特征 f_1 ,行人B具有的局部细粒度特征 f_2 。

[0095] 步骤B24:将 $f_s = (f_1 - f_2)^2$ 作为两个核大小为 $1 \times 1 \times 4096$ 的卷积层C的输入值,将softmax作为输出函数,输出一个二维向量 (q_1, q_2) ,表示输入两个对象属于现实世界中同一个人的概率值,作为外观适配度。

[0096] 其中,DR-ResNet基础网络包括两个权重共享的完全相同的深度卷积孪生神经基础网络模块R-ResNet;深度卷积孪生神经基础网络模块R-ResNet的结构包括四十九层卷积层、三个并行的卷积层、以及末端卷积层:

[0097] 其中,第一卷积层的卷积核大小为 $(7, 7, 64)$,max-pooling为 $(3, 3)$,滑动步长为2;

[0098] 第二卷积层到第四卷积层的卷积核大小分别为 $(1, 1, 64)$ 、 $(3, 3, 64)$ 、 $(1, 1, 256)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第二卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第五卷积层到第七卷积层,以及第八卷积层到第十卷积层都采用与第二卷积层到第四卷积层相同的结构;

[0099] 第十一卷积层到第十三卷积层的卷积核大小分别为 $(1, 1, 128)$ 、 $(3, 3, 128)$ 、 $(1, 1, 512)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第十一层卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第十四卷积层到第十六卷积层、第十七卷积层到第十九卷积层以及第二十卷积层到第二十二卷积层都采用了与第十一卷积层到第十三卷积层相同的结构;

[0100] 第二十三卷积层到第二十五卷积层的卷积核大小分别为 $(1, 1, 256)$ 、 $(3, 3, 256)$ 、 $(1, 1, 1024)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第二十三卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第二十六卷积层到第二十八卷积层、第二十九卷积层到三十一卷积层、第三十二卷积层到第三十四卷积层、第三十五卷积层到三十七卷积层、第三十八卷积层到第四十卷积层均采用与第二十三卷积层到第二十五卷积层相同的结构;

[0101] 第四十一卷积层到第四十三卷积层的卷积核大小分别为 $(1, 1, 512)$ 、 $(3, 3, 512)$ 、 $(1, 1, 2048)$,激活函数都采用ReLU函数;该三层卷积层和激活函数组成一个卷积块,将该卷积块的输入值既作为第四十一卷积层的输入值,也作为卷积块的第三层激活函数的输入值;第四十四卷积层到第四十六卷积层、第四十七卷积层到四十九卷积层采用与第四十一卷积层到第四十二卷积层相同的结构;

[0102] 在第四十九卷积层后为三个并行的卷积层,每个卷积层,使用2048个卷积核,第一并行卷积层到第三并行卷积层的尺寸大小分别为 $(3, 3, 1024)$ 、 $(5, 5, 1024)$ 和 $(7, 7, 1024)$,通过一个连接层将三个并行的卷积层的通道进行合并,其后的max-pooling为 $(4, 4)$;

[0103] 最后一层是使用1024个卷积核,且尺寸大小为 $(2, 2, 2048)$ 的末端卷积层;

[0104] 卷积层B使用2个卷积核,尺寸大小为 $(1, 1, 4096)$ 。

[0105] 如图5、图6所示,通过本实施例的方案,与现有的离线跨摄像头跟踪的方案均实现了跨摄像头跟踪的性能,区别在于,本实施例方案直接可以实现用于在线的跟踪。

[0106] 本专利不局限于上述最佳实施方式,任何人在本专利的启示下都可以得出其它各种形式的跨摄像头场景下行人身份对齐方法,凡依本发明申请专利范围所做的均等变化与修饰,皆应属本专利的涵盖范围。

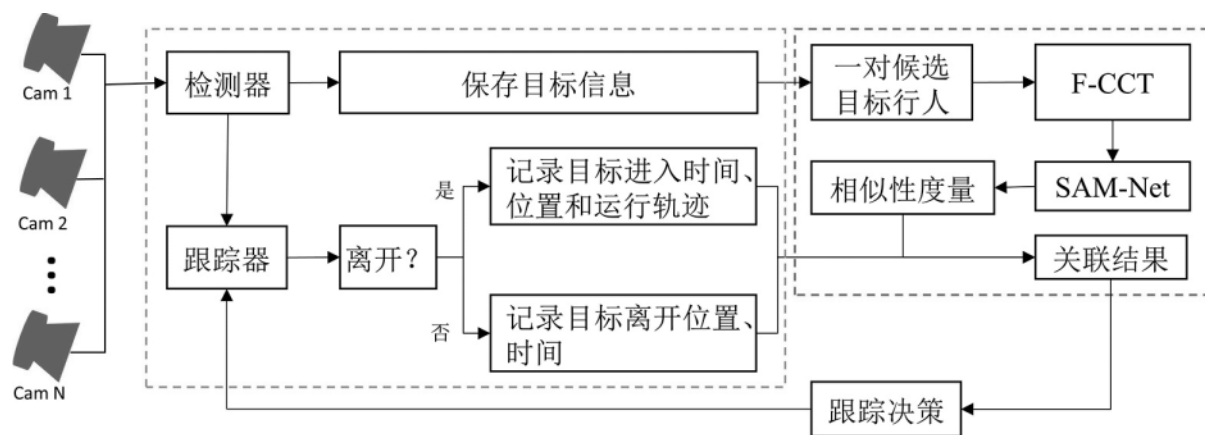


图1

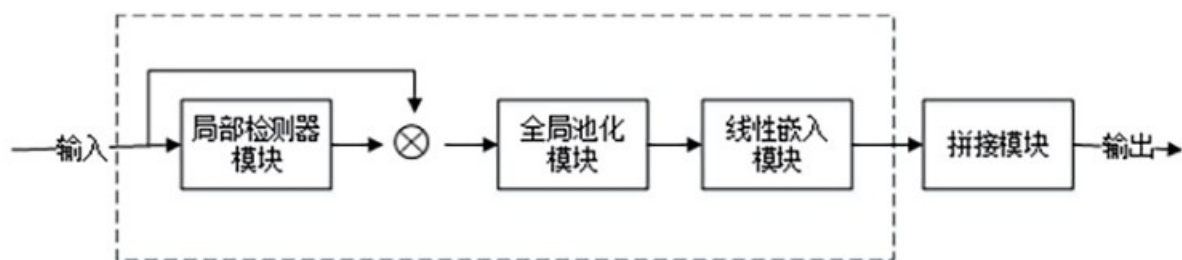


图2

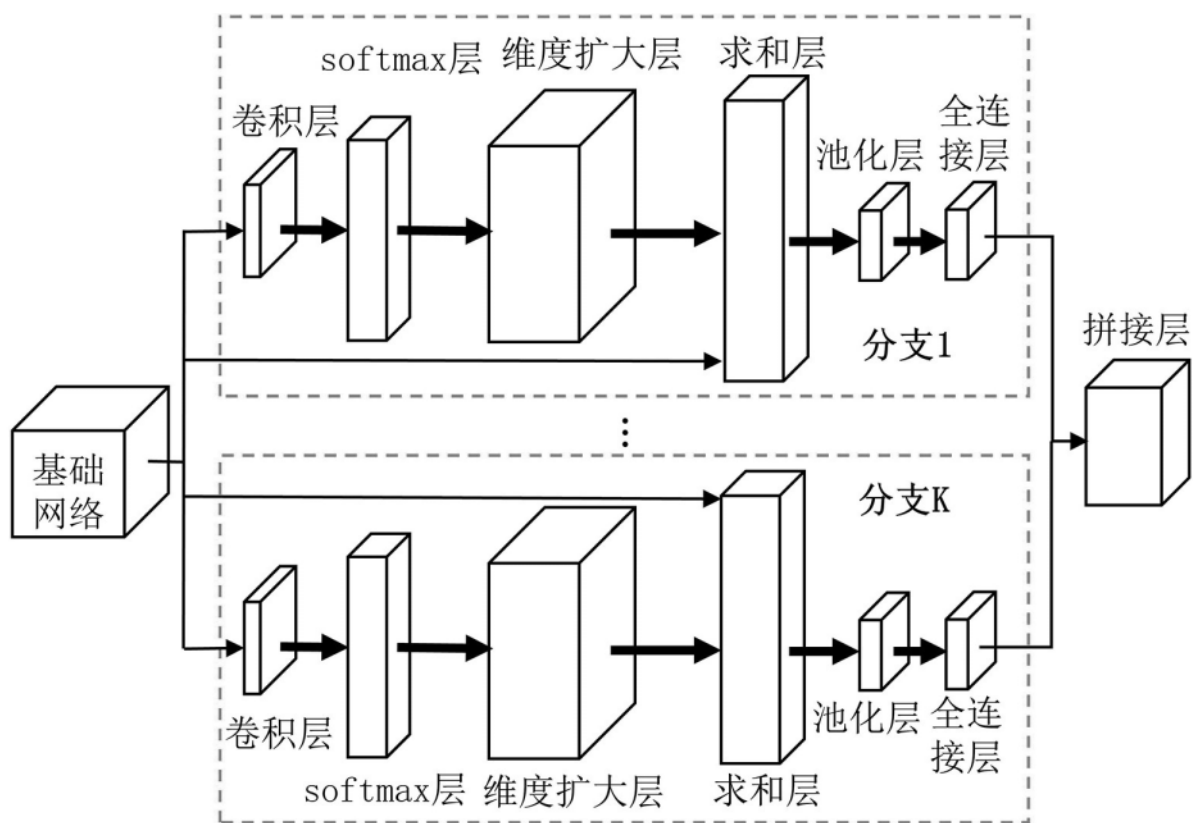


图3

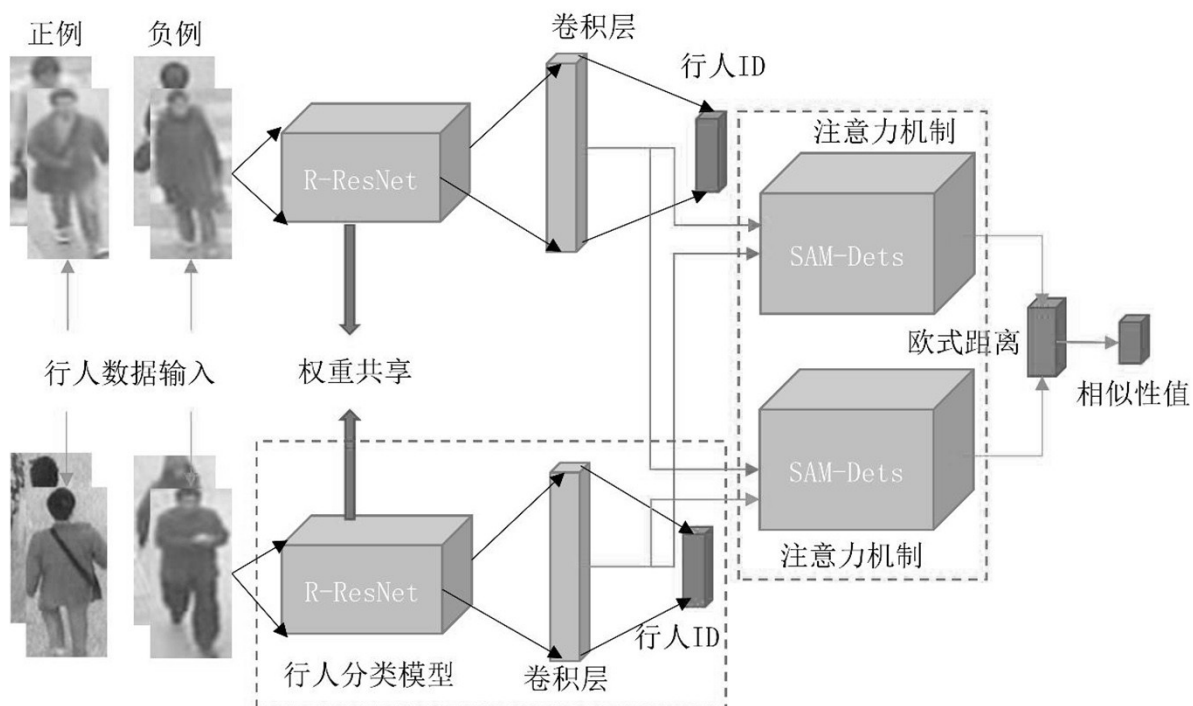


图4



图5



图6