



(12) 发明专利申请

(10) 申请公布号 CN 112200841 A

(43) 申请公布日 2021. 01. 08

(21) 申请号 202011055504.5

G06N 3/04 (2006.01)

(22) 申请日 2020.09.30

(71) 申请人 杭州海宴科技有限公司

地址 310052 浙江省杭州市滨江区六和路  
368号一幢(北)一层B1178-3室

(72) 发明人 叶利华 谢东福 羊海丰 牟森  
刘燕

(74) 专利代理机构 杭州奥创知识产权代理有限  
公司 33272

代理人 王佳健

(51) Int. Cl.

G06T 7/292 (2017.01)

G06T 7/194 (2017.01)

G06T 7/136 (2017.01)

G06K 9/62 (2006.01)

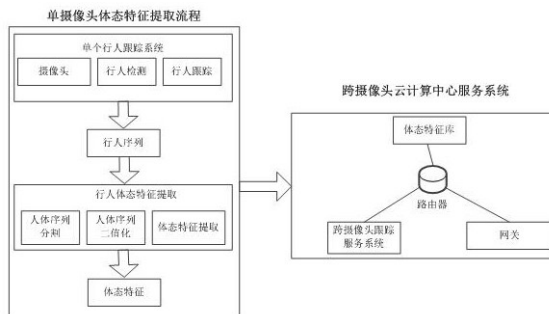
权利要求书2页 说明书6页 附图1页

(54) 发明名称

一种基于行人体态的跨域多摄像头跟踪方法和装置

(57) 摘要

本发明公开了一种基于行人体态的跨域多摄像头跟踪方法和装置。本发明包括多个位置关联摄像头;人体目标检测;单摄像头内行人跟踪;行人人体序列采集;行人人体序列分割;行人人体序列二值化;二值化行人序列特征提取;跨摄像头云计算中心服务系统检索各摄像头所获取的行人体态特征;根据时间关系和体态特征识别和标定行人轨迹。本发明通过采用行人体态为基础,通过提取序列人体体态的融合特征,实现监控跨域摄像头内行人的跟踪。该方法除避免现有方法受行人外观因素影响,如衣着颜色、款式、发型等,还不受光照、天气变化、穿着变化等因素影响,提高了在复杂条件和环境下的跟踪能力。



1. 一种基于行人体态的跨域多摄像头跟踪方法,通过非接触的方式采集行人体态生物特征,实现基于行人体态的跨域多摄像头行人跟踪,其特征在于该方法包括以下步骤:

- S1. 设置多个摄像头,且保证摄像头能以一定的角度拍摄到行人的侧身;
- S2. 各摄像头拍摄可视范围内的识别区域;
- S3. 各摄像头对识别区域进行行人检测;
- S4. 各摄像头对所检测到的行人进行跟踪;
- S5. 各摄像头对所跟踪的行人人体序列进行分割;
- S6. 各摄像头对分割的行人人体序列二值化;
- S7. 各摄像头对二值化后的行人人体轮廓序列进行特征提取;
- S8. 比对各摄像头获取的行人体态特征值并识别和标定行人;
- S9. 重复S2~S8上述步骤完成跨域多摄像头的行人识别;

其中s8具体是:将每个摄像头所提取的特征都传送到中心云服务器模块,中心云服务器模块中存储所有摄像头中所抓取的人体体态特征序列以及相应的时间轴;体态特征采用欧式距离算法求出距离,假如距离小于设定的阈值,则认定为同一个行人,并进行标定。

2. 根据权利要求1所述的一种基于行人体态的跨域多摄像头跟踪方法,其特征在于:s3中采用基于深度学习端到端的目标检测,并采用Tensor RT加速技术实现在监控视频中快速的检测人体位置,采用矩形框的方式标出行人。

3. 根据权利要求1所述的一种基于行人体态的跨域多摄像头跟踪方法,其特征在于:s4中采用交并比、局部色彩信息相似、行人语义特征相似度以及行人前序轨迹方向实现行人在当前摄像头内部的跟踪,得到同一行人的矩形框图像序列。

4. 根据权利要求1所述的一种基于行人体态的跨域多摄像头跟踪方法,其特征在于:S5~S6具体是:采用基于语义的深度学习人体分割方法,去除背景,并对人体图像二值化。

5. 根据权利要求1所述的一种基于行人体态的跨域多摄像头跟踪方法,其特征在于:s7具体是:输入一组或多组固定数量连续人体二值化体态,采用并行卷积神经网络分别提取特征并进行融合生成,输入的体态图像序列被映射成指定长度的一维特征序列。

6. 一种基于行人体态的跨域多摄像头跟踪装置,其特征在于,包括:

多个位置的关联摄像头,每个摄像头能以一定的角度拍摄到行人的侧身;

识别区域获取模块,由各摄像头拍摄可视范围内的识别区域;

行人检测模块,由各摄像头对识别区域进行行人检测;

行人跟踪模块,由各摄像头对所检测到的行人进行跟踪;

行人人体序列分割模块,由各摄像头对所跟踪的行人人体序列进行分割;

行人人体序列二值化模块,由各摄像头对分割的行人人体序列二值化;

行人人体轮廓序列特征提取模块,由各摄像头对二值化后的行人人体轮廓序列进行特征提取;

行人体态特征值识别和标定模块,比对各摄像头获取的行人体态特征值并识别和标定行人;

其中行人体态特征值识别和标定模块是将每个摄像头所提取的特征都传送到中心云服务器模块,中心云服务器模块中存储所有摄像头中所抓取的人体体态特征序列以及相应的时间轴,体态特征采用欧式距离算法求出距离,假如距离小于设定的阈值,则认定为同

一个行人,并进行标定。

7.根据权利要求6所述的一种基于行人体态的跨域多摄像头跟踪方法,其特征在于:所述行人检测模块采用基于深度学习端到端的目标检测,并采用Tensor RT加速技术实现在监控视频中快速的检测人体位置,并采用矩形框的方式标出行人。

8.根据权利要求6所述的一种基于行人体态的跨域多摄像头跟踪方法,其特征在于:所述行人跟踪模块采用交并比、局部色彩信息相似、行人语义特征相似度以及行人前序轨迹方向实现行人在当前摄像头内部的跟踪,得到同一行人的矩形框图像序列。

9.根据权利要求6所述的一种基于行人体态的跨域多摄像头跟踪方法,其特征在于:所述行人人体序列分割模块,采用基于语义的深度学习进行人体分割。

10.根据权利要求6所述的一种基于行人体态的跨域多摄像头跟踪方法,其特征在于:所述行人人体轮廓序列特征提取模块,输入一组或多组固定数量连续人体二值化体态,采用并行卷积神经网络分别提取特征并进行融合生成,输入的体态图像序列被映射成指定长度的一维特征序列。

## 一种基于行人体态的跨域多摄像头跟踪方法和装置

### 技术领域

[0001] 本发明属于图像处理领域,具体涉及一种基于行人体态的跨域多摄像头跟踪方法和装置。

### 背景技术

[0002] 当前摄像头被广泛的用于各类场景,由此产生了大量的视频数据,如何自动分析这些视频数据,成为迫切的需要。目标跟踪为自动视频处理提供了重要的数据基础,其在交通安全防卫、事故预警、监控等方面有广阔的应用前景。

[0003] 单个摄像头的视野范围有限,而外部环境中行人的活动范围较大,单个摄像头无法拍摄到行人在大范围内的整个活动过程。网络摄像头能将摄像头拍摄的视频实时的传输到视频监控中心,视频监控中心将相关的拍摄视频进行汇聚,可极大地扩展监控范围,实现对部分区域的全时段无死角的监控。然而,要获取目标行人在特定区域内的活动轨迹,采用人力方式观看监控视频,非常困难,尤其室外监控视频一般距离较远,行人目标较小,人眼查看效率低,工作量大。在面对海量视频资料时,要查找特定行人目标在此范围内的活动轨迹需要投入的人力物力成本非常巨大。因此,采用人工智能实现跨域跟踪技术实现多摄像头之间的行人目标跟踪,是有效解决海量视频下的行人目标检索与活动轨迹跟踪的方法。

[0004] 由于室外环境的复杂性、人员外形相似性、摄像头拍摄角度差异等因素,导致在不同摄像头和不同时间段所拍摄的同一行人的外貌特征差异性大,而行人中存在大量身高、衣着、发型等外貌特征相似的情况,在不同摄像头所拍摄的远景中,要进行有效地分辨很困难。因此,如何有效地实现跨域跟踪在视频监控领域是一个具有挑战性的问题。

[0005] 现有的跟踪方法包括单摄像头和多摄像头,主要提取行人的外观特征,其中衣着颜色、款式等对跟踪效果有很大影响,这些方法主要通过提取目标行人的外观特征。行人的衣着颜色、款式、发型等外观变化直接影响特征结果,也直接影响行人跟踪效果。除了行人自身外观变化对跟踪效果有影响外,还受光照、雨雾等外部因素变化的影响,在复杂的室外环境中跟踪效果有待提高。

### 发明内容

[0006] 本发明要解决的技术问题是提供一种基于行人体态的跨域多摄像头跟踪方法和装置,通过使用多种方法分别用于解决单个摄像头内部和跨摄像头之间的行人目标跟踪问题,本发明可以适用于不同天气条件和多种成像方式,如白天、夜晚以及可见光和红外等成像条件。

[0007] 为了达到上述目标,本发明采用了如下技术方案:

[0008] 本发明的一方面:提供了一种基于行人体态的跨域多摄像头跟踪方法,通过非接触的方式采集行人体态生物特征,实现基于行人体态的跨域多摄像头行人跟踪,该方法包括以下步骤:

[0009] S1. 设置多个摄像头,且保证摄像头能以一定的角度拍摄到行人的侧身;

- [0010] S2.各摄像头拍摄可视范围内的识别区域;
- [0011] S3.各摄像头对识别区域进行行人检测;
- [0012] S4.各摄像头对所检测到的行人进行跟踪;
- [0013] S5.各摄像头对所跟踪的行人人体序列进行分割;
- [0014] S6.各摄像头对分割的行人人体序列二值化;
- [0015] S7.各摄像头对二值化后的行人人体轮廓序列进行特征提取;
- [0016] S8.比对各摄像头获取的行人体态特征值并识别和标定行人;
- [0017] S9.重复S2~S8上述步骤完成跨域多摄像头的行人识别;
- [0018] 其中s8具体是:将每个摄像头所提取的特征都传送到中心云服务器模块,中心云服务器模块中存储所有摄像头中所抓取的人体体态特征序列以及相应的时间轴;体态特征采用欧式距离算法求出距离,假如距离小于设定的阈值,则认定为同一个行人,并进行标定。
- [0019] 进一步说,s3中采用基于深度学习端到端的目标检测,并采用Tensor RT加速技术实现在监控视频中快速的检测人体位置,采用矩形框的方式标出行人。
- [0020] 进一步说,s4中采用交并比、局部色彩信息相似、行人语义特征相似度以及行人前序轨迹方向实现行人在当前摄像头内部的跟踪,得到同一行人的矩形框图像序列。
- [0021] 进一步说,S5~S6具体是:采用基于语义的深度学习人体分割方法,去除背景,并对人体图像二值化;为提高解决人体大小对识别的结果影响,分别对X和Y两个轴方向进行投影,并按照特征提取的尺寸对二值图进行宽高比例填充,使图像缩放对人体外形特征不影响;对二值化后的人体分割序列中的每张图像分别进行缩放,缩放完成后,分别对宽和高分别加两个像素为最终输入图像的尺寸,用于提取行人体态特征。
- [0022] 进一步说,s7具体是:输入一组或多组固定数量连续人体二值化体态,采用并行卷积神经网络分别提取特征并进行融合生成;输入的体态图像序列被映射成指定长度的一维特征序列。
- [0023] 本发明的另一方面:提供了一种基于行人体态的跨域多摄像头跟踪装置,包括:
- [0024] 多个位置的关联摄像头,每个摄像头能以一定的角度拍摄到行人的侧身;
- [0025] 识别区域获取模块,由各摄像头拍摄可视范围内的识别区域;
- [0026] 行人检测模块,由各摄像头对识别区域进行行人检测;
- [0027] 行人跟踪模块,由各摄像头对所检测到的行人进行跟踪;
- [0028] 行人人体序列分割模块,由各摄像头对所跟踪的行人人体序列进行分割;
- [0029] 行人人体序列二值化模块,由各摄像头对分割的行人人体序列二值化;
- [0030] 行人人体轮廓序列特征提取模块,由各摄像头对二值化后的行人人体轮廓序列进行特征提取;
- [0031] 行人体态特征值识别和标定模块,比对各摄像头获取的行人体态特征值并识别和标定行人;
- [0032] 其中行人体态特征值识别和标定模块是将每个摄像头所提取的特征都传送到中心云服务器模块,中心云服务器模块中存储所有摄像头中所抓取的人体体态特征序列以及相应的时间轴,体态特征采用欧式距离算法求出距离,假如距离小于设定的阈值,则认定为同一个行人,并进行标定。

[0033] 进一步说,所述行人检测模块采用基于深度学习端到端的目标检测,并采用 Tensor RT加速技术实现在监控视频中快速的检测人体位置,并采用矩形框的方式标出行人。

[0034] 进一步说,所述行人跟踪模块采用交并比、局部色彩信息相似、行人语义特征相似度以及行人前序轨迹方向实现行人在当前摄像头内部的跟踪,得到同一行人的矩形框图像序列。

[0035] 进一步说,所述行人人体序列分割模块,采用基于语义的深度学习进行人体分割。

[0036] 进一步说,所述行人人体轮廓序列特征提取模块,输入一组或多组固定数量连续人体二值化体态,采用并行卷积神经网络分别提取特征并进行融合生成;输入的体态图像序列被映射成指定长度的一维特征序列。

[0037] 本发明的有益效果:

[0038] 本发明通过采用行人体态为基础,通过提取序列人体体态的融合特征,实现监控跨域摄像头内行人的跟踪。该方法除避免现有方法受行人外观因素影响,如衣着颜色、款式、发型等,还不受光照、天气变化、穿着变化等因素影响,提高了在复杂条件和环境下的跟踪能力。

## 附图说明

[0039] 图1为本发明的跨域多摄像头行人体态跟踪系统框架图;

[0040] 图2为本发明中的一种基于行人体态的跨域多摄像头跟踪处理流程图。

## 具体实施方式

[0041] 为了使本发明解决的技术问题、实施方案、优点更加清晰,下面结合系统实例以及上述附图进行详细描述。

[0042] 一种基于行人体态的跨域多摄像头跟踪方法,涉及以下技术内容:多个位置关联摄像头;人体目标检测;单摄像头内行人跟踪;行人人体序列采集;行人人体序列分割;行人人体序列二值化;二值化行人序列特征提取;跨摄像头云计算中心服务系统检索各摄像头所获取的行人体态特征;根据时间关系和体态特征识别和标定行人轨迹。

[0043] 所述摄像头是用于安装在一定高度位置用于拍摄行人视频图像的设备,所用设备成像可以为可见光、红外等方式。

[0044] 所述位置关联是指摄像头在地理位置上有联系,即这些摄像头在部署的空间位置上存在前后左右的关联关系。

[0045] 所述人体目标检测是指采用深度卷积神经网络的目标检测方法,对所拍摄视频图像内含的人体检出,给出位置,将人体采用矩形框框出。

[0046] 所述单摄像头内行人跟踪是指采用IOU方法对所检测的人体进行位置重合计算,当IOU的值超过设定阈值,则说明为同一个行人,假如小于阈值并且行人没有走出所拍摄的范围,则通过提取行人前一帧与当前帧的行人外部形状特征进行比较,找出相似度最高的值,当超过设定阈值则说明为同一行人。

[0047] 所述行人人体序列采集通过行人跟踪方法,按照时间顺序将所检出同一个行人的身体序列放入数组。

[0048] 所述行人人体序列分割是指对所取的行人人体序列,采用基于语义的深度卷积神经网络模型对所提取的人体矩形进行分割,将人体部分与背景分离并进行二值化,去除背景,得到二值化的行人人体序列。

[0049] 所述的二值化行人序列特征提取是指对所提取二值化的人体序列,采用合成特征提取方法将同一摄像头内所提取的同一个人序列采用模型方法映射生成一个特征序列,用于表示该行人的体态信息。

[0050] 所述跨摄像头云计算中心服务系统检索是将各摄像头所获取的行人体态特征进行比较,实现对不同摄像头同一个行人进行匹配。

[0051] 所述行人轨迹是指对所匹配不同摄像头的行人,基于时间和位置关系绘制出目标行人的活动轨迹,实现基于行人体态的跨域多摄像头行人跟踪目标。

[0052] 基于上述技术内容,本发明主要是通过非接触的方式采集行人体态生物特征,实现基于行人体态的跨域多摄像头行人跟踪,具体实施步骤如下:

[0053] S1. 设置多个摄像头,且保证摄像头能以一定的角度拍摄到行人的侧身;

[0054] S2. 各摄像头拍摄可视范围内的识别区域;

[0055] S3. 各摄像头对所拍摄的区域进行行人检测;

[0056] S4. 各摄像头对所检测到的行人进行跟踪;

[0057] S5. 各摄像头对所跟踪的行人人体序列进行分割;

[0058] S6. 各摄像头对分割的行人人体序列二值化;

[0059] S7. 各摄像头对二值化后的行人人体轮廓序列进行特征提取;

[0060] S8. 比对各摄像头获取的行人步态特征值并识别和标定行人;

[0061] S9. 重复S2~S8上述步骤完成跨域多摄像头的行人识别。

[0062] 1) 上述的S3步骤,所述的对所拍摄的区域进行行人检测,具体为采用基于深度学习端到端的目标检测,并采用Tensor RT加速技术实现在监控视频中快速的检测人体位置并采用矩形框的方式标出行人,用于进行行人跟踪。

[0063] 2) 上述的S4步骤,所述的对所检测到的行人进行跟踪,具体对检测所获得的行人矩形框为采用交并比、局部色彩信息相似、行人语义特征相似度以及行人前序轨迹方向实现行人在当前摄像头内部的跟踪,得到同一行人的矩形框图像序列。

[0064] 其中交并比的计算方法如公式(1):

$$[0065] \quad IoU = \frac{area(C) \cap area(G)}{area(C) \cup area(G)} \quad (1)$$

[0066] 公式(1)中,C和G分别为摄像视频中前后两帧中检测到任意两个人体矩形框,在下一帧中找到与C人体IoU值最大的G人体,当IoU的值超过设定阈值0.3后,从人体检测的中间部位裁剪一块为各自宽度的10%的图像,并以较小的块为参考,将大的块缩放为一样的大小,并比较两块图像色彩的相似性,比较方法如公式(2):

$$[0067] \quad R = \frac{abs(I_a - I_b)}{I_a} \quad (2)$$

[0068] 当R值小于0.2时,可以认定为同一个行人。假如不符合上述条件时,假如IoU的值大于0且R值大于0.2时,对检测的行人分别采用基于深度学习语义人体分割方法去除行人的背景信息,对分割后的行人采用深度卷积神经网络模型提取行人的语义特征信息,比较两个

行人语义特征相似度,当相似度大于0.7时,认为是同一人。

[0069] 假如最大IoU为0,则说明没有直接的交并人体,根据C的轨迹方向计算方法如公式(3):

$$[0070] \quad \alpha_c = \frac{I_{y_i} - I_{y_{i-1}}}{I_{x_i} - I_{x_{i-1}}} \quad (3)$$

[0071] 其中i代表帧的序号, $I_{y_i}$ 与 $I_{x_i}$ 分别为人体的中心点。在第i+1帧上,找到与 $\alpha_c$ 方向最接近的人体G,且角度差小于 $5^\circ$ ,假如满足上述条件,从人体检测的中间部位裁剪一块为各自宽度的10%的图像,并以较小的块为参考,将大的块缩放为一样的大小,并比较两块图像色彩的相似性,比较方法如公式(2),当R值小于0.2时,可以认定为同一个行人。假如不符合上述条件时,假如阈值大于0或R值大于0.2时,对检测的行人分别采用基于深度学习语义人体分割方法去除行人的背景信息,对分割后的行人采用深度卷积网络模型提取行人的语义特征信息,比较两个行人语义特征相似度,当相似度大于0.7时,认为是同一人。

[0072] 考虑到行人在行走过程中,其衣着特征一般不会变化,对行人的衣着色彩分布进行统计,用于辅助行人跟踪。当跟踪的行人丢失后,代表该行人走出该摄像头的拍摄范围,提取跟踪过程中体态最为明显的一段序列作为该行人在当前摄像头范围的体态序列用于提取体态特征。

[0073] 3) 上述的S5~S6步骤,对所检测的同一行人人体序列采用基于语义的深度学习人体分割方法,去除背景,并对人体图像二值化。为提高解决人体大小对识别的结果影响,分别对X和Y两个轴方向进行投影,并按照特征提取的尺寸对二值图进行宽高比例填充,使图像缩放对人体外形特征不影响。对二值化后的人体分割序列中的每张图像分别进行缩放,完成后分别对宽和高分别加两个像素为最终输入图像的尺寸,用于提取行人体态特征。

[0074] 4) 上述的S7步骤,人体的体态特征获得方法是输入一组或多组固定数量连续人体二值化体态,采用并行卷积神经网络分别提取特征并进行融合生成。模型对体态序列的时间间隔没有固定要求,仅要求体态按时间顺序输入即可。输入的体态图像序列被映射成指定长度的一维特征序列。

[0075] 5) 上述的S8步骤,每个摄像头所提取的特征都传送到中心云服务器模块,云服务器模块中存储所有摄像头中所抓取的人体体态特征序列以及相应的时间轴。摄像头的体态特征采用欧式距离算法求出距离,假如距离小于20认定为同一个行人,并进行标定。

[0076] 6) 上述的S2~S8步骤中所述的行人检测、跟踪、分割、体态序列特征提取操作,采用带有GPU处理能力嵌入式或服务器进行处理和特征提取。所述的跨域多摄像头行人跟踪通过CPU处理完成。

[0077] 结合附图进行再次说明,本发明提供的跨域多摄像头行人体态跟踪系统功能如图1所示,系统以多个单摄像头体态特征提取为基础,将单个摄像头以特定的角度分别部署在不同的区域,采用人体检测模型将所监视区域拍摄到视频中的人体采用矩形框出,并跟踪该行人在当前摄像头下的轨迹,得到行人人体序列。

[0078] 当体态序列数超过100或该行人离开监控区域,则该行人的体态序列采集结束。对所采集到的一个行人人体序列提取体态特征需要进行如下预处理流程,为排除背景环境对提取特征的影响,对获取的人体序列进行人体分割,去除背景。对分割后的行人矩形框进行



二值化处理,以消除背景和行人穿着变化的影响。单张行人轮廓所包含的信息量有限,通过对行人行走过程提取融合特征,获取信息量更大的每个行人特有的行走体态特征用于描述行人体态,达到更好的区分效果。

[0079] 每个摄像头内部提取的行人体态特征,在时间同步的前提下,将特征提交到跨摄像头云计算中心服务系统。跨摄像头跟踪服务系统中,将收到的体态特征进行归一化处理,并在体态特征库中进行搜索,查看该行人是否已经在历史记录中有过标注,假如有则做好记录并存储,假如没有则增加到体态特征库,其中比较的时候采用欧式距离计算体态特征的相似度,假如超过预设阈值则认为是一行人。跨摄像头跟踪服务系统根据行人在各摄像头出现的时间、摄像头的位置,绘制出目标行人的运动轨迹。上述内容可以采用图2的处理流程进行描述。

[0080] 综上,本发明提供了一种以人体体态特征实现跨域摄像头行人跟踪方法,本发明所采集的特征可以不受外部环境和行人自身外观变化的影响,适用于室外复杂环境下的行人跨域摄像头跟踪。对于摄像头内和摄像头之间分别采用不同的跟踪方法,降低了系统的整体计算机复杂度,提高了行人的跨域跟踪能力。

[0081] 以上为本发明的最佳实施方式,依据本发明公开的内容,本领域的普通技术人员能够显而易见地想到的一些雷同、替代方案,均应落入本发明保护的范围内。

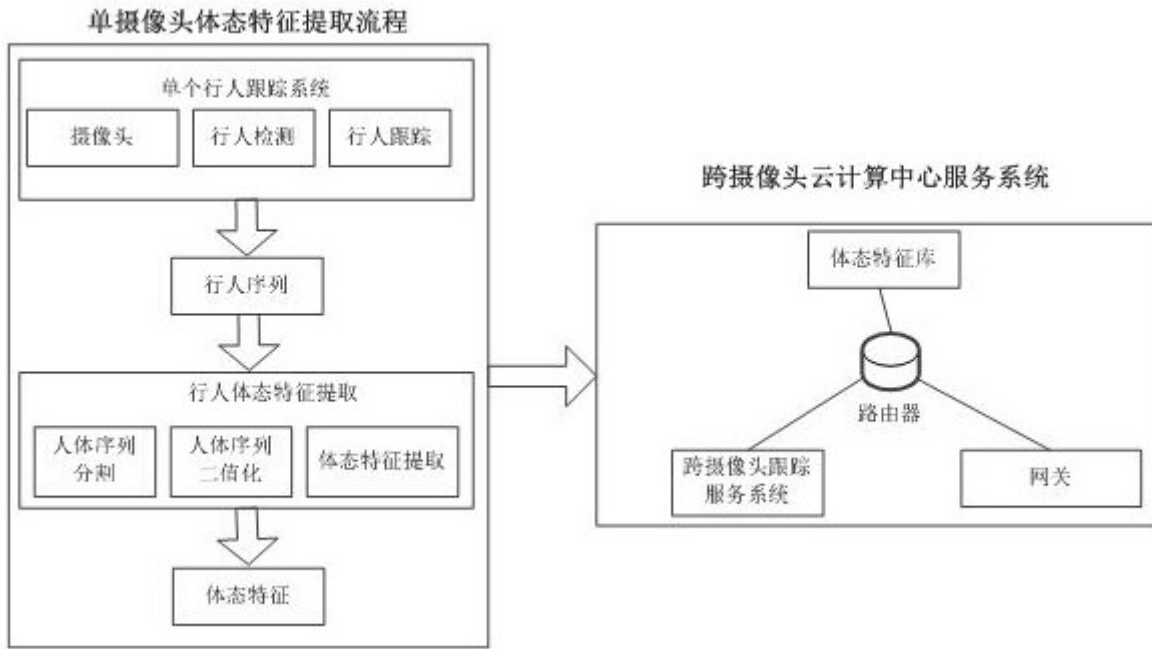


图 1

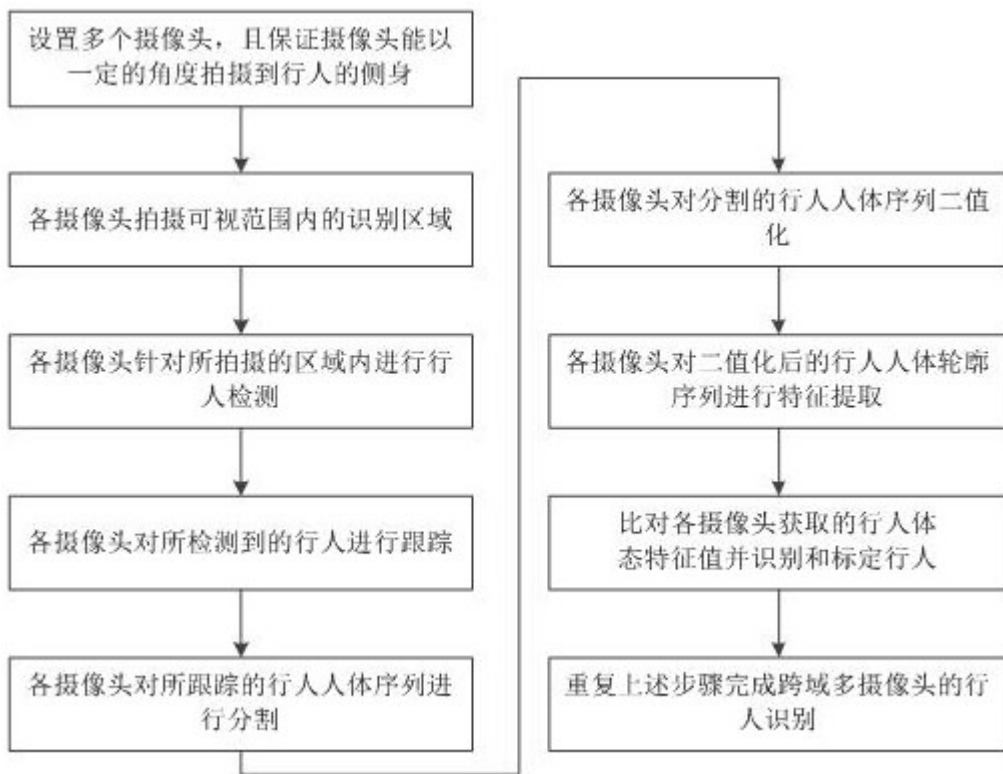


图 2