



(12) 发明专利申请

(10) 申请公布号 CN 105469059 A

(43) 申请公布日 2016. 04. 06

(21) 申请号 201510868820. 7

(22) 申请日 2015. 12. 01

(71) 申请人 上海电机学院

地址 200240 上海市闵行区江川路 690 号

(72) 发明人 牛震亚 赵雷 苏庆刚 田阔

(74) 专利代理机构 上海思微知识产权代理事务

所(普通合伙) 31237

代理人 菅秀君

(51) Int. Cl.

G06K 9/00(2006. 01)

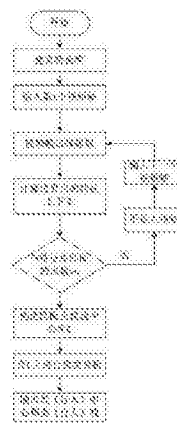
权利要求书2页 说明书4页 附图2页

(54) 发明名称

一种视频中的人物识别、定位和统计方法

(57) 摘要

本发明提供了一种视频中的人物识别、定位和统计方法,从视频图像中每个点的形状上下文的匹配的角度来实现对于视频中行人的识别、定位和统计,通过建立包含样本点的形状上下文特征的标准特征库,与输入的要求识别的视频中每一帧图像上的采样点进行匹配,这样无需对标准特征库进行大量的训练,只需建立包含典型的人体姿势的标准特征库,且即使人体具有复杂性,则将人体图像分割为每个采样点,由每个采样点与标准特征库中的样本点匹配,这种方法识别度较高,减少了识别误差,此外这种方法应用面较广,可以进一步应用于视频中的主体识别和主体跟踪等重要问题。



1. 一种视频中的人物识别、定位和统计方法,其特征在于,包括如下步骤:

步骤一:提供待训练的视频图像,对所述待训练的视频图像进行采样,得到各个样本点,计算每个样本点的形状上下文特征,并建立包含样本点的形状上下文特征的标准特征库;

步骤二:提供待测试的视频图像,在每一帧待测试的视频图像中生成采样点,将各个采样点与步骤一中的标准特征库中的样本点进行匹配,设定第一阈值,当找到匹配的样本点的采样点的大于该第一阈值时,则可认定该帧待测试的视频图像为人物帧,并计算得到所有待测试的视频图像中每个人物帧的候选中心集合;

步骤三:采用聚类算法对候选中心集合进行聚类分析,则聚类的数目为视频帧中人像的数目,每个聚类中心点为该人像的中心点。

2. 根据权利要求1所述的视频中的人物识别、定位和统计方法,其特征在于,步骤一中对待训练的视频图像进行边缘提取、均匀采样,得到人像形状的样本点集合 $I = \{p_1, p_2, \dots, p_k\}$,其中 p_1, p_2, \dots, p_k 为人像形状上的各个样本点,在样本点集合中对人像轮廓上各标准部位上的样本点进行形状上下文特征提取,建立包含有提取的样本点的形状上下文特征的标准特征库。

3. 根据权利要求2所述的视频中的人物识别、定位和统计方法,其特征在于,所述各标准部位为人物的头部、肩部、腿部和肘部。

4. 根据权利要求2所述的视频中的人物识别、定位和统计方法,其特征在于,对待训练的视频图像进行边缘提取的方法为:手工分割出每一帧待训练的视频图像中未知物体的前景掩码,得到前景掩码的边界及边界的信息。

5. 根据权利要求4所述的视频中的人物识别、定位和统计方法,其特征在于,建立标准特征库的方法为:在前景掩码的边界的各标准部位上提取样本点,构成包含所有所述提取样本点信息的标准特征库 $FB = \{f_i\}$, $f_i = \{s_i, \delta_i\}$,其中 f_i 为标准特征库中每个样本点的形状上下文特征值, s_i 是每个样本点的形状上下文, δ_i 是标准特征库中模型物体中心的坐标。

6. 根据权利要求5所述的视频中的人物识别、定位和统计方法,其特征在于,步骤二中对于每一帧待测试的视频图像,利用索贝尔算子提取每一帧待测试的视频图像中的边界及边界信息,并在提取的边界上等距离生成采样点,计算每个采样点的形状上下文特征并在标准特征库中查找与该采样点匹配的样本点。

7. 根据权利要求6所述的视频中的人物识别、定位和统计方法,其特征在于,采用 $\chi^2(X, Y)$ 来计算每个采样点的形状上下文特征与标准特征库中样本点的匹配程度,具体为设 X 和 Y 分别是待测试的视频图像的视频帧上的采样点和标准特征库中任意一个样本点的形状上

下文特征向量,其中 $X = \{x_1, x_2, \dots, x_i\}$, $Y = \{y_1, y_2, \dots, y_i\}$, $\chi^2(X, Y) = \frac{1}{2} \sum_{i=1}^n \frac{|x_i - y_i|^2}{x_i + y_i}$,

设定 $\chi^2(X, Y)$ 的第一阈值,当 $\chi^2(X, Y)$ 小于等于第一阈值时,则该待测试的视频图像的视频帧上的采样点和所述标准特征库中样本点匹配,设定在标准特征库中找到匹配的样本点的采样点的个数的第二阈值,当相匹配的点的个数大于设定的第二阈值时,则该待测试的视频图像的视频帧为视频帧为人物帧,即待测试的视频图像的视频帧上包含有人像。

8. 根据权利要求7所述的视频中的人物识别、定位和统计方法,其特征在于,计算在所

述标准特征库中找到匹配的样本点的采样点的形状上下文特征值 $\phi_i = \{\eta_i, \varepsilon_i\}$ ，其中 η_i 是该采样点的形状上下文， ε_i 则等于与该采样点所匹配的样本点对应的 δ_i ，每个 ε_i 成为所述人物帧中一个候选中心，所有的候选中心组成所述人物帧的候选中心集合 C 。

9. 根据权利要求1所述的视频中的人物识别、定位和统计方法，其特征在于，步骤三中所述聚类算法为硬聚类算法中的K-means算法。

一种视频中的人物识别、定位和统计方法

技术领域

[0001] 本发明涉及一种人物识别、定位和统计方法,尤其涉及一种视频中的人物识别、定位和统计方法。

背景技术

[0002] 人物识别是计算机图像和视频处理中的重要问题之一。图像和视频中的人物识别的本质是给定训练和测试图像集,根据给定的前提条件对图像中的主体进行分类和标定。有效的目标识别可以发现图像中的重要主体,包括人物或其他要进行识别的各类物体,从而为进一步的识别和跟踪提供依据和线索。目前,已经有很多文献提出和创建了图像或视频中的主体识别模型。例如,在识别具有纹理特征的物体时,可以采用常用的纹理特征算子对物体进行分类。在另一些方法中,通过给定一些固定的场景,利用全局或局部特征来分析相关图像。还有一些方法提出了模仿物体各个关节部位的空间分布来进行分析,例如Weber等人提出通过局部特征来计算目标物体关节部位空间分布的方法。Agarwal等创建了大量的物体部位库,利用特征分类来实现识别各个物体部位的空间分布。其他方法包括利用变形模板或主动外观模型等进行学习,以及结合图分割方法和识别同步进行识别的方法等。

[0003] 上述方法都源于使用目标物体的空间分布来进行识别和分类。其中全局特征和局部特征方法由于是较为被动地学习,需要大量训练数据,增加了识别的难度。而通过关节空间分布的方法由于人体先天的复杂性,因此仅限于较少部位之间识别。部位库方法需要不断重复观察空间中各个主体小部位的出现,增加了数据量和处理的复杂度。变形模板方法需要事先做出符合人体形状姿势的各种模板,而应用图分割方法目前只能识别较少的目标物体。

发明内容

[0004] 本发明提供一种人物识别、定位和统计方法,目的在于克服上述方法的不足,针对给定的视频文件,提出基于形状上下文特征算子匹配和基于聚类分析的行人识别、定位和统计的方法。

[0005] 为了实现上述目的,本发明提供了一种视频中的人物识别、定位和统计方法,包括如下步骤:

[0006] 步骤一:提供待训练的视频图像,对所述待训练的视频图像进行采样,得到各个样本点,计算每个样本点的形状上下文特征,并建立包含样本点的形状上下文特征的标准特征库;

[0007] 步骤二:提供待测试的视频图像,在每一帧待测试的视频图像中生成采样点,将各个采样点与步骤一中的标准特征库中的样本点进行匹配,设定第一阈值,当找到匹配的样本点的采样点的大于该第一阈值时,则可认定该帧待测试的视频图像为人物帧,并计算得到所有待测试的视频图像中每个人物帧的候选中心集合;

[0008] 步骤三:采用聚类算法对候选中心集合进行聚类分析,则聚类的数目为视频帧中

人像的数目,每个聚类中心点为该人像的中心点。

[0009] 作为优选,步骤一中对待训练的视频图像进行边缘提取、均匀采样,得到人像形状的样本点集合 $I = \{p_1, p_2, \dots, p_k\}$,其中 p_1, p_2, \dots, p_k 为人像形状上的各个样本点,在样本点集合中对人像轮廓上各标准部位上的样本点进行形状上下文特征提取,建立包含有提取的样本点的形状上下文特征的标准特征库。

[0010] 作为优选,所述各标准部位为人物的头部、肩部、腿部和肘部。

[0011] 作为优选,对待训练的视频图像进行边缘提取的方法为:手工分割出每一帧待训练的视频图像中未知物体的前景掩码,得到前景掩码的边界及边界的信息。

[0012] 作为优选,建立标准特征库的方法为:在前景掩码的边界的各标准部位上提取样本点,构成包含有所有所述提取样本点信息的标准特征库 $FB = \{f_i\}$, $f_i = \{s_i, \delta_i\}$,其中 f_i 为标准特征库中每个样本点的形状上下文特征值, s_i 是每个样本点的形状上下文, δ_i 是标准特征库中模型物体中心的坐标。

[0013] 作为优选,步骤二中对于每一帧待测试的视频图像,利用索贝尔算子提取每一帧待测试的视频图像中的边界及边界信息,并在提取的边界上等距离生成采样点,计算每个采样点的形状上下文特征并在标准特征库中查找与该采样点匹配的样本点。

[0014] 作为优选,采用 $x^2(X, Y)$ 来计算每个采样点的形状上下文特征与标准特征库中样本点的匹配程度,具体为设 X 和 Y 分别是待测试的视频图像的视频帧上的采样点和标准特征库中任意一个样本点的形状上下文特征向量,其中 $X = \{x_1, x_2, \dots, x_i\}$, $Y = \{y_1, y_2, \dots, y_i\}$, $x^2(X, Y) = \frac{1}{2} \sum_{i=1}^n \frac{|x_i - y_i|^2}{x_i + y_i}$,设定 $x^2(X, Y)$ 的第一阈值,当 $x^2(X, Y)$ 小于等于第一阈

值时,则该待测试的视频图像的视频帧上的采样点和所述标准特征库中样本点匹配,设定在标准特征库中找到匹配的样本点的采样点的个数的第二阈值,当相匹配的点的个数大于设定的第二阈值时,则该待测试的视频图像的视频帧为人物帧,即待测试的视频图像的视频帧上包含有人像。

[0015] 作为优选,计算在所述标准特征库中找到匹配的样本点的采样点的形状上下文特征值 $\varphi_i = \{\eta_i, \varepsilon_i\}$,其中 η_i 是该采样点的形状上下文, ε_i 则等于与该采样点所匹配的样本点对应的 δ_i ,每个 ε_i 成为所述人物帧中一个候选中心,所有的候选中心组成所述人物帧的候选中心集合 C 。

[0016] 作为优选,步骤三中所述聚类算法为硬聚类算法中的K-means算法。

[0017] 与现有技术相比,本发明提供了一种视频中的人物识别、定位和统计方法,包括如下步骤:

[0018] 步骤一:提供待训练的视频图像,对所述待训练的视频图像进行采样,得到各个样本点,计算每个样本点的形状上下文特征,并建立包含样本点的形状上下文特征的标准特征库;

[0019] 步骤二:提供待测试的视频图像,在每一帧待测试的视频图像中生成采样点,将各个采样点与步骤一中的标准特征库中的样本点进行匹配,设定第一阈值,当找到匹配的样本点的采样点的大于该第一阈值时,则可认定该帧待测试的视频图像为人物帧,并计算得到所有待测试的视频图像中每个人物帧的候选中心集合;

[0020] 步骤三:采用聚类算法对候选中心集合进行聚类分析,则聚类的数目为人物帧中

人像的数目,每个聚类中心点为该人像的中心点。

[0021] 本发明的从视频图像中每个点的形状上下文的匹配的角度来实现对于视频中行人的识别、定位和统计,通过建立包含样本点的形状上下文特征的标准特征库,与输入的要求识别的视频中每一帧图像上的采样点进行匹配,这样无需对标准特征库进行大量的训练,只需建立包含典型的人体姿势的标准特征库,且即使人体具有复杂性,则将人体图像分割为每个采样点,由每个采样点与标准特征库中的样本点匹配,这种方法识别度较高,减少了识别误差,此外这种方法应用面较广,可以进一步应用于视频中的主体识别和主体跟踪等重要问题。

附图说明

[0022] 图1为本发明提供的方法流程图;

[0023] 图2为训练过程与标准特征库的建立过程示意图。

具体实施方式

[0024] 为了使本发明的内容更加清楚和易懂,下面结合具体实施例对本发明的内容进行详细描述。

[0025] 目前,图像中常用的特征点有纹理特征、空间分布特征、形状特征等。本发明采用形状上下文特征,它是一种利用物体表面轮廓表征物体形状信息的特征算子,定义为特征点周围 n_r 个半径上的直方图区间和 n_θ 个角度方向上的直方图区间,其值是一个特征向量 $h = \{h_1, h_2, \dots, h_{n_r n_\theta}\}$,这里 $h_i (i=1, 2, \dots, n_r n_\theta)$ 是直方图中每个区间中像素点的数目。请参照图2中的特征点A,其中 $n_r=2, n_\theta=8$,图中以A为圆心形成两个圆,外圆所包含的平面空间即为特征点A的直方图区间, $n_r \times n_\theta = 16$,则特征点A的形状上下文特征即为 $h = \{h_1, h_2, \dots, h_{16}\}$ 。

[0026] 请参照图1,本发明提供的基于形状上下文特征的视频中行人识别与统计流程具体步骤如下:

[0027] 步骤一:使用训练图像建立标准特征库。

[0028] 提供若干待训练的视频图像,对这些视频图像的每一帧图像进行边缘提取、均匀采样,具体为对于每一帧待训练视频图像,手工分割出该视频图像上物体的前景掩码,只有处于前景掩码中的边界的信息才用来计算形状上下文直方图。然后利用上述前景掩码的边界构造标准特征库,也就是在前景掩码的边界上取样本点,得到一个行人形状图像上的样本点集合 $I = \{p_1, p_2, \dots, p_k\}$,其中 p_1, p_2, \dots, p_k 为人像形状上的各个样本点,并对该样本点集合中,位于行人形状或者轮廓上各标准部位上的样本点进行形状上下文特征提取,构成标准特征库 $FB = \{f_i\}$, $f_i = \{s_i, \delta_i\}$,其中 f_i 为标准特征库中每个样本点的形状上下文特征值, s_i 是每个样本点的形状上下文, δ_i 是标准特征库中模型物体中心的坐标,图2中A点箭头顶端的点即为模型物体的中心,标准特征库中所包含的信息对应若干个模型物体。

[0029] 这里的各标准部位为人物的头部、肩部、腿部和肘部。因为大部分图片中的行人与背景间的互相遮挡,会造成身体某些部位的缺失,因此在样本点集合I上只提取各标准部位点的形状上下文特征。

[0030] 步骤二:测试。首先进行测试输入的视频中人物帧的自动提取。这里输入的视频即

为待测试的视频图像,即给定一段视频文件,包含的视频帧中可能有人像的内容,也可能没有这些内容。

[0031] 对于每一帧待测试的视频图像,首先利用索贝尔算子提取边缘信息,即利用索贝尔算子计算出该帧视频图像中未知物体的边界信息以及边界点的相关坐标,对应形成未知物体的边界轮廓。其次,在该边界轮廓的边界线上等距离生成采样点,计算每个采样点的形状上下文特征并在标准特征库中查找与之最匹配的样本点,这里采用 χ^2 来计算二者之间的匹配程度,这是一种表征向量之间距离的参数,根据 $\chi^2(X, Y)$ 大小定义第一阈值,如 $\chi^2(X, Y)$ 大小接近于0,那么可设定第一阈值为0.5,设X和Y分别是待测试的视频图像的视频帧上的采样点和标准特征库中任意一个样本点的形状上下文特征向量,其中 $X = (x_1, x_2, \dots, x_i)$, $Y = (y_1, y_2, \dots, y_i)$,则

$$[0032] \quad \chi^2(X, Y) = \frac{1}{2} \sum_{i=1}^n \frac{|x_i - y_i|^2}{x_i + y_i}$$

[0033] 当 $\chi^2(X, Y)$ 的值小于等于第一阈值时,则可认为该采样点与标准特征库中的该样本点匹配,当 $\chi^2(X, Y)$ 的值大于第一阈值时,则可认为该采样点与标准特征库中的该样本点不匹配,有些采样点在标准特征库中无法找到与之相匹配的样本点,设定在标准特征库中找到匹配的样本点的采样点的个数的第二阈值,如n,若找到相匹配的样本点的采样点的个数大于设定的第二阈值即n时,则该帧待测试的视频图像为图像帧,即该帧视频图像中包含有人像。

[0034] 计算上述在标准特征库中找到匹配的样本点的采样点的特征值 $\varphi_i = \{\eta_i, \varepsilon_i\}$, 其中 η_i 是该采样点的形状上下文, ε_i 则等于与该采样点所匹配的样本点对应的 δ_i ,每个 ε_i 成为所述图像帧中一个候选中心,所有的候选中心组成所述图像帧的候选中心集合C。

[0035] 步骤三:采用K-means算法对候选中心集合C进行聚类分析,则聚类的数目为帧中图像或者行人的数目,每个聚类中心点为该帧待测试视频图像中该行人的中心点。

[0036] 本发明的从视频图像中每个点的形状上下文的匹配的角度来实现对于视频中行人的识别、定位和统计,通过建立包含样本点的形状上下文特征的标准特征库,与输入的要求识别的视频中每一帧图像上的采样点进行匹配,这样无需对标准特征库进行大量的训练,只需建立包含典型的人体姿势的标准特征库,且即使人体具有复杂性,则将人体图像分割为每个采样点,由每个采样点与标准特征库中的样本点匹配,这种方法识别度较高,减少了识别误差,此外这种方法应用面较广,可以进一步应用于视频中的主体识别和主体跟踪等重要问题。

[0037] 可以理解的是,虽然本发明已以较佳实施例披露如上,然而上述实施例并非用以限定本发明。对于任何熟悉本领域的技术人员而言,在不脱离本发明技术方案范围情况下,都可利用上述揭示的技术内容对本发明技术方案作出许多可能的变动和修饰,或修改为等同变化的等效实施例。因此,凡是未脱离本发明技术方案的内容,依据本发明的技术实质对以上实施例所做的任何简单修改、等同变化及修饰,均仍属于本发明技术方案保护的范围内。

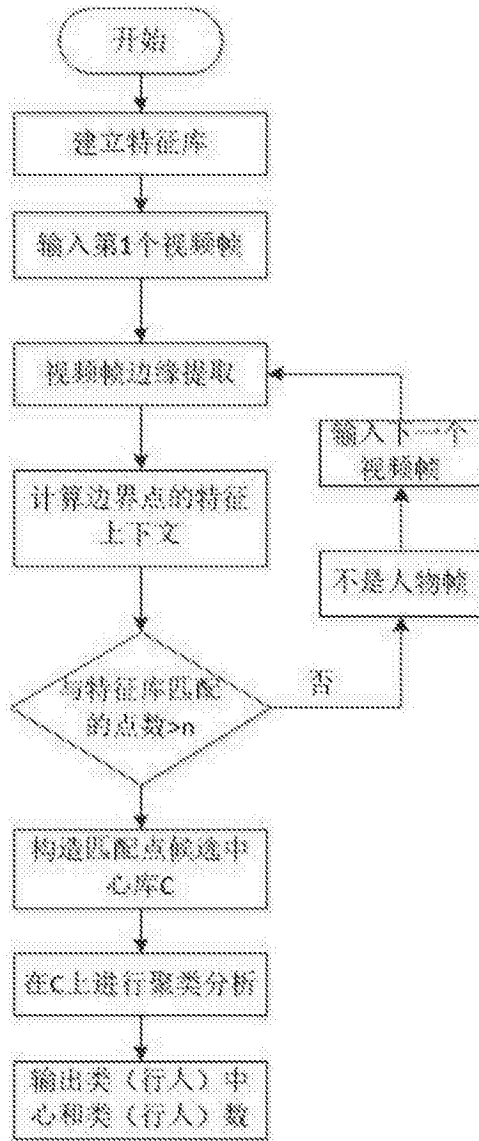


图1

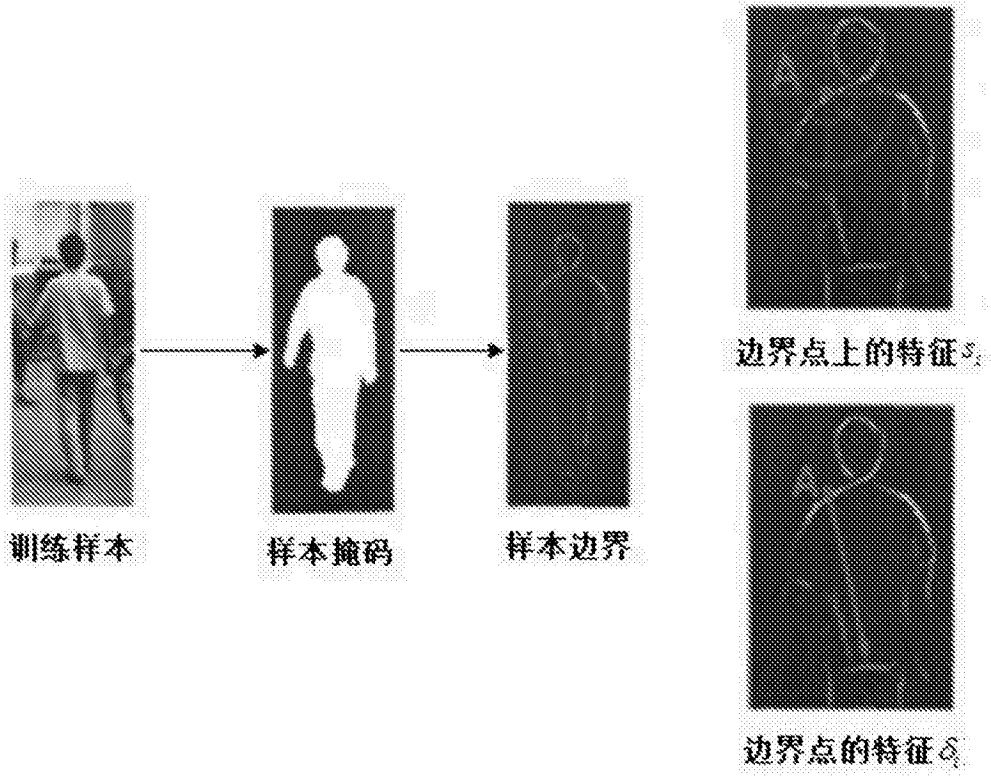


图2