

(19) 中华人民共和国国家知识产权局



(12) 发明专利申请

(10) 申请公布号 CN 103646257 A

(43) 申请公布日 2014.03.19

(21) 申请号 201310745952.1

G07C 9/00(2006.01)

(22) 申请日 2013.12.30

H04N 7/18(2006.01)

(71) 申请人 中国科学院自动化研究所

地址 100190 北京市海淀区中关村东路 95 号

申请人 东莞中国科学院云计算产业技术创新与育成中心

(72) 发明人 王飞跃 胡斌 熊刚 朱凤华
王生进

(74) 专利代理机构 中科专利商标代理有限责任公司 11021

代理人 宋焰琴

(51) Int. Cl.

G06K 9/66(2006.01)

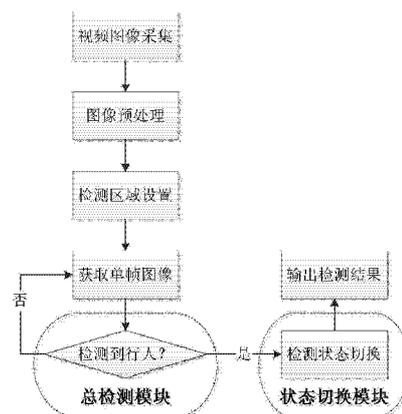
权利要求书2页 说明书4页 附图3页

(54) 发明名称

一种基于视频监控图像的行人检测和计数方法

(57) 摘要

本发明公开了一种行人检测和计数方法,该方法包括以下步骤:采集视频序列;获取视频序列中的每一帧图像;对于获得的每一帧视频图像进行预处理;对于经过预处理的视频图像设置检测区域;对于单帧视频图像中的检测区域进行行人个体检测,得到检测区域中所有可能存在的行人目标;基于多种检测状态的切换,对于行人目标进行准确的检测、跟踪和计数统计;输出行人检测结果,在每一帧图像中用明显的标示标记出每一个检测到的行人目标,并对每个行人目标设置唯一的编号。本发明可以提高行人检测的精度并能进行准确的计数,而且可以同时应用于静止或动态变化的背景中。



1. 一种行人检测和计数方法,其特征在于,该方法包括以下步骤:
 - 步骤 1,采集视频序列;
 - 步骤 2,获取视频序列中的每一帧图像;
 - 步骤 3,对于获得的每一帧视频图像进行预处理;
 - 步骤 4,对于经过预处理的视频图像设置检测区域;
 - 步骤 5,对于单帧视频图像中的检测区域进行行人个体检测,得到检测区域中所有可能存在的行人目标;
 - 步骤 6,基于多种检测状态的切换,对于行人目标进行准确的检测、跟踪和计数统计;
 - 步骤 7,输出行人检测结果,在每一帧图像中用明显的标示标记出每一个检测到的行人目标,并对每个行人目标设置唯一的编号。
2. 根据权利要求 1 所述的方法,其特征在于,所述预处理包括去噪和 / 或归一化。
3. 根据权利要求 1 所述的方法,其特征在于,所述步骤 3 进一步包括以下步骤:
 - 步骤 31,事先存储所有摄像机的参数;
 - 步骤 32,采集当前场景的环境参数;
 - 步骤 33,基于所述摄像机的参数和当前场景的环境参数,对于所述视频图像进行去噪和归一化处理。
4. 根据权利要求 1 所述的方法,其特征在于,所述检测区域为任意形状的多边形。
5. 根据权利要求 4 所述的方法,其特征在于,所述步骤 4 进一步包括以下步骤:
 - 步骤 41,对于所述视频图像,使用鼠标连续点击构成检测区域的多边形的各个顶点位置;
 - 步骤 42,自动连接这些点形成任意形状的多边形检测区域。
6. 根据权利要求 1 所述的方法,其特征在于,所述步骤 5 进一步包括以下步骤:
 - 步骤 51,对输入的单帧图像通过实时背景建模进行连通成分分析,并对其进行前景建模;
 - 步骤 52,通过离线采集的样本使用机器学习算法训练得到行人全身分类器和行人身体部位分类器;
 - 步骤 53,使用得到的分类器对于所述视频图像中的检测区域进行检测,得到所有可能存在的行人目标;
 - 步骤 54,结合前景检测的结果对人群密度进行估计;
 - 步骤 55,基于估计得到的人群密度和所有可能存在的行人目标得到最终行人检测结果并显示及存储。
7. 根据权利要求 1 所述的方法,其特征在于,所述多种检测状态包括检测态、存储态、前景提取态、统计态、删除态和显示态。
8. 根据权利要求 7 所述的方法,其特征在于,所述步骤 6 进一步包括以下步骤:
 - 步骤 61,对于所述步骤 5 检测到的当前帧中的每一个行人目标进行标定;
 - 步骤 62,跟踪处理每一个行人目标并与所述步骤 5 中后续帧检测到的结果进行交叉组合比较;
 - 步骤 63,经过检测态或前景提取态检测到的目标进入存储态,然后在统计态进行方向计算和数量标注,同时对不满足条件的目标通过删除态删除,满足条件的目标将进入显示

态进行显示；

步骤 64, 去除前景提取态和存储态中不满足条件的错误的行人检测目标, 保留正确的行人检测目标并对每个行人目标进行跟踪计数。

9. 根据权利要求 1 所述的方法, 其特征在于, 所述步骤 7 还包括标示出行人目标的不同运动方向的步骤。

10. 根据权利要求 1 所述的方法, 其特征在于, 所述步骤 7 还包括将所有行人目标的信息全部存入数据库的步骤。

一种基于视频监控图像的行人检测和计数方法

技术领域

[0001] 本发明属于智能监控和智能交通领域,具体涉及视频图像处理、机器学习、目标检测及跟踪技术,尤其涉及一种行人检测和计数方法。

背景技术

[0002] 如何在视频监控图像中准确有效地对行人进行检测和计数是一项很有意义的研究工作,在人机交互、机器人学、数字媒体内容分析、辅助驾驶、视频监控等很多领域都有着广泛的研究价值和应用前景。随着智能监控技术的发展,在视频监控领域的应用尤为重要和迫切。

[0003] 由于视频监控图像中的行人和人群的行为是一种非常复杂的随机行为,虽然在某些特定场合有一定的运动规律,但到目前为止还没有很好的可以对行人和人群进行智能化监控的方法。另外由于行人之间存在的遮挡和重叠现象、行人与周围环境的纹理融合问题都给图像处理和实际检测带来了相当大的难度。

发明内容

[0004] 为了解决上述存在的问题,本发明的主要目的在于提供一种基于多分类器检测和状态切换确认的行人检测和计数方法,该方法包含行人个体检测和整体确认计数两大部分。

[0005] 本发明提出的一种行人检测和计数方法包括以下步骤:

[0006] 步骤 1,采集视频序列;

[0007] 步骤 2,获取视频序列中的每一帧图像;

[0008] 步骤 3,对于获得的每一帧视频图像进行预处理;

[0009] 步骤 4,对于经过预处理的视频图像设置检测区域;

[0010] 步骤 5,对于单帧视频图像中的检测区域进行行人个体检测,得到检测区域中所有可能存在的行人目标;

[0011] 步骤 6,基于多种检测状态的切换,对于行人目标进行准确的检测、跟踪和计数统计;

[0012] 步骤 7,输出行人检测结果,在每一帧图像中用明显的标示标记出每一个检测到的行人目标,并对每个行人目标设置唯一的编号。

[0013] 本发明的优点和积极效果在于:

[0014] (1) 采用图像归一化预处理和检测区域设置,可以有效增加算法的适用范围和检测速度及准确性。

[0015] (2) 采用机器学习算法得到的多个分类器,可以有效提高算法检测的准确度。

[0016] (3) 多种状态切换的确认方法,既可以确认正确目标,去除错误目标,还可以对每个目标进行跟踪计数,并能得到每个目标运动的方向、速度等特征。

附图说明

[0017] 图 1 为本发明行人检测和计数方法的流程图；

[0018] 图 2 为本发明中行人目标检测的流程图；

[0019] 图 3 为本发明中行人目标准确检测的流程图。

具体实施方式

[0020] 为使本发明的目的、技术方案和优点更加清楚明白，以下结合具体实施例，并参照附图，对本发明进一步详细说明。

[0021] 图 1 为本发明行人检测和计数方法的流程图，如图 1 所示，所述行人检测和计数方法包括以下步骤：

[0022] 步骤 1，采集视频序列；

[0023] 视频序列的来源就是视频监控设备中的监控摄像机，由于摄像机硬件技术的发展，目前该步骤基本上没有什么问题，只要摄像机采集的图像达到基本标清的程度即可。

[0024] 步骤 2，获取视频序列中的每一帧图像；

[0025] 对于基于视频的图像处理而言，早期的目标检测跟踪算法多以跟踪为主，需要大量用到图像中的运动信息，只在一定数量的帧数上进行单帧检测用于消除累积误差和重置某些参数，这可以称之为基于跟踪的检测。但由于跟踪算法本身的不稳定性以及检测算法的介入时机不好确定，尤其是运动摄像机带来的大量运动信息的不确定性因素，该类方法在应用中遇到了很多问题。因此本发明采用另一类基于检测的跟踪方法，即对视频图像中的每一帧(或间隔很少的帧)都进行检测，而将连续视频图像中的运动信息和其它关联信息作为辅助信息参与到检测中，这类方法的最大优点就是基本不受摄像机运动和背景剧烈变化造成的运动信息不稳定的影响。

[0026] 步骤 3，对于获得的每一帧视频图像进行预处理；

[0027] 该步骤包括但不限于结合事先得到的当前摄像机的参数，并通过去噪、归一化等处理手段消除不同场景中光照、环境色调、噪声等的影响。

[0028] 具体来说，所述步骤 3 进一步包括以下步骤：

[0029] 步骤 31，事先存储所有摄像机的参数；

[0030] 步骤 32，采集当前场景的光照、色调、噪声等环境参数；

[0031] 步骤 33，基于所述摄像机的参数和当前场景的环境参数，对于所述视频图像进行去噪和归一化处理。

[0032] 步骤 4，对于经过预处理的视频图像设置检测区域，所述检测区域可以是任意形状的多边形；

[0033] 在通常无限定的检测中，需要对整个图像进行扫描检测，但实际场景中需要检测的往往只是图像中的一部分区域，检测区域设置好后，后续的检测及处理将只在该区域内进行，因此通过事先设定检测区域可以减少检测的范围，从而提高算法的运行速度和准确性；另外对于不同的场景，摄像机的视野和监控范围都会有所不同，通过设定大小位置相对固定的检测区域可以有效消除不同摄像机之间的差异，最后通过设定检测区域和各种辅助参数(如出入口设定、目标基本标定等)还可以在很大程度上提高算法检测的精度。

[0034] 具体来说，所述步骤 4 进一步包括以下步骤：

[0035] 步骤 41, 对于所述视频图像, 使用鼠标连续点击构成检测区域的多边形的各个顶点位置;

[0036] 步骤 42, 自动连接这些点形成任意形状的多边形检测区域。

[0037] 步骤 5, 对于单帧视频图像中的检测区域进行行人个体检测, 得到检测区域中所有可能存在的行人目标;

[0038] 该步骤采用事先通过机器学习方法得到的多个行人检测分类器进行行人个体检测, 由于本发明采用的基于检测的跟踪方法是以检测为主, 因此该步骤中使用的行人检测分类器应该具有尽可能高的检测精度和尽可能快的检测速度, 这可以通过很多训练方法进行改进, 比如采用多种分类器协同检测就可以有效提高检测的准确性。由于各处理部分的独立性, 该步骤中得到的分类器还可以根据实际场景的应用效果进行不断的更新。

[0039] 图 2 为本发明中行人个体检测的流程图, 如图 2 所示, 所述步骤 5 进一步包括以下步骤:

[0040] 步骤 51, 对输入的单帧图像通过实时背景建模进行连通成分分析, 并进行前景建模(前景检测子模块);

[0041] 步骤 52, 通过离线采集的样本使用机器学习算法训练得到行人全身分类器和行人身体部位分类器, 该步骤采用多种分类器是为了得到更准确的检测结果;

[0042] 步骤 53, 使用得到的分类器对于所述视频图像中的检测区域进行检测, 得到所有可能存在的行人目标(行人检测子模块);

[0043] 步骤 54, 结合前景检测的结果对人群密度进行估计(人群密度检测模块);

[0044] 步骤 55, 基于估计得到的人群密度和所有可能存在的行人目标通过场景状态评估模块得到最终行人检测结果(场景状态评估模块)并显示及存储。该步骤中, 如果估计得到的人群密度(如 80%)与所有可能存在的行人目标占总人数的比例(如 40 人左右, 满员为 50 人)基本相符, 则可以确定最终行人检测结果并显示及存储, 如果这两个结果相差过大, 则需要通过后面几帧的检测结果确认后再进行显示及存储。

[0045] 步骤 6, 基于多种检测状态的切换, 对于行人目标进行准确的检测、跟踪和计数统计;

[0046] 该步骤主要是根据所述步骤 5 单帧检测的结果, 结合连续视频帧之间的相关信息记录空间和时间的各个不同的关键状态, 通过不同关键状态的相互切换最大化地利用所能获得的所有环境信息来达到对目标的准确检测、跟踪和计数统计的目的。具体说来就是对于所述步骤 5 每一帧图像中检测到的所有可能的目标全部进行标定, 然后对每一个目标对象进行多帧跟踪处理, 并与所述步骤 5 中后续帧检测到的结果进行交叉组合比较, 每一个目标都要经过检测态、存储态、前景提取态、统计态、删除态、显示态等不同的状态确认, 从而去除错误的检测目标, 保留正确的检测目标并对其进行跟踪计数。

[0047] 图 3 为本发明中状态切换的流程图, 如图 3 所示, 所述步骤 6 进一步包括以下步骤:

[0048] 步骤 61, 对于所述步骤 5 检测到的当前帧中的每一个行人目标进行标定;

[0049] 步骤 62, 跟踪处理每一个行人目标并与所述步骤 5 中后续帧检测到的结果进行交叉组合比较;

[0050] 步骤 63, 经过检测态(分类器)或前景提取态检测到的目标进入存储态, 然后在统

计态进行方向计算和数量标注,同时对不满足条件的目标通过删除态删除,满足条件的目标将进入显示态进行显示;

[0051] 步骤 64,去除前景提取态和存储态中不满足条件的错误的行人检测目标,保留正确的行人检测目标并对每个行人目标进行跟踪计数。

[0052] 步骤 7,输出行人检测结果,在每一帧图像中用明显的标示标记出每一个检测到的行人目标,并对每个行人目标设置唯一的编号,根据需要还可以标示出行人目标的不同运动方向,另外将所有行人目标的信息全部存入数据库以用于后续可能的查询、统计、检索等。

[0053] 以上所述的具体实施例,对本发明的目的、技术方案和有益效果进行了进一步详细说明,所应理解的是,以上所述仅为本发明的具体实施例而已,并不用于限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

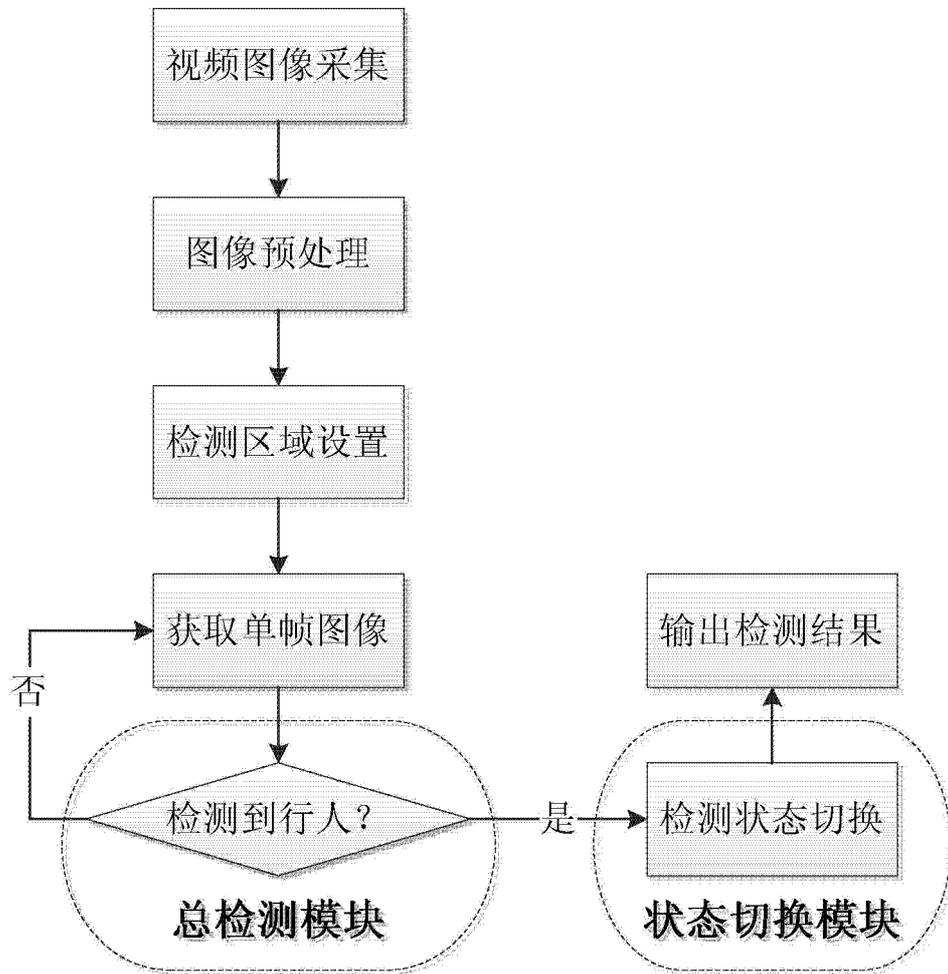


图 1

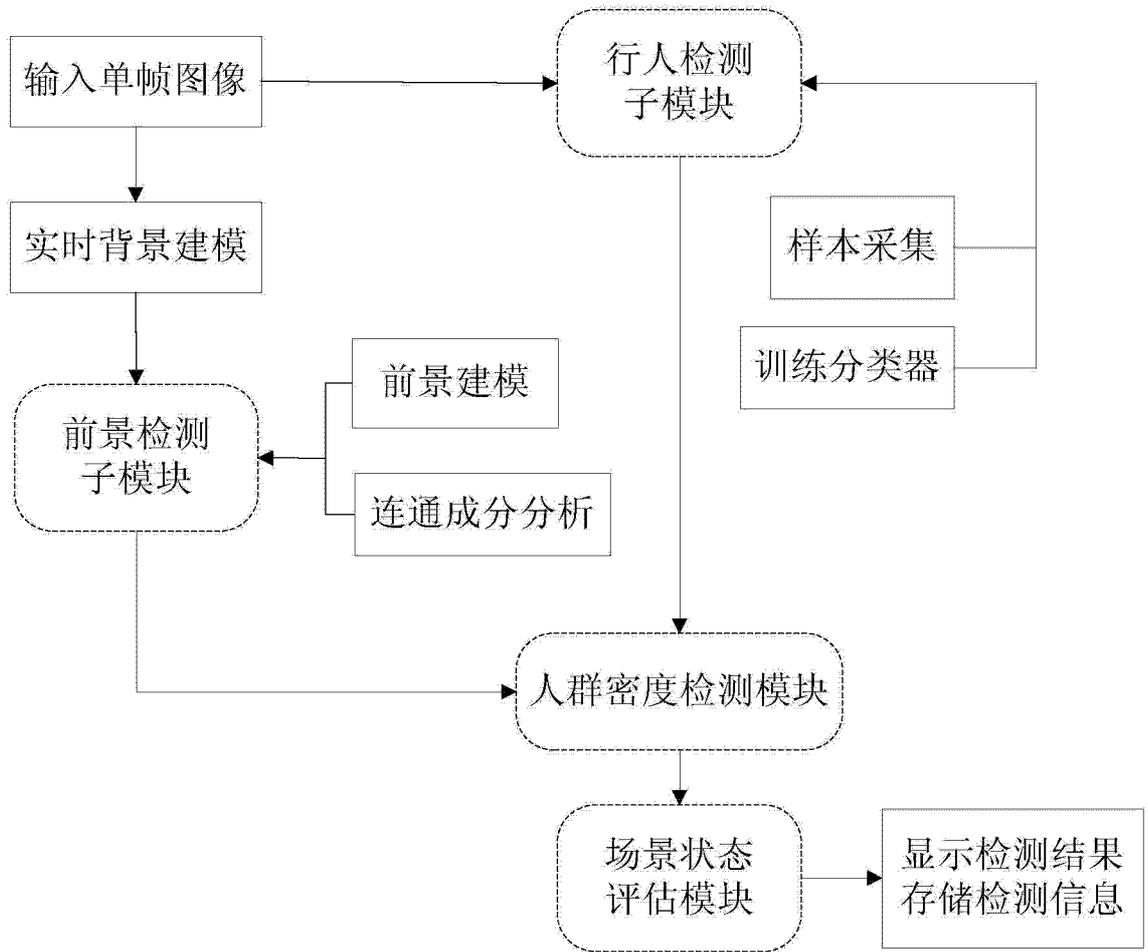


图 2

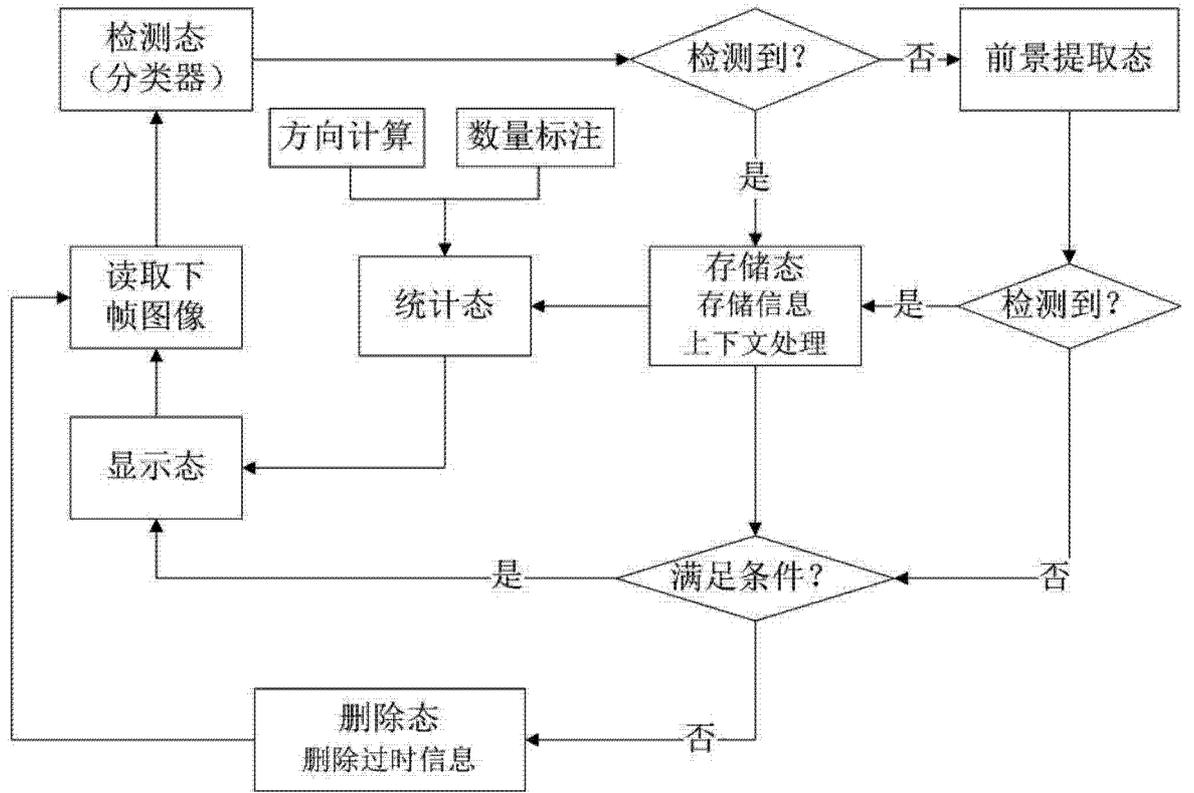


图 3