



(12) 发明专利

(10) 授权公告号 CN 108073890 B

(45) 授权公告日 2023.01.24

(21) 申请号 201711097624.X

(51) Int.Cl.

(22) 申请日 2017.11.09

G06V 40/20 (2022.01)

(65) 同一申请的已公布的文献号

审查员 周佳利

申请公布号 CN 108073890 A

(43) 申请公布日 2018.05.25

(30) 优先权数据

16198678.1 2016.11.14 EP

(73) 专利权人 安讯士有限公司

地址 瑞典,浪德

(72) 发明人 尼克拉·丹尼尔松 西蒙·莫林

(74) 专利代理机构 北京德琦知识产权代理有限公司 11018

专利代理人 康泉 宋志强

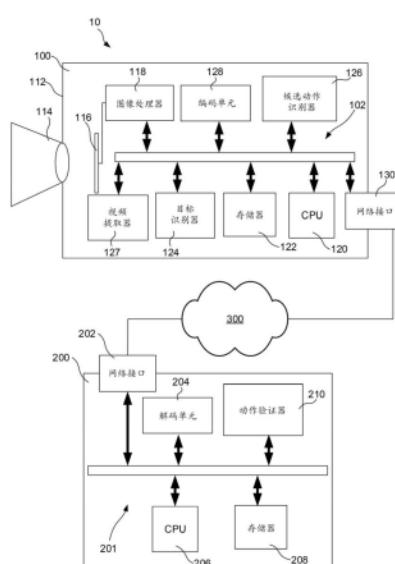
权利要求书2页 说明书8页 附图2页

(54) 发明名称

用于由相机所捕获的视频序列中的动作识别的方法和系统

(57) 摘要

本公开涉及视频序列中的动作识别。在视频序列中的动作识别的系统和方法中，该系统包括被配置为捕获视频序列的相机和被配置为执行动作识别的服务器。相机包括被配置为识别视频序列的目标图像帧中的兴趣目标的目标识别器；被配置为对目标图像帧应用第一动作识别算法并由此检测候选动作的存在的候选动作识别器；被配置为通过从视频序列提取与多个图像帧有关的视频数据来产生动作视频序列的动作图像帧的视频提取器，其中，提取出视频数据的多个图像帧的一个或多个图像帧包括兴趣目标；以及被配置为传输动作视频序列至服务器的网络接口。服务器包括被配置为对动作视频序列应用第二动作识别算法并由此验证或否决候选动作是动作的动作验证器。



1. 一种用于在由相机(100)所捕获的视频序列中的动作识别的方法,所述方法包括:
由所述相机(100)的电路(102):

识别所述视频序列的图像帧中的兴趣目标;

对所述图像帧应用第一动作识别算法,在所述第一动作识别算法中识别所述兴趣目标并由此检测候选动作的存在,其中,所述第一动作识别算法基于使用所述视频序列的单个图像帧中的情景和/或空间信息的静止图像情景和/或空间动作识别算法,所述单个图像帧包括所述兴趣目标;

基于检测到候选动作的存在,触发动作视频序列至服务器(200)的发送,以用于确定所述候选动作是否是真实动作,

通过从所述视频序列提取与多个图像帧有关的视频数据来产生所述动作视频序列的图像帧,其中,提取出所述视频数据的所述多个图像帧中的一个或多个图像帧包括所述兴趣目标;并且

将所述动作视频序列传输至被配置为执行动作识别的服务器(200);并且

由所述服务器的电路(201):

对所述动作视频序列应用第二动作识别算法并由此验证或否决所述候选动作是预定类型的动作,

其中,所述第二动作识别算法基于使用所述动作视频序列的多个图像帧的时域信息的时域动作识别算法。

2. 根据权利要求1所述的方法,其中,产生所述动作视频序列的所述图像帧的动作包括:裁剪所述视频序列的所述多个图像帧,以使得包括所述兴趣目标的所述图像帧包括所述兴趣目标的至少一部分。

3. 根据权利要求2所述的方法,其中,包括所述兴趣目标的所述动作视频序列的所述图像帧包括至少部分地围绕所述兴趣目标的背景的一部分。

4. 根据权利要求1所述的方法,其中,传输所述动作视频序列的动作包括将所述动作视频序列内的坐标传输至所述兴趣目标。

5. 根据权利要求1所述的方法,其中,所述方法进一步包括,由所述相机的所述电路:

检测所述视频序列中的兴趣目标,

其中,产生所述动作视频序列的所述图像帧的动作包括:提取与所述兴趣目标的检测之前的时间点相关的所述视频序列的第一预定数目的图像帧有关的视频数据。

6. 根据权利要求1所述的方法,其中,所述方法进一步包括,由所述相机的所述电路:

检测所述视频序列中的兴趣目标,

其中,产生所述动作视频序列的所述图像帧的动作包括:提取与所述兴趣目标的检测之后的时间点相关的所述视频序列的第二预定数目的图像帧有关的视频数据。

7. 根据权利要求1所述的方法,其中,所述相机和所述服务器是彼此以一距离定位并且被配置为经由数字网络(300)彼此通信的分立的物理实体。

8. 一种用于在视频序列中的动作识别的系统,所述系统包括:

相机(100),被配置为捕获所述视频序列,以及

服务器(200),被配置为执行动作识别;

所述相机包括:

目标识别器(124),被配置为识别所述视频序列的图像帧中的兴趣目标;

候选动作识别器(126),被配置为对所述图像帧应用第一动作识别算法,在所述第一动作识别算法中识别所述兴趣目标并由此检测候选动作的存在,其中,所述第一动作识别算法基于使用所述视频序列的单个图像帧中的情景和/或空间信息的静止图像情景和/或空间动作识别算法,所述单个图像帧包括所述兴趣目标;

视频提取器(127),被配置为通过从所述视频序列提取与多个图像帧有关的视频数据来产生动作视频序列的图像帧,其中,提取出所述视频数据的所述多个图像帧中的一个或多个图像帧包括所述兴趣目标;以及

网络接口(130),被配置为将所述动作视频序列传输至所述服务器;

所述服务器包括:

动作验证器(210),被配置为对所述动作视频序列应用第二动作识别算法并由此验证或否决所述候选动作是预定类型的动作,

其中,所述第二动作识别算法基于使用所述动作视频序列的多个图像帧的时域信息的时域动作识别算法。

9.根据权利要求8所述的系统,其中,所述视频提取器(127)被进一步配置为裁剪所述视频序列的所述多个图像帧,以使得包括所述兴趣目标的所述视频序列的所述图像帧包括所述兴趣目标的至少一部分。

10.根据权利要求8所述的系统,其中,所述视频提取器(127)被进一步配置为裁剪所述视频序列的所述多个图像帧,以使得包括所述兴趣目标的所述视频序列的所述图像帧包括至少部分地围绕所述兴趣目标的背景的一部分。

11.根据权利要求8所述的系统,其中,所述目标识别器(124)被进一步配置为检测所述视频序列中的兴趣目标,其中,所述视频提取器(127)被进一步配置为提取与所述兴趣目标的检测之前的时间点相关的所述视频序列的第一预定数目的图像帧有关的视频数据。

12.根据权利要求8所述的系统,其中,所述目标识别器(124)被进一步配置为检测所述视频序列中兴趣目标,其中,所述视频提取器(127)被进一步配置为提取与所述兴趣目标的检测之后的时间点相关的所述视频序列的第二预定数目的图像帧有关的视频数据。

用于由相机所捕获的视频序列中的动作识别的方法和系统

技术领域

[0001] 本发明涉及视频序列中的动作识别。

背景技术

[0002] 视频序列中的动作识别是检测视频序列中的特定预定义动作的任务。待检测的动作的示例可以例如是人员打斗、奔跑、进食、玩运动游戏。动作识别通常要求过多的处理能力,这使其不适用于在捕获视频序列的数字摄像机上执行。替代的通常由具有所需处理能力的服务器执行动作识别。然而,在数字网络之上发送大量视频需要大量带宽。

[0003] 因此,需要改进的视频序列中的动作识别。

发明内容

[0004] 考虑到以上,本发明的目标在于提供改进的视频序列中的动作识别。

[0005] 根据第一特征方面提供了一种用于在由相机所捕获的视频序列中的动作识别的方法。方法包括:由相机的电路:识别在视频序列的目标图像帧中的兴趣目标;对目标图像帧应用第一动作识别算法并由此检测候选动作的存在;通过从视频序列提取与多个图像帧有关的视频数据来产生动作视频序列的动作图像帧,其中,提取出视频数据的多个图像帧的一个或多个图像帧包括兴趣目标;并且将动作视频序列传输至被配置为执行动作识别的服务器;由服务器的电路:对动作视频序列应用第二动作识别算法并由此验证或否决候选动作是预定类型的动作。

[0006] 本发明的动作识别的方案是有利的,这是因为其允许对动作识别所需的处理能力分布在两个不同装置之上而没有在相机与服务器之间的通信中占据太多带宽的问题。另外,本发明的动作识别的方案使其能够执行处理器增强的动作识别而并不加载太多相机的处理器,并且同时不必不断地将整个视频序列流发送至服务器。因此,本发明的方案允许使用要求在相机本地进行的第一类型动作识别算法的较少的处理以及要求在服务器上进行的第二类型动作识别算法的较多的处理。通过提取并仅发送动作视频序列而不是连续的视频流,可以节省带宽。根据本发明的动作识别的方案可以在相机处检测候选动作。候选动作触发动作视频序列传输至服务器,在该服务器中执行更高级的动作识别分析。因此相机上的动作识别的目的是用于触发候选动作的检测。候选动作的检测触发将动作视频序列传输至服务器以用于进一步分析、验证、或否决候选动作作为真实动作。

[0007] 第一动作识别算法可以主要基于使用目标图像帧中的情景和/或空间信息的情景和/或空间动作识别算法。情景和/或空间动作识别算法通常不要求太多处理能力。因此,使用该类型算法的动作识别更易于在相机上本地地可。

[0008] 第二动作识别算法可以主要基于使用动作视频序列的多个图像帧的时域信息的时域动作识别算法。时域动作识别算法通常在识别动作方面更准确。因此,可以执行更准确的动作识别。

[0009] 产生动作图像帧的动作可以包括裁剪视频序列的多个图像帧,以使得包括兴趣目

标的动作图像帧包括该兴趣目标的至少一部分。这可以节省相机与服务器之间的带宽。

[0010] 产生动作图像帧的动作可以包括裁剪视频序列的多个图像帧,以使得包括兴趣目标的动作图像帧包括至少部分地围绕兴趣目标的背景的一部分。通过合并至少部分地围绕兴趣目标的背景,不仅可以执行动作视频序列的时域分析而且还可以执行动作视频序列的情景和/或空间分析。

[0011] 传输动作视频序列的动作可以包括将动作视频序列内的坐标传输至兴趣目标。坐标可以涉及哪个或哪些图像帧包括兴趣目标,和/或兴趣目标位于相应的图像帧中的何处。

[0012] 方法可以进一步包括,由相机的电路来检测视频序列中的兴趣目标。产生动作图像帧的动作可以包括提取与兴趣目标的检测之前的时间点相关的视频序列的第一预定数目的图像帧有关的视频数据。产生动作图像帧的动作可以包括提取与兴趣目标检测之后的时间点相关的视频序列的第二预定数目的图像帧有关的视频数据。这允许提高要在服务器处识别的动作发送至服务器的正确时域窗口的概率。

[0013] 相机和服务器可以是彼此相距一距离定位的分立的物理实体。相机和服务器可以被配置为经由数字网络来彼此通信。

[0014] 根据第二特征方面提供了一种用于在视频序列中的动作识别的系统。系统包括:被配置为捕获视频序列的相机,以及被配置为执行动作识别的服务器。该相机包括:被配置为在视频序列的目标图像帧中识别兴趣目标的目标识别器;被配置为对目标图像帧应用第一动作识别算法并由此检测候选动作的存在候选动作识别器;被配置为通过从视频序列提取与多个图像帧有关的视频数据来产生动作视频序列的动作图像帧的视频提取器,其中,提取出视频数据的多个图像帧的一个或多个包括兴趣目标;以及,被配置为将动作视频序列传输至服务器的网络接口。服务器包括被配置为对动作视频序列应用第二动作识别算法并由此验证或否决候选动作是预定类型动作的动作验证器。

[0015] 视频提取器可以进一步被配置为裁剪视频序列的多个图像帧,以使得包括兴趣目标的动作图像帧包括兴趣目标的至少一部分。

[0016] 视频提取器可以进一步被配置为裁剪视频序列的多个图像帧,以使得包括兴趣目标的动作图像帧包括至少部分地围绕兴趣目标的背景的一部分。

[0017] 目标识别器可进一步被配置为检测视频序列中的兴趣目标。视频提取器可进一步被配置为提取与兴趣目标的检测之前的时间点相关的视频序列的第一预定数目的图像帧有关的视频数据。视频提取器可进一步被配置为提取与兴趣目标的检测之后的时间点相关的视频序列的第二预定数目的图像帧有关的视频数据。

[0018] 当可应用时,上述方法的特征也适用于该第二特征方面。为了避免不必要的重复,参考以上内容。

[0019] 本发明的可应用性的进一步范围将从以下给出的详细说明书变得明显。然而,应该理解的是,当指示了本发明的优选实施例时,详细说明书和具体示例仅借由示意说明的方式给出,因为在本发明范围内的各种改变和修改将从该详细说明书而对于本领域技术人员变得明显。

[0020] 因此,应该理解的是,本发明不限于所述装置的特定部件部分或者所述方法的步骤,因为这些装置和方法可以改变。也应该理解的是,本文所使用的术语仅是为了描述特定实施例的目的,并且并非有意是限定性的。必须注意的是,如在说明书和所附权利要求中所

使用,冠词“一”、“一些”、“该”和“所述”意在意味着存在一个或多个元件,除非情景明确给出相反指示。因此,例如,对于“一单元”或“该单元”的参考可以包括数个装置等等。进一步,词语“包括”、“包含”、“含有”和类似词语并未排除其他元件或步骤。

附图说明

[0021] 现在将参照示出了本发明实施例的附图更详细描述本发明的以上和其他特征方面。附图不应视作将本发明限定于具体实施例;替代的它们用于解释并理解本发明。

[0022] 如图中所示,为了示意性说明的目的夸大了层和区域的大小,并且因此提供用于说明本发明的实施例的整体结构。相同的参考标号始终指相同的元件。

[0023] 图1说明了用于在视频序列中的动作识别的系统。

[0024] 图2是用于在视频序列中的动作识别的方法的方框图。

具体实施方式

[0025] 现在下文中将参照示出本发明当前优选实施例的附图更全面地描述本发明。然而,本发明可以以许多不同形式实施并且不应被构造为限定于本文所阐述的实施例;相反,提供这些实施例是为了全面性和完整性,并且向本领域技术人员完全传达本发明的范围。

[0026] 图1图示了用于视频序列中的动作识别的系统。系统包括数字网络相机100和服务器200。数字网络相机100经由数字网络300连接至服务器200。数字网络相机100和服务器200是彼此以一距离定位的分立的物理实体并被配置为经由数字网络300彼此通信。

[0027] 数字网络300涉及允许装置例如数字网络相机100和与其连接的服务器200交换数字数据的网络。使用电缆或者无线地建立连接至数字网络300的装置之间的连接。数字网络的非限定性示例是互联网、内联网、局域网和蜂窝网络。数字网络的一部分可以是私人数字网络。数字网络的一部分可以是公共数字网络。私人数字网络可以由网络访问限制装置(未示出)连接至公共数字网络。网络访问限制装置可以是用于保护私人数字网络而安装的防火墙。网络访问限制装置可以是执行网络地址转换(NAT)的装置。

[0028] 数字网络相机100被布置为捕获描绘场景的视频序列。数字网络相机100包括外壳112,透镜114和电路102。数字网络相机100被布置为捕获并处理(并且还可能存储)视频序列。电路102包括图像传感器116、图像处理单元118、目标识别器124、候选动作识别器126、视频提取器127、和网络接口130。电路102可以进一步包括中央处理单元CPU 120、数字数据存储介质(存储器)122、和编码单元128中的一个或多个。图像处理单元118、目标识别器124、候选动作识别器126、视频提取器127和/或编码单元128中的任意一个可以实施作为专用硬件电路和/或软件模块。在作为所实施的软件的情形中,软件可以在CPU 120上运行。CPU 120可以是用于执行数字数据处理的任意合适的CPU。还应该注意的是任意专用硬件电路可以部分地包括在专用处理器上或在CPU 120上运行的软件部分。

[0029] 存储器122可以是任意种类的易失性或非易失性存储器。另外,存储器122可以包括多个存储器单元。多个存储器单元中的至少一个存储器单元可以用作缓冲存储器,以用于在处理例如视频序列的内容的同时缓冲数据。

[0030] 数字网络相机100被布置为经由网络接口130连接至数字网络300。至数字网络的连接可以是有线或无线的。因此,网络接口130可以是适配于10/100/1000Mbps数据流量的

网络端口,诸如被布置为接收模块化连接器(例如RJ45连接器)的模块化端口的以太网端口。通常,该RJ45连接器被布置为接收诸如双绞线配对(例如cat 5,cat 5e或cat6)之类的网络电缆。备选地,网络端口的I/O机构可以是使用移动互联网通信标准(例如1G,2G,2.5G,2.75G,3G,3.5G,3.75G,3.96G,4G,5G)或使用WiFi的无线I/O机构。

[0031] 相机部件(即透镜114和图像传感器116)可被布置为捕获原始图像,其中每个原始图像可被描述为不同波长的光并且源自不同目标和目标部分。随后将这些原始图像从模拟信号转换为数字格式并传输至图像处理单元118中。根据此实施例,数字网络相机100是被布置为捕获摄影图像的相机。备选地或组合的,数字网络相机100的图像传感器116可被布置为捕获热图像。另外备选地或组合的,数字网络相机100的图像传感器116可被布置为捕获雷达图像。因此,由数字网络相机100所捕获的视频序列可以是摄影图像的表示、热图像的表示、雷达图像的表示、或者其组合。

[0032] 目标识别器124被配置为在由相机100所捕获的视频序列中检测兴趣目标。兴趣目标可以例如是人员、脸部、车辆、传输带上的产品、动物、地形特征、武器等。目标识别器124可进一步被配置为对检测到的兴趣目标进行分类。兴趣目标可以例如分类为属于特殊类型的目标。特殊类型目标的示例是:人员,脸部,车辆,特殊类型的产品。目标识别器124可进一步被配置为识别在视频序列中首次检测到兴趣目标的时间点。与此结合,存储器122可以进一步用作被配置为存储预定数目图像帧的图像帧缓冲器。因此,可以在用作图像帧缓冲器的存储器122中存储表示在视频序列中的首次检测到兴趣目标的时间点之前的图像帧的图像帧。

[0033] 进一步配置目标识别器124以在视频序列的一个或多个图像帧中识别兴趣目标,在本文中,在其中识别到兴趣目标的图像帧将被称为目标图像帧。

[0034] 视频序列中的动作识别是在视频序列中检测一个或多个预定类型动作的任务。预定类型的动作的示例是人员打斗、奔跑、进食、玩特定游戏等。预定类型动作的其他示例是酒驾检测、跳跃检测、愤怒检测、微笑检测、手势信号检测、跌落检测、空中巡逻检测、步法检测、危险行为检测、可疑行为检测(例如反常或非常规行为的检测)。

[0035] 可以通过静止图像情景和/或空间分析或时域分析(或两者的组合)来执行动作识别。对静止图像(例如视频序列的单个图像帧)执行情景和/或空间动作识别算法。对视频序列的多个图像帧执行时域动作识别算法。

[0036] 情景动作识别算法的示例由Georgia Gkioxari, Ross Girshick和Jitendra Malik在“Contextual Action Recognition with R*CNN”;arXiv:1505.01197中所公开。例如在CN102855462中、在CN103106394中以及由Karen Simonyan和Andrew Zisserman在“Two-Stream Convolutional Networks for Action Recognition in Videos”;arXiv:1406.2199中描述了动作识别算法的其他示例、情景和/或空间动作识别算法以及时域动作识别算法的两者。

[0037] 因此,动作识别具有两个主要步骤;静止图像情景和/或空间分析以及时域分析。尽管大多数有前途的步骤使用时域分析作为主要动作识别算法,但静止图像方案在一些情形中非常好用。然而,时域动作识别是困难的并且包括使用例如递归神经网络的处理强度很大的算法。这使得时域动作识别算法不适于在诸如数字网络相机100之类的嵌入式装置上执行。

[0038] 本发明涉及使用处理强度通常较为合理的情景和/或空间动作识别算法与处理增强的时域动作识别算法之间的协同合作。在数字网络相机100上运行第一动作识别算法以找到候选动作,以及在服务器200上运行第二动作识别算法以验证或否决候选动作是真实动作。第二动作识别算法比第一动作识别算法要求更强的处理性能。第一动作识别算法主要基于情景和/或空间动作识别。作为非限定性示例,第一动作识别算法可以单独地基于情景和/或空间动作识别算法。第二动作识别算法主要基于时域动作识别。然而,第二动作识别算法可以包括情景和/或空间动作识别的要素。

[0039] 候选动作识别器126被配置为对由目标识别器124识别到的目标图像帧中的至少一个一个图像帧应用第一动作识别算法。通过应用第一动作识别算法,检测到候选动作的存在。由候选动作识别器126执行的第一动作识别算法分析无需检测动作的类型。其仅需要检测一般的候选动作。然而,第一动作识别算法可被配置为对不同类型的动作进行过滤。因此,通过应用第一动作识别算法,可以检测到预定动作类型的候选动作的存在。

[0040] 通过应用第一动作识别算法,候选动作识别器126可被配置为检测可能指示动作的非常规姿势。因此,为了针对动作找到候选动作。候选动作识别器126可被进一步配置为过滤检测到的姿势以便于找到候选动作的预定类型的动作。第一动作识别算法是相对轻量的算法。可以接受一定程度的假阳性。候选动作的检测推动使用第二动作识别算法的更高级动作识别分析。因此,候选动作识别器126被配置为触发可能的动作建议,或者本文所涉及的候选动作。

[0041] 检测候选动作触发将作为包括候选动作的视频序列的一部分的动作视频序列发送至服务器200以用于通过对动作视频序列应用第二动作识别算法来确定候选动作是否是真实动作。

[0042] 视频提取器127被配置为产生动作视频序列的动作图像帧。动作视频序列是包括候选动作的视频序列的一部分。动作视频序列可以具有与视频序列相同的帧率。动作视频序列可以具有比视频序列的帧率较小的帧率,即,动作视频序列具有比视频序列的每秒帧数目fps较低的帧率。例如,视频序列的帧率可以是60fps而动作视频序列的帧率可以是30fps。

[0043] 通过从视频序列提取与多个图像帧有关的视频数据来产生动作视频序列的图像帧,本文称作动作图像帧。视频提取器127被配置为使得提取出视频数据的多个图像帧中的一个或多个图像帧包括兴趣目标。因此,多个动作图像帧中的至少一个或多个是目标图像帧。

[0044] 视频提取器127可被进一步配置为提取与在视频序列中检测到兴趣目标的所识别的时间点之前的时间点相关的视频序列的第一预定数目的图像帧有关的视频数据。因此,视频序列的图像帧可以缓存在存储器122中以稍后用于动作视频序列。这允许包括包含触发第一动作识别算法的运行的兴趣目标的目标图像帧、以及在与动作视频序列中将要包括的目标图像帧之前的图像帧的两者。这提高了动作视频序列的时域窗口包括与由兴趣目标所执行动作有关的所有相关信息的概率。作为非限定性示例,可以取决于以下项中的一个或多个来设置第一预定数目的图像帧:兴趣目标的类型、和候选动作的动作类型。此外,作为另一非限定性示例,目标以及可能其周围环境的运动分析可以用于揭示候选动作开始于哪个较早的图像帧。通过这种方式使其能够从相对较大的预缓冲仅提取可能与第二动作识别

别实际相关的图像帧。因此,可以动态地设置第一预定数目的图像帧。

[0045] 视频提取器127可被进一步配置为提取与在视频序列中检测到兴趣目标的所识别的时间点之后的时间点相关的视频序列的第二预定数目的图像帧有关的视频数据。这允许目标图像帧包括包含触发第一动作识别算法的运行的兴趣目标的目标图像帧、以及在动作视频序列中将要包括目标图像帧之后的图像帧的两者。这提高了动作视频序列的时域窗口包括与由兴趣目标所执行的动作有关的所有相关信息的概率。可以取决于以下项中的一个或多个设置第二预定数目的图像帧:兴趣目标的类型、和候选动作的动作类型。此外,作为另一非限定性示例,目标以及可能其周围环境的运动分析可以用于揭示候选动作结束于哪个图像帧。通过这种方式使其能够仅提取可能与第二动作识别实际相关的图像帧。因此,可以动态地设置第二预定数目的图像帧。

[0046] 视频提取器127可被进一步配置为提取与动作视频序列内坐标有关的信息至兴趣目标。坐标可以涉及哪个或哪些图像帧包括兴趣目标和/或兴趣目标在相应图像帧中所在的位置。这些坐标可以与动作视频序列一起传输至服务器。

[0047] 视频提取器127可被进一步配置为当产生动作图像帧时裁剪多个图像帧的视频数据。可以裁剪多个图像帧的视频数据,以使得包括兴趣目标的动作图像帧包括兴趣目标的至少一部分。此外,可以裁剪多个图像帧的视频数据,以使得包括兴趣目标的动作图像帧包括至少部分地围绕兴趣目标的背景的一部分。通过合并至少部分地围绕兴趣目标的背景,不仅可以执行动作视频序列的时域分析,并且还可以执行动作视频序列的情景和/或空间分析。

[0048] 编码单元128被布置为使用视频编码对视频序列的数字视频数据编码。视频编码的非限定性示例是ISO/MPEG或ITU-H.26X系列的视频编码标准。编码单元128被布置为编码数字视频数据的图像,以下称作已编码数字视频数据。已编码数字视频数据可以经由网络接口130直接通过数字网络300发送。备选地,已编码数字视频数据可以存储在存储器122中以用于经由网络接口130稍后通过数字网络300传输。编码单元128可被配置为在将动作视频序列传输至服务器200之前对动作视频序列进行编码。

[0049] 网络接口130被配置为将动作视频序列传输至服务器200。

[0050] 服务器200包括电路201,该电路201包括动作验证器210。电路201可以进一步包括网络接口202、解码单元204、中央处理单元CPU 206、以及数字数据存储介质(存储器)208中的一个或多个。解码单元204和/或动作验证器210的任意一个可以实施作为专用硬件电路和/或软件模块。在实施为软件的情形中,软件可以运行在CPU 206上。CPU 206可以是用于执行数字数据处理的任意合适的CPU。也应该注意的是任意专用硬件电路可以部分地包括运行在专用处理器上或在CPU 206上的软件部分。

[0051] 服务器200被布置为经由网络接口202连接至数字网络300。数字网络的连接可以是有线或无线的。因此,网络接口202可以是适用于10/100/1000Mbps数据流量的网络端口、诸如被布置为接收模块化连接器(例如RJ45连接器)的模块化端口以太网端口。通常,该RJ45连接器被布置为接收网络电缆,诸如双绞线电缆(例如cat 5,cat5e或cat6)。备选地,网络端口的I/O机构可以是使用移动互联网通信标准(例如1G、2G、2.5G、2.75G、3G、3.5G、3.75G、3.9G、4G、5G)或使用WiFi的无线I/O机构。

[0052] 在对动作视频序列编码的情形中,配置解码单元204以解码已编码的视频序列。因

此,解码单元204被布置为使用视频解码对视频序列的数字视频数据进行解码。

[0053] 存储器122可以是任意种类的易失性或非易失性存储器。另外,存储器122可以包括多个存储器单元。多个存储器单元中的至少一个存储单元可以用作缓冲存储器以在处理例如动作视频序列的同时缓冲数据。存储器122可以进一步存储动作视频序列的整个或一部分。

[0054] 动作验证器210被配置为对动作视频序列应用第二动作识别算法。由此可以验证或否决候选动作是真实动作。特别是可以验证或否决候选动作是预定动作类型的动作。由动作验证器210执行的动作识别并非必须要实时执行。这是由于动作是短时间实际而并不持续进行,并且重要的是判定是否应该发起对特定动作类型的警报。

[0055] 参照图2示出了用于在由相机100所捕获的视频序列中的动作识别的方法。方法包括由相机(100)的电路102:

[0056] 识别S501视频序列的目标图像帧中的兴趣目标;

[0057] 对目标图像帧应用S502第一动作识别算法并由此检测候选动作的存在;

[0058] 通过从视频序列提取与多个图像帧有关的视频数据来产生S504动作视频序列的动作图像帧,其中提取出视频数据的多个图像帧中的一个或多个图像帧包括兴趣目标;

[0059] 将动作视频序列传输S506至服务器200。

[0060] 方法进一步包括,由服务器200的电路201,对动作视频序列应用S508第二动作识别算法并由此验证或否决候选动作是预定类型的动作。

[0061] 产生S504动作图像帧的动作可以包括:裁剪视频序列的多个图像帧,以使得包括兴趣目标的动作图像帧包括兴趣目标的至少一部分。

[0062] 产生S504动作图像帧的动作可以包括:裁剪视频序列的多个图像帧,以使得包括兴趣目标的动作图像帧包括至少部分地围绕兴趣目标的背景的一部分。

[0063] 传输S506动作视频序列的动作可以包括:将动作视频序列内坐标传输至兴趣目标。坐标可以涉及哪个或哪些图像帧包括兴趣目标,和/或兴趣目标位于相应图像帧的何处。

[0064] 方法可以进一步包括由相机100的电路102检测S500视频序列中的兴趣目标。产生S504动作图像帧的动作可以包括提取与兴趣目标的检测之前的时间点相关的视频序列的第一预定数目的图像帧有关的视频数据。产生S504动作图像帧的动作可以包括提取与兴趣目标的检测之后的时间点相关的视频序列的第二预定数目的图像帧有关的视频数据。

[0065] 本领域技术人员认识到本发明绝非限定于上述优选实施例。相反,在所附权利要求的范围内做出许多修改和改变是可能的。

[0066] 例如,在验证候选动作是真实动作之后,可以配置服务器200以发出警报触发信号。警报触发信号可以被发送至数字网络相机100,以在其中进一步处理。例如,可以配置数字网络相机100,以在接收到警报触发信号之后改变相机设置。可以改变的相机的设置的非限定性示例是:帧率、分辨率、光敏度、切换至HDR、采用标准消息触发连接至相机的扬声器、开始兴趣目标的PTZ追踪、触发用于被检测目标的高级外观模型的产生以使得其可以在通过多个相机进行追踪、开始雷达追踪、切换至热模式、改变阈值用于进一步动作识别、或者检查相关类型的下一步动作(例如,如果人员跌落,则开始搜索“再次起身”动作,并且如果他在特定时间阈值内并未站起则触发警报)。

[0067] 备选地或组合的,可以将警报触发信号发送至视频管理中心,VMS。警报触发器可以用在VMS处以用于发送出已经发生了预定类型动作的通知。

[0068] 另外,服务器200可以实施在各种类型的装置中。实施作为服务器200的装置的非限定性示例是专用计算机、另一相机装置、视频管理系统、云服务器、相机附近的分析箱、访问控制单元、具有计算性能的输入输出测试(IoT)装置。另外,服务器功能可以分布在不同的装置上。特别是在至少部分地实施作为软件代码部分的动作验证器210在多个处理器上运行的情形中。

[0069] 此外,可以在视频序列的多个目标图像帧中识别兴趣目标。随后可以对多个目标图像帧中的每个目标图像帧单独地应用第一动作识别算法。第一动作识别算法的单独应用的结果随后可以用于找到候选动作。例如,多个图像的单独分析可以揭示人员的腿总是处于不同角度。这可以指示人员可能在行走或甚至奔跑。此后通过从视频序列提取与多个图像帧有关的视频数据而产生动作视频序列的动作图像帧,其中由此提取视频数据的多个图像帧中的一个或多个图像帧包括兴趣目标。

[0070] 额外地,通过学习附图、本公开和所附权利要求,实践所请求保护的本发明的本领域技术人员可以理解并实现所公开的实施例的改变。

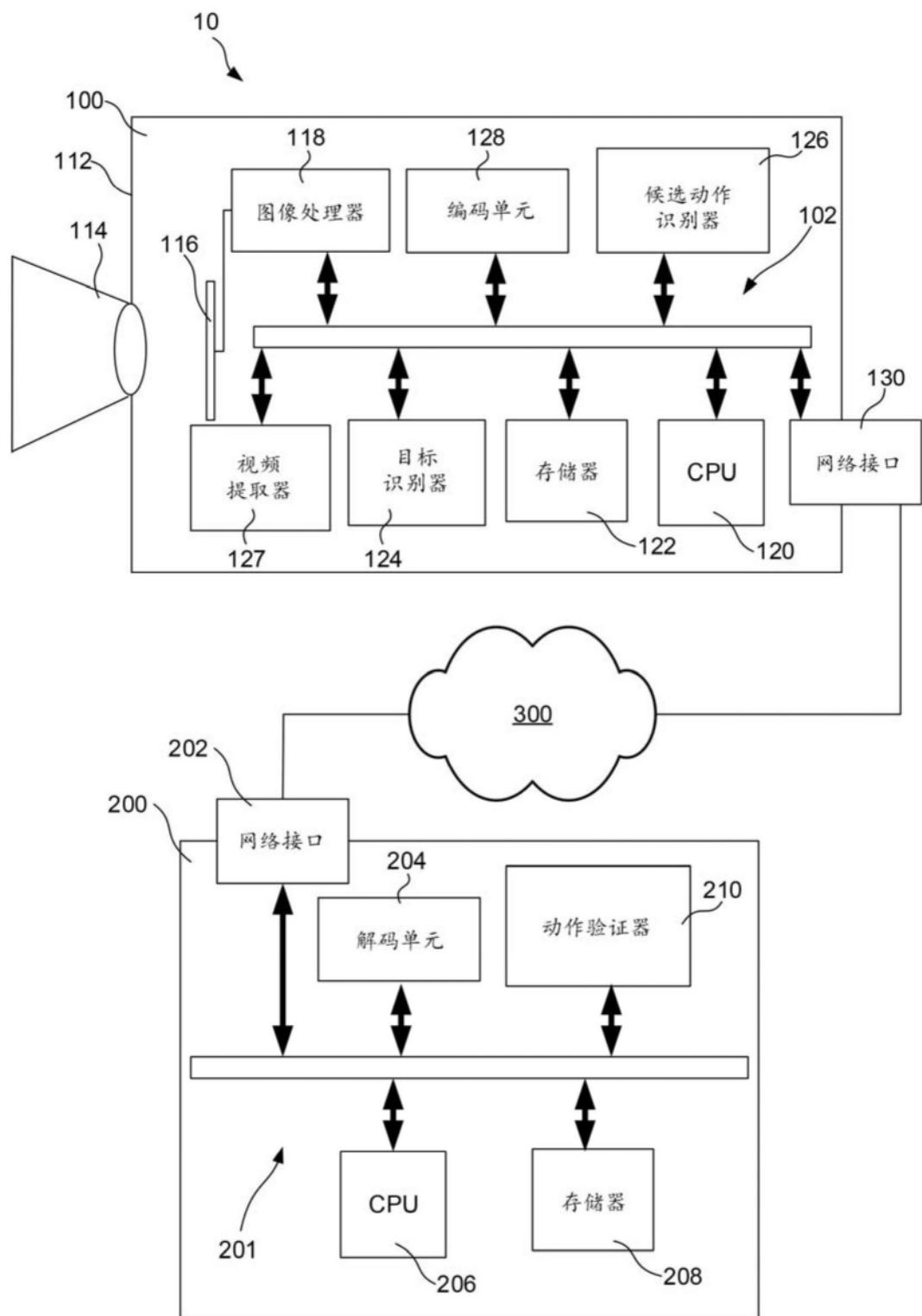


图1

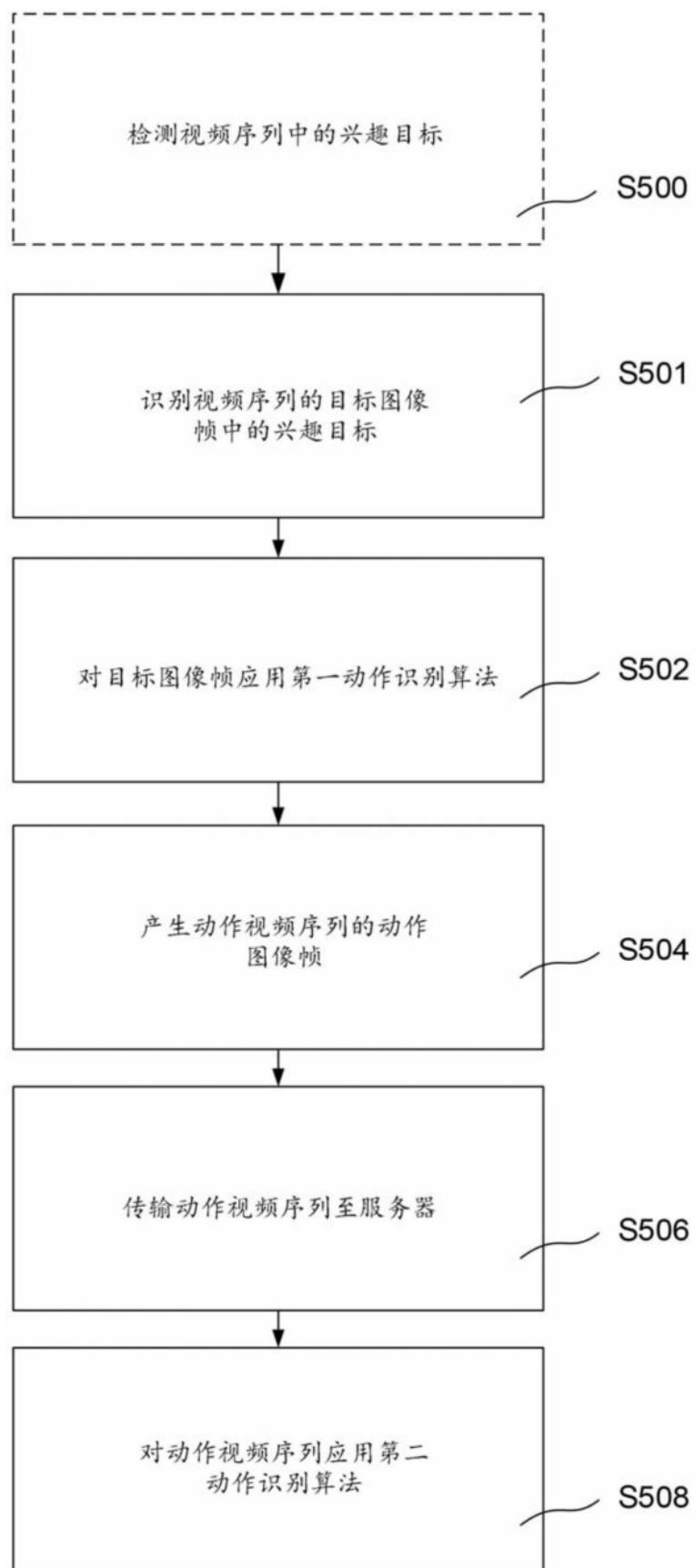


图2