



(12) 发明专利申请

(10) 申请公布号 CN 118786423 A

(43) 申请公布日 2024. 10. 15

(21) 申请号 202380023890.5

(22) 申请日 2023.02.16

(30) 优先权数据

202241011422 2022.03.03 IN

(85) PCT国际申请进入国家阶段日

2024.08.27

(86) PCT国际申请的申请数据

PCT/US2023/013252 2023.02.16

(87) PCT国际申请的公布数据

WO2023/167791 EN 2023.09.07

(71) 申请人 高通股份有限公司

地址 美国加利福尼亚州

(72) 发明人 S·D·帕特尔 P·A·布德瓦尼

S·C·纳迪帕里 S·孔达帕蒂

(74) 专利代理机构 上海专利商标事务所有限公  
司 31100

专利代理师 姚丹红

(51) Int.Cl.

G06F 16/783 (2006.01)

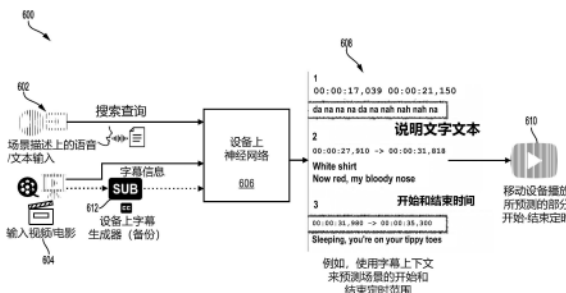
权利要求书3页 说明书15页 附图9页

(54) 发明名称

设备上人工智能视频搜索

(57) 摘要

一种用于使用人工神经网络(ANN)进行设备上视频查询和搜索的计算机实现的方法包括通过该ANN接收视频和搜索查询。视频包括帧序列和相关联字幕信息。在移动设备处通过ANN生成搜索查询中的第一组词语的第一表示和字幕信息中的第二组词语的第二表示。在移动设备处通过ANN基于第一表示和第二表示来确定搜索查询和字幕信息之间的相关性。ANN在移动设备处基于相关性来预测视频的包括响应于搜索查询的内容的部分。



1. 一种用于在移动设备上使用人工神经网络 (ANN) 来搜索视频的计算机实现的方法, 包括:

通过所述ANN接收所述视频和搜索查询, 所述视频包括帧序列和相关联字幕信息;

在所述移动设备处通过所述ANN生成所述搜索查询中的第一组词语的第一表示和所述字幕信息中的第二组词语的第二表示;

在所述移动设备处通过所述ANN基于所述第一表示和所述第二表示来确定相关性; 以及

在所述移动设备处通过所述ANN基于所述相关性来预测所述视频的包括响应于所述搜索查询的内容的部分。

2. 根据权利要求1所述的计算机实现的方法, 其中所述预测进一步指示所述视频的包括响应于所述搜索查询的所述内容的所述部分的开始时间和结束时间。

3. 根据权利要求2所述的计算机实现的方法, 还包括: 在所述开始时间显示所包括的所述视频部分直到所述结束时间为止。

4. 根据权利要求1所述的计算机实现的方法, 其中所述ANN包括变换器神经网络。

5. 根据权利要求1所述的计算机实现的方法, 还包括: 基于包括在所述视频中的隐藏式说明文字信息来生成所述相关联字幕信息。

6. 根据权利要求1所述的计算机实现的方法, 其中所述搜索查询包括场景的描述、事件、词语或短语中的一者或多者。

7. 根据权利要求1所述的计算机实现的方法, 其中所述搜索查询经由所述移动设备的语音输入文本输入来供应。

8. 一种用于在移动设备上使用人工神经网络 (ANN) 来搜索视频的装置, 包括:

存储器; 和

耦合到所述存储器的至少一个处理器, 所述至少一个处理器被配置为:

通过所述ANN接收所述视频和搜索查询, 所述视频包括帧序列和相关联字幕信息;

在所述移动设备处通过所述ANN生成所述搜索查询中的第一组词语的第一表示和所述字幕信息中的第二组词语的第二表示;

在所述移动设备处通过所述ANN基于所述第一表示和所述第二表示来确定相关性; 以及

在所述移动设备处通过所述ANN基于所述相关性来预测所述视频的包括响应于所述搜索查询的内容的部分。

9. 根据权利要求8所述的装置, 其中所述至少一个处理器被进一步配置为生成指示所述视频的包括响应于所述搜索查询的所述内容的所述部分的开始时间和结束时间的预测。

10. 根据权利要求9所述的装置, 其中所述至少一个处理器被进一步配置为在所述开始时间显示所包括的所述视频部分直到所述结束时间为止。

11. 根据权利要求8所述的装置, 其中所述ANN包括变换器神经网络。

12. 根据权利要求8所述的装置, 其中所述至少一个处理器被进一步配置为基于包括在所述视频中的隐藏式说明文字信息来生成所述相关联字幕信息。

13. 根据权利要求8所述的装置, 其中所述搜索查询包括场景的描述、事件、词语或短语中的一者或多者。

14. 根据权利要求8所述的装置,其中所述搜索查询经由所述移动设备的语音输入文本输入来供应。

15. 一种非暂态计算机可读介质,所述非暂态计算机可读介质上记录有用于在移动设备上使用人工神经网络(ANN)来搜索视频的程序代码,所述程序代码由处理器执行并且包括:

用于通过所述ANN接收所述视频和搜索查询的程序代码,所述视频包括帧序列和相关联字幕信息;

用于在所述移动设备处通过所述ANN生成所述搜索查询中的第一组词语的第一表示和所述字幕信息中的第二组词语的第二表示的程序代码;

用于在所述移动设备处通过所述ANN基于所述第一表示和所述第二表示来确定相关性的程序代码;和

用于在所述移动设备处通过所述ANN基于所述相关性来预测所述视频的包括响应于所述搜索查询的内容的部分的程序代码。

16. 根据权利要求15所述的非暂态计算机可读介质,其中所述程序代码还包括:用于生成指示所述视频的包括响应于所述搜索查询的所述内容的所述部分的开始时间和结束时间的预测的程序代码。

17. 根据权利要求16所述的非暂态计算机可读介质,其中所述程序代码还包括:用于在所述开始时间显示所包括的所述视频部分直到所述结束时间为止的程序代码。

18. 根据权利要求15所述的非暂态计算机可读介质,其中所述ANN包括变换器神经网络。

19. 根据权利要求15所述的非暂态计算机可读介质,其中所述程序代码还包括:用于基于包括在所述视频中的隐藏式说明文字信息来生成所述相关字幕信息的程序代码。

20. 根据权利要求15所述的非暂态计算机可读介质,其中所述搜索查询包括场景的描述、事件、词语或短语中的一者或多者。

21. 根据权利要求15所述的非暂态计算机可读介质,其中所述搜索查询经由所述移动设备的语音输入文本输入来供应。

22. 一种用于在移动设备上使用人工神经网络(ANN)来搜索视频的装置,包括:

用于通过所述ANN接收所述视频和搜索查询的部件,所述视频包括帧序列和相关联字幕信息;

用于在所述移动设备处通过所述ANN生成所述搜索查询中的第一组词语的第一表示和所述字幕信息中的第二组词语的第二表示的部件;

用于在所述移动设备处通过所述ANN基于所述第一表示和所述第二表示来确定相关性的部件;和

用于在所述移动设备处通过所述ANN基于所述相关性来预测所述视频的包括响应于所述搜索查询的内容的部分的部件。

23. 根据权利要求22所述的装置,还包括:用于生成指示所述视频的包括响应于所述搜索查询的所述内容的所述部分的开始时间和结束时间的预测的部件。

24. 根据权利要求23所述的装置,还包括:用于在所述开始时间显示所包括的所述视频部分直到所述结束时间为止的部件。

25. 根据权利要求22所述的装置,其中所述ANN包括变换器神经网络。
26. 根据权利要求22所述的装置,还包括:用于基于包括在所述视频中的隐藏式说明文字信息来生成所述相关联字幕信息的部件。
27. 根据权利要求22所述的装置,其中所述搜索查询包括场景的描述、事件、词语或短语中的一者或多者。
28. 根据权利要求22所述的装置,其中所述搜索查询经由所述移动设备的语音输入文本输入来供应。

## 设备上人工智能视频搜索

[0001] 相关申请的交叉引用

[0002] 本申请要求于2022年3月3日提交的并且名称为“设备上人工智能视频搜索 (ON-DEVICE ARTIFICIAL INTELLIGENCE VIDEO SEARCH)”的印度专利申请202241011422号的优先权,该专利申请的公开内容全文以引用方式明确并入。

### 技术领域

[0003] 本公开的各方面整体涉及神经网络,并且更具体地,涉及使用人工神经网络的设备上视频搜索。

### 背景技术

[0004] 人工神经网络可包括互连的人工神经元组(例如,神经元模型)。人工神经网络可以是计算设备或表示为要由计算设备执行的方法。卷积神经网络是一种前馈人工神经网络。卷积神经网络可包括神经元集合,其中每个神经元具有感受野并且共同地拼出一输入空间。卷积神经网络(CNN)(诸如深度卷积神经网络(DCN))具有众多应用。具体而言,这些神经网络架构被用于各种技术,诸如图像识别、模式识别、语音识别、自动驾驶和其他分类任务。

[0005] 例如,边缘设备诸如智能电话或其他移动设备广泛用于消费媒体诸如音乐或视频。考虑到神经网络的许多有用应用,对在边缘设备上使用的需求不断增加。搜索视频、歌曲或其他序列内的特定内容是用户的常见任务。例如,用户可能常常期望在不观看整个电影或视频的情况下播放例如电影中最喜爱或难忘的场景、重要事件(例如,目标)对话或视频中的谈话。然而,从功率的角度来看,自动搜索此类事件是麻烦、耗时且计算昂贵的。这在资源有限的设备诸如移动设备中尤为严重。

### 发明内容

[0006] 本公开在独立权利要求中分别阐述。本公开的一些方面在从属权利要求中描述。

[0007] 在本公开的一个方面,一种用于在移动设备上使用人工神经网络(ANN)来搜索视频的计算机实现的方法包括:通过该ANN接收视频和搜索查询。视频包括帧序列和相关联字幕信息。该计算机实现的方法还包括:在移动设备处通过ANN生成搜索查询中的第一组词语的第一表示和字幕信息中的第二组词语的第二表示。该计算机实现的方法还包括:在移动设备处通过ANN基于第一表示和第二表示来确定相关性。该计算机实现的方法还包括:在移动设备处通过ANN基于相关性来预测视频的包括响应于搜索查询的内容的部分。

[0008] 本公开的另一方面涉及一种装置,该装置包括:用于通过ANN接收视频和搜索查询的部件。视频包括帧序列和相关联字幕信息。装置还包括:用于在移动设备处通过ANN生成搜索查询中的第一组词语的第一表示和字幕信息中的第二组词语的第二表示的部件。此外,装置包括:用于在移动设备处通过ANN基于第一表示和第二表示来确定相关性的部件。装置还包括:用于在移动设备处通过ANN基于相关性来预测视频的包括响应于搜索查询的

内容的部分的部件。

[0009] 在本发明的另一方面,提出了一种非暂态计算机可读介质。该非暂态计算机可读介质上记录有用于在移动设备上使用人工神经网络(ANN)来搜索视频的程序代码。该程序代码由处理器执行并且包括:用于通过ANN接收视频和搜索查询的程序代码。视频包括帧序列和相关联字幕信息。程序代码还包括用于在移动设备处通过ANN生成搜索查询中的第一组词语的第一表示和字幕信息中的第二组词语的第二表示的程序代码。该程序代码还包括用于在该动设备处通过ANN基于第一表示和第二表示来确定相关性的程序代码。该程序代码还包括用于在移动设备处通过ANN基于相关性来预测视频的包括响应于搜索查询的内容的部分的程序代码。

[0010] 本公开的另一方面涉及一种用于在移动设备上使用人工神经网络(ANN)来搜索视频的装置。该装置具有存储器和耦合至该存储器的一个或多个处理器。该处理器被配置为:通过ANN接收视频和搜索查询。视频包括帧序列和相关联字幕信息。处理器还被配置为:在移动设备处通过ANN生成搜索查询中的第一组词语的第一表示和字幕信息中的第二组词语的第二表示。处理器被附加地配置为:在移动设备处通过ANN基于第一表示和第二表示来确定相关性。该处理器被进一步配置为:在移动设备处通过ANN基于相关性来预测视频的包括响应于搜索查询的内容的部分。

[0011] 本公开的附加特征和优点将在下文描述。本领域技术人员应当理解,本公开可容易被用作修改或设计用于实施与本公开相同的目的的其他结构的基础。本领域技术人员还应认识到,此类等效构造并不脱离所附权利要求中所阐述的本公开的教导内容。被认为是本公开的特性的新颖特征在其组织和操作方法两方面连同进一步的目的和优点在结合附图来考虑以下描述时将被更好地理解。然而,要明确理解的是,提供每一幅附图均仅用于例示和描述目的,并且无意作为对本公开的限定的定义。

## 附图说明

[0012] 在结合附图理解下面阐述的具体实施方式时,本公开的特征、实质和优点将变得更加明显,在附图中,相同参考字符始终相应地进行标识。

[0013] 图1例示了根据本公开的某些方面的使用片上系统(SOC)(包括通用处理器)的神经网络的示例具体实施。

[0014] 图2A、图2B和图2C是例示根据本公开的各方面的神经网络的示意图。

[0015] 图2D是例示根据本公开的各方面的示例性深度卷积网络(DCN)的示意图。

[0016] 图3是例示根据本公开的各方面的示例性深度卷积网络(DCN)的框图。

[0017] 图4是例示根据本公开的各方面的可使人工智能(AI)功能模块化的示例性软件架构的框图。

[0018] 图5是例示根据本公开的各方面的用于设备上视频查询和搜索的示例过程的高级框图。

[0019] 图6A是例示根据本公开的各方面的用于视频查询和搜索的示例架构的框图。

[0020] 图6B是例示根据本公开的各方面的示例设备上神经网络的框图。

[0021] 图7是例示根据本公开的各方面的用于使用人工神经网络的视频查询和搜索的示例过程的流程图。

## 具体实施方式

[0022] 以下结合附图阐述的具体实施方式旨在作为对各种配置的描述,而不旨在表示可实践所描述的概念的仅有配置。为了提供对各种概念的全面理解,具体实施方式包括具体细节。然而,对于本领域的技术人员来说显而易见的是,可以在没有这些具体细节的情况下实践这些概念。在一些情况下,为了避免模糊此类概念,公知的结构和组件是以框图形式示出的。

[0023] 基于教导内容,本领域技术人员应当认识到,本公开的范围旨在涵盖本公开的任何方面,无论该方面是独立于本公开的任何其他方面来实现的还是与任何其他方面结合地实现的。例如,使用所阐述的任意数量的方面可实现装置或可实践方法。此外,本公开的范围旨在覆盖使用作为所阐述的本公开的各个方面的补充或者与之不同的其他结构、功能性、或者结构和功能性来实践的这种装置或方法。应当理解,所公开的本公开的任何方面可以通过权利要求的一个或多个元素来体现。

[0024] 词语“示例性”用于意指“用作示例、实例、或例示”。描述为“示例性”的任何方面不必被解释为优于或胜过其他方面。

[0025] 尽管描述了特定方面,但这些方面的众多变体和置换落在本公开的范围之内。虽然提到了优选方面的一些益处和优点,但本公开的范围并非旨在被限定于特定益处、用途或目的。相反,本公开的各方面旨在能够广泛地应用于不同的技术、系统配置、网络和协议,其中一些作为示例在附图以及以下对优选方面的描述中例示。具体实施方式和附图仅仅是对本公开的例示而不是限制,本公开的范围由所附权利要求书及其等效物来限定。

[0026] 如所描述,搜索视频内的特定内容是用户的常见任务。用户可能常常期望在不观看整个电影或视频的情况下播放例如电影中最喜爱或难忘的场景、重要事件(例如,目标)对话或视频中的谈话。然而,从功率的角度来看,自动搜索此类事件是麻烦、耗时且计算昂贵的。这在资源有限的设备诸如移动设备中尤为严重。

[0027] 用于视频搜索的常规技术包括培养对视频帧的语义理解,并且涉及搜索视频帧中的每个视频帧以寻找在搜索中指示的正确视频时刻。为了产生更快的结果,一些常规技术还包括对所有视频帧进行预处理以及将学习存储在高速缓存的语料库/数据库中。然而,对视频帧的语义理解是困难的。附加地,此类常规技术是耗时的,在搜索视频上花费大量时间。此外,常规技术是计算昂贵的,这在资源有限的设备诸如移动设备中尤为严重。

[0028] 为了解决这些和其他挑战,本公开的各方面涉及设备上视频查询和搜索。人工神经网络可处理来自用户的搜索查询、来自视频的字幕信息,并且生成指示视频的最有可能包括所搜索的内容的部分的预测。在一些方面,人工神经网络可生成与搜索查询匹配或者可能包括所搜索的内容的可能事件的N个预测的列表(例如,如果搜索查询描述发生多次的事件),其中N为整数。

[0029] 在一些方面,该预测可包括用于标识包括所搜索的内容的部分的开始时间和结束时间。所预测的部分可在视频或媒体播放器中显示为文本或时间戳(例如,时间线位置),从而允许用户容易地导航到所预测的部分以播放所搜索的内容,或者在一些方面,移动设备可自动地播放所预测的部分。处理和预测可在设备上进行。设备上可指在不借助于云或远程计算的情况下进行处理和预测。

[0030] 图1例示了片上系统(SOC)100的示例具体实施,该片上系统(SOC)可包括被配置用

于使用人工神经网络(例如,神经端到端网络)来搜索视频的中央处理单元(CPU)102或多核CPU。变量(例如,神经信号和突触权重)、与计算设备相关联的系统参数(例如,具有权重的神经网络)、延迟、频率槽信息和任务信息可存储在与神经处理单元(NPU)108相关联的存储器块中、与CPU 102相关联的存储器块中、与图形处理单元(GPU)104相关联的存储器块中、与数字信号处理器(DSP)106相关联的存储器块中、存储器块118中,或者可跨多个块分布。在CPU 102处执行的指令可从与CPU 102相关联的程序存储器加载,或者可从存储器块118加载。

[0031] SOC 100还可包括针对特定功能定制的附加处理块,诸如GPU 104、DSP 106、连通性块110(其可包括第五代(5G)连通性、第四代长期演进(4G LTE)连通性、Wi-Fi连通性、USB连通性、蓝牙连通性等)和可例如检测和识别姿势的多媒体处理器112。在一种具体实施中,NPU 108在CPU 102、DSP 106和/或GPU 104中实现。SOC 100还可包括传感器处理器114、图像信号处理器(ISP)116和/或导航模块120,该导航模块可包括全球定位系统。

[0032] SOC 100可基于ARM指令集。在本公开的各方面,加载到通用处理器102中的指令可包括用于通过ANN接收视频和搜索查询的代码。视频包括帧序列和相关联字幕信息。通用处理器102还可包括用于在移动设备处通过ANN生成搜索查询中的第一组词语的第一表示和字幕信息中的第二组词语的第二表示的代码。通用处理器102可附加地包括用于在移动设备处通过ANN基于第一表示和第二表示来确定相关性的代码。通用处理器102还可包括用于在移动设备处通过ANN基于相关性来预测视频的包括响应于搜索查询的内容的部分的代码。

[0033] 深度学习架构可通过学习在每一层中以连续更高的抽象级别表示输入来执行对象识别任务,从而构建输入数据的有用特征表示。以这种方式,深度学习解决了传统机器学习的主要瓶颈。在深度学习出现之前,针对对象识别问题的机器学习方法可能严重依赖于人类工程化特征,可能与浅分类器组合。浅分类器可以是两类线性分类器,例如,其中可将特征向量分量的加权和与阈值进行比较以预测输入属于哪一类。人类工程化特征可以由具有领域专业知识的工程师针对特定问题领域定制的模板或内核。相反,虽然深度学习架构可学习表示与人类工程师可能设计的特征类似的特征,但是需要通过训练。此外,深度网络可学习表示和识别人类可能尚未考虑的新类型的特征。

[0034] 深度学习架构可学习特征的层次结构。例如,如果用视觉数据呈现,则第一层可学习识别输入流中相对简单的特征,诸如边缘。在另一个示例中,如果用听觉数据呈现,则第一层可学习识别特定频率中的频谱功率。第二层以第一层的输出作为输入,可学习识别特征的组合,诸如视觉数据的简单形状或听觉数据的声音组合。例如,较高层可学习在视觉数据中表示复杂的形状或在听觉数据中表示词语。再更高层可学习识别常见的视觉对象或口语短语。

[0035] 当应用于具有自然层次结构的问题时,深度学习架构可表现得尤其出色。例如,机动化交通工具的分类可受益于首先学习识别轮、挡风玻璃和其他特征。这些特征可在较高层以不同的方式进行组合,以识别汽车、卡车和飞机。

[0036] 神经网络可被设计成具有多种连通性模式。在前馈网络中,信息从较低层传递到较高层,其中给定层中的每个神经元与较高层中的神经元通信。如上所述,可在前馈网络的连续层中构建分层表示。神经网络还可具有循环或反馈(也被称为自上而下)连接。在循环

连接中,来自给定层中的神经元的输出可被传达给相同层中的另一神经元。循环架构可有助于识别跨越按顺序递送给神经网络的输入数据块中的多于一个输入数据块的模式。从给定层中的神经元到较低层中的神经元的连接被称为反馈(或自上而下)连接。当高层级概念的识别可辅助辨别输入的特定低层级特征时,具有许多反馈连接的网络可能是有帮助的。

[0037] 神经网络各层之间的连接可以是全连接的,或者是局部连接的。图2A例示了全连接的神经网络202的示例。在全连接的神经网络202中,第一层中的神经元可将其输出传达到第二层中的每个神经元,使得第二层中的每个神经元将从第一层中的每个神经元接收输入。图2B例示了局部连接的神经网络204的示例。在局部连接的神经网络204中,第一层中的神经元可连接到第二层中的有限数量的神经元。更一般地,局部连接的神经网络204的局部连接层可被配置为使得层中的每个神经元将具有相同或类似的连通性模式但具有可具有不同值的连接强度(例如,210、212、214和216)。局部连接的连通性模式可能会在较高层中产生空间上不同的感受野,因为给定区域中的较高层神经元可以接收输入,该输入通过训练被调谐到网络的总输入的受限部分的属性。

[0038] 局部连接的神经网络的一个示例是卷积神经网络。图2C例示了卷积神经网络206的示例。卷积神经网络206可被配置为使得与第二层中的每个神经元的输入相关联的连接强度是共享的(例如,208)。卷积神经网络可能非常适合于其中输入的空间位置有意义的问题。

[0039] 一种类型的卷积神经网络是深度卷积网络(DCN)。图2D例示了被设计成从自图像捕获设备230诸如车载相机输入的图像226识别视觉特征的DCN 200的详细示例。当前示例的DCN 200可被训练以标识交通标志和设置在该交通标志上的数字。当然,DCN 200可被训练用于其他任务,诸如标识车道标记或者标识交通灯。

[0040] DCN 200可利用监督学习进行训练。在训练期间,可向DCN 200呈现图像诸如限速标志的图像226,并且然后可计算前向传递以产生输出222。DCN 200可包括特征提取部和分类部。在接收到图像226时,卷积层232可将卷积内核(未示出)应用于图像226以生成第一组特征图218。作为一示例,卷积层232的卷积内核可以是生成28x28特征图的5x5内核。在本示例中,因为在第一组特征图218中生成四个不同特征图,所以在卷积层232处将四个不同卷积内核应用于图像226。卷积内核也可称为滤波器或卷积滤波器。

[0041] 第一组特征图218可由最大池化层(未示出)进行子采样以生成第二组特征图220。最大池化层减小第一组特征图218的大小。也就是说,第二组特征图220的大小(诸如14x14)小于第一组特征图218的大小(诸如28x28)。减小的大小向后续层提供类似的信息,同时减少存储器消耗。第二组特征图220可经由一个或多个后续卷积层(未示出)进一步卷积以生成一组或多组后续特征图(未示出)。

[0042] 在图2D的示例中,第二组特征图220被卷积以生成第一特征向量224。此外,第一特征向量224被进一步卷积以生成第二特征向量228。第二特征向量228中的每个特征可包括对应于图像226的可能特征的数字,诸如“标志”、“60”和“100”。softmax函数(未示出)可将第二特征向量228中的数字转换为概率。因此,DCN 200的输出222是图像226包括一个或多个特征的概率。

[0043] 在本示例中,输出222中“标志”和“60”的概率高于输出222中的其他数字(诸如“30”、“40”、“50”、“70”、“80”、“90”和“100”)的概率。在训练之前,由DCN 200产生的输出222

可能是不正确的。因此,可计算输出222和目标输出之间的误差。目标输出是图像226的真实值(例如,“标志”和“60”)。然后,可调整DCN 200的权重,使得DCN 200的输出222与目标输出更紧密地对准。

[0044] 为了调整权重,学习算法可计算权重的梯度向量。该梯度可指示权重被调整时误差将增加或减小的量。在顶层处,梯度可直接对应于连接倒数第二层中的激活神经元和输出层中的神经元的权重的值。在较低层中,梯度可取决于权重的值和较高层的计算的误差梯度。该权重可随后被调整以减小误差。这种调整权重的方式可称为“反向传播”,因为它涉及通过神经网络的“反向传递”。

[0045] 在实践中,可在少量示例来计算权重的误差梯度,使计算出的梯度接近真实的误差梯度。这种近似方法可称为随机梯度下降。可重复随机梯度下降,直到整个系统的可实现错误率停止下降或直到错误率达到目标水平。在学习之后,可向DCN呈现新图像,并且通过网络的前向传递可产生可被认为是DCN的推断或预测的输出222。

[0046] 深度置信网络(DBN)是包括多层隐藏节点的概率模型。DBN可用于提取训练数据集的分层表示。DBN可通过对受限玻尔兹曼机(RBM)的层进行堆叠来获得。RBM是一类可在输入集上学习概率分布的人工神经网络。因为RBM可在没有关于每个输入应该被分类到的类别的信息的情况下学习概率分布,所以RBM通常用于无监督学习。使用混合无监督和监督范式,DBN的底部RBM可以无监督的方式进行训练并且可用作特征提取器,并且顶部RBM可以监督的方式进行训练(在来自前一层的输入和目标类的联合分布上)并且可用作分类器。

[0047] 深度卷积网络(DCN)是被配置有附加池化和归一化层的卷积网络的网络。DCN在许多任务上都达到了最先进的性能。DCN可使用监督学习来训练,其中输入目标和输出目标两者对于许多范例是已知的并且被用于通过使用梯度下降方法来修改网络的权重。

[0048] DCN可以是前馈网络。此外,如上文所述,从DCN的第一层中的神经元到下一较高层中的一组神经元的连接跨第一层中的各神经元共享。DCN的前馈和共享连接可用于快速处理。例如,DCN的计算负担可能比包括循环或反馈连接的类似大小的神经网络的计算负担小得多。

[0049] 卷积网络的每一层的处理可被认为是空间不变的模板或基础投影。如果输入首先被分解为多个通道,诸如彩色图像的红、绿和蓝通道,那么在该输入上训练的卷积网络可以被认为是三维的,其中,两个空间维度沿着图像的轴,并且第三维度捕获颜色信息。卷积连接的输出可被认为在后续层中形成特征图,该特征图(例如,220)中的每个元素从前一层(例如,特征图218)中一定范围的神经元以及从该多个通道中的每个通道接收输入。特征图中的值可用非线性(诸如校正, $\max(0, x)$ )进一步处理。来自相邻神经元的值可进一步池化,这对应于下采样,并且可以提供附加的局部不变性和降维。对应于白化的归一化也可通过特征图中神经元之间的侧向抑制来应用。

[0050] 深度学习架构的性能可随着更多标记数据点变得可用或随着计算能力的增加而增加。现代深度神经网络通常用计算资源来训练,这些计算资源是仅十五年前的典型研究人员可用的计算资源的数千倍。新的架构和训练范式可进一步提升深度学习的性能。校正的线性单元可减少被称为梯度消失的训练问题。新的训练技术可减少过度拟合,从而使得更大的模型能够实现更好的泛化。封装技术可提取给定感受野中的数据,并进一步提高整体性能。

[0051] 图3是例示深度卷积网络350的框图。基于连通性和权重共享,深度卷积网络350可包括多个不同类型的层。如图3所示,深度卷积网络350包括卷积块354A、354B。卷积块354A、354B中的每个卷积块可被配置有卷积层(CONV) 356、归一化层(LNorm) 358和最大池化层(MAX POOL) 360。

[0052] 卷积层356可包括一个或多个卷积滤波器,该一个或多个卷积滤波器可应用于输入数据以生成特征图。尽管示出了仅两个卷积块354A、354B,但本公开不限于此,而是替代地,可根据设计偏好在深度卷积网络350中包括任何数量的卷积块354A、354B。归一化层358可对卷积滤波器的输出进行归一化。例如,归一化层358可提供白化或侧向抑制。最大池化层360可在空间上提供下采样聚合以实现局部不变性和降维。

[0053] 例如,深度卷积网络的并行滤波器组可被加载在SOC 100的CPU 102或GPU 104上以实现高性能和低功耗。在另选实施方案中,并行滤波器组可被加载在SOC 100的DSP 106或ISP 116上。此外,深度卷积网络350可访问可存在于SOC 100上的其他处理块,诸如分别专用于传感器和导航的传感器处理器114和导航模块120。

[0054] 深度卷积网络350还可包括一个或多个全连接层362 (FC1和FC2)。深度卷积网络350还可包括逻辑回归(LR)层364。深度卷积网络350的每个层356、358、360、362、364之间都有要更新的权重(未示出)。这些层(例如,356、358、360、362、364)中的每个层的输出可用作深度卷积网络350中这些层(例如,356、358、360、362、364)中的后一个层的输入,以从在第一个卷积块354A处供应的输入数据352(例如,图像、音频、视频、传感器数据和/或其他输入数据)学习分层特征表示。深度卷积网络350的输出是输入数据352的分类分数366。分类分数366可以是一组概率,其中每个概率是输入数据包括来自一组特征的特征的概率。

[0055] 图4是例示可使人工智能(AI)功能模块化的示例性软件架构400的框图。根据本公开的各方面,通过使用该架构,可以设计可使得SoC 420的各种处理块(例如,CPU 422、DSP 424、GPU 426和/或NPU 428)支持如所公开的用于针对AI应用402的后训练量化的自适应舍入的应用。

[0056] AI应用402可被配置为调用在用户空间404中定义的功能,这些功能可例如提供对指示设备当前操作的位置的场景的检测和识别。AI应用402可例如根据所识别的场景是办公室、演讲厅、餐厅还是户外环境诸如湖泊来以不同方式配置麦克风和相机。AI应用402可作出对与在AI功能应用程序编程接口(API) 406中定义的库相关联的编译程序代码的请求。该请求可最终依赖于被配置为基于例如视频和定位数据来提供推断响应的深度神经网络的输出。

[0057] 运行时引擎408(其可以是运行时框架的编译代码)可进一步可由AI应用402访问。AI应用402可使得运行时引擎例如按特定时间间隔或通过由应用的用户接口检测到的事件触发来请求推断。在使得运行时引擎提供推断响应时,运行时引擎可进而发送信号给在SOC 420上运行的操作系统(OS)空间410(诸如内核412)中的操作系统。操作系统进而可使得在CPU 422、DSP 424、GPU 426、NPU 428或它们的某种组合上执行连续量化松弛。CPU 422可由操作系统直接访问,并且其他处理块可通过驱动器(诸如分别用于DSP 424、GPU 426或NPU 428的驱动器414、416或418)来访问。在示例性示例中,深度神经网络可被配置为在处理块(诸如CPU 422、DSP 424和GPU 426)的组合上运行,或者可在NPU 428上运行。

[0058] 应用402(例如,AI应用)可被配置为调用在用户空间404中定义的功能,这些功能

可例如提供对指示设备当前操作的位置的场景的检测和识别。应用402可例如根据所识别的场景是办公室、演讲厅、餐厅还是户外环境诸如湖泊来以不同方式配置麦克风和相机。应用402可作出对与在场景检测应用程序编程接口(API)406中定义的库相关联的编译程序代码的请求,以提供当前场景的估计。该请求可最终依赖于被配置为基于例如视频和定位数据来提供场景估计的差分神经网络的输出。

[0059] 运行时引擎408(其可以是运行时框架的编译代码)可进一步可由应用402访问。应用402可使得运行时引擎例如按特定时间间隔或通过由应用的用户接口检测到的事件触发来请求场景估计。在使得运行时引擎估计场景时,运行时引擎可进而发送信号给在SOC 420上运行的操作系统410(诸如内核412)。操作系统410进而可使得在CPU 422、DSP 424、GPU 426、NPU 428或它们的某种组合上执行计算。CPU 422可由操作系统直接访问,并且其他处理块可通过驱动器(诸如用于DSP 424、GPU 426或NPU 428的驱动器414-418)来访问。在示例性示例中,差分神经网络可被配置为在处理块(诸如CPU 422和GPU 426)的组合上运行,或者可在NPU 428上运行。

[0060] 如所描述,本公开的各方面涉及设备上视频查询和搜索。人工神经网络可处理来自用户的搜索查询、来自视频的字幕信息,并且生成指示视频的最有可能包括所搜索的内容的部分的预测。在一些方面,人工神经网络可生成与搜索查询匹配或者可能包括搜索的内容的可能事件的N个预测的列表(例如,如果搜索查询描述发生多次的事件),其中N为整数。

[0061] 在一些方面,该预测可包括用于标识包括搜索的内容的部分的开始时间和结束时间。所预测的部分可在视频或媒体播放器中显示为文本或时间戳(例如,时间线位置)。因此,用户可更容易地导航到所预测的部分以播放所搜索的内容,或者在一些方面,移动设备可自动地播放所预测的部分。处理和预测可在设备上进行。设备上可指在不借助于云或远程计算的情况下使用移动设备的资源进行处理和预测。

[0062] 图5是例示根据本公开的各方面的用于设备上视频查询和搜索的示例过程500的高级框图。参考图5,在框502处,过程500可经由移动设备接收搜索查询和视频。搜索查询可包括例如视频中的场景的描述或事件。

[0063] 在框504处,可处理搜索查询和与视频相关联的字幕信息。人工神经网络(例如,图3的350)可处理搜索查询和字幕信息。字幕信息可包括视频中的场景或动作的描述。字幕信息可在逐帧的基础上以段落的形式提供。ANN可确定搜索查询和字幕信息的特征。在一些方面,ANN可确定搜索查询的突出部分并且搜索与视频相关联的字幕信息,以用于匹配特征。

[0064] 在框506处,ANN可生成视频的包括响应于搜索查询的主题材料的部分的预测。预测可包括最可能包括搜索查询中所描述的场景或事件的一组帧。在一些方面,该预测可包括视频的最可能包括搜索查询中所描述的场景或事件的部分的开始时间和结束时间。

[0065] 在框508处,可在移动设备的显示器处(例如,经由图1的多媒体处理器112)显示视频的所预测的部分。

[0066] 图6A是例示根据本公开的各方面的用于视频查询和搜索的示例架构600的框图。参考图6A,架构600可包括输入设备602。输入设备602可包括移动设备的麦克风(例如,图1的传感器114)、多媒体设备(例如,图1的多媒体处理器112)或用于接收文本输入的输入/输出设备。在一些方面,输入设备602可被配置为将语音信号或其他传感器输入转换为文本输

入。输入设备602可从用户接收搜索查询。搜索查询可以是包括在电影中的场景的描述或事件。搜索查询可呈词语、短语或句子的形式。

[0067] 架构600还可接收视频输入604。视频输入604可例如是电影。视频输入604可从存储装置或流媒体源接收。视频输入604可包括帧的时间序列。在一些方面,视频输入604可包括相关联隐藏式说明文字(CC)信息或字幕信息。隐藏式说明文字信息(其也可称为字幕信息)可包括电影中在视频的帧中显示的人物之间的对话的定时转录。在一些方面,字幕生成器612可任选地包括在设备上,并且例如在字幕信息不与视频输入604一起包括时用于生成视频输入604的帧的这种信息。例如,字幕信息可呈句子或段落的形式。

[0068] 搜索查询和视频的字幕信息可被供应给设备上神经网络606。例如,设备上神经网络606可以是变换器神经网络。变换器神经网络是使用自注意力并且提供输入序列内的任何位置的上下文信息的深度学习模型。在一些方面,设备上神经网络606可以是对词元替换进行准确分类的高效学习编码器(ELECTRA)小模型。当然,这仅仅是示例,并且还可采用其他架构,诸如来自变换器的双向编码器表示(BERT)、稳健优化的BERT方法(RoBERTa)、XLNet、变换器-XL和生成式预训练变换器(GPT)变换器系列。ELECTRA小模型是问题回答自然语言处理器(NLP)。ELECTRA小模型可被预训练以预测给定查询和输入段落的回答。

[0069] 设备上神经网络606可生成搜索查询中的词语的表示和字幕信息中的词语的表示。词语的表示中的每个表示可包括分别基于搜索查询和字幕信息中的词语中的每个词语的上下文信息。设备上神经网络606可基于词语的所生成的表示来确定搜索查询和字幕信息之间的相关性。进而,设备上神经网络606可基于相关性来生成视频输入604的包括响应于搜索查询的内容的部分的预测608。预测608可指示视频的最有可能包括所搜索的内容的部分。在一些方面,预测608可指示包括这种内容的一个或多个帧。例如,如图6A的示例所示,设备上神经网络606可生成三个候选部分,每个候选部分包括可满足搜索查询的字幕信息。在一些方面,预测608还可包括视频输入604的包括响应于搜索查询的内容的部分或片段的开始时间和结束时间。

[0070] 回放设备610可显示视频输入604的由开始时间和结束时间指示的所预测的部分的列表。此外,在一些方面,回放设备可导航(例如,快进)到所标识的开始时间,并且开始回放视频输入604的所预测的部分,直到所标识的结束时间为止。

[0071] 图6B是示例根据本公开的各方面的示例设备上神经网络606的框图。参考图6B,设备上神经网络606可包括搜索生成器网络620和鉴别器网络630。例如,搜索生成器网络620和鉴别器网络630可各自被配置为变换器网络。搜索生成器网络620可接收搜索查询作为输入,并且可将搜索查询映射到上下文向量表示 $h_{SQ}$ 。上下文向量表示 $h_{SQ}$ 可聚焦于搜索查询内的重要词语(例如,名词、动词)。上下文向量表示 $h_{SQ}$ 可被供应给鉴别器网络630。

[0072] 鉴别器网络630可接收字幕信息作为第一输入。鉴别器网络630可将与视频输入604的每个帧相关联的字幕信息映射到上下文向量表示 $h_{SI}$ 。在一些方面,鉴别器网络630可生成上下文向量表示 $h_{SI}$ ,使得该上下文向量表示聚焦于字幕信息内的重要词语。

[0073] 鉴别器网络630还可接收上下文向量表示 $h_{SQ}$ 。鉴别器网络630可将上下文向量表示 $h_{SQ}$ 与上下文向量表示 $h_{SI}$ 进行比较,以确定上下文向量表示 $h_{SQ}$ 与上下文向量表示 $h_{SI}$ 之间的相关性。也就是说,鉴别器网络630可被训练以将数据的词语(例如,仅在搜索查询中)与包括在字幕信息中的词语区分开。基于上下文向量表示 $h_{SQ}$ 与上下文向量表示 $h_{SI}$ 之间的相关

性,鉴别器网络630可生成搜索查询是否与字幕信息匹配并因此视频输入604的对应部分是否包括满足搜索查询的内容的预测(例如,608)。

[0074] 图7是例示根据本公开的各方面的用于使用人工神经网络(ANN)的视频查询和搜索的示例过程700的流程图。如图7所示,在框702处,过程700通过ANN接收视频和搜索查询,该视频包括帧序列和相关联字幕信息。例如,如参考图6A所描述,输入设备602可从用户接收搜索查询。搜索查询可以是包括在电影中的场景的描述或事件。搜索查询可呈词语、短语或句子的形式。架构600还可接收视频输入604。视频输入604可例如是电影。视频输入604可从存储装置或流媒体源接收。视频输入604可包括帧的时间序列。在一些方面,视频输入604可包括相关联隐藏式说明文字(CC)信息或字幕信息。隐藏式说明文字信息(其也可称为字幕信息)可包括电影中在视频的帧中显示的人物之间的对话的定时转录。

[0075] 在框704处,过程700在移动设备处通过ANN生成搜索查询中的第一组词语的第一表示和字幕信息中的第二组词语的第二表示。如参考图6B所描述,搜索生成器网络620可接收搜索查询作为输入,并且可将搜索查询映射到上下文向量表示 $h_{SQ}$ 。上下文向量表示 $h_{SQ}$ 可聚焦于搜索查询内的重要词语(例如,名词、动词)。类似地,鉴别器网络630可接收字幕信息。鉴别器网络630可将与视频输入604的每个帧相关联的字幕信息映射到上下文向量表示 $h_{SI}$ 。

[0076] 在框706处,过程700在移动设备处通过ANN基于第一表示和第二表示来确定相关性。例如,如参考图6B所描述,鉴别器网络630可将上下文向量表示 $h_{SQ}$ 与上下文向量表示 $h_{SI}$ 进行比较,以确定上下文向量表示 $h_{SQ}$ 与上下文向量表示 $h_{SI}$ 之间的相关性。也就是说,鉴别器网络630可被训练以将数据的词语(例如,仅在搜索查询中)与包括在字幕信息中的词语区分开。附加地,相关性可基于搜索查询的词语的上下文和字幕信息的词语的上下文的对应关系。

[0077] 在框708处,过程700在移动设备处通过ANN基于相关性来预测视频的包括响应于搜索查询的内容的部分。例如,如参考图6B所描述,基于上下文向量表示 $h_{SQ}$ 与上下文向量表示 $h_{SI}$ 之间的相关性,鉴别器网络630可生成搜索查询是否与字幕信息匹配并因此视频输入604的对应部分是否包括满足搜索查询的内容的预测(例如,608)。预测608可指示视频的最有可能包括所搜索的内容的部分。在一些方面,预测608可指示包括这种内容的一个或多个帧。在一些方面,预测608还可包括视频输入604的包括响应于搜索查询的内容的部分或片段的开始时间和结束时间。

[0078] 在以下编号的条款中包括各具体实施示例。

[0079] 1.一种用于在移动设备上使用人工神经网络(ANN)来搜索视频的计算机实现的方法,包括:

[0080] 通过所述ANN接收所述视频和搜索查询,所述视频包括帧序列和相关联字幕信息;

[0081] 在所述移动设备处通过所述ANN生成所述搜索查询中的第一组词语的第一表示和所述字幕信息中的第二组词语的第二表示;

[0082] 在所述移动设备处通过所述ANN基于所述第一表示和所述第二表示来确定相关性;以及

[0083] 在所述移动设备处通过所述ANN基于所述相关性来预测所述视频的包括响应于所述搜索查询的内容的部分。

[0084] 2. 根据条款1所述的计算机实现的方法,其中所述预测进一步指示所述视频的包括响应于所述搜索查询的所述内容的所述部分的开始时间和结束时间。

[0085] 3. 根据条款1或条款2所述的计算机实现的方法,还包括:在所述开始时间显示所包括的所述视频部分直到所述结束时间为止。

[0086] 4. 根据条款1-3中任一项所述的计算机实现的方法,其中所述ANN包括变换器神经网络。

[0087] 5. 根据条款1-4中任一项所述的计算机实现的方法,还包括:基于包括在所述视频中的隐藏式说明文字信息来生成所述相关联字幕信息。

[0088] 6. 根据条款1-5中任一项所述的计算机实现的方法,其中所述搜索查询包括场景的描述、事件、词语或短语中的一者或多者。

[0089] 7. 根据条款1-6中任一项所述的计算机实现的方法,其中所述搜索查询经由所述移动设备的语音输入文本输入来供应。

[0090] 8. 一种用于在移动设备上使用人工神经网络(ANN)来搜索视频的装置,包括:

[0091] 存储器;和

[0092] 耦合到所述存储器的至少一个处理器,所述至少一个处理器

[0093] 被配置为:

[0094] 通过所述ANN接收所述视频和搜索查询,所述视频包括帧序列和相关联字幕信息;

[0095] 在所述移动设备处通过所述ANN生成所述搜索查询中的第一组词语的第一表示和所述字幕信息中的第二组词语的第二表示;

[0096] 在所述移动设备处通过所述ANN基于所述第一表示和所述第二表示来确定相关性;以及

[0097] 在所述移动设备处通过所述ANN基于所述相关性来预测所述视频的包括响应于所述搜索查询的内容的部分。

[0098] 9. 根据条款8所述的装置,其中所述至少一个处理器被进一步配置为生成指示所述视频的包括响应于所述搜索查询的所述内容的所述部分的开始时间和结束时间的预测。

[0099] 10. 根据条款8或条款9所述的装置,其中所述至少一个处理器被进一步配置为在所述开始时间显示所包括的所述视频部分直到所述结束时间为止。

[0100] 11. 根据条款8-10中任一项所述的装置,其中所述ANN包括变换器神经网络。

[0101] 12. 根据条款8-11中任一项所述的装置,其中所述至少一个处理器被进一步配置为基于包括在所述视频中的隐藏式说明文字信息来生成所述相关联字幕信息。

[0102] 13. 根据条款8-12中任一项所述的装置,其中所述搜索查询包括场景的描述、事件、词语或短语中的一者或多者。

[0103] 14. 根据条款8-13中任一项所述的装置,其中所述搜索查询经由所述移动设备的语音输入文本输入来供应。

[0104] 15. 一种非暂态计算机可读介质,所述非暂态计算机可读介质上记录有用于在移动设备上使用人工神经网络(ANN)来搜索视频的程序代码,所述程序代码由处理器执行并且包括:

[0105] 用于通过所述ANN接收所述视频和搜索查询的程序代码,所述视频包括帧序列和相关联字幕信息;

[0106] 用于在所述移动设备处通过所述ANN生成所述搜索查询中的第一组词语的第一表示和所述字幕信息中的第二组词语的第二表示的程序代码;

[0107] 用于在所述移动设备处通过所述ANN基于所述第一表示和所述第二表示来确定相关性的程序代码;和

[0108] 用于在所述移动设备处通过所述ANN基于所述相关性来预测所述视频的包括响应于所述搜索查询的内容的部分的程序代码。

[0109] 16. 根据条款15所述的非暂态计算机可读介质,其中所述程序代码还包括:用于生成指示所述视频的包括响应于所述搜索查询的所述内容的所述部分的开始时间和结束时间的预测的程序代码。

[0110] 17. 根据条款15或条款16所述的非暂态计算机可读介质,其中所述程序代码还包括:用于在所述开始时间显示所包括的所述视频部分直到所述结束时间为止的程序代码。

[0111] 18. 根据条款15-17中任一项所述的非暂态计算机可读介质,其中所述ANN包括变换器神经网络。

[0112] 19. 根据条款15-18中任一项所述的非暂态计算机可读介质,其中所述程序代码还包括:用于基于包括在所述视频中的隐藏式说明文字信息来生成所述相关联字幕信息的程序代码。

[0113] 20. 根据条款15-19中任一项所述的非暂态计算机可读介质,其中所述搜索查询包括场景的描述、事件、词语或短语中的一者或多者。

[0114] 21. 根据条款15-20中任一项所述的非暂态计算机可读介质,其中所述搜索查询经由所述移动设备的语音输入文本输入来供应。

[0115] 22. 一种用于在移动设备上使用人工神经网络(ANN)来搜索视频的装置,包括:

[0116] 用于通过所述ANN接收所述视频和搜索查询的部件,所述视频包括帧序列和相关联字幕信息;

[0117] 用于在所述移动设备处通过所述ANN生成所述搜索查询中的第一组词语的第一表示和所述字幕信息中的第二组词语的第二表示的部件;

[0118] 用于在所述移动设备处通过所述ANN基于所述第一表示和所述第二表示来确定相关性的部件;和

[0119] 用于在所述移动设备处通过所述ANN基于所述相关性来预测所述视频的包括响应于所述搜索查询的内容的部分的部件。

[0120] 23. 根据条款22所述的装置,还包括:用于生成指示所述视频的包括响应于所述搜索查询的所述内容的所述部分的开始时间和结束时间的预测的部件。

[0121] 24. 根据条款22或条款23所述的装置,还包括:用于在所述开始时间显示所包括的所述视频部分直到所述结束时间为止的部件。

[0122] 25. 根据条款22-24中任一项所述的装置,其中所述ANN包括变换器神经网络。

[0123] 26. 根据条款22-25中任一项所述的装置,还包括:用于基于包括在所述视频中的隐藏式说明文字信息来生成所述相关联字幕信息的部件。

[0124] 27. 根据条款22-26中任一项所述的装置,其中所述搜索查询包括场景的描述、事件、词语或短语中的一者或多者。

[0125] 28. 根据条款22-27中任一项所述的装置,其中所述搜索查询经由所述移动设备的

语音输入文本输入来供应。

[0126] 在一个方面,接收部件、生成部件、确定部件、预测和/或显示部件可以是被配置为执行所叙述的功能的CPU 102、与CPU 102相关联的程序存储器、专用存储器块118、全连接层362、NPU 428和/或路由连接处理单元216。在另一配置中,前述部件可以是被配置为执行由前述部件所叙述的功能的任何模块或任何装置。

[0127] 上文所述的方法的各种操作可通过能够执行对应功能的任何合适的部件来执行。这些部件可包括各种硬件和/或软件组件和/或模块,包括但不限于电路、专用集成电路(ASIC)或处理器。一般来讲,在附图中所例示的操作的情况下,那些操作可具有带有类似编号的对应配对部件加功能组件。

[0128] 如本文所用,术语“确定”涵盖各种各样的动作。例如,“确定”可以包括计算、运算、处理、推导、调查、查找(例如,在表、数据库或另一数据结构中查找)、断定等等。附加地,“确定”可包括接收(例如,接收信息)、访问(例如,访问存储器中的数据)等。此外,“确定”可包括解析、选择、选取、建立等。

[0129] 如本文所用,提到条目列表“中的至少一项”的短语,指代这些条目的任意组合(其包括单一成员)。作为一个示例,“a、b或c中的至少一者”旨在涵盖:a、b、c、a-b、a-c、b-c和a-b-c。

[0130] 结合本公开所描述的各种例示性逻辑框、模块以及电路可用设计成执行所描述的功能的通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列信号(FPGA)或其他可编程逻辑器件(PLD)、分立门或晶体管逻辑、分立硬件组件或它们的任何组合来实现或执行。通用处理器可以是微处理器,但在另选方案中,处理器可以是任何市售的处理器、控制器、微控制器、或状态机。处理器还可以被实现为计算设备的组合,例如DSP与微处理器的组合、多个微处理器、与DSP核心结合的一个或多个微处理器,或任何其他此类配置。

[0131] 结合本公开描述的方法的步骤或者算法可以直接体现在硬件、由处理器执行的软件模块或二者的组合中。软件模块可驻留在本领域已知的任何形式的存储介质中。可使用的存储介质的一些示例包括随机存取存储器(RAM)、只读存储器(ROM)、闪存存储器、可擦除可编程只读存储器(EPROM)、电可擦除可编程只读存储器(EEPROM)、寄存器、硬盘、可移动磁盘、CD-ROM等。软件模块可包括单个指令或许多个指令,并且可分布在若干不同的代码段上,分布在不同的程序间以及跨多个存储介质分布。存储介质可以耦合到处理器,使得处理器能够从该存储介质读取信息,以及向该存储介质写入信息。在另选方案中,存储介质可与处理器成一整体。

[0132] 本文所公开的方法包括用于实现所描述方法的一个或多个步骤或动作。方法步骤和/或动作可以彼此互换而不偏离权利要求书的范围。换句话讲,除非指定了步骤或动作的特定顺序,否则在不脱离权利要求的范围的情况下,可以修改特定步骤和/或动作的顺序和/或使用。

[0133] 所描述的功能可以硬件、软件、固件或它们的任何组合来实现。如果以硬件实现,则示例硬件配置可包括设备中的处理系统。可以利用总线架构来实现该处理系统。取决于处理系统的具体应用和整体设计约束,总线可以包括任何数量的互连总线和桥接器。总线可以将各种电路链接在一起,包括处理器、机器可读介质和总线接口。此外,总线接口可用

于将网络适配器等经由总线连接到处理系统。网络适配器可用于实现信号处理功能。对于某些方面,用户接口(例如,小键盘、显示器、鼠标、操纵杆等)也可以连接到总线。总线还可以链接各种其他电路,诸如定时源、外围设备、稳压器、功率管理电路等,它们在本领域中是众所周知的,因此将不再进一步描述。

[0134] 处理器可以负责管理总线和通用处理,包括执行机器可读介质上存储的软件。处理器可以利用一个或多个通用处理器和/或专用处理器来实现。示例包括微处理器、微控制器、DSP处理器和可以执行软件的其他电路。软件应被广泛地解释为表示指令、数据或它们的任何组合,无论是被称为软件、固件、中间件、微代码、硬件描述语言还是其他。以举例的方式,机器可读介质可包括随机存取存储器(RAM)、闪存存储器、只读存储器(ROM)、可编程只读存储器(PROM)、可擦除可编程只读存储器(EPROM)、电可擦除可编程只读存储器(EEPROM)、寄存器、磁盘、光盘、硬盘驱动器、或任何其他合适的存储介质、或它们的任何组合。机器可读介质能够以计算机程序产品来体现。计算机程序产品可以包括包装材料。

[0135] 在硬件具体实施中,机器可读介质可以是与处理器分开的处理系统的一部分。然而,如本领域技术人员将容易理解的,机器可读介质或其任何部分可在处理系统外部。以举例的方式,机器可读介质可包括传输线、由数据调制的载波、和/或与设备分开的计算机产品,所有这些都可由处理器通过总线接口来访问。另选地或此外,机器可读介质或其任何部分可集成到处理器中,诸如在具有高速缓存和/或通用寄存器堆的情况下。虽然所讨论的各种组件可被描述为具有特定位置,诸如本地组件,但它们也可按各种方式来配置,诸如某些组件被配置为分布式计算系统的一部分。

[0136] 可以将处理系统配置为具有提供处理器功能性的一个或多个微处理器和提供机器可读介质的至少一部分的外部存储器的通用处理系统,所有这些组件通过外部总线架构与其他支持电路系统链接在一起。另选地,该处理系统可以包括一个或多个神经形态处理器以用于实现所描述的神经元模型和神经系统模型。作为另一另选方案,处理系统可以用具有集成在单个芯片中的处理器、总线接口、用户接口、支持电路以及机器可读介质的至少一部分的专用集成电路(ASIC)来实现,或者用一个或多个现场可编程门阵列(FPGA)、可编程逻辑器件(PLD)、控制器、状态机、门控逻辑、分立硬件组件、或者任何其他合适的电路、或者能够执行本公开通篇所描述的各种功能性的电路的任何组合来实现。本领域技术人员将认识到取决于特定应用和施加在整个系统上的总体设计约束如何最好地实现处理系统的所述功能性。

[0137] 机器可读介质可以包括多个软件模块。这些软件模块包括当由处理器执行时使处理系统执行各种功能的指令。软件模块可以包括传输模块和接收模块。每个软件模块可以驻留在单个存储设备中或者跨多个存储设备分布。以举例的方式,当触发事件发生时,软件模块可以从硬盘驱动器加载到RAM中。在软件模块的执行期间,处理器可以将指令中的一些指令加载到高速缓存中以提高访问速度。接着可将一个或多个高速缓存线加载到通用寄存器堆中以供处理器执行。当在下文提及软件模块的功能性时,将理解的是,此类功能性由处理器在执行来自该软件模块的指令时实现。此外,应当理解,本公开的各方面产生对处理器、计算机、机器或实现此类方面的其他系统的功能的改进。

[0138] 如果在软件中实现,则各功能可作为一个或多个指令或代码存储在计算机可读介质上或通过其进行传输。计算机可读介质包括计算机存储介质和通信介质两者,通信介质

包括促进计算机程序从一地到另一地传递的任何介质。存储介质可以是可由计算机访问的任何可用介质。以举例而非限制的方式,此类计算机可读介质可包括RAM、ROM、EEPROM、CD-ROM或其他光盘存储装置、磁盘存储装置或其他磁存储设备、或者可用于以指令或数据结构的形式携带或存储期望程序代码并且可由计算机访问的任何其他介质。附加地,任何连接也被适当地称为计算机可读介质。例如,如果软件是使用同轴电缆、光缆、双绞线、数字用户线(DSL)、或诸如红外(IR)、无线电、以及微波的无线技术从网站、服务器、或其他远程源发送的,则该同轴电缆、光缆、双绞线、DSL、或诸如红外、无线电、以及微波的无线技术被包括在介质的定义中。如本文所用,磁盘和光盘包括压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软盘和蓝光®光盘,其中磁盘通常磁性地复制数据,而光盘则用激光来光学地复制数据。因此,在一些方面,计算机可读介质可以包括非暂态计算机可读介质(例如,有形介质)。此外,对于其他方面,计算机可读介质可以包括暂态计算机可读介质(例如,信号)。上述的组合应也被包括在计算机可读介质的范围内。

[0139] 因此,特定方面可以包括用于执行本文给出的操作的计算机程序产品。例如,这种计算机程序产品可包括其上存储(和/或编码)有指令的计算机可读介质,这些指令可由一个或多个处理器执行以执行所描述的操作。对于某些方面,计算机程序产品可以包括包装材料。

[0140] 另外,应当理解,用于执行所描述的方法和技术的模块和/或其他适当的部件可由用户终端和/或基站在适用时下载和/或以其他方式获得。例如,这种设备可耦合到服务器以促进用于执行所描述的方法的部件的转移。另选地,所描述的各种方法可经由存储部件(例如,RAM、ROM、诸如压缩光盘(CD)或软盘的物理存储介质等)来提供,使得当将该存储部件耦合到或提供给用户终端和/或基站时,该设备可获得各种方法。此外,可利用适于向设备提供所描述的方法和技术的任何其他合适的技术。

[0141] 应理解,权利要求书不限于上文所例示的精确配置和组件。可在以上所描述的方法和装置的布局、操作和细节上作出各种改动、更换和变形而不会脱离权利要求的范围。

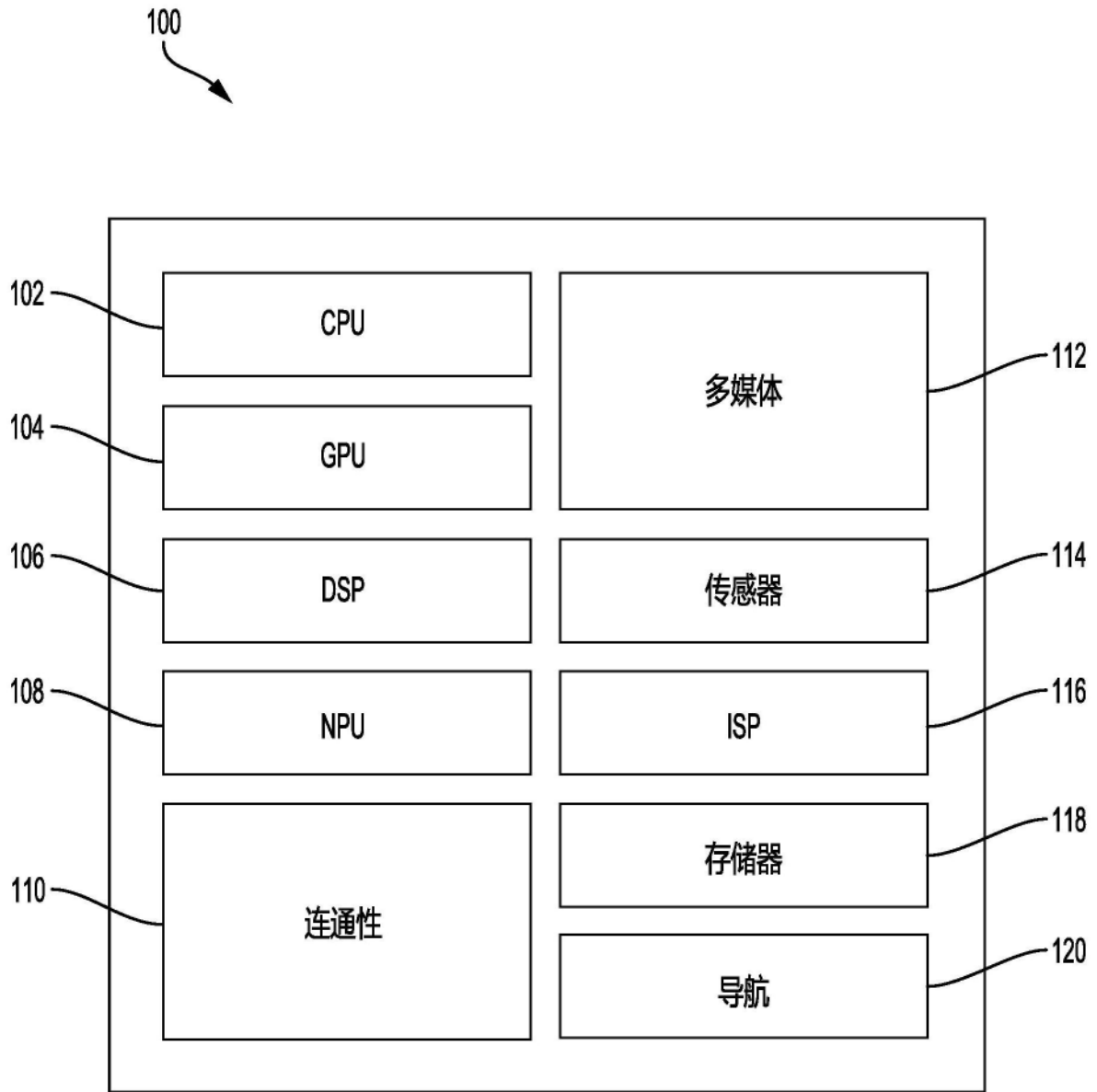


图1

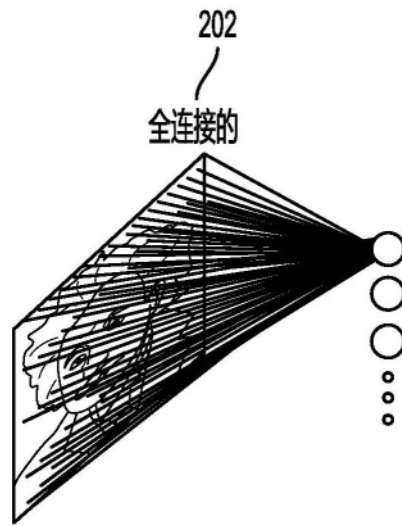


图2A

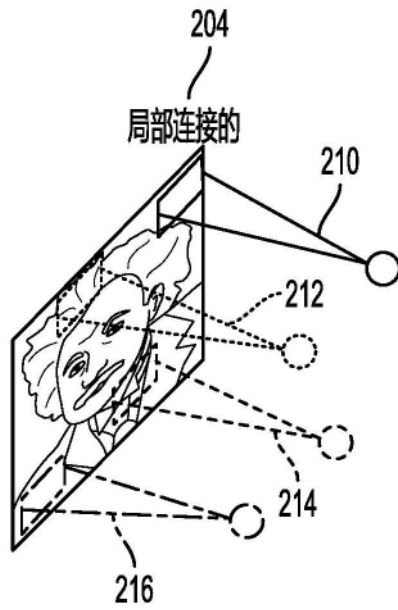


图2B

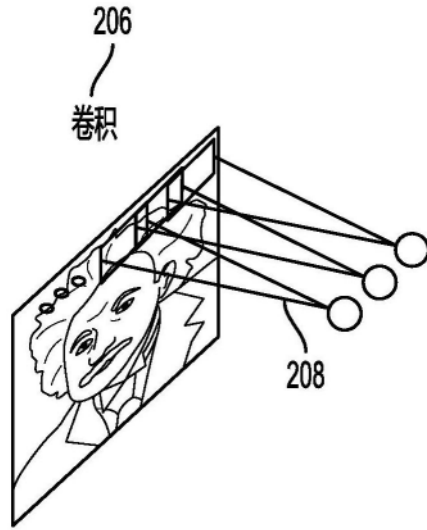


图2C

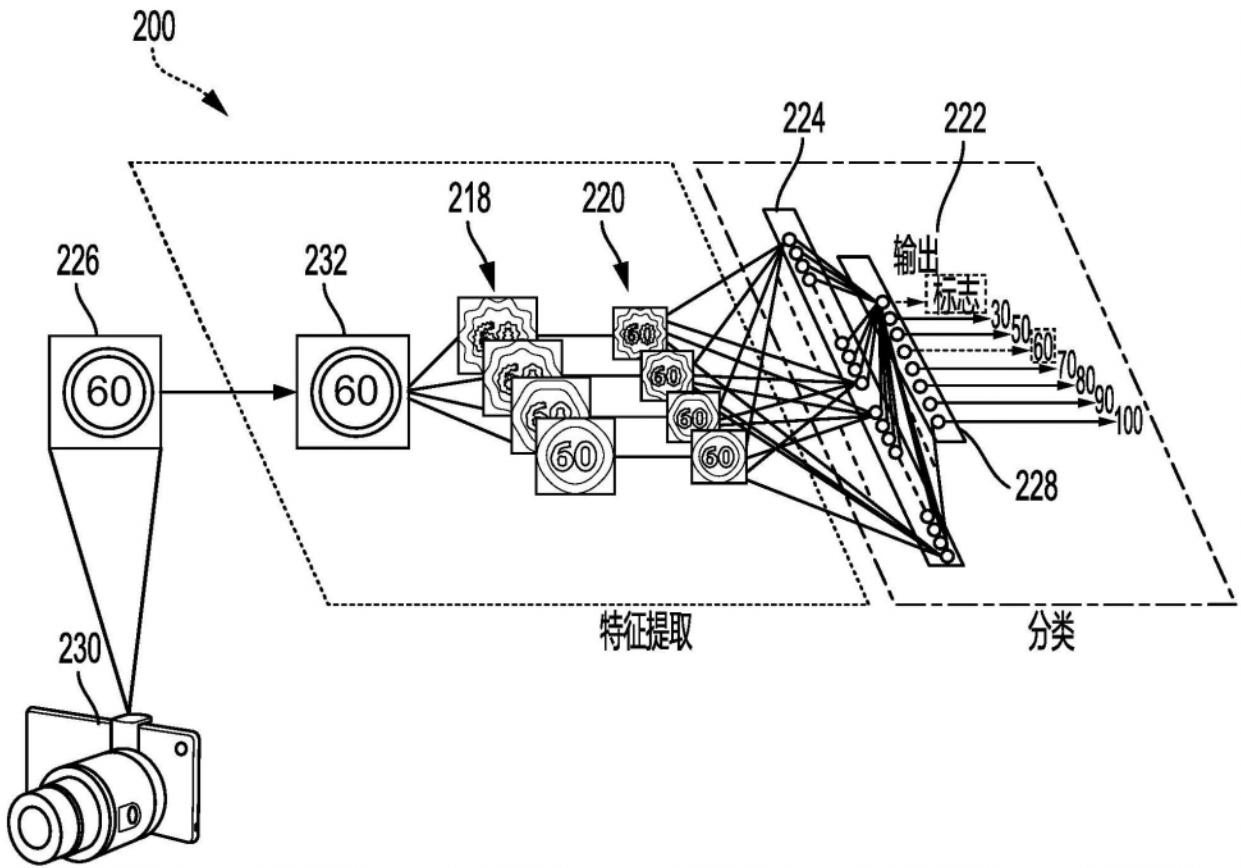


图2D

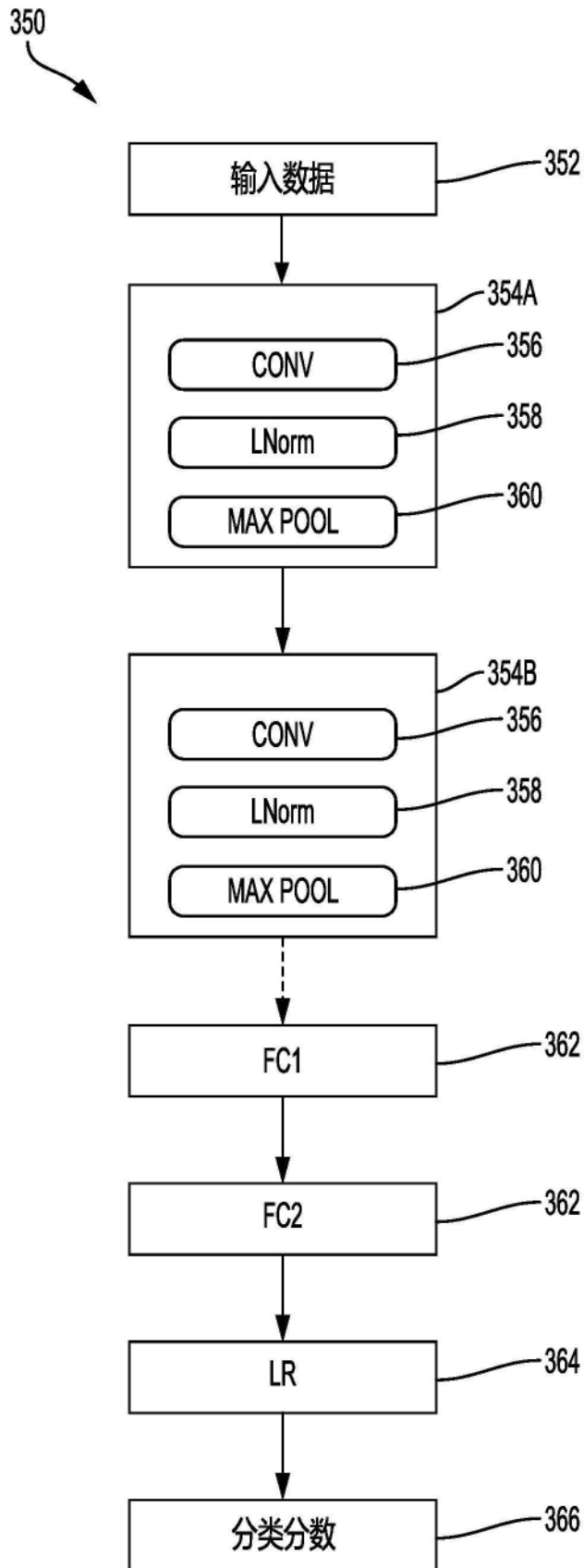
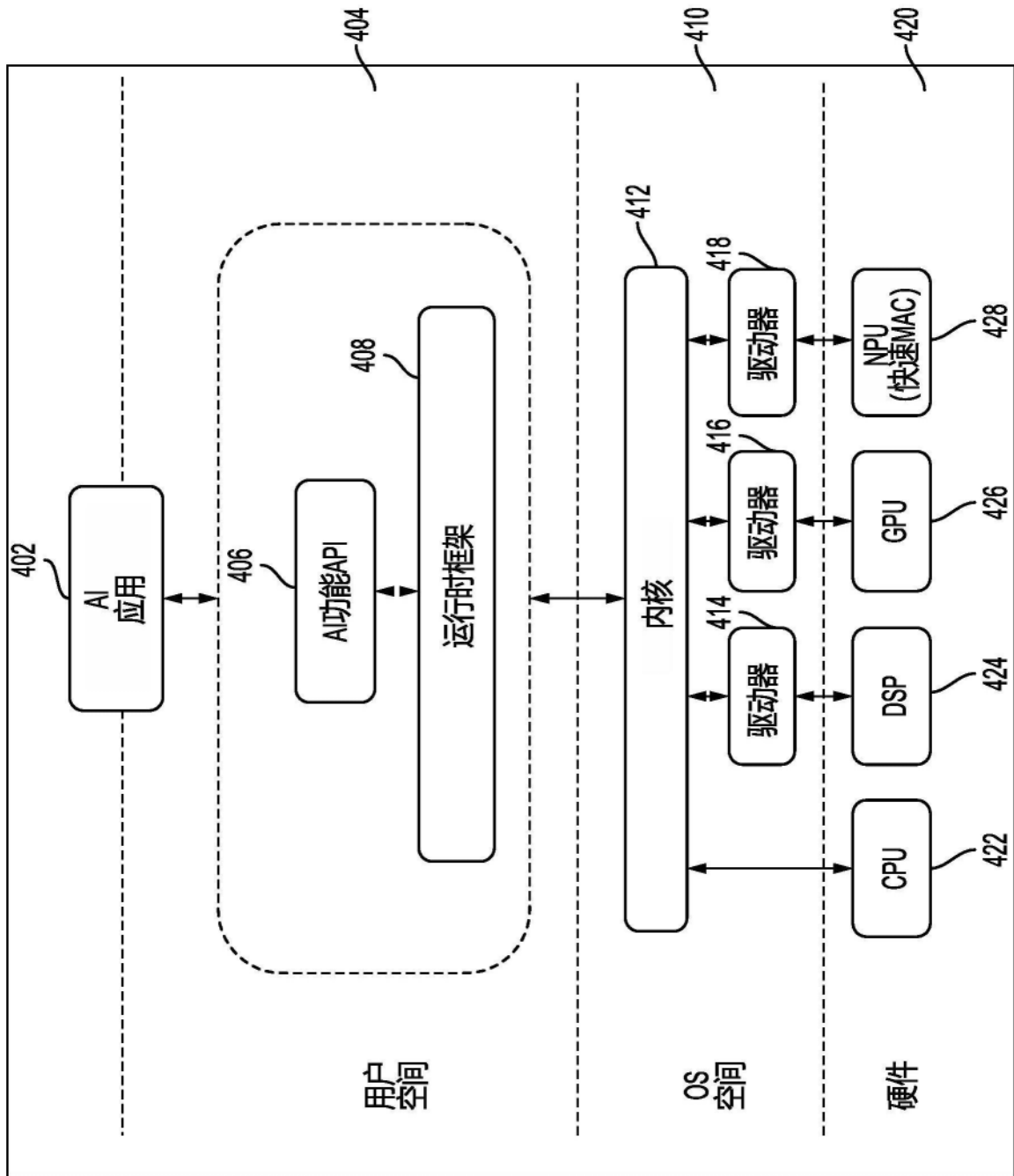


图3



400 ↗

图4

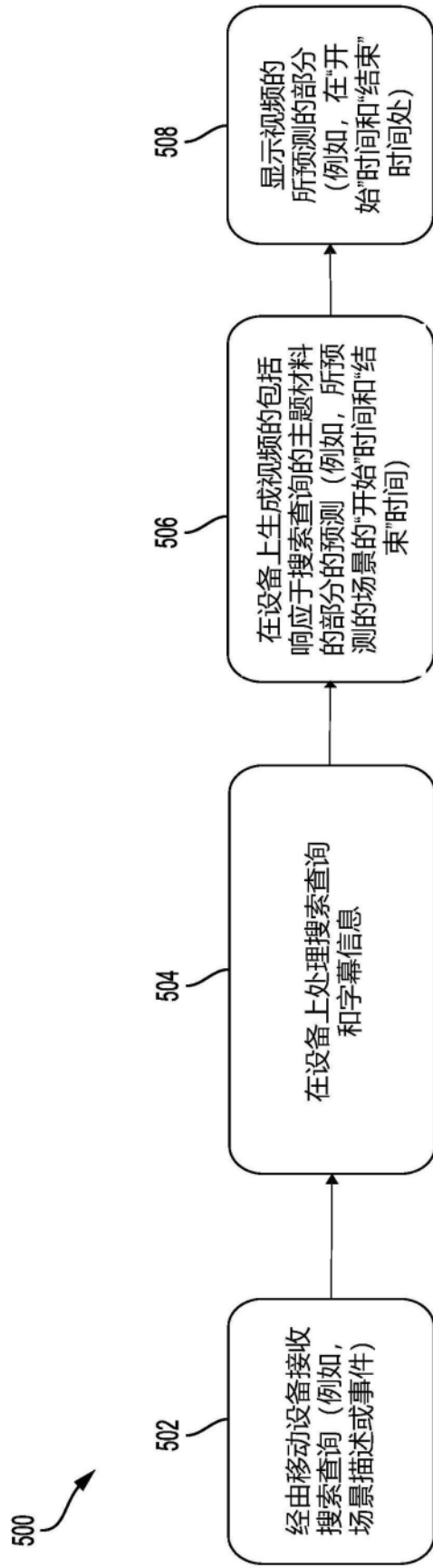


图5

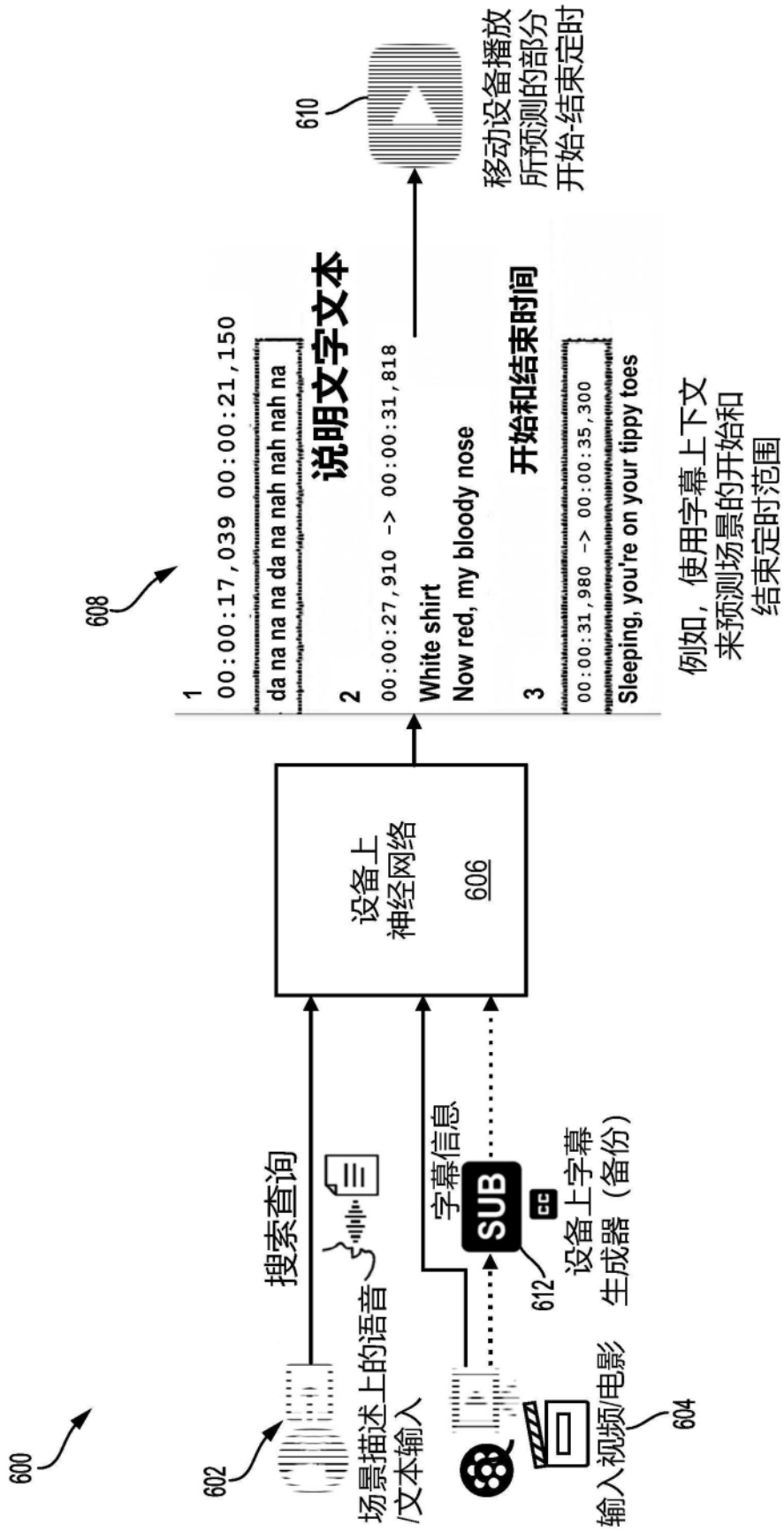


图6A

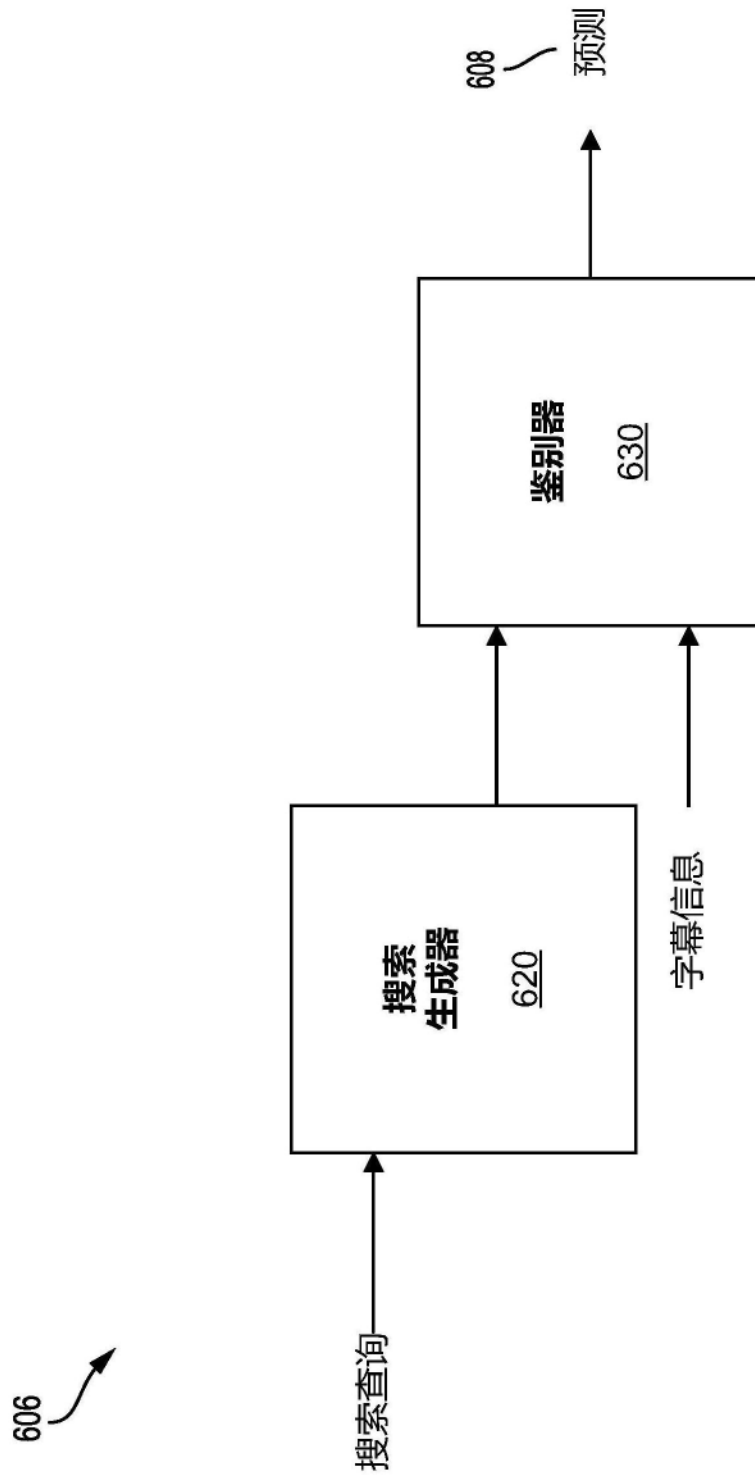


图6B

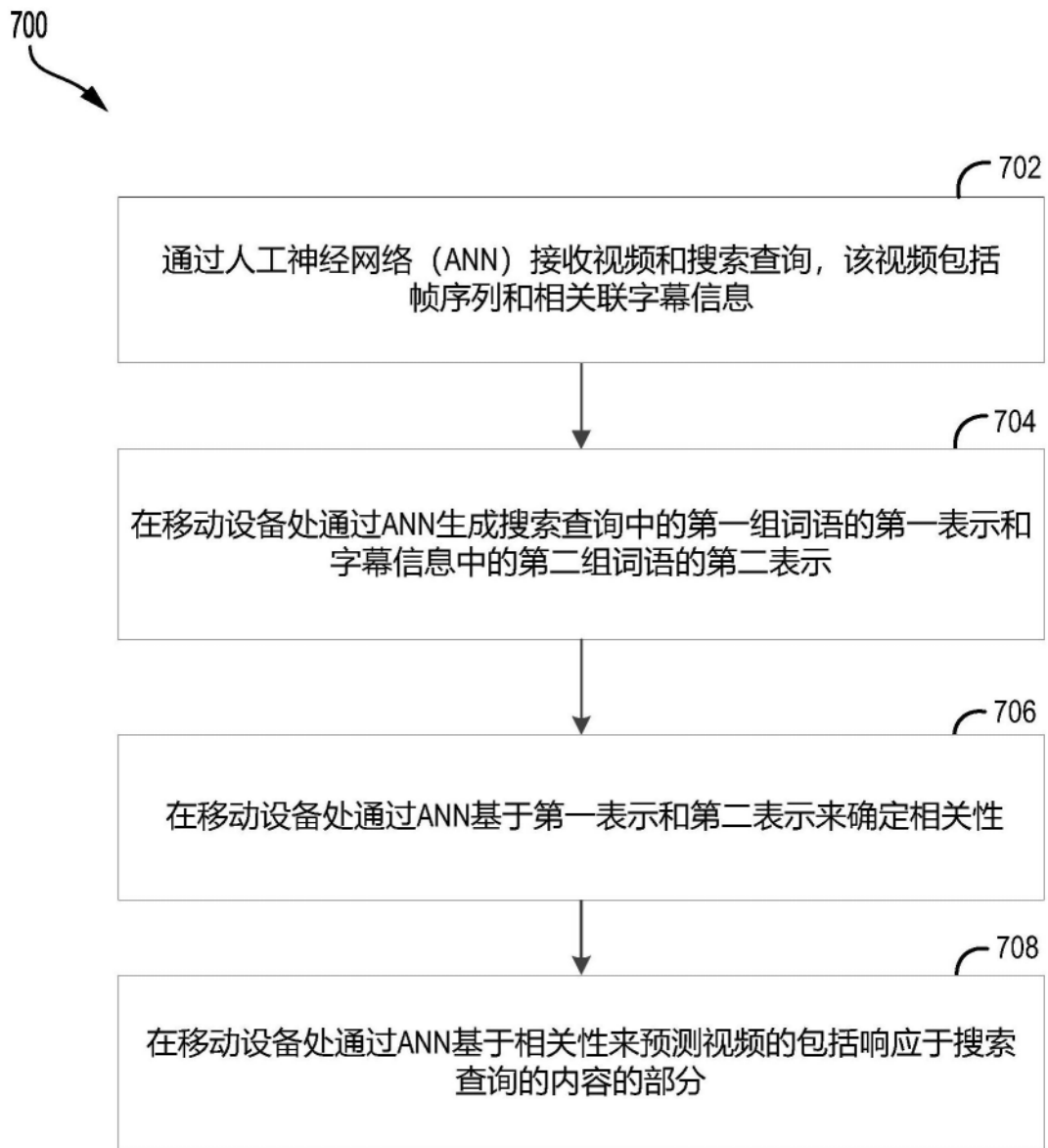


图7