



(12) 发明专利

(10) 授权公告号 CN 114080818 B

(45) 授权公告日 2024.07.16

(21) 申请号 202080011238.8

(22) 申请日 2020.06.22

(65) 同一申请的已公布的文献号
申请公布号 CN 114080818 A

(43) 申请公布日 2022.02.22

(85) PCT国际申请进入国家阶段日
2021.07.28

(86) PCT国际申请的申请数据
PCT/US2020/038963 2020.06.22

(87) PCT国际申请的公布数据
W02021/262137 EN 2021.12.30

(73) 专利权人 谷歌有限责任公司
地址 美国加利福尼亚州

(72) 发明人 N.J.弗雷 Z.孙

(74) 专利代理机构 北京市柳沈律师事务所
11105

专利代理师 金玉洁

(51) Int.Cl.

H04N 21/44 (2011.01)

H04N 21/4402 (2011.01)

H04N 21/435 (2011.01)

H04N 21/472 (2011.01)

(56) 对比文件

EP 3249651 A1, 2017.11.29

C. Luo.Video Summarization for Object Tracking in the Internet of Things.IEEE.2014, (第10.1109/NGMAST.2014.20期), 6.

C. Luo.Video Summarization for Object Tracking in the Internet of Things.IEEE.2014, (第10.1109/NGMAST.2014.20期), 6.

审查员 汤明皓

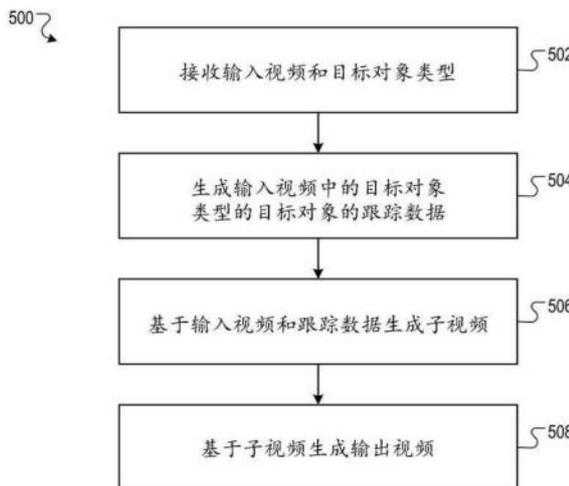
权利要求书2页 说明书14页 附图5页

(54) 发明名称

生成视频

(57) 摘要

用于生成视频的方法、系统和装置,包括在计算机存储介质上编码的计算机程序。在一方面,一种方法包括:接收:(i) 包括视频帧序列的输入视频,和(ii) 指示目标对象类型的的数据;处理输入视频以生成跟踪数据,跟踪数据识别和跟踪输入视频中的目标对象类型的目标对象的一个或多个实例的可视位置;基于输入视频和跟踪数据生成多个子视频,包括:对于每个子视频,生成子视频帧的相应序列,每个子视频帧从输入视频的相应视频帧中提取,以包括来自目标对象类型的识别目标对象中的给定目标对象的相应实例;以及生成包括多个子视频的输出视频。



1. 一种由一个或多个数据处理装置执行的方法,所述方法包括:
 - 接收:(i) 包括视频帧序列的输入视频,和(ii) 指示目标对象类型的数据;
 - 处理输入视频以生成跟踪数据,跟踪数据识别和跟踪输入视频中的目标对象类型的目标对象的一个或多个实例的可视位置;
 - 基于输入视频和跟踪数据生成多个子视频,包括:
 - 对于每个子视频,生成子视频帧的相应序列,每个子视频帧从输入视频的相应视频帧中提取,以包括来自目标对象类型的识别目标对象中的给定目标对象的相应实例,其中,子视频帧中的至少一个从输入视频的相应视频帧裁剪,以包括少于相应视频帧的所有内容;
 - 以及
 - 生成包括多个子视频的输出视频,
 - 其中,对于每个子视频,生成子视频帧的相应序列包括:
 - 基于跟踪数据在输入视频的多个视频帧中的每一个中确定包围与子视频相对应的给定目标对象的实例的相同边界框,包括确定相同边界框的尺寸和位置使得给定目标对象在多个视频帧中的每一个中被相同边界框包围;以及
 - 从输入视频的多个视频帧中的每一个裁剪相同边界框。
2. 根据权利要求1所述的方法,其中,处理输入视频以生成跟踪数据包括,对于输入视频中的目标对象类型的目标对象的每个实例:
 - 对于输入视频的多个视频帧中的每一个,确定包围视频帧中的目标对象的实例的相应边界框。
3. 根据权利要求1所述的方法,其中,第一子视频的子视频帧和第二子视频的子视频帧各自从输入视频的相同视频帧裁剪。
4. 根据权利要求1所述的方法,其中,生成包括多个子视频的输出视频包括:
 - 将多个子视频中的每个子视频分配到定义用于组合多个子视频的格式的视频模板中的相应时隙。
5. 根据权利要求4所述的方法,还包括:
 - 接收一个或多个附加数据元素;以及
 - 将每个附加数据元素分配给视频模板中的相应时隙。
6. 根据权利要求5所述的方法,其中,每个附加数据元素包括图像数据、文本数据或两者。
7. 根据权利要求1所述的方法,其中,输出视频同时显示多个子视频中的至少两个子视频。
8. 根据权利要求1所述的方法,还包括:
 - 接收对数字组件的请求;
 - 确定包括输出视频的数字组件响应于请求;以及
 - 提供数字组件,所述数字组件包括响应于将与搜索结果一起呈现或在第三方网页上呈现的请求的输出视频。
9. 根据权利要求1所述的方法,其中,接收指示目标对象类型的数据包括:
 - 接收指定关键词的数据;以及
 - 根据从关键词到预定义可能目标对象类型集合的映射,将关键词映射到目标对象类

型。

10. 一种用于生成视频的系统,所述系统包括一个或多个数据处理装置和存储指令的一个或多个存储设备,当由一个或多个数据处理装置执行指令时,使一个或多个数据处理装置执行包含任何前述权利要求所述方法的操作。

11. 一种计算机可读存储介质,存储指令,当由一个或多个数据处理装置执行指令时,使一个或多个数据处理装置执行包括权利要求1-9中任一权利要求所述的方法的操作。

生成视频

背景技术

[0001] 本说明书涉及生成视频。

[0002] 互联网促进了全球用户之间的信息交流。来自多个不同提供商的内容(包括视频)可以被集成到单个电子文档中,以创建复合文档。例如,在电子文档中包括的内容的一部分可以由电子文档的发布者选择(或指定)。内容的不同部分(例如,第三方内容,包括视频)可以由第三方(例如,不是电子文档的发布者的实体)提供,并被集成到电子文档中以形成包括来自多个不同来源的内容的复合文档。

发明内容

[0003] 本说明书描述了被实现为在一个或多个位置的一个或多个计算机上的计算机程序的系统,该系统可以处理输入视频以从输入视频中提取多个子视频,每个子视频显示目标对象类型的相应对象。系统可以组合显示目标对象类型的对象的子视频以生成“热门(topical)”视频(例如,其包括与给定主题相关的目标对象类型)。热门视频可以被集成到数字组件中,数字组件被发送以在用户设备与电子文档一起呈现,例如,在搜索结果旁边的块中或在第三方网站上。

[0004] 根据第一方面,提供了一种由一个或多个数据处理装置执行的方法,所述方法包括:接收:(i)包括视频帧序列的输入视频,以及(ii)指示目标对象类型的数据;处理输入视频以生成跟踪数据,跟踪数据识别和跟踪输入视频中目标对象类型的目标对象的一个或多个实例的可视位置;基于输入视频和跟踪数据生成多个子视频,包括:对于每个子视频,生成子视频帧的相应序列,每个子视频帧从输入视频的相应视频帧中提取,以包括来自目标对象类型的识别目标对象中的给定目标对象的相应实例,其中,子视频帧中的至少一个从输入视频的相应视频帧裁剪,以包括少于相应视频帧的所有内容;以及生成包括多个子视频的输出视频。

[0005] 在一些实现中,处理输入视频以生成跟踪数据包括,对于输入视频中的所述目标对象类型的目标对象的每个实例:对于输入视频的多个视频帧中的每一个,确定包围视频帧中的目标对象的实例的相应边界框。

[0006] 在一些实现中,对于每个子视频,生成子视频帧的相应序列包括:从输入视频的多个视频帧中的每一个裁剪包围与子视频相对应的给定目标对象的实例的相应边界框。

[0007] 在一些实现中,对于每个子视频,生成子视频帧的相应序列包括:基于跟踪数据确定包围与输入视频的多个视频帧中的每一个中的子视频相对应的给定目标对象的实例的相同边界框;以及从输入视频的多个视频帧中的每一个裁剪相同边界框。

[0008] 在一些实现中,第一子视频的子视频帧和第二子视频的子视频帧各自从输入视频的相同视频帧裁剪。

[0009] 在一些实现中,生成包括多个子视频的输出视频包括:将多个子视频中的每个子视频分配到定义用于组合多个子视频的格式的视频模板中的相应时隙。

[0010] 在一些实现中,所述方法还包括:接收一个或多个附加数据元素;以及将每个附加

数据元素分配给视频模板中的相应时隙。

[0011] 在一些实现中,每个附加数据元素包括图像数据、文本数据或两者。

[0012] 在一些实现中,输出视频同时显示多个子视频中的子视频中的至少两个。

[0013] 在一些实现中,所述方法还包括:接收对数字组件的请求;确定包括输出视频的数字组件响应于请求;以及提供数字组件,所述数字组件包括响应于将与搜索结果一起呈现或在第三方网页上呈现的请求的输出视频。

[0014] 在一些实现中,接收指示目标对象类型的数据包括:接收指定关键词的数据;以及根据从关键词到预定义可能目标对象类型集合的映射,将关键词映射到目标对象类型。

[0015] 根据另一方面,提供了一种系统,包括一个或多个数据处理装置和存储指令的一个或多个存储设备,当由一个或多个数据处理装置执行指令时,使一个或多个数据处理装置执行上述方法的操作。

[0016] 根据另一方面,提供了一种计算机可读存储介质,存储指令,当由一个或多个数据处理装置执行指令时,使一个或多个数据处理装置执行上述方法的操作。

[0017] 可以实现本说明书中描述的主题的特定实施例,以便实现以下一个或多个优点。

[0018] 本说明书描述了一种视频生成系统,该视频生成系统可以处理输入视频以提取多个子视频,每个子视频显示目标对象类型的相应对象,然后组合子视频以生成热门视频。视频生成系统可以组合子视频,例如,通过将子视频插入到可以同时显示多个子视频的视频模板中。热门视频可以有效地汇总与目标对象类型相关的输入视频的内容,同时潜在地具有比输入视频更短的持续时间、比输入视频占用更少的存储器空间和/或比输入视频需要更少的带宽在通信网络上传输(例如,互联网)。热门视频可以比输入视频占用更少的存储器空间和/或比输入视频需要更少的带宽在通信网络上传输,例如,因为(例如,构成热门视频的)每个子视频的每个子视频帧从输入视频的相应视频帧裁剪以包括少于视频帧的所有内容和/或因为每个子视频包括来自输入视频的视频帧的子集的内容。换句话说,每个子视频包括输入视频的内容的空间和/或时间子集。此外,该内容子集基于生成的跟踪数据,该跟踪数据识别和跟踪输入视频中的目标对象类型的目标对象的一个或多个实例的可视位置。这通过过滤输入视频的内容以识别和选择输入视频的目标部分来提供更有效的视频生成。代替将整个输入视频发送到分布式系统,可以生成包括输入视频的目标子集的热门视频并将其发送到特定系统。当接收到对在输入视频中包括的内容的请求时,这可以减少发送输入视频的内容所需的带宽。因此,视频生成系统能够更有效地使用计算资源,例如,存储器、带宽和计算能力。另外,通过组合子视频来汇总与目标对象类型相关的输入视频的内容,可以以更简洁的形式呈现来自输入视频的关于目标对象类型的信息,而不需要用户观看整个输入视频。这减少了向用户呈现关于目标对象类型的信息所需的网络带宽量,并且也减少了呈现关于目标对象类型的信息所需的时间量。

[0019] 在本说明书中描述的视频生成系统可以快速生成视觉上令人愉悦的热门视频(例如,在几分钟内),而只需要很少或不需要手动用户输入,例如,除了提供输入视频和指定目标对象类型之外。相比之下,手动生成热门视频可能昂贵且耗时,例如,需要数小时或数天,并且需要大量专业知识。因此,在本说明书中描述的视频生成系统能够通过自动化生成热门视频的任务来更有效地分配时间和资源。

[0020] 本说明书主题的一个或多个实施例的细节在附图和下面的描述中阐述。通过说明

书、附图和权利要求书,主题的其他特征、方面和优点将变得显而易见。

附图说明

[0021] 图1示出了示例视频生成系统。

[0022] 图2示出了示例视频模板的图示。

[0023] 图3示出了示例数字组件生成系统。

[0024] 图4是数字组件分发系统从数字组件数据库发送数字组件用于与电子文档一起呈现的示例环境的框图。

[0025] 图5是用于生成热门视频的示例处理的流程图。

[0026] 各种附图中类似的附图标记和名称表示类似的元件。

具体实施方式

[0027] 图1示出了示例视频生成系统100。视频生成系统100是被实现为在一个或多个位置中的一个或多个计算机上的计算机程序的系统的示例,其中实现下面描述的系统、组件和技术。

[0028] 系统100被配置为处理:(i) 输入视频102和(ii) 识别“目标”对象类型104的数据,以生成热门视频106。

[0029] 输入视频102可以包括视频帧序列,其中每个视频帧可以被表示为数值数组,并且可以与输入视频102中的相应时间点相关联。

[0030] 目标对象类型104可以是来自预定义可能对象类型集合的可能对象类型。可能对象类型集合可以包括任何适当的对象类型,例如,食品类型(例如,比萨饼、三明治等)、饮料类型(例如,葡萄酒、啤酒等)、服装类型(例如,连衣裙、裤子、男士衬衫、女士衬衫等)、配饰类型(例如,帽子、手提包、眼镜等)、电子设备类型(例如,计算机、智能手机等)、动物类型(例如,狗、猫等)、车辆类型(例如,卡车、轿车、运动型多用途车(SUV)等)、健康/美容类型(例如,口红、调色板、指甲油等)、运动物品类型(例如,橄榄球、足球、网球拍、游泳护目镜等)、家用电器类型(例如,洗碗机、冰箱等)、家具类型(例如,椅子、沙发、等)和/或人员类型(例如,男子、女子、儿童等)。在一些实现中,可能对象类型集合可以包括细粒度对象类型,例如,标识特定产品的对象类型,例如,具有特定品牌和型号的车辆。

[0031] 系统100处理输入视频102和目标对象类型104以生成热门视频106,该热门视频106包括从输入视频102提取(例如,裁剪)的多个“子视频”,其中每个子视频显示(例如,描绘)目标对象类型104的相应对象。也就是说,系统100生成热门视频106,该热门视频106包括显示目标对象类型104的对象的输入视频102的部分。

[0032] 系统100可以用于各种应用中的任何一种。例如,如将参考图3和图4更详细地描述的,系统100可以被用于生成集成到数字组件中的热门视频,所述数字组件被发送以在用户设备处与电子文档一起呈现,例如,在搜索结果旁边的块中或在第三方网站上。

[0033] 系统100包括跟踪引擎108、裁剪引擎110和合成引擎112,下面将更详细地描述这些引擎。

[0034] 跟踪引擎108被配置为处理输入视频102和识别目标对象类型104的数据,以生成跟踪数据114,跟踪数据114识别、标记和跟踪输入视频102中的目标对象类型104的对象的

一个或多个实例的可视位置。为了方便,在本说明书中,目标对象类型104的对象可以被称为“目标对象”。

[0035] 跟踪数据114可以包括定义一个或多个“目标对象轨迹”的数据,其中,每个目标对象轨迹对应于目标对象的相应实例,并且通过多个视频帧的序列中的每个视频帧来跟踪目标对象的位置。更具体地,对应于目标对象的目标对象轨迹可以例如通过指定与包围视频帧中的目标对象的每个视频帧相对应的相应边界框来指定目标对象在视频帧序列中的每个视频帧中的位置。可以例如通过边界框的顶点的坐标来定义与视频帧相对应的边界框,其中,坐标在视频帧的参考帧中表示。如在整个本说明书中使用的,视频帧中的“边界框”可以指定视频帧的具有任何适当形状的区域,例如,正方形、矩形或圆形。

[0036] 指定目标对象在视频帧中的位置的边界框可以包围少于所有视频帧,例如,使得该边界框排除视频帧的至少一部分(然而,在某些情况下,视频帧中的边界框可以包围整个视频帧)。此外,每个目标对象轨迹可以通过视频帧序列跟踪对应目标对象的位置,视频帧序列是输入视频的整个视频帧序列的适当子集。

[0037] 跟踪引擎108可以使用任何适当的对象检测和跟踪技术来生成跟踪数据114的目标对象轨迹。例如,跟踪引擎108可以使用在J.Henriques,R.Caseiro,P.Martins,J.Batista,“High-speed tracking with kernelized correlation filters,”IEEE Transactions on Pattern Analysis and Machine Intelligence,vol.37,no.3,pp.583-596(2014)中描述的对象检测和跟踪技术。作为另一示例,跟踪引擎108可以使用在N.Wojke,A.Bewley,D.Paulus,“Simple online and realtime tracking with a deep association metric,”2017IEEE International Conference on Image Processing (ICIP),pp.3645-3649(2017)中描述的对象检测和跟踪技术。

[0038] 在特定示例中,为了生成跟踪数据114,跟踪引擎108可以使用被配置为处理视频帧以生成对象检测输出的对象检测神经网络来处理输入视频102的每个视频帧。对象检测输出可以包括:(i)每个包围在视频帧中描绘的相应对象的一个或多个边界框,以及(ii)对于每个边界框,标识由边界框包围的对象的类型的标签。跟踪引擎108可以保持被标记为目标对象类型104的包围对象的边界框,同时丢弃(例如,避免进一步处理)任何其他边界框。跟踪引擎108可以通过确定第一帧中的第一边界框和第二(例如,后续)帧中的第二边界框之间的相似性度量超过阈值来确定第一边界框和第二边界框包围相同的对象。相似性度量可以基于:(i)第一边界框和第二边界框的内容之间的视觉相似性,以及(ii)第一边界框和第二边界框之间的重叠。响应于确定第一边界框和第二边界框包围相同的对象,跟踪引擎108可以确定第一边界框和第二边界框被包括在相同的目标对象轨迹中,例如,从而跟踪对象从第一帧到第二帧的位置。

[0039] 裁剪引擎110被配置为处理输入视频102和跟踪数据114以生成一个或多个子视频116。更具体地,裁剪引擎110可以生成与由跟踪数据114指定的每个目标对象轨迹相对应的相应子视频116。如在本说明书中使用的,“子视频”是指具有视频帧序列的视频,称为“子视频帧”,使得子视频的每个子视频帧从视频102的相应视频帧裁剪。从视频102的对应视频帧裁剪的子视频116的子视频帧可以包括少于所有对应视频帧,例如,使得子视频帧排除对应视频帧的至少一部分(然而,在一些情况下,子视频帧可以包括视频102的对应视频帧的整体)。在特定示例中,视频帧可以具有尺寸 $a \times b$,其中, a 是视频帧的宽度(例如,沿着视频帧

的宽度空间尺寸的像素数), b 是视频帧的高度(例如,沿着视频帧的高度空间尺寸的像素数)。在该示例中,从视频帧裁剪的子视频帧可以具有尺寸 $c \times d$,其中, c 小于 a 且 d 小于 b 。

[0040] 如整个说明书所使用的,视频帧(或子视频帧,或视频帧中的边界框)的“尺寸”可以是指指定以下的数据:(i)视频帧的宽度(例如,沿视频帧的宽度空间尺寸的像素数),和(ii)视频帧的高度(例如,沿视频帧的高度空间尺寸的像素数)。

[0041] 裁剪引擎110可以以各种方式生成对应于目标对象的目标对象轨迹的子视频116。接下来更详细地描述裁剪引擎110的一些示例实现。

[0042] 在一个实现中,裁剪引擎110可以通过基于目标对象在视频帧中的位置提取视频帧的一部分,来从与目标对象轨迹相对应的每个视频帧中提取相应的子视频帧。例如,目标对象轨迹可以对每个视频帧指定包围视频帧中的目标对象的相应边界框,并且裁剪引擎110可以提取每个视频帧中被视频帧的边界框包围的部分。也就是说,裁剪引擎110可以通过提取由目标对象轨迹指定为包围视频帧中的目标对象的边界框包围的对应视频帧的部分来生成每个子视频帧。在该示例中,跟踪引擎108可以被配置为生成目标对象轨迹,该目标对象轨迹为每个视频帧指定具有固定尺寸(例如,宽度和高度)的边界框,例如,使得每个提取的子视频帧具有相同的尺寸。

[0043] 在另一个实现中,为了生成与目标对象轨迹相对应的子视频116,裁剪引擎110可以确定“超级”边界框的尺寸(例如,宽度和高度)和位置(例如,中心)。具体地,裁剪引擎110可以确定超级边界框的尺寸和位置,使得目标对象在与目标对象轨迹相对应的每个视频帧中被超级边界框包围。例如,目标对象轨迹可以为每个视频帧指定包围视频帧中的目标对象的相应的边界框,并且裁剪引擎110可以识别包围由目标对象轨迹指定的每个边界框的超级边界框。也就是说,裁剪引擎110可以识别超级边界框,该超级边界框包围由目标对象轨迹指定的边界框集合的并集(同时仍然包围少于所有的每个视频帧)。在确定超级边界框的尺寸和位置之后,裁剪引擎110可以通过提取由超级边界框包围的对应视频帧的部分来生成每个子视频帧。使用对于每个视频帧相同的超级边界框生成子视频116可以例如通过更好地保持原始视频102的电影效果(例如,平移和缩放)来提高子视频116的视觉质量。

[0044] 在一些情况下,由裁剪引擎110生成的某些子视频116可以是“时间上重叠的”,例如,这样的两个或更多子视频包括从输入视频102的相同视频帧提取的相应子视频帧。例如,当跟踪引擎108通过视频中的相同视频帧跟踪目标对象的两个不同实例时,这可能发生。

[0045] 合成引擎112被配置为处理子视频116和视频模板118以生成热门视频106。视频模板118定义用于将多个子视频116组合成包括每个子视频的复合视频(例如,热门视频106)的格式。例如,视频模板可以将二维(2-D)区域(例如,具有正方形或矩形形状)的划分定义为一个或多个子区域,称为“时隙”。合成引擎112通过将每个子视频116分配给视频模板118中的相应时隙来生成热门视频106。视频模板118可以在热门视频106的持续时间内动态地改变,例如,在热门视频的持续时间内可以移除时隙、可以添加时隙、或者时隙的大小可以改变。

[0046] 参考图2示出了示例视频模板200。视频模板200包括三个时隙,即时隙202、时隙204和时隙206。合成引擎112可以通过向视频模板200的每个时隙分配相应的子视频116来生成热门视频106,例如,使得每个时隙中的子视频被同时显示。在参考图2所示的示例中,

目标对象类型可以是“车辆”，并且每个时隙可以被分配示出车辆实例的相应子视频。

[0047] 在另一示例中，可以定义视频模板118，使得每个子视频116例如一个接一个地连续显示，而不是同时显示任何子视频。

[0048] 作为生成热门视频106的一部分，合成引擎112可以在将子视频116分配给视频模板118中的时隙之前修改子视频116。例如，合成引擎112可以修改（例如，减小）子视频的子视频帧的尺寸（例如，宽度和/或高度），例如，使得子视频的尺寸与视频模板118中对应时隙的尺寸匹配。作为另一示例，合成引擎112可以减少子视频的持续时间（例如，通过移除子视频的一个或多个子视频帧）以使子视频的持续时间与时隙的持续时间匹配。时隙的持续时间可以是指时隙被包括在模板中的时间长度，例如，在从模板中移除之前。

[0049] 合成引擎112可以以各种方式中的任何一个来确定子视频116到视频模板118的时隙的分配。在一个示例中，合成引擎112可以将子视频116随机分配给视频模板118的相应时隙。在另一示例中，合成引擎112可以将每个子视频116分配给具有与子视频116的尺寸最相似的尺寸（例如，从所有时隙中）的对应时隙，例如，使应用于子视频的修改最小化，以使其适合分配的时隙。

[0050] 在一些情况下，在将子视频116分配给视频模板118中的时隙之前，合成引擎112可以确定每个子视频的相应分数，其中，子视频的分数表征子视频的质量。合成引擎112可以基于一个或多个标准——例如，子视频的锐度和子视频中的运动量——来确定每个子视频的分数。可以使用任何适当的锐度量来计算子视频的锐度，例如，通过在用拉普拉斯核卷积子视频帧之后计算子视频帧中的最大像素强度值。可以例如通过计算与子视频帧相对应的光流帧集合的平均值来计算子视频中的运动量。合成引擎112可以基于例如表征子视频的锐度的锐度分数和表征子视频中的运动量的运动分数的线性组合来生成子视频的总分数。合成引擎112可以为更清晰并且包括更多运动的子视频生成更高的分数。

[0051] 合成引擎112可以使用子视频116的分数来确定哪些子视频116应该被包括在热门视频106中。例如，合成引擎112可以确定只有具有最高分数的子视频116的适当子集应该被包括在热门视频106中。作为另一示例，合成引擎112可以确定具有不满足预定阈值（例如，通过超过阈值）的分数的任何子视频116不应被包括在热门视频106中。

[0052] 在一些实现中，系统100可以接收要包括在热门视频106中的一个或多个数据元素120，例如，除了从输入视频102中提取的子视频116之外。数据元素120可以包括，例如，图像或文本的部分。合成引擎112可以通过将数据元素分配给视频模板118的相应时隙来将数据元素120包括在热门视频106中。

[0053] 由视频生成系统100生成的热门视频106可以被用于各种应用中的任何一种。下面将更详细地描述几个示例应用。

[0054] 在一个示例中，由视频生成系统100生成的热门视频106可以包括在响应于数字组件请求而提供的数字组件中，例如，如参考图3和图4更详细地描述的。

[0055] 在另一示例中，视频生成系统100可以处理：(i) 例如由一个或多个监视相机（安全相机）生成的输入视频，以及 (ii) 由用户指定的目标对象类型104，以生成对应热门视频106。由监视相机生成的输入视频可以是长的，例如，具有数小时或数天的持续时间。视频生成系统100可以使得能够快速处理由监视相机生成的视频，以生成显示用户感兴趣的对象（例如，宠物、人或车辆）的热门视频。监视相机可以是指放置在给定位置（例如，家庭前门旁

边)并连续捕获给定位置附近区域的视频的相机。

[0056] 在另一示例中,视频生成系统100可以处理:(i)由用户设备、例如智能手机的摄像机生成的输入视频,以及(ii)由用户指定的目标对象类型104,以生成对应热门视频106。

[0057] 图3示出了示例性数字组件生成系统300。数字组件生成系统300是被实现为一个或多个位置的一个或多个计算机上的计算机程序的系统的示例,其中实现下面描述的系统、组件和技术。

[0058] 数字组件生成系统300被配置为从内容提供商接收一个或多个关键词302和输入视频102,并且生成一个或多个数字组件306,其包括与相应关键词302相关的热门视频106。内容提供商可以是,例如发布者、广告商或其他内容源。

[0059] 如在整个本说明书中使用的,短语数字组件是指数字内容或数字信息的离散单元,其可以包括例如图像、视频剪辑(例如,热门视频)、音频剪辑、多媒体剪辑、文本段或统一资源定位符(URL)中的一个或多个。数字组件可以作为单个文件或文件集合以电子方式存储在物理存储器设备中,且数字组件可以采取视频文件、音频文件、多媒体文件、图像文件或文本文件的形式,并且包括流视频、流音频、社交网络帖子、博客帖子和/或广告信息,使得广告是一种数字组件。可以使用来自一个或多个源的数据(例如,天气信息、实时事件信息或从其他源获得的其他信息)来增强数字组件。

[0060] 系统300包括关键词映射引擎308和视频生成系统100(如参考图1所述)。此外,系统300可以与使得内容提供商能够与系统300交互的用户界面(例如,图形用户界面或任何其他适当种类的用户界面)相关联,如下面将更详细地描述的。

[0061] 关键词映射引擎308被配置为例如根据从关键词到可能对象类型的预定义映射,将每个关键词302映射到可能对象类型集合的对应目标对象类型104。例如,关键词映射引擎308可以将关键词“spectacles”、“specs”、“reading glasses”、“sunglasses”和“eyeglasses”中的每一个映射到可能对象类型“glasses”。可能对象类型集合可以是视频生成系统100的跟踪引擎被训练(或以其他方式配置)以在视频中识别和跟踪的对象类型的集合。在一些实现中,系统300可以向内容提供商提供与关键词302相对应的目标对象类型104的指示,并请求内容提供商“批准”(例如,执行动作以接受)每个目标对象类型104。

[0062] 系统300可以使用视频生成系统100生成与每个目标对象类型104相对应的一个或多个相应热门视频106,例如,通过使用视频生成系统100处理目标对象类型104和输入视频102。视频生成系统100可以生成与单个目标对象类型104相对应的多个热门视频106,例如,通过使用来自可能视频模板310集合的不同视频模板生成热门视频106,通过使用从输入视频提取的子视频的不同组合来生成热门视频106,通过生成输入视频102的子视频被分配给视频模板中的不同间隙组合的热门视频106,或者通过在热门视频106上覆盖不同的音频声道。系统300可以使内容提供商能够指定(例如,通过用户界面)控制热门视频106的生成的各种因素,例如,视频生成系统100应当使用哪些视频模板310和/或音频声道来生成热门视频106。

[0063] 除了关键词302和输入视频102之外,内容提供商也可以向系统300提供一个或多个其他数据元素(例如,图像或文本的一部分)以被包括在热门视频106中。视频生成系统100可以将提供的数据元素包括在热门视频中,例如,如参考图1所述。

[0064] 系统300可以使得内容提供商能够选择要包括在相应数字组件306中的一个或多

个热门视频106。每个数字组件306可以包括热门视频106和可能由内容提供商提供的其他数据元素(例如,视频、图像、文本部分、URL)。

[0065] 内容提供商可以将系统300生成的数字组件306提供给数字组件分发系统,例如,参考图4描述数字组件分发系统410。数字组件分发系统可以响应于在用户设备上将数字组件与电子文档一起呈现(例如,在搜索结果旁边的块中或在第三方网站上)的请求来发送数字组件306。

[0066] 在说明系统300的使用的特定示例中,内容提供商可以向系统300提供输入视频102,该输入视频102包括时装秀的几个小时的镜头和关键词302“handbag”。系统300可以处理输入视频102和关键词302以生成对应的热门视频106集合,每个热门视频106包括显示手提包的实例的输入视频的多个子视频,其中,子视频被自动编辑成视觉上令人愉悦的视频模板,并且省略输入视频的部分,使得填充的视频模板具有比输入视频102的持续时间(例如,时装秀的几个小时的镜头)更短的持续时间(例如,30秒、1分钟、2分钟或其他持续时间)。然后,内容提供商可以选择要包括在提供给数字组件分发系统的数字组件中的一个或多个热门视频106。

[0067] 图4是示例环境400的框图,其中数字组件分发系统410从数字组件数据库416发送数字组件以与电子文档一起呈现。内容提供商可以使用数字组件生成系统300(参考图3描述)生成数字组件,并将生成的数字组件提供给数字组件分发系统410。

[0068] 示例环境400包括网络402,诸如局域网(LAN)、广域网(WAN)、互联网或其组合。网络402连接电子文档服务器404、客户端设备406、数字组件服务器408和数字组件分发系统410(也被称为“分发系统”410)。示例环境400可以包括许多不同的电子文档服务器404、客户端设备406和数字组件服务器408。

[0069] 客户端设备406是能够通过网络402请求和接收资源的电子设备。示例客户端设备406包括个人计算机、移动通信设备(例如,移动电话),以及可以通过网络402发送和接收数据的其他设备。客户端设备406通常包括用户应用,诸如web浏览器,以便于通过网络402发送和接收数据,但是,由客户端设备406执行的本原应用也可以促进通过网络402发送和接收数据。

[0070] 电子文档是在客户端设备406呈现内容集合的数据。电子文档的示例包括网页、字处理文档、可移植文档格式(PDF)文档、图像、视频、搜索结果页和提要源。本原应用(例如,“app”),诸如安装在移动设备、平板电脑或桌面计算设备上的应用也是电子文档的示例。电子文档可以由电子文档服务器404(“Electronic Doc Server”)提供给客户端设备406。例如,电子文档服务器404可以包括托管发布者网站的服务器。在该示例中,客户端设备406可以发起对给定发布者网页的请求,并且托管给定发布者网页的电子服务器404可以通过发送在客户端设备406处发起给定网页的呈现的机器可执行指令来响应该请求。

[0071] 在另一示例中,电子文档服务器404可以包括应用服务器,客户端设备406可以从其下载应用。在该示例中,客户端设备406可以下载在客户端设备406处安装应用所需的文件,然后在本地执行下载的应用。

[0072] 电子文档可以包含各种内容。例如,电子文档可以包括静态内容(例如,文本或其他指定内容),这些内容在电子文档本身中并且/或者不随时间而改变。电子文档也可以包括动态内容,这些内容可能会随着时间的推移或根据请求而变化。例如,给定电子文档的发

布者可以维护用于填充电子文档部分的数据源。在该示例中,给定电子文档可以包括一个或多个标签或脚本,当客户端设备406处理给定电子文档(例如,渲染或执行)时,这些标签或脚本使客户端设备406从数据源请求内容。客户端设备406将从数据源获得的内容整合到给定电子文档中以创建包括从数据源获得的内容的复合电子文档。

[0073] 在某些情况下,给定电子文档可以包括引用数字组件分发系统410的一个或多个数字组件标签或数字组件脚本。在这些情况下,当客户端设备406处理给定电子文档时,客户端设备406执行数字组件标签或数字组件脚本。数字组件标签或数字组件脚本的执行将客户端设备406配置为生成对一个或多个数字组件412的请求(称为“组件请求”),其通过网络402被发送到数字组件分配系统410。例如,数字组件标签或数字组件脚本可以使得客户端设备406能够生成包括报头和有效载荷数据的分组化数据请求。组件请求412可以包括指定特征的事件数据,诸如从其请求数字组件的服务器的名称(或网络位置)、请求设备(例如,客户端设备406)的名称(或网络位置)和/或数字组件分发系统410可以用于选择响应于请求而提供的一个或多个数字组件的信息。组件请求412由客户端设备406通过网络402(例如,电信网络)发送到数字组件分发系统410的服务器。

[0074] 组件请求412可以包括指定其他事件特征的事件数据,诸如被请求的电子文档和可以呈现数字组件的电子文档的位置的特性。例如,可以向数字组件分发系统410提供指定对数字组件将被呈现的电子文档(例如,网页)的引用(例如URL)的事件数据、可用于呈现数字组件的电子文档的可用位置、可用位置的大小和/或有资格在这些位置呈现的媒体类型。类似地,指定与电子文档相关联的关键词(“文档关键词”)或电子文档引用的实体(例如,人、地点或事物)的事件数据也可以被包括在组件请求412中(例如,作为有效载荷数据),并提供给数字组件分发系统410,以便于识别有资格与电子文档一起呈现的数字组件。事件数据也可以包括从客户端设备406提交以获得搜索结果页的搜索查询,和/或指定搜索结果和/或包括在搜索结果中的文本、声音或其他视觉内容的数据。

[0075] 组件请求412也可以包括与其他信息相关的事件数据,诸如客户端设备的用户已经提供的信息、指示提交组件请求的状态或区域的地理信息或为提供关于将显示数字组件的环境的情境的其他信息(例如,组件请求的时间、组件请求的星期几、将显示数字组件的设备的类型,诸如移动设备或平板设备)。组件请求412可以例如通过分组化网络被发送,并且组件请求412本身可以被格式化为具有报头和有效负载数据的分组化数据。报头可以指定分组的目的地,并且有效负载数据可以包括上面讨论的任何信息。

[0076] 组件分发系统410响应于接收组件请求412和/或使用包括在组件请求412中的信息来选择将与给定电子文档一起呈现的数字组件。在一些实现中,在不到一秒的时间内(使用本文所描述的技术)选择数字组件以避免可能由数字组件的延迟选择引起的错误。例如,响应于组件请求412而提供数字组件的延迟可能导致客户端设备406处的页面加载错误,或导致即使在电子文档的其他部分在客户端设备406处呈现之后,电子文档的部分仍保持未填充状态。另外,当向客户端设备406提供数字组件的延迟增加时,当数字组件被传递到客户端设备406时,电子文档很可能不再在客户端设备406处呈现,从而对用户使用电子文档的体验产生负面影响。此外,例如,如果当提供数字组件时电子文档不再在客户端设备406处呈现,则在提供数字组件时的延迟可能导致数字组件的失败传递。

[0077] 在一些实现方式中,数字组件分发系统410在分布式计算系统中被实现,该分布式

计算系统包括例如服务器和多个计算设备414的集合,这些计算设备相互连接并响应于请求412识别和分发数字组件。多个计算设备414的集合一起操作,以从数百万个可用数字组件(DC1-x)的语料库中识别数字组件的集合,这些数字组件有资格在电子文档中被呈现。例如,可以在数字组件数据库416中对数百万个可用数字组件编索引。每个数字组件索引条目可以引用对应的数字组件和/或包括有助于(例如,条件或限制)对应数字组件的分配/传输的分发参数(DP1-DPx)。例如,分发参数可以通过要求组件请求包括与数字组件的分发参数中的一个匹配(例如,精确地或以具有一些预先指定的相似性水平)的至少一个准则来有助于数字组件的传输。

[0078] 在一些实现中,用于特定数字组件的分发参数可以包括必须匹配的分发关键词(例如,通过在组件请求412中指定的电子文档、文档关键词或术语),以便数字组件有资格呈现。换句话说,分发参数用于触发数字组件在网络402上的分发(例如,传输)。分发参数也可以要求组件请求412包括指定特定地理区域的信息(例如,国家或州)和/或指定组件请求412起源于特定类型的客户端设备(例如,移动设备或平板设备)的信息,以便数字组件有资格呈现。

[0079] 分发参数也可以指定用于评估用于分发/传输的数字组件(例如,在其他可用数字组件中)的资格的资格值(例如,排名分数、出价或一些其他指定值),例如,通过组件评估过程。在一些情况下,资格值可以指定数字组件的提供商响应于数字组件的传输而愿意提交的最大补偿量(例如,对于归因于数字组件的呈现的特定事件的每个实例,诸如用户与数字组件的交互)。

[0080] 有资格数字组件的标识可以被分割成多个任务417a-417c,然后在多个计算设备414的集合内的计算设备之间被分配。例如,集合414中的不同计算设备各自可以分析数字组件数据库416的不同部分,以识别具有与在组件请求412中包括的信息匹配的分发参数的各种数字组件。在一些实现中,集合414中的每个给定计算设备可以分析不同的数据尺寸(或尺寸集合),并且将分析的结果(Res 1-Res 3)418a-418c传递(例如,发送)回数字组件分发系统410。例如,由集合414中的每个计算设备提供的结果418a-418c可以识别响应于组件请求而有资格分发的数字组件的子集和/或具有特定分发参数的数字组件的子集。数字组件的子集的认识可以包括,例如,将事件数据与分发参数进行比较,以及识别具有匹配事件数据的至少一些特征的分发参数的数字组件子集。

[0081] 数字组件分发系统410聚合从多个计算设备414的集合接收的结果418a-418c,并使用与聚合结果相关联的信息来:(i)选择将响应于请求412而提供的一个或多个数字组件,以及(ii)确定一个或多个数字组件的传输要求。例如,数字组件分发系统410可以基于一个或多个组件评估过程的结果来选择获胜数字组件(一个或多个数字组件)的集合。进而,数字组件分发系统410可以通过网络402生成和发送应答数据420(例如,表示应答的数字数据),其使得客户端设备406能够将获胜数字组件的集合集成到给定电子文档中,使得在客户端设备406的显示器上一起呈现获胜数字组件的集合和电子文档的内容。

[0082] 在一些实现中,客户端设备406执行包括在应答数据420中的指令,其配置并使得客户端设备406能够从一个或多个数字组件服务器获得获胜数字组件的集合。例如,应答数据420中的指令可以包括网络位置(例如,统一资源定位符(URL))和使客户端设备406向数字组件服务器408发送服务器请求(SR)421以从数字组件服务器408获得给定获胜数字组件

的脚本。响应于该请求,数字组件服务器408将识别在服务器请求421中指定的给定获胜数字组件(例如,在存储多个数字组件的数据库中),并向客户端设备406发送在客户端设备406处在电子文档中呈现给定获胜数字组件的数字组件数据(DC数据)422。

[0083] 为了便于搜索电子文档,环境400可以包括搜索系统450,其通过对电子文档进行爬网和编索引来识别电子文档(例如,基于电子文档的爬网内容来编索引)。可以基于与数据相关联的电子文档来对关于电子文档的数据编索引。电子文档的索引副本和可选的高速缓存副本被存储在搜索索引452(例如,硬件存储设备(或多个))中。与电子文档相关联的数据是表示在电子文档中包括的内容和/或电子文档的元数据的数据。

[0084] 客户端设备406可以通过网络402向搜索系统450提交搜索查询。作为响应,搜索系统450访问搜索索引452以识别与搜索查询相关的电子文档。搜索系统450以搜索结果的形式识别电子文档,并在搜索结果页中将搜索结果返回给客户端设备406。搜索结果是由搜索系统450生成的数据,该搜索系统450识别响应于特定搜索查询(例如,与其相关的)的电子文档,并且包括使客户端设备响应于与搜索结果的用户交互从指定的网络位置(例如,URL)请求数据的活动链接(例如,超文本链接)。示例搜索结果可以包括网页标题、从网页中提取的文本片段或图像的一部分以及网页的URL。另一示例搜索结果可以包括可下载应用的标题、描述可下载应用的文本片段、描述可下载应用的用户界面的图像和/或到可将应用下载到客户端设备406的位置的URL。在一些情况下,搜索系统450可以是应用商店(或在线门户)的一部分,或者与应用商店(或在线门户)交互,从应用商店(或在线门户)可以下载应用以在客户端设备406处安装,以便呈现关于与提交的搜索查询相关的可下载应用的信息。与其他电子文档一样,搜索结果页面可以包括其中可以呈现数字组件(例如,广告、视频剪辑、音频剪辑、图像或其他数字组件)的一个或多个时隙。

[0085] 为了选择要响应于组件请求而发送的数字组件,分发系统410可以识别有资格响应于组件请求而发送的数字组件的集合。然后,分发系统410可以选择一个或多个有资格的数字组件以通过例如拍卖过程来发送。在一些实现中,分发系统410通过根据有资格的数字组件的相应资格值对它们进行排名,并选择响应于组件请求要发送的一个或多个最高排名的数字组件来执行拍卖过程。

[0086] 图5是用于生成热门视频的示例过程500的流程图。为了方便,将过程500描述为由位于一个或多个位置的一个或多个计算机的系统执行。例如,根据本说明书适当编程的视频生成系统(例如,图1的视频生成系统100)可以执行处理500。

[0087] 系统接收:(i) 包括视频帧序列的输入视频,以及(ii) 指示目标对象类型的数据(502)。例如,系统可以接收指定关键词(例如,“sunglasses”)的数据,然后根据从关键词到预定义可能目标对象类型集合的映射,将接收的关键词映射到对应目标对象类型(例如,“glasses”)。可选地,系统也可以接收一个或多个附加数据元素,例如,图像、文本部分或两者。

[0088] 系统处理输入视频以生成跟踪数据,跟踪数据识别和跟踪输入视频中的目标对象类型的目标对象的一个或多个实例的可视位置(504)。例如,对于输入视频中的目标对象类型的目标对象的每个实例,系统可以确定在输入视频的多个视频帧中的每一个中包围目标对象的实例的相应边界框。系统可以使用任何适当的对象检测和跟踪技术来生成跟踪数据,例如,如参考图1所述。

[0089] 系统基于输入视频和跟踪数据生成多个子视频(506)。对于每个子视频,系统可以生成从输入视频的相应视频帧裁剪的相应子视频帧序列,以包括(i)小于相应视频帧的所有内容和(ii)来自目标对象类型的识别目标对象中的给定目标对象的相应实例。在一些实现中,代替将子视频帧中的每一个裁剪成包括少于相应视频帧的全部内容,从输入视频的至少一个相应视频帧裁剪至少一个子视频帧以包括少于至少一个相应视频帧的全部内容。在这样的实现中,子视频帧中的一个或多个可以包括相应视频帧的所有内容,例如,因为目标对象占据了那些相应视频帧中的所有相应视频帧或大部分相应视频帧。子视频帧可以包括相应视频帧的所有内容的子集,该子集小于或等于相应视频帧的所有内容。在一些实现中,系统可以通过从输入视频的多个视频帧中的每一个裁剪包围给定目标对象的实例的相应边界框来生成与给定目标对象的实例相对应的子视频帧的序列。包围给定目标对象实例的边界框可以由跟踪数据指定。在一些实现中,为了生成与给定目标对象的实例相对应的子视频的子视频帧序列,系统可以基于跟踪数据确定在输入视频的多个视频帧中的每一个中包围给定目标对象的实例的相同的边界框。然后,系统可以通过从输入视频的多个视频帧中的每一个裁剪相同的边界框来生成子视频帧的序列。

[0090] 在一些情况下,特定子视频可以是“时间上重叠的”,例如,这样的两个或更多子视频包括从输入视频的相同视频帧提取的相应子视频帧。例如,当系统通过输入视频中的相同视频帧跟踪目标对象的两个不同实例时,这可能发生。

[0091] 系统生成包括子视频的输出视频(508)。例如,为了生成输出视频,系统可以将每个子视频分配给定义用于组合子视频的格式的视频模板中的相应时隙。作为生成输出视频的一部分,系统可以确定每个子视频的相应分数(例如,基于子视频的锐度),并选择子视频的适当子集以包括在输出视频中(例如,具有最高分数的预定义数量的子视频)。在一些情况下,输出视频可以同时显示子视频中的至少两个,而在其他情况下,输出视频可以连续地——例如,一个接一个地——显示每个子视频。作为生成输出视频的一部分,系统可以将一个或多个附加数据元素(例如,文本或图像数据元素,如参考502所述)中的每一个分配给视频模板中的相应时隙。

[0092] 由系统生成的输出视频可以被包括在数字组件中,数字组件响应于数字组件请求而提供,并且随后与搜索结果一起被呈现或在第三方网页上呈现。

[0093] 本说明书在系统和计算机程序组件中使用术语“被配置为”。对于一个或多个被配置为执行特定操作或动作的计算机的系统,意味着系统已在其上安装了软件、固件、硬件或它们的组合,这些软件、固件、硬件或它们的组合在操作中导致系统执行操作或动作。对于将被配置为执行特定操作或动作的一个或多个计算机程序,意味着一个或多个程序包括指令,该指令在由数据处理装置执行时使该装置执行该操作或动作。

[0094] 本说明书中描述的主题和功能操作的实施例可以在数字电子电路中,在有形体现的计算机软件或固件中,在计算机硬件中实现,包括本说明书中公开的结构及其等同结构,或它们中的一个或多个的组合。本说明书中描述的主题的实施例可以被实现为一个或多个计算机程序,即,在有形非暂时性存储介质上编码的计算机程序指令的一个或多个模块,由数据处理装置执行或控制操作。计算机存储介质可以是机器可读存储设备、机器可读存储基板、随机或串行访问存储设备或它们中的一个或多个的组合。可选地或附加地,程序指令可以被编码在人工产生的传播信号上,例如机器产生的电、光或电磁信号,其被产生以对

信息进行编码用于传输到合适的接收器装置由数据处理装置执行。

[0095] 术语“数据处理装置”是指数据处理硬件,并且包括用于处理数据的各种装置、设备和机器,例如包括可编程处理器、计算机或多个处理器或计算机。装置还可以是或进一步包括专用逻辑电路,例如,FPGA(现场可编程门阵列)或ASIC(专用集成电路)。除了硬件之外,装置可以可选地包括为计算机程序创建执行环境的代码,例如,构成处理器固件、协议栈、数据库管理系统、操作系统或其一种或多种的组的代码。

[0096] 也可以被称为或描述为程序、软件、软件应用、app、模块、软件模块、脚本或代码的计算机程序可以以任何形式的编程语言来编写,包括编译或解释语言,声明性或程序性语言;它可以以任何形式进行部署,包括作为独立程序或作为模块、组件、子例程或适用于计算环境的其他单元进行部署。程序可以但不必对应于文件系统中的文件。程序可以存储在保存其他程序或数据的文件的一部分中,例如,存储在标记语言文档中的一个或多个脚本,专用于所讨论程序的单个文件或多个协调文件(例如,存储一个或多个模块、子程序或部分代码的文件。可以将计算机程序部署为在一个计算机上执行,或者在位于一个站点上或分发给多个站点上并通过数据通信网络互连的多个计算机上执行。

[0097] 在本说明书中,术语“引擎”广泛地用于指代被编程为执行一个或多个特定功能的基于软件的系统、子系统或处理。通常,引擎将被实现为安装在一个或多个位置的一个或多个计算机上的一个或多个软件模块或组件。在一些情况下,一个或多个计算机将专用于特定引擎;在其他情况下,可以在相同的一个或多个计算机上安装并运行多个引擎。

[0098] 本说明书中描述的处理和逻辑流程可以由执行一个或多个计算机程序以通过对输入数据进行操作并生成输出来执行功能的一个或多个可编程计算机来执行。处理和逻辑流程还可以通过专用逻辑电路(例如,FPGA或ASIC)或者通过专用逻辑电路与一个或多个编程计算机的组合来执行。

[0099] 适用于执行计算机程序的计算机可以基于通用或专用微处理器或两者,或者基于任何其他类型的中央处理器。通常,中央处理单元将从只读存储器或随机存取存储器或两者接收指令和数据。计算机的基本元件是用于进行或执行指令的中央处理单元以及用于存储指令和数据的一个或多个存储设备。中央处理单元和存储器可以由专用逻辑电路补充或并入专用逻辑电路中。通常,计算机还将包括或可操作地耦合以从一个或多个用于存储数据的大容量存储设备(例如,磁、磁光盘或光盘)接收数据或将数据传输到一个或多个大容量存储设备或两者。然而,计算机不必具有此类设备。此外,计算机可以嵌入到另一设备中,例如,移动电话、个人数字助理(PDA)、移动音频或视频播放器、游戏机、全球定位系统(GPS)接收器或便携式存储设备,例如,通用串行总线(USB)闪存驱动器,仅举几例。

[0100] 适用于存储计算机程序指令和数据的计算机可读介质包括所有形式的非易失性存储器、介质和存储设备,包括例如半导体存储设备,例如,EPROM、EEPROM和闪存设备;磁盘,例如,内部硬盘或可移动磁盘;磁光盘;以及CD-ROM和DVD-ROM磁盘。

[0101] 为了提供与用户的交互,可以在具有用于向用户显示信息的显示设备(例如,CRT(阴极射线管)或LCD(液晶显示器)监视器)以及键盘和指示设备(例如,鼠标或轨迹球,用户可通过其向计算机提供输入)的计算机上实现本说明书中描述的主题的实施例。其他类型的设备也可以用于提供与用户的交互,例如,提供给用户的反馈可以是任何形式的感觉反馈,例如,视觉反馈、听觉反馈或触觉反馈;并且可以以任何形式接收来自用户的输入,包括

声音、语音或触觉输入。另外,计算机可以通过向用户使用的设备发送文档和从用户使用的设备接收文档来与用户进行交互,例如,通过响应从网络浏览器收的请求,将网页发送到用户设备上的网络浏览器。此外,计算机可以通过将文本消息或其他形式的消息发送到个人设备(例如,运行消息收发应用的智能手机),并从用户接收响应消息来与用户交互。

[0102] 用于实现机器学习模型的数据处理设备还可以包括例如专用硬件加速器单元,用于处理机器学习训练或生产(即推理、工作负载)的公共和计算密集型部分。

[0103] 可以使用机器学习框架(例如,TensorFlow框架、Microsoft Cognitive Toolkit框架、Apache Singa框架或Apache MXNet框架)来实现和部署机器学习模型。

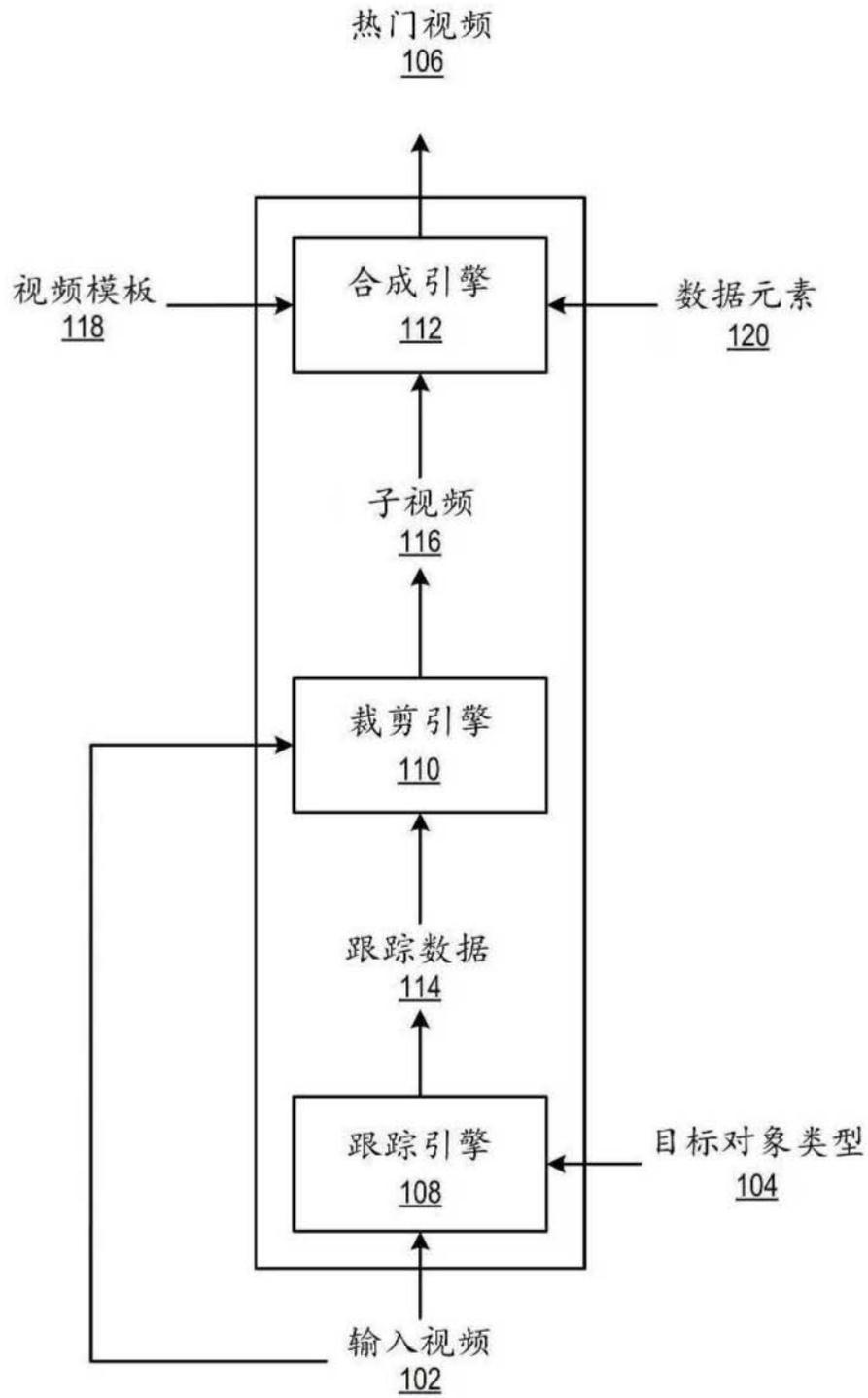
[0104] 本说明书中描述的主题的实施例可以实现在包括后端组件(例如,作为数据服务器)或包括中间件组件(例如,应用服务器)或包括前端组件(例如,具有图形用户界面或Web浏览器的客户端计算机,用户可以通过图形用户界面或Web浏览器与本说明书中描述的主题的实现进行交互)或者包括一个或多个这种后端组件,中间件组件或前端组件的任意组合的计算系统中。系统的组件可以通过数字数据通信的任何形式或介质(例如,通信网络)互连。通信网络的示例包括局域网(LAN)和广域网(WAN),例如,互联网。

[0105] 计算系统可以包括客户端和服务端。客户端和服务端通常彼此远离,并且通常通过通信网络进行交互。客户端和服务端之间的关系通过在相应计算机上运行并彼此具有客户端-服务器关系的计算机程序产生。在一些实施例中,服务器例如为了向与作为客户端的设备交互的用户显示数据并从中接收用户输入的目的,向用户设备发送例如HTML页面的数据。可以在服务器上从设备接收在用户设备处生成的数据,例如,用户交互的结果。

[0106] 尽管本说明书包含许多特定实现细节,但是这些细节不应被解释为对任何发明或所要求保护的范围的限制,而应解释为对特定发明的特定实施例而言特定的特征的描述。在单独的实施例的上下文中在本说明书中描述的特定特征也可以在单个实施例中组合实现。相反,在单个实施例的上下文中描述的各种特征也可以分别在多个实施例中或以任何合适的子组合来实现。此外,尽管上面可以将特征描述为以特定组合起作用并且甚至最初如此宣称,但是在一些情况下,可以从组合中删除所要求保护的组合中的一个或多个特征,并且可以将所要求保护的组合用于子组合或子组合的变型。

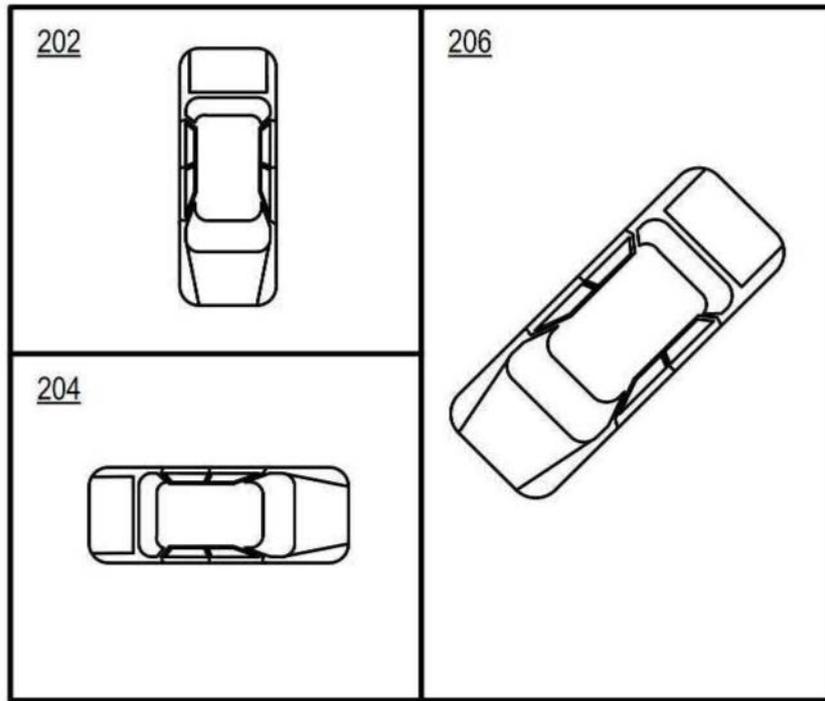
[0107] 类似地,尽管在附图中以特定顺序描绘了操作并且在权利要求中对其进行了叙述,但是这不应理解为要求以所示的特定顺序或以连续的顺序执行这样的操作,或者执行所有示出的操作以实现期望的结果。在特定情况下,多任务和并行处理可能是有利的。此外,上述实施例中的各种系统模块和组件的分离不应被理解为在所有实施例中都需要这种分离,并且应当理解,所描述的程序组件和系统通常可以一起集成在单个软件产品或打包成多个软件产品。

[0108] 已经描述了主题的特定实施例。其他实施例在所附权利要求的范围内。例如,权利要求中记载的动作可以以不同的顺序执行并且仍然实现期望的结果。作为一个示例,附图中描绘的处理不一定需要所示的特定顺序或连续顺序来实现期望的结果。在特定情况下,多任务和并行处理可能是有利的。



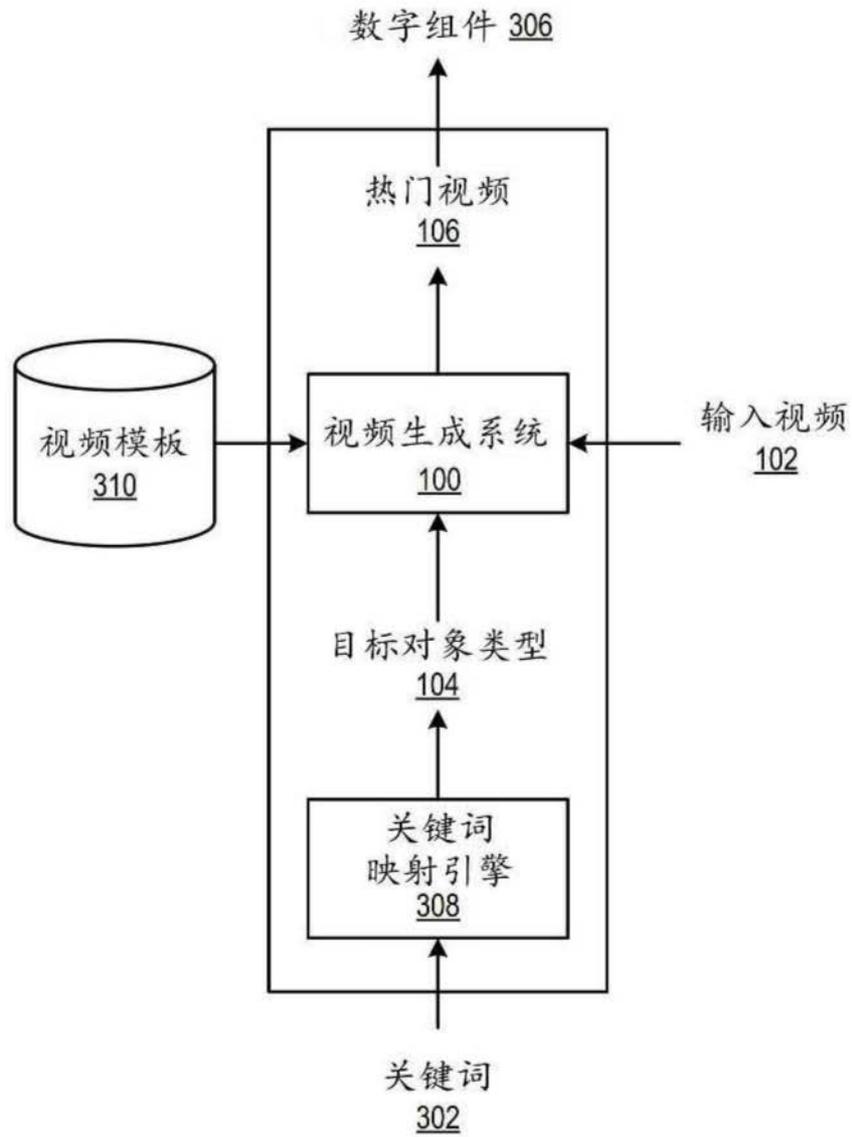
视频生成系统 **100**

图1



视频模板 200

图2



数字组件生成系统 **300**

图3

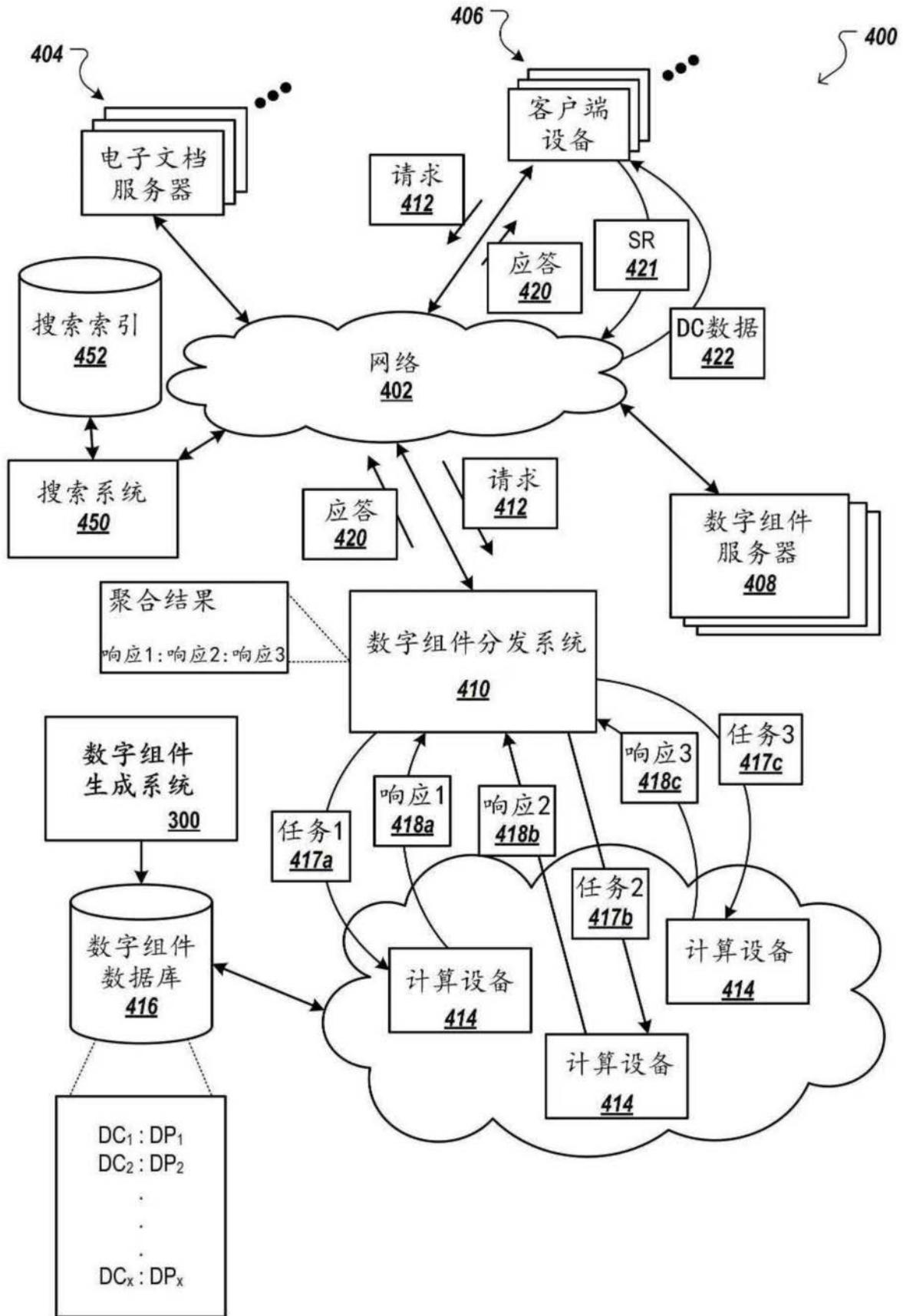


图4

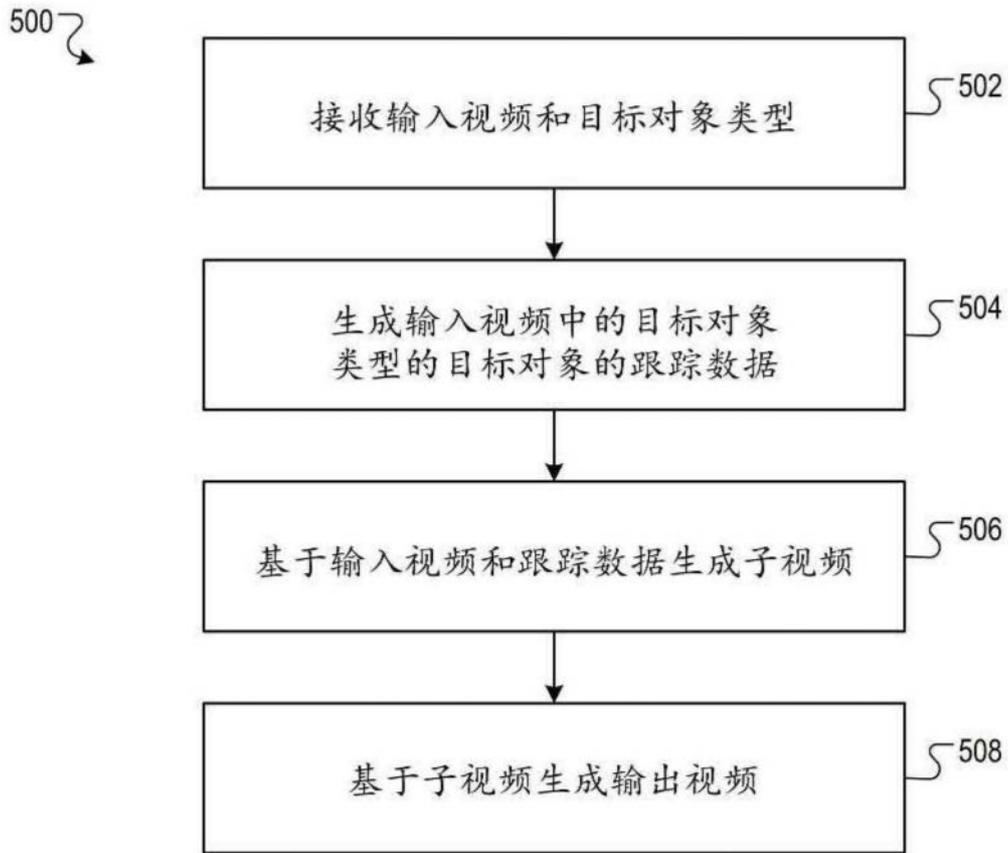


图5