



(12) 发明专利申请

(10) 申请公布号 CN 118132258 A

(43) 申请公布日 2024. 06. 04

(21) 申请号 202410234080.0

(22) 申请日 2024.02.29

(71) 申请人 山东云海国创云计算装备产业创新中心有限公司

地址 250102 山东省济南市高新区港西路2177号港盛大厦4层401室

(72) 发明人 张磊

(74) 专利代理机构 北京三聚阳光知识产权代理有限公司 11250

专利代理师 熊飞雪

(51) Int. Cl.

G06F 9/50 (2006.01)

G06T 1/20 (2006.01)

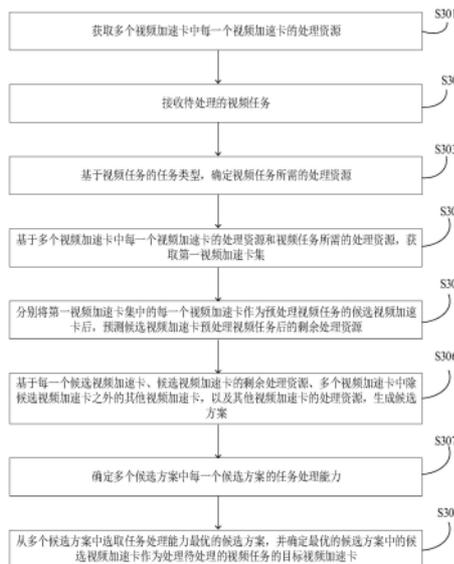
权利要求书3页 说明书15页 附图5页

(54) 发明名称

视频加速卡选取方法、装置、计算机设备及存储介质

(57) 摘要

本发明涉及视频处理技术领域,公开了视频加速卡选取方法、装置、计算机设备及存储介质,方法包括:基于每一个视频加速卡的处理资源和视频任务所需的处理资源,获取第一视频加速卡集;分别将第一视频加速卡集中的每一个视频加速卡作为预处理视频任务的候选视频加速卡后,预测候选视频加速卡预处理视频任务后的剩余处理资源;基于候选视频加速卡、候选视频加速卡的剩余处理资源、其他视频加速卡,以及其他视频加速卡的处理资源,生成候选方案;确定每一个候选方案的任务处理能力;选取任务处理能力最优的候选方案,并确定最优的候选方案中的候选视频加速卡作为目标视频加速卡。本发明实现了视频加速卡资源的自动分配,降低了用户开发成本。



1. 一种视频加速卡选取方法,其特征在于,所述方法包括:

获取多个视频加速卡中每一个视频加速卡的处理资源,所述处理资源用于处理视频任务;

接收待处理的视频任务;

基于所述视频任务的任务类型,确定所述视频任务所需的处理资源;

基于多个所述视频加速卡中每一个所述视频加速卡的处理资源和所述视频任务所需的处理资源,获取第一视频加速卡集;

分别将所述第一视频加速集中的每一个视频加速卡作为预处理所述视频任务的候选视频加速卡后,预测所述候选视频加速卡预处理所述视频任务后的剩余处理资源;

基于每一个所述候选视频加速卡、所述候选视频加速卡的剩余处理资源、多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡,以及所述其他视频加速卡的处理资源,生成候选方案;

确定多个所述候选方案中每一个候选方案的任务处理能力,所述任务处理能力通过所述候选视频加速卡的剩余处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量,以及多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡的处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量确定;

从多个所述候选方案中选取任务处理能力最优的候选方案,并确定所述最优的候选方案中的候选视频加速卡作为处理所述待处理的视频任务的目标视频加速卡。

2. 根据权利要求1所述的方法,其特征在于,所述基于多个所述视频加速卡中每一个所述视频加速卡的处理资源和所述视频任务所需的处理资源,获取第一视频加速卡集,包括:

从所述多个所述视频加速卡中筛选处理资源大于等于所述视频任务所需的处理资源的视频加速卡,构成所述第一视频加速卡集。

3. 根据权利要求1所述的方法,其特征在于,所述确定多个所述候选方案中每一个候选方案的任务处理能力,包括:

基于不同类型视频任务中每一种类型视频任务所需的处理资源,确定不同类型视频任务中每一种类型视频任务的优先级;

确定多个所述候选方案中每一个候选方案下,所述候选视频加速卡的剩余处理资源处理所述优先级最高的视频任务的数量,以及多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡的处理资源处理所述优先级最高的视频任务的数量,进而确定每一个所述候选方案下处理所述优先级最高的视频任务的总量;

基于处理所述优先级最高的视频任务的总量确定每一个所述候选方案的任务处理能力,其中,处理所述优先级最高的视频任务的总量的数量越多,对应候选方案的任务处理能力越大;

当存在至少两个候选方案处理优先级最高的视频任务的总量最多时,从处理优先级最高的视频任务的总量最多的候选方案中,获取处理次级优先级的视频任务的总量,并确定处理次级优先级的视频任务的总量最多的,任务处理能力最大;

或者,在处理次级优先级的视频任务的总量最多的候选方案包括至少两个时,按照所述优先级的排序顺序,进一步确定处理下一优先级的视频任务的总量,直至确定每一个所述候选方案的任务处理能力。

4. 根据权利要求3所述的方法,其特征在于,所述基于不同类型视频任务中每一种类型视频任务所需的处理资源,确定不同类型视频任务中每一种类型视频任务的优先级,包括:

按照不同类型视频任务中每一种类型视频任务所需的处理资源由多到少的顺序对所述不同类型视频任务中每一种类型视频任务进行由高到低的优先级的排列。

5. 根据权利要求3所述的方法,其特征在于,所述从多个所述候选方案中选取任务处理能力最优的候选方案,包括:

当按照所述优先级的排序顺序,确定多个候选方案中每一个候选方案处理第i优先级的视频任务之前的任一优先级的视频任务的总量均相同,则选取处理所述第i优先级的视频任务的总量最多的候选方案作为所述最优的候选方案;

或者,当存在至少两个候选方案处理每一个优先级的视频任务的总量均相同时,则从处理每一个优先级的视频任务的总量均相同的至少两个候选方案中随机选取一个候选方案作为所述最优的候选方案。

6. 根据权利要求3所述的方法,其特征在于,所述视频加速卡的处理资源、所述候选视频加速卡的剩余处理资源、所述不同类型视频任务所需的处理资源和所述待处理的视频任务所需的处理资源都是子处理资源的整数倍,其中,优先级最低的视频任务类型所需的处理资源为所述子处理资源。

7. 根据权利要求1所述的方法,其特征在于,在初始化启动的情况下,所述多个视频加速卡中每一个视频加速卡的处理资源为满状态。

8. 一种视频加速卡选取装置,其特征在于,所述装置包括:

处理资源获取模块,用于获取多个视频加速卡中每一个视频加速卡的处理资源,所述处理资源用于处理视频任务;

待处理视频任务接收模块,用于接收待处理的视频任务;

处理资源确定模块,用于基于所述视频任务的任务类型,确定所述视频任务所需的处理资源;

第一视频加速卡集获取模块,用于基于多个所述视频加速卡中每一个所述视频加速卡的处理资源和所述视频任务所需的处理资源,获取第一视频加速卡集;

剩余处理资源预测模块,用于分别将所述第一视频加速卡集中的每一个视频加速卡作为预处理所述视频任务的候选视频加速卡后,预测所述候选视频加速卡预处理所述视频任务后的剩余处理资源;

候选方案生成模块,用于基于每一个所述候选视频加速卡、所述候选视频加速卡的剩余处理资源、多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡,以及所述其他视频加速卡的处理资源,生成候选方案;

任务处理能力确定模块,用于确定多个所述候选方案中每一个候选方案的任务处理能力,所述任务处理能力通过所述候选视频加速卡的剩余处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量,以及多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡的处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量确定;

目标加速卡确定模块,用于从多个所述候选方案中选取任务处理能力最优的候选方案,并确定所述最优的候选方案中的候选视频加速卡作为处理所述待处理的视频任务的目

标视频加速卡。

9. 一种计算机设备,其特征在于,包括:

存储器和处理器,所述存储器和所述处理器之间互相通信连接,所述存储器中存储有计算机指令,所述处理器通过执行所述计算机指令,从而执行权利要求1至7中任一项所述的视频加速卡选取方法。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有计算机指令,所述计算机指令用于使计算机执行权利要求1至7中任一项所述的视频加速卡选取方法。

## 视频加速卡选取方法、装置、计算机设备及存储介质

### 技术领域

[0001] 本发明涉及视频处理技术领域,具体涉及视频加速卡选取方法、装置、计算机设备及存储介质。

### 背景技术

[0002] 视频加速卡是一种专门用于加速视频处理和渲染的硬件设备。它们通常与计算机或服务器配合使用,可以提供更快的视频处理速度和更流畅的播放体验。

[0003] 一般情况下,一路视频处理任务由一个视频加速卡运行,开发或运维人员需要在应用层手动或者通过配置文件的方式指定视频处理任务在哪一块视频加速卡上运行。这种方式需要用户对视频加速卡资源进行维护,增加了用户开发成本。

### 发明内容

[0004] 有鉴于此,本发明提供了一种视频加速卡选取方法、装置、计算机设备及存储介质,以解决现有视频加速卡选取方式需要用户对视频加速卡资源进行维护,用户开发成本高的问题。

[0005] 第一方面,本发明提供了一种视频加速卡选取方法,所述方法包括:

[0006] 获取多个视频加速卡中每一个视频加速卡的处理资源,所述处理资源用于处理视频任务;

[0007] 接收待处理的视频任务;

[0008] 基于所述视频任务的任务类型,确定所述视频任务所需的处理资源;

[0009] 基于多个所述视频加速卡中每一个所述视频加速卡的处理资源和所述视频任务所需的处理资源,获取第一视频加速卡集;

[0010] 分别将所述第一视频加速集中的每一个视频加速卡作为预处理所述视频任务的候选视频加速卡后,预测所述候选视频加速卡预处理所述视频任务后的剩余处理资源;

[0011] 基于每一个所述候选视频加速卡、所述候选视频加速卡的剩余处理资源、多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡,以及所述其他视频加速卡的处理资源,生成候选方案;

[0012] 确定多个所述候选方案中每一个候选方案的任务处理能力,所述任务处理能力通过所述候选视频加速卡的剩余处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量,以及多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡的处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量确定;

[0013] 从多个所述候选方案中选取任务处理能力最优的候选方案,并确定所述最优的候选方案中的候选视频加速卡作为处理所述待处理的视频任务的目标视频加速卡。

[0014] 本实施例提供的视频加速卡选取方法,通过待处理的视频任务所需的处理资源、每一个视频加速卡的处理资源、每一种类型视频任务所需的处理资源,确定最优的目标视频加速卡,利用目标视频加速卡处理待处理的视频任务,实现了视频加速卡资源的自动分

配,降低了用户开发成本。

[0015] 在一种可选的实施方式中,所述基于多个所述视频加速卡中每一个所述视频加速卡的处理资源和所述视频任务所需的处理资源,获取第一视频加速卡集,包括:

[0016] 从所述多个所述视频加速卡中筛选处理资源大于等于所述视频任务所需的处理资源的视频加速卡,构成所述第一视频加速卡集。

[0017] 本实施例提供的视频加速卡选取方法,实现了视频加速卡资源的自动分配,降低了用户开发成本。

[0018] 在一种可选的实施方式中,所述确定多个所述候选方案中每一个候选方案的任务处理能力,包括:

[0019] 基于不同类型视频任务中每一种类型视频任务所需的处理资源,确定不同类型视频任务中每一种类型视频任务的优先级;

[0020] 确定多个所述候选方案中每一个候选方案下,所述候选视频加速卡的剩余处理资源处理所述优先级最高的视频任务的数量,以及多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡的处理资源处理所述优先级最高的视频任务的数量,进而确定每一个所述候选方案下处理所述优先级最高的视频任务的总量;

[0021] 基于处理所述优先级最高的视频任务的总量确定每一个所述候选方案的任务处理能力,其中,处理所述优先级最高的视频任务的总量的数量越多,对应候选方案的任务处理能力越大;

[0022] 当存在至少两个候选方案处理优先级最高的视频任务的总量最多时,从处理优先级最高的视频任务的总量最多的候选方案中,获取处理次级优先级的视频任务的总量,并确定处理次级优先级的视频任务的总量最多的,任务处理能力最大;

[0023] 或者,在处理次级优先级的视频任务的总量最多的候选方案包括至少两个时,按照所述优先级的排序顺序,进一步确定处理下一优先级的视频任务的总量,直至确定每一个所述候选方案的任务处理能力。

[0024] 本实施例提供的视频加速卡选取方法,通过待处理的视频任务所需的处理资源、每一个视频加速卡的处理资源、每一种类型视频任务所需的处理资源以及每一个候选方案的任务处理能力,确定最优的目标视频加速卡,利用目标视频加速卡处理待处理的视频任务,实现了视频加速卡资源的自动分配,降低了用户开发成本。

[0025] 在一种可选的实施方式中,所述基于不同类型视频任务中每一种类型视频任务所需的处理资源,确定不同类型视频任务中每一种类型视频任务的优先级,包括:

[0026] 按照不同类型视频任务中每一种类型视频任务所需的处理资源由多到少的顺序对所述不同类型视频任务中每一种类型视频任务进行由高到低的优先级的排列。

[0027] 在一种可选的实施方式中,所述从多个所述候选方案中选取任务处理能力最优的候选方案,包括:

[0028] 当按照所述优先级的排序顺序,确定多个候选方案中每一个候选方案处理第*i*优先级的视频任务之前的任一优先级的视频任务的总量均相同,则选取处理所述第*i*优先级的视频任务的总量最多的候选方案作为所述最优的候选方案;

[0029] 或者,当存在至少两个候选方案处理每一个优先级的视频任务的总量均相同时,则从处理每一个优先级的视频任务的总量均相同的至少两个候选方案中随机选取一个候

选方案作为所述最优的候选方案。

[0030] 本实施例提供的视频加速卡选取方法,通过待处理的视频任务所需的处理资源、每一个视频加速卡的处理资源、每一种类型视频任务所需的处理资源以及每一个候选方案的任务处理能力,确定最优的目标视频加速卡,利用目标视频加速卡处理待处理的视频任务,实现了视频加速卡资源的自动分配,降低了用户开发成本。

[0031] 在一种可选的实施方式中,所述视频加速卡的处理资源、所述候选视频加速卡的剩余处理资源、所述不同类型视频任务所需的处理资源和所述待处理的视频任务所需的处理资源都是子处理资源的整数倍,其中,优先级最低的视频任务类型所需的处理资源为所述子处理资源。

[0032] 在一种可选的实施方式中,在初始化启动的情况下,所述多个视频加速卡中每一个视频加速卡的处理资源为满状态。

[0033] 第二方面,本发明提供了一种视频加速卡选取装置,所述装置包括:

[0034] 处理资源获取模块,用于获取多个视频加速卡中每一个视频加速卡的处理资源,所述处理资源用于处理视频任务;

[0035] 待处理视频任务接收模块,用于接收待处理的视频任务;

[0036] 处理资源确定模块,用于基于所述视频任务的任务类型,确定所述视频任务所需的处理资源;

[0037] 第一视频加速卡集获取模块,用于基于多个所述视频加速卡中每一个所述视频加速卡的处理资源和所述视频任务所需的处理资源,获取第一视频加速卡集;

[0038] 剩余处理资源预测模块,用于分别将所述第一视频加速卡集中的每一个视频加速卡作为预处理所述视频任务的候选视频加速卡后,预测所述候选视频加速卡预处理所述视频任务后的剩余处理资源;

[0039] 候选方案生成模块,用于基于每一个所述候选视频加速卡、所述候选视频加速卡的剩余处理资源、多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡,以及所述其他视频加速卡的处理资源,生成候选方案;

[0040] 任务处理能力确定模块,用于确定多个所述候选方案中每一个候选方案的任务处理能力,所述任务处理能力通过所述候选视频加速卡的剩余处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量,以及多个所述视频加速卡中除所述候选视频加速卡之外的其他视频加速卡的处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量确定;

[0041] 目标加速卡确定模块,用于从多个所述候选方案中选取任务处理能力最优的候选方案,并确定所述最优的候选方案中的候选视频加速卡作为处理所述待处理的视频任务的目标视频加速卡。

[0042] 第三方面,本发明提供了一种计算机设备,包括:存储器和处理器,存储器和处理器之间互相通信连接,存储器中存储有计算机指令,处理器通过执行计算机指令,从而执行上述第一方面或其对应的任一实施方式的视频加速卡选取方法。

[0043] 第四方面,本发明提供了一种计算机可读存储介质,该计算机可读存储介质上存储有计算机指令,计算机指令用于使计算机执行上述第一方面或其对应的任一实施方式的视频加速卡选取方法。

## 附图说明

[0044] 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案,下面将对具体实施方式或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施方式,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0045] 图1是根据相关技术中视频加速卡运行架构的示意图;

[0046] 图2是根据本发明实施例的视频加速卡运行架构的示意图;

[0047] 图3是根据本发明实施例的视频加速卡选取方法的流程示意图;

[0048] 图4是根据本发明实施例的另一视频加速卡选取方法的流程示意图;

[0049] 图5是根据本发明实施例的视频加速卡选取装置的结构框图;

[0050] 图6是本发明实施例的计算机设备的硬件结构示意图。

## 具体实施方式

[0051] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0052] 视频加速卡是一种专门用于加速视频处理和渲染的硬件设备。它们通常与计算机或服务器配合使用,可以提供更快的视频处理速度和更流畅的播放体验。

[0053] 目前市场上提供各种类型的视频加速卡,视频加速卡主要应用于以下几个领域:

[0054] 专业视频编辑:视频剪辑、特效处理和色彩校正等任务需要大量的计算资源,视频加速卡可以帮助编辑人员提高工作效率。

[0055] 游戏直播和内容创作:许多游戏玩家和内容创作者在进行游戏直播或录制时需要同时处理视频和游戏画面,视频加速卡可以减轻CPU(Central Processing Unit,中央处理器)负担,提供更流畅的直播和录制体验。

[0056] 云视频服务:随着云技术的发展,许多视频服务提供商开始采用视频加速卡来提供更高质量的视频内容和更好的用户体验。

[0057] 虽然视频加速卡在性能和效率方面取得了显著的进步,但目前市场上仍然存在一些挑战。首先是价格问题,视频加速卡通常比较昂贵,对于普通消费者来说可能不太容易承担。

[0058] 总体而言,视频加速卡在视频处理和渲染领域有着广阔的应用前景,并且随着技术的不断进步,我们可以期待更高性能和更多创新的视频加速卡产品出现。

[0059] 视频加速卡用于处理视频任务,视频任务可以是视频编解码任务,也可以是视频转码任务。以下以视频任务为视频编解码任务来描述,其中,视频编解码任务与视频分辨率之间存在一定的关系,视频分辨率是影响视频编解码任务资源占用的重要因素之一。一般而言,随着视频分辨率的增加,视频编解码任务对视频加速卡资源的需求也会相应增加。

[0060] 具体来说,视频分辨率对视频编解码任务资源占用的主要影响包括以下方面:

[0061] GPU(Graphics Processing Unit,图形处理器) 计算资源:视频分辨率的增加会导致视频编解码任务中的计算量增加。例如,在视频编码过程中,需要对更多的像素点进行压

缩处理；在解码过程中，需要还原更多的像素点。这将对视频加速卡的GPU计算资源产生更高的需求，可能需要更多的计算核心和更大的内存带宽。

[0062] 显存需求：视频分辨率越高，视频数据的大小也会相应增加。视频编解码任务需要将视频数据加载到视频加速卡的显存中进行处理。因此，随着视频分辨率的增加，视频编解码任务对显存容量的需求也会增加。较高分辨率的视频可能需要更大的显存来存储视频帧、码流以及相关的解码参数等。

[0063] 硬件解码器/编码器效率：视频加速卡上的硬件解码器和编码器通常针对特定的视频编解码格式进行优化。较高分辨率的视频编解码任务会对硬件解码器/编码器的效率提出更高要求，需要确保能够快速且稳定地处理大量高分辨率的视频数据。

[0064] 需要注意的是，视频编解码任务对视频加速卡资源的占用不仅与视频分辨率有关，还与视频编解码算法、压缩格式以及系统性能等因素有关。因此，在实际应用中，开发者需要综合考虑这些因素，并进行适当的系统配置和资源调配，以满足视频编解码任务的需求。

[0065] 图1是根据相关技术中视频加速卡运行架构的示意图。如图1所示，相关技术中，视频加速卡运行架构包括计算机硬件层、视频加速卡驱动层、视频加速卡软件栈、应用层和视频处理框架层，其中，多个视频加速卡与计算机硬件层物理连接，操作系统为视频加速卡驱动层、视频加速卡软件栈、应用层和视频处理框架层提供运行环境。

[0066] 相关技术中，用户利用视频加速卡处理视频编解码任务的方式为：应用层接收视频编解码任务，从用户预先设置的配置文件中确定运行视频编解码任务的视频加速卡，通过调用视频加速卡软件栈获取对应的视频加速卡，在该确定的视频加速卡上运行该视频编解码任务。

[0067] 上述应用层从用户预先设置的配置文件中确定运行视频编解码任务的视频加速卡之后，还可以通过视频处理框架调用视频加速卡软件栈，进而获取对应的视频加速卡，在该确定的视频加速卡上运行该视频编解码任务。示例性地，视频处理框架可以是FFmpeg。

[0068] 一般地，一路视频编解码任务由一个视频加速卡运行，用户通过在应用层手动或者通过配置文件的方式指定运行视频编解码任务的视频加速卡。其中，用户可以是开发或运维人员，也可以是使用视频加速卡的任意人员。

[0069] 示例性地，运行ffmpeg命令时进行镜像1080P视频编码，通过参数“insvid=dev0:/dev/transcoder0”指定业务使用/dev/transcoder0视频加速卡的资源进行处理。对应的代码可以包括：

```
[0070] ffmpeg-y-vsync 0-init_hw_device
```

```
[0071] insvid=dev0:/dev/transcoder0,priority=vod,insvidloglevel=0-s  
1920x1080-pix_fmt yuv420p-iinput.yuv-filter_complex'insvid_pp'-c:v h264enc_  
insvid-preset superfast-b:v 10000000out0.h264.
```

[0072] 使用代码进行业务开发时，在视频加速卡初始化时需要指定业务所用的视频加速卡，视频加速卡可以是/dev/transcoder4。对应的代码可以包括：

```
static const char* device = "/dev/transcoder4";
static int init_insvideo_device(CodecContext* ctx)
{
    int ret = 0;
    AVDictionary *opts = NULL;

    for(int i = 0; i < FF_ARRAY_ELEMS(device_options); i++){
        av_dict_set(&opts, device_options[i].key, device_options[i].value, 0);
    }
[0073]         ret = av_hwdevice_ctx_create(&ctx->hw_device_ctx
AV_HWDEVICE_TYPE_INSVID, device, opts, 0);
    if (ret < 0){
        av_log(NULL, AV_LOG_ERROR,
            "Failed to create a insvideo device. Error code: %s\n",
            av_err2str(ret));
        return -1;
    }
    return 0;
}
```

[0074] 可见,相关技术中,视频任务绑定视频加速卡需要应用层用户指定,需要用户对视频加速卡资源进行维护,增加用户开发成本。用户无法获取视频加速卡资源,视频加速卡到达任务处理上限之后,用户仍然可以对该卡下发视频任务,会导致业务资源混乱,影响其他正在运行的任务。

[0075] 图2是根据本发明实施例的视频加速卡运行架构的示意图。如图2所示,本发明实施例中,视频加速卡运行架构包括计算机硬件层、视频加速卡驱动层、视频加速卡软件栈、视频加速卡资源管理模块、应用层和视频处理框架层,其中,多个视频加速卡与计算机硬件层物理连接,操作系统为视频加速卡驱动层、视频加速卡软件栈、视频加速卡资源管理模块、应用层和视频处理框架层提供运行环境。

[0076] 相较于图1,本发明实施例的视频加速卡运行架构增加了视频加速卡资源管理模块,利用该模块执行视频加速卡选取方法,能够为待处理的视频任务自动分配其对应的视频加速卡,无需用户手动在应用层进行预先配置。

[0077] 本发明实施例提供了一种视频加速卡选取方法,该视频加速卡选取方法在视频加速卡资源管理模块上执行。视频加速卡处理视频编解码任务的方式为:应用层接收视频编解码任务,视频加速卡资源管理模块执行视频加速卡选取方法,获得处理视频编解码任务的视频加速卡,通过该视频加速卡处理该视频编解码任务。本发明实施例提供的视频加速卡选取方法通过获取多个视频加速卡中每一个视频加速卡的处理资源、待处理的视频任务

所需的处理资源以及不同类型视频任务所需的处理资源,为待处理的视频任务分配最优的视频加速卡,达到了为视频任务自动分配视频加速卡的效果,且保证了视频加速卡资源的合理分配。

[0078] 根据本发明实施例,提供了一种视频加速卡选取方法实施例,需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行,并且,虽然在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0079] 在本实施例中提供了一种视频加速卡选取方法,可用于上述的移动终端,如中央处理单元、服务器等,图3是根据本发明实施例的视频加速卡选取方法的流程图,如图3所示,该流程包括如下步骤:

[0080] 步骤S301,获取多个视频加速卡中每一个视频加速卡的处理资源。

[0081] 其中,处理资源用于处理视频任务,多个视频加速卡与计算机硬件物理连接。

[0082] 获取的是多个视频加速卡中每一个视频加速卡的当前处理资源。可以理解的是,当前多个视频加速卡中可能存在部分视频加速卡正在处理之前下发的视频任务,则视频加速卡的当前处理资源为视频加速卡的总处理资源与该视频加速卡正在处理的视频任务所占处理资源的差值。

[0083] 步骤S302,接收待处理的视频任务。

[0084] 其中,接收应用层下发的待处理的视频任务,或者,接收应用层经由视频处理框架下发的待处理的视频任务,该视频任务可以是视频编码任务,也可以是视频解码任务,也可以是视频转码任务。

[0085] 步骤S303,基于视频任务的任务类型,确定视频任务所需的处理资源。

[0086] 其中,在获取到待处理的视频任务后,根据待处理的视频任务的任务类型,确定该视频任务所需的处理资源。

[0087] 需要说明的是,视频任务的任务类型可以包括:视频分辨率、视频帧率、视频编解码算法、压缩格式以及系统性能等。视频任务的任务类型所涵盖的信息应与视频任务所需的处理资源相关。

[0088] 步骤S304,基于多个视频加速卡中每一个视频加速卡的处理资源和视频任务所需的处理资源,获取第一视频加速卡集。

[0089] 其中,在获取到多个视频加速卡中每一个视频加速卡的当前处理资源和下发的视频任务所需的处理资源之后,将多个视频加速卡中每一个视频加速卡的当前处理资源与下发的视频任务所需的处理资源进行比较,获得第一视频加速卡集。

[0090] 步骤S305,分别将第一视频加速卡集中的每一个视频加速卡作为预处理视频任务的候选视频加速卡后,预测候选视频加速卡预处理视频任务后的剩余处理资源。

[0091] 其中,在获取到第一视频加速卡集之后,分别将第一视频加速卡集中的每一个视频加速卡作为预处理视频任务的候选视频加速卡,获取该候选视频加速卡预处理视频任务后的剩余处理资源。

[0092] 候选视频加速卡预处理视频任务后的剩余处理资源通过该候选视频加速卡的当前处理资源与视频任务所需的处理资源做差值获得。

[0093] 步骤S306,基于每一个候选视频加速卡、候选视频加速卡的剩余处理资源、多个视

频加速卡中除候选视频加速卡之外的其他视频加速卡,以及其他视频加速卡的处理资源,生成候选方案。

[0094] 其中,在确定了候选视频加速卡和该候选视频加速卡的剩余处理资源后,将候选视频加速卡,候选视频加速卡的剩余处理资源、多个视频加速卡中除候选视频加速卡之外的其他视频加速卡,以及其他视频加速卡的处理资源作为一个候选方案。

[0095] 可以理解的是,候选方案的个数与第一视频加速卡集中视频加速卡的数量相同且小于与计算机硬件物理连接的视频加速卡的总数量。也就是说,第一个视频加速卡集中的每一个视频加速卡作为候选视频加速卡时,对应一个候选方案。

[0096] 多个视频加速卡中除候选视频加速卡之外的其他视频加速卡的处理资源与获取的其他视频加速卡的当前处理资源相同。

[0097] 步骤S307,确定多个候选方案中每一个候选方案的任务处理能力。

[0098] 其中,在获取到多个候选方案后,通过候选视频加速卡的剩余处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量,以及多个视频加速卡中除候选视频加速卡之外的其他视频加速卡的处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量确定每一个候选方案的任务处理能力。

[0099] 需要说明的是,不同类型视频任务可以是视频加速卡可处理的多种类型的视频任务。

[0100] 假设多个视频加速卡中除候选视频加速卡之外的一个视频加速卡的总处理资源为32个子处理资源,当前处理资源为21个子处理资源,该视频加速卡能处理四种类型的视频编码任务,分别是720p30fps、1080p30fps、4k30fps、4k60fps,在影响视频编解码任务资源占用的因素中除视频分辨率和视频帧率之外的因素相同的情况下,上述四种类型的视频编码任务中,720p30fps对应1个子处理资源,1080p30fps对应2个子处理资源,4k30fps对应8个子处理资源,4k60fps对应16个子处理资源。

[0101] 则,该视频加速卡的当前处理资源按照优先级的顺序对四种类型的视频编码任务进行预处理,即预处理的顺序为4k60fps、4k30fps、1080p30fps、720p30fps。该视频加速卡的当前处理资源除以4k60fps任务所需的处理资源,商为1,余数为5,则该视频加速卡可以预处理1个4k60fps任务,剩余5个子处理资源。利用剩余的5个子处理资源除以4k30fps任务所需的处理资源,商为0,余数为5,则该视频加速卡可以预处理0个4k30fps任务,剩余5个子处理资源。利用剩余的5个子处理资源除以1080p30fps任务所需的处理资源,商为2,余数为1,则该视频加速卡可以预处理2个1080p30fps任务,剩余1个子处理资源。利用剩余的1个子处理资源除以720p30fps任务所需的处理资源,商为1,余数为0,则该视频加速卡可以预处理1个720p30fps任务,无剩余处理资源。

[0102] 可见,该视频加速卡的当前处理资源可以预处理1个4k60fps任务、0个4k30fps任务、2个1080p30fps任务和1个720p30fps任务。

[0103] 步骤S308,从多个候选方案中选取任务处理能力最优的候选方案,并确定最优的候选方案中的候选视频加速卡作为处理待处理的视频任务的目标视频加速卡。

[0104] 其中,在获得多个候选方案中每一个候选方案的任务处理能力之后,从多个候选方案中选取任务处理能力最优的候选方案,并确定最优的候选方案中的候选视频加速卡作为处理待处理的视频任务的目标视频加速卡。

[0105] 可以理解的是,任务处理能力越大,其对应的候选方案越优。

[0106] 在确定了目标视频加速卡后,利用目标视频加速卡处理待处理的视频任务,更新目标视频加速卡的处理资源为目标视频加速卡的当前处理资源与待处理的视频任务所需的处理资源的差值,也可以更新为最优的候选方案中的候选视频加速卡的剩余处理资源。

[0107] 进一步需要说明的是,在目标视频加速卡处理完该待处理的视频任务后,再次更新目标视频加速卡的处理资源,将目标视频加速卡的处理资源更新为处理该视频任务之前的处理资源值。

[0108] 本实施例提供的视频加速卡选取方法,通过获取多个视频加速卡中每一个视频加速卡的处理资源;接收待处理的视频任务;基于视频任务的任务类型,确定视频任务所需的处理资源;基于多个视频加速卡中每一个视频加速卡的处理资源和视频任务所需的处理资源,获取第一视频加速卡集;分别将第一视频加速卡集中的每一个视频加速卡作为预处理视频任务的候选视频加速卡后,预测候选视频加速卡预处理视频任务后的剩余处理资源;基于每一个候选视频加速卡、候选视频加速卡的剩余处理资源、多个视频加速卡中除候选视频加速卡之外的其他视频加速卡,以及其他视频加速卡的处理资源,生成候选方案;确定多个候选方案中每一个候选方案的任务处理能力,从多个候选方案中选取任务处理能力最优的候选方案,并确定最优的候选方案中的候选视频加速卡作为处理待处理的视频任务的目标视频加速卡。实现了视频加速卡资源的自动分配和自动维护,无需用户手动配置,降低用户开发成本。

[0109] 在本实施例中提供了一种视频加速卡选取方法,可用于上述的移动终端,如中央处理单元、服务器等,图4是根据本发明实施例的视频加速卡选取方法的流程图,如图4所示,该流程包括如下步骤:

[0110] 步骤S401,获取多个视频加速卡中每一个视频加速卡的处理资源。详细请参见图3所示实施例的步骤S301,在此不再赘述。

[0111] 步骤S402,接收待处理的视频任务。详细请参见图3所示实施例的步骤S302,在此不再赘述。

[0112] 步骤S403,基于视频任务的任务类型,确定视频任务所需的处理资源。详细请参见图3所示实施例的步骤S303,在此不再赘述。

[0113] 步骤S404,基于多个视频加速卡中每一个视频加速卡的处理资源和视频任务所需的处理资源,获取第一视频加速卡集。

[0114] 具体地,上述步骤S404包括:

[0115] 步骤S4041,从多个视频加速卡中筛选处理资源大于等于视频任务所需的处理资源的视频加速卡,构成第一视频加速卡集。

[0116] 其中,根据待处理的视频任务所需的处理资源,遍历每一张视频加速卡,假设待处理的视频任务下发到第n张视频加速卡上,判断该视频加速卡的处理资源是否大于等于视频任务所需的处理资源,若该视频加速卡的处理资源小于视频任务所需的处理资源,跳过该视频加速卡,继续遍历下一张视频加速卡直至遍历结束,遍历次数为视频加速卡的总数。

[0117] 遍历结束后,获取处理资源大于等于视频任务所需的处理资源的视频加速卡,构成第一视频加速卡集。

[0118] 步骤S405,分别将第一视频加速卡集中的每一个视频加速卡作为预处理视频任务

的候选视频加速卡后,预测候选视频加速卡预处理视频任务后的剩余处理资源。详细请参见图3所示实施例的步骤S305,在此不再赘述。

[0119] 步骤S406,基于每一个候选视频加速卡、候选视频加速卡的剩余处理资源、多个视频加速卡中除候选视频加速卡之外的其他视频加速卡,以及其他视频加速卡的处理资源,生成候选方案。详细请参见图3所示实施例的步骤S306,在此不再赘述。

[0120] 步骤S407,确定多个候选方案中每一个候选方案的任务处理能力。详细请参见图3所示实施例的步骤S307,在此不再赘述。

[0121] 步骤S408,从多个候选方案中选取任务处理能力最优的候选方案,并确定最优的候选方案中的候选视频加速卡作为处理待处理的视频任务的目标视频加速卡。详细请参见图3所示实施例的步骤S308,在此不再赘述。

[0122] 本实施例提供的视频加速卡选取方法,实现了视频加速卡资源的自动分配和自动维护,无需用户手动配置,降低用户开发成本。

[0123] 在一些可选的实施方式中,上述步骤S407包括:

[0124] 步骤a1,基于不同类型视频任务中每一种类型视频任务所需的处理资源,确定不同类型视频任务中每一种类型视频任务的优先级。

[0125] 其中,不同类型视频任务可以是视频加速卡可处理的多种类型的视频任务。

[0126] 步骤a2,确定多个候选方案中每一个候选方案下,候选视频加速卡的剩余处理资源处理优先级最高的视频任务的数量,以及多个视频加速卡中除候选视频加速卡之外的其他视频加速卡的处理资源处理优先级最高的视频任务的数量,进而确定每一个候选方案下处理优先级最高的视频任务的总量。

[0127] 其中,在确定了不同类型视频任务中每一种类型视频任务的优先级之后,获取每一个候选方案下,候选视频加速卡的剩余处理资源处理优先级最高的视频任务的数量与多个视频加速卡中除候选视频加速卡之外的其他视频加速卡的处理资源处理优先级最高的视频任务的数量之和,为每一个候选方案处理优先级最高的视频任务的总量。

[0128] 步骤a3,基于处理优先级最高的视频任务的总量确定每一个候选方案的任务处理能力。

[0129] 其中,处理优先级最高的视频任务的总量的数量越多,对应候选方案的任务处理能力越大。

[0130] 步骤a4,当存在至少两个候选方案处理优先级最高的视频任务的总量最多时,从处理优先级最高的视频任务的总量最多的候选方案中,获取处理次级优先级的视频任务的总量,并确定处理次级优先级的视频任务的总量最多的,任务处理能力最大。

[0131] 或者,在处理次级优先级的视频任务的总量最多的候选方案包括至少两个时,按照优先级的排序顺序,进一步确定处理下一优先级的视频任务的总量,直至确定每一个候选方案的任务处理能力。

[0132] 可以理解的是,在处理优先级最高的视频任务的总量最多的候选方案为至少两个的情况下,按照优先级的排序顺序,比较该处理优先级最高的视频任务的总量最多的候选方案中次级优先级的视频任务的总量,获取其中处理次级优先级的视频任务的总量最多的候选方案,其任务处理能力最大。如果处理次级优先级的视频任务的总量最多的方案包括至少两个,则继续按照优先级的顺序比较处理下一优先级的视频任务的总量,直至确定每

一个候选方案的任务处理能力。

[0133] 示例性地,第一视频加速卡集中包括三个视频加速卡,将三个视频加速卡分别作为候选视频加速卡,生成每一个候选视频加速卡对应的候选方案。

[0134] 获取每一个候选方案处理不同类型视频任务中每一个类型视频任务的总量,其中,视频任务类型可以包括720p30fps、1080p30fps、4k30fps、4k60fps,三个候选方案处理不同类型视频任务中每一个类型视频任务的总量可以如表1所示。

[0135] 表1

视频任务类型	4k60fps	4k30fps	1080p30fps	720p30fps
所需处理资源	16	8	2	1
候选方案1	1	0	2	1
[0136] 可运行视频任务总数				
候选方案2	1	0	1	3
可运行视频任务总数				
候选方案3	0	2	2	1
可运行视频任务总数				

[0137] 其中,候选方案1与候选方案2处理优先级最高的视频任务的总数比候选方案3的总数多且相同,如表1,候选方案1与候选方案2处理4k60fps任务的总数都是1。

[0138] 在处理优先级最高的视频任务的总数最多的候选方案包括至少两个的情况下,比较候选方案1和候选方案2处理次级优先级的视频任务的总数,即比较候选方案1和候选方案2处理4k30fps任务的总数。

[0139] 如表1所示,候选方案1和候选方案2处理4k30fps任务的总数都为0,则继续比较候选方案和候选方案2处理下一优先级的视频任务的总数,即比较候选方案1和候选方案2处理1080p30fps任务的总数。

[0140] 如表1所示,候选方案1相较于候选方案2处理1080p30fps任务的总数更多,则确定候选方案1的任务处理能力最大,选取候选方案1为最优的候选方案,候选方案1中的候选视频加速卡为处理待处理的视频任务的目标视频加速卡。

[0141] 在一些可选的实施方式中,上述步骤S407还包括:

[0142] 当存在至少两个候选方案处理优先级最高的视频任务的总量最多时,从处理优先级最高的视频任务的总量最多的候选方案中,获取处理不同类型视频任务中除优先级最高的视频任务之外的其他优先级的视频任务的总量,并确定处理其他优先级的视频任务的总量最多的,任务处理能力最大。

[0143] 或者,在处理不同类型视频任务中除优先级最高的视频任务之外的其他优先级的视频任务的总量最多的候选方案包括两个时,按照优先级的排序顺序,获取处理次级优先

级的视频任务的总量,并确定处理次级优先级的视频任务的总量最多的,任务处理能力最大。

[0144] 在处理次级优先级的视频任务的总量最多的候选方案包括至少两个时,按照优先级的排序顺序,获取低于次级优先级的其他优先级的视频任务的总量,并确定处理低于次级优先级的其他优先级的视频任务的总量最多的,任务处理能力最大。

[0145] 可以理解的是,在处理低于次级优先级的其他优先级的视频任务的总量最多的候选方案包括至少两个时,按照上述方式继续比较,直至确定每一个候选方案的任务处理能力。

[0146] 本实施例提供的视频加速卡选取方法,在实现视频加速卡资源的自动分配的同时,还实现了视频加速卡资源的合理分配,能够保证后续业务中优先级高的视频任务的分配,基于现有视频加速卡的数量达到最优的使用效果,减少了视频加速卡的成本。

[0147] 在一些可选的实施方式中,上述步骤a1包括:

[0148] 步骤a11,按照不同类型视频任务中每一种类型视频任务所需的处理资源由多到少的顺序对不同类型视频任务中每一种类型视频任务进行由高到低的优先级的排列。

[0149] 其中,视频任务所需的处理资源越多,其对应的优先级越高。

[0150] 在一些可选的实施方式中,上述步骤S408包括:

[0151] 步骤b1,当按照优先级的排序顺序,确定多个候选方案中每一个候选方案处理第i优先级的视频任务之前的任一优先级的视频任务的总量均相同,则选取处理第i优先级的视频任务的总量最多的候选方案作为最优的候选方案。

[0152] 或者,当存在至少两个候选方案处理每一个优先级的视频任务的总量均相同时,则从处理每一个优先级的视频任务的总量均相同的至少两个候选方案中随机选取一个候选方案作为最优的候选方案。

[0153] 在一些可选的实施方式中,视频加速卡的处理资源、候选视频加速卡的剩余处理资源、不同类型视频任务所需的处理资源和待处理的视频任务所需的处理资源都是子处理资源的整数倍,其中,优先级最低的视频任务类型所需的处理资源为子处理资源。

[0154] 按照前述方式选取视频加速卡进行视频任务的处理,可能会存在视频加速卡处理资源浪费的情况,为解决视频加速卡处理资源存在浪费的情况,本实施例通过将优先级最低的视频任务类型所需的处理资源作为子处理资源,也即视频任务类型所需的处理资源最少的处理资源值作为子处理资源。视频加速卡的处理资源、候选视频加速卡的剩余处理资源、不同类型视频任务所需的处理资源和待处理的视频任务所需的处理资源都是子处理资源的整数倍,使得视频加速卡的处理资源能够被合理利用,不会造成资源浪费。

[0155] 在一些可选的实施方式中,在初始化启动的情况下,多个视频加速卡中每一个视频加速卡的处理资源为满状态。

[0156] 可以理解的是,在视频加速卡运行架构每次初始启动的时候,多个视频加速卡中每一个视频加速卡的处理资源都处于未被使用状态,即满状态。

[0157] 本实施例提供的视频加速卡选取方法,通过将优先级最低的视频任务类型所需的处理资源作为子处理资源,视频加速卡的处理资源、候选视频加速卡的剩余处理资源、不同类型视频任务所需的处理资源和待处理的视频任务所需的处理资源都是子处理资源的整数倍,在实现视频加速卡资源的自动分配的同时,进一步实现了视频加速卡资源的合理分

配。

[0158] 为使得本实施例提供的视频加速卡选取方法更加清楚,以下以在4张处理资源值为32个子处理资源的视频加速卡上编码四种类型的视频任务为例进行说明,其中,四种类型的视频任务分别是720p30fps、1080p30fps、4k30fps、4k60fps,在影响视频编解码任务资源占用的因素中除视频分辨率和视频帧率之外的因素相同的情况下,上述四种类型的视频编码任务中,720p30fps对应1个子处理资源,1080p30fps对应2个子处理资源,4k30fps对应8个子处理资源,4k60fps对应16个子处理资源。

[0159] 初始化函数init()函数对所有视频加速卡的处理资源数据进行初始化,使得每张视频加速卡的处理资源值为32,已使用处理资源值为0。任务创建通过int create\_task(task\_struct\*task)函数,通过select\_card(int need\_ability\_value)接口获取处理该任务最优的视频加速卡。select\_card函数将遍历四张视频加速卡,分别计算将视频任务下发至当前候选视频加速卡后4张视频加速卡的剩余处理资源,比较四种方案,对比四组剩余处理资源,以可运行的高优先级视频任务总数最多的方案为最优方案,选取该最优方案中的视频加速卡进行视频任务下发。将视频任务下发至该视频加速卡后将该视频加速卡的剩余处理资源减少;视频任务删除通过stop\_task(task\_struct\*task)函数将视频任务停止后更新对应视频加速卡的处理资源值。

[0160] 在本实施例中还提供了一种视频加速卡选取装置,该装置用于实现上述实施例及优选实施方式,已经进行过说明的不再赘述。如以下所使用的,术语“模块”可以实现预定功能的软件和/或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0161] 本实施例提供一种视频加速卡选取装置,如图5所示,包括:

[0162] 处理资源获取模块501,用于获取多个视频加速卡中每一个视频加速卡的处理资源,处理资源用于处理视频任务。

[0163] 待处理视频任务接收模块502,用于接收待处理的视频任务。

[0164] 处理资源确定模块503,用于基于视频任务的任务类型,确定视频任务所需的处理资源。

[0165] 第一视频加速卡集获取模块504,用于基于多个视频加速卡中每一个视频加速卡的处理资源和视频任务所需的处理资源,获取第一视频加速卡集。

[0166] 剩余处理资源预测模块505,用于分别将第一视频加速卡集中的每一个视频加速卡作为预处理视频任务的候选视频加速卡后,预测候选视频加速卡预处理视频任务后的剩余处理资源。

[0167] 候选方案生成模块506,用于基于每一个候选视频加速卡、候选视频加速卡的剩余处理资源、多个视频加速卡中除候选视频加速卡之外的其他视频加速卡,以及其他视频加速卡的处理资源,生成候选方案。

[0168] 任务处理能力确定模块507,用于确定多个候选方案中每一个候选方案的任务处理能力,任务处理能力通过候选视频加速卡的剩余处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量,以及多个视频加速卡中除候选视频加速卡之外的其他视频加速卡的处理资源分别处理不同类型视频任务中一种或多种类型视频任务的数量确定。

[0169] 目标加速卡确定模块508,用于从多个候选方案中选取任务处理能力最优的候选

方案,并确定最优的候选方案中的候选视频加速卡作为处理待处理的视频任务的目标视频加速卡。

[0170] 在一些可选的实施方式中,第一视频加速卡集获取模块504,包括:

[0171] 第一视频加速卡集获取子单元,用于从多个视频加速卡中筛选处理资源大于等于视频任务所需的处理资源的视频加速卡,构成第一视频加速卡集。

[0172] 在一些可选的实施方式中,任务处理能力确定模块507包括:

[0173] 任务优先级确定单元,用于基于不同类型视频任务中每一种类型视频任务所需的处理资源,确定不同类型视频任务中每一种类型视频任务的优先级。

[0174] 任务总量确定单元,用于确定多个候选方案中每一个候选方案下,候选视频加速卡的剩余处理资源处理优先级最高的视频任务的数量,以及多个视频加速卡中除候选视频加速卡之外的其他视频加速卡的处理资源处理优先级最高的视频任务的数量,进而确定每一个候选方案下处理优先级最高的视频任务的总量。

[0175] 第一任务处理能力确定子单元,用于基于处理优先级最高的视频任务的总量确定每一个候选方案的任务处理能力,其中,处理优先级最高的视频任务的总量的数量越多,对应候选方案的任务处理能力越大。

[0176] 第二任务处理能力确定子单元,用于当存在至少两个候选方案处理优先级最高的视频任务的总量最多时,从处理优先级最高的视频任务的总量最多的候选方案中,处理次级优先级的视频任务的总量,并确定处理次优先级的视频任务的总量最多的,任务处理能力最大。

[0177] 或者,在处理次级优先级的视频任务的总量最多的候选方案包括至少两个时,按照优先级的排序顺序,进一步确定处理下一优先级的视频任务的总量,直至每一个候选方案的任务处理能力。

[0178] 在一些可选的实施方式中,任务优先级确定单元包括:

[0179] 任务优先级确定子单元,用于按照不同类型视频任务中每一种类型视频任务所需的处理资源由多到少的顺序对不同类型视频任务中每一种类型视频任务进行由高到低的优先级的排列。

[0180] 在一些可选的实施方式中,目标加速卡确定模块508包括:

[0181] 最优候选方案确定单元,用于当按照优先级的排序顺序,确定多个候选方案中每一个候选方案处理第i优先级的视频任务之前的任一优先级的视频任务的总量均相同,则选取处理第i优先级的视频任务的总量最多的候选方案作为最优的候选方案。

[0182] 或者,当存在至少两个候选方案处理每一个优先级的视频任务的总量均相同时,则从处理每一个优先级的视频任务的总量均相同的至少两个候选方案中随机选取一个候选方案作为最优的候选方案。

[0183] 上述各个模块和单元的更进一步的功能描述与上述对应实施例相同,在此不再赘述。

[0184] 本实施例中的视频加速卡选取装置是以功能单元的形式来呈现,这里的单元是指ASIC(Application Specific Integrated Circuit,专用集成电路)电路,执行一个或多个软件或固定程序的处理器和存储器,和/或其他可以提供上述功能的器件。

[0185] 本发明实施例还提供一种计算机设备,具有上述图5所示的视频加速卡选取装置。

[0186] 请参阅图6,图6是本发明可选实施例提供的一种计算机设备的结构示意图,如图6所示,该计算机设备包括:一个或多个处理器601、存储器602,以及用于连接各部件的接口,包括高速接口和低速接口。各个部件利用不同的总线互相通信连接,并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在计算机设备内执行的指令进行处理,包括存储在存储器中或者存储器上以在外部输入/输出装置(诸如,耦合至接口的显示设备)上显示GUI的图形信息的指令。在一些可选的实施方式中,若需要,可以将多个处理器和/或多条总线与多个存储器和多个存储器一起使用。同样,可以连接多个计算机设备,各个设备提供部分必要的操作(例如,作为服务器阵列、一组刀片式服务器、或者多处理器系统)。图6中以一个处理器601为例。

[0187] 处理器601可以是中央处理器,网络处理器或其组合。其中,处理器601还可以进一步包括硬件芯片。上述硬件芯片可以是专用集成电路,可编程逻辑器件或其组合。上述可编程逻辑器件可以是复杂可编程逻辑器件,现场可编程逻辑门阵列,通用阵列逻辑或其任意组合。

[0188] 其中,所述存储器602存储有可由至少一个处理器601执行的指令,以使所述至少一个处理器601执行实现上述实施例示出的方法。

[0189] 存储器602可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储根据计算机设备的使用所创建的数据等。此外,存储器602可以包括高速随机存取存储器,还可以包括非瞬时存储器,例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些可选的实施方式中,存储器602可选包括相对于处理器601远程设置的存储器,这些远程存储器可以通过网络连接至该计算机设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0190] 存储器602可以包括易失性存储器,例如,随机存取存储器;存储器也可以包括非易失性存储器,例如,快闪存储器,硬盘或固态硬盘;存储器602还可以包括上述种类的存储器的组合。

[0191] 该计算机设备还包括通信接口603,用于该计算机设备与其他设备或通信网络通信。

[0192] 本发明实施例还提供了一种计算机可读存储介质,上述根据本发明实施例的方法可在硬件、固件中实现,或者被实现为可记录在存储介质,或者被实现通过网络下载的原始存储在远程存储介质或非暂时机器可读存储介质中并将被存储在本地存储介质中的计算机代码,从而在此描述的方法可被存储在使用通用计算机、专用处理器或者可编程或专用硬件的存储介质上的这样的软件处理。其中,存储介质可为磁碟、光盘、只读存储记忆体、随机存储记忆体、快闪存储器、硬盘或固态硬盘等;进一步地,存储介质还可以包括上述种类的存储器的组合。可以理解,计算机、处理器、微处理器控制器或可编程硬件包括可存储或接收软件或计算机代码的存储组件,当软件或计算机代码被计算机、处理器或硬件访问且执行时,实现上述实施例示出的方法。

[0193] 虽然结合附图描述了本发明的实施例,但是本领域技术人员可以在不脱离本发明的精神和范围的情况下做出各种修改和变型,这样的修改和变型均落入由所附权利要求所限定的范围之内。

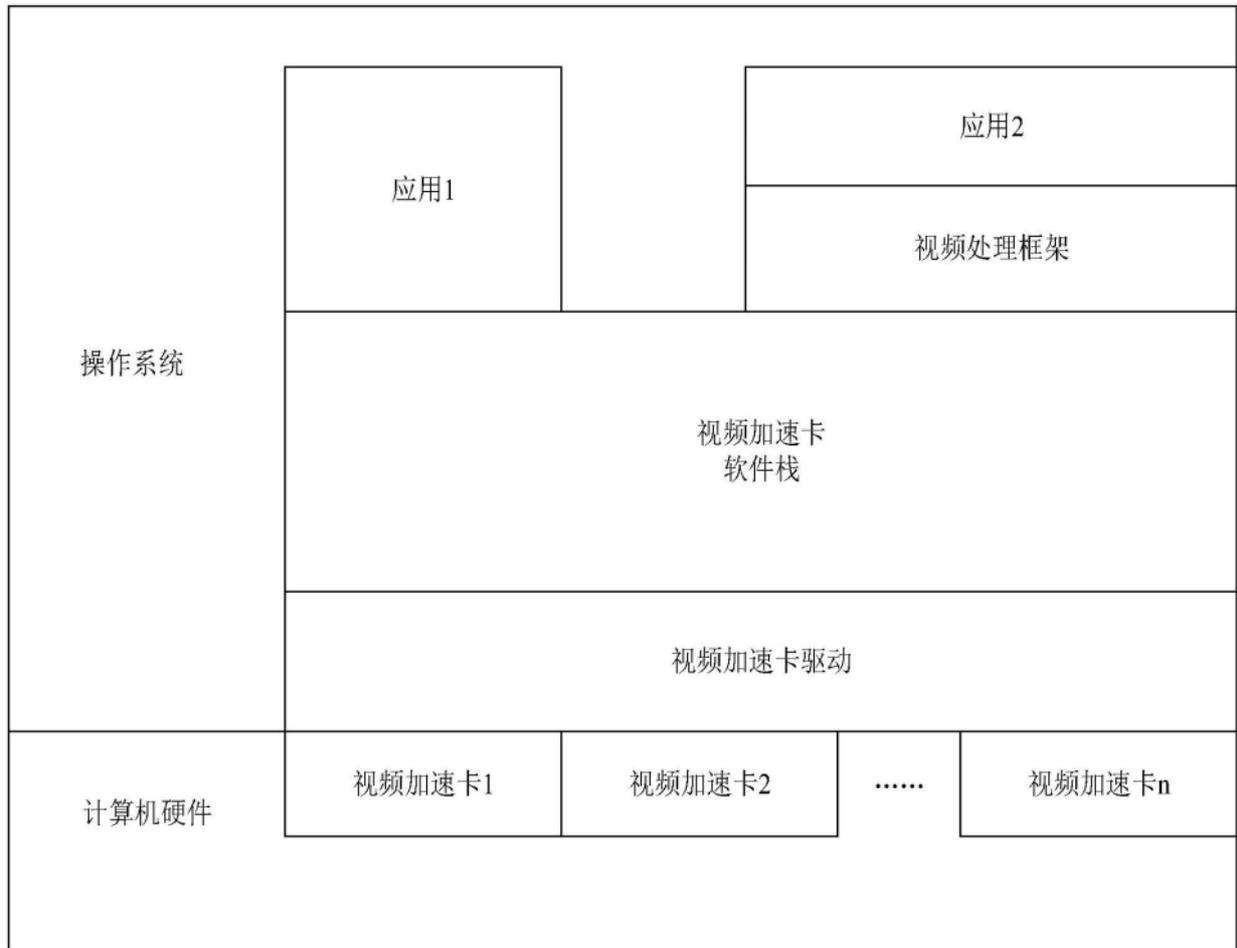


图1

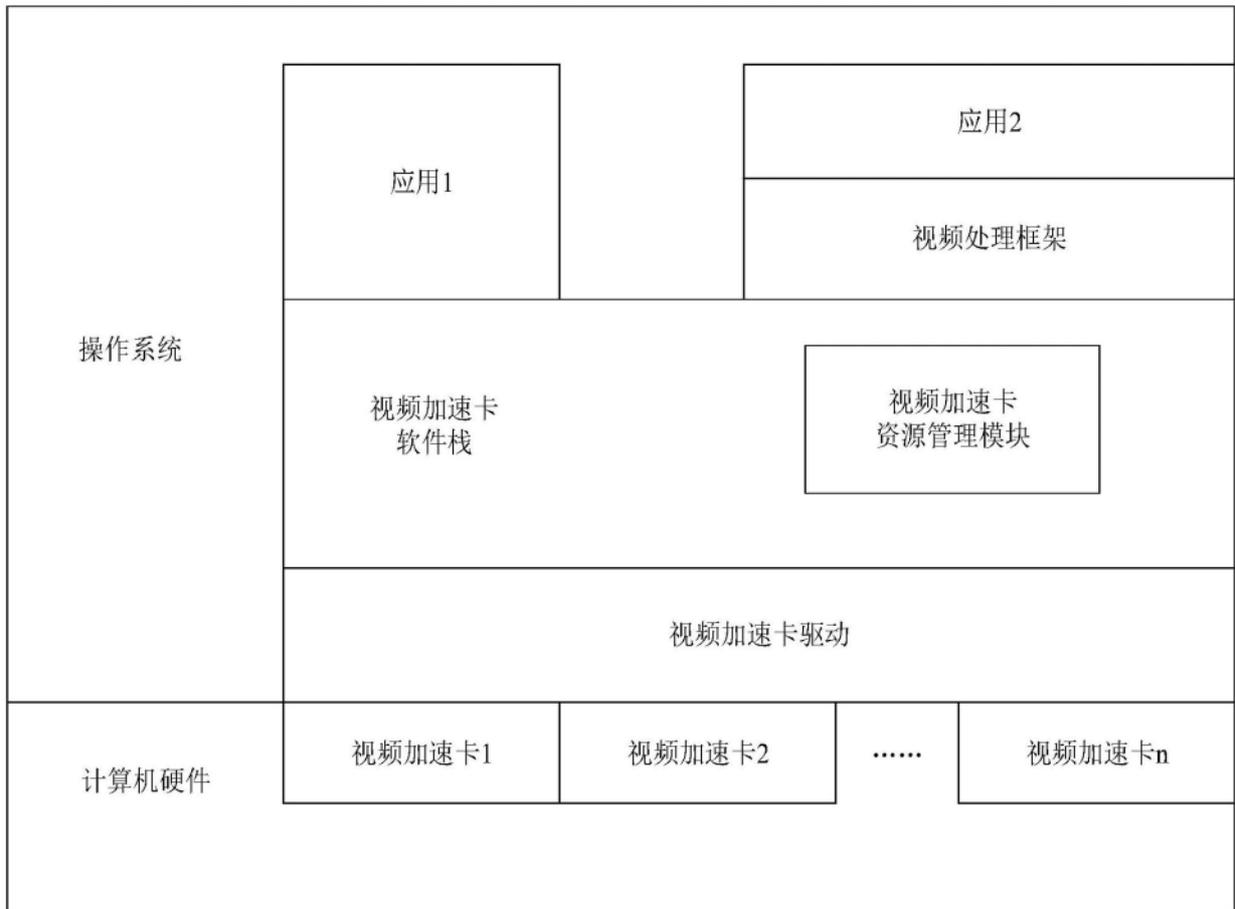


图2

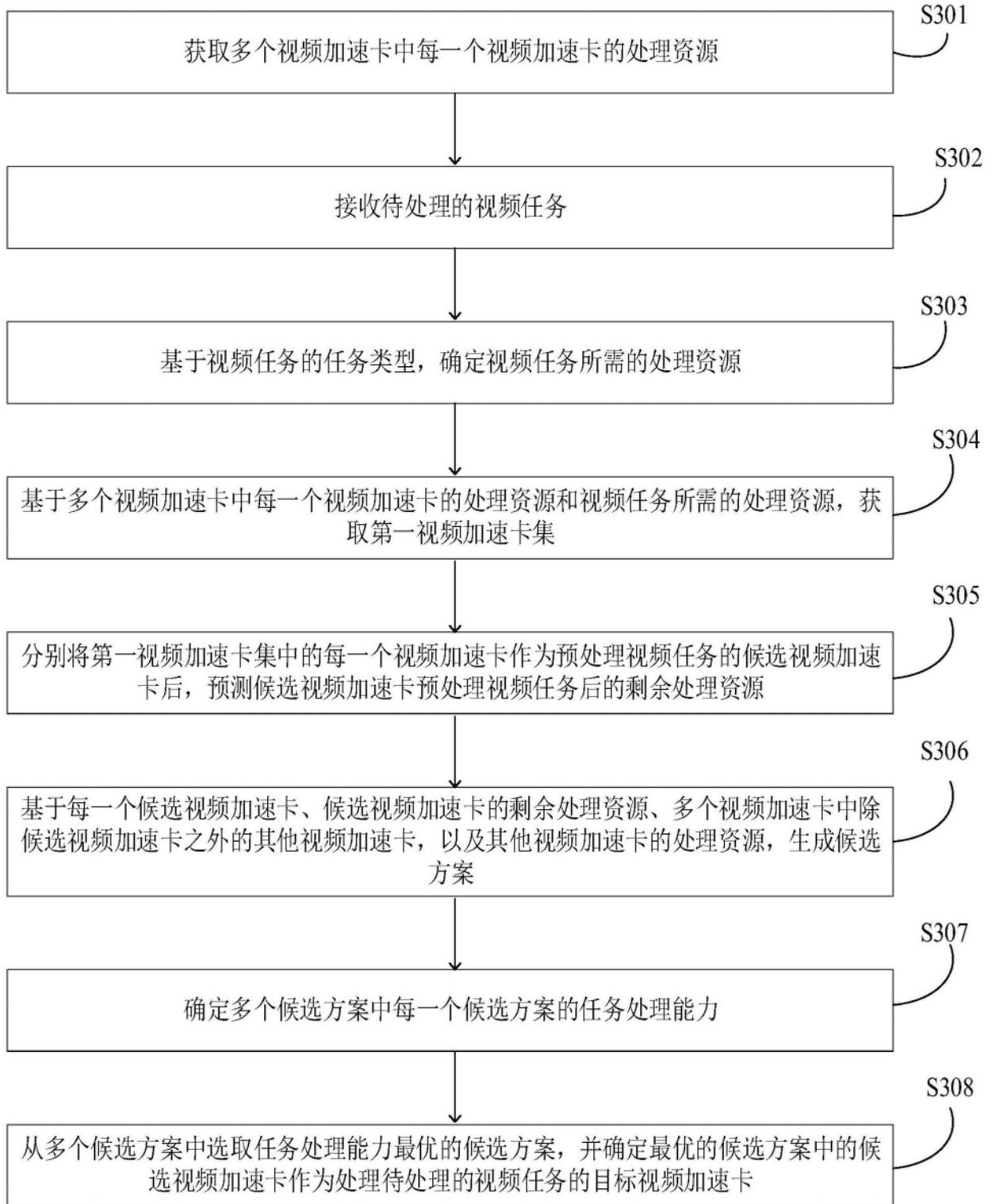


图3

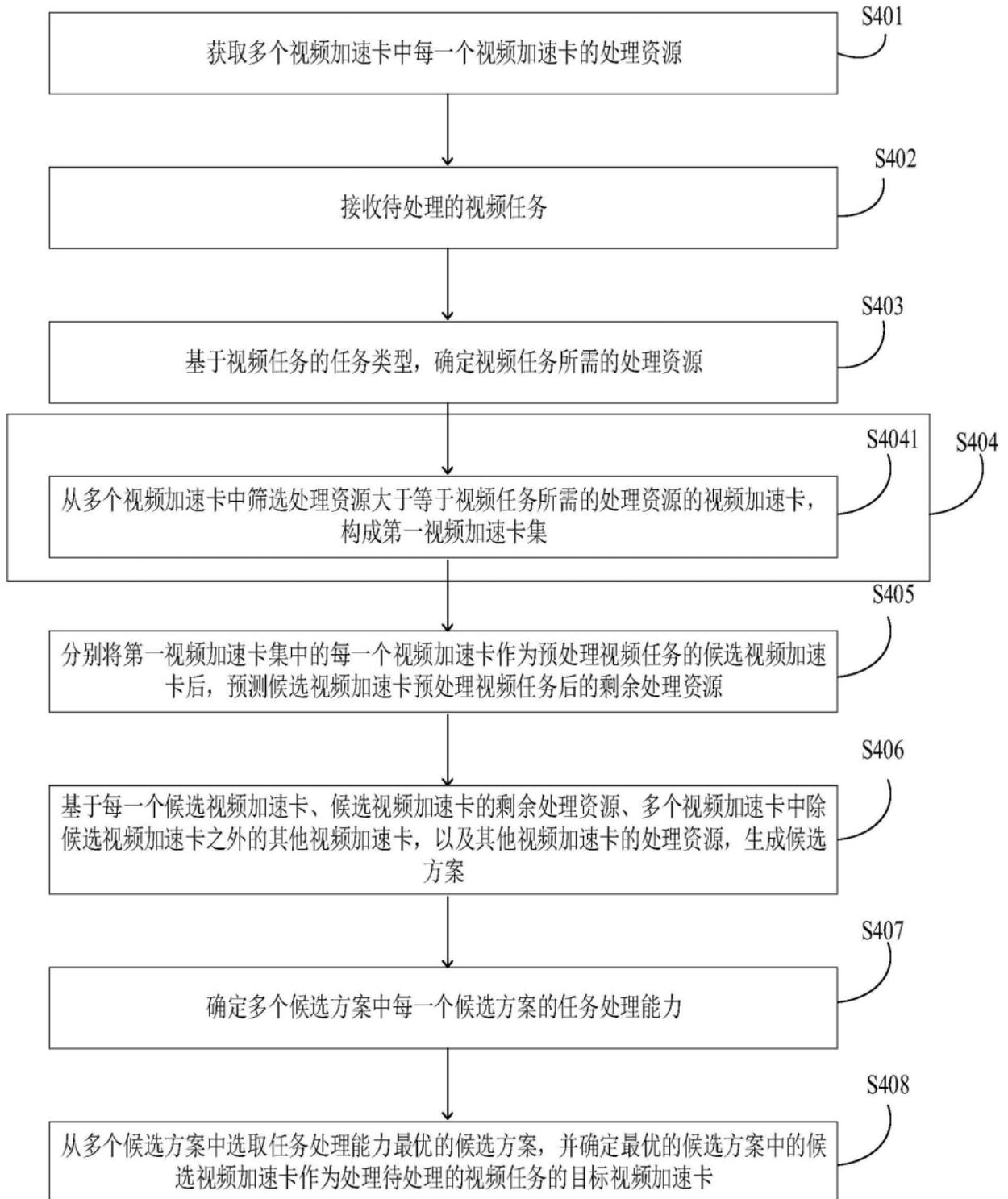


图4

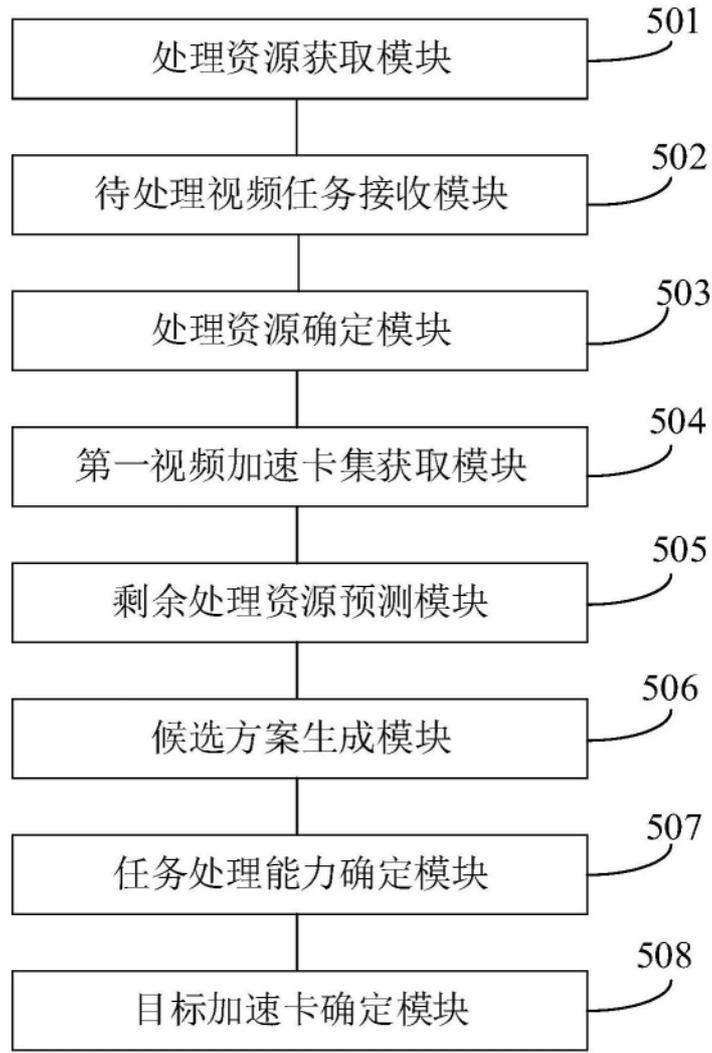


图5

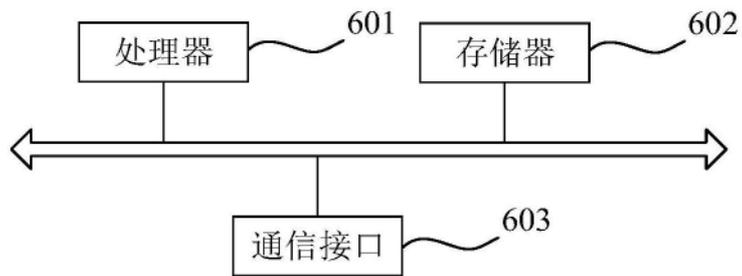


图6