



(12) 发明专利申请

(10) 申请公布号 CN 116055723 A

(43) 申请公布日 2023. 05. 02

(21) 申请号 202211690147.9

(22) 申请日 2022.12.27

(71) 申请人 上海哔哩哔哩科技有限公司

地址 200433 上海市杨浦区四平路1945号
1641室

(72) 发明人 蔡春磊 尹壮 王一 叶天晓

(74) 专利代理机构 北京市汉坤律师事务所
11602

专利代理师 魏小微 吴丽丽

(51) Int.Cl.

H04N 19/146 (2014.01)

H04N 19/142 (2014.01)

H04N 19/14 (2014.01)

H04N 19/503 (2014.01)

H04N 19/593 (2014.01)

权利要求书4页 说明书17页 附图5页

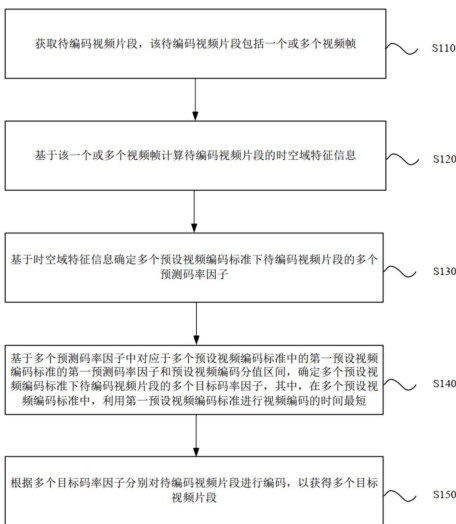
(54) 发明名称

视频编码方法及装置、电子设备和存储介质

(57) 摘要

本公开提供了一种视频编码方法及装置、电子设备和存储介质。该方法包括：获取待编码视频片段，该待编码视频片段包括一个或多个视频帧；基于该一个或多个视频帧计算待编码视频片段的时空域特征信息；基于时空域特征信息确定多个预设视频编码标准下待编码视频片段的多个预测码率因子；基于多个预测码率因子中对应于多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间，确定多个预设视频编码标准下待编码视频片段的多个目标码率因子，其中，在多个预设视频编码标准中，利用第一预设视频编码标准进行视频编码的时间最短；以及根据多个目标码率因子分别对待编码视频片段进行编码，以获得多个目标视频片段。

100



1. 一种视频编码方法,包括:

获取待编码视频片段,所述待编码视频片段包括一个或多个视频帧;

基于所述一个或多个视频帧计算所述待编码视频片段的时空域特征信息;

基于所述时空域特征信息确定多个预设视频编码标准下所述待编码视频片段的多个预测码率因子;

基于所述多个预测码率因子中对应于所述多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间,确定所述多个预设视频编码标准下所述待编码视频片段的多个目标码率因子,其中,在所述多个预设视频编码标准中,利用所述第一预设视频编码标准进行视频编码的时间最短;以及

根据所述多个目标码率因子分别对所述待编码视频片段进行编码,以获得多个目标视频片段。

2. 根据权利要求1所述的方法,其中,所述时空域特征信息包括时空复杂度信息,并且其中,基于所述一个或多个视频帧计算所述待编码视频片段的时空域特征信息包括:

基于所述一个或多个视频帧计算所述待编码视频片段的灰度共生信息和归一化信息作为所述时空复杂度信息,其中,所述灰度共生信息用于描述所述待编码视频片段的空间复杂度,并且所述归一化信息用于描述所述待编码视频片段的时域复杂度;以及

对所述灰度共生信息和所述归一化信息进行组合,以生成所述待编码视频片段的所述时空复杂度信息。

3. 根据权利要求1或2所述的方法,其中,所述时空域特征信息包括编码特征信息,并且其中,基于所述一个或多个视频帧计算所述待编码视频片段的时空域特征信息包括:

对所述一个或多个视频帧进行预处理,以生成新的视频帧序列,所述新的视频帧序列包括视频帧对集合;以及

针对所述视频帧对集合中的每一个视频帧对,分别进行帧内编码和帧间编码,以获得所述待编码视频片段的所述编码特征信息。

4. 根据权利要求3所述的方法,其中,对所述一个或多个视频帧进行预处理,以生成新的视频帧序列包括:

对所述一个或多个视频帧中除首帧和尾帧之外的视频帧进行复制,以生成一个或多个复制视频帧;以及

对所述一个或多个视频帧和所述一个或多个复制视频帧进行排序,使得所述一个或多个复制视频帧中的每一个复制视频帧位于对应的视频帧之后。

5. 根据权利要求3所述的方法,其中,针对所述视频帧对集合中的每一个视频帧对,分别进行帧内编码和帧间编码,以获得所述待编码视频片段的所述编码特征信息包括:

对所述视频帧对集合中的每一个视频帧对中的第一视频帧进行帧内编码,以获得帧内编码信息;

对所述视频帧对集合中的每一个视频帧对进行帧间编码,以获得帧间编码信息;以及基于所述帧内编码信息和所述帧间编码信息获得所述编码特征信息。

6. 根据权利要求3-5中任一项所述的方法,其中,所述编码特征信息基于以下项中的一个或多个:编码比特数、帧内预测模式占比值、帧间运动矢量幅度分布信息。

7. 根据权利要求1-6中任一项所述的方法,其中,基于所述时空域特征信息确定多个预

设视频编码标准下所述待编码视频片段的多个预测码率因子包括：

将所述待编码视频片段的所述时空域特征信息输入到第一码率因子预测模型，以经由所述第一码率因子预测模型确定所述多个预设视频编码标准下所述待编码视频片段的所述多个预测码率因子。

8. 根据权利要求7所述的方法，其中，基于所述多个预测码率因子中对应于所述多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间，确定所述多个预设视频编码标准下所述待编码视频片段的多个目标码率因子包括：

基于所述第一预测码率因子对所述待编码视频片段进行第一次编码，以获得所述第一预设视频编码标准下的第一预编码视频片段；

确定所述第一预编码视频片段的视频编码分值是否在所述预设视频编码分值区间内；以及

响应于确定所述第一预编码视频片段的视频编码分值在所述预设视频编码分值区间内：

确定所述多个预测码率因子作为所述多个目标码率因子。

9. 根据权利要求8所述的方法，其中，基于所述多个预测码率因子中对应于所述多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间，确定所述多个预设视频编码标准下所述待编码视频片段的多个目标码率因子还包括：

响应于确定所述第一预编码视频片段的视频编码分值不在所述预设视频编码分值区间内：

基于所述待编码视频片段的所述时空域特征信息、所述多个预测码率因子和所述第一预编码视频片段的视频编码分值，更新所述多个预测码率因子以获得所述多个目标码率因子。

10. 根据权利要求9所述的方法，其中，基于所述待编码视频片段的所述时空域特征信息、所述多个预测码率因子和所述第一预编码视频片段的视频编码分值，更新所述多个预测码率因子以获得所述多个目标码率因子包括：

将所述待编码视频片段的所述时空域特征信息、所述多个预测码率因子和所述第一预编码视频片段的视频编码分值输入到所述第二码率因子预测模型，以经由所述第二码率因子预测模型确定所述多个预设视频编码标准下所述待编码视频片段的所述多个目标码率因子。

11. 根据权利要求10所述的方法，其中，所述第一码率因子预测模型通过以下操作被训练：

针对所述多个预设视频编码标准中的每一个预设视频编码标准：

获取样本视频片段的样本时空域特征信息以及该预设视频编码标准下所述样本视频片段的样本码率因子；

将所述样本时空域特征信息输入到所述第一码率因子预测模型，以经由所述第一码率因子预测模型确定该预设视频编码标准下所述样本视频片段的样本预测码率因子；以及

基于所述样本码率因子和所述样本预测码率因子来计算第一模型损失值；以及基于所述每一个预设视频编码标准下的所述第一模型损失值来调整所述第一码率因子预测模型的参数，直到达到第一模型训练停止条件。

12. 根据权利要求11所述的方法, 其中, 所述第二码率因子预测模型通过以下操作被训练:

将所述样本时空域特征信息、所述每一个预设视频编码标准下的所述样本预测码率因子、和与所述第一预设视频编码标准下的样本预测码率因子相对应的第一样本视频编码分值输入到所述第二码率因子预测模型, 以经由所述第二码率因子预测模型确定所述多个预设视频编码标准下所述样本视频片段的多个样本目标码率因子;

针对所述多个预设视频编码标准中的所述每一个预设视频编码标准, 基于所述样本码率因子和对应于该预设视频编码标准的样本目标码率因子来计算第二模型损失值; 以及

基于所述每一个预设视频编码标准下的所述第二模型损失值来调整所述第二码率因子预测模型的参数, 直到达到第二模型训练停止条件。

13. 根据权利要求1所述的方法, 其中, 获取待编码视频片段包括:

获取待编码视频;

对所述待编码视频进行场景检测;

基于场景检测结果将所述待编码视频分成一个或多个子视频片段; 以及

识别所述一个或多个视频片段中的每一个视频子片段作为待编码视频片段。

14. 根据权利要求13所述的方法, 还包括:

针对所述多个预设视频编码标准中的每一个预设视频编码标准, 获取该预设视频编码标准下的所有目标视频片段; 以及

基于所述一个或多个子视频片段在所述待编码视频中的顺序, 对所述所有目标视频片段进行组合, 以获得该预设视频编码标准下对应于所述待编码视频的目标视频。

15. 一种视频编码装置, 包括:

获取模块, 被配置为获取待编码视频片段, 所述待编码视频片段包括一个或多个视频帧;

计算模块, 被配置为基于所述一个或多个视频帧计算所述待编码视频片段的时空域特征信息;

确定预测码率因子模块, 被配置为基于所述时空域特征信息确定多个预设视频编码标准下所述待编码视频片段的多个预测码率因子;

确定目标码率因子模块, 被配置为基于所述多个预测码率因子中对应于所述多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间, 确定所述多个预设视频编码标准下所述待编码视频片段的多个目标码率因子, 其中, 在所述多个预设视频编码标准中, 利用所述第一预设视频编码标准进行视频编码的时间最短; 以及

编码模块, 被配置为根据所述多个目标码率因子分别对所述待编码视频片段进行编码, 以获得多个目标视频片段。

16. 一种电子设备, 包括:

至少一个处理器; 以及

与所述至少一个处理器通信连接的至少一个存储器,

其中, 所述至少一个存储器存储有计算机程序, 所述计算机程序在被所述至少一个处理器执行时实现如权利要求1-14中任一项所述的方法。

17.一种存储有计算机程序的非瞬时计算机可读存储介质,其中,所述计算机程序在被处理器执行时实现如权利要求1-14中任一项所述的方法。

18.一种计算机程序产品,包括计算机程序,其中,所述计算机程序在被处理器执行时实现如权利要求1-14中任一项所述的方法。

视频编码方法及装置、电子设备和存储介质

技术领域

[0001] 本公开涉及计算机技术领域，具体涉及一种视频编码方法及装置、电子设备、计算机可读存储介质和计算机程序产品。

背景技术

[0002] 随着互联网技术的发展，视频平台也得到了迅猛发展，越来越多的用户分享、观看视频。对于平台而言，随着用户数量的不断增长，用于视频传输的带宽成本也越来越大。为了降低该成本，需要压缩性能更好的转码系统，在保证视频质量的同时，显著降低压缩后视频的大小，从而降低传输所需的网络流量。

[0003] 在此部分中描述的方法不一定是之前已经设想到或采用的方法。除非另有指明，否则不应假定此部分中描述的任何方法仅因其包括在此部分中就被认为是现有技术。类似地，除非另有指明，否则此部分中提及的问题不应认为在任何现有技术中已被公认。

发明内容

[0004] 本公开提供了一种视频编码方法及装置、电子设备、计算机可读存储介质和计算机程序产品。

[0005] 根据本公开的一方面，提供一种视频编码方法，包括：获取待编码视频片段，该待编码视频片段包括一个或多个视频帧；基于该一个或多个视频帧计算待编码视频片段的时空域特征信息；基于时空域特征信息确定多个预设视频编码标准下待编码视频片段的多个预测码率因子；基于多个预测码率因子中对应于多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间，确定多个预设视频编码标准下待编码视频片段的多个目标码率因子，其中，在多个预设视频编码标准中，利用第一预设视频编码标准进行视频编码的时间最短；以及根据多个目标码率因子分别对待编码视频片段进行编码，以获得多个目标视频片段。

[0006] 根据本公开的另一方面，还提供一种视频编码装置，包括：获取模块，被配置为获取待编码视频片段，该待编码视频片段包括一个或多个视频帧；计算模块，被配置为基于该一个或多个视频帧计算待编码视频片段的时空域特征信息；确定预测码率因子模块，被配置为基于时空域特征信息确定多个预设视频编码标准下待编码视频片段的多个预测码率因子；确定目标码率因子模块，被配置为基于多个预测码率因子中对应于多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间，确定多个预设视频编码标准下待编码视频片段的多个目标码率因子，其中，在多个预设视频编码标准中，利用第一预设视频编码标准进行视频编码的时间最短；以及编码模块，被配置为根据多个目标码率因子分别对待编码视频片段进行编码，以获得多个目标视频片段。

[0007] 根据本公开的另一方面，还提供一种电子设备，包括：至少一个处理器；以及与所述至少一个处理器通信连接的至少一个存储器，其中，所述至少一个存储器存储有计算机程序，所述计算机程序在被所述至少一个处理器执行时实现上述视频编码方法。

[0008] 根据本公开的另一方面,还提供一种存储有计算机程序的非瞬时计算机可读存储介质,其中,所述计算机程序在被处理器执行时实现上述视频编码方法。

[0009] 根据本公开的另一方面,还提供一种计算机程序产品,包括计算机程序,其中,所述计算机程序在被处理器执行时实现上述视频编码方法。

[0010] 根据本公开的一个或多个实施例,基于待编码视频片段的时空域特征信息对多个预设视频编码标准下的码率因子进行第一次预测,并且基于预测得到的码率因子和预设视频编码分值对多个预设视频编码标准下的码率因子进行第二次预测,由此可以提高码率因子的准确度,从而使得视频编码的准确度也得到了有效提升,保证了视频质量。另外,由于可以同时多个预设视频编码标准下的码率因子进行预测,因此可以使得编码系统同时输出多路不同视频编码标准但均满足视频质量要求的编码视频,避免了针对多个视频编码标准重复执行预测操作,由此节约了时间成本。进一步地,对码率因子进行的第二次预测是基于第一次预测结果中对应于编码时间最短的视频编码标准的码率因子的,由于编码时间最短的视频编码标准的计算复杂度最低,因此该视频编码标准下的码率因子可以在一定程度上反应编码时间较长的其他视频编码标准下的码率因子的准确度,从而将不同视频编码标准关联起来,有利于进一步减少预测码率因子所需的计算量。

[0011] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0012] 附图示例性地示出了实施例并且构成说明书的一部分,与说明书的文字描述一起用于讲解实施例的示例性实施方式。所示出的实施例仅出于例示的目的,并不限制权利要求的范围。在所有附图中,相同的附图标记指代类似但不一定相同的要素。

[0013] 图1示出了根据本公开的一些实施例的视频编码方法的流程图;

[0014] 图2示出了根据本公开的一些实施例的训练第一码率因子预测模型的方法的流程图;

[0015] 图3示出了根据本公开的一些实施例的第一码率因子预测模型的结构框图;

[0016] 图4示出了根据本公开的一些实施例的确定多个预设视频编码标准下待编码视频片段的多个目标码率因子的流程图;

[0017] 图5示出了根据本公开的实施例的视频编码装置的结构框图;

[0018] 图6示出了能够用于实现本公开的实施例的示例性电子设备的结构框图。

具体实施方式

[0019] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本公开的范围。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0020] 在本公开中,除非另有说明,否则使用术语“第一”、“第二”等来描述各种要素不意图限定这些要素的位置关系、时序关系或重要性关系,这种术语只是用于将一个元件与另一元件区分开。在一些示例中,第一要素和第二要素可以指向该要素的同一实例,而在某些

情况下,基于上下文的描述,它们也可以指代不同实例。

[0021] 在本公开中对各种所述示例的描述中所使用的术语只是为了描述特定示例的目的,而并非旨在进行限制。除非上下文另外明确地表明,如果不特意限定要素的数量,则该要素可以是一个也可以是多个。此外,本公开中所使用的术语“和/或”涵盖所列出的项目中的任何一个以及全部可能的组合方式。

[0022] 随着互联网技术的发展,视频平台也得到了迅猛发展,越来越多的用户分享、观看视频。对于平台而言,随着用户数量的不断增长,用于视频传输的带宽成本也越来越大。为了降低该成本,需要压缩性能更好的转码系统,在保证视频质量的同时,显著降低压缩后视频的大小,从而降低传输所需的网络流量。

[0023] 目前,转码系统常见的编码模式包括固定量化参数模式、恒定码率模式、可变码率模式等,其中,最常用的编码模式是恒定码率模式。虽然在该模式下输出的视频流具有稳定的码率,可以改善播放卡顿的问题,但是,在该模式下由于对所有的视频使用同一个码率因子进行编码,因此,对于场景内容丰富的视频,无法确保所有视频的画质体验,也无法避免码率浪费。例如,对于场景复杂的视频,使用平均码率因子对视频进行编码,将使得输出视频的码率不足并导致画质明显下降,而对于场景简单的视频,使用平均码率因子对视频进行编码,将可能造成不必要的码率浪费。

[0024] 为了确保画质体验并避免码率浪费,还可以采用恒定质量编码模式。具体而言,可以设定一个目标质量,并使用编码器对视频进行编码,以获得满足该目标质量的视频。该编码模式的目标是使编码后视频的质量尽可能与设定的目标质量接近,并且编码后视频的质量与设定的目标质量越接近,编码效果约好。当实际质量低于目标质量时,将可能降低视频的画质体验,而当实际质量超过目标质量时,将可能造成码率浪费。

[0025] 但是在使用这种编码模式对视频进行编码之前,编码使用的码率因子与编码结果的实际质量之间的关系是未知的,即,无法通过设定的目标质量直接得到准确的码率因子。常见的解决方法为:

[0026] (1) 对视频进行多次编码,通过后验方式,经由搜索得到最优的码率因子,但是这种方法需要消耗巨大的计算资源;

[0027] (2) 对码率因子进行预测,例如,可以通过机器学习的方式,在对视频进行编码之前基于该视频的时空复杂度信息对码率因子进行预测。但是,该解决方法通常是一次预测和编码的方式,即,根据预设的目标,基于视频的时空复杂度信息对码率因子进行预测,并基于预测的码率因子对视频进行编码。这可能导致对码率因子的预测结果准确度较低,尤其对于场景复杂的视频或热点视频,可能无法稳定地确保画质体验并同时避免码率浪费。

[0028] 此外,发明人还发现,为了适配各种终端播放器,转码系统需要输出满足不同视频编码标准的视频码流。然而,如上所述,在现有技术中通常采用一次预测和编码的方式,这将导致针对每种视频编码标准,转码系统均需要进行一次码率因子的预测和视频的编码,从而增加了计算量和时间成本。

[0029] 有鉴于此,本公开的实施例提供了一种视频编码方法,该方法基于待编码视频片段的时空域特征信息对多个预设视频编码标准下的码率因子进行第一次预测,并且基于预测得到的码率因子和预设视频编码分值对多个预设视频编码标准下的码率因子进行第二次预测,由此可以提高码率因子的准确度,进而使得视频编码的准确度也得到了有效提升,

保证了视频质量。另外,由于可以同时多个预设视频编码标准下的码率因子进行预测,因此可以使得编码系统同时输出多路不同视频编码标准但均满足视频质量要求的编码视频,避免了针对多个视频编码标准重复执行预测操作,由此节约了时间成本。

[0030] 进一步地,在该方法中,对码率因子进行的第二次预测是基于第一次预测结果中对应于编码时间最短的视频编码标准的码率因子的。由于编码时间最短的视频编码标准的计算复杂度最低,因此该视频编码标准下的码率因子可以在一定程度上反应编码时间较长的其他视频编码标准下的码率因子的准确度,从而将不同视频编码标准关联起来,有利于进一步减少预测码率因子所需的计算量。

[0031] 下面将结合附图详细描述本公开的实施例。

[0032] 图1示出了根据本公开的一些实施例的视频编码方法100的流程图。如图1所示,方法100可以包括:步骤S110、获取待编码视频片段,该待编码视频片段包括一个或多个视频帧;步骤S120、基于该一个或多个视频帧计算待编码视频片段的时空域特征信息;步骤S130、基于时空域特征信息确定多个预设视频编码标准下待编码视频片段的多个预测码率因子;步骤S140、基于多个预测码率因子中对应于多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间,确定多个预设视频编码标准下待编码视频片段的多个目标码率因子,其中,在多个预设视频编码标准中,利用第一预设视频编码标准进行视频编码的时间最短;以及步骤S150、根据多个目标码率因子分别对待编码视频片段进行编码,以获得多个目标视频片段。

[0033] 通过利用待编码视频片段的时空域特征信息确定多个预设视频编码标准下的预测码率因子,并且利用预测码率因子和预设视频编码分值确定多个预设视频编码标准下的目标码率因子,可以实现对码率因子的修正,由此提高了码率因子的准确度,进而使得视频编码的准确度也得到了有效提升,保证了视频质量。另外,由于可以同时多个预设视频编码标准下的码率因子进行预测,因此可以使得编码系统同时输出多路不同视频编码标准但均满足视频质量要求的编码视频,避免了针对多个视频编码标准重复执行预测操作,由此节约了时间成本。进一步地,对码率因子的修正是基于第一次预测结果中对应于编码时间最短的视频编码标准的码率因子的。由于编码时间最短的视频编码标准的计算复杂度最低,因此该视频编码标准下的码率因子可以在一定程度上反应编码时间较长的其他视频编码标准下的码率因子的准确度,从而将不同视频编码标准关联起来,实现了针对不同视频编码标准的码率因子的联合预测,进一步减少了预测码率因子所需的计算量。

[0034] 根据本公开的一些实施例,步骤S110、获取待编码视频片段可以包括:获取待编码视频;对所述待编码视频进行场景检测;基于场景检测结果将所述待编码视频分成一个或多个子视频片段;以及识别所述一个或多个视频片段中的每一个视频子片段作为待编码视频片段。

[0035] 根据一些实施例,可以从适当的存储装置(本地的和/或远程的)读取已存储或缓存的待编码视频。可替代地,也可以经由有线或无线通信链路从外部其他设备接收待编码视频。待编码视频可以指任何一段完整的视频文件。例如,其可以是用户自己录制的视频文件,也可以是用户从其他视频文件中截取的视频文件,也可以是用户由多个视频文件为基础创作的视频文件,本公开要求保护的的主题的范围在这方面不受限制。

[0036] 在获取到待编码视频后,可以对该待编码视频进行场景检测,根据检测结果对待

编码视频进行分割,并且将分割出的每个视频片段均作为待编码视频片段。

[0037] 根据一些实施例,可以基于开源编码器对待编码视频进行场景检测。例如,在响应于接收到用户上传的视频V而确定视频V为待编码视频的场景中,可以调用开源编码器x264对视频V进行编码,并打开场景检测开关,得到对应的场景切换检测结果。根据该检测结果,可以将视频V分割成t个独立的单镜头视频片段(V1、V2……Vt),并且每个单镜头视频片段均可以作为一个待编码视频片段。

[0038] 根据另一些实施例,可以基于分帧操作来对待编码视频进行分割。继续上述待编码视频为V的示例,可以对该待编码视频V进行分帧操作以获得多个视频帧,然后针对其中的每一个视频帧确定该视频帧与相邻视频帧之间的相似度,并且基于计算出的相似度确定该视频帧是否为待处理视频的用于区分相邻镜头的边界视频帧。应当理解,可以利用任何合适的解码器或其他技术(例如,Open CV、FFmpeg等)来进行分帧操作,也可以利用任何合适的计算方法还确定相邻视频帧之间的相似度(例如,计算帧差、灰度直方图等),本公开要求保护的的主题的范围在这方面不受限制。

[0039] 相较于完整的视频文件,一个单镜头视频片段中所展示的内容通常是相似的并且连贯性更强,因此,在对视频进行编码时,针对每个单镜头视频片段使用特定的、相同的码率因子将更加合理,从而确保了各个场景下稳定的画质体验。

[0040] 根据本公开的一些实施例,时空域特征信息可以包括时空复杂度信息,并且其中,步骤S120、基于一个或多个视频帧计算待编码视频片段的时空域特征信息可以包括:基于该一个或多个视频帧计算待编码视频片段的灰度共生信息和归一化信息作为时空复杂度信息,其中,灰度共生信息用于描述待编码视频片段的空间复杂度,并且归一化信息用于描述待编码视频片段的时域复杂度;以及对灰度共生信息和归一化信息进行组合,以生成待编码视频片段的时空复杂度信息。

[0041] 灰度共生信息(也被称为灰度共生矩阵,Gray Level Co-occurrence Matrix, GLCM),是一种用于描述视频空间纹理复杂度的特征。在一些实施例中,针对待编码视频片段的一个或多个视频帧中的每一个视频帧,可以计算获得一个 $m \times m$ 的二维GLCM矩阵,其中该GLCM矩阵中每个元素的取值范围是0-1。然后,可以针对该矩阵进一步求解信息熵(entropy)、对比度(contrast)、相关度(inverse different moment)、能量、自相关(correlation)等5个特征值,并且对所有视频帧对应的这些特征值计算均值、方差、偏度和峰度等4个统计值,例如,计算每个待处理视频帧的信息熵的均值、方差、偏度和峰度等。最终将获得共计20(5×4)个特征值。进一步地,为了考量视频多尺度特性,可以将每一个视频帧缩小4倍和16倍,并且分别重复上述步骤(即,进行两次GLCM特征提取)。因此,针对待编码视频片段的视频帧最终将获得60个特征值,即60个GLCM特征值。

[0042] 归一化信息(也被称为归一化相关系数,Normalized Correlation Coefficient, NCC),是一种用于描述时域复杂度的特征。在一些实施例中,针对待编码视频片段的一个或多个视频帧中的每对相邻的视频帧,可以计算NCC以获得长宽和原视频帧一样大的二维相关系数矩阵,其中,矩阵中每个元素的取值范围是0-1,表示相邻两个视频帧对应位置的时域相关性。对于包含n个视频帧的待编码视频片段,将存在n-1个相邻视频帧对,因此可以获得n-1个相关系数矩阵。然后,计算这n-1个相关系数矩阵的均值、方差、峰度、偏度和信息熵等共计5个特征值,并且对于每个特征值,再从时域上计算均值和方差,最终可以得到10个

(5×2)特征值,即10个NCC特征值。

[0043] 在通过例如如上所述的方法计算出GLCM特征值和NCC特征值之后,可以对两者进行组合,即可获得70个特征值作为待编码视频片段的时空复杂度信息,其中,该时空复杂度信息可以通过向量形式进行表示。

[0044] 由于灰度共生信息和归一化信息可以分别从空间和时间角度来表征待编码视频的复杂度,因此,基于这些空间复杂度信息来预测不同视频编码标准下待编码视频片段的多个码率因子,可以使预测得到的码率因子将更为准确,从而有利于确保输出视频的质量。

[0045] 根据本公开的一些实施例,时空域特征信息可以包括编码特征信息,并且其中,步骤S120、基于一个或多个视频帧计算待编码视频片段的时空域特征信息可以包括:对一个或多个视频帧进行预处理,以生成新的视频帧序列,其中,新的视频帧序列包括视频帧对集合;以及针对视频帧对集合中的每一个视频帧对,分别进行帧内编码和帧间编码,以获得待编码视频片段的编码特征信息。

[0046] 通过对待编码视频中的每一视频帧进行帧内编码或帧间编码,可以从另一个维度反映待编码视频片段的时空域特征信息,即空间编码复杂度和时域编码复杂度,这是因为帧内编码采用了待编码视频帧的图像内部数据,而帧间编码采用了待编码视频帧与其相邻视频帧的图像数据,由此可以更为全面的表征待编码视频片段。利用该编码特征信息与上述由GLCM特征值和NCC特征值组成的时空复杂度信息的组合来预测不同视频编码标准下待编码视频片段的多个码率因子,可以进一步提高预测结果的准确度,这对于期望更高质量、编码更为精细的热点视频将尤为有益。

[0047] 根据本公开的一些实施例,在步骤S120中,对一个或多个视频帧进行预处理,以生成新的视频帧序列可以包括:对一个或多个视频帧中除首帧和尾帧之外的视频帧进行复制,以生成一个或多个复制视频帧;以及对一个或多个视频帧和一个或多个复制视频帧进行排序,使得一个或多个复制视频帧中的每一个复制视频帧位于对应的视频帧之后。

[0048] 根据一些实施例,在生成对应于除首帧和尾帧之外的视频帧的一个或多个复制帧后,可以将每一个复制帧放在对应的原始视频帧之后,以生成新的视频帧序列。举例而言,假设待编码视频片段包括n个视频帧并且该n个视频帧组成的序列为 $(f_1, f_2, f_3, \dots, f_{n-2}, f_{n-1}, f_n)$,则除首帧 f_1 和尾帧 f_n 之外,分别对视频帧 f_2 至 f_{n-1} 中的每个视频帧进行复制,以生成新的视频帧序列 $(f_1, f_2, f_2, f_3, f_3, \dots, f_{n-2}, f_{n-2}, f_{n-1}, f_{n-1}, f_n)$,其中, (f_1, f_2) 、 (f_2, f_3) …… (f_{n-2}, f_{n-1}) 、 (f_{n-1}, f_n) 为视频帧对。

[0049] 与未生成复制帧的视频帧序列相比,通过对除首帧和尾帧之外的视频帧进行复制并生成一个或多个有序视频帧对,可以确保每个视频帧内以及相邻视频帧间的数据均能够得到利用,保留了数据完整性,有利于实现对待编码视频片段的每一视频帧的更为准确的帧内编码和帧间编码,从而提高预测出的码率因子的准确度。

[0050] 根据本公开的另一些实施例,对一个或多个视频帧进行预处理还可以包括例如对每个视频帧进行裁剪以满足尺寸要求、去除视频帧中的非必要(干扰)特征等。由此,可以加快计算和预测速度,快速获得编码特征信息并且预测待编码视频的码率因子,同时也有利于提高预测出的码率因子的准确度。

[0051] 在对待编码视频片段进行预处理获得视频帧对集合后,可以采用例如显卡内置的硬件H.264编码器对其进行预编码。预编码可以采用例如固定量化参数模式,并且任选地,

量化参数可以为28、29、30等任何适合的值。编码帧结构可以采用例如IP结构,其中,I表示帧内编码并且P表示帧间编码。在这种编码模式下,除首帧和尾帧之外,待编码视频片段的每个视频帧均被编码了两次(即一次采用帧内编码I,一次采用帧间编码P),其中,帧内编码可以反映每一个视频帧的空间编码复杂度,并且帧间编码可以反映每一个视频帧的时域编码复杂度。在对每一个视频帧对集合进行IP编码之后,可以获得待编码视频片段的编码特征信息。

[0052] 根据本公开的一些实施例,针对视频帧对集合中的每一个视频帧对,分别进行帧内编码和帧间编码,以获得待编码视频片段的编码特征信息可以包括:对视频帧对集合中的每一个视频帧对中的第一视频帧进行帧内编码,以获得帧内编码信息;对视频帧对集合中的每一个视频帧对进行帧间编码,以获得帧间编码信息;以及基于帧内编码信息和帧间编码信息获得编码特征信息。

[0053] 待编码视频片段的每个视频帧对中存在两个视频帧,在本申请中,第一视频帧指代视频帧对中在先的视频帧。例如,对于视频帧对 (f_1, f_2) ,则 f_1 为第一视频帧;对于待编码视频帧对 (f_2, f_3) ,则 f_2 为第一视频帧,以此类推。

[0054] 根据一些示例,对每一个视频帧对中的第一视频帧进行帧内编码,以获得帧内编码信息可以包括:在对每个第一视频帧进行帧内编码后提取每个帧内编码的特征值,例如编码比特数、帧内预测模式的占比值等34个特征值;以及计算所有帧内编码的对应特征值的均值、方差、偏度、峰度和信息熵5个统计值,以得到170(即 34×5)个帧内编码信息。

[0055] 对于每个视频帧对,还可以提取每个视频帧对的帧间编码信息,例如对于视频帧对 (f_1, f_2) ,则计算 f_1 与 f_2 的帧间编码信息;对于视频帧对 (f_2, f_3) ,则计算 f_2 与 f_3 的帧间编码信息,依此类推。

[0056] 根据一些示例,对每个视频帧对进行帧间编码,以获得帧间编码信息可以包括:在对每个视频帧对进行帧间编码后提取每个帧间编码的特征值,例如编码比特数、帧间运动矢量幅度分布等26个特征值;以及计算所有帧间编码的对应特征值计算均值、方差、偏度、峰度和信息熵5个统计值,以得到130(即 26×5)个帧间编码信息。

[0057] 继续上述示例,将170个帧内编码信息和130个帧间编码信息进行拼接,从而获得300个待编码视频片段的编码特征信息。与空间复杂度信息类似,编码特征信息也可以通过向量形式进行表示。

[0058] 将理解的是,在上述实施例中,虽然帧内编码的步骤被描述为在帧间编码的步骤之后执行,但是这不应解释为要求帧内编码和帧间编码必须以此特定顺序执行。例如,帧内编码的步骤可以和帧间编码的步骤并行执行、或者帧内编码的步骤可以在帧间编码的步骤之后执行。

[0059] 还将理解的是,在上述实施例中,虽然帧内编码信息和帧间编码信息被描述为诸如编码比特数、帧内预测模式的占比值、帧间运动矢量幅度分布等特征值的均值、方差、偏度、峰度和信息熵,但是也可以直接确定这些特征值作为帧内编码信息或帧间编码信息。

[0060] 根据本公开的一些实施例,步骤S130、基于时空域特征信息确定多个预设视频编码标准下待编码视频片段的多个预测码率因子可以包括:将待编码视频片段的时空域特征信息输入到第一码率因子预测模型,以经由第一码率因子预测模型确定多个预设视频编码标准下待编码视频片段的多个预测码率因子。

[0061] 码率因子(Rate Factor, RF)是指恒定码率因子编码模式下的编码参数,用于控制编码文件的码率和质量。码率因子的取值通常在0-50的范围之间。码率因子的取值越大,则码率越低,编码质量也越差,反之,码率因子的取值越小,则码率越高,编码质量也越好。

[0062] 多个预设视频编码标准的示例可以包括H.264视频编码标准、H.265视频编码标准、H.266视频编码标准、AV1视频编码标准。将理解的是,预设视频标准还可以包括任何其他现有的或者随着技术的发展而生成的视频编码标准。还将理解的是,多个预设视频编码标准的数量可以满足实际需求的任意整数值。本公开要求保护的的主题的范围在这两方面均不受限制。

[0063] 根据本公开的一些实施例,在时空域特征信息仅包括时空复杂度信息的情况下,可以将待编码视频片段的灰度共生信息和归一化信息组合后的特征信息输入到第一码率因子预测模型,其中,该第一码率因子预测模型被训练用于根据时空复杂度信息来确定多个预设视频编码标准下每个待编码视频片段的预测码率因子。

[0064] 根据本公开的另一些实施例,在时空域特征信息仅包括编码特征信息的情况下,可以将待编码视频片段的帧内编码信息和帧间编码信息组合后的特征信息输入到第一码率因子预测模型,其中,该第一码率因子预测模型被训练用于根据编码特征信息来确定多个预设视频编码标准下每个待编码视频片段的预测码率因子。

[0065] 根据本公开的又一些实施例,在时空域特征信息包括时空复杂度信息和编码特征信息两者的情况下,可以将时空复杂度信息和编码特征信息进行组合以生成以一组特征向量进行表示的特征信息。继续上述时空复杂度信息为包含有70个特征值的向量、编码特征信息为包含有300个特征值的向量的示例,则将这两个向量进行组合后将获得包含有370个特征值的特征向量,从而更为完整地表征待编码视频片段的时空特性。在这种情况下,可以将生成的一组特征向量输入到第一码率因子预测模型,其中,该第一码率因子预测模型被训练用于根据时空复杂度信息和编码特征信息两者来确定多个预设视频编码标准下每个待编码视频片段的预测码率因子。

[0066] 通过利用经训练的第一码率因子预测模型来预测多个预设视频编码标准下每个待编码视频片段的预测码率因子,可以直接确定满足质量要求(例如,编码后视频片段的视频编码分值较高)的码率因子,而无需在每次编码时经过多次编码和搜索获得,从而节约了大量时间成本。同时,在样本数据较为充足的情况下,也可以在一定程度上确保预测出的码率因子的准确度。

[0067] 图2示出了根据本公开的一些实施例的训练第一码率因子预测模型的方法200的流程图。如图2所示,第一码率因子预测模型可以通过方法200的以下步骤被训练:步骤S210、针对多个预设视频编码标准中的每一个预设视频编码标准,获取样本视频片段的样本时空域特征信息以及该预设视频编码标准下样本视频片段的样本码率因子;步骤S220、将样本时空域特征信息输入到第一码率因子预测模型,以经由第一码率因子预测模型确定该预设视频编码标准下样本视频片段的样本预测码率因子;步骤S230、基于样本码率因子和样本预测码率因子来计算第一模型损失值;以及步骤S240、基于每一个预设视频编码标准下的第一模型损失值来调整第一码率因子预测模型的参数,直到达到第一模型训练停止条件。

[0068] 通过利用样本数据对第一码率因子预测模型进行预训练和调整,可以提高预测出

的码率因子的准确度。

[0069] 根据一些实施例,在步骤S210中,可以采用与步骤S110中获取待编码视频片段类似的方法来获取样本视频片段,并且采用与步骤S120中计算待编码视频片段的时空域特征信息类似的方法来计算样本时空域特征信息,故此不再赘述。为了确保训练后的第一码率因子预测模型的准确度,可以在兼顾计算资源和时间成本的前提下获取更多的样本视频片段,样本视频片段的数量例如可以为至少5万条、至少10万条等。

[0070] 根据一些实施例,在步骤S210中,样本码率因子可以为真实码率因子,即采用该码率因子对样本视频片段进行编码后得到的经编码视频的质量满足目标质量要求。经编码视频的质量可以例如通过视频编码分值(也被称为视频多维度混合评价因子,Video Multimethod Assessment Fusion,VMAF)来评价,VMAF的值越大,则表明编码后视频的质量越高,反之,VMAF的值越小,则表明编码后视频的质量越低,其中,VMAF的取值范围在通常情况下可以为0-100。真实码率因子可以通过如下步骤来获得:针对该视频编码标准,选择初始样本码率因子和预设样本视频编码分值区间;基于该初始样本码率因子,使用真实转码系统中的对应编码器(例如,H.264编码器、H.265编码器、H.266编码器、AV1编码器等)对样本视频片段进行编码,以获得该视频编码标准下的样本预编码视频片段;将获得的样本预编码视频片段的样本视频编码分值与预设样本视频编码分值区间进行比较;如果该样本视频编码分值在预设样本视频编码分值区间之内(即,编码后的视频满足目标质量要求),则表明该视频编码标准下初始样本码率因子是可行的,可以采用该初始样本码率因子作为真实码率因子;如果该样本视频编码分值在预设样本视频编码分值区间之外(即,编码后的视频未满足目标质量要求),则表明该视频编码标准下初始样本码率因子不可行,可以使用诸如二分法等搜索方法对初始样本码率因子进行调整,直到该样本视频编码分值在预设样本视频编码分值区间之内为止。通常情况下,针对同一样本视频片段,不同视频编码标准下的码率因子一般不同,因此,将针对每一视频编码标准重复上述步骤,以搜索获得多个视频编码标准下样本视频片段的样本码率因子(真实码率因子)。

[0071] 在上述实施例中,预设样本视频编码分值区间例如可以为基于VMAF的、用户根据实际需求自定义的一个数值范围(例如[91,93]、[90,95]),也可以为为了使第一码率因子预测模型适配于多种目标质量要求而关于每一个VMAF整数值设置的区间(例如[91.5,92.5]、[92.5,93.5]……[95.5,96.5]等)。本公开要求保护的的主题的范围在这方面不受限制。

[0072] 根据一些实施例,在步骤S220中,在初始阶段,可以针对多个视频编码标准选择第一码率因子预测模型的初始参数,并在此初始参数的基础上利用获取的样本时空域特征信息确定该多个视频编码标准下样本视频片段的多个样本预测码率因子。

[0073] 根据一些实施例,在步骤S230中,针对每一个预设视频编码标准,可以例如经由损失函数来基于对应的样本预测码率因子和样本预测码率因子来计算模型损失值。损失函数的示例包括但不限于交叉熵损失函数、最大损失函数、平均值损失函数、0-1损失函数等。将理解的是,在本公开中,还可以使用其他适合的方法来计算模型损失值,本公开要求保护的的主题的范围在这方面不受限制。

[0074] 根据一些实施例,在步骤S240中,在计算出每个预设视频编码标准下的模型损失值后,可以反向传播这些模型损失值,并调整第一码率因子预测模型的参数,然后重复步骤

S220-S240,直至达到模型训练停止条件。具体的,模型训练停止条件可以是模型损失值低于预设阈值,和/或模型训练的轮次达到了预设轮次。将理解的是,本公开对模型训练停止条件不做限定,可以根据实际需求来进行调整。

[0075] 图3示出了根据本公开的一些实施例的第一码率因子预测模型300的结构框图。如图3所述,该第一码率因子预测模型300可以包括批归一化模块310、第一注意力机制模块320、残差模块330、第二注意力机制模块340和全连接模块350,并且该第一码率因子预测模型300以待编码视频片段的时空域特征作为输入,输出不同视频编码标准下该待编码视频片段的预测码率因子。

[0076] 批归一化模块310用于将原始输入的特征归一化为具有均值为0、方差为1、高斯分布的特征,从而消除特征之间量纲差异带来的影响。然后,归一化后的特征被输入到第一注意力机制模块320,该模块可以为每个特征计算一个权重,其中,为具有更高辨识度的特征赋予数值更大的权重并且为信息量较低的特征赋予数值较小的权重,由此对特征进行自动筛选,从而提高预测结果的准确度。然后,经过筛选后的特征被输入到残差模块330,该模块可以包括多个卷积层,这些卷积层对作为输入的筛选后特征进行变化,同时未经变化的筛选后特征将跳过这些卷积层传输到后序层,然后对其整体进行处理以得到残差模块的输出结果,由此减少了信息丢失。残差模块330的输出可以进一步被输入到第二注意力机制模块340,以实现特征的进一步筛选,从而进一步提高预测结果的准确度。最后,经由全连接模块350可以对输入的特征进行聚合,以生成并输出预测码率因子。

[0077] 图4示出了根据本公开的一些实施例的确定多个预设视频编码标准下待编码视频片段的多个目标码率因子的流程图。如图4所示,步骤S140、基于多个预测码率因子中对应于多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间,确定多个预设视频编码标准下待编码视频片段的多个目标码率因子可以包括:步骤S442、基于第一预测码率因子对待编码视频片段进行第一次编码,以获得第一预设视频编码标准下的第一预编码视频片段;步骤S444、确定第一预编码视频片段的视频编码分值是否在预设视频编码分值区间内;以及步骤S446、响应于确定第一预编码视频片段的视频编码分值在预设视频编码分值区间内:确定多个预测码率因子作为多个目标码率因子。

[0078] 在本公开中,第一预设视频编码标准是指采用其进行视频编码时耗时最短的视频编码标准。举例而言,在视频编码标准H.264、H.265、H.266和AV1中,由于采用视频编码标准H.264进行视频编码时花费的时间最短,因此视频编码标准H.264为第一预设视频编码标准。

[0079] 在步骤S442中,除了第一预测码率因子之外,对待编码视频进行编码时利用的编码参数还可以包括但不限于preset、GOP(Group of Picture)、size等。由于码率因子对于编码后视频的质量起到重要作用,因此在使用本公开的方法对视频进行编码时,可以设定这些编码参数的值为固定常数(例如,默认预设值)。

[0080] 在步骤S444中,类似于预设样本视频编码分值区间的设定,预设视频编码分值区间也可以为用户根据实际需求自定义的一个数值范围,故在此不在赘述。

[0081] 由于编码时间最短的视频编码标准的计算复杂度最低,因此该视频编码标准下的码率因子可以在一定程度上反应编码时间较长的其他视频编码标准下的码率因子的准确

度。换言之,如果基于对应于第一预设视频编码标准(例如,在上述示例中为视频编码标准H.264)的第一预测码率因子进行编码的视频片段的视频编码分值在预设视频编码分值区间内(满足视频质量要求),则表明对应于其他预设视频编码标准(例如,在上述示例中为视频编码标准H.265、H.266和AV1)的其他预测码率因子的准确度也较高(例如,如果视频编码标准H.264下的经编码视频的视频编码分值有99%的概率在预设视频编码分值区间内,则视频编码标准H.265、H.266和AV1下的经编码视频的视频编码分值也有超过95%的概率在预设视频编码分值区间内)。利用这些预测码率因子编码后的视频片段的视频编码分值通常也在预设视频编码分值区间内(满足视频质量要求),从而可以将不同视频编码标准关联起来,有利于减少预测码率因子所需的计算量。在一些示例中,经过多次试验数据显示,当视频编码标准H.264下的预测码率因子满足视频质量要求时,视频编码标准H.265、H.266和AV1下各自的预测码率因子满足视频质量要求的概率为95%,因此可以将这些预测码率因子直接作为用来对待编码视频进行编码的目标码率因子。

[0082] 继续参考图4,根据本公开的另一些实施例,步骤S140、基于多个预测码率因子中对应于多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间,确定多个预设视频编码标准下待编码视频片段的多个目标码率因子还可以包括:步骤S448、响应于确定第一预编码视频片段的视频编码分值不在预设视频编码分值区间内:基于待编码视频片段的时空域特征信息、多个预测码率因子和第一预编码视频片段的视频编码分值,更新多个预测码率因子以获得多个目标码率因子。

[0083] 继续上述多个预设视频编码标准为H.264、H.265、H.266和AV1的示例,当基于视频编码标准H.264下的预测码率因子获得的预编码视频片段的视频编码分值不在预设视频编码分值区间内时,可以将视频编码标准H.265、H.266和AV1下各自的预测码率因子和视频编码标准为H.264下的预编码视频片段的视频编码分值作为反馈信息,连同待编码视频片段的时空域特征信息一起对这些视频编码标准下的码率因子进行第二次预测,以获得多个目标码率因子。

[0084] 通过大量的实验数据表明,经过第一次预测并且利用预测码率因子对待编码视频片段进行编码后,视频编码分值满足视频质量要求的概率为40%。因此,60%的待编码视频片段将经历上述第二次预测更新码率因子的操作。当第一预编码视频片段的视频编码分值不在预设视频编码分值区间内时,通过执行对不同视频编码标准下的码率因子的第二次预测,可以进一步提高目标码率因子的准确度,避免针对上述60%的待编码视频片段预测的码率因子误差较大而不能稳定地保障画质体验。

[0085] 此外,对第一次预测获得的预测码率因子的更新是基于第一预编码视频片段的视频编码分值,从而无需利用其他预设视频编码标准下的预测码率因子分别对待编码视频片段进行编码,这将大大节约编码所需的时间成本和计算资源。

[0086] 根据本公开的一些实施例,步骤S448、基于待编码视频片段的时空域特征信息、多个预测码率因子和第一预编码视频片段的视频编码分值,更新多个预测码率因子以获得多个目标码率因子可以包括:将待编码视频片段的时空域特征信息、多个预测码率因子和第一预编码视频片段的视频编码分值输入到第二码率因子预测模型,以经由第二码率因子预测模型确定多个预设视频编码标准下待编码视频片段的多个目标码率因子。

[0087] 根据本公开的一些实施例,类似于第一码率因子预测模型的训练方法,第二码率

因子预测模型可以通过以下操作被训练：将样本时空域特征信息、每一个预设视频编码标准下的样本预测码率因子、和与第一预设视频编码标准下的样本预测码率因子相对应的第一样本视频编码分值输入到第二码率因子预测模型，以经由第二码率因子预测模型确定多个预设视频编码标准下样本视频片段的多个样本目标码率因子；针对多个预设视频编码标准中的每一个预设视频编码标准，基于样本码率因子和对应于该预设视频编码标准的样本目标码率因子来计算第二模型损失值；以及基于第二模型损失值来调整第二码率因子预测模型的参数，直到达到第二模型训练停止条件。

[0088] 将理解的是，除了增加每一个预设视频编码标准下的样本预测码率因子和第一样本视频编码分值作为第二码率因子预测模型的输入，上述训练第二码率因子预测模型的方法中步骤与训练第一码率因子预测模型的方法200类似，故此不再赘述。

[0089] 相应地，与第一码率因子预测模型相比，除了增加每一个预设视频编码标准下的样本预测码率因子和第一样本视频编码分值作为输入，第二码率因子预测模型的结构框架也与第一码率因子预测模型的结构框架相同，故此也不再赘述。

[0090] 还将理解的是，对第二码率因子预测模型的训练是依赖于第一码率因子预测模型的训练结果的。因此，在完成对第一码率因子预测模型的训练之后执行对第二码率因子预测模型的训练。

[0091] 在确定多个目标码率因子之后，根据该多个码率因子分别使用转码系统中的对应编码器（基于对应的预设视频编码标准）对待编码视频片段进行编码。经过多次试验数据显示，使用第二次预测获得的目标码率因子对视频片段进行编码之后，经编码视频片段的视频编码分值有99%的概率在预设视频编码分值区间之内，因此，可以直接信任基于第二次预测获得的目标码率因子的视频片段编码结果。

[0092] 在利用多个目标码率因子对每个待编码视频片段进行编码之后，方法100还可以包括：针对多个预设视频编码标准中的每一个预设视频编码标准，获取该预设视频编码标准下的所有目标视频片段；以及基于一个或多个子视频片段在待编码视频中的顺序，对所有目标视频片段进行组合，以获得该预设视频编码标准下对应于待编码视频的目标视频。

[0093] 对一个待编码视频中的每个待编码视频片段进行上述的编码操作之后，可以获得针对每个待编码视频片段的多个目标视频片段。然后，针对每一个预设视频编码标准，可以将针对每个待编码视频片段的多个目标视频片段中的对应目标视频片段写入到最终的视频码流文件中，并且根据每个待编码视频片段在该待编码视频中的顺序对写入的目标视频片段进行拼接，获得该预设视频编码标准下对应于该待编码视频的目标视频。

[0094] 如上所述，每个待编码视频片段均为展示内容相似且连贯的一个单镜头视频片段，因此，针对每个待编码视频片段进行编码并获得对应的目标视频片段可以确保每个经编码的视频片段的质量，进而确保了各个场景下稳定的画质体验。此外，基于这些待编码视频片段在待编码视频中的顺序对获得的目标视频片段进行拼接组合，可以还原完整的视频文件，避免目标视频片段拼接错误导致的视频文件不连续。

[0095] 图5示出了根据本公开的实施例的视频编码装置500的结构框图。如图5所示，装置500可以包括：获取模块510，被配置为获取待编码视频片段，该待编码视频片段包括一个或多个视频帧；计算模块520，被配置为基于该一个或多个视频帧计算待编码视频片段的时空域特征信息；确定预测码率因子模块530，被配置为基于时空域特征信息确定多个预设视频

编码标准下待编码视频片段的多个预测码率因子;确定目标码率因子模块540,被配置为基于多个预测码率因子中对应于多个预设视频编码标准中的第一预设视频编码标准的第一预测码率因子和预设视频编码分值区间,确定多个预设视频编码标准下待编码视频片段的多个目标码率因子,其中,在多个预设视频编码标准中,利用第一预设视频编码标准进行视频编码的时间最短;以及编码模块550,被配置为根据多个目标码率因子分别对待编码视频片段进行编码,以获得多个目标视频片段。

[0096] 根据本公开的一些实施例,时空域特征信息可以包括时空复杂度信息,并且其中,计算模块530可以包括:被配置为基于一个或多个视频帧计算待编码视频片段的灰度共生信息和归一化信息作为时空复杂度信息的模块,其中,灰度共生信息用于描述待编码视频片段的空间复杂度,并且归一化信息用于描述待编码视频片段的时域复杂度;以及被配置为对灰度共生信息和归一化信息进行组合,以生成待编码视频片段的时空复杂度信息的模块。

[0097] 根据本公开的一些实施例,时空域特征信息可以包括编码特征信息,并且其中,计算模块530可以包括:预处理模块,被配置为对一个或多个视频帧进行预处理,以生成新的视频帧序列,新的视频帧序列包括视频帧对集合;以及帧内帧间编码模块,被配置为针对视频帧对集合中的每一个视频帧对,分别进行帧内编码和帧间编码,以获得待编码视频片段的编码特征信息。

[0098] 根据本公开的一些实施例,预处理模块可以包括:被配置为对一个或多个视频帧中除首帧和尾帧之外的视频帧进行复制,以生成一个或多个复制视频帧的模块;以及被配置为对一个或多个视频帧和一个或多个复制视频帧进行排序,使得一个或多个复制视频帧中的每一个复制视频帧位于对应的视频帧之后的模块。

[0099] 根据本公开的一些实施例,帧内帧间编码模块可以包括:被配置为对视频帧对集合中的每一个视频帧对中的第一视频帧进行帧内编码,以获得帧内编码信息的模块;被配置为对视频帧对集合中的每一个视频帧对进行帧间编码,以获得帧间编码信息的模块;以及被配置为基于帧内编码信息和帧间编码信息获得编码特征信息的模块。

[0100] 根据本公开的一些实施例,编码特征信息可以基于以下项中的一个或多个:编码比特数、帧内预测模式占比值、帧间运动矢量幅度分布信息。

[0101] 根据本公开的一些实施例,确定预测码率因子模块530可以包括:被配置为将待编码视频片段的时空域特征信息输入到第一码率因子预测模型,以经由第一码率因子预测模型确定多个预设视频编码标准下待编码视频片段的多个预测码率因子的模块。

[0102] 根据本公开的一些实施例,确定目标码率因子模块540可以包括:被配置为基于第一预测码率因子对待编码视频片段进行第一次编码,以获得第一预设视频编码标准下的第一预编码视频片段的模块;被配置为确定第一预编码视频片段的视频编码分值是否在预设视频编码分值区间内的模块;以及被配置为响应于确定第一预编码视频片段的视频编码分值在预设视频编码分值区间内而确定多个预测码率因子作为多个目标码率因子的模块。

[0103] 根据本公开的一些实施例,确定目标码率因子模块540还可以包括:更新模块,被配置为响应于确定第一预编码视频片段的视频编码分值不在预设视频编码分值区间内,基于待编码视频片段的时空域特征信息、多个预测码率因子和第一预编码视频片段的视频编码分值,更新多个预测码率因子以获得多个目标码率因子。

[0104] 根据本公开的一些实施例,更新模块可以包括:被配置为将待编码视频片段的时空域特征信息、多个预测码率因子和第一预编码视频片段的视频编码分值输入到第二码率因子预测模型,以经由第二码率因子预测模型确定多个预设视频编码标准下待编码视频片段的多个目标码率因子的模块。

[0105] 根据本公开的一些实施例,其中,第一码率因子预测模型可以通过以下操作被训练:针对多个预设视频编码标准中的每一个预设视频编码标准:获取样本视频片段的样本时空域特征信息以及该预设视频编码标准下样本视频片段的样本码率因子;将样本时空域特征信息输入到第一码率因子预测模型,以经由第一码率因子预测模型确定该预设视频编码标准下样本视频片段的样本预测码率因子;以及基于样本码率因子和样本预测码率因子来计算第一模型损失值;以及基于每一个预设视频编码标准下的第一模型损失值来调整第一码率因子预测模型的参数,直到达到第一模型训练停止条件。

[0106] 根据本公开的一些实施例,第二码率因子预测模型通过以下操作被训练:将样本时空域特征信息、每一个预设视频编码标准下的样本预测码率因子、和与第一预设视频编码标准下的样本预测码率因子相对应的第一样本视频编码分值输入到第二码率因子预测模型,以经由第二码率因子预测模型确定多个预设视频编码标准下样本视频片段的多个样本目标码率因子;针对多个预设视频编码标准中的每一个预设视频编码标准,基于样本码率因子和对应于该预设视频编码标准的样本目标码率因子来计算第二模型损失值;以及基于每一个预设视频编码标准下的第二模型损失值来调整第二码率因子预测模型的参数,直到达到第二模型训练停止条件。

[0107] 根据本公开的一些实施例,获取模块510可以包括:被配置为获取待编码视频的模块;被配置为对待编码视频进行场景检测的模块;被配置为基于场景检测结果将待编码视频分成一个或多个子视频片段的模块;以及被配置为识别一个或多个视频片段中的每一个视频子片段作为待编码视频片段的模块。

[0108] 根据本公开的一些实施例,装置500还可以包括:被配置为针对多个预设视频编码标准中的每一个预设视频编码标准,获取该预设视频编码标准下的所有目标视频片段的模块;以及被配置为基于一个或多个子视频片段在待编码视频中的顺序,对所有目标视频片段进行组合,以获得该预设视频编码标准下对应于待编码视频的目标视频的模块。

[0109] 应当理解,图5中所示装置500的各个模块510-550可以与参考图1描述的方法100中的各个步骤S110-S150相对应。由此,上面针对方法100所描述的操作、特征和优点同样适用于装置500及其所包括的模块。为了简洁起见,某些操作、特征和优点在此不再赘述。

[0110] 还应当理解,本文可以在软件硬件元件或程序模块的一般上下文中描述各种技术。上面关于图5描述的各个模块可以在硬件中或在结合软件和/或固件的硬件中实现。例如,这些模块可以被实现为计算机程序代码/指令,该计算机程序代码/指令被配置为在一个或多个处理器中执行并存储在计算机可读存储介质中。可替换地,这些模块可以被实现为硬件逻辑/电路。例如,在一些实施例中,获取模块510、计算模块520、确定预测码率因子模块530、确定目标码率因子模块540和编码模块550中的一个或多个可以一起被实现在片上系统(System on Chip, SoC)中。SoC可以包括集成电路芯片(其包括处理器(例如,中央处理单元(Central Processing Unit, CPU)、微控制器、微处理器、数字信号处理器(Digital Signal Processor, DSP)等)、存储器、一个或多个通信接口、和/或其他电路中的一个或多

个部件),并且可以可选地执行所接收的程序代码和/或包括嵌入式固件以执行功能。

[0111] 根据本公开的另一方面,还提供一种电子设备,包括:至少一个处理器;以及与所述至少一个处理器通信连接的至少一个存储器;其中所述至少一个存储器存储有计算机程序,所述计算机程序在被所述至少一个处理器执行时实现上述视频编码方法。

[0112] 根据本公开的另一方面,还提供一种存储有计算机程序的非瞬时计算机可读存储介质,其中,所述计算机程序在被处理器执行时实现上述视频编码方法。

[0113] 根据本公开的另一方面,还提供一种计算机程序产品,包括计算机程序,其中,所述计算机程序在被处理器执行时实现上述视频编码方法。

[0114] 参见图6,现将描述可以作为本公开的服务器的电子设备600的结构框图,其是可以应用于本公开的各方面的硬件设备的示例。电子设备可以是不同类型的计算机设备,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0115] 如图6所示,电子设备600可以包括能够通过系统总线630彼此通信的至少一个处理器610、工作存储器620、输入单元640、显示单元650、扬声器660、存储单元670、通信单元680以及其它输出单元690。

[0116] 处理器610可以是单个处理单元或多个处理单元,所有处理单元可以包括单个或多个计算单元或者多个核心。处理器610可以被实施成一个或更多微处理器、微型计算机、微控制器、数字信号处理器、中央处理单元、状态机、逻辑电路和/或基于操作指令来操纵信号的任何设备。处理器610可以被配置成获取并且执行存储在工作存储器620、存储单元670或者其他计算机可读介质中的计算机可读指令,诸如操作系统620a的程序代码、应用程序620b的程序代码等。

[0117] 工作存储器620和存储单元670是用于存储指令的计算机可读存储介质的示例,指令由处理器610执行来实施前面所描述的各种功能。工作存储器620可以包括易失性存储器和非易失性存储器二者(例如RAM、ROM等等)。此外,存储单元670可以包括硬盘驱动器、固态驱动器、可移除介质、包括外部和可移除驱动器、存储器卡、闪存、软盘、光盘(例如CD、DVD)、存储阵列、网络附属存储、存储区域网等等。工作存储器620和存储单元670在本文中都可以被统称为存储器或计算机可读存储介质,并且可以是能够把计算机可读、处理器可执行程序指令存储为计算机程序代码的非暂态介质,计算机程序代码可以由处理器610作为被配置成实施在本文的示例中所描述的操作和功能的特定机器来执行。

[0118] 输入单元660可以是能向电子设备600输入信息的任何类型的设备,输入单元660可以接收输入的数字或字符信息,以及产生与电子设备的用户设置和/或功能控制有关的键信号输入,并且可以包括但不限于鼠标、键盘、触摸屏、轨迹板、轨迹球、操作杆、麦克风和/或遥控器。输出单元可以是能呈现信息的任何类型的设备,并且可以包括但不限于显示单元650、扬声器660以及其它输出单元690,其它输出单元690可以但不限于包括视频/音频输出终端、振动器和/或打印机。通信单元680允许电子设备600通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据,并且可以包括但不限于调制解调器、网

卡、红外通信设备、无线通信收发机和/或芯片组,例如蓝牙™设备、802.11设备、Wi-Fi设备、WiMAX设备、蜂窝通信设备和/或类似物。

[0119] 工作寄存器620中的应用程序620b可以被加载执行上文所描述的各个方法及其步骤,例如图1中的步骤S110-S150、图2中的步骤S210-S240、图4中的步骤S442-S448。例如,在一些实施例中,上文描述的各种方法可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元670。在一些实施例中,计算机程序的部分或者全部可以经由存储单元670和/或通信单元680而被载入和/或安装到电子设备600上。当计算机程序被加载并由处理器610执行时,可以执行上文描述的方法100、200的一个或多个步骤。备选地,在其他实施例中,处理器610可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行方法100、200。

[0120] 本文中以上描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、芯片上系统的系统(SOC)、负载可编程逻辑设备(CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0121] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器,使得程序代码当由处理器或控制器执行时使流程图和/或框图所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

[0122] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0123] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还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传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0124] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算

系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术的实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0125] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务端关系的计算机程序来产生客户端和服务端的关系。

[0126] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本公开中记载的各步骤可以并行地执行、也可以顺序地或以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0127] 虽然已经参照附图描述了本公开的实施例或示例,但应理解,上述的方法、系统和设备仅仅是示例性的实施例或示例,本发明的范围并不由这些实施例或示例限制,而是仅由授权后的权利要求书及其等同范围来限定。实施例或示例中的各种要素可以被省略或者可由其等同要素替代。此外,可以通过不同于本公开中描述的次序来执行各步骤。进一步地,可以以各种方式组合实施例或示例中的各种要素。重要的是随着技术的演进,在此描述的很多要素可以由本公开之后出现的等同要素进行替换。

100

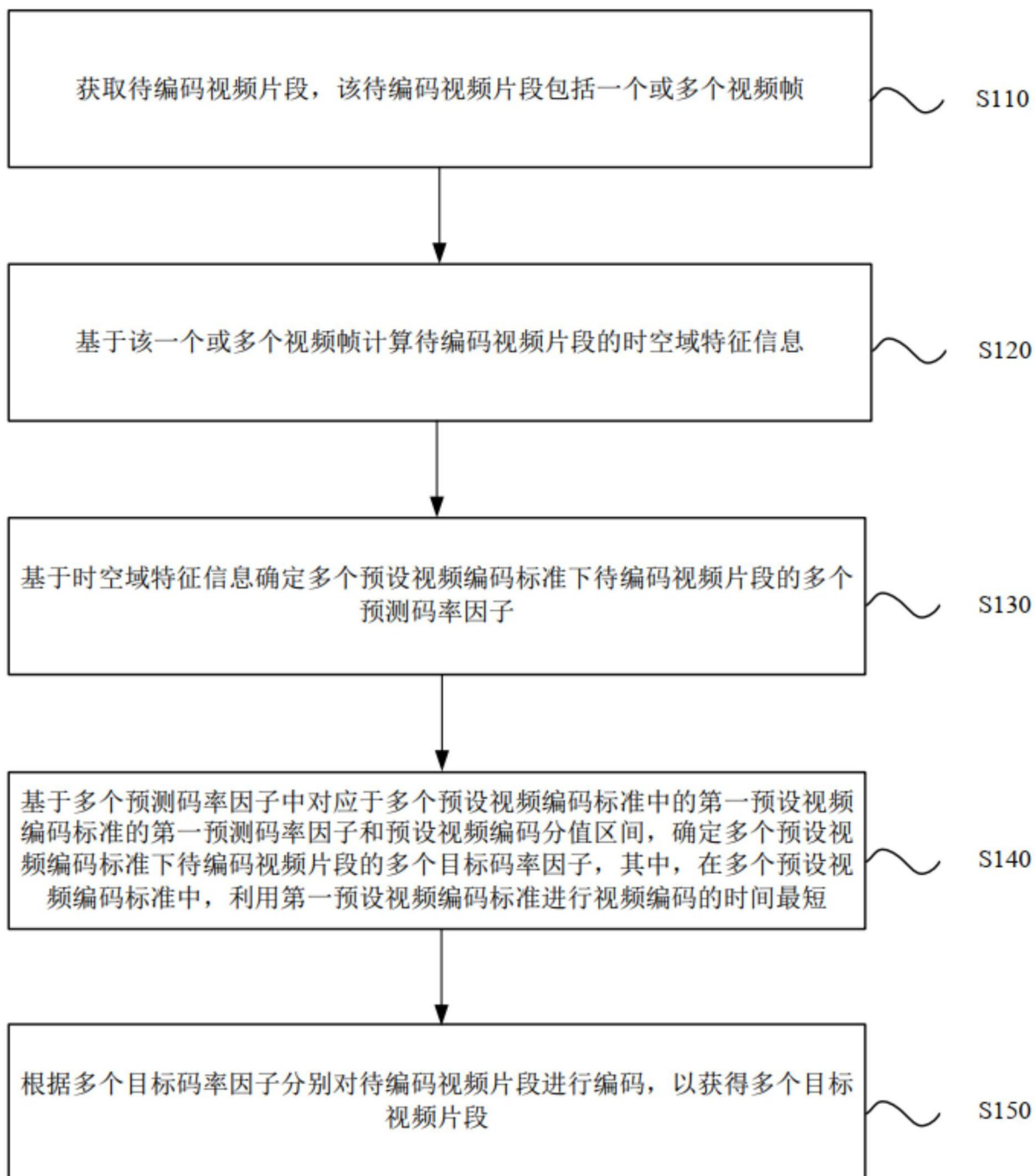


图1

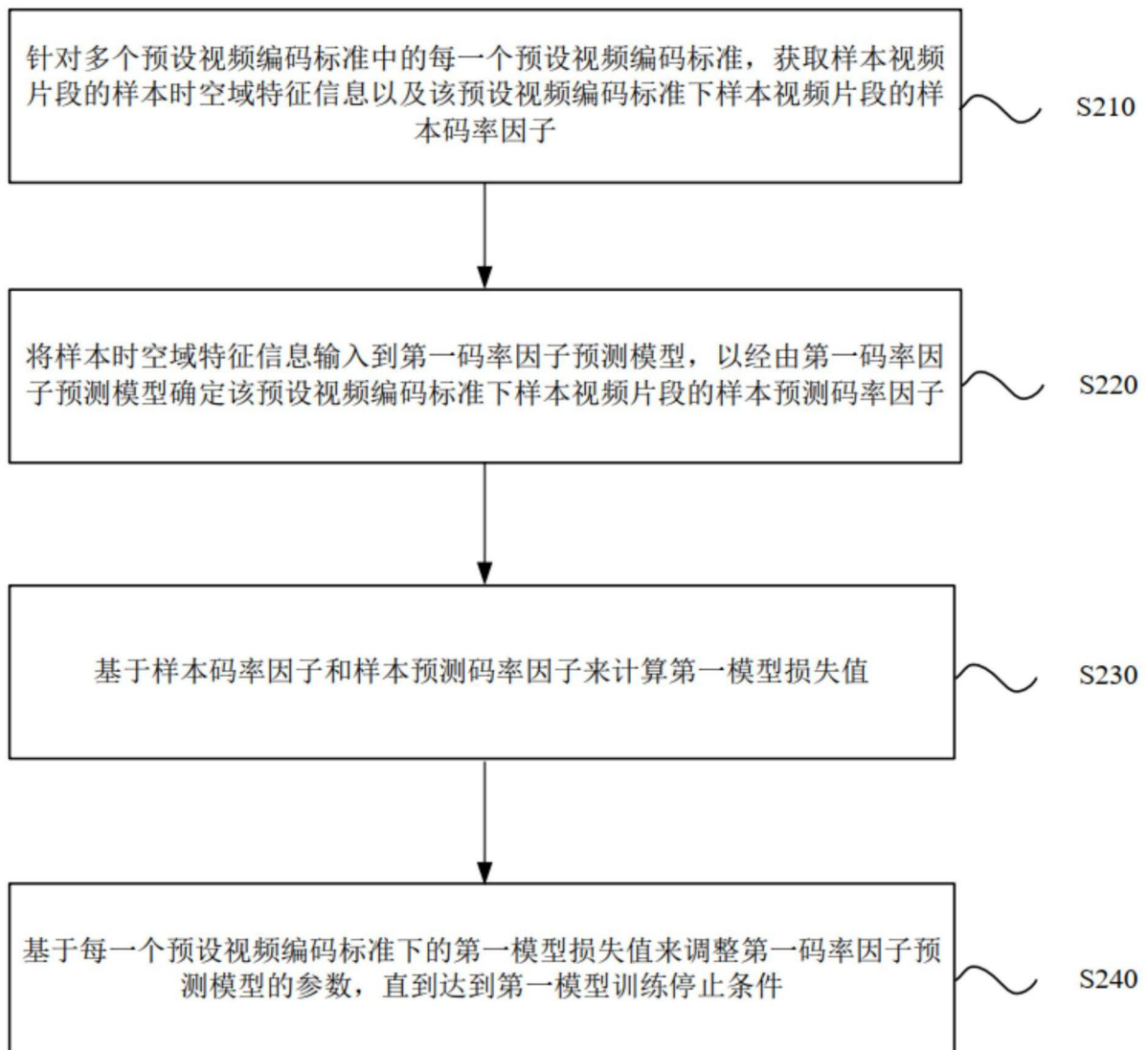
200

图2

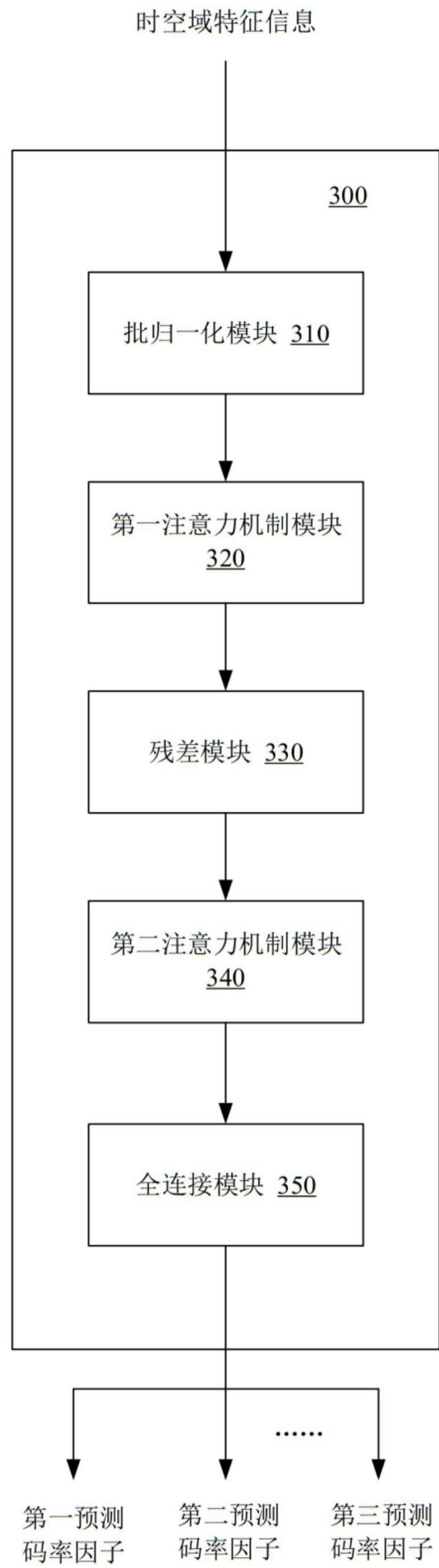


图3

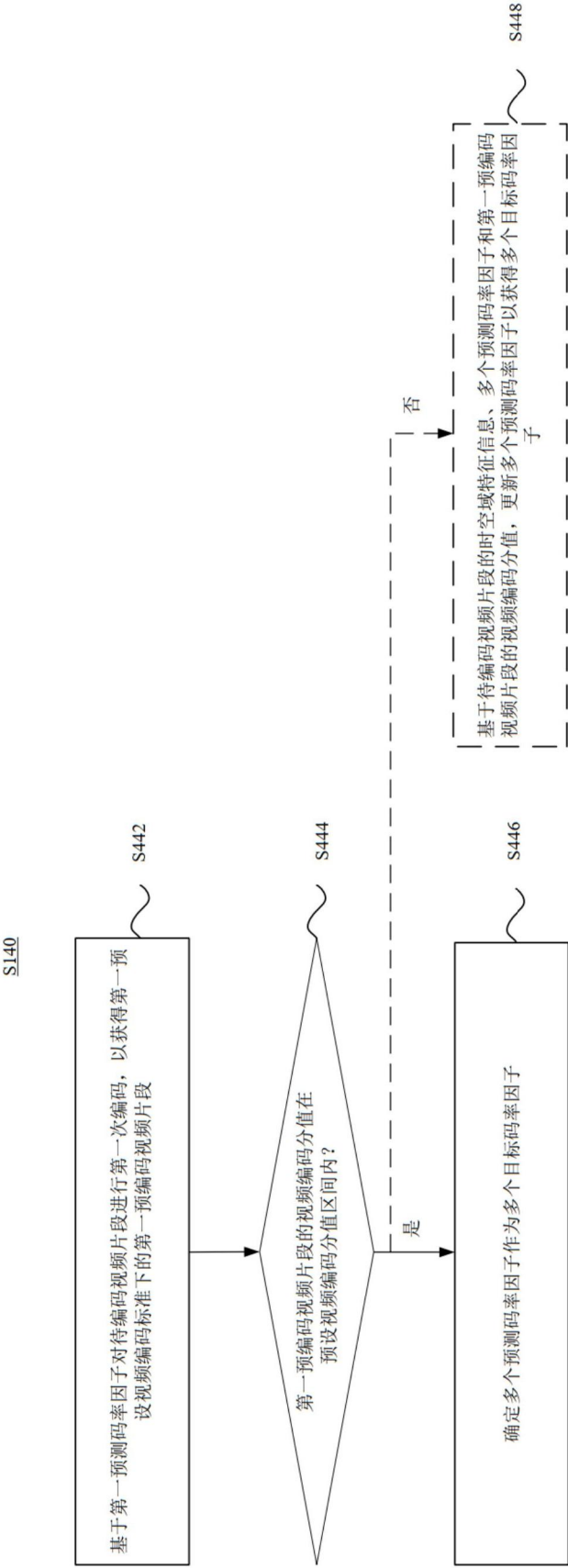


图4



图5

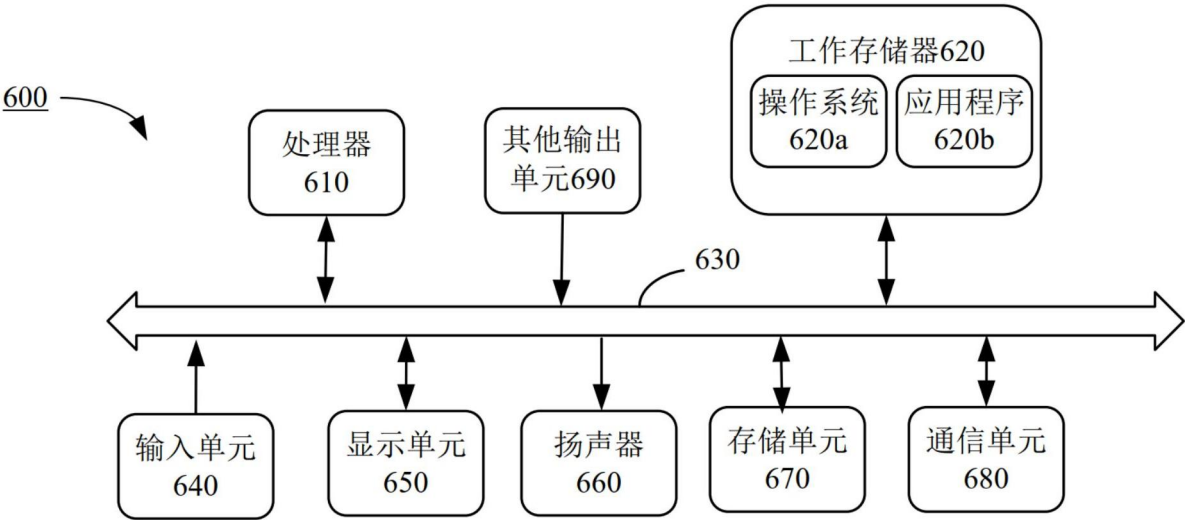


图6