

一种视频处理方法、视频处理装置、智能设备及存储介质

本申请要求于2020年11月9日提交中国专利局、申请号为202011239333.1、名称为“一种视频处理方法、视频处理装置、智能设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本申请涉及计算机技术领域，尤其涉及一种视频处理方法、视频处理装置、智能设备以及计算机可读存储介质。

背景技术

10 目前，视频编码技术广泛应用于视频会话、视频点播等场景中。例如，采用视频编码技术对视频会话场景中涉及的会话视频进行处理，采用视频编码技术对视频点播场景中涉及的点播视频进行处理等等。视频编码技术是指按照编码模式对视频进行压缩编码的技术。对视频进行压缩编码能够有效节省视频的存储空间，提升视频的传输效率。

在视频编码过程中准确地确定较为合适的编码模式，并按照该编码模式对视频进行编码能够加速整个视频编码过程，提升视频编码效率，因此，在视频编码过程中如何较为准确地确定编码模式进行视频帧编码成为当前研究的热点问题。

发明内容

本申请实施例提供了一种视频处理方法、装置、设备及存储介质，可以较为准确地选择合适的编码模式进行视频帧编码。

20 一方面，本申请实施例提供一种视频处理方法，由智能设备执行，该视频处理方法包括：

从待编码视频中获取目标视频帧，并从目标视频帧中确定待编码的目标数据块；
对目标数据块进行场景复杂度分析，得到数据块指标信息；
将目标数据块划分成N个子数据块，并分别对每一个子数据块进行场景复杂度分析，得到子块指标信息，N为大于或等于2的整数；
25 根据数据块指标信息和子块指标信息，确定对目标数据块的编码模式；
按照确定的编码模式对目标数据块进行编码。

另一方面，本申请实施例提供一种视频处理装置，该视频处理装置包括：
获取单元，用于从待编码视频中获取目标视频帧，并从目标视频帧中确定待编码的目标数据块；

30 处理单元，用于对目标数据块进行场景复杂度分析，得到数据块指标信息；将目标数据块划分成N个子数据块，并分别对每一个子数据块进行场景复杂度分析，得到子块指标信息，N为大于或等于2的整数；根据数据块指标信息和子块指标信息，确定对目标数据块的编码模式；以及，按照确定的编码模式对目标数据块进行编码。

另一方面，本申请实施例提供一种智能设备，该智能设备包括：

处理器，适于实现计算机程序；以及，
存储器，所述存储器存储有计算机程序，所述计算机程序由处理器加载并运行，实现上述的视频处理方法。

另一方面，本申请实施例提供一种计算机可读存储介质，该计算机可读存储介质存储有计算机程序，该计算机程序被计算机设备的处理器读取并执行时，实现上述的视频处理方法。

另一方面，本申请实施例提供了一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器执行该计算机指令，实现上述的视频处理方法。

附图简要说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例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申请的一些实施例，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。

- 图 1a 是本申请实施例提供的一种视频帧的递归划分过程的示意图；
- 图 1b 是本申请实施例提供的一种视频帧的划分结果的示意图；
- 图 2 是本申请实施例提供的一种编码器的结构示意图；
- 图 3 是本申请实施例提供的一种关联数据块的示意图；
- 图 4 是本申请实施例提供的一种视频处理系统的架构示意图；
- 图 5 是本申请实施例提供的一种视频处理方法的流程示意图；
- 图 6 是本申请实施例提供的另一种视频处理方法的流程示意图；
- 图 7 是本申请实施例提供的另一种视频处理方法的流程示意图；
- 图 8 是本申请实施例提供的一种视频处理装置的结构示意图；
- 图 9 是本申请实施例提供的一种智能设备的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本申请实施例涉及云技术（Cloud Technology），本申请实施例可以利用云技术实现视频处理过程。云技术是指在广域网或局域网内将硬件、软件、网络等系列资源统一起来，实现数据的计算、储存、处理和共享的一种托管技术。云技术是基于云计算商业模式应用的网络技术、信息技术、整合技术、管理平台技术、应用技术等的总称，可以组成资源池，按需所用，灵活便利。云计算技术将变成重要支撑。技术网络系统的后台服务需要大量的计算、存储资源，如视频网站、图片类网站和更多的门户网站。伴随着互联网行业的高度发展和应用，将来每个物品都有可能存在自己的识别标志，都需要传输到后台系统进行逻

辑处理，不同程度级别的数据将会分开处理，各类行业数据皆需要强大的系统后盾支撑，这只能通过云计算来实现。

云计算（Cloud Computing）是一种计算模式，它将计算任务分布在大量计算机构成的资源池上，使各种应用系统能够根据需要获取计算力、存储空间和信息服务。提供资源的网络被称为“云”。“云”中的资源在使用者看来是可以无限扩展的，并且可以随时获取、按需使用、随时扩展、按使用付费。作为云计算的基础能力提供商，会建立云计算资源池（简称云平台），一般称为 IaaS（Infrastructure as a Service，基础设施即服务）平台，并在云计算资源池中部署多种类型的虚拟资源，供外部客户选择使用。云计算资源池中主要包括：计算设备（为虚拟化机器，包含操作系统）、存储设备、网络设备。按照逻辑功能划分，在 IaaS 层上可以部署 PaaS（Platform as a Service，平台即服务）层，PaaS 层之上再部署 SaaS（Software as a Service，软件即服务）层，也可以直接将 SaaS 部署在 IaaS 上。PaaS 为软件（例如数据库、web 容器等）运行的平台。SaaS 为各式各样的业务软件（例如 web 门户网站、短信群发器等）。一般来说，SaaS 和 PaaS 相对于 IaaS 是上层。云计算还指 IT（Internet Technology，互联网技术）基础设施的交付和使用模式，指通过网络以按需、易扩展的方式获得所需资源。广义云计算指服务的交付和使用模式，指通过网络以按需、易扩展的方式获得所需服务。这种服务可以是 IT 和软件、互联网相关，也可能是其他服务。云计算是网格计算（Grid Computing）、分布式计算（Distributed Computing）、并行计算（Parallel Computing）、效用计算（Utility Computing）、网络存储（Network Storage Technologies）、虚拟化（Virtualization）、负载均衡（Load Balance）等传统计算机和网络技术发展融合的产物。

云计算可以用于云会议领域，云会议是基于云计算技术的一种高效、便捷、低成本的会议形式。使用者只需要通过互联网界面，进行简单易用的操作，便可快速高效地与全球各地团队及客户同步分享语音、数据文件及视频，而会议中数据的传输、处理等复杂技术由云会议服务商帮助使用者进行操作。目前国内云会议主要集中在以 SaaS 模式为主体的服务内容，包括电话、网络、视频等服务形式，基于云计算的视频会议称为云会议。在云会议时代，数据的传输、处理、存储全部由视频会议厂家的计算机资源处理，使用者无需再购置昂贵的硬件和安装繁琐的软件，只需打开浏览器，登录相应界面，就能进行高效的远程会议。云会议系统支持多服务器动态集群部署，并提供多台高性能服务器，大大提升了会议稳定性、安全性、可用性。近年来，视频会议因能大幅提高沟通效率，持续降低沟通成本，带来内部管理水平升级，而获得众多用户欢迎，已广泛应用在政府、军队、交通、运输、金融、运营商、教育、企业等各个领域。毫无疑问，视频会议运用云计算以后，在方便性、快捷性、易用性上具有更强的吸引力，必将激发视频会议应用新高潮的到来。

本申请实施例涉及视频（Video），视频是由连续的视频帧构成的视频帧序列。由于人眼的视觉暂留效应，当视频帧序列以一定的速率播放时，我们看到的动作就是连续的视频。由于连续的视频帧之间相似性极高，每一个视频帧内部以及连续的视频帧之间均存在大量的冗余信息，因此在对视频进行存储或传输之前，往往需要采用视频编码（Video Encoding）技术对视频进行编码，去除视频在空间、时间等维度上的冗余信息，以节省存储空间、提升视频传输效率。

视频编码技术也可以称为视频压缩技术，是指按照编码模式对视频中的每一个视频帧进行压缩编码的技术。具体地，在对视频进行编码的过程中，需要将视频中的每一个视频

帧递归地划分成多种尺寸的数据块, 然后分别将每一个划分得到的数据块输入编码器中进行编码。下面结合图 1a 对本申请实施例涉及的递归划分过程进行介绍。图 1a 是本申请实施例提供的一种视频帧的递归划分过程的示意图。如图 1a 所示, 视频帧 10 由若干第一尺寸的数据块 101 组成; 视频帧 10 的递归划分方式可以包括: ①直接将第一尺寸的数据块 101 输入编码器中进行编码, 也就是说, 不对第一尺寸的数据块 101 进行划分, 直接将第一尺寸的数据块输入编码器中进行编码; ②将第一尺寸的数据块 101 划分为 4 个尺寸相同的第二尺寸的数据块 102, 不对第二尺寸的数据块 102 进行进一步划分, 将第二尺寸的数据块 102 输入编码器中进行编码; ③对第二尺寸的数据块 102 进行进一步划分, 划分得到 4 个尺寸相同的第三尺寸的数据块 103, 不对第三尺寸的数据块 103 进行进一步划分, 将第三尺寸的数据块 103 输入编码器中进行编码; 以此类推, 还可以对第三尺寸的数据块 103 进行进一步划分, 从而将视频帧 10 递归地划分成多种尺寸的数据块, 视频帧 10 的划分结果可参见图 1b。图 1b 是本申请实施例提供的一种视频帧的划分结果的示意图。如图 1b 所示, 划分后的视频帧 10 由三种尺寸的数据块组成, 分别是第一尺寸的数据块 101、第二尺寸的数据块 102 和第三尺寸的数据块 103。

在实际视频编码的过程中, 视频帧中包含的多个数据块的划分策略不同, 对视频帧进行编码的编码模式也不相同, 对视频帧的编码速度也不相同。本申请实施例提供一种视频处理方案, 该视频处理方案提供两种不同的数据块划分策略, 分别是自顶向下的数据块划分策略和自底向上的数据块划分策略。这两种数据块划分策略均能确定出视频帧中每一个数据块的最优划分方式。最优划分方式可以理解为按照该数据块划分方式对数据块进行划分后, 对划分得到的数据块进行编码时的编码质量较好、编码码率较低。编码码率可以是指对编码得到的视频数据流进行传输时单位时间 (例如 1 秒) 内传输的数据量, 单位时间内传输的数据量越多, 编码码率越高。下面分别对以上两种数据块划分策略进行详细介绍。

(1) 自顶向下的数据块划分策略

在对视频进行编码的过程中, 对视频帧中的每一个数据块进行自顶向下的、递归的划分。具体地, 自顶向下的、递归的划分可以是指从视频帧中数据块的最大尺寸向视频帧中数据块的最小尺寸进行划分, 直至找到数据块的最优划分方式的过程。举例来说, 视频帧中数据块的最大尺寸为 $64\text{px} \times 64\text{px}$ (Pixel, 像素), 视频帧中数据块的最小尺寸为 $4\text{px} \times 4\text{px}$ 。数据块的最大尺寸为 $64\text{px} \times 64\text{px}$, 数据块的最小尺寸为 $4\text{px} \times 4\text{px}$ 仅用于举例, 并不构成对本申请实施例的限定, 例如数据块的最大尺寸还可以为 $128\text{px} \times 128\text{px}$, 数据块的最小尺寸还可以为 $8\text{px} \times 8\text{px}$ 等等。可以将 $64\text{px} \times 64\text{px}$ 的数据块划分为 4 个 $32\text{px} \times 32\text{px}$ 的数据块。分别计算对 4 个 $32\text{px} \times 32\text{px}$ 的数据块中的每一个数据块进行编码的率失真参数 (Rate-Distortion Cost, RDCost), 并计算 4 个 $32\text{px} \times 32\text{px}$ 的数据块的 RDCost 之和 (即计算 4 个 $32\text{px} \times 32\text{px}$ 的数据块的第一总和 RDCost), 以及计算直接对 $64\text{px} \times 64\text{px}$ 的数据块进行编码的 RDCost。如果第一总和 RDCost 大于 $64\text{px} \times 64\text{px}$ 的数据块的 RDCost, 则确定数据块的划分尺寸为 $64\text{px} \times 64\text{px}$, 即无需对 $64\text{px} \times 64\text{px}$ 的数据块进行划分; 如果第一总和 RDCost 小于或等于 $64\text{px} \times 64\text{px}$ 的数据块的 RDCost, 则确定数据块的划分尺寸为 $32\text{px} \times 32\text{px}$ 。随后, 还可以继续将 $32\text{px} \times 32\text{px}$ 的数据块划分为 4 个 $16\text{px} \times 16\text{px}$ 的数据块进行编码, 以此类推, 直至找到数据块的最优划分方式。自顶向下的数据块划分策略通常情况下都能够较为准确的确定出最优划分方式。对于场景复杂程度较低的数据块, 数据块的最优划分尺寸比较大, 因此对于场景复杂程度较低的数据块, 通过自顶向下的数据块划分策

略能够快速确定该数据块的最优划分方式。对于场景复杂程度较高的数据块，数据块的最优划分尺寸比较小，因此对于场景复杂程度较高的数据块，通过自顶向下的数据块划分策略确定该数据块的最优划分方式的过程需要耗费大量的时间成本，进而影响了该数据块的编码速度。

- 5 数据块的最优划分尺寸可以理解为当划分得到的数据块的尺寸为该尺寸时，对划分得到的数据块进行编码时的编码质量较好、编码码率较低。一般而言，场景复杂程度较低的数据块对应的最优划分尺寸比较大，场景复杂程度较高的数据块对应的最优划分尺寸比较小。数据块的场景复杂程度可以采用空间信息（Spatial Information, SI）或时间信息（Temporal Information）等方法进行衡量。空间信息可以用于表征数据块的空间细节量，
- 10 数据块中包含的元素越多，数据块的空间信息的数值越高，数据块的场景复杂程度越高。例如，数据块 A 中包含 5 个元素（一只猫、一条狗、一棵树、一朵花和一个太阳），数据块 B 中包含 2 个元素（一只猫和一条狗），数据块 A 中包含的元素多于数据块 B 中包含的元素，那么数据块 A 的空间信息的数值高于数据块 B 的空间信息的数值，数据块 A 的场景复杂程度高于数据块 B 的场景复杂程度。时间信息可以用于表征数据块的时间变化量，
- 15 当前正在处理的目标视频帧中的目标数据块相对于目标视频帧的参考视频帧中的参考数据块的运动程度越高，目标数据块的时间信息的数值越大，目标数据块的场景复杂程度越高。参考视频帧是视频帧序列中编码顺序位于目标视频帧之前的视频帧，目标数据块在目标视频帧中的位置与参考数据块在参考视频帧中的位置相同。例如，目标数据块中包含一个元素（例如汽车元素），参考数据块中同样包含该元素，目标数据块中汽车元素相对于
- 20 参考数据块中汽车元素的位移越大，目标视频帧中的目标数据块相对于参考视频帧中的参考数据块的运动程度越高，目标数据块的时间信息的数值越大，目标数据块的场景复杂程度越高。

（2）自底向上的视频帧划分策略

- 在对视频进行编码的过程中，对视频帧中的每一个数据块进行自底向上的、递归的划分。具体地，自底向上的、递归的划分可以是指从视频帧中数据块的最小尺寸向视频帧中数据块的最大尺寸进行划分，直至找到数据块的最优划分方式的过程。举例来说，视频帧中数据块的最小尺寸为 $4\text{px} \times 4\text{px}$ ，视频帧中数据块的最大尺寸为 $64\text{px} \times 64\text{px}$ 。分别对 4 个 $4\text{px} \times 4\text{px}$ 的数据块中的每一个数据块进行编码，再对由 4 个 $4\text{px} \times 4\text{px}$ 的数据块组成的 $8\text{px} \times 8\text{px}$ 的数据块进行编码，分别计算对 4 个 $4\text{px} \times 4\text{px}$ 的数据块中的每一个数据块进行
- 30 编码的 RDCost，并计算 4 个 $4\text{px} \times 4\text{px}$ 的数据块的 RDCost 之和（即计算 4 个 $4\text{px} \times 4\text{px}$ 的数据块的第二总和 RDCost）；以及计算对 $8\text{px} \times 8\text{px}$ 的数据块进行编码的 RDCost。如果第二总和 RDCost 小于或等于 $8\text{px} \times 8\text{px}$ 的数据块的 RDCost，则确定数据块的划分尺寸为 $4\text{px} \times 4\text{px}$ ；如果第二总和 RDCost 大于 $8\text{px} \times 8\text{px}$ 的数据块的 RDCost，则确定数据块的划分尺寸为 $8\text{px} \times 8\text{px}$ 。随后，还可以继续对由 4 个 $8\text{px} \times 8\text{px}$ 的数据块组成的 $16\text{px} \times 16\text{px}$ 的数据块进行编码，以此类推，直至找到数据块的最优划分方式。自底向上的视频帧划分策略通常情况下都能够较为准确的确定出最优划分方式。对于场景复杂程度较高的数据块，数据块的最优划分尺寸比较小，因此对于场景复杂程度较高的数据块，通过自底向上的视频帧划分策略能够快速确定该数据块的最优划分方式。对于场景复杂程度较低的数据块，数据块的最优划分尺寸比较大，因此对于场景复杂程度较低的数据块，通过自底向上的视频帧
- 40 划分策略确定该数据块的最优划分方式的过程需要耗费大量的时间成本，进而影响了该数

据块的编码速度。

由此可见，自顶向下的数据块划分策略和自底向上的视频帧划分策略通常情况下都能较为准确地确定出视频帧中每一个数据块的最优划分方式，自顶向下的数据块划分策略更适用于场景复杂程度较低的数据块，自底向上的视频帧划分策略更适用于场景复杂程度较高的数据块。

在此基础上，本申请实施例提供了进一步的视频处理方案，该视频处理方案对待编码视频帧中待编码的目标数据块进行场景复杂度分析；将目标数据块划分成多个子数据块，并分别对每一个子数据块进行场景复杂度分析；通过目标数据块的场景复杂度分析结果，以及划分得到的多个子数据块中每一个子数据块的场景复杂度分析结果，确定对目标数据块的较优划分方式，进而确定对目标数据块的编码模式，从而可以按照确定的编码模式对目标数据块进行编码。本申请实施例中的场景复杂度分析可以采用相关技术中的任意场景复杂度分析方法来执行，例如误差平方和（SSE，Sum of Squared Error）方法、绝对差和（SAD，Sum of Absolute Difference）方法等。本申请实施例提及的较优划分方式是指在编码质量与编码码率之间取得平衡的前提下，能够在一定程度上提升编码速度的数据块划分方式。该视频处理方案能够根据目标数据块的场景复杂度制定与目标数据块的场景复杂度相适应的编码模式，有效提升了对目标数据块的编码速度，从而提升了视频编码速度；并且该视频处理方案更具普适性，适用于任意视频编码场景，既能确定出适用于场景复杂程度较低的数据块的编码模式，又能确定出适用于场景复杂程度较高的数据块的编码模式。

基于上述描述，本申请实施例提供一种编码器，该编码器可以用于实现上述的视频处理方案，该视频编码器可以是 AV1（Alliance for Open Media Video 1，开放媒体视频联盟 1）标准编码器。图 2 是本申请实施例提供的一种编码器的结构示意图。如图 2 所示，该编码器 20 包括场景复杂度分析模块 201、自适应划分决策模块 202 和划分提前终止模块 203。

场景复杂度分析模块 201 可以用于对待编码视频帧中待编码的目标数据块进行场景复杂度分析，得到数据块指标信息。场景复杂度分析模块 201 还可以用于将目标数据块划分为 N 个子数据块，并对 N 个子数据块中的每一个子数据块进行场景复杂度分析，得到子块指标信息，N 为大于或等于 2 的整数。在实际编码场景中，N 的取值一般为 4，本申请实施例对此不作限定。将目标数据块划分为 N 个子数据块可以是指将目标数据块划分为 4 个尺寸相同的子数据块。

自适应划分决策模块 202 可以用于根据数据块指标信息和子块指标信息确定对目标数据块的编码模式。编码模式可以包括第一模式或第二模式中的任意一种，即自适应划分决策模块 202 可以根据数据块指标信息和子块指标信息确定出对目标数据块的编码模式为第一模式或第二模式。第一模式是指将数据块划分为 N 个子数据块，并分别对 N 个子数据块中的每一个子数据块进行编码的编码模式；第二模式是指不对目标数据块进行划分，直接对目标数据块进行编码的编码模式。若自适应划分决策模块 202 不能根据数据块指标信息和子块指标信息确定出对目标数据块的编码模式，那么自适应划分决策模块 202 还可以用于获取与目标数据块相关的 M 个关联数据块的关联块指标信息，并根据关联块指标信息确定对目标数据块的尝试编码顺序，M 为正整数。在实际编码场景中，M 的取值一般为 3，本申请实施例对此不作限定。图 3 是本申请实施例提供的一种关联数据块的示意图。如图 3 所示，与目标数据块相关的 M 个关联数据块可以是指位于目标数据块 301 左部位置的关联数据块 302、位于目标数据块 301 顶部位置的关联数据块 303、位于目标数据块 301 左

上部位置的关联数据块 304。尝试编码顺序可以包括第一尝试编码顺序或第二尝试编码顺序中的任意一种,即自适应划分决策模块 202 可以根据关联块指标信息确定对目标数据块的尝试编码顺序为第一尝试编码顺序或第二尝试编码顺序。第一尝试编码顺序是指优先尝试按照第一模式对目标数据块进行编码,再尝试按照第二模式对目标数据块进行编码的编码顺序;第二尝试编码顺序是指优先尝试按照第二模式对目标数据块进行编码,再尝试按照第一模式对目标数据块进行编码的编码顺序。

划分提前终止模块 203 用于在自适应划分决策模块 202 无法确定出对目标数据块的编码模式时设置划分终止条件,并根据设置的划分终止条件确定对目标数据块的编码模式。具体地,若自适应划分决策模块 202 确定对目标数据块的尝试编码顺序为第一尝试编码顺序,则划分提前终止模块 203 获取按照第一模式对目标数据块进行编码得到的 N 个子数据块的编码信息,若 N 个子数据块的编码信息满足第一划分终止条件(即下文中的第五条件),则划分提前终止模块 203 确定对目标数据块的编码模式为第一模式。若自适应划分决策模块 202 确定对目标数据块的尝试编码顺序为第二尝试编码顺序,则划分提前终止模块 203 获取按照第二模式对目标数据块进行编码得到的目标数据块的编码信息,若目标数据块的编码信息满足第二划分终止条件(即下文中的第六条件),则划分提前终止模块 203 确定对目标数据块的编码模式为第二模式。

图 2 所示的编码器 20 能够根据目标数据块的场景复杂度制定与目标数据块的场景复杂度相适应的编码模式,有效提升了对目标数据块的编码速度,从而提升了视频编码速度。此外,编码器 20 中的划分提前终止模块 203 能够在编码信息满足划分终止条件时确定出目标数据块的编码模式,并终止对目标数据块的进一步划分,从而在目标数据块的编码速度与编码码率取得平衡的前提下进一步提升目标数据块的编码速度,并进一步提升了视频编码效率。下面将结合图 4 所示的视频处理系统对本申请实施例提供的视频处理方案以及图 2 所示实施例提供的编码器 20 的具体应用场景进行介绍。

图 4 是本申请实施例提供的一种视频处理系统的架构示意图。如图 4 所示,该视频处理系统包括 P 个终端(例如第一终端 401、第二终端 402 等)和服务器 403, P 为大于 1 的整数。其中, P 个终端中的任一个终端可以是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表、智能可穿戴设备等带摄像功能的设备,但并不局限于此。 P 个终端中的任一个终端可以支持各种应用程序的安装及运行,此处的应用程序可以包括但不限于社交应用程序(例如即时通信应用程序、音频会话应用程序、视频会话应用程序等等)、音视频应用程序(例如音视频点播应用程序、音视频播放器等等)、游戏应用程序等等。服务器可以是独立的物理服务器,也可以是多个物理服务器构成的服务器集群或者分布式系统,还可以是提供云计算服务的云服务器,本申请实施例在此不作限制。 P 个终端与服务器 403 可以通过有线通信或者无线通信方式进行直接或间接地连接,本申请在此不做限制。下面以视频会话场景为例,对本申请实施例提供的视频处理方案进行介绍。

(1) 编码器 20 分别部署于 P 个终端中,由 P 个终端进行视频处理。

P 个用户分别使用 P 个终端中运行的视频会话应用程序参与视频会话,例如用户 1 使用第一终端 401 中运行的视频会话应用程序参与视频会话,用户 2 使用第二终端 402 中运行的视频会话应用程序参与视频会话等等。服务器 403 用于对 P 个终端在视频会话过程中产生的目标视频进行传输。其中,目标视频可以包括 P 个终端在视频会话过程中产生的会话视频。第一终端 401 是 P 个终端中的任一个终端,此处以第一终端 401 对目标视频进行

处理为例进行说明，P 个终端中除第一终端 401 之外的其他终端对目标视频的处理过程与第一终端 401 对目标视频的处理过程相同。

目标视频由多个连续的视频帧组成，目标视频包含的每一个视频帧中包括多个待编码的数据块。编码器 20 部署于第一终端 401 中，第一终端 401 调用编码器 20 对目标视频包含的每一个视频帧中的各个数据块的场景复杂度进行分析，确定出对每一个数据块的编码模式，并分别按照确定得到的编码模式对每一个数据块进行编码，最终得到编码后的目标视频。第一终端 401 可以将编码后的目标视频发送至服务器 403，由服务器 403 将编码后的目标视频传输至参与视频会话的其他终端中；或者，第一终端 401 还可以直接将编码后的目标视频传输至参与视频会话的其他终端中；其他终端接收到编码后的目标视频，解析并播放目标视频，以实现 P 个终端参与的视频会话。

(2) 编码器 20 部署于服务器 403 中，由服务器 403 进行视频处理。

P 个用户分别使用 P 个终端中运行的视频会话应用程序参与视频会话，例如用户 1 使用第一终端 401 中运行的视频会话应用程序参与视频会话，用户 2 使用第二终端 402 中运行的视频会话应用程序参与视频会话等等。服务器 403 用于对 P 个终端在视频会话过程中产生的目标视频进行处理和传输。其中，目标视频可以包括 P 个终端在视频会话过程中产生的会话视频。

目标视频可以是第一终端 401 在参与视频会话的过程中产生的会话视频，第一终端 401 将目标视频传输至服务器 403，目标视频由多个连续的视频帧组成，目标视频包含的每一个视频帧中包括多个待编码的数据块。编码器 20 部署于服务器 403 中，服务器 403 调用编码器 20 对目标视频包含的每一个视频帧中的各个数据块的场景复杂度进行分析，确定出对每一个数据块的编码模式，并分别按照确定得到的编码模式对每一个数据块进行编码，最终得到编码后的目标视频。服务器 403 将编码后的目标视频传输至参与视频会话的其他终端中；其他终端接收到编码后的目标视频，解析并播放目标视频，以实现 P 个终端参与的视频会话。

本申请实施例中，参与视频会话的各个终端或服务器可以调用编码器对视频会话过程中产生的目标视频进行编码，并且各个终端或服务器对目标视频中包含的各个数据块的编码模式是根据各个数据块的场景复杂度确定得到的，有效提升了对数据块的编码速度，从而提升了对目标视频的编码速度，在保障目标视频质量的同时，大幅度加快了目标视频的编码速度，从而提升了目标视频的流畅性，在一定程度上提升了视频会话的会话质量，优化了视频会话效果，提升用户体验。

图 5 是本申请实施例提供的一种视频处理方法的流程示意图，该视频处理方法可以由一个智能设备来执行，该智能设备可以是用户终端或者服务器，该用户终端可以是智能手机、平板电脑、智能可穿戴设备等带摄像功能的设备，该智能设备例如可以是图 4 所示的视频处理系统中的任一个终端或服务器，该视频处理方法包括以下步骤 S501 至步骤 S505：

S501，从待编码视频中获取目标视频帧，并从目标视频帧中确定待编码的目标数据块。待编码视频由多个连续的视频帧组成，从待编码视频中获取目标视频帧，目标视频帧是待编码视频中的任一个视频帧。目标视频帧中包含多个数据块，多个数据块中可能存在已编码的数据块和待编码的数据块，从目标视频帧中确定待编码的目标数据块，目标数据块是目标视频帧中任一个待编码的数据块。

S502，对目标数据块进行场景复杂度分析，得到数据块指标信息。对目标数据块进行

场景复杂度分析得到的数据块指标信息可以包括但不限于以下至少一种：目标数据块的预估失真参数、目标数据块的空间信息参数、目标数据块的时间信息参数。其中，预估失真（Distortion, Dist）参数是对帧内预测编码或帧间预测编码过程中的失真程度进行估计得到的。目标数据块的预估失真参数可以用于衡量重建后的目标数据块相比于原始的目标数据块的失真程度。原始的目标数据块是指未进行编码的目标数据块，重建后的目标数据块是指对目标数据块进行帧内预测编码或帧间预测编码后得到的编码后的目标数据块。目标数据块的空间信息参数可以是指对目标数据块进行计算得到的空间信息的数值；目标数据块的时间信息参数可以是指对目标数据块进行计算得到的时间信息的数值。

5 S503，将目标数据块划分成N个子数据块，并分别对每一个子数据块进行场景复杂度分析，得到子块指标信息，N为大于或等于2的整数。

子块指标信息可以包括N个子块指标数据，第i个子数据块是N个子数据块中的任一个子数据块，第i个子块指标数据是N个子块指标数据中的任一个子块指标数据，第i个子数据块与第i个子块指标数据对应，第i个子块指标数据是对第i个子数据块进行场景复杂度分析得到的， $i \in [1, N]$ 。

15 第i个子块指标数据可以包括但不限于以下至少一种：第i个子数据块的失真估计参数、第i个子数据块的空间信息参数、第i个子数据块的时间信息参数。其中，第i个子数据块的预估失真参数可以用于衡量重建后的第i个子数据块相比于原始的第i个子数据块的失真程度。原始的第i个子数据块是指未进行编码的第i个子数据块，重建后的第i个子数据块是指对第i个子数据块进行帧内预测编码或帧间预测编码后得到的编码后的第i个子数据块。第i个子数据块的空间信息参数可以是指对第i个子数据块进行计算得到的空间信息的数值；第i个子数据块的时间信息参数可以是指对第i个子数据块进行计算得到的时间信息的数值。

S504，根据数据块指标信息和子块指标信息，确定对目标数据块的编码模式。

25 在一个实施例，可以将数据块指标信息和子块指标信息输入联合统计模型，采用联合统计模型对数据块指标信息和子块指标信息进行计算，获取联合统计模型对数据块指标信息和子块指标信息进行计算后得到的输出值，并根据联合统计模型的输出值确定对目标数据块的编码模式。联合统计模型可以采用已确定编码模式的数据块的数据块指标信息，以及对已确定编码模式的数据块进行划分得到的N个子数据块的子块指标信息进行训练得到。在实施例中，联合统计模型可以对数据块指标信息和子块指标信息进行加权计算，得到输出值，其中加权因子可以通过对已确定编码模式的数据块及其子块的相关信息进行训练得到。

35 若联合统计模型的输出值满足第一条件，即输出值大于第一划分阈值，则确定对目标数据块的编码模式为第一模式。联合统计模型的输出值满足第一条件表明，对目标数据块进行划分得到的N个子数据块之间的关联性较弱，更倾向于将目标数据块划分为N个子数据块后分别对每一个子数据块进行编码。若联合统计模型的输出值满足第二条件，即输出值小于第二划分阈值，则确定对目标数据块的编码模式为第二模式。联合统计模型的输出值满足第二条件表明，对目标数据块进行划分得到的N个子数据块之间的关联性较强，更倾向于不对目标数据块进行划分，而直接对目标数据块进行编码。需要说明的是，第一划分阈值和第二划分阈值可以在联合统计模型的训练过程中训练得到。

40 S505，按照确定的编码模式对目标数据块进行编码。若编码模式为第一模式，则按照

确定的编码模式对目标数据块进行编码可以是指按照第一模式对目标数据块进行编码，具体可以是指将目标数据块划分成N个子数据块，并分别将每一个子数据块输入编码器中进行编码。若编码模式为第二模式，则按照确定的编码模式对目标数据块进行编码可以是指按照第二模式对目标数据块进行编码，具体可以是指直接将目标数据块输入编码器中进行编码。

本申请实施例中，将对目标数据块进行场景复杂度分析的场景复杂度分析结果（即数据块指标信息和子块指标信息）输入联合统计模型中进行计算，并根据联合统计模型计算后的输出值确定对目标数据块的编码模式。本申请实施例提供的视频处理方案适用于任意视频场景中，能够根据任意视频中待编码的目标数据块的场景复杂度自适应地对目标数据块的编码模式进行调整，从而确定出与目标数据块的场景复杂度相适应的编码模式，有效提升了对目标数据块的编码速度，进而提升了视频编码速度，并且能够获得视频的编码速度与视频的编码码率之间的最佳平衡。

图6是本申请实施例提供的另一种视频处理方法的流程示意图，该视频处理方法可以由一个智能设备来执行，该智能设备可以是用户终端或者服务器，该用户终端可以是智能手机、平板电脑、智能可穿戴设备等带摄像功能的设备，该智能设备例如可以是图4所示的视频处理系统中的任一个终端或服务器，该视频处理方法包括以下步骤S601至步骤S608：

S601，从待编码视频中获取目标视频帧，并从目标视频帧中确定待编码的目标数据块。

S602，对目标数据块进行场景复杂度分析，得到数据块指标信息。

S603，将目标数据块划分成N个子数据块，并分别对每一个子数据块进行场景复杂度分析，得到子块指标信息，N为大于或等于2的整数。

本申请实施例中步骤S601的执行过程与图5所示实施例中步骤S501的执行过程相同，步骤S602的执行过程与图5所示实施例中步骤S502的执行过程相同，步骤S603的执行过程与图5所示实施例中步骤S503的执行过程相同，具体执行过程可参见图5所示实施例的描述，在此不再赘述。

S604，将数据块指标信息和子块指标信息输入联合统计模型。

S605，获取联合统计模型对数据块指标信息和子块指标信息进行计算后得到的输出值。

S606，若输出值满足第三条件，则对与目标数据块相关的M个关联数据块进行场景复杂度分析，得到M个关联数据块的关联块指标信息，M为正整数。

S607，根据关联块指标信息确定对目标数据块的编码模式。

S606至S607中，若联合统计模型的输出值满足第三条件，即输出值小于或等于第一划分阈值，且输出值大于或等于第二划分阈值，表明对目标数据块进行划分得到的N个子数据块之间的关联性处于较强和较弱之间，没有明显倾向于将目标数据块划分为N个子数据块后分别对每一个子数据块进行编码，也没有明显倾向于不对目标数据块进行划分而直接对目标数据块进行编码。在这种情况下，确定与与目标数据块相关的M个关联数据块，并对M个关联数据块进行场景复杂度分析，得到M个关联数据块的关联块指标信息，并根据M个关联数据块的关联块指标信息确定对目标数据块的编码模式，M为正整数。M个关联数据块的关联块指标信息可以包括M个关联数据块中被划分为多个子数据块进行编码的关联数据块的第一数量。

在一个实施例中，若第一数量满足第四条件，即第一数量大于或等于第一数量阈值，

则表明与目标数据块相关的 M 个关联数据块的场景复杂程度较高,可以先尝试将目标数据块划分成 N 个待分析子数据块 (即 N 个子数据块),并对 N 个待分析子数据块进行场景复杂度分析,得到 N 个待分析子数据块的待分析子数据块指标信息,并根据待分析子数据块指标信息进一步确定对目标数据块的编码模式。需要说明的是,第一数量阈值可以根据经验值设定,例如,与目标数据块相关的关联数据块的数量为 3 个,可以将第一数量阈值设置为 2。以图 3 为例,与目标数据块 301 相关的 3 个关联数据块分别是关联数据块 302、关联数据块 303 和关联数据块 304,关联数据块 302 和关联数据块 303 被划分为多个子数据块进行编码,关联数据块 304 未被划分而直接进行编码,即 3 个关联数据块中被划分为多个子数据块进行编码的关联数据块的第一数量为 2 个。第一数量满足第四条条件,那么可以确定与目标数据块 301 相关的 3 个关联数据块的场景复杂程度较高,可以先尝试将目标数据块 301 划分成 N 个待分析子数据块。待分析子数据块指标信息可以包括 N 个待分析子数据块中满足进一步划分条件的待分析子数据块的第二数量。具体地,待分析子数据块指标信息可以是指采用联合统计模型对 N 个待分析子数据块中的每一个待分析子数据块的数据块指标信息和将每一个待分析数据块划分成的 N 个子数据块的子块指示信息进行计算得到的 N 个输出值中,满足第一条件的输出值的第二数量。

在一个实施例中,若第二数量满足第五条条件,即第二数量大于或等于第二数量阈值,可以确定对目标数据块的编码模式为第一模式。若第二数量不满足第五条条件,即第二数量小于第二数量阈值,尝试直接对目标数据块进行编码,并根据直接对目标数据块进行编码得到的目标数据块的编码信息,以及根据将目标数据块划分成 N 个子数据块并分别对每一个子数据块进行编码得到的 N 个子数据块的编码信息,确定对目标数据块的编码模式。需要说明的是,第二数量阈值也可以根据经验值设定。例如,将目标数据块划分成 4 个待分析子数据块,可以将第二数量阈值设置为 3。若 4 个待分析子数据块全部满足进一步划分条件,即 4 个待分析子数据块中满足进一步划分条件的待分析子数据块的第二数量为 4 个,第二数量满足第五条条件,可以确定对目标数据块的编码模式为第一模式。

在一个实施例中,若第一数量不满足第四条条件,即第一数量小于第一数量阈值,则表明与目标数据块相关的 M 个关联数据块的场景复杂程度较低,可以先尝试不对数据块进行划分而直接对目标数据块进行编码,并获取对目标数据块进行编码得到的目标数据块的编码信息,目标数据块的编码信息可以包括目标数据块的编码失真参数。需要说明的是,目标数据块的失真估计参数可以是指目标数据块的编码过程中失真程度的估计值,而目标数据块的编码失真参数可以用于衡量编码后的目标数据块 (即重建后的目标数据块) 相比于编码前的目标数据块 (即原始的目标数据块) 的实际失真程度。进一步地,根据目标数据块的编码失真参数和对目标数据块进行量化的量化参数 (Quantization Parameter, QP) 计算得到编码参数,编码参数的计算过程可参见下述公式 1:

$$\text{Code} = \text{Dist}/\text{QP}^2 \quad \text{公式 1}$$

如上述公式 1, Code 表示编码参数, Dist 表示目标数据块的编码失真参数, QP 表示量化参数。

在一个实施例中,若编码参数满足第六条条件,即编码参数小于第三划分阈值,则可以确定对目标数据块的编码模式为第二模式。在另一个实施例中,若编码参数不满足第六条条件,即编码参数大于或等于第三划分阈值,尝试将目标数据块划分成 N 个子数据块,并分别对每一个子数据块进行编码,并根据直接对目标数据块进行编码得到的目标数据块的编

码信息, 以及根据将目标数据块划分成N个子数据块并分别对每一个子数据块进行编码得到的N个子数据块的编码信息, 确定对目标数据块的编码模式。需要说明的是, 第三划分阈值可以在联合统计模型的训练过程中训练得到。

在以上两个实施例, 目标数据块的编码信息还可以包括目标数据块的第一编码率失真损失参数; N个子数据块的编码信息可以包括N个子数据块的第二编码率失真损失参数。第二编码率失真损失参数是根据N个子数据块中的每一个子数据块的第三编码率失真损失参数计算得到的, 例如第二编码率失真损失参数可以是N个子数据块中的每一个子数据块的第三编码率失真损失参数的和。其中, 第一编码率失真损失参数是根据按照直接对目标数据块进行编码的编码模式的编码码率与编码失真参数计算得到的, 例如第一编码率失真损失参数可以是编码码率与编码失真参数的比值。第一编码率失真损失参数可以用于衡量直接对目标数据块进行编码的编码效果, 第一编码率失真损失参数越小, 对目标数据块进行编码的效果越好。目标数据块的编码效果好可以是指在保证编码后的目标数据块相比于编码前的目标数据块的失真程度较低的情况下, 目标数据块的编码码率较低。类似的, 第二编码率失真损失参数可以用于衡量将目标数据块划分成N个子数据块并分别对每一个子数据块进行编码的编码效果。第二编码率失真损失参数越小, 对目标数据块进行编码的效果越好。若第一编码率失真损失参数大于或等于第二编码率失真损失参数, 则确定对目标数据块的编码模式为第一模式; 若第一编码率失真损失参数小于第二编码率失真损失参数, 则确定对目标数据块的编码模式为第二模式。

S608, 按照确定的编码模式对目标数据块进行编码。

本申请实施例中步骤S608的执行过程与图5所示实施例中步骤S505的执行过程相同, 具体执行过程可参见图5所示实施例的描述, 在此不再赘述。

本申请实施例中, 与目标数据块相关的关联数据块的场景复杂度在一定程度上也能反映目标数据块的场景复杂程度, 因此, 本申请实施例根据目标数据块的场景复杂度(即联合统计模型的输出值)以及与目标数据块相关的关联数据块的场景复杂度(即关联块指示信息)联合确定对目标数据块的编码模式。本申请实施例提供的视频处理方案具有普适性, 能够根据任意视频中待编码的目标数据块的场景复杂度自适应地对目标数据块的编码模式进行调整, 从而确定出与目标数据块的场景复杂度相适应的编码模式, 有效提升了对目标数据块的编码速度, 进而提升了视频编码速度, 并且能够在不损失视频编码效率的情况下, 有效提升视频编码速度, 并获得视频编码速度与视频编码码率之间的最佳平衡。

图7是本申请实施例提供的另一种视频处理方法的流程示意图, 该视频处理方法可以由一个智能设备来执行, 该智能设备可以是用户终端或者服务器, 该用户终端可以是智能手机、平板电脑、智能可穿戴设备等带摄像功能的设备, 该智能设备例如可以是图4所示的视频处理系统中的任一个终端或服务器, 该视频处理方法包括以下步骤S701至步骤S720:

S701, 从待编码视频中获取目标视频帧, 并从目标视频帧中确定待编码的目标数据块。

S702, 对目标数据块进行场景复杂度分析, 得到数据块指标信息。

S703, 将目标数据块划分成N个子数据块, 并分别对每一个子数据块进行场景复杂度分析, 得到子块指标信息。

S704, 将数据块指标信息和子块指标信息输入联合统计模型中进行计算, 得到联合统计模型的输出值。

S705, 判断输出值是否满足第一条条件。若输出值满足第一条条件, 则执行步骤 S706; 若输出值不满足第一条条件, 则执行步骤 S707。

S706, 若输出值满足第一条条件, 则确定对目标数据块的编码模式为第一模式。步骤 S706 执行结束后, 执行步骤 S720。

5 S707, 若输出值不满足第一条条件, 则判断输出值是否满足第二条条件。若输出值满足第二条条件, 则执行步骤 S708; 若输出值不满足第二条条件, 则执行步骤 S709。

S708, 若输出值满足第二条条件, 则确定对目标数据块的编码模式为第二模式。步骤 S708 执行结束后, 执行步骤 S720。

10 S709, 输出值不满足第二条条件, 则判断输出值是否满足第三条条件。若输出值满足第三条条件, 则执行步骤 S710。

S710, 若输出值满足第三条条件, 则确定与目标数据块相关的 M 个关联数据块中被划分为多个子数据块进行编码的关联数据块的第一数量。

S711, 判断第一数量是否满足第四条条件。若第一数量满足第四条条件, 则执行步骤 S712; 若第一数量不满足第四条条件, 则执行步骤 S716。

15 S712, 若第一数量满足第四条条件, 则优先尝试将目标子数据块划分成 N 个子数据块, 并分别对每一个子数据块进行编码。

S713, 确定 N 个子数据块中满足进一步划分条件的子数据块的第二数量。

20 S714, 判断第二数量是否满足第五条条件。若第二数量满足第五条条件, 则确定对目标数据块的编码模式为第一模式, 并执行步骤 S720。若第二数量不满足第五条条件, 则执行步骤 S715。

25 S715, 尝试不对目标数据进行划分, 直接对目标数据块进行编码, 并根据编码信息确定对目标数据块的编码模式。具体地, 获取直接对目标数据块进行编码得到的目标数据块的编码信息, 以及获取将目标数据块划分成 N 个子数据块并分别对每一个子数据块进行编码得到的 N 个子数据块的编码信息, 并根据目标数据块的编码信息和 N 个子数据块的编码信息确定对目标数据块的编码模式。具体执行过程可参见图 6 所示实施例, 步骤 S715 执行结束后, 执行步骤 S720。

S716, 若第一数量不满足第四条条件, 则优先尝试不对目标数据块进行划分, 直接对目标数据块进行编码, 并确定目标数据块的编码失真参数。

S717, 根据编码失真参数和量化参数计算得到编码参数。

30 S718, 判断编码参数是否满足第六条条件。若编码参数满足第六条条件, 则确定对目标数据块的编码模式为第二模式, 并执行步骤 S720; 若编码参数不满足第六条条件, 则执行步骤 S719。

35 S719, 尝试将目标子数据块划分成 N 个子数据块, 并分别对每一个子数据块进行编码, 并根据编码信息确定对目标数据块的编码模式。具体地, 获取直接对目标数据块进行编码得到的目标数据块的编码信息, 以及获取将目标数据块划分成 N 个子数据块并分别对每一个子数据块进行编码得到的 N 个子数据块的编码信息, 并根据目标数据块的编码信息和 N 个子数据块的编码信息确定对目标数据块的编码模式。具体执行过程可参见图 6 所示实施例, 步骤 S719 执行结束后, 执行步骤 S720。

40 S720, 按照确定得到的编码模式对目标数据块进行编码。按照确定得到的编码模式对目标数据块进行编码后, 可以从待编码视频帧中确定除目标数据块之外的另一个待编码的

数据块,并根据该待编码的数据块的场景复杂度确定对该待编码的数据块的编码模式,并按照确定得到的编码模式,对该待编码的数据块进行编码。根据待编码的数据块的场景复杂度确定对该待编码的数据块的编码模式的过程,与根据目标数据块的场景复杂度确定对目标数据块的编码模式的过程相同。

- 5 本申请实施例中,与目标数据块相关的关联数据块的场景复杂度在一定程度上也能反映目标数据块的场景复杂程度,因此,本申请实施例根据目标数据块的场景复杂度(即联合统计模型的输出值)以及与目标数据块相关的关联数据块的场景复杂度(即关联块指示信息)联合确定对目标数据块的编码模式。本申请实施例设置了6个判断条件,对目标数据块的场景复杂度进行多角度综合分析,从而使得最终确定得到的对目标数据块的编码模式与目标数据块的场景复杂度的适应程度较高,有效提升了对目标数据块的编码速度,进而提升了视频编码速度,并且能够在保证编码后的视频的清晰度与流畅度的同时,在不损失视频编码质量与视频编码效率的情况下,有效提升视频编码速度。

10 请参见图8,图8是本申请实施例提供的一种视频处理装置的结构示意图,本申请实施例的所述视频处理装置80可以设置在智能设备中,该智能设备可以是智能终端,也可以是服务器,该视频处理装置80可以用于执行图5、图6或图7所示的视频处理方法中的相应步骤,该视频处理装置80包括如下单元。

获取单元801,用于从待编码视频中获取目标视频帧,并从目标视频帧中确定待编码的目标数据块;

20 处理单元802,用于对目标数据块进行场景复杂度分析,得到数据块指标信息;将目标数据块划分成N个子数据块,并分别对每一个子数据块进行场景复杂度分析,得到子块指标信息,N为大于或等于2的整数;根据数据块指标信息和子块指标信息,确定对目标数据块的编码模式;以及,按照确定的编码模式对目标数据块进行编码。

在一个实施例中,数据块指标信息包括:目标数据块的失真估计参数、目标数据块的空间信息参数、目标数据块的时间信息参数中的任意一种或多种。

25 子块指标信息包括:N个子块指标数据,其中,N个子块指标数据中的第i个子块指标数据包括:N个子数据块中的第i个子数据块的失真估计参数、第i个子数据块的空间信息参数、第i个子数据块的时间信息参数中的任意一种或多种, $i \in [1, N]$ 。

在一个实施例中,处理单元802,具体用于:

将数据块指标信息和子块指标信息输入联合统计模型;

30 获取联合统计模型对数据块指标信息和子块指标信息进行计算后得到的输出值;

根据输出值确定对目标数据块的编码模式。

在一个实施例中,处理单元802,具体用于:

若输出值满足第一条件,则确定编码模式为第一模式;

将目标数据块划分成N个子数据块,并分别将每一个子数据块输入编码器中进行编码;

35 其中,输出值满足第一条件是指:输出值大于第一划分阈值。

在一个实施例中,处理单元802,具体用于:

若输出值满足第二条件,则确定编码模式为第二模式;

将目标数据块输入编码器中进行编码;

其中,输出值满足第二条件是指:输出值小于第二划分阈值。

40 在一个实施例中,处理单元802,具体用于:

若输出值满足第三条件,则对与目标数据块相关的M个关联数据块进行场景复杂度分析,得到M个关联数据块的关联块指标信息,M为正整数;

根据关联块指标信息确定对目标数据块的编码模式;

其中,输出值满足第三条件是指:输出值小于或等于第一划分阈值,且输出值大于或等于第二划分阈值。

在一个实施例中,处理单元802,具体用于:

获取关联块指标信息,关联块指标信息包括:M个关联数据块中被划分为多个子数据块进行编码的关联数据块的第一数量;

若第一数量满足第四条件,则将目标数据块划分成N个待分析子数据块;

对N个待分析子数据块进行场景复杂度分析,确定N个待分析子数据块的待分析子数据块指标信息;

根据待分析子数据块指标信息确定对目标数据块的编码模式;

其中,第一数量满足第四条件是指:第一数量大于或等于第一数量阈值。

在一个实施例中,处理单元802,具体用于:

获取待分析子数据块指标信息,待分析子数据块指标信息包括:N个待分析子数据块中满足进一步划分条件的待分析子数据块的第二数量;

若第二数量满足第五条件,则确定编码模式为第一模式;

若第二数量不满足第五条件,则将目标数据块输入编码器中进行编码,并根据编码得到的目标数据块的编码信息确定对目标数据块的编码模式;

其中,第二数量满足第五条件是指:第二数量大于或等于第二数量阈值;第二数量不满足第五条件是指:第二数量小于第二数量阈值。

在一个实施例中,处理单元802,还用于:

若第一数量不满足第四条件,则将目标数据块输入编码器中进行编码;

根据编码得到的目标数据块的编码信息确定对目标数据块的编码模式;

其中,第一数量不满足第四条件是指:第一数量小于第一数量阈值。

在一个实施例中,目标数据块的编码信息包括目标数据块的编码失真参数;处理单元802,具体用于:

根据目标数据块的编码失真参数和对目标数据块进行量化的量化参数计算得到编码参数;

若编码参数满足第六条件,则确定编码模式为第二模式;

若编码参数不满足第六条件,则将目标数据块划分成N个子数据块,并分别将N个子数据块中的每一个子数据块输入编码器中进行编码;

根据目标数据块的编码信息和编码得到的N个子数据块的编码信息确定对目标数据块的编码模式;

其中,编码参数满足第六条件是指:编码参数小于第三划分阈值;编码参数不满足第六条件是指:编码参数大于或等于第三划分阈值。

在一个实施例中,目标数据块的编码信息还包括目标数据块的第一编码率失真损失参数;N个子数据块的编码信息包括N个子数据块的第二编码率失真损失参数,第二编码率失真损失参数是根据N个子数据块中的每一个子数据块的第三编码率失真损失参数计算得

到的;处理单元802,具体用于:

若第一编码率失真损失参数大于或等于第二编码率失真损失参数，则确定对目标数据块的编码模式为第一模式；

若第一编码率失真损失参数小于第二编码率失真损失参数，则确定对目标数据块的编码模式为第二模式。

5 根据本申请的一个实施例，图 8 所示的视频处理装置 80 中的各个单元可以分别或全部合并为一个或若干个另外的单元来构成，或者其中的某个（些）单元还可以再划分为功能上更小的多个单元来构成，这可以实现同样的操作，而不影响本申请的实施例的技术效果的实现。上述单元是基于逻辑功能划分的，在实际应用中，一个单元的功能也可以由多个单元来实现，或者多个单元的功能由一个单元实现。在本申请的其它实施例中，该视频
10 处理装置 80 也可以包括其它单元，在实际应用中，这些功能也可以由其它单元协助实现，并且可以由多个单元协作实现。根据本申请的另一个实施例，可以通过在包括例如中央处理单元（CPU）、随机存取存储介质（RAM）、只读存储介质（ROM）等处理元件和存储元件的通用计算机的通用计算设备上运行能够执行如图 5、图 6 或图 7 中所示的相应方法所涉及的各步骤的计算机程序（包括程序代码），来构造如图 8 中所示的视频处理装置
15 80，以及来实现本申请实施例的视频处理方法。计算机程序可以记载于例如计算机可读存储介质上，并通过计算机可读存储介质装载于图 4 所示的视频处理系统的任一个终端（例如第一终端 401 或第二终端 402 等等）或服务器 403 中，并在其中运行。

本申请实施例中，通过对目标视频帧中待编码的目标数据块进行场景复杂度分析，分别得到对整个待编码的目标数据块的分析结果和将待编码的目标数据块划分成多个子数据块后的分析结果，并基于这些分析结果选择编码模式，可以为当前待编码的目标数据块
20 较为准确地确定出合适的编码模式，能够有效提升对目标数据块的编码速度，提升视频编码效率。

请参见图 9，图 9 是本申请实施例提供的一种智能设备的结构示意图，该智能设备 90
25 至少包括处理器 901 以及存储器 902。其中，处理器 901 以及存储器 902 可通过总线或者其它方式连接。

处理器 901 可以是中央处理器（Central Processing Unit，CPU）。处理器 901 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（Application-Specific Integrated Circuit，ASIC），可编程逻辑器件（Programmable Logic Device，PLD）等。上述 PLD 可以是现场可编程逻辑门阵列（Field-Programmable Gate Array，FPGA），通用阵列逻辑（Generic
30 Array Logic，GAL）等。

存储器 902 可以包括易失性存储器（Volatile Memory），例如随机存取存储器（Random-Access Memory，RAM）；存储器 902 也可以包括非易失性存储器（Non-Volatile Memory），例如快闪存储器（Flash Memory），固态硬盘（Solid-State Drive，SSD）等；存储器 902 还可以包括上述种类的存储器的组合。

35 存储器 902 用于存储计算机程序，计算机程序包括计算机指令，处理器 901 用于执行计算机指令。处理器 901（或称 CPU（Central Processing Unit，中央处理器））是智能设备 90 的计算核心以及控制核心，其适于实现一条或多条计算机指令，具体适于加载并执行一条或多条计算机指令从而实现相应方法流程或相应功能。

该智能设备 90 可以是图 4 所示的视频处理系统中的任一个终端（例如第一终端 401
40 或第二终端 402 等等）或服务器 403；该存储器 902 中存储有计算机程序，计算机程序包

括一条或多条计算机指令;由处理器 901 加载并执行一条或多条计算机指令,以实现图 5、图 6 或图 7 所示方法实施例中的相应步骤;具体实现中,存储器 902 中的计算机指令由处理器 901 加载并执行如下步骤:

从待编码视频中获取目标视频帧,并从目标视频帧中确定待编码的目标数据块;

5 对目标数据块进行场景复杂度分析,得到数据块指标信息;

将目标数据块划分成 N 个子数据块,并分别对每一个子数据块进行场景复杂度分析,得到子块指标信息, N 为大于或等于 2 的整数;

根据数据块指标信息和子块指标信息,确定对目标数据块的编码模式;

按照确定的编码模式对目标数据块进行编码。

10 在一个实施例中,数据块指标信息包括:目标数据块的失真估计参数、目标数据块的空间信息参数、目标数据块的时间信息参数中的任意一种或多种;

子块指标信息包括: N 个子块指标数据,其中, N 个子块指标数据中的第 i 个子块指标数据包括: N 个子数据块中的第 i 个子数据块的失真估计参数、第 i 个子数据块的空间信息参数、第 i 个子数据块的时间信息参数中的任意一种或多种, $i \in [1, N]$ 。

15 在一个实施例中,存储器 902 中的计算机指令由处理器 901 加载时具体执行如下步骤:

将数据块指标信息和子块指标信息输入联合统计模型;

获取联合统计模型对数据块指标信息和子块指标信息进行计算后得到的输出值;

根据输出值确定对目标数据块的编码模式。

在一个实施例中,存储器 902 中的计算机指令由处理器 901 加载时具体执行如下步骤:

20 若输出值满足第一条件,则确定编码模式为第一模式;

将目标数据块划分成 N 个子数据块,并分别将每一个子数据块输入编码器中进行编码;其中,输出值满足第一条件是指:输出值大于第一划分阈值。

在一个实施例中,存储器 902 中的计算机指令由处理器 901 加载时具体执行如下步骤:

若输出值满足第二条件,则确定编码模式为第二模式;

25 将目标数据块输入编码器中进行编码;

其中,输出值满足第二条件是指:输出值小于第二划分阈值。

在一个实施例中,存储器 902 中的计算机指令由处理器 901 加载时具体执行如下步骤:

若输出值满足第三条件,则对与目标数据块相关的 M 个关联数据块进行场景复杂度分析,得到 M 个关联数据块的关联块指标信息, M 为正整数;

30 根据关联块指标信息确定对目标数据块的编码模式;

其中,输出值满足第三条件是指:输出值小于或等于第一划分阈值,且输出值大于或等于第二划分阈值。

在一个实施例中,存储器 902 中的计算机指令由处理器 901 加载时具体执行如下步骤:

35 获取关联块指标信息,关联块指标信息包括: M 个关联数据块中被划分为多个子数据块进行编码的关联数据块的第一数量;

若第一数量满足第四条件,则将目标数据块划分成 N 个待分析子数据块;

对 N 个待分析子数据块进行场景复杂度分析,确定 N 个待分析子数据块的待分析子数据块指标信息;

根据待分析子数据块指标信息确定对目标数据块的编码模式;

40 其中,第一数量满足第四条件是指:第一数量大于或等于第一数量阈值。

在一个实施例中,存储器 902 中的计算机指令由处理器 901 加载时具体执行如下步骤:

获取待分析子数据块指标信息,待分析子数据块指标信息包括:N 个待分析子数据块中满足进一步划分条件的待分析子数据块的第二数量;

若第二数量满足第五条件,则确定编码模式为第一模式;

- 5 若第二数量不满足第五条件,则将目标数据块输入编码器中进行编码,并根据编码得到的目标数据块的编码信息确定对目标数据块的编码模式;

其中,第二数量满足第五条件是指:第二数量大于或等于第二数量阈值;第二数量不满足第五条件是指:第二数量小于第二数量阈值。

在一个实施例中,存储器 902 中的计算机指令由处理器 901 加载时还执行如下步骤:

- 10 若第一数量不满足第四条件,则将目标数据块输入编码器中进行编码;

根据编码得到的目标数据块的编码信息确定对目标数据块的编码模式;

其中,第一数量不满足第四条件是指:第一数量小于第一数量阈值。

在一个实施例中,目标数据块的编码信息包括目标数据块的编码失真参数;存储器 902 中的计算机指令由处理器 901 加载时具体执行如下步骤:

- 15 根据目标数据块的编码失真参数和对目标数据块进行量化的量化参数计算得到编码参数;

若编码参数满足第六条件,则确定编码模式为第二模式;

若编码参数不满足第六条件,则将目标数据块划分成 N 个子数据块,并分别将 N 个子数据块中的每一个子数据块输入编码器中进行编码;

- 20 根据目标数据块的编码信息和编码得到的 N 个子数据块的编码信息确定对目标数据块的编码模式;

其中,编码参数满足第六条件是指:编码参数小于第三划分阈值;编码参数不满足第六条件是指:编码参数大于或等于第三划分阈值。

- 25 在一个实施例中,目标数据块的编码信息还包括目标数据块的第一编码率失真损失参数;N 个子数据块的编码信息包括 N 个子数据块的第二编码率失真损失参数,第二编码率失真损失参数是根据 N 个子数据块中的每一个子数据块的第三编码率失真损失参数计算得到的;存储器 902 中的计算机指令由处理器 901 加载时具体执行如下步骤:

若第一编码率失真损失参数大于或等于第二编码率失真损失参数,则确定对目标数据块的编码模式为第一模式;

- 30 若第一编码率失真损失参数小于第二编码率失真损失参数,则确定对目标数据块的编码模式为第二模式。

- 35 本申请实施例中,通过对目标视频帧中待编码的目标数据块进行场景复杂度分析,分别得到对整个待编码的目标数据块的分析结果和将待编码的目标数据块划分成多个子数据块后的分析结果,并基于这些分析结果选择编码模式,可以为当前待编码的目标数据块较为准确地确定出合适的编码模式,能够有效提升对目标数据块的编码速度,提升视频编码效率。

- 40 根据本申请的一个方面,提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行上述各种可选方式中提供的视频处理方法。

本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通过程序来指令相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，计算机可读存储介质可以包括：闪存盘、只读存储器（Read-Only Memory, ROM）、随机存取器（Random Access Memory, RAM）、磁盘或光盘等。

- 5 以上所揭露的仅为本申请一种较佳实施例而已，当然不能以此来限定本申请之权利范围，本领域普通技术人员可以理解实现上述实施例的全部或部分流程，并依本申请权利要求所作的等同变化，仍属于本申请所涵盖的范围。

权利要求书

- 1、一种视频处理方法，由智能设备执行，所述方法包括：
从待编码视频中获取目标视频帧，并从所述目标视频帧中确定待编码的目标数据块；
对所述目标数据块进行场景复杂度分析，得到数据块指标信息；
5 将所述目标数据块划分成 N 个子数据块，并分别对每一个子数据块进行场景复杂度分析，得到子块指标信息， N 为大于或等于 2 的整数；
根据所述数据块指标信息和所述子块指标信息，确定对所述目标数据块的编码模式；
按照确定的所述编码模式对所述目标数据块进行编码。
- 2、如权利要求 1 所述的方法，其中，所述数据块指标信息包括：所述目标数据块的失真估计参数、所述目标数据块的空间信息参数、所述目标数据块的时间信息参数中的任意一种或多种；
10 所述子块指标信息包括： N 个子块指标数据，其中，所述 N 个子块指标数据中的第 i 个子块指标数据包括：所述 N 个子数据块中的第 i 个子数据块的失真估计参数、所述第 i 个子数据块的空间信息参数、所述第 i 个子数据块的时间信息参数中的任意一种或多种， $i \in [1, N]$ 。
- 3、如权利要求 1 所述的方法，其中，所述根据所述数据块指标信息和所述子块指标信息，确定对所述目标数据块的编码模式，包括：
将所述数据块指标信息和所述子块指标信息输入联合统计模型；
获取所述联合统计模型对所述数据块指标信息和所述子块指标信息进行计算后得到的
20 的输出值；
根据所述输出值确定对所述目标数据块的编码模式。
- 4、如权利要求 3 所述的方法，其中，所述根据所述输出值确定对所述目标数据块的编码模式，包括：
若所述输出值满足第一条件，则确定所述编码模式为第一模式；
25 所述按照确定的所述编码模式对所述目标数据块进行编码，包括：
将所述目标数据块划分成所述 N 个子数据块，并分别将每一个子数据块输入编码器中进行编码；
其中，所述输出值满足所述第一条件是指：所述输出值大于第一划分阈值。
- 5、如权利要求 3 所述的方法，其中，所述根据所述输出值确定对所述目标数据块的
30 编码模式，包括：
若所述输出值满足第二条件，则确定所述编码模式为第二模式；
所述按照确定的所述编码模式对所述目标数据块进行编码，包括：
将所述目标数据块输入编码器中进行编码；
其中，所述输出值满足所述第二条件是指：所述输出值小于第二划分阈值。
- 6、如权利要求 3 所述的方法，其中，所述根据所述输出值确定对所述目标数据块的
35 编码模式，包括：
若所述输出值满足第三条件，则对与所述目标数据块相关的 M 个关联数据块进行场景复杂度分析，得到所述 M 个关联数据块的关联块指标信息， M 为正整数；

根据所述关联块指标信息确定对所述目标数据块的编码模式;

其中,所述输出值满足所述第三条件是指:所述输出值小于或等于第一划分阈值,且所述输出值大于或等于第二划分阈值。

7、如权利要求6所述的方法,其中,所述根据所述关联块指标信息确定对所述目标数据块的编码模式,包括:

获取所述关联块指标信息,所述关联块指标信息包括:所述M个关联数据块中被划分为多个子数据块进行编码的关联数据块的第一数量;

若所述第一数量满足第四条件,则将所述目标数据块划分成N个待分析子数据块;

对所述N个待分析子数据块进行场景复杂度分析,确定所述N个待分析子数据块的待分析子数据块指标信息;

根据所述待分析子数据块指标信息确定对所述目标数据块的编码模式;

其中,所述第一数量满足所述第四条件是指:所述第一数量大于或等于第一数量阈值。

8、如权利要求7所述的方法,其中,所述根据所述待分析子数据块指标信息确定对所述目标数据块的编码模式,包括:

获取所述待分析子数据块指标信息,所述待分析子数据块指标信息包括:所述N个待分析子数据块中满足进一步划分条件的待分析子数据块的第二数量;

若所述第二数量满足第五条件,则确定所述编码模式为第一模式;

若所述第二数量不满足所述第五条件,则将所述目标数据块输入编码器中进行编码,并根据编码得到的所述目标数据块的编码信息确定对所述目标数据块的编码模式;

其中,所述第二数量满足所述第五条件是指:所述第二数量大于或等于第二数量阈值;所述第二数量不满足所述第五条件是指:所述第二数量小于所述第二数量阈值。

9、如权利要求7所述的方法,其中,所述根据所述关联块指标信息确定对所述目标数据块的编码模式,还包括:

若所述第一数量不满足所述第四条件,则将所述目标数据块输入编码器中进行编码;根据编码得到的所述目标数据块的编码信息确定对所述目标数据块的编码模式;

其中,所述第一数量不满足所述第四条件是指:所述第一数量小于所述第一数量阈值。

10、如权利要求8或9所述的方法,其中,所述目标数据块的编码信息包括所述目标数据块的编码失真参数;所述根据编码得到的所述目标数据块的编码信息确定对所述目标数据块的编码模式,包括:

根据所述目标数据块的编码失真参数和对所述目标数据块进行量化的量化参数计算得到编码参数;

若所述编码参数满足第六条件,则确定所述编码模式为第二模式;

若所述编码参数不满足所述第六条件,则将所述目标数据块划分成所述N个子数据块,并分别将所述N个子数据块中的每一个子数据块输入所述编码器中进行编码;

根据所述目标数据块的编码信息和编码得到的所述N个子数据块的编码信息确定对所述目标数据块的编码模式;

其中,所述编码参数满足所述第六条件是指:所述编码参数小于第三划分阈值;所述编码参数不满足所述第六条件是指:所述编码参数大于或等于所述第三划分阈值。

11、如权利要求10所述的方法,其中,所述目标数据块的编码信息还包括所述目标数据块的第一编码率失真损失参数;所述N个子数据块的编码信息包括所述N个子数据块的

第二编码率失真损失参数,所述第二编码率失真损失参数是根据所述N个子数据块中的每一个子数据块的第三编码率失真损失参数计算得到的;所述根据所述目标数据块的编码信息和所述N个子数据块的编码信息确定对所述目标数据块的编码模式,包括:

- 5 若所述第一编码率失真损失参数大于或等于所述第二编码率失真损失参数,则确定对所述目标数据块的编码模式为所述第一模式;

若所述第一编码率失真损失参数小于所述第二编码率失真损失参数,则确定对所述目标数据块的编码模式为所述第二模式。

12、一种视频处理装置,包括:

- 10 获取单元,用于从待编码视频中获取目标视频帧,并从所述目标视频帧中确定待编码的目标数据块;

- 15 处理单元,用于对所述目标数据块进行场景复杂度分析,得到数据块指标信息;将所述目标数据块划分成N个子数据块,并分别对每一个子数据块进行场景复杂度分析,得到子块指标信息,N为大于或等于2的整数;根据所述数据块指标信息和所述子块指标信息,确定对所述目标数据块的编码模式;以及,按照确定的所述编码模式对所述目标数据块进行编码。

13、一种智能设备,包括:

处理器,适于实现计算机程序;以及,

存储器,所述存储器存储有计算机程序,所述计算机程序被所述处理器运行时,实现如权利要求1至11任一项所述的视频处理方法。

- 20 14、一种计算机可读存储介质,其中,所述计算机可读存储介质存储有计算机程序,所述计算机程序被处理器读取并运行时,实现如权利要求1至11任一项所述的视频处理方法。