



(12)发明专利申请

(10)申请公布号 CN 111684800 A

(43)申请公布日 2020.09.18

(21)申请号 201980012173.6

(22)申请日 2019.02.07

(30)优先权数据

15/890,710 2018.02.07 US

(85)PCT国际申请进入国家阶段日

2020.08.06

(86)PCT国际申请的申请数据

PCT/US2019/017134 2019.02.07

(87)PCT国际申请的公布数据

W02019/157235 EN 2019.08.15

(71)申请人 奈飞公司

地址 美国加利福尼亚州

(72)发明人 李智 克里斯托斯·巴姆皮斯

(74)专利代理机构 北京东方亿思知识产权代理
有限责任公司 11258

代理人 林强

(51)Int.Cl.

H04N 19/154(2014.01)

H04N 19/136(2014.01)

G06K 9/00(2006.01)

G06K 9/03(2006.01)

G06T 7/00(2017.01)

G06N 20/20(2019.01)

G06K 9/62(2006.01)

G06N 3/08(2006.01)

权利要求书3页 说明书16页 附图5页

(54)发明名称

用于基于互补感知质量模型来预测感知视频
频质量的技术

(57)摘要

在各种实施例中,集成预测应用计算从视觉内容导出的重构视觉内容的质量得分。集成预测应用基于第一特征集的第一组值和第一模型来计算重构视频内容的第一质量得分,该第一模型将第一组值与第一质量得分相关联。集成预测应用基于第二特征集的第二组值和第二模型来计算重构视频内容的第二质量得分,该第二模型将第二组值与第二质量得分相关联。随后,集成预测应用基于第一质量得分和第二质量得分来确定重构视频内容的整体质量得分。整体质量得分指示与流式视频内容相关联的视觉质量的水平。

500



1. 一种计算机实现的方法,包括:

基于第一特征集的第一组值和第一模型来计算重构视频内容的第一质量得分,所述第一模型将所述第一组值与所述第一质量得分相关联,其中,所述第一模型是经由第一机器学习算法生成的;

基于第二特征集的第二组值和第二模型来计算所述重构视频内容的第二质量得分,所述第二模型将所述第二组值与所述第二质量得分相关联;以及

基于所述第一质量得分和所述第二质量得分来确定所述重构视频内容的整体质量得分,其中,所述整体质量得分指示与流式视频内容相关联的视觉质量的水平。

2. 根据权利要求1所述的计算机实现的方法,其中,所述第一质量得分和所述第二质量得分与所述重构视频内容中包括的第一帧相关联,并且确定所述整体质量得分包括:

基于所述第一质量得分和所述第二质量得分来确定与所述第一帧相关联的第一聚合质量得分;以及

对所述第一聚合质量得分和第二聚合质量得分执行一个或多个时间池化操作以计算所述整体质量得分,所述第二聚合质量得分与所述重构视频内容中包括的第二帧相关联。

3. 根据权利要求2所述的计算机实现的方法,其中,确定所述第一聚合质量得分包括计算所述第一质量得分和所述第二质量得分的平均值或加权平均值。

4. 根据权利要求2所述的计算机实现的方法,其中,确定所述第一聚合质量得分包括基于第三模型将所述第一质量得分和所述第二质量得分与所述第一聚合质量得分相关联,所述第三模型是经由所述第一机器学习算法或第二机器学习算法生成的。

5. 根据权利要求1所述的计算机实现的方法,其中,所述第一特征集包括所述第二特征集中未包括的第一特征,并且所述第二特征集包括所述第一特征集中未包括的第二特征。

6. 根据权利要求1所述的计算机实现的方法,其中,所述第一特征集包括时间特征和空间特征。

7. 根据权利要求1所述的计算机实现的方法,其中,所述第一特征集包括以下各项中的至少一项:附加损伤测量特征、盲或无参考图像空间质量评估器特征以及视觉信息保真度特征。

8. 根据权利要求1所述的计算机实现的方法,还包括:基于所述第一机器学习算法和多个主观得分来生成所述第一模型,其中,所述多个主观得分中包括的每个主观得分与不同的重构测试视频内容相关联。

9. 根据权利要求8所述的计算机实现的方法,还包括:基于所述多个主观得分来生成所述第二模型。

10. 一种计算机可读存储介质,所述计算机可读存储介质包括指令,所述指令在由处理器执行时使得所述处理器执行以下步骤:

基于第一特征集的第一组值和第一模型来计算重构视频内容的第一质量得分,所述第一模型将所述第一组值与所述第一质量得分相关联,其中,所述第一模型是经由第一机器学习算法生成的;

基于第二特征集的第二组值和第二模型来计算所述重构视频内容的第二质量得分,所述第二模型将所述第二组值与所述第二质量得分相关联,其中,所述第二模型是经由所述第一机器学习算法或第二机器学习算法生成的;以及

基于所述第一质量得分和所述第二质量得分来确定所述重构视频内容的整体质量得分,其中,所述整体质量得分指示与流式视频内容相关联的视觉质量的水平。

11.根据权利要求10所述的计算机可读存储介质,其中,所述第一质量得分和所述第二质量得分与所述重构视频内容中包括的第一帧相关联,并且确定所述整体质量得分包括:

基于所述第一质量得分和所述第二质量得分来确定与所述第一帧相关联的第一聚合质量得分;以及

对所述第一聚合质量得分和第二聚合质量得分执行一个或多个时间池化操作以计算所述整体质量得分,所述第二聚合质量得分与所述重构视频内容中包括的第二帧相关联。

12.根据权利要求11所述的计算机可读存储介质,其中,确定所述第一聚合质量得分包括计算所述第一质量得分和所述第二质量得分的平均值或加权平均值。

13.根据权利要求11所述的计算机可读存储介质,其中,执行所述一个或多个时间池化操作包括对所述第一聚合质量得分和所述第二聚合质量得分执行以下各项中的至少一项:线性低通操作、非线性等级排列加权操作和算术平均操作。

14.根据权利要求10所述的计算机可读存储介质,其中,所述第一特征集包括所述第二特征集中未包括的第一特征,并且所述第二特征集包括所述第一特征集中未包括的第二特征。

15.根据权利要求10所述的计算机可读存储介质,其中,所述第一特征集包括时间特征和空间特征。

16.根据权利要求10所述的计算机可读存储介质,其中,所述第一特征集包括多个时间特征,并且所述多个时间特征中包括的每个时间特征与不同的缩放比例相关联。

17.根据权利要求10所述的计算机可读存储介质,还包括:在计算所述第一质量得分之前,将整体特征集中包括的每个特征分配给所述第一特征集和所述第二特征集中的至少一个。

18.根据权利要求10所述的计算机可读存储介质,还包括:基于所述第一机器学习算法和多个主观得分来生成所述第一模型,其中,所述多个主观得分中包括的每个主观得分与不同的重构测试视频内容相关联。

19.一种系统,包括:

存储器,所述存储器存储指令;以及

处理器,所述处理器耦合到所述存储器,并且所述处理器在执行所述指令时,被配置为:

基于第一特征集的第一组值和第一模型来计算重构视频内容的第一质量得分,所述第一模型将所述第一组值与所述第一质量得分相关联,其中,所述第一模型是基于多个主观得分生成的;

基于第二特征集的第二组值和第二模型来计算所述重构视频内容的第二质量得分,所述第二模型将所述第二组值与所述第二质量得分相关联,其中,所述第二模型是基于所述多个主观得分生成的;以及

基于所述第一质量得分和所述第二质量得分来确定所述重构视频内容的整体质量得分,其中,所述整体质量得分指示与流式视频内容相关联的视觉质量的水平。

20.根据权利要求19所述的系统,其中,所述第一质量得分和所述第二质量得分与所述

重构视频内容中包括的第一帧相关联,并且所述处理器被配置为通过以下方式来确定所述整体质量得分:

基于所述第一质量得分和所述第二质量得分来确定与所述第一帧相关联的第一聚合质量得分;以及

对所述第一聚合质量得分和第二聚合质量得分执行一个或多个时间池化操作以计算所述整体质量得分,所述第二聚合质量得分与所述重构视频内容中包括的第二帧相关联。

用于基于互补感知质量模型来预测感知视频质量的技术

[0001] 相关申请的交叉引用

[0002] 本申请要求于2018年2月7日提交的、序列号为15/890,710的美国专利申请的权益,其通过引用合并于此。

技术领域

[0003] 本发明的实施例总体上涉及视频技术,并且更具体地,涉及用于基于互补感知质量模型来预测感知视频质量的技术。

背景技术

[0004] 有效且准确地编码源视频内容对于实时传送高质量视频内容至关重要。由于经编码的视频内容质量的变化,期望实施质量控制以确保从经编码的源视频内容解码的重构视频内容的视觉质量是可接受的。手动验证重构视频内容的视觉质量非常耗时。因此,自动视频内容质量评估通常是编码和流式基础结构的组成部分。例如,在评估编码器和微调流式比特率以优化流式视频内容时,可以采用自动视频质量评估。

[0005] 用于自动评估重构视频内容的质量的一些方法涉及基于感知质量模型来计算质量得分,该感知质量模型是基于对视觉质量的人类评估而训练的。通常,感知质量模型将重构视频内容的特征值与不同质量得分相关联。经由感知质量模型来预测重构视频内容的质量的一个限制是感知质量模型通常缺乏鲁棒性。例如,如果与感知质量模型相关联的特征均未反映某些类型的失真,则感知质量模型无法准确地预测人类对包括这种失真的重构视频内容的质量的感知。

[0006] 特别地,感知质量模型通常无法可靠地预测包括在公共域的实时视频质量数据库中的重构视频内容的感知视觉质量。对关于实时视频质量数据库的感知质量模型的缺陷的分析表明,与感知质量模型相关联的特征都无法准确地对可归因于许多类型的复杂时间视频失真的感知质量下降进行建模。复杂时间视频失真的示例包括但不限于运动估计不匹配、闪烁和重影。

[0007] 如前所述,本领域需要更有效的技术来预测重构视频内容的感知质量。

发明内容

[0008] 本发明的一个实施例阐述了一种用于预测重构视频内容的感知视频质量的计算机实现的方法。该方法包括:基于第一特征集的第一组值和第一模型来计算重构视频内容的第一质量得分,该第一模型将第一组值与第一质量得分相关联,其中,第一模型是经由第一机器学习算法生成的;基于第二特征集的第二组值和第二模型来计算重构视频内容的第二质量得分,该第二模型将第二组值与第二质量得分相关联;以及基于第一质量得分和第二质量得分来确定重构视频内容的整体质量得分,其中,整体质量得分指示与流式视频内容相关联的视觉质量的水平。

[0009] 所公开的用于预测感知视频质量的技术的至少一个优点是,整体质量得分准确地

预测了跨多种重构视频内容的感知视频质量。特别地,相对于现有技术的至少一项技术改进是,所公开的技术聚合了经由与不同特征集相关联的不同感知质量模型而获得的质量得分,以确定整体质量得分。因为特征在感知质量模型之间分布,所以所公开的技术降低了增加特征的总数导致归因于过度拟合的不准确的可能性。

附图说明

[0010] 为了可以详细地理解本发明的上述特征的方式,通过参考实施例,可以获得上面简要概述的本发明的更具体的描述,其中一些实施例在附图中示出。然而,应当注意,附图仅示出了本发明的典型实施例,并且因此不应当被视为对本发明范围的限制,因为本发明可以允许其他等效实施例。

[0011] 图1是被配置为实现本发明的一个或多个方面的系统的概念图;

[0012] 图2是根据本发明的各种实施例的图1的时间提取器的更详细图示;

[0013] 图3是根据本发明的各种实施例的用于计算重构视频内容的质量得分的方法步骤的流程图;

[0014] 图4是被配置为实现本发明的一个或多个方面的集成系统的概念图;以及

[0015] 图5是根据本发明的各种实施例的用于基于两个不同特征集来计算重构视频内容的整体质量得分的方法步骤的流程图。

具体实施方式

[0016] 在下面的描述中,阐述了许多具体细节以提供对本发明的更透彻的理解。然而,对于本领域技术人员显而易见的是,可以在没有这些具体细节中的一个或多个的情况下实践本发明。

[0017] 所公开的技术经由将特征值与质量得分相关联的感知质量模型来有效且可靠地预测重构视频内容的质量得分。应当注意,感知质量模型与多个时间特征相关联,其中每个时间特征与不同的缩放比例(scale)相关联。在操作中,时间提取器基于重构视频内容、关联视频内容以及与时间特征相关联的缩放比例来计算特定时间特征的每帧值。对于给定的帧,时间提取器计算一个帧间差分矩阵,该帧间差分矩阵指定了帧与相邻帧之间的亮度差。然后,时间提取器基于缩放比例对帧间差分矩阵进行下采样以生成经缩放的帧间差分矩阵。随后,时间提取器对经缩放的帧间差分矩阵执行时间滤波操作和局部均值减法运算,以计算局部均值相减矩阵。然后,时间提取器对经缩放的帧间差分矩阵执行调节操作以生成与重构视频内容的帧相关联的第一加权熵矩阵。

[0018] 以类似的方式,时间提取器基于视频内容和缩放比例来得出第二加权熵矩阵。时间提取器在第一加权熵矩阵和第二加权熵矩阵之间执行减法运算以计算熵差矩阵。最后,时间提取器对熵差矩阵执行空间池化操作,以生成重构视频内容的帧的时间特征的单个值。因为以多个缩放比例计算熵差矩阵可以模拟人类视觉系统如何处理视觉信息的各个方面,所以感知质量模型可以准确地对归因于许多类型的复杂时间视频失真的感知质量下降进行建模。

[0019] 在一些实施例中,集成预测引擎基于两个感知质量模型来计算用于重构视频内容的集成质量得分。感知质量模型中的每一个都将不同的特征集的值与不同质量得分相关

联。集成预测引擎将通过两个感知质量模型而获得的质量得分进行合成,以获得总质量得分。有利地,在两个预测质量模型之间分配特征降低了增加特征的总数导致归因于过度拟合的不准确的可能性。

[0020] 系统概述

[0021] 图1是被配置为实现本发明的一个或多个方面的系统100的概念图。如图所示,系统100包括但不限于任何数量的计算实例110、训练数据库120、客观数据库150、主观数据库182和目标特征数据库170。出于说明的目的,相似对象的多个实例是用标识对象的附图标记和在需要的情况下标识实例的括号内的数字表示的。在各种实施例中,系统100的任何数量的组件可以分布在多个地理位置上或以任何组合被包括在一个或多个云计算环境(即,经封装的共享资源、软件、数据等)中。

[0022] 如图所示,计算实例110中的每一个包括但不限于处理器112和存储器116。处理器112可以是能够执行指令的任何指令执行系统、装置或设备。例如,处理器112可以包括中央处理单元(CPU)、图形处理单元(GPU)、控制器、微控制器、状态机或前述的任何组合。存储器116存储诸如软件应用和数据之类的内容,以供计算实例110的处理器112使用。

[0023] 存储器116可以是随时可用的存储器中的一个或多个,例如随机存取存储器(RAM)、只读存储器(ROM)、软盘、硬盘或本地或远程的任何其他形式的数字存储装置。在一些实施例中,存储装置(未示出)可以补充或替换存储器116。存储装置可以包括处理器112可访问的任何数量和类型的外部存储器。例如,但不限于,存储装置可以包括安全数字卡、外部闪存、便携式光盘只读存储器(CD-ROM)、光学存储设备、磁性存储设备或前述的任何合适组合。

[0024] 通常,计算实例110被配置为实现一个或多个应用。仅出于说明的目的,每个应用被描绘为驻留在单个计算实例110的存储器116中并且在单个计算实例110的处理器112上执行。然而,如本领域技术人员将认识到的,每个应用的功能可以在驻留在任何数量的计算实例110的存储器116中并且在任何组合的任何数量的计算实例110的处理器112上执行的其他数量的其他应用上分布。此外,可以将任何数量的应用的功能合并到单个应用或子系统中。

[0025] 特别地,计算实例110被配置为提供用于重构视频内容的自动化视频内容质量评估。视频内容质量评估可以用于实现与编码和流式基础结构相关联的质量控制。例如,为了确保可接受的观看体验,通常在设计和传递过程中的各个点处基于视觉质量得分来评估重构视频内容的质量以及间接地评估包括在编码和流式基础结构中的元素的口径。然后,视觉质量得分用于指导应用(例如,编码器)的开发和内容传递的实时优化,例如质量感知的流切换算法。

[0026] 如图所示,训练特征引擎130、训练引擎180、预测引擎192、特征提取器140的实例以及时间池化引擎(temporal pooling engine)194驻留在存储器116中并且在不同计算实例110的处理器112上执行。在替代实施例中,训练特征引擎130、训练引擎180、特征提取器140、预测引擎192、时间池化引擎194、训练特征引擎130和训练引擎180的任何数量的实例可以驻留在任何数量和类型的存储器中,并且在任何组合的任何数量和类型的处理设备上行。

[0027] 在用虚线描绘的学习阶段中,训练特征引擎130和训练引擎180协作以生成感知质

量模型190。对于重构视频内容,感知质量模型190将特征集的一组值与质量得分184相关联,该质量得分184与观看重构视频内容时的主观人类视觉体验相关。

[0028] 在操作中,训练特征引擎130获取训练数据库120并生成客观数据库150。训练数据库120包括但不限于任何数量的源122和任何数量的重构源124。源122中的每一个包括但不限于任何数量和类型的视频内容。视频内容的示例包括但不限于例如特征长度电影、电视节目集和音乐视频的任何部分(包括全部)。重构源124中的每一个包括但不限于从包括在源122之一中的视频内容中导出的重构视频内容。

[0029] 一般情况下,重构源124中的每一个复制包括在关联源122中的视频内容,该视频内容通过流式基础结构和显示设备传送给观看者。可以基于任何数量的编码参数并且以任何技术上可行的方式从源122导出重构源124。例如,在一些实施例中,从源122中的每一个导出三个不同的重构源124。对于三个不同的编码参数集中的每一个,编码引擎(图1中未示出)独立地对源122(x)进行编码。对于三个所得经编码的源中的每一个,解码引擎(图1中未示出)独立地对经编码的源进行解码,以生成从源122(x)导出的重构源124中的一个。

[0030] 如图所示,训练特征引擎130包括但不限于特征提取器140。特征提取器140计算特征值。特征中的每一个都是测量空间视频失真的“空间特征”,或者是测量时间视觉失真的“时间特征”。空间视频失真的示例包括但不限于压缩伪像、缩放伪像和胶片颗粒噪声(film grain noise)。时间视频失真的示例包括但不限于运动估计不匹配、闪烁和重影。

[0031] 特征提取器140包括但不限于空间提取器148和时间提取器146。对于每个重构源124(x)和每个空间特征,空间提取器148基于重构源124(x)和可选地与重构源124(x)相关联的源122来计算空间阵列172(x)。空间阵列172(x)包括但不限于包括在重构源124(x)中的每个帧的不同帧值。

[0032] 类似地,对于每个重构源124(x)和每个时间特征,时间提取器146基于重构源124(x)和与重构源124(x)相关联的源122来计算时间阵列174(x)。时间阵列174(x)包括但不限于包括在重构源124(x)中的每个帧的不同帧对值。如本文所提及的,“帧对”包括两个连续的帧,并且帧对值与两个连续的帧中的第一个相关联。在替代实施例中,特征提取器140可以以任何技术上可行的方式来计算特定重构源124的特定特征的任何数量的值。

[0033] 对于每个空间特征和每个重构源124(x),训练特征引擎130对与空间特征相关联的空间阵列172(x)执行时间池化操作,以确定重构源124(x)的特征的空间值162。训练特征引擎130可以以任何技术上可行的方式来执行任何数量和类型的时间池化操作。例如,在一些实施例中,训练特征引擎130可以将重构源124(x)的给定空间特征的空间值162设置为等于包括在与特征相关联的空间阵列172(x)中的帧值的平均值。类似地,对于每个时间特征和每个重构源124(x),训练特征引擎130可以执行将重构源124(x)的给定时间特征的时间值164设置为等于包括在与特征相关联的时间阵列174(x)中的帧对值的平均值。

[0034] 对于重构源124中的每一个,训练特征引擎130生成训练特征数据库152,该训练特征数据库152包括但不限于用于重构源124的特征的空间值162和时间值164。训练特征引擎130然后生成客观数据库150,该客观数据库150包括但不限于训练特征数据库152。因此,如果训练数据库120包括N个重构源124,则客观数据库150包括N个训练特征数据库152。此外,如果特征提取器140计算S个空间特征的S个空间阵列172和T个时间特征的T个时间阵列174,则训练特征数据库152中的每一个均包括S个空间值162和T个时间值164。

[0035] 训练引擎180基于客观数据库150和主观数据库182来生成感知质量模型190。主观数据库182包括但不限于重构源142中的每一个的质量得分184。基于人类对重构源142(x)的观看来生成重构源142(x)的质量得分184(x)。更具体地,当观看者观看显示设备上的重构源142(x)时,观看者亲自对视觉质量进行评级-分配个性化得分。然后将个性化得分聚合以生成质量得分184(x)。

[0036] 训练引擎180以任何技术上可行的方式实施任何数量和类型的机器学习算法,以基于客观数据库150和主观数据库182来生成感知质量模型190。机器学习算法的示例包括但不限于例如支持向量机算法、人工神经网络算法和随机森林算法。通常,对于重构源124(x)中的每一个,基于包括在客观数据库150(x)中的空间值162和时间值164与包括在主观数据库182中的质量得分184(x)之间的关联来训练感知质量模型190。在替代实施例中,训练引擎180可以以任何技术上可行的粒度级别来训练感知质量模型190(例如,基于将特征的帧值与帧质量得分相关联)。

[0037] 应当注意,不管训练引擎180训练感知质量模型190的粒度如何,感知质量模型190都将任何给定粒度的特征集的一组值与相同粒度的得分相关联。更精确地,感知质量模型190将用于S个空间特征的集合的一组S个值和用于T个时间特征的集合的一组T个值与预测相关联的感知视觉质量的得分相关联。例如,感知质量模型190将用于第一帧的空间特征的一组S个帧值和用于第一帧的时间特征的一组T个帧对值与表示第一帧的感知视觉质量的帧得分相关联。

[0038] 在用实线描绘的预测阶段期间,特征提取器140、预测引擎192和时间池化引擎194协作以基于感知质量模型190、重构目标144和目标142来预测重构目标144的质量得分184。特征提取器140、预测引擎192和时间池化引擎194在本文中统称为“预测应用”。重构目标144包括但不限于从包括在目标142中的视频内容导出的重构视频内容。目标142包括但不限于任何数量和类型的视频内容。

[0039] 重构目标144复制包括在关联源122中的视频内容,该视频内容经由流式基础结构和显示设备传送给观看者。可以以任何技术上可行的方式来生成重构目标144。例如,在一些实施例中,编码器基于恒定速率编码策略对目标142进行编码以生成经编码的目标。随后,解码器对经编码的目标进行解码以生成重构目标144。

[0040] 对于与感知质量模型190相关联的每个空间特征,空间提取器148基于重构目标144和可选地目标142来计算空间阵列172。对于与感知质量模型190相关联的每个时间特征,时间提取器146基于重构目标144和目标142来计算时间阵列174。以这种方式,对于包括在重构目标144中的每个帧,特征提取器140计算用于每个空间特征的不同的帧值以及用于每个时间特征的不同的帧对值。

[0041] 对于包括在重构目标144中的每个帧,预测引擎192基于与帧相关联的帧值、与帧相关联的帧对值以及感知质量模型190来生成帧得分。例如,对于包括在重构目标144中的第一帧,预测引擎192基于空间阵列172的第一条目和时间阵列174的第一条目来确定特征的第一组值。预测引擎192然后基于特征的第一组值和感知质量模型190来确定第一帧的帧得分。在替代实施例中,预测引擎192可以基于相同粒度级别的特征的值和感知质量模型190来计算任何粒度级别的感知视觉质量的得分。

[0042] 随后,时间池化引擎194基于帧得分的算术平均值来确定重构目标144的质量得分

184。更具体地,时间池化引擎194将所有帧得分求和,并且然后将总和除以帧的总数以计算质量得分184。因此,质量得分184表示跨包括在重构目标144中的帧的平均视觉质量。最后,时间池化引擎194发送或显示重构目标144的质量得分184,以评估流式视频内容。

[0043] 在替代实施例中,时间池化引擎194可以以任何技术上可行的方式基于帧得分来确定重构目标144的质量得分184。例如,在一些实施例中,时间池化引擎194可以响应于视频质量的变化来执行模仿人类观点得分的相对平滑的方差的滞后池化操作。更精确地,时间池化引擎194可以对帧得分执行线性低通操作和非线性(等级排列(rank-order))加权操作两者,以生成重构目标144的质量得分184。

[0044] 在一些替代实施例中,时间池化引擎194确定重构目标144的每个块的不同质量得分184,而不是重构目标144的单个质量得分184。如本文所提及的,“块”通常对应于场景,并且场景是时空特性近似恒定的帧序列。为了计算重构目标144的特定块的质量得分184,时间池化引擎194计算包括在块中的帧的子集的帧得分的算术平均值。

[0045] 通过常规感知质量模型来预测重构视频内容的质量的一个限制是常规感知质量模型通常缺乏鲁棒性。例如,如果与常规感知质量模型相关联的特征均未反映出某些类型的失真,则常规感知质量模型无法准确地预测人类对包括这种失真的重构视频内容的质量的感知。

[0046] 常规感知质量模型通常与多个空间特征和单个时间特征相关联。空间特征可以包括但不限于例如任何数量的附加损伤测量(ADM)特征、盲或无参考图像空间质量评估器特征以及视觉信息保真度(VIF)特征。时间特征是帧间差分特征。对于给定的帧对,常规时间提取器计算两个帧之间的每个像素的亮度差。然后,时间提取器将帧间差分特征的帧对值设置为等于每个像素的亮度差的平均值。

[0047] 然而,如通过使用公共域的实时视频质量数据库进行基准测试所证明的,帧间差分特征无法准确地测量归因于许多类型的复杂时间视频失真的感知质量下降。复杂时间视频失真的示例包括但不限于运动估计不匹配、闪烁和重影。因为帧间差分特征是复杂时间视频失真的不可靠度量,所以常规感知质量模型无法可靠地预测重构视频内容的感知视觉质量。

[0048] 实现准确的时间特征

[0049] 为了解决以上问题,系统100建立了多个时间特征,其模仿了人类视觉系统如何处理视觉信息的各个方面。例如,如结合图2所描述的,因为人类视觉系统以多个缩放比例处理视觉信息,所以时间提取器146计算多个时间特征的帧对值,其中每个时间特征与不同的缩放比例相关联。在替代实施例中,时间提取器146可以以确保表示多个缩放比例的任何技术上可行的方式来计算任何数量的时间特征的帧对值。通常,时间提取器146计算在感知上相关的时间特征的帧对值,并且一起精确地测量归因于复杂时间视频失真的感知质量下降。

[0050] 应当理解,本文所示的系统100是说明性的,并且可以进行变化和修改。例如,由本文所述的训练特征引擎130、训练引擎180、特征提取器140、空间提取器148、时间提取器146、预测引擎192和时间池化引擎194提供的功能可以被集成到任何数量的软件应用(包括一个)和任何数量的系统100的组件中或者分布在其中。此外,可以根据需要来修改图1中各个单元之间的连接拓扑。

[0051] 图2是根据本发明的各种实施例的图1的时间提取器146的更详细图示。如图所示,时间提取器146包括但不限于帧比较引擎220、下采样引擎240、空间滤波引擎250、熵引擎270、熵比较引擎280和空间池化引擎290。

[0052] 通常,时间提取器146基于重构视频内容和从其导出重构视频内容的视频内容,来生成重构视频内容的时间特征的值。应当注意,时间特征中的每一个与不同的缩放比例250相关联。重构视频内容可以是重构源124或重构目标144中的任何一个。仅出于说明的目的,在图2中,在生成重构目标144的第一时间特征的时间阵列174(1)的上下文中描述时间提取器146。

[0053] 如图所示,重构目标144包括但不限于帧210(1)-210(F),其中F是大于1的任何整数。以互补的方式,目标142包括但不限于帧210'(1)-210'(F)。仅出于说明的目的,与目标142相关联的对象与可比较对象相区别,该可比较对象与具有修饰附图标标记的撇号(即,')的重构目标144相关联。如结合图1所描述的,重构目标144和目标142中的每一个与F个帧210和F个帧对相关联。

[0054] 在操作中,帧比较引擎220(1)基于重构目标144来计算帧间差分矩阵230(1)-230(F)。每个帧间差分矩阵230与不同的帧对相关联。通常,为了计算帧间差分矩阵230(x),帧比较引擎220(1)在帧210(x)和210(x+1)之间执行每个像素的比较操作以确定每个像素的亮度差。因此,包括在帧间差分矩阵230(x)中的每个像素位置指定了不同的每个像素的亮度差异。以类似的方式,帧比较引擎220(2)基于目标142来独立地计算帧间差分矩阵230'(1)-230'(F)。

[0055] 下采样引擎240(1)基于与第一时间特征相关联的缩放比例250(1)来独立地对帧间差分矩阵230(1)-230(F)中的每一个执行下采样操作,以分别生成经缩放的帧间差分(FD)矩阵242(1)-242(F)。如图所示,缩放比例250(1)等于2。在替代实施例中,第一时间特征可以以任何技术上可行的方式与任何缩放比例250相关联。下采样引擎240(1)可以基于缩放比例250(1)来执行任何数量和类型的下采样操作。通常,如果帧间差分矩阵230中的每一个包括 $N \times M$ 个像素并且缩放比例240(1)是Q,则经缩放的帧间差分矩阵242中的每一个包括 $N/Q \times M/Q$ 个像素。以类似的方式,下采样引擎240(2)基于缩放比例240(1)来独立地对帧间差分矩阵230'(1)-230'(F)中的每一个执行下采样操作,以分别生成经缩放的帧间差分矩阵242'(1)-242'(F)。

[0056] 空间滤波引擎250(1)独立地处理经缩放的帧间差分矩阵242(1)-242(F)中的每一个,以分别生成局部均值相减矩阵262(1)-262(F)。为了处理经缩放的帧间差分矩阵242(x),空间滤波引擎250(1)执行两步局部空间滤波处理,其近似经缩放的帧间差分矩阵242(x)的带通滤波。首先,空间滤波引擎250(1)利用空间各向同性高斯滤波器对经缩放的帧间差分矩阵242(x)进行滤波,以生成局部平均的经缩放的帧间差分矩阵(未示出)。然后,空间滤波引擎250(1)从经缩放的帧间差分矩阵242(x)中减去局部平均的经缩放的帧间差分矩阵,以生成局部均值相减矩阵262(x)。以类似的方式,空间滤波引擎250(2)独立地处理经缩放的帧间差分矩阵242'(1)-242'(F)中的每一个,以分别生成局部均值相减矩阵262'(1)-262'(F)。

[0057] 熵引擎270(1)独立地对局部均值相减矩阵262(1)-262(F)中的每一个执行调节操作,以分别生成加权熵矩阵272(1)-272(F)。应当注意,调节操作与已知在初级视觉皮层(区

域V1)中发生的除法归一化过程有关。为了处理局部均值相减矩阵262(x),空间滤波引擎250(1)将包括在局部均值相减矩阵262(x)中的样本分配给重叠块,其中每个块包括相邻样本的bxb集。然后,空间滤波引擎250(1)基于重叠块来估计bxb协方差矩阵。空间滤波引擎250(1)还将包括在局部均值相减矩阵262(x)中的样本划分为非重叠bxb块。对于每个非重叠块,空间滤波引擎250(1)计算非重叠块的方差。

[0058] 随后,对于每个非重叠块,空间滤波引擎250(1)基于协方差矩阵来计算非重叠块的熵值。最后,对于每个非重叠块,空间滤波引擎250(1)通过对数函数用非重叠块的方差对非重叠块的熵值进行加权,以计算加权熵值。因此,加权熵矩阵272(x)包括但不限于每个非重叠块的加权熵值。以类似的方式,空间滤波引擎250(2)独立地处理局部均值相减矩阵262'(1)-262'(F)中的每一个,以分别生成加权熵矩阵272'(1)-272'(F)。

[0059] 熵差引擎280基于与重构目标144相关联的加权熵矩阵272(1)-272(F)和与目标142相关联的加权熵矩阵272'(1)-272'(F)来计算熵差矩阵282(1)-282(F)。更具体地,为了计算熵差矩阵282(x),熵差引擎280计算包括在加权熵矩阵272(x)中的加权熵值中的每一个与包括在加权熵矩阵272'(x)中的相应加权熵值之间的差。因此,熵差矩阵282(x)包括但不限于每个非重叠块的“熵差值”。

[0060] 空间池化引擎290基于熵差异矩阵282来执行空间池化操作以生成时间阵列174(1)。如图所示,时间阵列174(1)包括但不限于帧对值264(1)-264(x)。包括在时间阵列174(1)中的帧对值264中的每一个指定包括在重构目标144中的帧对中的不同的一个的第一时间特征的值。为了计算帧对值264(x),空间池化引擎290对包括在熵差矩阵282(x)中的所有非重叠块上的熵差值求平均值。

[0061] 在替代实施例中,空间池化引擎290可以以任何技术上可行的方式来聚合包括在熵差矩阵282(x)中的熵差值以生成帧对值264(x)。有利地,重构视频内容与从其导出重构视频内容的视频内容之间的熵差与重构视频内容的质量的主观人类评估相关。

[0062] 应当注意,第二时间特征与缩放比例250(2)为3相关联,并且第三时间特征与缩放比例240(3)为4相关联。当计算时间阵列174(2)时,下采样引擎240(1)和240(2)基于缩放比例240(2)为3来执行下采样操作。当计算时间阵列174(3)时,下采样引擎240(1)和240(2)基于4缩放比例240(3)为4来执行下采样操作。帧比较引擎220、空间滤波引擎250、熵引擎270、熵比较引擎280和空间池化引擎290的操作不基于缩放比例250而改变。有利地,以多个缩放比例来计算时间特征反映了人类视觉系统中固有的多缩放比例处理。

[0063] 应当注意,本文描述的技术是说明性的而非限制性的,并且在不脱离本发明的更广泛的精神和范围的情况下可以对其进行改变。在不脱离所描述的实施例的范围和精神的情况下,对由时间提取器146、帧比较引擎220、下采样引擎240、空间滤波引擎250、熵引擎270、熵比较引擎280和空间池化引擎290提供的功能的许多修改和变化,对于本领域普通技术人员而言将是显而易见的。

[0064] 例如,在各个实施例中,时间提取器146可以被配置为以任何数量的缩放比例250基于任意矩阵来计算任何数量的时间特征的值。例如,时间提取器146可以基于帧间差分矩阵230和230'来计算第一帧间差分特征的帧对值。时间提取器146可以基于与2的缩放比例250相关联的帧间差分矩阵242和242'来计算第二帧间差分特征的帧对值。时间提取器可以基于与2的缩放比例250相关联的熵差矩阵282来计算第一熵特征的帧对值。最后,时间提取

器可以基于与3的缩放比例250相关联的熵差矩阵282来计算第二熵特征的帧对值。

[0065] 图3是根据本发明的各种实施例的用于计算重构视频内容的质量得分的方法步骤的流程图。尽管参考图1-图2的系统描述了方法步骤,但是本领域技术人员应当理解的是,被配置为以任何顺序实现方法步骤的任何系统都落在本发明的范围内。

[0066] 如图所示,方法300开始于步骤302,在步骤302中,帧比较引擎220基于重构目标144来生成帧间差分矩阵230,并且基于目标142来生成帧间差分矩阵230'。在步骤304,时间提取器146选择第一时间特征和与第一时间特征相关联的缩放比例250(1)。在步骤306,下采样引擎240基于选定缩放比例250对帧间差分矩阵230和帧间差分矩阵230'进行下采样,以分别生成经缩放的帧间差分矩阵242和经缩放的帧间差分矩阵242'。

[0067] 在步骤308,对于经缩放的帧间差分矩阵242中的每一个,空间滤波引擎250独立地对帧间差分矩阵242执行局部空间滤波操作,以生成相应局部均值相减矩阵262。以类似的方式,对于经缩放的帧间差分矩阵242'中的每一个,空间滤波引擎250独立地对经缩放的帧间差分矩阵242'执行局部空间滤波操作,以生成相应局部均值相减矩阵262'。

[0068] 在步骤310,对于局部均值相减矩阵262中的每一个,熵引擎270对局部均值相减矩阵262执行调节操作,以计算加权熵矩阵272。以类似的方式,对于局部均值相减矩阵262'中的每一个,熵引擎270对局部均值相减矩阵262'执行调节操作,以计算加权熵矩阵272'。

[0069] 在步骤312,熵差引擎280将与重构目标144相关联的加权熵矩阵272和与目标142相关联的加权熵矩阵272'进行比较,以计算熵差矩阵282。更具体地,为了计算熵差矩阵282(x),熵差引擎280计算包括在加权熵矩阵272(x)中的加权熵值中的每一个与包括在加权熵矩阵272'(x)中的相应加权熵值之间的差。在步骤314,对于熵差矩阵282中的每一个,空间池化引擎290执行空间池化操作,以计算包括在与所选时间特征相关联的时间阵列174中的帧对值264(x)。

[0070] 在步骤316,时间提取器146确定选定时间特征是否是最后的时间特征。如果在步骤316,时间提取器146确定选定时间特征不是最后的时间特征,则方法300前进到步骤318。在步骤318,时间提取器146选择下一个特征和相关联的缩放比例250,并且方法300返回到步骤306,在步骤306中,下采样引擎240基于选定缩放比例250对帧间差分矩阵230和帧间差分矩阵230'进行下采样。

[0071] 然而,如果在步骤316,时间提取器146确定选定时间特征是最后的时间特征,则方法300直接前进到步骤320。在步骤320,对于每个空间特征,空间提取器146基于重构目标144和可选地目标142来计算帧值。在步骤322,对于帧210中的每一个,预测引擎192基于与帧相关联的帧值、与帧相关联的帧对值以及感知质量模型190来生成帧得分。

[0072] 在步骤324,时间池化引擎194对帧得分执行时间池化操作,以确定重构目标144的质量得分184。时间池化引擎194可以执行任何数量和类型的时间池化操作,以聚合帧得分。例如,在一些实施例中,时间池化引擎194基于帧得分的算术平均值来确定重构目标144的质量得分184。然后,时间池化引擎194发送或显示重构目标144的质量得分184,以评估流式视频内容。然后,方法300终止。

[0073] 有利地,作为评估流式媒体内容的视觉质量的一部分,可以聚合任何数量的质量得分184以生成聚合质量得分。例如,每天监测流式服务的仪表板应用可以基于聚合质量得分来指示流式服务的健康状况。通常,除非流式服务发生改变,否则聚合质量得分将保持稳

定。因此,特定日期和紧接的前一天之间的聚合质量得分的偏差指示流式服务的改变(例如,问题或改进),该改变会影响流式媒体内容的视觉质量。相反,如果对流式服务的改变不旨在影响流式媒体内容的视觉质量,则聚合质量得分可以提供对流式媒体内容的视觉质量不受影响的确认。例如,如果用于智能手机的新编码在节省带宽的同时又保持了视觉质量,则聚合质量得分在保持稳定的同时,时间加权比特率将下降。

[0074] 在另一示例中,A/B测试应用将流式会话的控制组与一个或多个流式会话的处理组进行比较。处理组与控制组的区别仅在于一个关键参数。聚合质量得分指示流式媒体内容的视觉质量是否受到关键参数差异的影响。例如,如果处理组使用新的编码,而控制组使用当前编码,则A/B测试应用可以比较控制组的聚合质量得分和处理组的聚合质量得分。比较结果指示新的编码是否影响处理组的流式媒体内容的视觉质量。

[0075] 基于互补感知质量模型来预测视觉质量

[0076] 如本领域技术人员将认识到的,随着与感知质量模型190相关联的特征的数量的增加,过度拟合感知质量模型190的可能性也增加。当感知质量模型190过度拟合时,感知质量模型190对训练数据(例如,客观数据库150和主观数据库184)中的特质进行建模。因此,过度拟合感知质量模型190通常无法准确地预测未包括在训练数据库120中的重构视频内容的质量得分184。为了减少过度拟合单个感知质量模型190的可能性,本发明的一些实施例实现了多个互补感知质量模型190。

[0077] 图4是被配置为实现本发明的一个或多个方面的集成系统400的概念图。如图所示,集成系统400包括但不限于任何数量的计算实例110。如图所示,集成训练引擎410驻留在存储器112中,并且在计算机实例110中的一个的处理器上执行。集成预测引擎490驻留在存储器112中,并且在另一计算机实例110的处理器116上执行。在替代实施例中,集成训练引擎410和集成预测引擎490可以驻留在任何数量和类型的存储器中,并且在任何组合的任何数量和类型的处理设备执行。集成预测引擎490在本文中也称为“集成预测应用”。

[0078] 集成训练引擎410包括但不限于特征分配器420和训练引擎180的两个实例。特征分配器420接收包括任何数量的特征的整体特征集412,并且生成两个特征集422(1)和422(2)。更准确地说,特征分配器420将包括在整体特征集412中的特征中的每一个分配给特征集422中的至少一个。

[0079] 特征分配器420可以被配置为以任何技术上可行的方式将包括在整体特征集412中的特征中的每一个分配给特征集422中的至少一个。例如,在一些实施例中,特征分配器420基于经由图形用户界面(未示出)接收的用户输入来生成特征集422。在其他实施例中,特征分配器420基于优化任何数量和类型的准则的任何数量的试探法来生成特征集422。

[0080] 仅出于解释的目的,图4中描绘的整体特征集412包括五个空间特征和四个时间特征。五个空间特征包括一个VIF特征和四个ADM特征(与四个不同缩放比例250相关联)。四个时间特征包括一个帧间差分特征和三个熵特征(与三个不同缩放比例250相关联)。特征分配器420将五个空间特征分配给特征集422(1)和特征集422(2),将帧间差分特征分配给特征集422(1),以及将三个熵特征分配给特征集422(2)。

[0081] 随后,集成训练引擎410配置训练引擎180(1)以基于主观数据库182和客观数据库150中指定包括在特征集422(1)中的特征的值的部分来生成感知质量模型190(1)。相反,集成训练引擎410配置训练引擎180(2)以基于主观数据库182和客观数据库150中指定包括在

特征集422中的特征的值的的部分(2)来生成感知质量模型190(2)。

[0082] 在替代实施例中,集成训练引擎410包括实现不同机器学习算法的两个不同训练引擎。集成训练引擎410配置第一训练引擎以基于主观数据库182和客观数据库150中指定包括在特征集422(1)中的特征的值的的部分来生成感知质量模型190(1)。集成训练引擎410配置第二训练引擎以基于主观数据库182和客观数据库150中指定包括在特征集422(2)中的特征的值的的部分来生成感知质量模型190(2)。

[0083] 通常,感知质量模型190(1)将特征集422(1)的一组值与第一质量得分相关联,并且感知质量模型190(2)将特征集422(2)的一组值与第二质量得分相关联。有利地,因为没有感知质量模型190与包括在整体特征集422中的所有特征相关联,所以集成训练引擎410生成过度拟合感知质量模型190的可能性被减轻。

[0084] 集成预测引擎490包括但不限于预测引擎192(1)和192(2)的两个实例、聚合引擎440和时间池化引擎194。集成预测引擎490接收感知质量模型190(1)、感知质量模型190(2)和与重构目标144相关联的目标特征数据库170。在替代实施例中,集成预测引擎490配置特征提取器140以计算目标特征数据库170。

[0085] 集成预测引擎490配置预测引擎192(1)以基于感知质量模型190(1)和目标特征数据库170中与包括在特征集422(1)中的特征相关联的部分来计算帧预测数据库430(1)。对于包括在重构目标144中的每个帧210(x),帧预测数据库430(1)包括但不限于帧得分434(x),该帧得分434(x)表示帧210的预测感知视觉质量。在操作中,对于帧210(x),集成预测引擎490基于目标特征数据库170来确定特征集422(1)的一组值。然后,集成预测引擎490配置预测引擎192(1),以基于特征集422(1)的该组值和感知质量模型190(1)来计算包括在帧预测数据库430(1)中的帧得分434(x)。

[0086] 以互补的方式,集成预测引擎490配置预测引擎192(2)以基于感知质量模型190(2)和目标特征数据库170中与包括在特征集422(2)中的特征相关联的部分来计算帧预测数据库430(2)。在操作中,对于帧210(x),集成预测引擎490基于目标特征数据库170来确定特征集422(2)的一组值。然后,集成预测引擎490配置预测引擎192(2)以基于特征集422(2)的该组值和感知质量模型190(2)来计算包括在帧预测数据库430(2)中的帧得分434(x)。

[0087] 在接收到帧预测数据库430(1)和帧预测数据库430(2)时,聚合引擎440生成聚合预测数据库450,该聚合预测数据库450包括帧210中的每一个的不同聚合帧得分454。更准确地说,聚合引擎440对包括在帧预测数据库430(1)中的帧得分434(x)和包括在帧预测数据库430(2)中的帧得分434(x)执行任何数量的聚合操作,以生成帧210(x)的聚合帧得分454(x)。

[0088] 聚合引擎440可以以任何技术上可行的方式来计算聚合帧得分454。例如,在一些实施例中,聚合引擎440计算包括在帧预测数据库430(1)中的帧得分434(x)和包括在帧预测数据库430(2)中的帧得分434(x)的平均值,以生成帧210(x)的聚合帧得分454(x)。在其他实施例中,聚合引擎440计算包括在帧预测数据库430(1)中的帧得分434(x)和包括在帧预测数据库430(2)中的帧得分434(x)的加权平均值,以生成帧210(x)的聚合帧得分454(x)。

[0089] 在其他实施例中,聚合引擎440可以执行任何数量的机器学习算法,以生成聚合模型,该聚合模型将与特征集422(1)相关联的帧得分434和与特征集422(2)相关联的相应帧

得分434与聚合帧质量得分454相关联。聚合引擎440可以执行的机器学习算法的示例包括但不限于基于贝叶斯的方法、装袋 (bagging)、提升、贝叶斯参数平均、贝叶斯模型组合、模型桶、堆叠等。聚合模型可以实现或可以不实现训练引擎180实现来生成感知质量模型190的相同的机器学习算法。随后,聚合引擎440可以基于聚合模型、帧预测数据库430 (1) 和帧预测数据库430 (2) 来计算包括聚合帧得分454的聚合预测数据库450。

[0090] 在聚合引擎440生成聚合预测数据库450之后,时间池化引擎194对聚合预测数据库450执行任何数量的时间池化操作,以计算重构目标144的质量得分184。如结合图1所详细描述,时间池化操作的一些示例包括算术平均操作和滞后池化操作。此外,如还结合图1所详细描述,在替代实施例中,时间池化引擎194可以为重构目标144的每个块生成不同的质量得分184。

[0091] 应当注意,质量得分184的可靠性与感知质量模型190 (1) 的可靠性、与感知质量模型190 (2) 的可靠性、以及与感知质量模型190彼此互补的程度相关。因此,特征分配器420可以被配置为基于优化质量得分184的准确度和/或鲁棒性的任何数量的试探法来生成特征集422。

[0092] 例如,在一些实施例中,特征分配器420实现优化包括在特征集422中的每一个中的特征的试探法,以捕获感知视觉质量的时空方面,产生准确且鲁棒的感知质量模型190,并且在特征集422内以及跨特征集422是多样的。

[0093] 应当注意,本文描述的技术是说明性的而不是限制性的,并且可以在不脱离本发明的更宽的精神和范围的情况下进行改变。在不脱离所描述的实施例的范围和精神的情况下,对由以下各项提供的功能的许多修改和变化对于本领域的普通技术人员而言将是显而易见的:集成训练引擎410、特征分配器420、训练引擎180、集成预测引擎490、预测引擎192、聚合引擎440和时间池化引擎194。

[0094] 图5是根据本发明的各种实施例的用于基于两个不同的特征集来计算重构视频内容的整体质量得分的方法步骤的流程图。尽管参考图1、图2和图4的系统描述了方法步骤,但是本领域技术人员应当理解,被配置为以任何顺序实现方法步骤的任何系统均落入本发明的范围内。

[0095] 如图所示,方法500开始于步骤502,在步骤502中对于包括在整体特征集412中的每个特征,特征分配器420将特征分配给特征集422 (1) 和特征集422 (2) 中的至少一个。应当注意,包括在特征组422中的每一个中的至少一个特征不被包括在另一特征集422中。在步骤504,集成训练引擎410配置训练引擎180 (1) 以基于主观数据库182和客观数据库150中与特征集422 (1) 相关联的部分来生成感知质量模型190 (1)。在步骤506,集成训练引擎410配置训练引擎180 (2) 以基于主观数据库182和客观数据库150中与特征集422 (2) 相关联的部分来生成感知质量模型190 (2)。

[0096] 在步骤508,集成预测引擎490接收目标142和重构目标144。在步骤510,对于包括在重构目标144中的每个帧210,集成预测引擎490计算特征集422 (1) 的第一组值和特征集422 (2) 的第二组值。更准确地说,集成预测引擎490配置特征提取器140,以基于重构目标144和目标142来计算目标特征数据库170。对于帧210中的每一个,然后,集成预测引擎490基于目标特征数据库170来识别特征集422 (1) 的第一组值和特征集422 (2) 的第二组值。在可选实施例中,对于帧210中的每一个,集成预测引擎490可以以任何技术上可行的方式来

计算特征集422 (1) 的第一组值和特征集422 (2) 的第二组值。

[0097] 在步骤512,对于每个帧210 (x),集成预测引擎490配置预测引擎192 (1) 以基于特征集422 (1) 的相关联的一组值和感知质量模型190 (1) 来计算第一帧得分434 (x)。在步骤514,对于每个帧210 (x),集成预测引擎490配置预测引擎192 (2) 以基于特征集422 (2) 的相关联的一组值和感知质量模型190 (2) 来计算第二帧得分434 (x)。

[0098] 在步骤514,对于每个帧210 (x),聚合引擎440基于第一帧得分434 (x) 和第二帧得分434 (x) 来计算聚合帧得分454。在步骤516,时间池化引擎194基于帧210的聚合帧得分454来计算重构目标144的质量得分184。时间池化引擎194可以执行任何数量和类型的时间池化操作以对帧得分434进行聚合。例如,在一些实施例中,时间池化引擎194基于包括在重构目标144中的帧210的帧得分434的算术平均值来确定重构目标144的质量得分184。

[0099] 时间池化引擎194然后传送或显示重构目标144的质量得分184,以便评估流式视频内容。然后方法500终止。有利地,如结合图3更详细地描述的,作为评估流式媒体内容的视觉质量的一部分,可以聚集任何数量的质量得分184以生成聚合质量得分。

[0100] 总之,所公开的技术可以用于高效且可靠地预测重构视频内容的质量得分。在训练阶段期间,训练引擎生成将特征的值映射到质量得分的感知质量模型。随后,在预测阶段期间,预测引擎基于感知质量模型和重构视频内容的特征的值来计算重构视频内容的每帧质量得分。时间池化引擎然后将每帧质量得分合成为单个质量得分。应当注意,在训练阶段和预测阶段两者期间,时间提取器计算多个时间特征的每帧值,其中每个时间特征与不同的缩放比例相关联。

[0101] 时间提取器基于重构视频内容、视频内容和与时间特征相关联的缩放比例来计算特定时间特征的每帧值。为了计算特定帧的每帧值,时间提取器计算指定该帧与相邻帧之间的亮度差的帧间差分矩阵。时间提取器随后基于该缩放比例来下采样帧间差分矩阵以生成经缩放的帧间差分矩阵。随后,时间提取器对经缩放的帧间差分矩阵执行时间滤波操作和局部均值相减操作,以计算与重构视频内容的帧相关联的第一加权熵矩阵。以类似的方式,时间提取器基于视频内容和缩放比例来导出与视频内容的相应帧相关联的第二加权熵矩阵。时间提取器执行第一加权熵矩阵和第二加权熵矩阵之间的相减操作以计算熵差矩阵。最后,时间提取器对熵差矩阵执行空间池化操作以生成针对特征的单个帧特定值。

[0102] 在一些实施例中,集成预测引擎基于两个感知质量模型来计算重构视频内容的质量得分。感知质量模型中的每一个将不同的特征集的值映射到不同的质量得分。对于包括在重构视频内容中的每个帧,集成预测引擎基于第一感知质量模型和第一特征集的第一组值来计算第一帧得分。对于包括在重构视频内容中的每个帧,集成预测引擎基于第二感知质量模型和第二特征集的第二组值来计算第二帧得分。对于每个帧,集成预测引擎然后聚合第一帧得分和第二帧得分以生成聚合帧得分。最后,时间池化引擎基于包括在重构视频内容中的帧的聚合帧得分来生成单个质量得分。

[0103] 相对于现有技术的至少一个技术改进是,以多个缩放比例计算熵差矩阵模仿人类视觉系统如何处理视觉信息的各个方面。因此,时间特征准确地对归因于许多类型的复杂时间视频失真的所感知质量下降进行建模。此外,基于与不同的特征集相关联的不同感知质量模型来计算质量得分降低了增加特征的总数导致归因于过度拟合的不精确性的可能性。

[0104] 1. 在一些实施例中, 一种计算机实现的方法, 包括: 基于第一特征集的第一组值和第一模型来计算重构视频内容的第一质量得分, 该第一模型将第一组值与第一质量得分相关联, 其中, 第一模型经由第一机器学习算法生成; 基于第二特征集的第二组值和第二模型来计算重构视频内容的第二质量得分, 该第二模型将第二组值与第二质量得分相关联; 以及基于第一质量得分和第二质量得分来确定重构视频内容的整体质量得分, 其中, 整体质量得分指示与流式视频内容相关联的视觉质量的水平。

[0105] 2. 根据条款1所述的计算机实现的方法, 其中, 第一质量得分和第二质量得分与包括在重构视频内容中的第一帧相关联, 并且确定整体质量得分包括: 基于第一质量得分和第二质量得分来确定与第一帧相关联的第一聚合质量得分; 以及对第一聚合质量得分和第二聚合质量得分执行一个或多个时间池化操作以计算整体质量得分, 该第二聚合质量得分与包括在重构视频内容中的第二帧相关联。

[0106] 3. 根据条款1或2所述的计算机实现的方法, 其中, 确定第一聚合质量得分包括计算第一质量得分和第二质量得分的平均值或加权平均值。

[0107] 4. 根据条款1至3中任一项所述的计算机实现的方法, 其中, 确定第一聚合质量得分包括基于第三模型将第一质量得分和第二质量得分与第一聚合质量得分相关联, 该第三模型经由第一机器学习算法或第二机器学习算法生成。

[0108] 5. 根据条款1至4中任一项所述的计算机实现的方法, 其中, 第一特征集包括未包括在第二特征集中的第一特征, 并且第二特征集包括未包括在第一特征集中的第二特征。

[0109] 6. 根据条款1至5中任一项所述的计算机实现的方法, 其中, 第一特征集包括时间特征和空间特征。

[0110] 7. 根据条款1至6中任一项所述的计算机实现的方法, 其中, 第一特征集包括以下各项中的至少一项: 附加损伤测量特征、盲或无参考图像空间质量评估器特征以及视觉信息保真度特征。

[0111] 8. 根据条款1至7中任一项所述的计算机实现的方法, 还包括: 基于第一机器学习算法和多个主观得分来生成第一模型, 其中, 包括在多个主观得分中的每个主观得分与不同的重构测试视频内容相关联。

[0112] 9. 根据条款1至8中任一项所述的计算机实现的方法, 还包括: 基于多个主观得分来生成第二模型。

[0113] 10. 在一些实施例中, 一种计算机可读存储介质, 该计算机可读存储介质包括指令, 该指令在由处理器执行时, 使得处理器执行以下步骤: 基于第一特征集的第一组值和第一模型来计算重构视频内容的第一质量得分, 该第一模型将第一组值与第一质量得分相关联, 其中, 第一模型经由第一机器学习算法生成; 基于第二特征集的第二组值和第二模型来计算重构视频内容的第二质量得分, 该第二模型将第二组值与第二质量得分相关联, 其中, 该第二模型经由第一机器学习算法或第二机器学习算法生成; 以及基于第一质量得分和第二质量得分来确定重构视频内容的整体质量得分, 其中, 整体质量得分指示与流式视频内容相关联的视觉质量的水平。

[0114] 11. 根据条款10所述的计算机可读存储介质, 其中, 第一质量得分和第二质量得分与包括在重构视频内容中的第一帧相关联, 并且确定整体质量得分包括: 基于第一质量得分和第二质量得分来确定与第一帧相关联的第一聚合质量得分; 以及对第一聚合质量得分

和第二聚合质量得分执行一个或多个时间池化操作以计算整体质量得分,该第二聚合质量得分与包括在重构视频内容中的第二帧相关联。

[0115] 12.根据条款10或11所述的计算机可读存储介质,其中,确定第一聚合质量得分包括计算第一质量得分和第二质量得分的平均值或加权平均值。

[0116] 13.根据条款10至12中任一项所述的计算机可读存储介质,其中,执行一个或多个时间池化操作包括对第一聚合质量得分和第二聚合质量得分执行以下各项中的至少一项:线性低通操作、非线性等级排列加权操作和算术平均操作。

[0117] 14.根据条款10至13中任一项所述的计算机可读存储介质,其中,第一特征集包括未包括在第二特征集中的第一特征,并且第二特征集包括未包括在第一特征集中的第二特征。

[0118] 15.根据条款10至14中任一项所述的计算机可读存储介质,其中,第一特征集包括时间特征和空间特征。

[0119] 16.根据条款10至15中任一项所述的计算机可读存储介质,其中,第一特征集包括多个时间特征,并且包括在多个时间特征中的每个时间特征与不同的缩放比例相关联。

[0120] 17.根据条款10至16中任一项所述的计算机可读存储介质,还包括:在计算第一质量得分之前,将包括在整体特征集中的每个特征分配给第一特征集和第二特征集的至少一个。

[0121] 18.根据条款10至17中任一项所述的计算机可读存储介质,还包括:基于第一机器学习算法和多个主观得分来生成第一模型,其中,包括在多个主观得分中的每个主观得分与不同的重构测试视频内容相关联。

[0122] 19.在一些实施例中,一种系统,包括:存储器,该存储器存储指令;以及处理器,该处理器耦合到存储器,并且该处理器在执行指令时,被配置为执行以下操作:基于第一特征集的第一组值和第一模型来计算重构视频内容的第一质量得分,该第一模型将第一组值与第一质量得分相关联,其中,第一模型基于多个主观得分生成;基于第二特征集的第二组值和第二模型来计算重构视频内容的第二质量得分,该第二模型将第二组值与第二质量得分相关联,其中,该第二模型基于多个主观得分生成;以及基于第一质量得分和第二质量得分来确定重构视频内容的整体质量得分,其中,整体质量得分指示与流式视频内容相关联的视觉质量的水平。

[0123] 20.根据条款19所述的系统,其中,第一质量得分和第二质量得分与包括在重构视频内容中的第一帧相关联,并且处理器被配置为通过以下方式来确定整体质量得分:基于第一质量得分和第二质量得分来确定与第一帧相关联的第一聚合质量得分;以及对第一聚合质量得分和第二聚合质量得分执行一个或多个时间池化操作以计算整体质量得分,该第二聚合质量得分与包括在重构视频内容中的第二帧相关联。

[0124] 在任何权利要求中叙述的任何权利要求元素和/或在本申请中描述的任何元素的任何和所有组合,以任何方式,都落入本发明和保护的预期范围内。

[0125] 已经出于说明的目的呈现了对各种实施例的描述,但是不旨在是穷举的或限于所公开的实施例。在不脱离所描述的实施例的范围和精神的情况下,许多修改和变化对于本领域的普通技术人员而言将是显而易见的。

[0126] 本发明实施例的各个方面可以被体现为系统、方法或计算机程序产品。因此,本公

开的各个方面可以采取完全硬件实施例、完全软件实施例(包括固件、驻留软件、微代码等)或组合软件和硬件方面(其在本文通常可以被统称为“模块”或“系统”)的实施例的形式。此外,本公开的各个方面可以采取在一个或多个计算机可读介质中体现的计算机程序产品的形式,该计算机可读介质具有在其上体现的计算机可读程序代码。

[0127] 可以利用一个或多个计算机可读介质的任何组合。计算机可读介质可以是计算机可读信号介质或计算机可读存储介质。计算机可读存储介质可以是,例如但不限于,电、磁、光、电磁、红外或半导体系统、装置或设备,或前述的任何合适的组合。计算机可读存储介质的更具体的示例(非穷举列表)将包括以下各项:具有一个或多个导线的电连接、便携式计算机磁盘、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或闪存)、光纤、便携式光盘只读存储器(CD-ROM)、光存储设备、磁存储设备、或前述的任何合适的组合。在本文档的上下文中,计算机可读存储介质可以是任何有形介质,其可以包含或存储由指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合使用的程序。

[0128] 以上参考根据本公开的实施例的方法、装置(系统)和计算机程序产品的流程图图示和/或框图描述了本公开的各个方面。应当理解,流程图图示和/或框图的每个框以及流程图图示和/或框图中的框的组合可以由计算机程序指令实现。这些计算机程序指令可以被提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器以产生机器。指令在经由计算机或其他可编程数据处理装置的处理器执行时,能够实现流程图和/或框图的一个或多个框中指定的功能/动作。这样的处理器可以是但不限于通用处理器、专用处理器、特定于应用的处理器或现场可编程门阵列。

[0129] 附图中的流程图和框图示出了根据本公开的各种实施例的系统、方法和计算机程序产品的可能实现方式的架构、功能和操作。在这点上,流程图或框图中的每个框可以表示代码的模块、段或部分,其包括用于实现(一个或多个)指定的逻辑功能的一个或多个可执行指令。还应当注意,在一些替代实现方式中,框中所标注的功能可以不按图中标注的顺序发生。例如,连续示出的两个框实际上可以基本上同时执行,或者这些框有时可以以相反的顺序执行,这取决于所涉及的功能。还应当注意,框图和/或流程图图示的每个框以及框图和/或流程图图示中的框的组合可以由执行特定功能或动作的基于专用硬件的系统或专用硬件和计算机指令的组合来实现。

[0130] 虽然前述内容涉及本公开的实施例,但是在不脱离本公开的基本范围的情况下,可以设计本公开的其他和进一步的实施例,并且本公开的范围由所附权利要求确定。

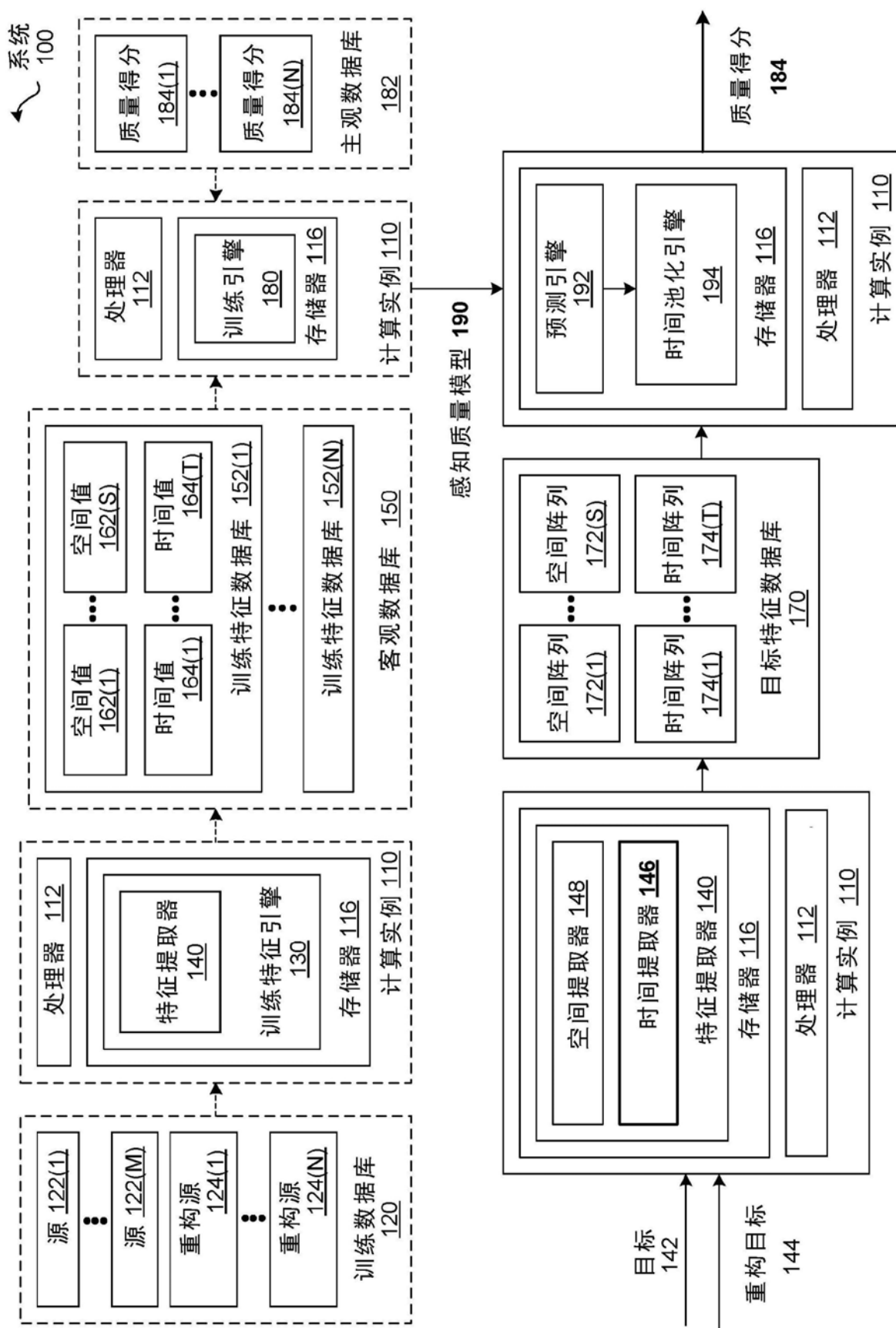


图1

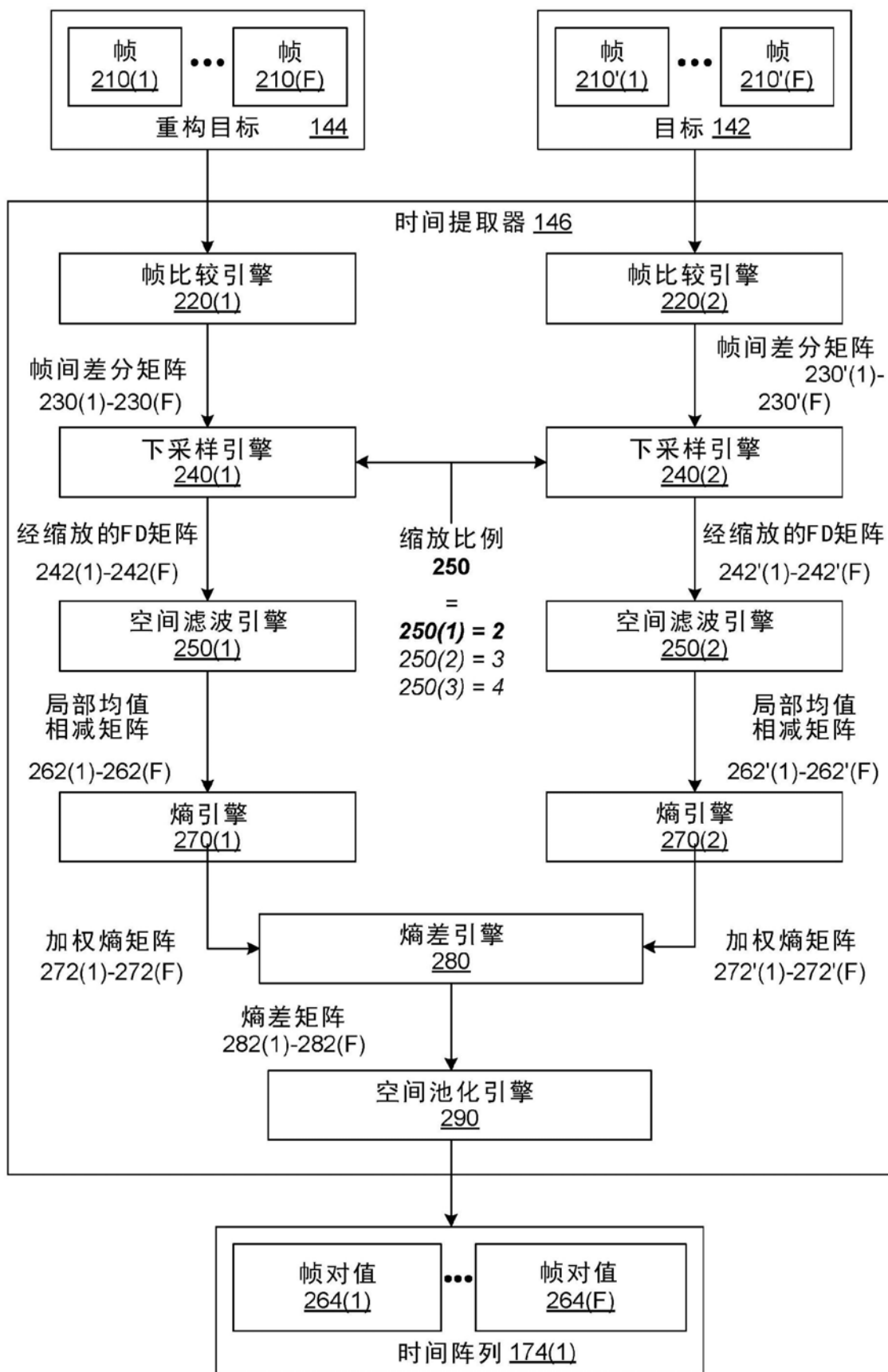


图2

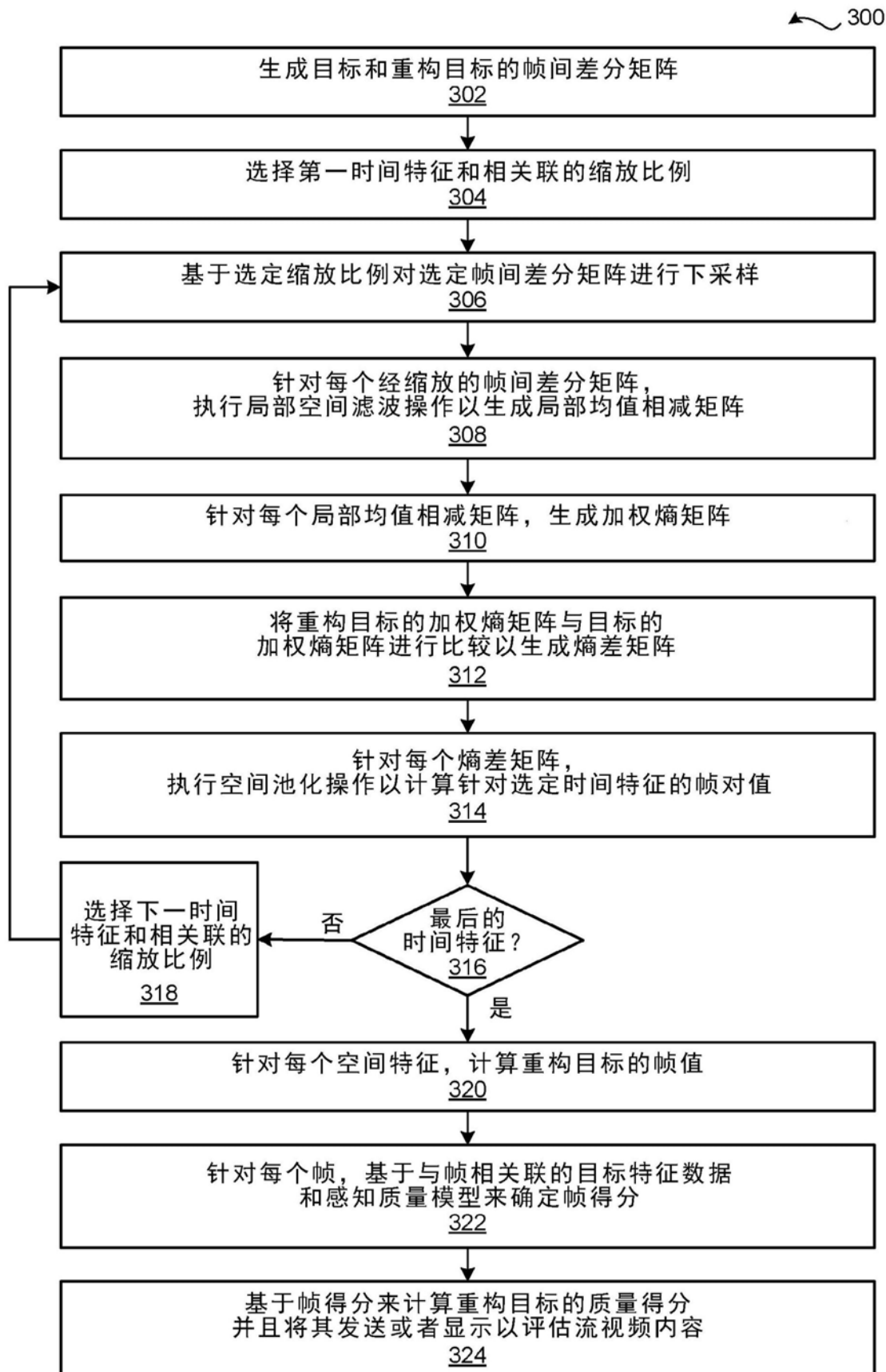


图3

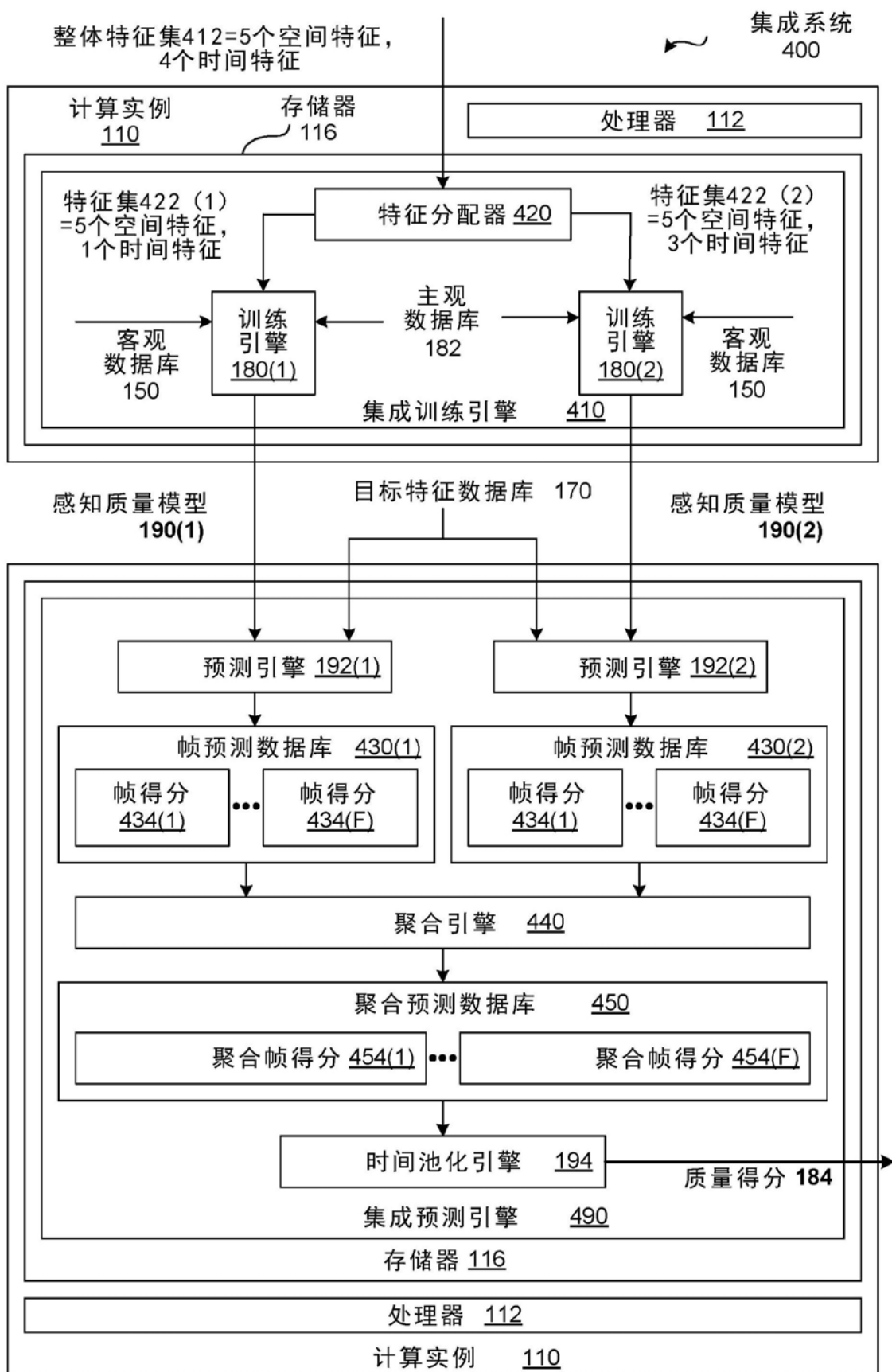


图4

500

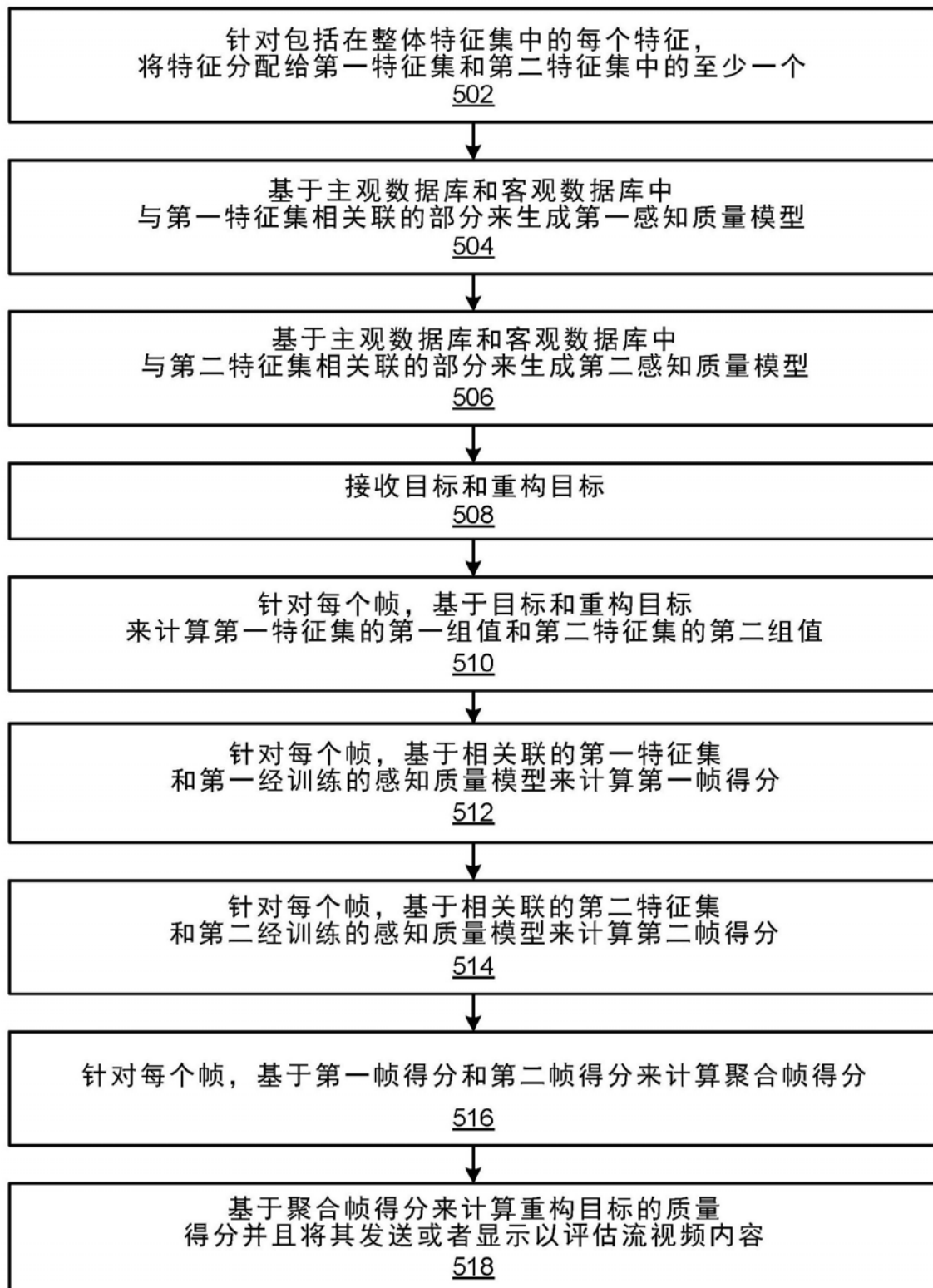


图5