



(12) 发明专利

(10) 授权公告号 CN 103999471 B

(45) 授权公告日 2016. 06. 15

(21) 申请号 201280062237. 1

(51) Int. Cl.

(22) 申请日 2012. 10. 08

H04N 21/2662(2006. 01)

(30) 优先权数据

H04N 21/234(2006. 01)

13/275, 090 2011. 10. 17 US

审查员 刘昶

(85) PCT国际申请进入国家阶段日

2014. 06. 16

(86) PCT国际申请的申请数据

PCT/US2012/059194 2012. 10. 08

(87) PCT国际申请的公布数据

W02013/059015 EN 2013. 04. 25

(73) 专利权人 谷歌公司

地址 美国加利福尼亚州

(72) 发明人 杨蕾 D·穆克赫吉

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 鄧迅 辛鸣

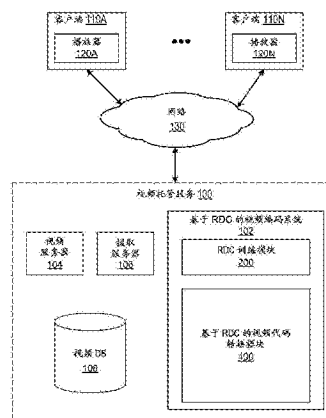
权利要求书3页 说明书15页 附图7页

(54) 发明名称

由视频描述长度引导的视频编码的速率-失真-复杂性优化

(57) 摘要

一种系统和方法提供了具有比特率约束的由视频描述长度 (VDL) 引导的恒定质量视频编码策略和一种用于优化输入视频的编码比特率、失真和复杂性的视频编码系统。该方法获得输入视频的总 VDL、时间 VDL 和空间 VDL 并且比较输入视频的总 VDL、时间 VDL 和空间 VDL 与参考 VDL、时间 VDL 和空间 VDL。基于比较, 该方法调整输入视频的编码比特率、总编码复杂性、时间编码复杂性和空间编码复杂性, 并且用输入视频的调整的编码比特率、总编码复杂性、时间编码复杂性和空间编码复杂性对输入视频编码。



1. 一种用于优化输入视频的编码比特率、失真和复杂性的计算机系统,所述系统包括: 视频描述长度(VDL)模块,被配置为:

获得所述输入视频的总VDL、时间VDL和空间VDL,所述输入视频的所述总VDL指示所述输入视频的总内容复杂性,所述输入视频的所述时间VDL指示所述输入视频的时间复杂性并且所述输入视频的所述空间VDL指示所述输入视频的空间复杂性;并且

比较所述输入视频的所述总VDL、所述时间VDL和所述空间VDL与对应的参考总VDL、参考时间VDL和参考空间VDL;

基于所述比较调整所述输入视频的编码比特率、总编码复杂性、时间编码复杂性和空间编码复杂性;以及

视频代码转换器,与所述视频描述长度模块耦合,被配置为:

用所述输入视频的调整的所述编码比特率、所述总编码复杂性、所述时间编码复杂性和所述空间编码复杂性对所述输入视频编码。

2. 根据权利要求1所述的系统,其中所述输入视频的所述总VDL由用预定恒定速率因子和计算复杂性水平编码的所述输入视频的实际比特率代表。

3. 根据权利要求1所述的系统,其中所述输入视频的所述时间VDL由在用相同预定恒定速率因子和不同计算复杂性水平编码的所述输入视频的实际比特率之间的差代表。

4. 根据权利要求1所述的系统,其中所述输入视频的所述空间VDL由在用不同预定恒定速率因子和相同计算复杂性水平编码的所述输入视频的实际比特率之间的差代表。

5. 根据权利要求1所述的系统,其中所述参考总VDL、时间VDL和空间VDL来自动态复杂性参考表或者静态复杂性参考表的至少一行。

6. 根据权利要求5所述的系统,其中所述动态复杂性参考表维持根据对所述输入视频编码而生成的所述参考总VDL、时间VDL和空间VDL。

7. 根据权利要求5所述的系统,其中所述静态复杂性参考表维持在视频全集中存储的视频的平均总VDL、时间VDL和空间VDL。

8. 根据权利要求1所述的系统,其中所述视频代码转换器还被配置为:

用基于与所述输入视频关联的目标比特率从恒定速率因子(CRF)参考比特率模型获得的CRF对所述输入视频编码;

根据所述编码计算修正因子;

用所述修正因子修改所述CRF参考比特率模型;以及

用从修改的所述CRF参考比特率模型获得的所述恒定速率因子对所述输入视频编码。

9. 根据权利要求8所述的系统,其中所述视频代码转换器还被配置为:

验证用从修改的所述CRF参考比特率模型获得的所述恒定速率因子对所述输入视频编码的实际比特率是否在所述输入视频的所述目标比特率的预定范围内。

10. 根据权利要求8所述的系统,其中所述CRF参考比特率模型被配置为:

对在对多个视频的视频编码的所述恒定速率因子与速率-失真性能之间的关系建模; 并且

从所述建模获得在视频编码的恒定速率因子与平均比特率之间的一个或者多个关系。

11. 根据权利要求10所述的系统,其中在视频编码的所述恒定速率因子与所述速率-失真性能之间的所述关系对在对多个视频编码的恒定速率因子与平均比特率之间的所述关

系建模。

12. 一种用于优化输入视频的编码比特率、失真和复杂性的方法,所述方法包括:

获得所述输入视频的总视频描述长度(VDL)、时间VDL和空间VDL,所述输入视频的所述总VDL指示所述输入视频的总内容复杂性,所述输入视频的所述时间VDL指示所述输入视频的时间复杂性并且所述输入视频的所述空间VDL指示所述输入视频的空间复杂性;并且

比较所述输入视频的所述总VDL、所述时间VDL和所述空间VDL与对应的参考总VDL、参考时间VDL和参考空间VDL;

基于所述比较调整所述输入视频的编码比特率、总编码复杂性、时间编码复杂性和空间编码复杂性;以及

用所述输入视频的调整的所述编码比特率、所述总编码复杂性、所述时间编码复杂性和所述空间编码复杂性对所述输入视频编码。

13. 根据权利要求12所述的方法,其中所述输入视频的所述总VDL由用预定恒定速率因子和计算复杂性水平编码的所述输入视频的实际比特率代表。

14. 根据权利要求12所述的方法,其中所述输入视频的所述时间VDL由在用相同预定恒定速率因子和不同计算复杂性水平编码的所述输入视频的实际比特率之间的差代表。

15. 根据权利要求12所述的方法,其中所述输入视频的所述空间VDL由在用不同预定恒定速率因子和相同计算复杂性水平编码的所述输入视频的实际比特率之间的差代表。

16. 根据权利要求12所述的方法,其中所述参考总VDL、时间VDL和空间VDL来自动态复杂性参考表或者静态复杂性参考表的至少一行。

17. 根据权利要求12所述的方法,其中用所述输入视频的调整的所述总编码复杂性、所述时间编码复杂性和所述空间编码复杂性对所述输入视频编码还包括:

用基于与所述输入视频关联的目标比特率从恒定速率因子(CRF)参考比特率模型获得的CRF对所述输入视频编码;

根据所述编码计算修正因子;

用所述修正因子修改所述CRF参考比特率模型;以及

用从修改的所述CRF参考比特率模型获得的所述恒定速率因子对所述输入视频编码。

18. 根据权利要求17所述的方法,还包括:

验证用从修改的所述CRF参考比特率模型获得的所述恒定速率因子对所述输入视频编码的实际比特率是否在所述输入视频的所述目标比特率的预定范围内。

19. 根据权利要求17所述的方法,还包括:

对在对多个视频的视频编码的所述恒定速率因子与速率-失真性能之间的关系建模;以及

从所述建模获得在视频编码的恒定速率因子与平均比特率之间的一个或者多个关系。

20. 根据权利要求19所述的方法,其中在视频编码的所述恒定速率因子与所述速率-失真性能之间的所述关系对在对多个视频编码的恒定速率因子与平均比特率之间的所述关系建模。

21. 一种用于优化输入视频的编码比特率、失真和复杂性的设备,所述设备包括:

用于获得所述输入视频的总视频描述长度(VDL)、时间VDL和空间VDL的装置,所述输入视频的所述总VDL指示所述输入视频的总内容复杂性,所述输入视频的所述时间VDL指示所

述输入视频的时间复杂性并且所述输入视频的所述空间VDL指示所述输入视频的空间复杂性;并且

用于比较所述输入视频的所述总VDL、所述时间VDL和所述空间VDL与对应的参考总VDL、参考时间VDL和参考空间VDL的装置;

用于基于所述比较调整所述输入视频的编码比特率、总编码复杂性、时间编码复杂性和空间编码复杂性的装置;以及

用于用所述输入视频的调整的所述编码比特率、所述总编码复杂性、所述时间编码复杂性和所述空间编码复杂性对所述输入视频编码的装置。

22. 根据权利要求21所述的设备,其中所述输入视频的所述总VDL由用预定恒定速率因子和计算复杂性水平编码的所述输入视频的实际比特率代表。

23. 根据权利要求21所述的设备,其中所述输入视频的所述时间VDL由在用相同预定恒定速率因子和不同计算复杂性水平编码的所述输入视频的实际比特率之间的差代表。

24. 根据权利要求21所述的设备,其中所述输入视频的所述空间VDL由在用不同预定恒定速率因子和相同计算复杂性水平编码的所述输入视频的实际比特率之间的差代表。

25. 根据权利要求21所述的设备,其中用于用所述输入视频的调整的所述总编码复杂性、所述时间编码复杂性和所述空间编码复杂性对所述输入视频编码的装置还包括:

用于用基于与所述输入视频关联的目标比特率从恒定速率因子(CRF)参考比特率模型获得的CRF对所述输入视频编码的装置;

用于根据所述编码计算修正因子的装置;

用于用所述修正因子修改所述CRF参考比特率模型的装置;以及

用于用从修改的所述CRF参考比特率模型获得的所述恒定速率因子对所述输入视频编码的装置。

26. 根据权利要求25所述的设备,还包括:

用于对在对多个视频的视频编码的所述恒定速率因子与速率-失真性能之间的关系建模的装置;以及

用于从所述建模获得在视频编码的恒定速率因子与平均比特率之间的一个或者多个关系的装置。

27. 根据权利要求26所述的设备,其中在视频编码的所述恒定速率因子与所述速率-失真性能之间的所述关系对在对多个视频编码的恒定速率因子与平均比特率之间的所述关系建模。

由视频描述长度引导的视频编码的速率-失真-复杂性优化

技术领域

[0001] 本发明总体上涉及视频托管系统,并且更特别地涉及一种用于在编码时间和比特率约束之下用优化的视觉质量对视频进行代码转换的视频编码系统。

背景技术

[0002] 视频托管服务(比如YOUTUBE)允许用户发布视频。大多数视频托管服务将原有源视频从它的原生编码格式(例如,MOV)代码转换成一个或者多个输出格式(例如,ADOBE FLASH或者Windows Media Video(WMV))。代码转换包括使用用于原生格式的视频编码解码器将源视频从原生格式解码成未编码的表示,并且然后利用用于输出格式的视频编码解码器对未编码的表示编码。代码转换可以用来减少存储要求,并且也减少用于向客户端供应视频的带宽要求。

[0003] 在为具有数百万个视频的视频托管服务设计视频编码系统时的一个挑战是用可接受的视觉质量并且以合理计算成本对视频进行代码转换和存储。一个特别的问题是高效分配编码比特和计算以实现源视频的优化的速率-失真(R-D)和计算时间。一般而言,给定目标分辨率和帧速率,视频的视觉质量由它的使用速率控制算法而计算出的编码比特率确定。常规视频编码系统使用多种编码策略以获得源视频的优化的速率-失真,这些编码策略包括一次通过和多次通过平均比特率编码(ABR)、恒定比特率编码(CBR)、恒定量化器编码(CQP)和恒定速率因子编码(CRF)。

[0004] 常规编码策略无法向编码的视频提供恒定视觉质量而又满足与视频关联的比特率约束并且未联合地优化比特率、失真和复杂性。例如,ABR编码策略使用伸缩因子以及长期和短期补偿以实现目标比特率并且满足网络带宽约束。但是ABR编码的视觉质量可能在视频场景改变时波动。CBR编码策略被设计用于由具有固定大小的存储缓冲器控制的具有恒定比特率的实时流传输。CBR在以上提到的常规编码策略之中提供最高编码速度但是最低R-D性能。CQP编码策略维持恒定量化器并且使用相同量化参数(QP)来压缩每个帧。CQP可能尤其在它对具有密集场景改变的视频使用大型量化器时引起编码的视频的时间感知波动。CRF编码策略以用恒定速率因子实现恒定视觉质量为目标。CRF用标称量化器对视频编码,但是在场景具有大量动作和运动时增加QP并且反之亦然。CRF编码的缺点在于输出视频文件大小由于视频内容中的可变场景而不可预测。因此,难以选择适当恒定速率因子值以满足网络或者存储系统的所需比特率约束。

发明内容

[0005] 一种方法、系统和计算机程序产品为视频托管服务提供由视频描述长度引导的具有恒定质量和比特率约束的优化的视频代码转换。

[0006] 在一个实施例中,该方法为输入视频获得总复杂性、时间复杂性和空间复杂性的测量并且比较总复杂性、时间和空间复杂性测量与参考复杂性、时间和空间复杂性测量。输入视频的复杂性(总、时间和空间)的测量由输入视频的某个视频描述长度(VDL)代表。VDL

是以某个可视质量对视频编码而需要的比特率。总VDL指示输入视频的总内容复杂性。输入视频的时间VDL反映视频的时间复杂性。输入视频的空间VDL反映视频的空间复杂性。基于比较,该方法调整输入视频的总编码复杂性、时间和空间编码复杂性并且用输入视频的调整的总编码复杂性、时间和空间编码复杂性对输入视频编码。

[0007] 可以通过使用训练的CRF参考比特率模型和修正的CRF比特率模型来进一步增强用输入视频的调整的总编码复杂性、时间和空间编码复杂性对输入视频编码。对大型视频全集训练CRF参考比特率模型以对在CRF编码的CRF值与平均比特率之间的关系建模。修正的CRF比特率模型被适配为用优化对输入视频编码。

[0008] 另一实施例包括一种用于优化输入视频的编码比特率、失真和复杂性的包括视频描述长度模块和代码转换模块的计算机系统。又一实施例包括一种存储用于按照以上描述的方式压缩图像的可执行计算机程序指令的非瞬态计算机可读介质。

[0009] 在说明书中描述的特征和优点并非囊括性的,并且特别地,许多附加特征和优点将鉴于附图、说明书和权利要求而为本领域普通技术人员所清楚。另外,应当注意,在说明书中使用的言语已经主要出于可读性和指导目的而被选择并且可能尚未被选择用于界定或者限制公开的主题内容。

附图说明

[0010] 图1是图示了具有视频编码系统的视频托管服务的系统视图的框图,该视频编码系统具有速率-失真-复杂性(RDC)优化。

[0011] 图2是图示了利用RDC优化在比特率约束之下的由视频描述长度引导的恒定质量视频编码策略的流程图。

[0012] 图3是利用RDC优化在比特率约束之下的输入视频的由视频描述长度引导的恒定质量视频编码的流程图。

[0013] 图4是在比特率约束之下的多次通过恒定质量视频编码策略的编码步骤的流程图。

[0014] 图5是调整用于示例视频的修正的CRF比特率模型的因子、目标比特率、实际比特率的示例数据。

[0015] 图6是用于RDC优化的离线训练编码计算简档(ECP)映射模块的流程图。

[0016] 图7是用于训练CRF参考比特率模型的相对于恒定速率因子(CRF)和空间分辨率的平均比特率的示例数据。

[0017] 附图仅出于例示的目的而描绘本发明的各种实施例,并且本发明不限于这些所示的实施例。本领域技术人员将从以下讨论容易地认识到可以运用这里所示的结构和方法的备选实施例而未脱离这里描述的本发明的原理。

具体实施方式

[0018] I. 系统概述

[0019] 图1是图示了具有基于速率-失真-复杂性(RDC)的视频编码系统102的视频托管服务100的系统视图的框图。多个用户/查看者使用客户端110A-110N以向视频托管服务100发送视频托管请求,比如向视频托管网站上传视频,以及从视频托管服务100接收请求的服

务。视频托管服务100经由网络130与一个或者多个客户端110A-110N通信。视频托管服务100从客户端110A-110N接收视频托管服务请求、通过基于RDC的视频编码系统102对源视频进行代码转换并且向客户端110A-110N返回代码转换的源视频。视频托管服务100可以驻留于云计算网络内。

[0020] 转向图1中所示的单独的实体,每个客户端110由用户用来请求视频托管服务。例如,用户使用客户端110以发送对于上传视频以用于共享或者播放视频的请求。客户端110可以是任何类型的计算机设备(比如个人计算机(例如,台式、笔记本、膝上型)计算机)以及设备(比如移动电话、个人数字助理、具有IP功能的视频播放器)。客户端110通常包括处理器、显示设备(或者向显示设备的输出)、客户端110将用户在执行任务时使用的数据存储在本地存储装置(比如硬驱动器或者闪存设备)以及用于经由网络130耦合到视频托管服务100的网络接口。

[0021] 客户端110还具有用于播放视频流的视频播放器120(例如,来自Adobe Systems, Inc.的Flasher™播放器等)。视频播放器120可以是单独应用或者是向另一应用(比如网络浏览器)的插件。可以在硬件或者硬件和软件的组合中实施播放器120。所有这些实现方式关于本发明在功能上等效。播放器120包括用于选择视频馈送、开始、停止和倒回视频馈送的用户界面控件(和对应的应用编程接口)。播放器120也可以在它的用户界面中包括被配置用于指示哪个视频显示格式(例如,标准清晰度TV或者高清晰度TV)的视频显示格式选择。其它类型的用户界面控件(例如,按钮、键盘控件)也可以用来控制播放器120的回放和视频格式选择功能。

[0022] 网络130支持在客户端110与视频托管服务100之间的通信。在一个实施例中,网络130是因特网,并且使用现在已知或者随后开发的标准化的网间通信技术和协议,这些网间通信技术和协议使客户端110能够与视频托管服务100通信。

[0023] 视频托管服务100包括基于RDC的视频编码系统102、视频服务器104、摄取服务器106和视频数据库108。摄取服务器106接收用户上传的视频并且在视频数据库108中存储视频。视频服务器104响应于用户视频托管服务请求从视频数据库108供应视频。视频数据库108存储用户上传的视频和由基于RDC视频编码系统102代码转换的视频。在一个实施例中,视频数据库108存储用于基于RDC的视频编码系统102的大型视频全集(例如,20,000个由用户上传的视频)以训练速率-失真-复杂性优化模型。

[0024] 基于RDC的视频编码系统102包括RDC训练模块200和基于RDC的视频代码转换视频400。RDC训练模块200从存储于视频数据库108中的视频离线训练恒定速率因子(CRF)比特率模型和编码计算简档(ECP)映射模块。如以下进一步说明的那样,基于RDC的视频代码转换模块400利用RDC优化在比特率约束之下用视频描述长度(VDL)引导的恒定质量视频编码策略对给定的源视频进行代码转换。通过经由CRF比特率模型和ECP映射模块用帕雷托最优集合发现用于每个编码复杂性级别的RDC最优编码策略来实现RDC优化。作为有益结果,将源视频代码转换成具有在与源视频的目标输出视频格式关联的比特率约束之下的优化的视觉质量的输出视频。

[0025] 为了进一步例示视频托管服务100的系统操作,图2是图示了在视频托管服务100内的基于RDC的视频编码系统102的流程图。用户向视频托管服务100发送对于上传源视频150的请求。基于RDC的视频编码系统102按照源视频150的原生格式接收它并且通过基于

RDC的视频代码转换模块400处理源视频150。源视频150在它的原生格式中具有多个视频编码参数和例如包括帧速率、空间分辨率和比特率的视频源信息。

[0026] 源视频150也可以被存储于视频数据库(例如,图1的视频数据库180)中并且变成被存储于视频数据库中的视频全集160的一部分。视频全集160由RDC训练模块200处理。在一个实施例中,RDC训练模块200包括视频编码器210、CRF参考比特率模型220和ECP映射模块230。RDC训练模块200使用CRF编码对视频全集160中的视频编码并且对视频全集160离线训练CRF参考比特率模型220和ECP映射模块230。CRF参考比特率模型220对在CRF编码的恒定速率因子与比特率之间的关系建模。ECP映射模块230将计算复杂性级别从CRF编码映射到编码计算简档和编码通过次数。

[0027] 基于RDC的视频代码转换模块400对由RDC训练模块200的CRF参考比特率模型220和ECP模块230引导的源视频150进行代码转换。在一个实施例中,基于RDC的视频代码转换模块400包括修正的CRF比特率模型410、视频描述长度模块420和基于RDC的代码转换器430。修正的CRF比特率模型410描述在恒定速率因子与实际比特率之间的关系以对源视频150编码。修正的CRF比特率模型410基于用于对输入视频编码的CRF参考比特率模型220被修正。视频描述长度模型420描述对源视频150编码的包括总体、时间和空间内容复杂性的计算复杂性。基于RDC的视频代码转换模块400基于在恒定速率因子和实际比特率与由源视频150的视频描述长度代表的计算复杂性之间的关系对源视频150进行代码转换。

[0028] II. 对视频比特率-质量关系建模:CRF参考比特率模型

[0029] CRF编码策略以实现使用恒定速率因子对视频编码的恒定视觉质量为目标。一般而言,为视频选择的CRF代表用于在一次编码通过中对视频编码的希望的质量级别。用于对视频编码的大CRF值造成大量化参数(“QP”)、高压缩比,并且因此造成低视觉质量。小CRF造成小QP、低压缩比,并且因此造成高视觉质量。为零的CRF值意味着无损压缩,并且因此意味着最高视觉质量。CRF编码通过根据视频帧中的运动数量动态调整QP来不同地压缩视频的不同帧。

[0030] 使用CRF编码的视频压缩性能可以由CRF参考比特率模型(例如,图2的CRF参考比特率模型220)建模,其中视频压缩质量可以由适当CRF值代表。使用训练的CRF参考比特率模型,基于RDC的视频编码系统102选择用于源视频的适当CRF值,该CRF值使在对源视频编码时使用的实际比特率能够接近与源视频的目标输出视频格式关联的目标比特率。另外,训练的CRF参考比特率模型提供关于需要多少平均比特率以使用CRF编码策略对视频编码的指示。用于用从训练的CRF参考比特率获得的CRF对源视频编码的实际比特率可以用来在对给定的源视频编码时修正CRF比特率模型以在与源视频的目标输出视频格式关联的比特率约束内实现恒定视觉质量。

[0031] 使用图2的RDC训练模块200作为示例,RDC训练模块200对视频全集160中的视频训练CRF参考比特率模型220。视频全集160存储具有多个空间分辨率、帧速率和比特率的大量多场景视频序列。RDC训练模块200的视频编码器210使用CRF编码用预先选择的CRF值对从视频全集160选择的每个视频编码并且从编码中获得实际比特率。本领域普通技术人员已知的任何存在的视频编码方案可以由视频编码器210用来对视频编码。在一个实施例中,视频编码器210是将选择的视频编码成HD.264/MPEG-4AVC格式的x.264视频编码器。视频编码器210的其它实施例可以使用VP8编码器或者其它可用视频编码器。

[0032] CRF参考比特率模型220使用来自对视频全集160中的视频编码的(crf, bitrate)作为训练采样以对在CRF与CRF编码的比特率之间的关系建模。在CRF与视频全集的视频的平均比特率之间的关系依赖于视频的空间分辨率、时间分辨率和在默认x.264CRF编码中固定的编码算法。在平均比特率和空间分辨率、时间分辨率与CRF之间的关系被描述如下：

$$[0033] \quad B = f(\text{crf}, M, T), \quad (1)$$

[0034] 其中B是平均比特率, M是视频的帧的亮度分量的像素数目(即, 空间分辨率), 并且T是每秒的帧数目(即, 时间分辨率)。由于参数crf、M和T通常相互独立, 所以在等式(1)中描述的平均比特率可以被表示如下：

$$[0035] \quad B = f_1(\text{crf}) \times f_2(M) \times f_3(T). \quad (2)$$

[0036] CRF参考比特率模型220将拟合采样表示为通过最小二乘法发现的四元组(crf, M, T, B)。

[0037] 对于对单场景和多场景视频编码的视频编码的平均性能, CRF参考比特率模型220按照视频中的场景数目对来自对视频编码的所得比特率加权并且按照在选择的数目的训练视频中的场景总数归一化比特率如下：

$$[0038] \quad B = \frac{\sum_i B_i \times S_i}{\sum_i S_i} \quad (3)$$

[0039] 其中B是平均比特率, B_i 是对第i个视频编码的实际比特率, 并且 S_i 是第i个视频中的场景数目。模型的残值误差是 $r_i = B_i - f(\text{crf}_i, M_i, T_i)$ 。CRF参考比特率模型220的拟合度由平方和误差(SSE)和均方根误差(RMSE)评估如下：

$$[0040] \quad SSE = \sum_{i=1}^n r_i^2 \quad (4)$$

$$[0041] \quad RMSE = \sqrt{\frac{\sum_{i=1}^n r_i^2}{n}} \quad (5)$$

[0042] 其中 r_i 是残值误差, 并且n是拟合采样数目。

[0043] 一般而言, 按照以帧每秒(fps)为单位的帧速率评估时间分辨率。为了研究时间分辨率对CRF参考比特率模型220的影响, 沿着时间轴对选择的用于训练CRF参考比特率模型220的视频进行下采样, 并且用视频的原有帧速率的一半对下采样的视频编码。以这一方式, 原有视频及其下采样的版本具有相同显示时间。CRF参考比特率模型220被训练以关于每个空间分辨率获得对视频的平均比特率, 例如, $\text{fps} = 12.5$ 和 $\text{fps} = 25$ 、 $\text{fps} = 15$ 和 $\text{fps} = 30$ 。

[0044] 训练结果表明比特率随着编码帧速率(fps)几乎线性增加, 换言之：

$$[0045] \quad \frac{B_1}{B_2} = \frac{\text{fps}_1}{\text{fps}_2}. \quad (6)$$

[0046] 基于在等式(6)中所示的线性关系, 在平均比特率与帧速率之间的关系可以被描述如下：

$$[0047] \quad y = a \times T, \quad (7)$$

[0048] 其中参数a包括来自空间分辨率和CRF的影响。

[0049] 图7图示了具有固定帧速率(即, $\text{fps} = 25$)的CRF参考比特率模型220的示例性训练

结果。在图7中所示的训练结果还表明平均比特率在固定CRF值和平均比特率时是空间分辨率的近似幂函数,并且在固定空间分辨率时是CRF的近似指数函数。在帧速率不同于fps=25时,可以用因子fps/25缩放模型。训练CRF参考比特率模型220的一个实施例是使用帧速率fps=25。

[0050] 为了探索在等式(4)中描述在平均比特率与空间分辨率M之间的关系,用通过对原有视频的低通滤波和下采样而获得的分级空间分辨率对仿真/推导的视频训练CRF参考比特率模型220。例如,视频编码器210将低通滤波器应用于具有空间分辨率704x576的原有视频以用于防混叠。视频编码器210对这些视频下采样以获得视频的具有大小352x288和176x144的更低分辨率版本。视频编码器210用相同CRF值对原有704x576视频及其下采样的更低分辨率版本编码以确定相对于每个CRF值在比特率与空间分辨率之间的关系。训练结果表明与不同CRF值对应的比特率-空间分辨率折线接近平行。比特率增加速率在减少CRF值时与空间分辨率的增加一起逐渐减少。用具有多种空间分辨率的视频进一步训练CRF参考比特率模型220,并且训练结果表明与用于仿真的视频的在比特率、空间分辨率和CRF值之间的关系相似的关系。

[0051] 基于训练结果,CRF参考比特率模型220对在平均比特率与空间分辨率之间的关系建模如下:

$$[0052] \quad y = b \times x^c, \quad (8)$$

[0053] 其中对于每个CRF值为 $0 < c < 1$ 。函数 $y = b \times x^c$ 通过坐标系的原点,这意味着零空间分辨率造成零比特率。 $0 < c < 1$ 指示在比特率与空间分辨率之间的单调增加关系,并且平均比特率具有比空间分辨率更小的增加速率,因为具有更高空间分辨率的视频在运动估计中具有更有效的帧内和帧间预测。在一个实施例中,参数c在对训练采样拟合时平均为0.65。可以从对使用在等式(8)中描述的幂函数对在平均比特率与空间分辨率之间的映射建模中获得用于每个CRF值的参数b。

[0054] CRF值可以取浮点值。在一个实施例中,CRF值在用于x.264CRF编码的范围[0,51]中。CRF=0用于无损编码,并且CRF=51用于具有最小文件大小和最小峰值信噪比(PSNR)的有损压缩。使用浮点RF值以训练CRF参考比特率模型220,仅改变CRF值,并且默认固定其它编码参数(例如帧速率)以用于将视频编码成某些目标空间和时间分辨率。采样CRF值12、14、...和34用于将视频编码成28-55dB的PSNR而具有视觉上可接受和令人满意的压缩质量。

[0055] 可以在以下等式(9)中描述在某个空间和时间分辨率之下在平均速率与CRF之间的关系:

$$[0056] \quad y = m \times e^{nx}. \quad (9)$$

[0057] 为了对等式(9)建模,将参数m设置成1380,并且将参数n设置成-0.20。拟合误差由SSE=540.3和RMSE=7.351评估。在等式(4)中描述的平均比特率B和CRF关系(crf,M,T,B)可以由以下等式(10)表示:

$$\begin{aligned}
 B &= f_1(crf) \times f_2(M) \times f_3(T) \\
 &= m \times e^{n \cdot crf} \times M^c \times \frac{T}{25} \\
 [0058] \quad &= 1380 \times e^{-0.2 \cdot crf} \times M^{0.65} \times \frac{T}{25}
 \end{aligned} \quad (10)$$

[0059] 假定经常用默认帧速率对视频编码,忽略参数T获得修改的等式(10)如下:

$$[0060] \quad B = f_1(\text{crf}) \times f_2(M) = m \times e^{n \cdot \text{crf}} \times M^c, \quad (11)$$

[0061] 因而,可以使用以下等式(12)从比特率B获得CRF:

$$[0062] \quad \begin{aligned} \text{crf} &= f_1^{-1} \left(\frac{B}{f_2(M) \times f_3(T)} \right) \\ &= 5 \times \ln \left(\frac{55.2 \times M^{0.65} \times T}{B} \right) \end{aligned} \quad (12)$$

[0063] 相对拟合误差由以下等式每空间分辨率地评估:

$$[0064] \quad E_r(M) = \frac{|B_a^d(M) - B_e^d(M)|}{B_a^d(M)} \quad (13)$$

[0065] 其中:

$$[0066] \quad B_a^d(M) = \frac{\left(\sum_{\text{crf}=1}^M \sum_{\text{video} \in \Omega_M} B_r(\text{crf}, M) \right)}{|\Omega_M| \times 12} \quad (14)$$

[0067] M是空间分辨率, $B_a^d(M)$ 是对具有空间分辨率M的视频的CRF平均的实际比特率, Ω_M 是具有空间分辨率M的视频集合,并且 $|\Omega_M|$ 是 Ω_M 的基数。 E_r 代表相对误差, $B_a^d(M)$ 代表平均实际比特率,并且 $B_e^d(M)$ 代表根据等式(10)估计的比特率。

[0068] III. 使用修正的CRF比特率模型的视频编码

[0069] 为了用从训练的CRF参考比特率模型220获得的CRF对输入视频编码,来自对输入视频编码的实际比特率可以从与输入视频关联的目标比特率偏离。需要在CRF与实际比特率之间的修正的CRF比特率模型以朝着用于对输入视频编码的目标比特率调整实际比特率。在一个实施例中,基于RDC的视频代码转换模块400使用修正的CRF比特率模型410以适配用于对输入视频编码的CRF值。修正的CRF比特率模型410可以使用以下等式(15)被修正:

$$[0070] \quad B = k \times f_1(\text{crf}) \times f_2(M) = k \times m \times e^{n \cdot \text{crf}} \times M^c, \quad (15)$$

[0071] 其中k是与将被编码的输入视频关联的修正因子。

[0072] 为了进一步例示修正,修正的CRF比特率模型410先通过用与输入视频关联的目标比特率替换参数B来从在等式(12)中描述的在CRF与平均比特率之间的映射中获得CRF。对输入视频编码的所得实际比特率及其对应CRF(即,实际比特率、CRF)用来确定等式(15)的修正因子k如下:

$$[0073] \quad k = \frac{B}{m \times e^{n \cdot \text{crf}} \times M^c} \quad (16)$$

[0074] 图5是用于示例视频序列的调整因子k、目标比特率 B_t 、实际比特率 B_a 的示例性数据,其中参数M代表视频序列的空间分辨率。

[0075] 使用训练的CRF参考比特率模型220和修正的CRF比特率模型410,基于RDC的视频代码转换模块400可以在与输入视频关联的比特率约束内用恒定视觉质量对输入视频编码。以下算法代表在与输入视频关联的比特率约束内用恒定视觉质量对输入视频编码的一

个实施例。

[0076] // *输入: 未压缩的视频及其目标比特率 B_t

[0077] // *输出: 在它的目标比特率内具有恒定视觉质量的压缩的视频

[0078] 1: 通过用 B_t 替换 B 来从等式(12)中的CRF参考比特率模型220发现 crf_t ;

[0079] 2: 用 crf_t 对视频编码, 并且获得实际比特率 B_a ;

[0080] 3: 通过(crf_t, B_a)对在等式(15)中获得修正的CRF比特率模型410;

[0081] 4: 通过用 B_t 替换 B 来从等式(15)的修正的CRF比特率模型发现 crf_a ;

[0082] 5: 用 crf_a 对视频编码并且获得修正的实际比特率 B'_a ;

[0083] 6: 如果 B'_a 并不接近 B_t , 则重复步骤3直至收敛。

[0084] 这里描述的算法是用从训练和修正的CRF比特率模型评估的CRF值的一个类型的多次通过CRF编码。步骤6中的收敛判据是在实际比特率 B'_a 与目标比特率 B_t 之间的差落在目标比特率的 $\pm 10\%$ 中。

[0085] 图4是将输入视频编码成在目标比特率内具有恒定视觉质量的输出视频的流程图。初始地, 基于RDC的视频代码转换模块400接收401输入视频, 并且输入视频具有用于对输入视频编码的作为比特率约束的目标比特率。在一个实施例中, 输入视频的目标比特率与输入视频的目标输出视频格式关联。在另一实施例中, 基于以下描述的视频描述长度(VDL)分析来修正输入视频的目标比特率。视频代码转换模块400相对于输入视频的目标比特率发现403恒定速率因子(例如, 使用等式(12)而用目标比特率替换平均比特率参数 B)。代码转换模块400用获得的CRF对输入视频编码405并且获得407用于对输入视频编码的实际比特率。视频代码转换模块用实际比特率计算409修正因子(例如, 使用等式(15))并且用修正因子发现411修改的恒定速率因子。视频代码转换模块400用修改的恒定速率因子对输入视频编码413并且确定415来自编码413的实际比特率是否在目标比特率的范围内。响应于实际比特率在目标比特率的范围内, 视频代码转换模块400输出417编码的视频。响应于实际比例特未在目标比特率的范围内, 视频代码转换模块400重复操作步骤409至417。

[0086] 朝着与输入视频关联的目标比特率对输入视频编码, 测试结果表明编码通过次数平均为1.8。如果输入视频具有与平均值接近的内容复杂性, 则利用由CRF参考比特率模型220发现的CRF值的一次通过CRF编码将朝着目标比特率对视频编码。如果输入视频具有从平均值偏离的内容复杂性, 则需要利用由自适应地修正的CRF比特率模型410发现的CRF值的多次通过CRF编码。

[0087] IV. 编码计算简档(ECP)

[0088] 在由视频序列捕获的场景中的可变内容造成在视频序列中包含的各种数量的信息。给定视频托管服务的大型视频全集, 将相同编码比特率应用于不同视频内容造成不同视频质量。另外, 视频托管服务100为向视频供应服务100上传的每个视频供应多种视频输出格式。不同输出格式在目标回放平台(例如, 移动电话和高清晰度电势)、递送约束(例如, 实时视频流传输和点播视频)和用户预期方面具有不同视频处理目标。

[0089] 一种对具有多种视频格式和内容复杂性的视频进行代码转换的解决方案是将不同比特率控制策略应用于不同视频。在一个实施例中, 基于RDC的视频编码系统102为给定的视频输出格式选择三个不同比特率控制策略和编码参数中的至少一个比特率控制策略和编码参数:

[0090] 控制_比特率,

[0091] 控制_质量,

[0092] 控制_质量_有界_比特率。

[0093] 控制_比特率策略是实现与视频的目标比特率尽可能接近的对视频编码的实际比特率。这一比特率控制策略适合用于移动视频流传输和使用拨号因特网连接的视频流传输。

[0094] 控制_质量策略是实现恒定目标视觉质量而无论实现它而需要的比特率如何。用于用可变复杂性对视频编码的控制_质量策略面临可变实际比特率,例如,用于对复杂视频编码的高比特率和用于简单视频的低比特率。这一比特率控制策略适合用于在分布式视频处理环境中的视频处理和流传输以及用于高视觉编码质量代码转换(例如,无损编码质量)。

[0095] 控制_质量_有界_比特率策略是尽可能实现恒定目标视觉质量而又维持对比特率的上界。因此,允许视觉质量仅对于复杂视频降至目标质量级别以下,其中这一质量在分配的最大比特率约束内不可实现。对于更简单的视频,只要实现了希望的目标质量级别,比特率就可以低于这一最大比特率。

[0096] 为了进一步提高视频代码转换性能,基于RDC的视频编码系统102对于使用给定的比特率控制策略的代码转换器支持等级计算级别。具体而言,对于比特率控制策略(例如,控制_质量_有界_比特率策略),可以例如有由基于RDC的视频编码系统102支持的五个不同编码计算级别:

[0097] 最快:提供最快代码转换时间,但是对于相同比特率未必产生最紧凑输出比特流或者最高视觉质量。对于大多数比特率控制策略和输出格式,进行一次通过编码。

[0098] 快速:提供合理快速编码。对于大多数比特率控制策略和输出格式,进行一次通过编码。

[0099] 中等:提供适合用于大多数比特率控制策略和输出格式的编码时间比对编码效率的折衷。

[0100] 慢速:使用更高计算成本的编码工具以便以更高编码效率(即,在相同比特率的更高视觉质量)产生代码转换的视频。例如,视频编码解码器x.264设置压缩效率控制参数“预设=更慢”或者“预设=placebo”以用高视觉质量对视频编码。

[0101] 最慢:使用计算要求高和复杂的编码工具以便产生最紧凑输出比特流和最高感知质量视频,比如电影工作室质量视频。当在PSNR方面被测量时,在这一编码计算级别的视觉质量可以等于或者低于由慢速模式实现的视觉质量。

[0102] 有了可用于代码转换器的多个编码计算级别,基于RDC的视频编码系统102实施各种代码转换策略以根据因素(比如在网络(例如,云计算网络)中的服务器上的当前系统负荷、对于实况流传输的需求或者可允许的代码转换延时)适配编码计算级别。例如,可以在低计算级别进行多个初始流传输代码转换以减少用于实况视频流传输的延时。在另一方面,可以用具有可接受的延时的更高计算级别进行视频流传输重新运行,其中重新运行是对视频的增加的受欢迎度做出响应。

[0103] 为了优化各种比特率控制策略并且支持用于每个比特率控制策略的等级编码计算级别,基于RDC的视频编码系统102遵循系统过程,该系统过程不仅评估多种实施不同比

特率控制策略的方式,而且还考虑用于每个比特率控制策略的最优计算分配。在一个实施例中,基于RDC的视频编码系统102训练编码计算简档(ECP)映射模块230以按照细粒度方式对计算效率折衷建模。

[0104] EPC是可以在x.264编码中被容易地集成的计算复杂性分配控制参数。ECP参数取在0与1之间的浮点值并且可以用来微调在比特率、失真和计算复杂性之间的折衷。现有比特率控制策略聚焦于在视频编码的速率与失真之间的折衷。然而,在编码计算被视为用于评估编码效率的判据时,需要与编码的计算一起调整在速率与失真之间的折衷。

[0105] 例如,在两次通过编码方案中,通常比第二次编码通过更快地进行第一次编码通过,因为第一次编码通过用来获得视频的一些统计信息。整个两次通过编码一般消耗比一次通过编码方案更高的计算。然而,并不清楚具有高计算复杂性的一次通过编码是否比具有更少计算复杂性的两次通过编码效率更高或者效率更低。基于RDC的视频编码系统102的ECP映射模块230被设计用于发现用于基于多次通过的代码转换(例如,基于x.264的多次通过代码转换)的最优计算分配以及用于给定的编码计算级别的编码通过次数。

[0106] 现在转向图2,对存储于视频全集160中的视频训练RDC训练模块200的ECP映射模块230。对于每个编码策略,针对用于多次通过代码转换的不同可能实现方式和不同可能计算级别训练ECP映射模块230以获得原始速率-失真-复杂性数据。例如,使用对原始速率、失真、复杂性数据的多判据帕雷托最优集合分析来进一步优化ECP映射模块230以推导在多个相等间距的编码计算级别用于每个比特率控制策略的最佳代码转换实现方式和计算分配(例如,在一个实施例中,可以有五个编码计算级别:最快、快速、中等、慢速和最慢计算级别)。

[0107] 取多次通过x.264代码转换作为示例,ECP映射模块230通过对 $\{ecp_0, ecp_1\}$ 代表用于两次通过编码的计算分配。 ecp_0 代表一次通过比特率编码或者两次通过比特率编码的第一次通过的计算复杂性。 ecp_1 代表两次通过比特率编码的第二次通过的计算复杂性。例如, $\{0.07, \text{NULL}\}$ 代表一次通过比特率编码实现方式的计算复杂性,并且 $\{0.07, 0.71\}$ 是两次通过比特率编码实现方式的示例计算复杂性。

[0108] 对于每个比特率控制策略,对来自视频全集160的视频集合离线训练ECP映射模块230。每个训练采样与速率-失真-复杂性性能、 $\{R, D, C\}$ 三元组关联,其中从训练中获得实际编码比特率(R),失真(D)由视频帧的亮度分量的均方误差(MSE)测量,并且复杂性(C)由每帧编码时间代表。以下是用于利用多次通过x.264代码转换方案对于控制_比特率比特率控制策略训练ECP映射模块230的算法:

[0109] 1:用所有可能的 $\{ecp_0, ecp_1\}$ 对和不同目标比特率使用控制_比特率策略对训练视频编码;

[0110] 2:从训练中获得 $\{R, D, C\}$ 性能采样;

[0111] 3:对于训练视频对于每个 $\{ecp_0, ecp_1\}$ 对针对速率R对失真D和计算时间C平滑地插值;

[0112] 4:发现在 $\{R, D, C\}$ 采样之中的帕雷托最优集合并且对帕雷托集合插值以获得分段连续 $\{R, D, C\}$ 表面;

[0113] 5:将计算时间C划分成5个均匀间隔的连续级别;

[0114] 6:选择速率值并且在选择的比特率值对于五个计算级别中的每个计算级别在帕

雷托最优表面上选择{ecp₀,ecp₁}；

[0115] 7:通过回归测试来验证优化的策略对视频的有效性。

[0116] 在下面的表1中例示利用控制_比特率控制策略的用于五个计算级别的优化的计算分配：

[0117] 表1:利用x.264的用于控制_比特率控制策略的ECP参数

[0118]

计算级别	ecp ₀	ecp ₁
最快	0.14	空值
快速	0.35	空值
中等	0.07	0.50
慢速	0.07	0.71
最慢	0.42	0.71

[0119] 如表1中所示,对于每个计算级别,ECP映射模块230提供一对ECP值{ecp₀,ecp₁} ,该对ECP值指示编码通过次数和所需计算复杂性分配控制参数值。例如,对于使用控制_比特率控制策略来对具有中等计算复杂性的视频编码,两次编码通过可以用来实现优化的速率-失真-复杂性折衷,并且对于第一次编码通过,ECP值为0.07,而对于第二次编码通过为0.50。为了对用于实时流传输的视频编码,编码时间需要快速而无不可接受的延迟。基于RDC的视频编码系统102可以选择具有ECP值0.35的一次通过编码以平衡速率-失真-复杂性折衷。

[0120] 在另一实施例中,使用另一视频编码方案VP8(也被称为WebM)来训练ECP映射模块230。取代使用编码计算简档参数(即,ecp₀,ecp₁) ,VP8使用编码速度控制(ES)参数以管理在速率、失真和计算复杂性之间的折衷。表2示出了使用VP8编码方案、利用控制_比特率控制策略的用于五个计算级别的计算分配的示例：

[0121] 表2:利用VP8的用于控制_比特率控制策略的ECP参数

[0122]

计算级别	ES ₀	ES ₁
最快	219	空值
快速	218	空值
中等	220	217
慢速	220	218
最慢	218	216

[0123] 对于其它比特率控制策略(比如控制_质量策略和控制_质量_有界_比特率策略)对来自视频全集160的视频相似地训练ECP映射模块230。表3示出了利用x.264编码方案、利用控制_质量控制策略的用于五个计算级别的优化的计算分配和用于VP8编码方案的对应的计算分配。为了实现恒定目标质量,ECP映射模块230向基于RDC的视频代码转换仅建议一次通过编码。表4示出了利用x.264编码方案、利用控制_质量_有界_比特率控制策略的用于五个计算级别的优化的计算分配和用于VP8编码方案的对应的计算分配。

[0124] 表3:用于控制_质量控制策略的ECP参数和ES参数

[0125]

计算级别	x.264	VP8
	<i>ecp</i>	ES
最快	0.14	219
快速	0.35	218
中等	0.50	217
慢速	0.71	216
最慢	0.78	216

[0126] 表4:用于控制_质量_有界_比特率控制策略的ECP参数和ES参数

[0127]

计算级别	x.264		VP8	
	<i>ecp</i> ₀	<i>ecp</i> ₁	ES ₀	ES ₁
最快	0.14	空值	220	219
快速	0.35	空值	220	218
中等	0.14	0.50	220	218
慢速	0.14	0.71	220	217
最慢	0.50	0.71	220	216

[0128] 图6是离线训练EPC映射模块230以用于速率-失真复杂性优化的流程图。初始地，RDC训练模块200从视频全集160接收610视频。RDC训练模块200的视频编码器210对选择的视频的集合编码620并且从训练中获得630{R,D,C}采样。EPC映射模块230从获得的{R,D,C}采样中发现640最优{R,D,C}采样。例如，ECP映射模块230发现在{R,D,C}采样之中的帕雷托最优集合并且对帕雷托集合插值以获得分段连续{R,D,C}表面。ECP映射模块230将最优{R,D,C}采样空间划分650成多个计算复杂性级别(例如，以上描述的五個计算级别)。ECP映射模块230选择660用于每个计算级别的RDC最优编码策略(例如，控制_质量_有界_比特率策略)并且获得670用于每个计算级别的ECP参数和编码通过次数。

[0129] 基于RDC的视频编码系统102可以通过基于先前编码通过调整后续编码通过的编码参数来优化多次通过编码性能。例如，在用于对输入视频编码的两次通过编码方案中，基于RDC的视频编码系统102可以基于第一次编码通过的编码参数(例如，比特率、运动估计方法和运动矢量搜索范围)适配第二次编码通过的编码参数。如果第一次编码通过的实际比特率小于用相同ECP值编码的平均比特率，则它指示输入视频具有比平均视频更简单的视频内容，可以减少第二次编码通过的ECP值以节省用于对输入视频编码的计算而保持相同速率-失真性能。如果第一次编码通过的实际比特率大于用相同ECP值编码的平均比特率，则可以增加第二次编码通过的ECP值以实现用于对输入视频编码的更多速率-失真性能增益。

[0130] V. 用视频描述长度(VDL)编码

[0131] 在最小成本用可接受的视觉质量对源视频进行代码转换需要有效地评估源视频的视频编码复杂性。忽略视频内容复杂性对代码转换视频的影响降低代码转换的视频的视觉质量和用户体验。关于需要多少比特率以在某个视觉质量对视频编码的信息对于视频编码效率优化有用,因为用于对输入视频编码的比特率信息反映源视频的视频内容复杂性。

[0132] 在一个实施例中,基于RDC的视频编码系统102使用视频描述长度以代表视频的视频内容复杂性。VDL是在某个视觉质量对视频编码而需要的比特率。不同视频具有不同VDL。具有慢速运动和平滑图像的视频具有短VDL,这指示视频易于用小比特率、但是高视觉质量来编码。在另一方面,具有密集运动和复杂文本内容的视频具有长VDL,这指示视频难以编码。

[0133] 视频描述长度对于许多应用(比如自适应视频编码和解码)有用。对于视频代码转换,VDL引导基于RDC的视频代码转换器以选择目标比特率或者目标恒定速率因子以将视频高效地编码成它的输出视频格式。基于RDC的视频代码转换器可以将输入视频的复杂性与参考视频在它们的关联VDL方面进行比较,并且通过引用参考视频的现有合理编码参数来确定用于输入视频的恰当编码参数。

[0134] 视频的VDL信息可以具有至少三个形式:

[0135] 总VDL,

[0136] 时间VDL,以及

[0137] 空间VDL。

[0138] 总VDL是视频在用‘CRF=26,复杂性=最快’对它编码时的实际比特率。CRF值是可配置的参数,并且实验数据表明视频在用CRF编码、用CRF=26对视频编码时有良好视觉质量。

[0139] 时间VDL是视频在它用两个不同复杂性(例如,‘CRF=26,复杂性=快速’和‘CRF=26,复杂性=最快’)对它编码时的实际比特率的差。视频的时间VDL反映视频的时间复杂性。具有不同时间复杂性的视频主要不同在于多么快地处理在视频中被检测到的运动。将多么快地处理视频中的检测到的运动依赖于一个或者多个因素,比如用于运动估计的运动估计方法和参考帧数目。表示时间VDL的一个实施例是使用视频当在不同复杂性水平对视频编码时的比特率差。

[0140] 空间VDL是视频在它用两个不同CRF(例如,‘CRF=26,复杂性=最快’和‘CRF=28,复杂性=最快’)对它编码时的实际比特率的差。一般而言,由运动估计算法预测的残值信号复杂性反映视频帧的复杂性,并且残值信号的变换系数经历由CRF值代表的量化。因此,用不同CRF值、但是相同复杂性级别编码的视频的比特率差减少时间因素影响,并且反映视频的空间复杂性。具有更大空间VDL的视频在视频中具有更复杂的空间内容。

[0141] 视频的总VDL、时间VDL和空间VDL可以用来引导基于RDC的视频代码转换。例如,参照图4,基于RDC的视频代码转换模块400的视频描述长度模块420从对输入视频编码中获得输入视频的总VDL、时间VDL和空间VDL。输入视频的获得的VDL数据被存储于动态VDL数据库中作为用于输入视频的参考VDL。例如,视频描述长度模块420维持用于总VDL的动态参考表,其中每个表条目存储输入视频的标识、输入视频的总VDL、输入视频的时间分辨率和空间分辨率。在对视频进行代码转换之后,视频描述长度模块420删除输入视频的VDL参考。视频描述长度模块420相似地分别维持用于时间VDL和空间VDL的动态参考表。为输入视频的

一个视频输出格式获得的总VDL可以用来引导输入视频的用于其它视频输出格式的编码。为输入视频的一个视频输出格式获得的时间和空间VDL可以相似地用于输入视频的其它视频输出格式。

[0142] 在另一实施例中,基于RDC的视频编码系统102基于存储于动态VDL参考表中的VDL数据为视频全集160的视频的总VDL、时间VDL和空间VDL生成静态VDL参考表。例如,用于总VDL的静态VDL参考表对于选择的时间分辨率存储用于每个个别输入数据的平均总VDL数据,并且另一静态VDL参考表对于选择的空间分辨率用于总VDL。可以相似地生成对于选择的时间分辨率或者空间分辨率的用于时间VDL和空间VDL的静态VDL参考表。可以每当生成或者更新在动态VDL参考表中的记录时更新静态VDL参考表。

[0143] 对于给定的输入视频,视频描述长度模块420如以上描述的那样计算输入视频的总VDL、时间VDL和空间VDL并且与基于RDC的代码转换器430通信以用于速率-失真复杂性优化。例如,响应于输入视频的总VDL小于输入视频的目标比特率,视频描述长度模块420设置输入视频的总VDL作为目标比特率。设置总VDL作为输入视频的目标比特率帮助获得良好视觉质量。响应于输入视频的空间VDL大于参考视频的空间VDL,视频描述长度模块420指令基于RDC的代码转换器430增加关于空间处理的编码的复杂性水平(例如,将复杂性水平从例如“中等”改变成“快速”)。相似地,视频描述长度模块420指令基于RDC的代码转换器430基于在输入视频的时间VDL与参考视频的时间VDL之间的比较关于输入视频的时间处理来调整编码的复杂性水平。

[0144] 图3进一步图示了以上描述的VDL引导的视频代码转换。初始地,基于RDC的视频代码转换模块400接收310输入视频并且经由修正的CRF比特率模型410对输入视频编码。视频描述长度模块420获得320输入视频的总VDL、时间VDL和空间VDL。视频描述长度模块420从修正的CRF比特率模型420的编码中获得输入视频的VDL信息或者自行计算输入视频的VDL信息。

[0145] 视频描述长度模块420比较330输入视频的总VDL与输入视频的目标比特率并且基于比较来调整340输入视频的目标比特率(例如,响应于总VDL小于目标比特率来设置输入视频的总VDL作为输入视频的目标比特率)。视频描述长度模块420比较350输入视频的时间VDL与参考时间VDL并且基于该比较在时间上调整360视频编码复杂性。视频描述长度模块420比较370输入视频的空间VDL与参考空间VDL并且基于该比较在空间上调整380视频编码复杂性。基于RDC的代码转换器430如图4中所示用调整的目标比特率、时间编码复杂性和空间编码复杂性对输入视频编码390。调整的目标比特率、时间或者空间编码复杂性指示需要不同ECP值和编码通过次数以反映目标比特率和编码复杂性更新。

[0146] 包括以上描述以例示优选实施例的操作而并非为了限制本发明的范围。本发明的范围将仅由所附权利要求限制。根据以上讨论,许多变化将为相关领域技术人员所清楚而仍将为本发明的精神实质和范围所涵盖。

[0147] 已经关于一个可能的实施例用具体细节描述了本发明。本领域技术人员将领会到可以在其它实施例中实践本发明。首先,部件的特定命名、术语的大写、属性、数据结构或者任何其它编程或者结构方面并非必需或者重要的,并且实施本发明或者它的特征的机制可以具有不同名称、格式或者协议。另外,可以如描述的那样经由硬件与软件的组合或者完全在硬件单元中实施系统。在这里描述的各种系统部件之间的特定功能划分也仅为示例性的

而非必需;由单个系统部件执行的功能可以作为替代由多个部件执行,并且由多个部件执行的功能可以作为替代由单个部件执行。

[0148] 以上描述的一些部分在对信息的操作的算法和符号表示方面呈现本发明的特征。这些算法描述和表示是本领域技术人员用来向本领域其他技术人员最有效地传达他们的工作实质的手段。这些操作当在功能或者逻辑上被描述之时被理解为由计算机程序实施。另外,将这些操作布置称为模块或者冠以功能名称也已经证实有时是便利的而不失一般性。

[0149] 除非如从当前讨论中清楚的那样另有具体明示,应领会到,贯穿说明书,利用诸如“处理”或者“计算”或者“运算”或者“确定”或者“显示”等术语的讨论指代计算机系统或者相似电子计算设备的动作和过程,该计算机系统或者相似电子计算设备操控和变换在计算机系统存储器或者寄存器或者其它这样的信息存储装置、传输或者显示设备内被表示为物理(电子)数量的数据。

[0150] 本发明的某些方面以算法的形式包括这里描述的过程步骤和指令。应当注意,可以在软件、固件或者硬件中体现本发明的过程步骤和指令,并且当在软件中被体现时,可以下载这些过程步骤和指令以驻留于由实时网络操作系统使用的不同平台上并且从这些不同平台操作这些过程步骤和指令。

[0151] 本发明也涉及一种用于执行这里的操作的装置。这一装置可以被具体构造用于所需目的,或者它可以包括在计算机可以访问的计算机可读介质上存储的计算机程序有选择地激活或者重新配置的通用计算机。这样的计算机程序可以被存储于计算机可读存储介质(比如但不限于包括软盘、光盘、DVD、CD-ROM、光磁盘的任何类型的盘、只读存储器(ROM)、随机存取存储器(RAM)、EPROM、EEPROM、磁卡或者光卡、专用集成电路(ASIC)或者适合用于存储电子指令并且各自耦合到计算机系统总线的任何类型的介质)中。另外,在说明书中指代的计算机可以包括单个处理器或者可以是运用多处理器设计以用于增加的计算能力的架构。

[0152] 这里呈现的算法和显示并非固有地与任何特定计算机或者其它装置有关。各种通用系统也可以与根据这里的教导的程序一起被使用,或者构造更专门化的装置以执行方法步骤可以证实是便利的。用于多种这些系统的结构将与等效变化一起为本领域技术人员所清楚。此外,未参照任何特定编程语言描述本发明。将领会到,多种编程语言可以用来实施如这里描述的本发明的教导,并且提供对具体语言的任何引用以便公开本发明的实现和最佳实施方式。

[0153] 本发明良好的适合于在许多拓扑之上的广泛多种计算机网络系统。在这一领域内,大型网络的配置和管理包括存储设备和通过网络(比如因特网)通信地耦合到不相似计算机和存储设备的计算机。

[0154] 最后,应当注意,在说明书中使用的言语已经主要出于可读性和教导的目的而被选择并且可以未被选择用来界定或者限制发明主题内容。因而,本发明的公开内容旨在于举例说明而非限制在所附权利要求中阐述的本发明的范围。

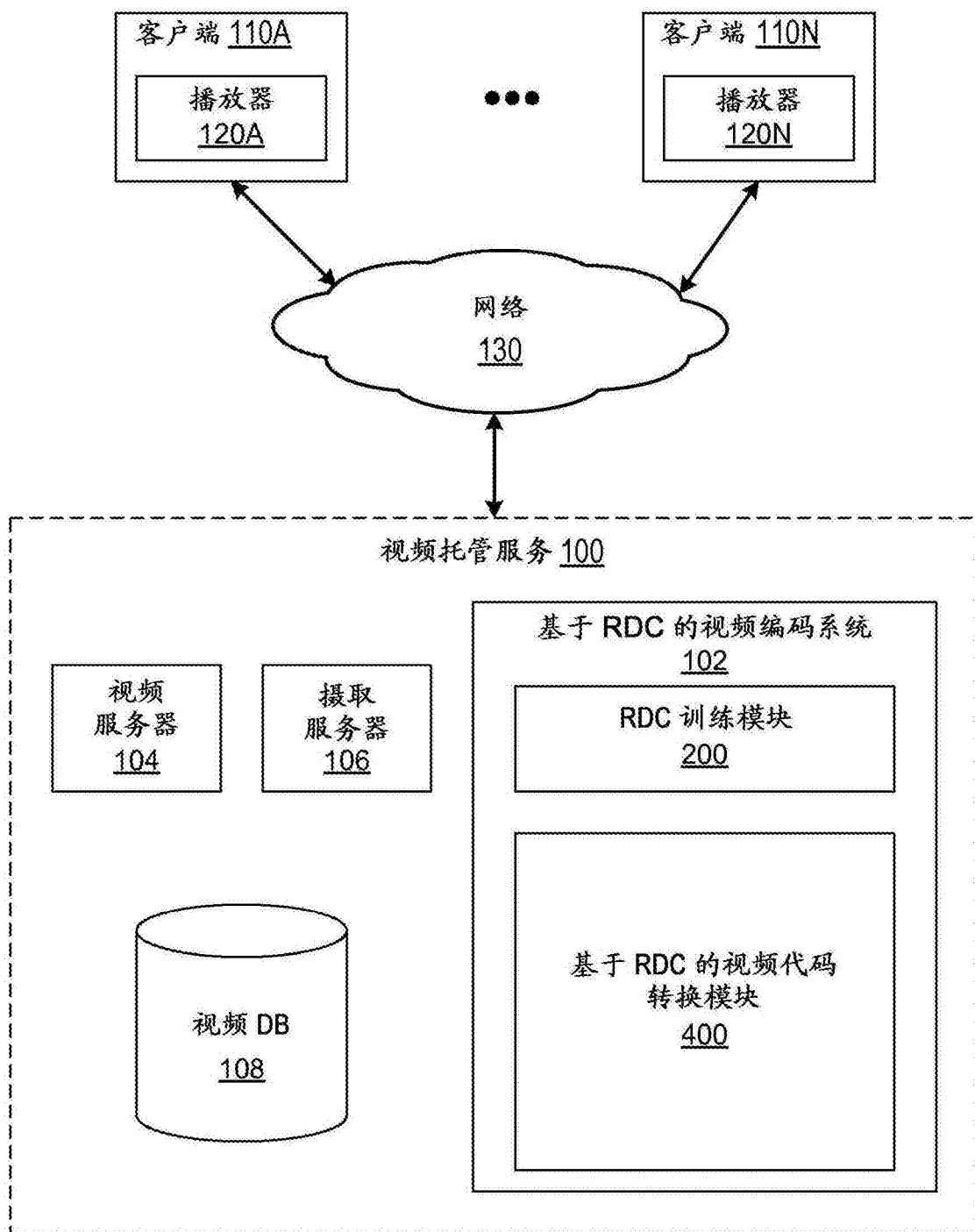


图1

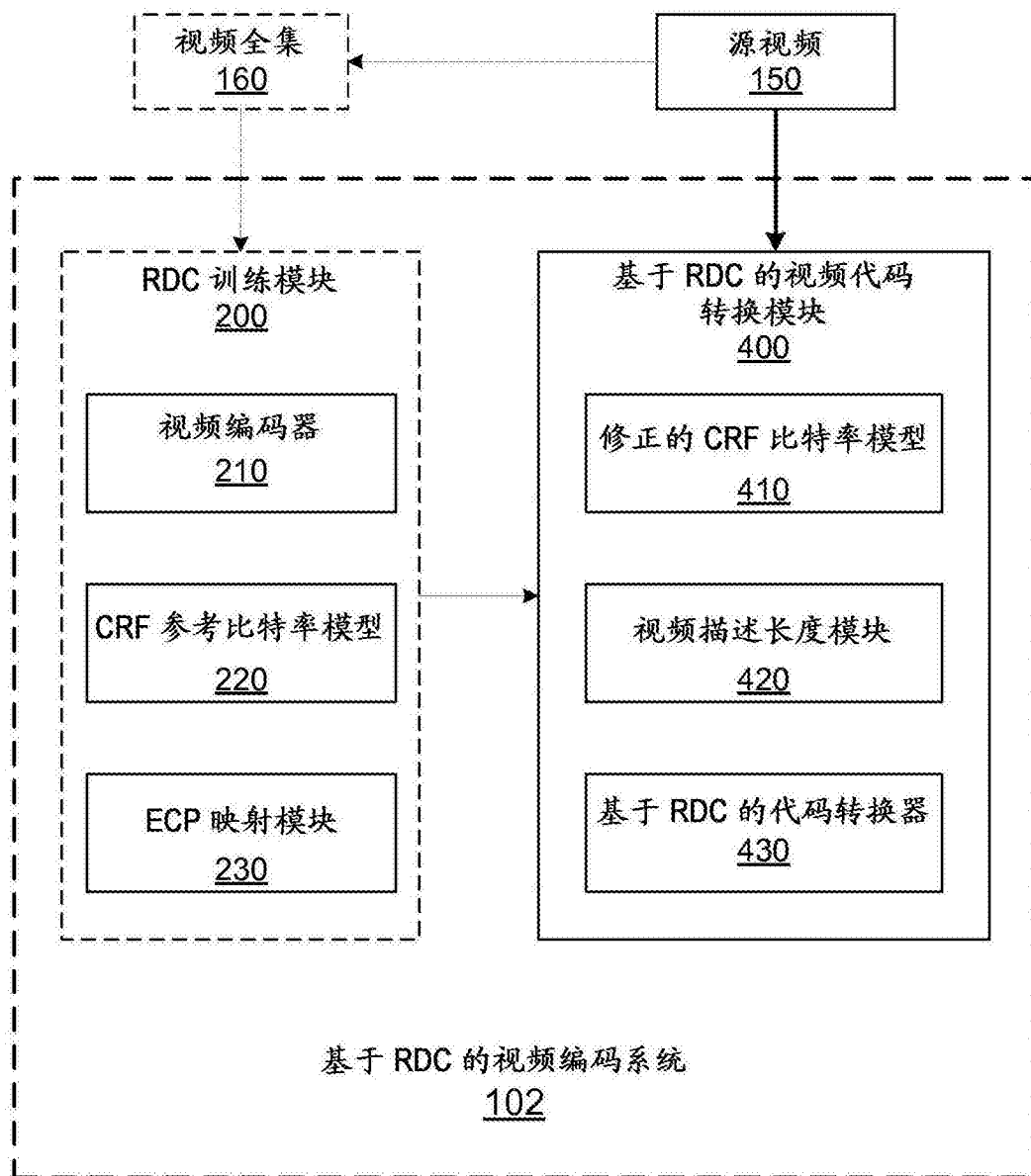


图2

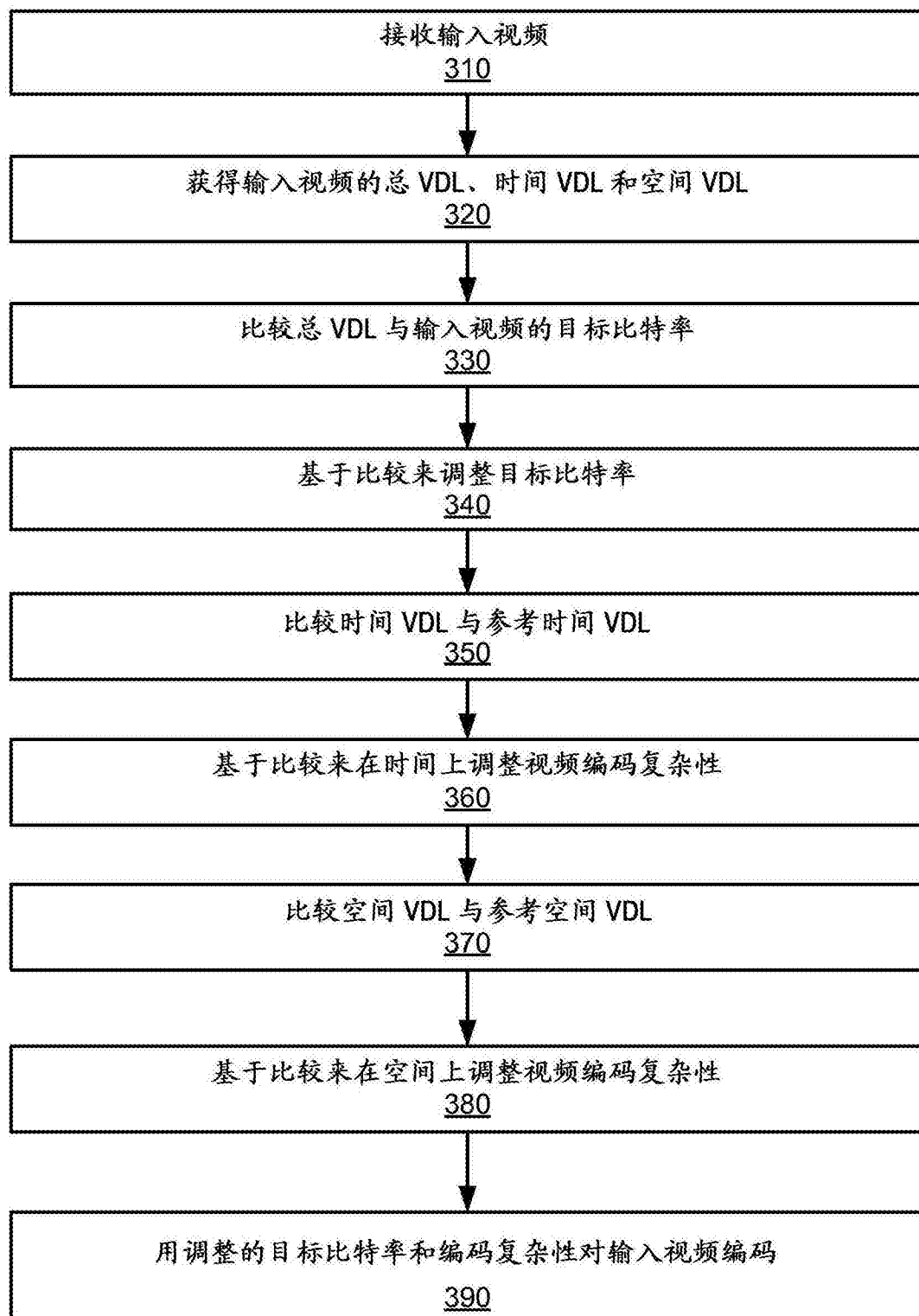


图3

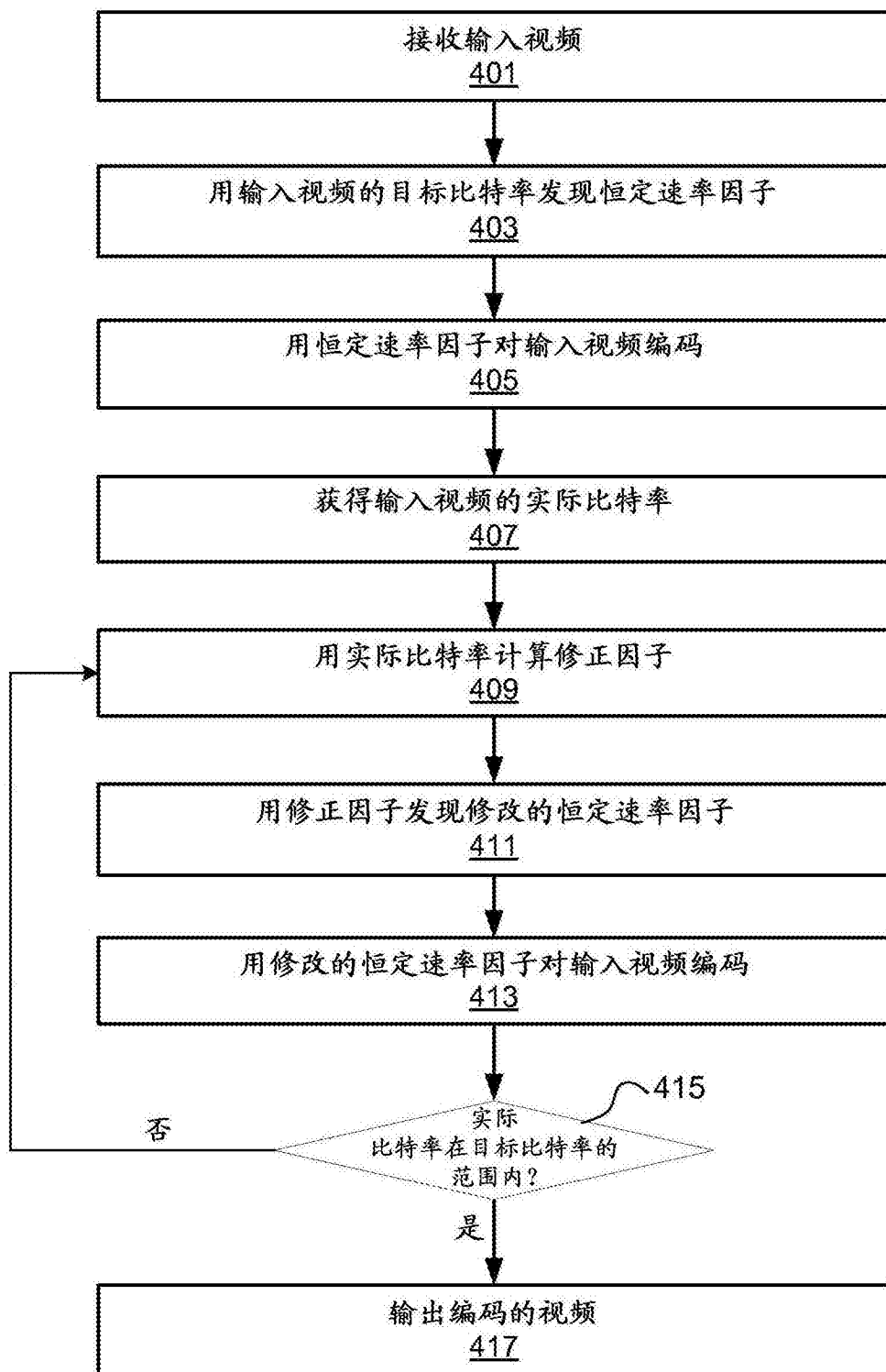


图4

视频	M	B_t	k	B_a
移动台	176x144	100	0.50	91.53
花朵	352x288	300	0.75	293.95
网球	352x240	300	0.66	291.84
公园跑道	640x360	600	0.80	622.16
海港	704x576	1500	0.69	1457.02
公园跑道	1280x720	2500	1.05	2534.98
行人	1920x1080	3500	0.82	3313.23

图5

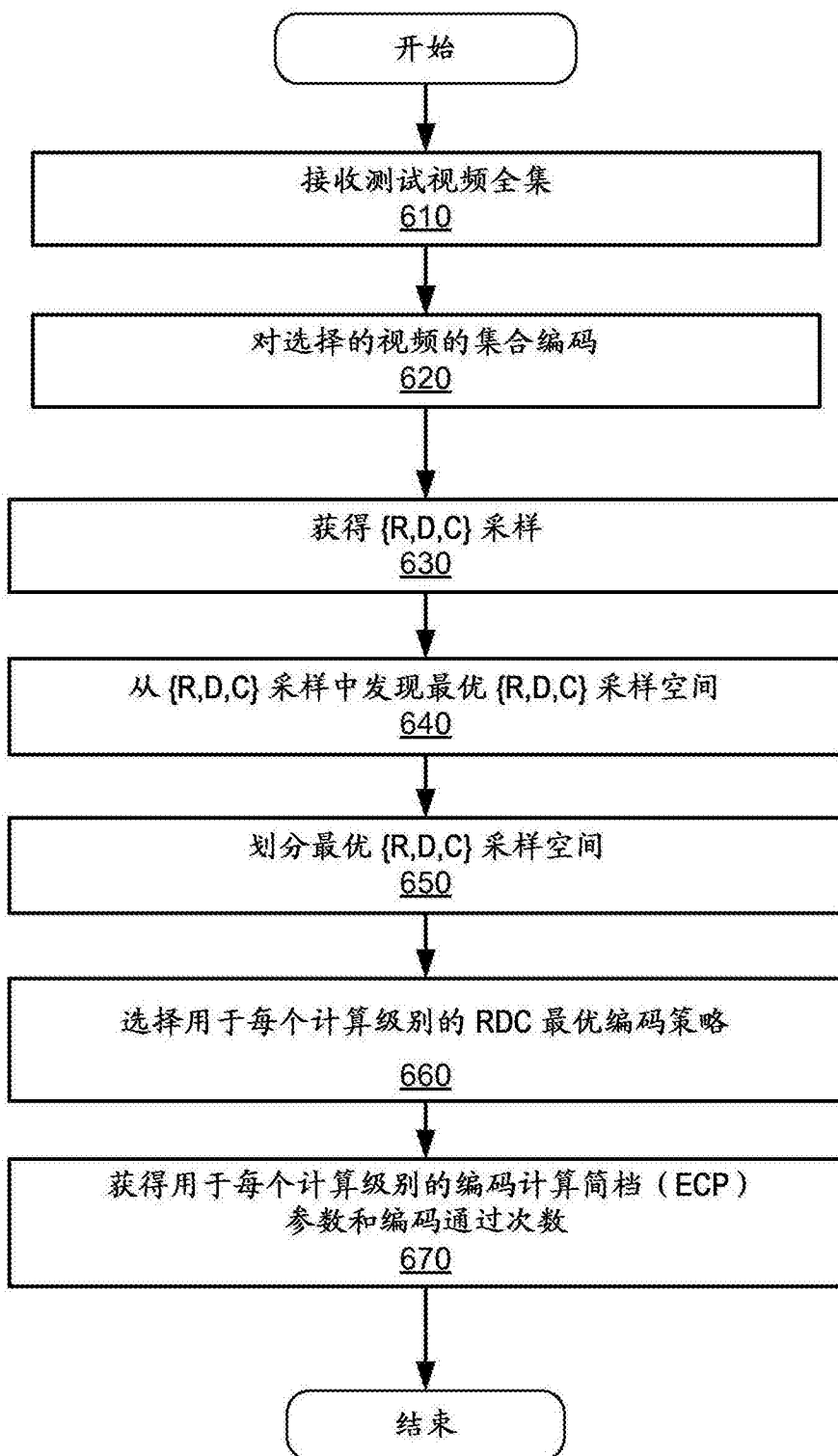


图6

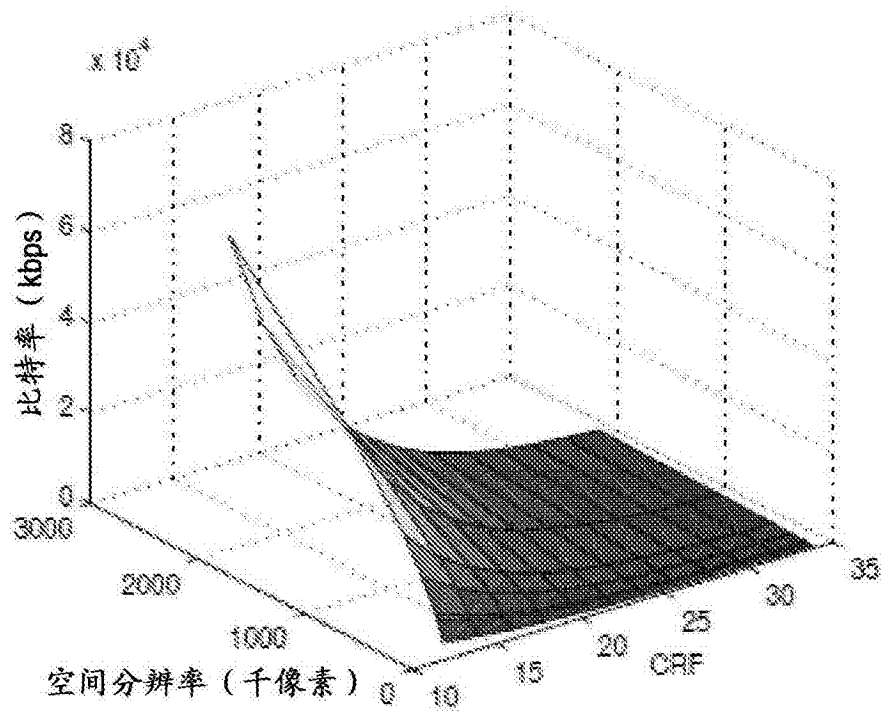


图7