

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2018 年 5 月 24 日 (24.05.2018)



(10) 国际公布号
WO 2018/090774 A1

(51) 国际专利分类号:
H04N 21/2662 (2011.01) **H04N 21/24** (2011.01)

(21) 国际申请号: PCT/CN2017/106444

(22) 国际申请日: 2017 年 10 月 17 日 (17.10.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201611032465.0 2016 年 11 月 16 日 (16.11.2016) CN

(71) 申请人: 上海交通大学 (SHANGHAI JIAO TONG UNIVERSITY) [CN/CN]; 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。

(72) 发明人: 熊红凯 (XIONG, Hongkai); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。
李成林 (LI, Chenglin); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。

(74) 代理人: 上海恒慧知识产权代理事务所 (特殊普通合伙) (SHANGHAI HENGHUI INTELLECTUAL

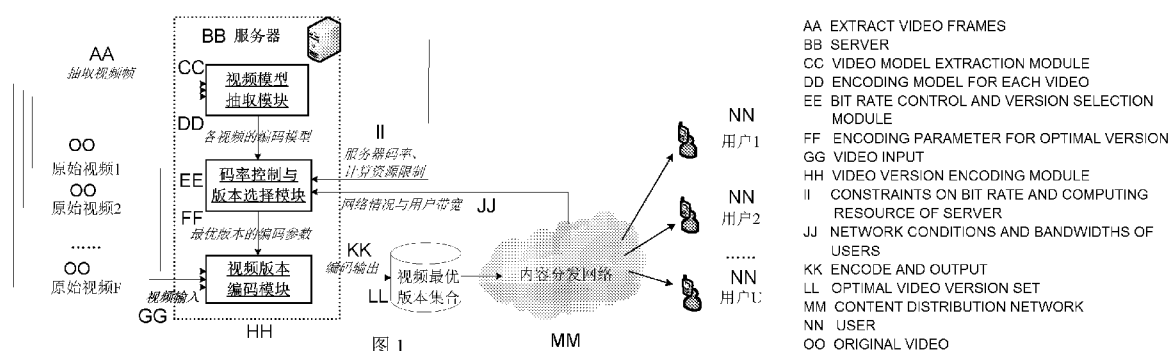
PROPERTY AGENCY); 中国上海市闵行区剑川路 953 弄 322 号 1 楼, Shanghai 200240 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,

(54) Title: METHOD AND SYSTEM FOR BIT RATE CONTROL AND VERSION SELECTION FOR DYNAMIC ADAPTIVE VIDEO STREAMING MEDIA

(54) 发明名称: 动态自适应视频流媒体的码率控制与版本选择方法及系统



(57) Abstract: The present invention provides a method and system for bit rate control and version selection for dynamic adaptive video streaming media at an encoding end. The method adopts a dynamic adaptive streaming media encoding technology to encode respective original videos into a plurality of versions with different bit rates at a server, and determines an original video version subset to be encoded for each original video and a specific encoding parameter of each video version by taking into consideration a complexity-bit rate-distortion encoding model for different original video content, constraints on an encoding bit rate and a computing resource of the video server, network connection conditions of different users, and a video on demand probability distribution, and the video server finally encodes and outputs an optimal video version set, so as to maximize the overall quality of videos viewed by users. The method of the present invention improves the encoding efficiency of a dynamic adaptive streaming media encoding technology when an encoding bit rate and a computing resource are limited, and provides better video service quality for users.

RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

(57) 摘要：本发明提供了一种动态自适应视频流媒体的编码端码率控制与版本选择的方法及系统，所述方法结合动态自适应流媒体编码技术，在服务器处将各原始视频编码为多个不同码率的版本，同时兼顾各不同原始视频内容的编码复杂度-码率-失真模型、视频服务器的编码码率和计算资源限制、不同用户的网络连接情况以及视频点播概率分布，确定各原始视频所需编码的视频版本子集以及每个视频版本的具体编码参数，视频服务器最终通过编码输出视频最优版本集合，以实现用户观看视频整体质量的最大化。本发明所述方法提高了在编码码率和计算资源受限情况下动态自适应流媒体编码技术的编码效率，为用户提供更佳的视频服务质量。

动态自适应视频流媒体的码率控制与版本选择方法及系统

技术领域

本发明涉及一种数据通信技术领域的方法和系统，具体地，涉及一种适用于动态自适应视频流媒体的编码码率控制与版本选择方法及系统。

背景技术

随着移动数据流量的迅速增长以及智能终端设备的日益普及，以移动视频服务为代表的无线视频流媒体技术在近几年得到了越来越广泛的应用。与此同时，移动用户所使用的移动设备终端、点播内容以及网络连接性等方面呈现出更为复杂的异构特性，从而大大增加了视频流媒体服务复杂度和难度。动态自适应流媒体技术能够为用户提供同一视频内容的不同版本，以提高异构网络中用户的视频观看满意度。其中，每个视频版本以给定的码率和/或分辨率进行编码，从而使得各用户能够根据自己的视频点播需求和网络状况来确定下载最合适的视频版本。

一方面，单一码率的视频编码技术需要消耗很高的编码复杂度以实现高更高的视频压缩性能。对于执行动态自适应流媒体技术的编码服务器而言，往往受到自身物理功耗的限制，无法通过编码得到过多的视频码率版本以适应不同用户的需求，因此需要将有限的计算资源合理地分配给各个视频的不同版本。另一方面，收到服务器本身存储空间的限制以及网络传输中瓶颈带宽的限制，所有视频的各不同版本的码率总和也收到限制，因此还需要将有限的码率资源合理的分配给各个视频的不同版本。

经过对现有技术的检索发现，L. Toni 等人在《ACM Transactions on Multimedia Computing, Communications, and Applications, Feb. 2015, (计算机协会多媒体计算、通信与应用学报, 2015 年 2 月)》上发表了题为“Optimal selection of adaptive streaming representations (自适应流媒体版本的最优化选择)”的文章，以及 R. Aparicio-Pardo 等人在《Proceedings of ACM Multimedia Systems Conference, Mar. 2015, pp. 49-60, (计算机协会多媒体系统会议论文集, 2015 年 3 月, 第 49-60 页)》上发表了题为“Transcoding live adaptive video streams at a massive scale in the cloud (大规模云端转码直播自适应视频流)”的文章，这两篇文章分别研究了服务器在编码码率和功耗受限情况下，如何最优

地选择各视频版本的编码码率和分辨率。但是，上述工作的主要方法是将最优版本选择问题建模为一个极为复杂的整数线性规划问题，通过求解该问题得到最优的视频编码版本。这一方法的复杂度随系统的规模呈指数趋势增长，因此需要消耗极高的复杂度和计算资源，从而占用了服务器进行视频编码的计算资源，具有很大的局限性。另一方面，上述工作假设各视频的不同视频版本已经通过预编码获得，其最终目标是从这些已知版本中选取最优的版本子集，这一假设在实际系统中并不成立。例如在实时视频通信系统中，所有视频流通过实时编码后将码率传输给用户，因此没有足够的时间对所有视频进行预编码。

发明内容

针对现有技术中的缺陷，本发明的目的是提供一种适用于动态自适应视频流媒体的编码码率控制与版本选择的方法及系统。

为实现以上目的，本发明采用的技术方案是：

根据本发明的一个方面，提供一种动态自适应视频流媒体的编码码率控制与版本选择方法，所述方法结合动态自适应流媒体编码技术，在服务器处将各原始视频编码为多个不同码率的版本，同时兼顾各不同原始视频内容的编码复杂度-码率-失真模型、服务器的编码码率和计算资源限制、不同用户的网络连接情况（比如用户带宽资源）以及视频点播概率分布，确定各原始视频所需编码的视频版本子集以及每个视频版本的具体编码参数，服务器最终通过编码输出视频最优版本集合，以实现用户观看视频整体质量的最大化。

具体的，所述动态自适应视频流媒体的编码码率控制与版本选择方法，包括以下步骤：

第一步、视频模型抽取步骤：

抽取各原始视频的前若干帧，得到各原始视频的编码复杂度-码率-失真模型；

第二步、码率控制与版本选择步骤：

根据服务器处可编码得到的各视频不同版本组成的全集、第一步中得到的各原始视频的编码复杂度-码率-失真模型、服务器的编码码率和计算资源限制、以及用户的网络连接情况和视频点播概率分布，建立适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题，采用最优版本集合码率控制与版本选择方法，得到各视频各最优版本的编码参数；

第三步、视频版本编码步骤：

使用动态自适应流媒体编码技术, 并采用第二步中得到的各视频各最优版本的编码参数, 将各原始视频编码为多个不同码率的视频版本, 输出得到视频最优版本集合。

优选地, 第一步中, 所述的前若干帧为任意大于二的正整数帧。

优选地, 第一步中, 所述各原始视频由于内容的不同呈现出不同的编码复杂度-码率-失真性能。

优选地, 第一步中, 所述的编码复杂度-码率-失真模型将编码复杂度、码率和失真表示为具体编码参数的函数。

优选地, 第二步中, 所述的适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题, 结合服务器处可编码得到的各视频不同版本组成的全集、第一步中得到的各视频的编码复杂度-码率-失真模型、服务器的编码码率和计算资源限制、以及用户的网络连接情况和视频点播概率分布, 并使用网络效用最大化建模方法得到。

更优选地, 所述的建立适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题, 包括:

$$\text{目标优化问题: } \max_{\mathcal{A} \subseteq \mathcal{V}} D(\mathcal{A}) = \sum_{u \in \mathcal{U}} \bar{D}_u(\mathcal{A}),$$

$$\begin{aligned} \text{约束条件: } & \sum_{f=1}^F \sum_{m=1}^M \mathbf{1}_{v_{f,m} \in \mathcal{A}} \cdot R_f(\lambda_m, Q_m) \leq R_{\max}, \\ & \sum_{f=1}^F \sum_{m=1}^M \mathbf{1}_{v_{f,m} \in \mathcal{A}} \cdot C_f(\lambda_m, Q_m) \leq C_{\max}, \end{aligned}$$

其中: 视频文件集合 $\mathcal{F} = \{1, 2, \dots, F\}$; 视频版本集合 $\mathcal{M} = \{1, 2, \dots, M\}$; 用户集合 $\mathcal{U} = \{1, 2, \dots, U\}$; F 、 M 、 U 为任意大于二的正整数; λ_m 表示视频版本 m 在编码时运动估计模块中的搜索范围, Q_m 表示视频版本 m 在编码时量化模中块的量化步长, $R_f(\lambda_m, Q_m)$ 表示视频文件 f 的第 m 个视频版本的编码码率, $C_f(\lambda_m, Q_m)$ 表示视频文件 f 的第 m 个视频版本的编码复杂度; 定义基集合 $\mathcal{V} = \{v_{f,m} | \forall f \in \mathcal{F}, \forall m \in \mathcal{M}\}$ 表示所有服务器处可编码得到的各视频不同版本组成的全集, 其中元素 $v_{f,m}$ 表示服务器编码输出视频文件 f 的第 m 个视频版本; 对于每个用户 $u \in \mathcal{U}$, 将其通过链路带宽 B_u 能够下载的所有视频的所有版本的集合记为: $\Omega_u = \{v_{f,m} \in \mathcal{V} | R_f(\lambda_m, Q_m) \leq B_u, \forall f \in \mathcal{F}, \forall m \in \mathcal{M}\}$;

优化变量为: $\mathcal{A}$ 表示实际编码输出的各视频版本组成的集合; 具体地, 某一个元素 $v_{f,m} \in \mathcal{A}$ 表示视频文件 f 的第 m 个视频版本由服务器编码输出;

优化目标为: 最大化所有用户的预期视频失真减少量之和 $D(\mathcal{A}) = \sum_{u \in \mathcal{U}} \bar{D}_u(\mathcal{A})$, 其中所有服务器处可编码得到的各原始视频不同版本集合 $\mathcal{V}$ 是所述优化问题的基集合; 当

用户请求视频 f 时, 将从集合 $\mathcal{A} \cap \Omega_u$ 中选择该视频码率最高的版本进行下载观看; 因此, $\bar{D}_u(\mathcal{A})$ 表示已知用户 u 请求视频文件 f 的概率为 $P_{u,f}$, 服务器实际编码输出的视频版本集合 $\mathcal{A}$ 能够造成用户 u 观看视频时的预期视频失真减少量, 即:

$$\bar{D}_u(\mathcal{A}) = \sum_{f=1}^F \sum_{m=1}^M \left[\prod_{j=1}^{m-1} (1 - \mathbf{1}_{v_{f,j} \in (\mathcal{A} \cap \Omega_u)}) \right] \cdot \mathbf{1}_{v_{f,m} \in (\mathcal{A} \cap \Omega_u)} \cdot P_{u,f} \cdot \left[D_{max} - D_f(\lambda_m, Q_m) \right]$$

上式中: 指示函数 $\mathbf{1}_{v_{f,j} \in (\mathcal{A} \cap \Omega_u)}$ 的值在 $v_{f,j} \in (\mathcal{A} \cap \Omega_u)$ 的时候为 1、在 $v_{f,j} \notin (\mathcal{A} \cap \Omega_u)$ 的时候为 0; D_{max} 分别表示当视频无法解码时的最大失真常量, $D_f(\lambda_m, Q_m)$ 表示视频文件 f 的第 m 个视频版本的编码失真, $D_{max} - D_f(\lambda_m, Q_m)$ 表示用户在成功解码视频文件 f 的第 m 个视频版本之后的视频失真减少量;

约束条件为: 服务器编码码率限制条件, 即要求所有视频版本的码率之和不超过服务器最大码率资源 R_{max} ; 以及服务器计算限制条件, 即要求所有视频版本的编码复杂度之和不超过服务器最大计算资源 C_{max} 。

优选地, 第二步中, 所述的最优版本集合码率控制与版本选择方法, 在为各原始视频确定最优版本的编码参数时, 采用具有多项式时间复杂度和高近似优化性能的高成本效益贪心算法, 最终快速高效地实现各原始视频最优版本的码率控制与版本选择。

更优选地, 所述的最优版本集合码率控制与版本选择方法的步骤为:

(1) 初始化: 设置初始局部最优解集合为空集、初始搜索集合为服务器处可编码得到的各视频不同版本组成的全集, 以及初始步数为 1;

(2) 迭代搜索步骤: 根据已有的局部最优解集合, 在剩余搜索集合中寻找使得边际增量与码率代价的比值与边际增量与复杂度代价的比值加权和最大的元素; 其中: 所述剩余搜索集合为局部最优解对于搜索集合的补集, 所述剩余搜索集合中的一个元素对应于某一视频编码得到的某一码率版本;

(3) 更新步骤: 如果添加(2)中搜索得到的元素仍能满足服务器处的编码码率和计算资源约束, 则将该元素添加至局部最优解集合, 并且搜索集合保持不变; 如果添加该元素不能满足服务器处的编码码率和计算资源约束, 则局部最优解集合保持不变, 并且将该元素从搜索集合中移除;

(4) 判定步骤: 如果剩余搜索集合不是空集, 则令搜索步数加一并返回(2)迭代搜索步骤; 否则, 停止迭代, 并将当前的局部最优解集合输出为最优结果。

优选地, 第三步中:

所述的动态自适应流媒体编码技术将任意一个原始视频文件编码为多个具有不同编码码率的视频版本;

所述的视频最优版本集合由服务器使用动态自适应流媒体编码技术并且采用第二步中得到的各视频各最优版本的编码参数编码获得。

根据本发明的另一个方面，提供一种动态自适应视频流媒体的编码码率控制与版本选择系统，所述系统包括：

视频模型抽取模块：用于抽取各原始视频的前若干帧，得到各原始视频的编码复杂度-码率-失真模型；

码率控制与版本选择模块：利用服务器处可编码得到的各视频不同版本组成的全集、视频模型抽取模块中得到的各原始视频的编码复杂度-码率-失真模型、服务器的编码码率和计算资源限制、以及用户的网络连接情况和视频点播概率分布，建立适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题，采用最优版本集合码率控制与版本选择方法，得到各视频各最优版本的编码参数；

视频版本编码模块：使用动态自适应流媒体编码技术，并采用码率控制与版本选择模块中得到的各视频各最优版本的编码参数，将各原始视频编码为多个不同码率的视频版本，输出得到视频最优版本集合。

与现有技术相比，本发明具有如下的有益效果：

本发明为适应动态自适应流媒体技术的需要，提供了一种在服务器编码码率和计算资源受限情况下高效快速确定最优版本码率控制与版本选择的方法，提高了服务器处编码视频内容的码率和计算资源的利用率，为用户提供更佳的视频服务质量。

附图说明

通过阅读参照以下附图对非限制性实施例所作的详细描述，本发明的其它特征、目的和优点将会变得更明显：

图 1 为本发明一实施例的系统原理及方法流程图；

图 2 为本发明一实施例编码参数选择方法的流程图；

图 3 为本发明一实施例视频编码复杂度-码率-失真性能的示意图；

图 4 为本发明一实施例最优视频版本码率控制与版本选择方法性能的示意图。

具体实施方式

下面结合具体实施例对本发明进行详细说明。以下实施例将有助于本领域的技术人员进一步理解本发明，但不以任何形式限制本发明。应当指出的是，对本领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干变形和改进。这些都属于

本发明的保护范围。

如图 1 所示, 一种用于实现动态自适应视频流媒体的码率控制与版本选择的系统, 包括设置于服务器内的视频模型抽取模块、码率控制与版本选择模块、视频版本编码模块, 其中:

视频模型抽取模块: 抽取各原始视频的前若干帧, 得到各原始视频的编码复杂度-码率-失真模型;

码率控制与版本选择模块: 利用服务器处可编码得到的各视频不同版本组成的全集、视频模型抽取模块中得到的各原始视频的编码复杂度-码率-失真模型、服务器的编码码率和计算资源限制、以及用户的网络连接情况和视频点播概率分布, 建立适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题, 采用最优版本集合码率控制与版本选择方法, 得到各视频各最优版本的编码参数;

视频版本编码模块: 使用动态自适应流媒体编码技术, 并采用码率控制与版本选择模块中得到的各视频各最优版本的编码参数, 将各原始视频编码为多个不同码率的视频版本, 输出得到视频最优版本集合。

如图 1 所示, 对动态自适应流媒体的内容分发网络进行实例分析。假设服务器处有 F 个原始视频文件 (F 为任意大于 2 的正整数), 记为视频文件集合 $\mathcal{F} = \{1, 2, \dots, F\}$ 。使用动态自适应流媒体编码技术, 将任意一个视频文件 $f \in \mathcal{F}$ 编码为 M 个不同码率的视频版本 (M 为任意大于 2 的正整数), 记为视频版本集合 $\mathcal{M}$ 。每一个视频版本 $m \in \mathcal{M}$ 对应于一对具体的编码参数 (λ_m, Q_m) , 其中: λ_m 表示视频版本 m 在编码时运动估计模块中的搜索范围, Q_m 表示视频版本 m 在编码时量化模块中的量化步长。将视频文件 f 的第 m 个视频版本的编码码率记为 $R_f(\lambda_m, Q_m)$, 并且将视频版本集合 $\mathcal{M}$ 按照编码码率呈降序排列, 即 $R_f(\lambda_i, Q_i) > R_f(\lambda_j, Q_j), \forall 1 \leq i < j \leq M$ 。从而, 可以定义基集合 $\mathcal{V} = \{v_{f,m} | \forall f \in \mathcal{F}, \forall m \in \mathcal{M}\}$ 表示所有服务器处可编码得到的各视频不同版本组成的全集, 其中元素 $v_{f,m}$ 表示服务器编码输出视频文件 f 的第 m 个视频版本。

网络中随机分布的 U 个用户, 记为用户集合 $\mathcal{U} = \{1, 2, \dots, U\}$; 用户可以向服务器提出视频点播请求, 并且根据其链路带宽 B_u 选择从服务器下载码率小于 B_u 的视频版本; 对于每个用户 $u \in \mathcal{U}$ 而言, 将其通过链路带宽 B_u 能够下载的所有视频的所有版本的集合记为 $\Omega_u = \{v_{f,m} \in \mathcal{V} | R_f(\lambda_m, Q_m) \leq B_u, \forall f \in \mathcal{F}, \forall m \in \mathcal{M}\}$ 。

基于上述, 一种适用于动态自适应视频流媒体的码率控制与版本选择方法, 具体实现包括如下步骤:

1、视频模型抽取

服务器抽取每个原始视频 $f \in \mathcal{F}$ 的前 K 帧 (K 为任意大于 2 的正整数), 通过编码和参数分析过程将视频版本 $m \in \mathcal{M}$ 的编码复杂度、编码码率和编码失真表示为编码参数对 (λ_m, Q_m) 的函数, 分别记为: $C_f(\lambda_m, Q_m)$ 、 $R_f(\lambda_m, Q_m)$ 和 $D_f(\lambda_m, Q_m)$ 。此步骤可以通过图 1 中所示视频模型抽取模块来实现。

2、码率控制与版本选择

此步骤可以通过图 1 中所示码率控制与版本选择模块来实现。

建立适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题如下 (其中每个参数的含义可在上下文中对应获取):

$$\text{目标优化问题: } \max_{\mathcal{A} \subseteq \mathcal{V}} D(\mathcal{A}) = \sum_{u \in \mathcal{U}} \bar{D}_u(\mathcal{A}),$$

$$\begin{aligned} \text{约束条件: } & \sum_{f=1}^F \sum_{m=1}^M \mathbf{1}_{|v_{f,m} \in \mathcal{A}} \cdot R_f(\lambda_m, Q_m) \leq R_{\max}, \\ & \sum_{f=1}^F \sum_{m=1}^M \mathbf{1}_{|v_{f,m} \in \mathcal{A}} \cdot C_f(\lambda_m, Q_m) \leq C_{\max}. \end{aligned}$$

其中:

优化变量为: $\mathcal{A}$ 表示实际编码输出的各视频版本组成的集合, 具体地, 某一个元素 $v_{f,m} \in \mathcal{A}$ 表示视频文件 f 的第 m 个视频版本由服务器编码输出;

优化目标为: 最大化所有用户的预期视频失真减少量之和 $D(\mathcal{A}) = \sum_{u \in \mathcal{U}} \bar{D}_u(\mathcal{A})$, 其中, 所有服务器处可编码得到的各视频不同版本集合 $\mathcal{V}$ 是所述优化问题的基集合。当用户请求视频 f 时, 将从集合 $\mathcal{A} \cap \Omega_u$ 中选择该视频码率最高的版本进行下载观看; 因此, $\bar{D}_u(\mathcal{A})$ 表示已知用户 u 请求视频文件 f 的概率为 $P_{u,f}$, 服务器实际编码输出的视频版本集合 $\mathcal{A}$ 能够造成用户 u 观看视频时的预期视频失真减少量, 即:

$$\bar{D}_u(\mathcal{A}) = \sum_{f=1}^F \sum_{m=1}^M \left[\prod_{j=1}^{m-1} (1 - \mathbf{1}_{|v_{f,j} \in (\mathcal{A} \cap \Omega_u)}) \right] \cdot \mathbf{1}_{|v_{f,m} \in (\mathcal{A} \cap \Omega_u)} \cdot P_{u,f} \cdot \left[D_{\max} - D_f(\lambda_m, Q_m) \right]$$

上式中: 指示函数 $\mathbf{1}_{|v_{f,j} \in (\mathcal{A} \cap \Omega_u)}$ 的值在 $v_{f,j} \in (\mathcal{A} \cap \Omega_u)$ 的时候为 1, 在 $v_{f,j} \notin (\mathcal{A} \cap \Omega_u)$ 的时候为 0; $D_{\max}$ 分别表示当视频无法解码时的最大失真常量, $D_{\max} - D_f(\lambda_m, Q_m)$ 表示用户在成功解码视频文件 f 的第 m 个视频版本之后的视频失真减少量;

约束条件为: 服务器编码码率限制条件, 即要求所有视频版本的码率之和不超过服务器最大码率资源 $R_{\max}$; 以及服务器计算限制条件, 即要求所有视频版本的编码复杂度之和不超过服务器最大计算资源 $C_{\max}$ 。

如图 2 所示, 给出具有多项式时间复杂度和高近似优化性能的高成本效益贪心算法, 最终快速高效地实现视频最优版本集合的码率控制与版本选择。

所述的最优版本集合的码率控制与版本选择方法的执行过程如下 (其中每个参数的含义可在上下文中对应获取):

(1) 初始化: 设置初始局部最优解集合 $\mathcal{A}^0 = \emptyset$, 初始搜索集合 $\mathcal{V}^0 = \mathcal{V}$, 初始步数 $t = 1$;

(2) 迭代搜索步骤 ($t = 1, 2, 3, \dots$):

根据已有的局部最优解集合 $\mathcal{A}^{t-1}$, 在剩余搜索集合 $\mathcal{V}^{t-1} \setminus \mathcal{A}^{t-1}$ (即 $\mathcal{A}^{t-1}$ 对于 $\mathcal{V}^{t-1}$ 的补集 $\mathcal{V}^{t-1} - \mathcal{A}^{t-1}$) 中寻找使得边际增量 $D(\mathcal{A}^{t-1} \cup \{v_{f,m}\}) - D(\mathcal{A}^{t-1})$ 与码率代价 $R_f(\lambda_m, Q_m)$ 和复杂度代价 $C_f(\lambda_m, Q_m)$ 的比值的加权和最大的元素 v_{f_t, m_t} , 即:

$$v_{f_t, m_t} = \arg \max_{v_{f,m} \in \mathcal{V}^{t-1} \setminus \mathcal{A}^{t-1}} \omega \cdot \frac{D(\mathcal{A}^{t-1} \cup \{v_{f,m}\}) - D(\mathcal{A}^{t-1})}{R_f(\lambda_m, Q_m)} + (1-\omega) \cdot \frac{D(\mathcal{A}^{t-1} \cup \{v_{f,m}\}) - D(\mathcal{A}^{t-1})}{C_f(\lambda_m, Q_m)}$$

上式中: 元素 $v_{f,m}$ 表示在服务器编码输出视频文件 f 的第 m 个视频版本, 元素 v_{f_t, m_t} 表示在第 t 个搜索步骤中搜索得到的边际增量与码率代价和复杂度代价的比值加权和最大的元素, 权重为 ω , $R_f(\lambda_m, Q_m)$ 表示视频文件 f 的第 m 个视频版本的编码码率, $C_f(\lambda_m, Q_m)$ 表示视频文件 f 的第 m 个视频版本的编码复杂度;

(3) 更新步骤:

如果添加元素 v_{f_t, m_t} 仍能满足服务器的编码码率和计算资源约束, 即:

$$\begin{aligned} \sum_{f=1}^F \sum_{m=1}^M \mathbf{1}_{v_{f,m} \in (\mathcal{A}^{t-1} \cup \{v_{f_t, m_t}\})} \cdot R_f(\lambda_m, Q_m) &\leq R_{max}, \\ \sum_{f=1}^F \sum_{m=1}^M \mathbf{1}_{v_{f,m} \in (\mathcal{A}^{t-1} \cup \{v_{f_t, m_t}\})} \cdot C_f(\lambda_m, Q_m) &\leq C_{max}, \end{aligned}$$

则将该元素 v_{f_t, m_t} 添加至局部最优解集合, 并且搜索集合保持不变, 即令 $\mathcal{A}^t = \mathcal{A}^{t-1} \cup \{v_{f_t, m_t}\}$ 以及 $\mathcal{V}^t = \mathcal{V}^{t-1}$;

如果添加该元素不能满足服务器的编码码率和计算资源约束, 则局部最优解集合保持不变, 并且将该元素从搜索集合中移除, 即令 $\mathcal{A}^t = \mathcal{A}^{t-1}$ 以及 $\mathcal{V}^t = \mathcal{V}^{t-1} \setminus \{v_{f_t, m_t}\}$;

(4) 判定步骤:

如果剩余搜索集合 $\mathcal{V}^t \setminus \mathcal{A}^t$ 不是空集, 则令 $t = t + 1$ 并返回 (2) 迭代搜索步骤; 否则, 停止迭代并且将当前的局部最优解集合 $\mathcal{A}^t$ 输出为最优结果 $\mathcal{A}^*$ 。

3、视频版本编码

此部分可以采用图 1 所示的视频版本编码模块来实现。

对应于第二步中求解得到的最优结果 $\mathcal{A}^*$ ，服务器进一步使用动态自适应流媒体编码技术将各原始视频编码为多个不同码率的视频版本，输出得到视频最优版本集合；

具体地：

$\forall v_{f,m} \in \mathcal{A}^*$ ，服务器将采用编码参数 (λ_m, Q_m) 对原始视频 f 进行实际编码操作，得到该视频 f 的第 m 个码率版本；

$\forall v_{f,m} \notin \mathcal{A}^*$ ，服务器将不会采用编码参数 (λ_m, Q_m) 对原始视频 f 进行实际编码操作。

如图 3 所示，服务器处三个具体视频文件（Crowd Run、Tractor 和 Sunflower）的编码复杂度-码率-失真性能曲面，其中三个视频的空间分辨率均为 1080p（1920×1080），编码帧率均为 30 帧每秒。

如图 4 所示，在服务器码率约束 $R_{max} = 30$ Mbps和计算资源约束 $C_{max} = 30$ GHz的设置下，所述的视频最优版本集合的码率控制与版本选择方法的视频失真性能随权重 ω 变化的情况，其中最大视频失真为 $D_{max} = 500$ 。

本发明为适应动态自适应视频流媒体多码率版本传输以及异构网络用户差异化带宽条件和点播需求的需要，建立了服务器编码码率和计算资源受限情况下的最优版本码率控制与版本选择问题，并且相应地提供了一种高效快速的最优版本集合的码率控制与版本选择方法，实现了用户下载观看视频整体质量的最大化。本发明提高了服务器处编码视频内容的码率和计算资源的利用率，为用户提供更佳的视频服务质量。

以上对本发明的具体实施例进行了描述。需要理解的是，本发明并不局限于上述特定实施方式，本领域技术人员可以在权利要求的范围内做出各种变形或修改，这并不影响本发明的实质内容。

权 利 要 求 书

1、一种动态自适应视频流媒体的编码码率控制与版本选择方法，其特征在于，包括以下步骤：

第一步、视频模型抽取步骤：

抽取各原始视频的前若干帧，得到各视频的编码复杂度-码率-失真模型；

第二步、码率控制与版本选择步骤：

根据服务器处可编码得到的各视频不同版本组成的全集、第一步中得到的各原始视频的编码复杂度-码率-失真模型、服务器的编码码率和计算资源限制、以及用户的网络连接情况和视频点播概率分布，建立适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题，采用最优版本集合码率控制与版本选择方法，得到各视频各最优版本的编码参数；

第三步、视频版本编码步骤：

使用动态自适应流媒体编码技术，并采用第二步中得到的各视频各最优版本的编码参数，将各原始视频编码为多个不同码率的视频版本，输出得到视频最优版本集合。

2、根据权利要求 1 所述的一种动态自适应视频流媒体的编码码率控制与版本选择方法，其特征在于，第一步中，所述的前若干帧为任意大于二的正整数帧。

3、根据权利要求 1 所述的一种动态自适应视频流媒体的编码码率控制与版本选择方法，其特征在于，第一步中，所述的各原始视频由于内容的不同呈现出不同的编码复杂度-码率-失真性能；

所述的编码复杂度-码率-失真模型将编码复杂度、码率和失真表示为具体编码参数的函数。

4、根据权利要求 1 所述的一种动态自适应视频流媒体的编码码率控制与版本选择方法，其特征在于，第二步中，所述的适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题，使用网络效用最大化建模方法得到。

5、根据权利要求 4 所述的一种动态自适应视频流媒体的编码码率控制与版本选择方法，其特征在于，所述的建立适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题，包括：

$$\text{目标优化问题：} \max_{\mathcal{A} \subseteq \mathcal{V}} D(\mathcal{A}) = \sum_{u \in \mathcal{U}} \bar{D}_u(\mathcal{A}),$$

$$\begin{aligned} \text{约束条件: } & \sum_{f=1}^F \sum_{m=1}^M \mathbf{1}_{v_m \in \mathcal{A}} \cdot R_f(\lambda_m, Q_m) \leq R_{max}, \\ & \sum_{f=1}^F \sum_{m=1}^M \mathbf{1}_{v_{f,m} \in \mathcal{A}} \cdot C_f(\lambda_m, Q_m) \leq C_{max}, \end{aligned}$$

其中：视频文件集合 $\mathcal{F} = \{1, 2, \dots, F\}$ ；视频版本集合 $\mathcal{M} = \{1, 2, \dots, M\}$ ；用户集合 $\mathcal{U} = \{1, 2, \dots, U\}$ ； F 、 M 、 U 为任意大于二的正整数； λ_m 表示视频版本 m 在编码时运动估计中的搜索范围， Q_m 表示视频版本 m 在编码时量化中的量化步长， $R_f(\lambda_m, Q_m)$ 表示视频文件 f 的第 m 个视频版本的编码码率， $C_f(\lambda_m, Q_m)$ 表示视频文件 f 的第 m 个视频版本的编码复杂度；定义基集合 $\mathcal{V} = \{v_{f,m} | \forall f \in \mathcal{F}, \forall m \in \mathcal{M}\}$ 表示所有服务器处可编码得到的各视频不同版本组成的全集，其中元素 $v_{f,m}$ 表示服务器编码输出视频文件 f 的第 m 个视频版本；对于每个用户 $u \in \mathcal{U}$ ，将其通过链路带宽 B_u 能够下载的所有视频的所有版本的集合记为 $\Omega_u = \{v_{f,m} \in \mathcal{V} | R_f(\lambda_m, Q_m) \leq B_u, \forall f \in \mathcal{F}, \forall m \in \mathcal{M}\}$ ；

优化变量为： $\mathcal{A}$ 表示实际编码输出的各视频版本组成的集合；具体地，某一个元素 $v_{f,m} \in \mathcal{A}$ 表示视频文件 f 的第 m 个视频版本由服务器编码输出；

优化目标为：最大化所有用户的预期视频失真减少量之和 $D(\mathcal{A}) = \sum_{u \in \mathcal{U}} \bar{D}_u(\mathcal{A})$ ，其中，所有服务器处可编码得到的各视频不同版本集合 $\mathcal{V}$ 是所述优化问题的基集合；当用户请求视频 f 时，将从集合 $\mathcal{A} \cap \Omega_u$ 中选择该视频码率最高的版本进行下载观看；因此， $\bar{D}_u(\mathcal{A})$ 表示已知用户 u 请求视频文件 f 的概率为 $P_{u,f}$ ，服务器实际编码输出的视频版本集合 $\mathcal{A}$ 能够造成用户 u 观看视频时的预期视频失真减少量，即：

$$\bar{D}_u(\mathcal{A}) = \sum_{f=1}^F \sum_{m=1}^M \left[\prod_{j=1}^{m-1} (1 - \mathbf{1}_{v_{f,j} \in (\mathcal{A} \cap \Omega_u)}) \right] \cdot \mathbf{1}_{v_{f,m} \in (\mathcal{A} \cap \Omega_u)} \cdot P_{u,f} \cdot \left[D_{max} - D_f(\lambda_m, Q_m) \right]$$

上式中：指示函数 $\mathbf{1}_{v_{f,j} \in (\mathcal{A} \cap \Omega_u)}$ 的值在 $v_{f,j} \in (\mathcal{A} \cap \Omega_u)$ 的时候为 1、在 $v_{f,j} \notin (\mathcal{A} \cap \Omega_u)$ 的时候为 0； D_{max} 分别表示当视频无法解码时的最大失真常量， $D_f(\lambda_m, Q_m)$ 表示视频文件 f 的第 m 个视频版本的编码失真， $D_{max} - D_f(\lambda_m, Q_m)$ 表示用户在成功解码视频文件 f 的第 m 个视频版本之后的视频失真减少量；

约束条件为：服务器编码码率限制条件，即要求所有视频版本的码率之和不超过服务器最大码率资源 R_{max} ；以及服务器计算限制条件，即要求所有视频版本的编码复杂度之和不超过服务器最大计算资源 C_{max} 。

6、根据权利要求 1 所述的一种动态自适应视频流媒体的编码码率控制与版本选择方法，其特征在于，第二步中，所述的最优版本集合码率控制与版本选择方法，是指：

在为各原始视频确定最优版本的编码参数时，采用具有多项式时间复杂度和高近似优化性能的高成本效益贪心算法，最终快速高效地实现各原始视频最优版本的码率控制与版本选择。

7、根据权利要求 6 所述的一种动态自适应视频流媒体的编码码率控制与版本选择方法，其特征在于，所述的最优版本集合码率控制与版本选择方法，步骤为：

(1) 初始化：设置初始局部最优解集合为空集、初始搜索集合为服务器处可编码得到的各视频不同版本组成的全集，以及初始步数为 1；

(2) 迭代搜索步骤：根据已有的局部最优解集合，在剩余搜索集合中寻找使得边际增量与码率代价的比值与边际增量与复杂度代价的比值加权和最大的元素；其中：所述剩余搜索集合为局部最优解对于搜索集合的补集，所述剩余搜索集合中的一个元素对应于某一视频编码得到的某一码率版本；

(3) 更新步骤：如果添加 (2) 中搜索得到的元素仍能满足服务器处的编码码率和计算资源约束，则将该元素添加至局部最优解集合，并且搜索集合保持不变；如果添加该元素不能满足服务器处的编码码率和计算资源约束，则局部最优解集合保持不变，并且将该元素从搜索集合中移除；

(4) 判定步骤：如果剩余搜索集合不是空集，则令搜索步数加一并返回 (2) 迭代搜索步骤；否则，停止迭代，并将当前的局部最优解集合输出为最优结果。

8、根据权利要求 1 所述的一种动态自适应视频流媒体的编码码率控制与版本选择方法，其特征在于，第三步中：

所述的动态自适应流媒体编码技术，将任意一个原始视频文件编码为多个具有不同码率的视频版本。

9、一种用于实现上述权利要求 1-8 任一项所述方法的动态自适应视频流媒体的编码码率控制与版本选择系统，其特征在于，所述系统包括：

视频模型抽取模块：用于抽取各原始视频的前若干帧，得到各原始视频的编码复杂度-码率-失真模型；

码率控制与版本选择模块：利用服务器处可编码得到的各视频不同版本组成的全集、视频模型抽取模块中得到的各原始视频的编码复杂度-码率-失真模型、服务器的编码码率和计算资源限制、以及用户的网络连接情况和视频点播概率分布，建立适用于动态自适应视频流媒体的编码码率控制与版本选择的优化问题，采用最优版本集合码率控制与版本选择方法，得到各视频各最优版本的编码参数；

视频版本编码模块：使用动态自适应流媒体编码技术，并采用码率控制与版本选择模块中得到的各视频各最优版本的编码参数，将各原始视频编码为多个不同码率的视频版本，输出得到视频最优版本集合。

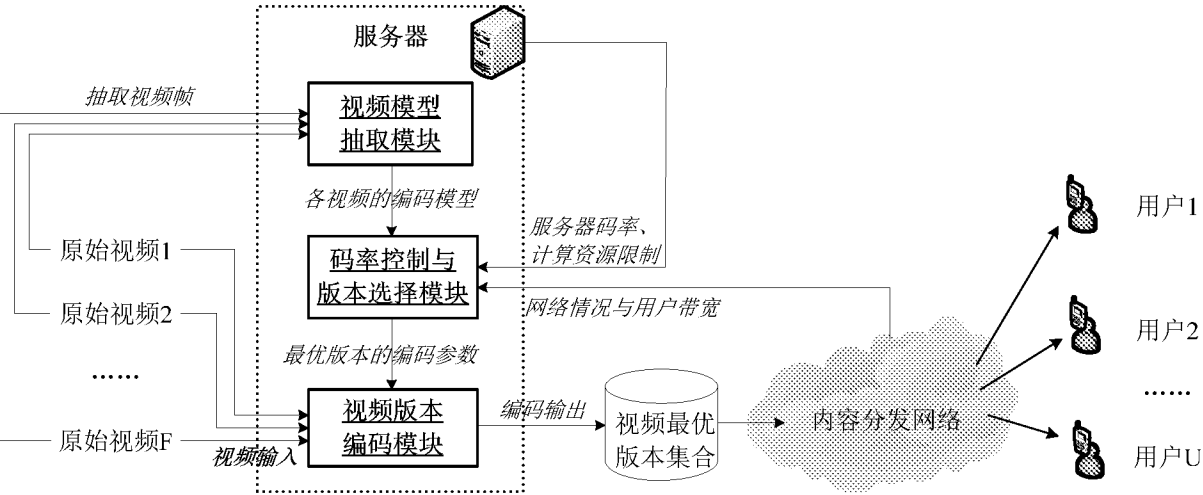


图 1

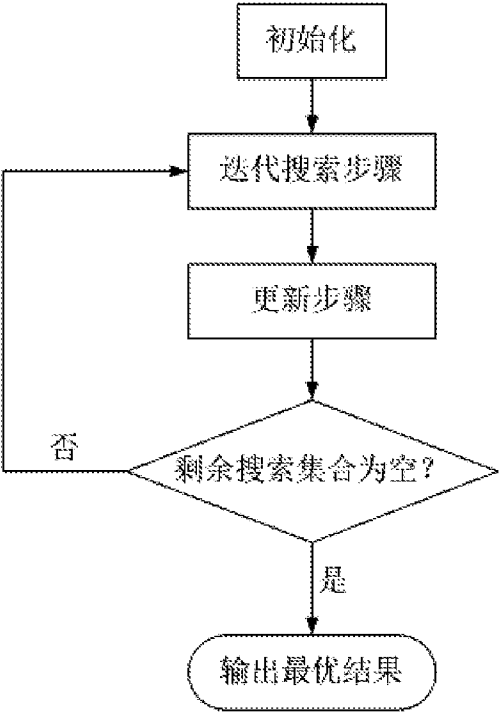


图 2

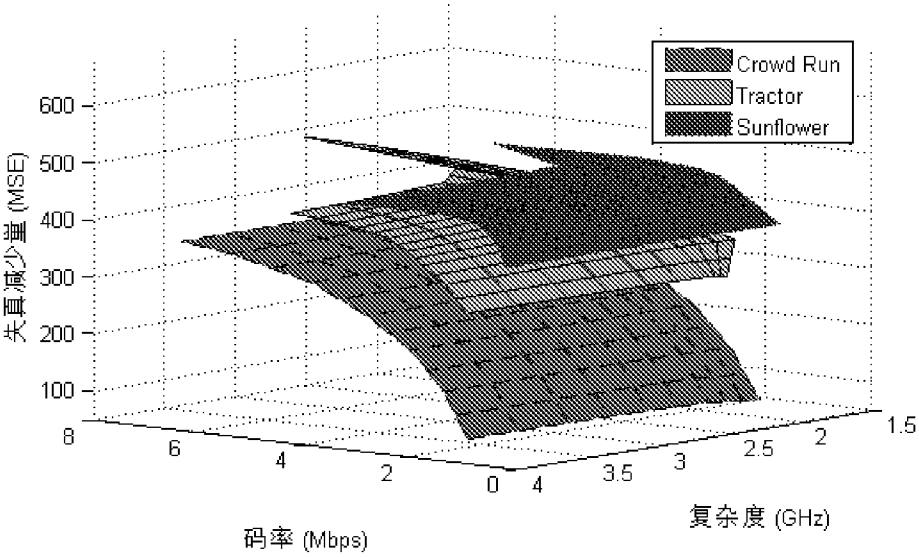


图 3

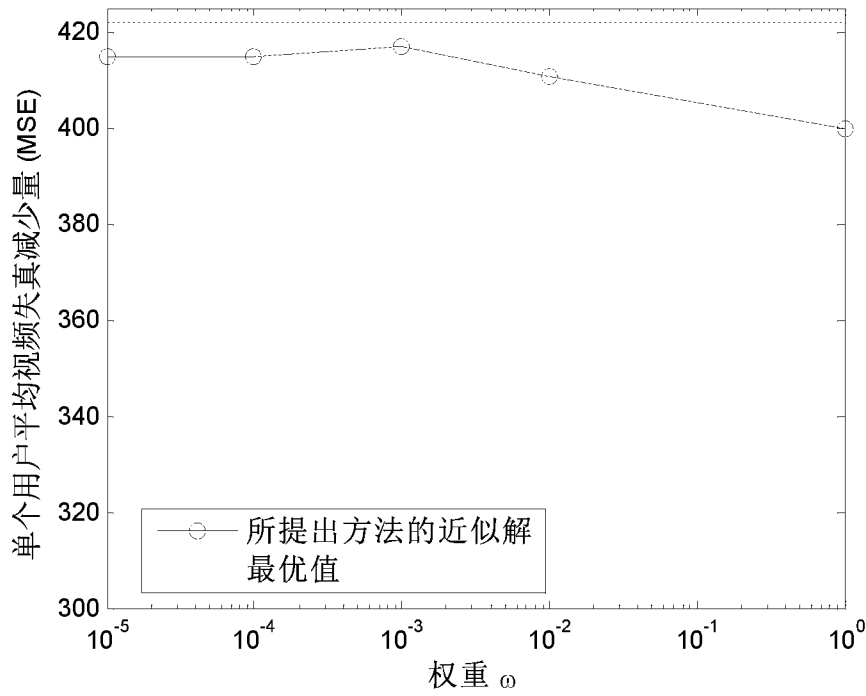


图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/106444

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/2662 (2011.01) i; H04N 21/24 (2011.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI: 自适应, 编码, 码率, 比特率, 版本, 服务器, 参数;

VEN, USTXT: adaptive, cod+, encod+, bit, rate, version?, representation?, server, parameter?

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 105979274 A (SHANGHAI JIAO TONG UNIVERSITY), 28 September 2016 (28.09.2016), description, paragraphs [0002], [0009]-[0024] and [0036]-[0064]	1-9
Y	TONI, L. et al., "Optimal Selection of Adaptive Streaming Representations", ACM Trans. Multimedia Comput. Commun. Appl., 28 February 2015 (28.02.2015), 2s(11), sections 1 and 3	1-9
A	CN 103702139 A (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY), 02 April 2014 (02.04.2014), entire document	1-9
A	US 2014223502 A1 (GEN. INSTRUMENT CORP.), 07 August 2014 (07.08.2014), entire document	1-9
PX	CN 106713956 A (SHANGHAI JIAO TONG UNIVERSITY), 24 May 2017 (24.05.2017), entire document	1-9

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 06 December 2017	Date of mailing of the international search report 04 January 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer YAO, Nan Telephone No. (86-10) 62089578

International application No.
PCT/CN2017/106444

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2017/106444

A. 主题的分类

H04N 21/2662(2011.01)i; H04N 21/24(2011.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

H04N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, CNTXT, CNKI: 自适应, 编码, 码率, 比特率, 版本, 服务器, 参数; VEN, USTXT:adaptive, cod+, encod+, bit, rate, version?, representation?, server, parameter?

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 105979274 A (上海交通大学) 2016年 9月 28日 (2016 - 09 - 28) 说明书第[0002], [0009]-[0024], [0036]-[0064]段	1-9
Y	L.Toni等. "Optimal Selection of Adaptive Streaming Representations" ACM Trans. Multimedia Comput. Commun. Appl., 第2s卷, 第11期, 2015年 2月 28日 (2015 - 02 - 28), 第1, 3节	1-9
A	CN 103702139 A (华中科技大学) 2014年 4月 2日 (2014 - 04 - 02) 全文	1-9
A	US 2014223502 A1 (GEN. INSTRUMENT CORP.) 2014年 8月 7日 (2014 - 08 - 07) 全文	1-9
PX	CN 106713956 A (上海交通大学) 2017年 5月 24日 (2017 - 05 - 24) 全文	1-9

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2017年 12月 6日

国际检索报告邮寄日期

2018年 1月 4日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

姚楠

传真号 (86-10)62019451

电话号码 (86-10)62089578

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/106444

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	105979274	A	2016年 9月 28日	无			
CN	103702139	A	2014年 4月 2日	CN	103702139	B	2017年 2月 1日
US	2014223502	A1	2014年 8月 7日	WO	2014124058	A1	2014年 8月 14日
CN	106713956	A	2017年 5月 24日	无			