(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
2024 年 8 月 29 日 (29.08.2024)

WIPO | PCT

(10) 国际公布号
WO 2024/174679 A1

(51) 国际专利分类号:
H04N 19/147 (2014.01) *H04N 19/62* (2014.01)

(21) 国际申请号: PCT/CN2023/137731

(22) 国际申请日: 2023 年 12 月 9 日 (09.12.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310183247.0 2023年2月22日 (22.02.2023) CN

(71) 申请人: 中国科学院深圳先进技术研究院 (**SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY CHINESE ACADEMY OF SCIENCES**) [CN/CN]; 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(72) 发明人: 丁克勤 (**DING, Keqin**); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。张云 (**ZHANG, Yun**); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。

(74) 代理人: 北京中巡通大知识产权代理有限公司 (**BEIJING ZHONG XUN TONG DA INTELLECTUAL PROPERTY AGENCY CO., LTD.**); 中国北京市西城区西直门外大街 1 号院 2 号楼7层7C1, Beijing 100044 (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,

(54) **Title:** POINT-CLOUD RATE-DISTORTION CODING METHOD BASED ON PERCEPTION WEIGHTING, AND RELATED APPARATUS

(54) 发明名称: 一种基于感知加权的点云率失真编码方法及相关装置

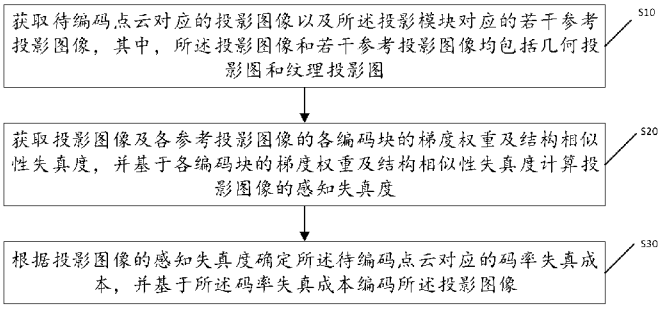


图 1

S10 Acquire projection images corresponding to a point cloud to be subjected to coding, and several reference projection images corresponding to a projection module, wherein the projection images and the several reference projection images each comprise a geometric projection image and a texture projection image

S20 Acquire gradient weights and structural-similarity distortion degrees of coding blocks of the projection images and the reference projection images, and calculate perceptual distortion degrees of the projection images on the basis of the gradient weights and structural-similarity distortion degrees of the coding blocks

S30 According to the perceptual distortion degrees of the projection images, determine a rate distortion cost corresponding to said point cloud, and encode the projection images on the basis of the rate distortion cost

(57) **Abstract:** Disclosed in the present application are a point-cloud rate-distortion coding method based on perceptual weighting, and a related apparatus. The method comprises: acquiring projection images corresponding to a point cloud to be subjected to coding, and several reference projection images corresponding to a projection module; acquiring gradient weights and structural-similarity distortion degrees of coding blocks of the projection images and the reference projection images, and calculating perceptual distortion degrees of the projection images on the basis of the gradient weights and structural-similarity distortion degrees of the coding blocks;

WO 2024/174679 A1

[见续页]

CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, 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))。

and according to the perceptual distortion degrees of the projection images, determining a rate distortion cost corresponding to said point cloud, and encoding the projection images on the basis of the rate distortion cost. In the present application, two-dimensional projection images are generated by means of projection, a perceptual distortion degree corresponding to each projection image is then determined by taking a coding block as a unit, and encoding decision-making is performed on the basis of a perceptual-loss degree, such that matching between point-cloud coding and visual perception quality is improved, thereby improving the coding efficiency.

(57) 摘要: 本申请公开了一种基于感知加权的点云率失真编码方法及相关装置，所述方法包括获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像；获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度；根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。本申请通过投影生成二维的投影图像，然后计算以编码块为单位确定每个投影图像对应的感知失真度，并基于感知损失度进行编码决策，提高了点云编码与视觉感知质量的匹配性，从而提高了编码效率。

一种基于感知加权的点云率失真编码方法及相关装置

技术领域

本申请涉及图像处理技术领域，特别涉及一种基于感知加权的点云率失真编码方法及相关装置。

背景技术

近年来，虚拟现实（VR）、增强现实（AR）和混合现实（MR）等三维因能够为用户提供独特的六自由度（6DoF）交互、逼真和沉浸式3D视觉体验，而在许多应用中开始流行，如3D电影观看、遗产保护、导航、沉浸式远程通话和远程手术等。其中，动态点云（DPC）因其所具有的逼真的表现能力，已经成为新兴的沉浸式VR、AR和MR媒体的主流表现形式之一。

动态点云（DPC）用大量非结构化的高维点来表示三维场景，其中，DPC是一系列时间上连续的点云，其不但可以反映运动和时间上的光线变化，而且其中的每个点都均包括用于识别三维空间中的位置的几何成分以及用于反映光线和物体属性的光度信息，如RGB颜色、反射和透明度等。然而，动态点云由于包括大量高维点也产生了庞大的数据量，这需要巨大的存储空间和庞大的网络带宽来传输。

为了解决这个问题，在对动态点云进行存储和传输之前，需要对动态点云进行压缩以减少动态点云所需的存储空间以及传输所需的带宽。目前动态点云普遍应用的压缩技术为视频点云压缩(V-PCC)模型,V-PCC模型是用原始点云和失真的点云之间的信号差来衡量失真，如，点对点（D1）、点对面（D2）和点对网（P2mesh）。然而，现有的V-PCC模型均未考虑动态点云的视觉感知质量，无法利用动态点云的感知冗余，进而影响编码效率。

因而现有技术还有待改进和提高。

技术问题

本申请要解决的技术问题在于，针对现有技术的不足，提供一种

基于感知加权的点云率失真编码方法及相关装置。

技术解决方案

为了解决上述技术问题，本申请实施例第一方面提供了一种基于感知加权的点云率失真编码方法，所述方法包括：

获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像，其中，所述投影图像和若干参考投影图像均包括几何投影图和纹理投影图；

获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度；

根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。

所述基于感知加权的点云率失真编码方法，其中，所述参考投影图像为通过对投影图像进行下采样得到的，并且投影图像和若干参考投影图像的图像尺度互不相同。

所述基于感知加权的点云率失真编码方法，其中，所述梯度权重的获取过程具体包括：

计算所述编码块中的各像素点的梯度值，并基于各像素点的梯度值计算所述编码块的梯度权重。

所述基于感知加权的点云率失真编码方法，其中，所述根据投影图像的感知失真度确定所述待编码点云对应的率失真成本具体包括：

将所述感知失真度中的结构相似性失真度转换为均方误差失真度，以得到所述投影图像对应的感知系数；

基于所述感知系数确定所述投影图像对应的感知拉格朗日乘数；

基于所述感知失真度和所述感知拉格朗日乘数确定所述待编码点云对应的率失真成本。

所述基于感知加权的点云率失真编码方法，其中，所述结构相似

性失真度与所述均方误差失真度的对应关系为：

$$D_{SIMM_{k,i}} = \Phi_{k,i} D_{MSE_{k,i}}$$

$$\Phi_{k,i} = \frac{\sum_{n-1}^{N} \left(\rho_n \frac{1}{16} \sum_{i \in blk_n} E_i \right)}{\sum_{n=1}^{N} \rho_n}$$

其中，$D_{SIMM_{k,i}}$表示第k个目标投影图像的编码块i的结构相似性失真度，$D_{MSE_{k,i}}$表示第k个目标投影图像的编码块i的均方误差失真度，$\Phi_{k,i}$表示第k个目标投影图像的编码块i的感知系数，blk_n表示编码块i的第n个子图像块，ρ_n表示编码块i的第n个子图像块的线性模型参数，N表示子图像块的数量，E_i表示像素内容权重，第k个目标投影图像为由投影图像和若干参考投影图像形成的投影图像集中的一投影图像。

所述基于感知加权的点云率失真编码方法，其中，当感知拉格朗日乘数用于编码时，所述基于所述感知系数确定所述投影图像对应的感知拉格朗日乘数具体包括：

基于所述投影图像的每个编码块的感知系数计算第一目标感知系数，并计算所述第一目标感知系数与所述编码块的感知系数的比值，以得到第一感知系数比值；

计算所述第一感知系数比值与所述投影图像的均方误差拉格朗日乘数的乘积，以得到所述编码块对应的感知拉格朗日乘数。

所述基于感知加权的点云率失真编码方法，其中，当感知拉格朗日乘数用于模式决策时，所述基于所述感知系数确定所述投影图像对应的感知拉格朗日乘数具体包括：

基于所述投影图像的每个编码块的感知系数计算第二目标感知系数，并计算第二目标感知系数与编码块的感知系数的平方根的比值，以得到第二感知系数比值；

计算所述第二感知系数比值与所述投影图像的所述均方误差拉格朗日乘数的平方根的乘积，以得到所述编码块对应的感知拉格朗日

乘数。

本申请实施例第二方面提供了一种基于感知加权的点云率失真编码系统，所述系统包括：

获取模块，用于获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像，其中，所述投影图像和若干参考投影图像均包括几何投影图和纹理投影图；

计算模块，用于获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度；

编码模块，用于根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。

本申请实施例第三方面提供了一种计算机可读存储介质，所述计算机可读存储介质存储有一个或者多个程序，所述一个或者多个程序可被一个或者多个处理器执行，以实现如上任一所述的基于感知加权的点云率失真编码方法中的步骤。

本申请实施例第四方面提供了一种终端设备，其包括：处理器、存储器及通信总线；所述存储器上存储有可被所述处理器执行的计算机可读程序；

所述通信总线实现处理器和存储器之间的连接通信；

所述处理器执行所述计算机可读程序时实现如上任一所述的基于感知加权的点云率失真编码方法中的步骤。

有益效果

与现有技术相比，本申请提供了一种基于感知加权的点云率失真编码方法及相关装置，方法包括获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像；获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度

权重及结构相似性失真度计算投影图像的感知失真度；根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。本申请通过投影生成二维的投影图像，然后计算以编码块为单位确定每个投影图像对应的感知失真度，并基于感知损失度进行编码决策，提高了点云编码与视觉感知质量的匹配性，从而提高了编码效率。

附图说明

为了更清楚地说明本申请实施例中的技术方案，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，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员而言，在不符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 1 为本申请提供的基于感知加权的点云率失真编码方法的流程图。

图 2 为本申请提供的基于感知加权的点云率失真编码方法的原理流程图。

图 3 为近似模型与原模型的纹理投影图上的相关性的散点图。

图 4 为近似模型与原模型的几何投影图上的相关性的散点图。

图 5 为本申请提供的基于感知加权的点云率失真编码方法的结构原理图。

图 6 为本申请提供的终端设备的结构原理图。

本发明的最佳实施方式

本申请提供一种基于感知加权的点云率失真编码方法及相关装置，为使本申请的目的、技术方案及效果更加清楚、明确，以下参照附图并举实施例对本申请进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。

本技术领域技术人员可以理解，除非特意声明，这里使用的单数

形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是，本申请的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和/或组件，但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。应该理解，当我们称元件被“连接”或“耦接”到另一元件时，它可以直接连接或耦接到其他元件，或者也可以存在中间元件。此外，这里使用的“连接”或“耦接”可以包括无线连接或无线耦接。这里使用的措辞“和/或”包括一个或更多个相关联的列出项的全部或任一单元和全部组合。

本技术领域技术人员可以理解，除非另外定义，这里使用的所有术语(包括技术术语和科学术语)，具有与本申请所属领域中的普通技术人员的一般理解相同的意义。还应该理解的是，诸如通用字典中定义的那些术语，应该被理解为具有与现有技术的上下文中的意义一致的意义，并且除非像这里一样被特定定义，否则不会用理想化或过于正式的含义来解释。

应理解，本实施例中各步骤的序号和大小并不意味着执行顺序的先后，各过程的执行顺序以其功能和内在逻辑确定，而不应对本申请实施例的实施过程构成任何限定。

经过研究发现，近年来，虚拟现实（VR）、增强现实（AR）和混合现实(MR)等三维因能够为用户提供独特的六自由度(6DoF)交互、逼真和沉浸式3D视觉体验，而在许多应用中开始流行，如3D电影观看、遗产保护、导航、沉浸式远程通话和远程手术等。其中，动态点云(DPC)因其所具有的逼真的表现能力，已经成为新兴的沉浸式VR、AR和MR媒体的主流表现形式之一。

动态点云（DPC）用大量非结构化的高维点来表示三维场景，其中，DPC是一系列时间上连续的点云，其不但可以反映运动和时间上的光线变化，而且其中的每个点都均包括用于识别三维空间中的位置

的几何成分以及用于反映光线和物体属性的光度信息，如 RGB 颜色、反射和透明度等。然而，动态点云由于包括大量高维点也产生了庞大的数据量，这需要巨大的存储空间和庞大的网络带宽来传输。

为了解决这个问题，在对动态点云进行存储和传输之前，需要对动态点云进行压缩以减少动态点云所需的存储空间以及传输所需的带宽。目前动态点云普遍应用的压缩技术为视频点云压缩(V-PCC)模型,V-PCC模型是用原始点云和失真的点云之间的信号差来衡量失真，如，点对点（D1)、点对面（D2）和点对网（P2mesh)。然而，现有的V-PCC 模型均未考虑动态点云的视觉感知质量，无法利用动态点云的感知冗余，进而影响编码效率。

为此，将视觉感知与编码相结合一直是视频和图像处理领域的热门课题。在已有的研究工作当中，Jiang 等人提出了一种新的符合 H.265/HEVC 标准的感知编码方案，通过使用时空显著性。Wu 等人提出了感知加权平均平方误差（PWMSE）的码率失真（RD）模型，并根据等效失真得出了码率失真优化（RDO）过程的拉格朗日乘数。基于 WS-PSNR，Li 等人用球面域的 360 度视频评价指标优化了 RDO 任务。Li 等人提出了一种针对球域视频的 RDO 方法，以解决二维图像平面畸变计算不准确的问题。上述方法的基本思路是利用 HVS 模型或其近似值作为编码模块中的失真，并以最小化感知失真为约束条件，选择最佳的编码模式或参数。然而，上述方法是针对传统的 2D 或 360 度视频编码提出的，无法用于动态点云。

为了利用点云中的视觉冗余，Li 等人采用了一种基于占用图的 RDO 方法，该方法设计了一个占用图指导的视频压缩框架，利用占用图来提升 DPC 压缩性能。由于目前的 RDO 和几何失真与评估标准 D1 和 D2 不一致，Xiong 等人提出了一种基于 EPM 的 RDO 方法，该方法首先描述了现有失真模型和几何质量测量之间的关系，然后通过估计 CU 的法向量来估计 D1 和 D2 的法向量，从而修改三维几何距离。然

而，这些率失真编码方式中重建的点云质量仍然是用 D1 和 D2 来衡量的，不能真实反映人类的感知。因此，现有的点云编码方案中，没有充分考虑到点云的感知特性，应然没有利用点云的感知冗余。

为了解决上述问题，在本申请实施例中，获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像；获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度；根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。本申请通过投影生成二维的投影图像，然后计算以编码块为单位确定每个投影图像对应的感知失真度，并基于感知损失度进行编码决策，提高了点云编码与视觉感知质量的匹配性，从而提高了编码效率。

本发明的实施方式

下面结合附图，通过对实施例的描述，对申请内容作进一步说明。

本实施例提供了一种基于感知加权的点云率失真编码方法，如图 1 和图 2 所示，所述方法包括：

S10、获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像。

具体地，所述投影图像和若干参考投影图像均包括几何投影图和纹理投影图，其中，投影图像为对待编码点云进行将待编码点云进行补片分割和打包所形成的，几何投影图和纹理投影均为二维图像，并且几何投影图和纹理投影图的图像尺度相同。若干参考投影图像中的每个参考投影图像均为一个图像组，图像组包括参考几何投影图和参考纹理投影图，其中，各参考投影图像中的各几何投影图的图像尺度互不相同，且均不同等于几何投影图的图像尺度；各参考投影图像中的各纹理投影图的图像尺度互不相同，且均不同等于纹理投影图的图像尺度。

参考投影图像为通过对投影图像进行下采样得到的，其中，参考投影图像为通过对投影图像进行下采样得到的指的是参考几何投影图为对几何投影图进行下采样得到，考纹理投影图为对几何投影图进行下采样得到。在一个实现方式中，若干参考投影图像可以采用高斯金字塔进行下采样操作，即将投影图像包括的几何投影图和纹理投影图输入高斯金字塔，获取高斯金字塔各层的输出项，以得到若干参考投影图。可以理解的是，通过高斯金字塔的每一层，将该层输出项的分辨率降低为输入项的二分之一，例如，用k来表示不同的尺度，$k=1$表示投影图像对应的图像尺度，$k=2$表示对投影图像进行下采样一次，下采样得到的参考投影图像对应的图像尺度为投影图像对应的图像尺度的二分之一，以此类推，以得到若干投影图像。

当然，在实际应用中，在通过投影生成投影图像后，若干参考投影图像的获取方式还可以采用其他方式，例如，将若干投影图像按照图像尺度从大到小排序形成第一参考投影图像、第二参考投影图像，、、、，第 N 参考投影图像，其中，第一参考投影图像通过对投影图像进行 2 倍下采样得到的，第二参考投影图像通过对投影图像进行 4 倍下采样得到的，依次类推，得到第 N 参考投影图像。此外，各第 N 参考投影图像的图像尺度的对应关系也可以不同，例如，投影图像对应的图像尺度为第一参考投影图像的图像尺度 3 倍，第一参考投影图像对应的图像尺度为第二参考投影图像对应的图像尺度的 3 倍等。

S20、获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度。

具体地，梯度权重为基于编码块中的各像素点的梯度确定，结构相似性失真度用于反映编码块对应的失真块的失真度，其中，失真块为基于待编码点云对应的重构点云所形成的几何重构图和纹理重构图确定的。也就是说，在获取投影图像和若干参考投影图像时，会同

步获取待编码点云对应的重构点云的重构投影图像，重构投影图像包括几何重构图和纹理重构图；然后再对重构投影图像进行下采样操作以得若干参考重构投影图像。其中，重构投影图像和若干参考重构投影图像的确定过程与投影图像和若干参考投影图像的确定过程相同，这里就不具体说明，仅是说明投影图像与重构投影图像相对应，若干参考投影图像与若干参考重构投影图像一一对应，并且参考投影图像的图像尺度与其对应的参考重构投影图像的图像尺度相同，即参考投影图像中的参考几何投影图的图像尺度等于其对应的参考重构投影图像中的参考几何重构图的图像尺度，参考投影图像中的参考纹理投影图的图像尺度等于其对应的参考重构投影图像中的参考纹理重构图的图像尺度。

在一个实现方式中，所述梯度权重的获取过程具体包括：

计算所述编码块中的各像素点的梯度值，并基于各像素点的梯度值计算所述编码块的梯度权重。

具体地，编码块中的各像素点的梯度值包括水平方向梯度和竖直方向梯度，水平方向梯度和竖直方向梯度的计算公式分别为：

$$G_{i,j}^{H} = F_H * I_i,\ G_{i,j}^{V} = F_V * I_i$$

$$F_H = \frac{1}{16}\begin{bmatrix} 3 & 0 & -3 \\ 10 & 0 & -10 \\ 3 & 0 & -3 \end{bmatrix}, F_V = \frac{1}{16}\begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix}$$

其中，F_H表示水平方向的梯度算子，I_i表示第i个编码块，尺寸大小为M×N，j代表编码块中的第j个像素点。

在得到各像素点的像素梯度值后，基于编码块中的各像素点的像素梯度值计算编码块的梯度值，其中，编码块的候选梯度值的计算公式可以为：

$$\overline{G_i} = \frac{\sum_j \sqrt{G_{i,j}^{H\,2} + G_{i,j}^{V\,2}}}{MN}$$

其中，$\overline{G_i}$表示候选梯度值。

在各编码块的候选梯度值后，对各候选梯度值进行归一化操作，以得到各编码块的梯度值，编码块的梯度值用于反映该编码块相对于所有编码块的重要程度。

其中，梯度值的计算公式为：

$$W_i = \mathrm{Nomal}(\bar{G}_i)$$

其中，W_i表示编码块的梯度值，$Nomal(\cdot)$表示归一化操作，其具体计算方式为：

$$Nomal(x_i) = C\left(\frac{x_i - \mu_x}{6\sigma_x} + \frac{1}{2}\right)$$

其中，$C(\cdot)$操作表示将计算得到的最大值和最小值约束在$[0,1]$之间，即小于 0 的值置 0，大于 1 的数置 1，μ_x表示x_i的均值，σ_x表示x_i的标准差。

进一步，结构相似性失真度为编码块与其对应的重构编码块件的结构相似性失真度。在获取到各编码块的梯度权重及结构相似性失真度，可以基于各编码块的梯度权重及结构相似性失真度确定投影图像对应的感知失真度。其中，将由投影图像和若干参考投影图像形成的投影图像集，该投影图像集中的第k个目标投影图像中第i个纹理失真图像块的纹理图像块失真度可以表示为：

$$D_{PPCM_{k,i}} = W_{k,i} D_{SSIM_{k,i}}$$

其中，$W_{k,i}$表示第k组图像组中第i个编码块的梯度权重，$D_{SSIM_{k,i}}$表示第k组图像组中第i个编码块相对于第i个重构编码块的SSIM失真度。

当然，值得说明的，在计算编码块的感知失真度时，投影图像中的几何投影图和纹理投影图分别计算各自对应的感知失真度，并且两者的计算过程相同，从而本实施例以投影图像对感知失真度的确定过程进行说明。

S30、根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。

具体地，率失真成本的目标函数为：

$$\underset{\{P\}}{\operatorname{argmin}} J\ ,\ J = D + \lambda \mathrm{R}$$

其中，J表示总的码率失真（RD）成本，D表示待编码点云和待编码单元对应的重构点云之间的失真差异，R表示编码比特。λ表示用于平衡失真和比特率的拉格朗日乘数。

基于此，根据投影图像的感知失真度确定所述待编码点云对应的率失真成本的目标函数可以表示为：

$$\underset{\{P\}}{\operatorname{argmin}} J\ ,\ J = D_{PPCM} + \lambda R$$

其中，D_{PPCM}表示感知失真度。

由于 V-PCC 标准利用 HEVC 对几何视频和纹理视频分别进行编码，从而将几何视频和纹理视频的压缩看作两个独立的过程，并且纹理视频的失真只与纹理视频编码参数有关，几何视频的失真只与几何视频编码参数有关。由此，纹理投影图像的率失真成本的目标函数和几何投影图像的率失真成本的目标函数分别表示为：

$$\underset{\{P_G\}}{\operatorname{argmin}} J_G\ ,\ J_G = D_G + \lambda_G R_G$$

$$\underset{\{P_T\}}{\operatorname{argmin}} J_T\ ,\ J_T = D_T + \lambda_T R_T$$

其中，J_G表示纹理投影图像的独立约束，J_T表示几何投影图像的的独立约束，$\mathrm{R_T}$对于几何视频编码器来说是一个常数，$\mathrm{R_G}$对于纹理视频编码器。

此外，由于纹理投影图像的编码决策过程和几何投影图像的编码决策过程相同，这里仅对编码决策过程进行说明，不做纹理投影图像和几何投影图像说明，纹理投影图像和几何投影图像进步可以采用下述编码决策过程。

在一个实现方式中，所述根据投影图像的感知失真度确定所述待编码点云对应的码率失真成本具体包括：

将所述感知失真度中的结构相似性失真度转换为均方误差失真度，以得到所述投影图像对应的感知系数；

基于所述感知系数确定所述投影图像对应的感知拉格朗日乘数；

基于所述感知失真度和所述感知拉格朗日乘数确定所述待编码点云对应的码率失真成本。

具体地，在高分辨率量化近似下，编码过程将尽可能多地保留亮度信息。此外，对于不同的视频内容，即使在高 QP 设置下，μ_{x_i}和μ之间的皮尔逊相关系数(PCC)值也超过 0.99。基于此，可以将$i^{t\square}$像素的 SSIM 值重写为：

$$D_{SSIM} = 1 - \frac{1}{M}\sum_{i\in Unit} SSIM_i = \frac{1}{M}\sum_{i\in CTU} d_i^2\, E_i$$

$$E_i = \sum_{t=1}^{T} \frac{\varepsilon_t}{\sigma_{x_{i-t}}^2 + \sigma_{y_{i-t}}^2 + C_2}$$

其中，$d_i^2 = (x_i - y_i)^2$表示像素级平方误差，E_i表示像素内容权重，M表示编码块的数量，C_2是 SSIM 中的C_2常数，ε_t表示当前编码块第 t 个位置的滤波系数，$\sigma^2_{x_{i-t}}$和$\sigma^2_{y_{i-t}}$表示编码块的编码前后在 i-t 处的方差。

进一步，为了减少计算量，编码块的 SSIM 失真度设置为 16 个子块的平均值，n表示编码块的子图像块，即：

$$D_{SSIM} \approx \frac{1}{M}\sum_{n=1}^{N}\left(\sum_{i\in blk_n} d_i^2 \frac{1}{16}\sum_{i\in blk_n} E_i\right)$$

已知 HEVC 中的编码单元存在以下关系

$$\sum_{i\in blk_n} d_i^2 \approx \rho_n Q_n$$

其中，ρ_n是与图像内容有关的线性模型参数，其中，前一帧的编码同位块的结果作为参考图像内容（第一帧除外），Q_n是子块的量化步长，通常以相同的数值应用于所有子块。

对于编码块来说，D_{MSE}可以被表示为：

$$D_{MSE} = \frac{1}{M}\sum_{n=1}^{N}\rho_n Q_n$$

$$D_{SSIM} = \frac{\sum_{n=1}^{N}\left(\rho_n \frac{1}{16}\sum_{i\in blk_n} E_i\right)}{\sum_{n=1}^{N}\rho_n} D_{MSE} = \Phi_{k,i} D_{MSE_{k,i}}$$

$$\Phi_{k,i} = \frac{\sum_{n-1}^{N}\left(\rho_n \frac{1}{16}\sum_{i\in blk_n} E_i\right)}{\sum_{n=1}^{N}\rho_n}$$

其中，$D_{SIMM_{k,i}}$表示第k个投影图像的编码块i的结构相似性失真度，$D_{MSE_{k,i}}$表示第k个投影图像的编码块i的均方误差失真度，$\Phi_{k,i}$表示第k个投影图像的编码块i的感知系数，blk_n表示编码块i的第n个子图像块，ρ_n表示编码块i的第n个子图像块的线性模型参数，N表示子图像块的数量，E_i表示像素内容权重，第k个投影图像为由投影图像和若干参考投影图像形成的投影图像集中的一投影图像。

基于此，视频压缩导致的图像 MSE 数值损失经过适当的变换可以近似为领域图像的 SSIM 损失，即D_{PPCM}可以表示为：

$$D_{PPCM_{k,i}} \approx \widetilde{D}_{PPCM_{k,i}} = W_{k,i}\Phi_{k,i} D_{MSE_{k,i}} \quad (1)。$$

进一步，在获取到感知系数确定所述投影图像对应的感知拉格朗日乘数，其中，感知拉格朗日乘数的确定过程可以为：

当把基于 PPCM 的失真度量D_{PPCM}应用于 V-PCC 中的 RD 目标函数时，需要对拉格朗日乘数λ_{MODE}^{ϕ}和λ_{ME}^{ϕ}进行调整，以便在通道ϕ（其中，通道包括纹理通道和几何通道，这里以ϕ作为表示）的感知失真D_{PPCM}^{ϕ}和比特率R^{ϕ}之间进行最佳折中。基于 MSE 的率失真成本J_{MSE_i}在每一个编码块i的计算为：

$$J_{MSE_i} = D_{MSE_i} + \lambda_{MSE} R_{MSE_i}$$

其中，D_{MSE_i}、R_{MSE_i}和λ_{MSE}是 HEVC 中的基于 MSE 的失真、比特率

和拉格朗日乘数，码率R_{MSE_i}和失真D_{MSE_i}之间的关系可以建模为

$$R_{MSE_i} = \alpha log\left(\frac{\sigma_i^2}{D_{MSE_i}}\right)$$

其中，σ_i^2是在帧间或帧内预测之后的编码残差的方差，α是比例常数。将J_{MSE_i}对D_{MSE_i}求偏导数，并在偏导数为零时可以得到：

$$\frac{\partial J_{MSE}}{\partial D_{MSE_i}} = 1 + \lambda_{MSE}\frac{\partial R_{MSE_i}}{\partial D_{MSE_i}} \equiv 0$$

求解上式可以得到最优的D_{MSE_i}和R_{MSE_i}为：

$$\begin{cases} D_{MSE_i} = \alpha\lambda_{MSE} \\ R_{MSE_i} = \alpha log\left(\frac{\sigma_i^2}{\alpha\lambda_{MSE}}\right) \end{cases}$$

而一个视频的总比特率是所有块的比特数目总和，可以计算为：

$$R_{MSE} = \sum_{i=1}^{M} R_{MSE_i} = \alpha\sum_{i=1}^{M} log\left(\frac{\sigma_i^2}{\alpha\lambda_{MSE}}\right)$$

其中，M 是一帧/视频中的编码块数。

同样地，基于率编码失真的目标函数对D_{MSE}取偏导数，并在偏导数为零时可以得到：

$$\frac{\partial J_{\phi}}{\partial D_{MSE_i}} = \frac{\partial D_{PPCM_i}^{\phi}}{\partial D_{MSE_i}} + \lambda_{\phi}\frac{\partial R_i^{\phi}}{\partial D_{MSE_i}} \equiv 0$$

对上式进行求解可以得到

$$R_i^{\phi} = \alpha log\left(\frac{\Phi_i^{\phi}\sigma_i^2}{\alpha\lambda_{\phi}}\right)$$

编码通道ϕ的总比特数计算为：

$$R^{\phi} = \sum_{i=1}^{M} R_i^{\phi} = \alpha\sum_{i=1}^{M} log\left(\frac{\Phi_i^{\phi}\sigma_i^2}{\alpha\lambda_{\phi}}\right)$$

基于此，基于编码通道ϕ的总比特数可以确定感知拉格朗日参数。

在一个实现方式中，当感知拉格朗日乘数用于编码时，所述基于

所述感知系数确定所述投影图像对应的感知拉格朗日乘数具体包括：

基于所述投影图像的每个编码块的感知系数计算第一目标感知系数，并计算所述第一目标感知系数与所述编码块的感知系数的比值，以得到第一感知系数比值；

计算所述第一感知系数比值与所述投影图像的均方误差拉格朗日乘数的乘积，以得到所述编码块对应的感知拉格朗日乘数。

具体地，在 V-PCC 模型中，属性和几何视频是基于 MSE 的编码器编码的。，基于 MSE 的 V-PCC 的总比特率R_{MSE}与基于感知失真度的 V-PCC 的总比特率R^{ϕ}相同，在不同的点云编码通道ϕ，即$R_{MSE} = R^{\phi}$，由此可以得到λ_{ϕ}和λ_{MSE}之间的关系为：

$$\lambda_{\phi} = \lambda_{MSE}\left(\prod_{i=1}^{M} \Phi_{i}^{\phi}\right)^{\frac{1}{M}} \tag{2}$$

由此，基于公式（1）和公式（2），以得到模式决策的 RD 成本为：

$$J_{MSE_i}^{\phi} = D_{MSE_i} + \frac{\left(\prod_{i=1}^{M} \Phi_{i}^{\phi}\right)^{\frac{1}{M}}}{\Phi_{i}^{\phi}} \lambda_{MSE} R_{i}^{\phi}$$

因此，感知拉格朗日系数为：

$$\lambda_{MODE_i}^{\phi} = \frac{\left(\prod_{i=1}^{M} \Phi_{i}^{\phi}\right)^{\frac{1}{M}}}{\Phi_{i}^{\phi}} \lambda_{MSE} \quad 。$$

在一个实现方式中，当感知拉格朗日乘数用于模式决策时，所述基于所述感知系数确定所述投影图像对应的感知拉格朗日乘数具体包括：

基于所述投影图像的每个编码块的感知系数计算第二目标感知系数，并计算第二目标感知系数与编码块的感知系数的平方根的比值，以得到第二感知系数比值；

计算所述第二感知系数比值与所述投影图像的所述平均绝对误

差拉格朗日乘数的平方根的乘积，以得到所述编码块对应的感知拉格朗日乘数。

在传统的基于 MSE 的视频编码器中，SAD/MAD 被用作率失真优化中的失真项，以避免平方运算以保持低复杂度。RD 成本被计算为：

$$J_{MAD_i} = D_{MAD_i} + \lambda_{MAD} R_{MAD_i}$$

其中D_{MAD_i}，R_{MAD_i}和λ_{MAD_i}分别为 HEVC 中基于 MAD 的失真、比特率和拉格朗日乘数。ME 的拉格朗日乘数$\lambda_{MAD} \approx \sqrt{\lambda_{MSE}}$。同样，基于$D^{\phi}_{PPCM,ME_{k,i}} = \sqrt{D^{\phi}_{PPCM_{k,i}}}$的关系，有$\lambda^{\phi}_{MAD_i} \approx \sqrt{\lambda^{\phi}_{MODE_i}}$。

由此，RD 成本为：

$$J^{\phi}_{MAD_i} = D^{\phi}_{MAD_i} + \frac{\left(\prod_{i=1}^{M} \Phi_i^{\phi}\right)^{\frac{1}{2M}}}{\sqrt{\Phi_i^{\phi}}} \lambda_{MAD} R^{\phi}_{ME_i} \quad ;$$

因此，基于感知失真度的感知拉格朗日乘数$\lambda^{\phi}_{ME_i}$被更新为：

$$\lambda^{\phi}_{ME_i} = \frac{\left(\prod_{i=1}^{M} \Phi_i^{\phi}\right)^{\frac{1}{2M}}}{\sqrt{\Phi_i^{\phi}}} \lambda_{MAD} \quad 。$$

为了进一步说明本实施例的编码过程，V-PCC 的编码标准将三维点云的几何特征和纹理特征分别映射为两路视频序列，分别记为几何投影图序列和纹理投影图序列，然后使用已有的 HEVC/VVC 或其他的视频编码器压缩动态点云的视频序列，其中，在编码过程中还需要产生一些元信息，如：占位图和辅助分块信息，用于说明这两个视频序列，其编码过程与现有过程相同，这里不做具体说明。

V-PCC 中的点云投影面片（Patch）是一组信息的集合，包括点云的三维包围盒、相关的几何和纹理信息以及进行三维重构时所需的图集信息。点云投影面片的划分过程是将每一帧点云从三维空间投影到给定的二维平面上，并尽可能少地将点云划分为边界平滑的面片块，

以使重构误差最小。其中，点云投影过程可以为首先计算出输入点云中每一点的法向量，然后按照法向量将整块点云聚类并投影到立方体6个面上形成初始聚类。接着，根据法向量及最邻近点的索引，通过不断迭代更新每个点的聚类索引，从而实现更精细的面片划分。最后，应用连通分量提取法提取连通的面片，并将它们合并成较大的面片块，从而得到最终的面片集合。由此，通过点云投影面片集合在二维平面上打包重组，将三维空间中的几何和纹理信息分别转换为二维图像，从而实现几何和纹理图像的生成和填充过程，以得到几何投影图序列和纹理投影图序列。

下面以已经获取到几何投影图序列和纹理投影图序列为前提进行说明，几何投影图序列和纹理投影图序列的编码过程包括：

a) 进行几何投影图分析，获取几何投影图中的每个编码块的感知系数Φ_i^G，其中，除了第一个帧参考帧，基于当前编码块和前一帧的编码同位块获取Φ_i^G的中的ρ_n参数；

b) 基于感知系数确定感知拉格朗日乘数；

c) 对当前的几何投影帧进行编码；如果所有的视频帧对应的几何投影图像均被编码，进入步骤 d)进行纹理编码，否则，转到步骤a)进行下一帧的编码；

d) 进行纹理投影图分析，获取几何投影图中的每个编码块的感知系数Φ_i^T，其中，除了第一个帧参考帧，基于当前编码块和前一帧的编码同位块获取Φ_i^T的中的ρ_n参数；

e) 基于感知系数确定感知拉格朗日乘数；

f) 对当前纹理投影图进行编码，然后转到步骤 d)对下一个帧纹理投影图进行编码，直到所有纹理投影图均被编码。

综上所述，本实施例提供了一种基于感知加权的点云率失真编码方法，所述方法包括获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像；获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及

结构相似性失真度计算投影图像的感知失真度；根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。本申请通过投影生成二维的投影图像，然后计算以编码块为单位确定每个投影图像对应的感知失真度，并基于感知损失度进行编码决策，提高了点云编码与视觉感知质量的匹配性，从而提高了编码效率。

进一步，为了说明本实施例的准确性，对 Longdress 和 Andrew 两个点云进行了统计分析。每个点云都被压缩了 17 个量化参数(QP)对，范围从 20 到 32 和 27 到 42 不等，每个点云都被投射到二维图像上并分解成 400 个块，即 400×17 个采样块。如图 3 和图 4 所示，采用R^2来计算$\tilde{D}^{\phi}_{PPCM_{k,i}}$和$D^{\phi}_{PPCM_{k,i}}$之间的近似模型准确度，结果表示在属性和几何通道上平均R^2为 0.9732 和 0.9556,感知中的近似值是准确的。

为了验证所提出的编码效率，编码方法在 V-PCC 参考软件 TMC2-10.0 和相应的 HEVC 参考软件 HM16.20-SCM8.8 上实现。最新的 V-PCC（TMC2-10.0+HM16.20-SCM8.8）被用来作为比较的锚定方法。此外,两个最先进的 V-PCC 的 RDO,基于占用图的 RDO,表示为"OC-RDO",和基于 EPM 的 RDO 方法，表示为"EPM-RDO"，也被用作比较的基准。由于提出的编码方法可以应用于 V-PCC 的帧内和帧间编码，将 V-PCC 配置为几何和纹理编码的随机存取编码方式（RA）。符合 V-PCC 通用测试条件（CTC）的低（r1）到高（r5）的五个比特率点被选取测试。为了进行公平的比较，GraphSIM（独立于提议的 PPCM）被用来测量不同编码方案的压缩点云的感知质量。然后，Bjontegaard Delta Bit Rate（BDBR）被用来比较所提出的方案和基准方案的 RD 性能。

表 1 显示了在被测试的 6 个 DPC 上的 RD 比较，其中压缩视频的视觉质量是用 GraphSIM 测量的，R_T表示总比特率（Mbps）。可以看出，与 V-PCC 相比，OC-RDO 可以实现 0.34%到 17.05%的比特率降低，平均为 9.61%。EPM-RDO 能够实现从-4.08%到 5.61%，平均 1.17%的

BDBR 节省。至于所提出的 PWRDO，它实现了 BDBR 从 7.09%到 22.79%的降低，与 V-PCC 相比，平均降低了 13.52%，明显超过了 OC-RDO 和 EPM-RDO 这两个编码优化的对比方案。为了评估所提出的 PWRDO 的感知编码效率，五个不同的感知 PCQA 指标，包括 D1、GraphSIM、MPED、SIAT_PCQA 和提出的 PPCM，被用来衡量不同方案中压缩 DPC 的视觉质量。然后，为每个 PCQA 计算 BDBR，其中包括几何位、属性位和元数据位在内的总编码位的数量。负的 BDBR 意味着节省比特，而正的则表示与锚点相比，编码效率下降。表 1 显示了提出的 PWRDO 和基准 RDO 方案在五个不同的感知 PCQA 指标下的 RD 比较。这些比较结果表明，所提出的 PWRDO 在所有 PCQA 中都能稳定地获得较高的编码收益，这证明了所提出的 PWRDO 的有效性。

表 1. 所提方法与 2 种点云编码率失真优化方法在 5 种评价指标上的增益

	Sequences	*BDBR Gains in Different PCQAs*				
		D1	GraphSIM	MPED	SIAT	PPCM
EPM-RDO	Loot	5.29	-5.61	7.61	-7.31	4.40
	Redandblack	14.84	-2.92	18.81	-1.81	1.85
	Soldier	-11.06	-2.16	-10.69	-8.94	-7.08
	Longdress	9.64	4.08	10.77	-0.53	6.17
	Basketball.	12.43	1.02	11.51	-0.26	7.56
	Dancer	8.43	-1.41	6.29	-2.84	6.74
	Average	**6.60**	**-1.17**	**7.38**	**-3.62**	**3.27**
OC-RDO	Loot	2.25	-9.74	2.62	0.55	-7.66
	Redandblack	-0.62	-15.97	-0.70	-0.65	0.03

	Soldier	-8.66	-11.07	-8.76	-6.96	-18.86
	Longdress	-0.62	-17.05	1.37	-1.15	2.37
	Basketball.	-0.20	-3.05	-2.31	2.71	-0.64
	Dancer	-3.21	-0.34	-4.07	-1.18	-5.14
	Average	**-1.84**	**-9.61**	**-1.97**	**-1.11**	**-4.98**
proposed PWRDO	Loot	-2.91	-11.33	-8.02	-13.73	-8.53
	Redandblack	-5.78	-22.79	-9.17	-8.87	-8.87
	Soldier	-13.61	-9.67	-14.07	-13.35	-9.75
	Longdress	-9.88	-21.58	-7.91	-14.34	-12.96
	Basketball.	-2.53	-7.09	-4.30	-5.99	-9.44
	Dancer	-2.73	-8.65	-5.49	-7.07	-7.67
	Average	**-6.24**	**-13.52**	**-8.16**	**-10.56**	**-9.54**

基于上述基于感知加权的点云率失真编码方法，本实施例提供了一种基于感知加权的点云率失真编码系统，如图 5 所示，所述系统包括：

获取模块 100，用于获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像，其中，所述投影图像和若干参考投影图像均包括几何投影图和纹理投影图；

计算模块 200，用于获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度；

编码模块 300，用于根据投影图像的感知失真度确定所述待编码

点云对应的码率失真成本，并基于所述码率失真成本编码所述投影图像。

基于上述基于感知加权的点云率失真编码方法，本实施例提供了一种计算机可读存储介质，所述计算机可读存储介质存储有一个或者多个程序，所述一个或者多个程序可被一个或者多个处理器执行，以实现如上述实施例所述的基于感知加权的点云率失真编码方法中的步骤。

基于上述基于感知加权的点云率失真编码方法，本申请还提供了一种终端设备，如图 6 所示，其包括至少一个处理器（processor）20；显示屏 21；以及存储器（memory）22，还可以包括通信接口（Communications Interface）23 和总线 24。其中，处理器 20、显示屏 21、存储器 22 和通信接口 23 可以通过总线 24 完成相互间的通信。显示屏 21 设置为显示初始设置模式中预设的用户引导界面。通信接口 23 可以传输信息。处理器 20 可以调用存储器 22 中的逻辑指令，以执行上述实施例中的方法。

此外，上述的存储器 22 中的逻辑指令可以通过软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。

存储器 22 作为一种计算机可读存储介质，可设置为存储软件程序、计算机可执行程序，如本公开实施例中的方法对应的程序指令或模块。处理器 20 通过运行存储在存储器 22 中的软件程序、指令或模块，从而执行功能应用以及数据处理，即实现上述实施例中的方法。

存储器 22 可包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序；存储数据区可存储根据终端设备的使用所创建的数据等。此外，存储器 22 可以包括高速随机存取存储器，还可以包括非易失性存储器。例如，U 盘、移动硬盘、只读存储器(Read-Only Memory，ROM)、随机存取存储器(Random

Access Memory，RAM）、磁碟或者光盘等多种可以存储程序代码的介质，也可以是暂态存储介质。

此外，上述存储介质以及终端设备中的多条指令处理器加载并执行的具体过程在上述方法中已经详细说明，在这里就不再一一陈述。

最后应说明的是：以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

权利要求书

1. 一种基于感知加权的点云率失真编码方法，其特征在于，所述方法包括：

获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像，其中，所述投影图像和若干参考投影图像均包括几何投影图和纹理投影图；

获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度；

根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。

2. 根据权利要求 1 所述基于感知加权的点云率失真编码方法，其特征在于，所述参考投影图像为通过对投影图像进行下采样得到的，并且投影图像和若干参考投影图像的图像尺度互不相同。

3. 根据权利要求 1 所述基于感知加权的点云率失真编码方法，其特征在于，所述梯度权重的获取过程具体包括：

计算所述编码块中的各像素点的梯度值，并基于各像素点的梯度值计算所述编码块的梯度权重。

4. 根据权利要求 1 所述基于感知加权的点云率失真编码方法，其特征在于，所述根据投影图像的感知失真度确定所述待编码点云对应的率失真成本具体包括：

将所述感知失真度中的结构相似性失真度转换为均方误差失真度，以得到所述投影图像对应的感知系数；

基于所述感知系数确定所述投影图像对应的感知拉格朗日乘数；

基于所述感知失真度和所述感知拉格朗日乘数确定所述待编码点云对应的率失真成本。

5. 根据权利要求 4 所述基于感知加权的点云率失真编码方法，其

特征在于，所述结构相似性失真度与所述均方误差失真度的对应关系为：

$$D_{SIMM_{k,i}} = \Phi_{k,i} D_{MSE_{k,i}}$$

$$\Phi_{k,i} = \frac{\sum_{n-1}^{N} \left(\rho_n \frac{1}{16} \sum_{i \in blk_n} E_i \right)}{\sum_{n=1}^{N} \rho_n}$$

其中，$D_{SIMM_{k,i}}$表示第k个目标投影图像的编码块i的结构相似性失真度，$D_{MSE_{k,i}}$表示第k个目标投影图像的编码块i的均方误差失真度，$\Phi_{k,i}$表示第k个目标投影图像的编码块i的感知系数，blk_n表示编码块i的第n个子图像块，ρ_n表示编码块i的第n个子图像块的线性模型参数，N表示子图像块的数量，E_i表示像素内容权重，第k个目标投影图像为由投影图像和若干参考投影图像形成的投影图像集中的一投影图像。

6. 根据权利要求 4 所述基于感知加权的点云率失真编码方法，其特征在于，当感知拉格朗日乘数用于编码时，所述基于所述感知系数确定所述投影图像对应的感知拉格朗日乘数具体包括：

基于所述投影图像的每个编码块的感知系数计算第一目标感知系数，并计算所述第一目标感知系数与所述编码块的感知系数的比值，以得到第一感知系数比值；

计算所述第一感知系数比值与所述投影图像的均方误差拉格朗日乘数的乘积，以得到所述编码块对应的感知拉格朗日乘数。

7. 根据权利要求 4 所述基于感知加权的点云率失真编码方法，其特征在于，当感知拉格朗日乘数用于模式决策时，所述基于所述感知系数确定所述投影图像对应的感知拉格朗日乘数具体包括：

基于所述投影图像的每个编码块的感知系数计算第二目标感知系数，并计算第二目标感知系数与编码块的感知系数的平方根的比值，以得到第二感知系数比值；

计算所述第二感知系数比值与所述投影图像的所述均方误差拉

格朗日乘数的平方根的乘积，以得到所述编码块对应的感知拉格朗日乘数。

8. 一种基于感知加权的点云率失真编码系统，其特征在于，所述系统包括：

获取模块，用于获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像，其中，所述投影图像和若干参考投影图像均包括几何投影图和纹理投影图；

计算模块，用于获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度；

编码模块，用于根据投影图像的感知失真度确定所述待编码点云对应的率失真成本，并基于所述率失真成本编码所述投影图像。

9. 一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有一个或者多个程序，所述一个或者多个程序可被一个或者多个处理器执行，以实现如权利要求 1-7 任意一项所述的基于感知加权的点云率失真编码方法中的步骤。

10. 一种终端设备，其特征在于，包括：处理器、存储器及通信总线；所述存储器上存储有可被所述处理器执行的计算机可读程序；

所述通信总线实现处理器和存储器之间的连接通信；

所述处理器执行所述计算机可读程序时实现如权利要求 1-7 任意一项所述的基于感知加权的点云率失真编码方法中的步骤。

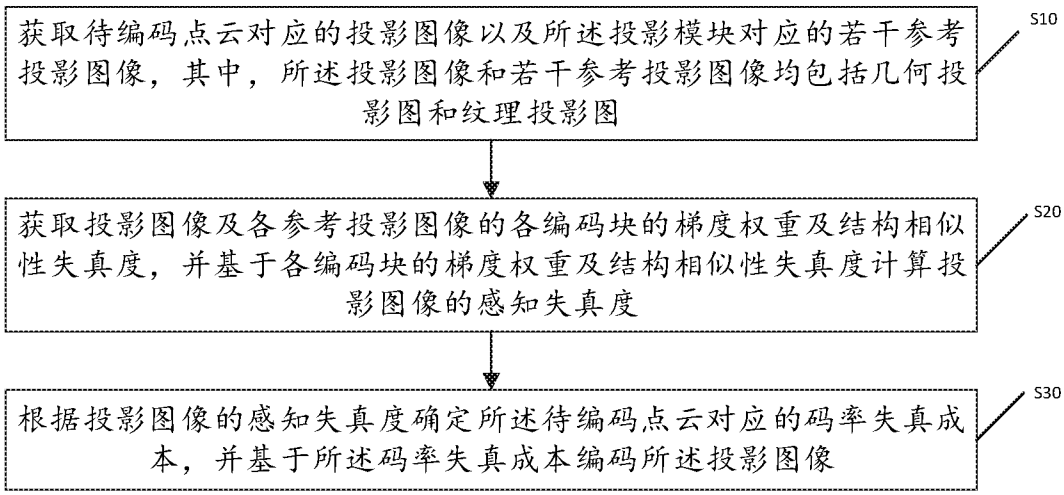

获取待编码点云对应的投影图像以及所述投影模块对应的若干参考投影图像，其中，所述投影图像和若干参考投影图像均包括几何投影图和纹理投影图
S10
获取投影图像及各参考投影图像的各编码块的梯度权重及结构相似性失真度，并基于各编码块的梯度权重及结构相似性失真度计算投影图像的感知失真度
S20
根据投影图像的感知失真度确定所述待编码点云对应的码率失真成本，并基于所述码率失真成本编码所述投影图像
S30

图 1

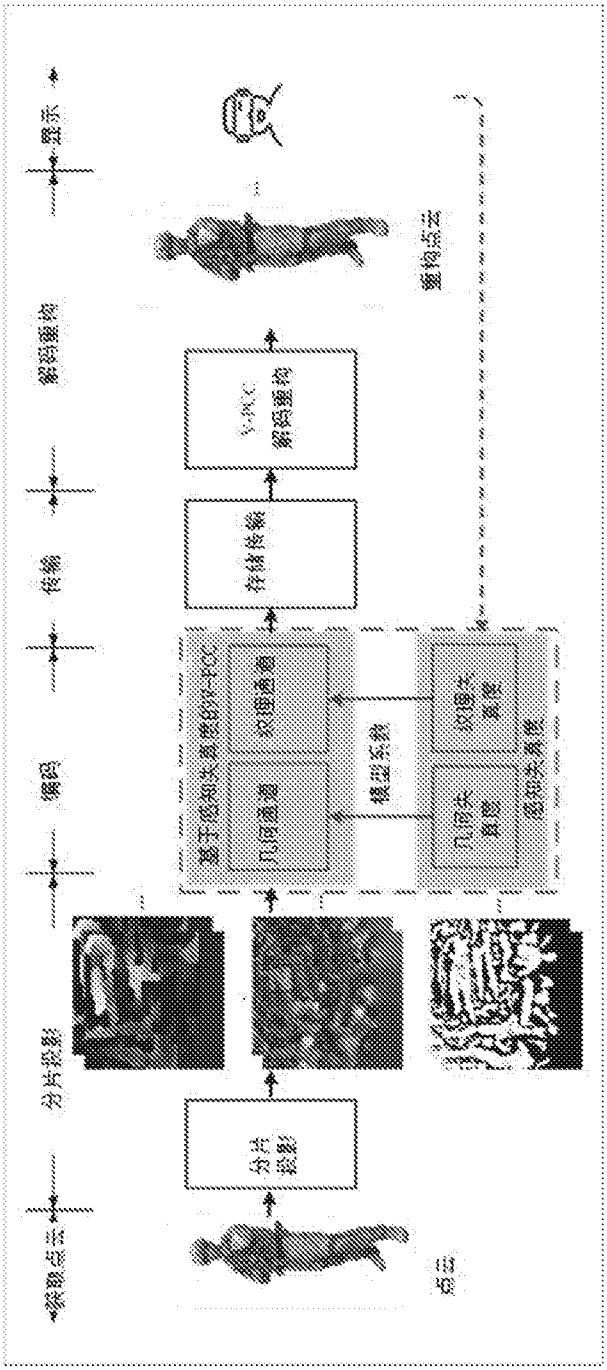

获取点云
分片投影
编码
传输
解码重构
显示
分片投影
几何通道
纹理通道
基于感知失真模型的V-PCC
模型系数
存储传输
V-PCC
解码重构
重构点云
几何失真度
纹理失真度
感知失真度

图 2

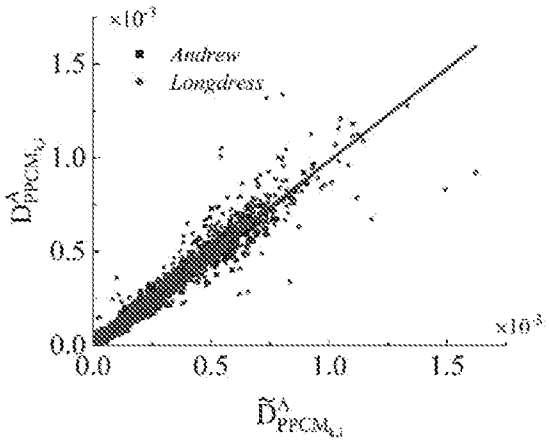

×10⁻³
Andrew
Longdress
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
×10⁻³

图 3

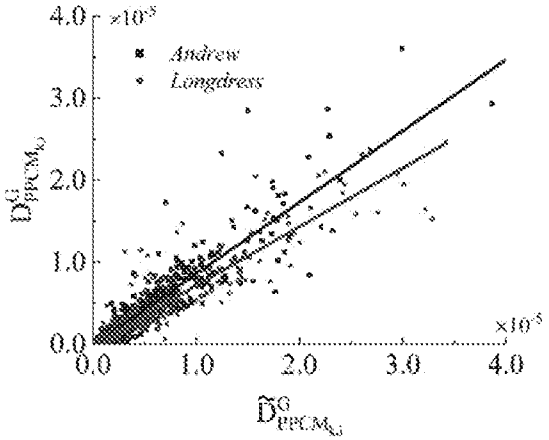

×10⁻⁵
Andrew
Longdress
4.0
3.0
2.0
1.0
0.0
0.0
1.0
2.0
3.0
4.0
×10⁻⁵

图 4

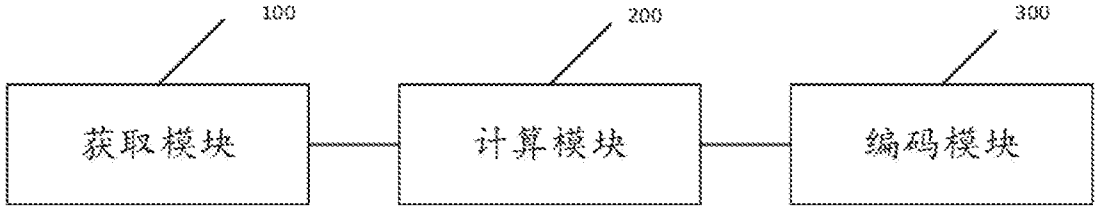

100
200
300
获取模块
计算模块
编码模块

图 5

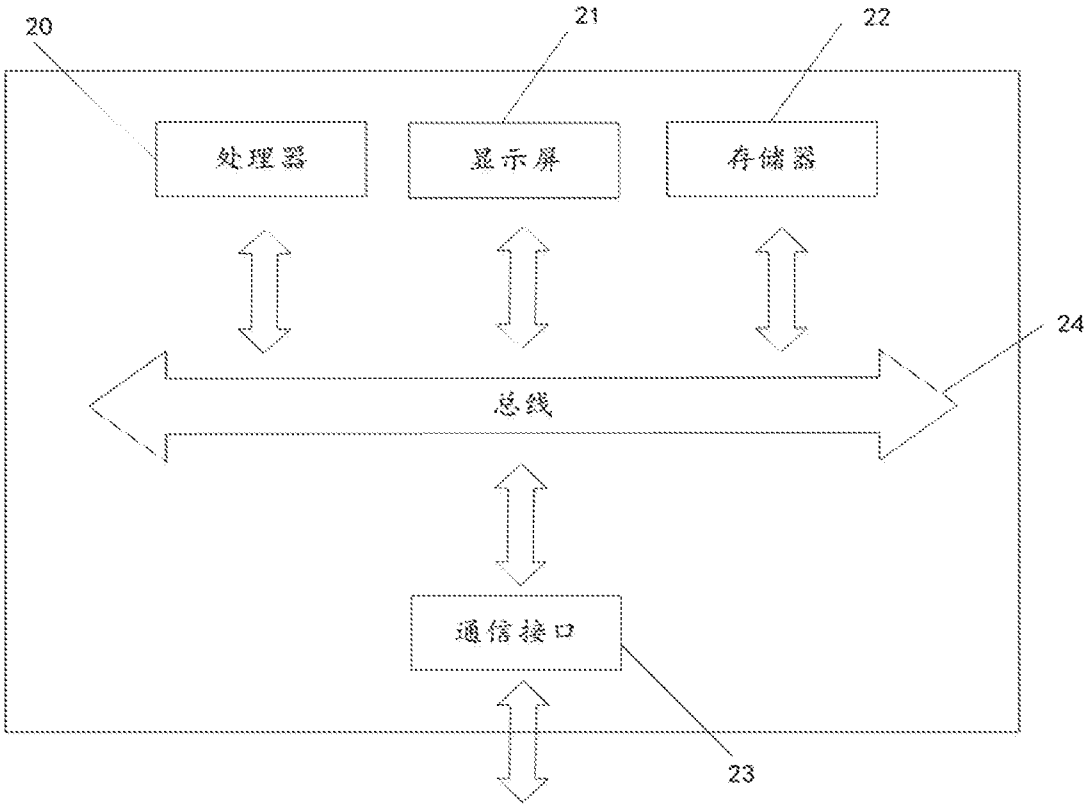


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2023/137731

A. CLASSIFICATION OF SUBJECT MATTER

H04N19/147(2014.01)i; H04N19/62(2014.01)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)

IPC: H04N G06T

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)

CNTXT, CNABS, WPABS, WPABSC, ENTXT, ENTXTC, DWPI, CNKI, IEEE, JCTVC: 点云, 二维映射, 结构相似性, 梯度, 权重, 投影, 拉格朗日, 感知失真度, 投影面片, 损失, 动态点云, 均方差, SSIM, V-PCC, MSE, point cloud, two-dimensional map, structural similarity, gradient, weight, projection, lagrangian, distortion, projected patches, loss, dynamic point cloud, mean square error

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 115018753 A (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 06 September 2022 (2022-09-06) description, paragraphs [0014]-[0130]	1-10
A	CN 103607590 A (BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 26 February 2014 (2014-02-26) entire document	1-10
A	CN 103918271 A (WANG ZHOU et al.) 09 July 2014 (2014-07-09) entire document	1-10
A	CN 112601082 A (NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 02 April 2021 (2021-04-02) entire document	1-10
A	CN 115631126 A (QINGDAO UNIVERSITY) 20 January 2023 (2023-01-20) entire document	1-10

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"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)
"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
"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
"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
"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
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 February 2024	**04 March 2024**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2012171113 A1 (WANG ZHOU) 20 December 2012 (2012-12-20) entire document	1-10
A	KARCZEWICZ, Marta et al. "Intra Frame Rate Control Based on SATD" *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/ IEC JTC 1/SC 29/WG 11 13th Meeting: Incheon, KR, 18–26 Apr. 2013, JCTVC-M0257,* 26 April 2013 (2013-04-26), entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2023/137731

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 115018753 A	06 September 2022	None	
CN 103607590 A	26 February 2014	None	
CN 103918271 A	09 July 2014	EP 2716048 A1	09 April 2014
		WO 2012162806 A1	06 December 2012
		KR 20140037880 A	27 March 2014
		US 2014140396 A1	22 May 2014
		CA 2837755 A1	06 December 2012
		CA 2837755 C	09 April 2019
CN 112601082 A	02 April 2021	None	
CN 115631126 A	20 January 2023	None	
WO 2012171113 A1	20 December 2012	US 2014119432 A1	01 May 2014
		EP 2721826 A1	23 April 2014
		KR 20140042845 A	07 April 2014
		CA 2839345 A1	20 December 2012

国际检索报告

国际申请号
PCT/CN2023/137731

A. 主题的分类

H04N19/147(2014.01)i；H04N19/62(2014.01)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)

IPC：H04N G06T

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））

CNTXT,CNABS,WPABS,WPABSC,ENTXT,ENTXTC,DWPI,CNKI,IEEE，JCTVC:点云,二维映射,结构相似性,梯度,权重,投影,拉格朗日,感知失真度,投影面片,损失,动态点云，均方差，SSIM, V-PCC, MSE,point cloud, two-dimensional map, structural similarity, gradient, weight, projection, lagrangian, distortion, projected patches, loss, dynamic point cloud, mean square error

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
A	CN 115018753 A (中国科学院深圳先进技术研究院) 2022年9月6日 (2022 - 09 - 06) 说明书第[0014]段-[0130]段	1-10
A	CN 103607590 A (北京邮电大学) 2014年2月26日 (2014 - 02 - 26) 全文	1-10
A	CN 103918271 A (王舟等) 2014年7月9日 (2014 - 07 - 09) 全文	1-10
A	CN 112601082 A (南京邮电大学) 2021年4月2日 (2021 - 04 - 02) 全文	1-10
A	CN 115631126 A (青岛大学) 2023年1月20日 (2023 - 01 - 20) 全文	1-10
A	WO 2012171113 A1 (WANG， Zhou) 2012年12月20日 (2012 - 12 - 20) 全文	1-10

☑ 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

"D" 申请人在国际申请中引证的文件

"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	国际检索报告邮寄日期
2024年2月23日	2024年3月4日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	受权官员 龚锦玲 电话号码 (+86) 010-53961810

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
A	KARCZEWICZ，Marta等. "Intra Frame Rate Control Based on SATD" Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 13th Meeting: Incheon, KR, 18 - 26 Apr. 2013，JCTVC-M0257, 2013年4月26日 (2013 - 04 - 26)， 全文	1-10

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/137731

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	115018753	A	2022年9月6日	无			
CN	103607590	A	2014年2月26日	无			
CN	103918271	A	2014年7月9日	EP	2716048	A1	2014年4月9日
				WO	2012162806	A1	2012年12月6日
				KR	20140037880	A	2014年3月27日
				US	2014140396	A1	2014年5月22日
				CA	2837755	A1	2012年12月6日
				CA	2837755	C	2019年4月9日
CN	112601082	A	2021年4月2日	无			
CN	115631126	A	2023年1月20日	无			
WO	2012171113	A1	2012年12月20日	US	2014119432	A1	2014年5月1日
				EP	2721826	A1	2014年4月23日
				KR	20140042845	A	2014年4月7日
				CA	2839345	A1	2012年12月20日