



(12) 发明专利

(10) 授权公告号 CN 101888566 B

(45) 授权公告日 2012. 02. 15

(21) 申请号 201010222351. 9

审查员 杨浩

(22) 申请日 2010. 06. 30

(73) 专利权人 清华大学

地址 100084 北京市 100084-82 信箱

(72) 发明人 季向阳 汪启扉 戴琼海

(74) 专利代理机构 北京清亦华知识产权代理事

务所(普通合伙) 11201

代理人 黄德海

(51) Int. Cl.

H04N 13/00(2006. 01)

H04N 7/26(2006. 01)

(56) 对比文件

CN 101729891 A, 2010. 06. 09,

CN 101466038 A, 2009. 06. 24,

CN 1471319 A, 2004. 01. 28,

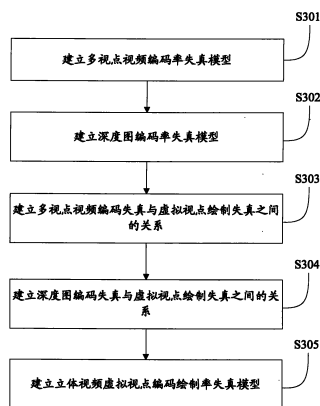
权利要求书 2 页 说明书 8 页 附图 3 页

(54) 发明名称

立体视频编码率失真性能估计方法

(57) 摘要

本发明提出一种立体视频编码率失真性能估计方法,包括以下步骤:获得多视点视频,并根据所述多视点视频获得对应的多视点深度图;分别获得多视点视频编码率失真模型和多视点深度图编码率失真模型;分别获得所述多视点视频编码失真与虚拟视点绘制失真之间的关系,以及所述多视点深度图编码失真与虚拟视点绘制失真之间的关系;联合上述关系建立立体视频虚拟视点编码绘制率失真模型。通过本发明提出的立体视频中虚拟视点编码绘制率失真分析方法,能够准确快速地估计立体视频中虚拟视点编码绘制率失真性能,从而为立体视频编码参数的选择与码率分配等问题提供了模型指导和解决方案。



1. 一种立体视频编码率失真性能估计方法,其特征在于,包括以下步骤:

获得多视点视频,并根据所述多视点视频获得对应的多视点深度图;

根据所述多视点视频和所述多视点深度图分别获得多视点视频编码率失真模型和多视点深度图编码率失真模型;

根据所述多视点视频和所述多视点深度图分别获得所述多视点视频编码失真与虚拟视点绘制失真之间的关系,以及所述多视点深度图编码失真与虚拟视点绘制失真之间的关系;

根据所述多视点视频编码失真与虚拟视点绘制失真之间的关系,以及所述多视点深度图编码失真与虚拟视点绘制失真之间的关系获得虚拟视点绘制失真与多视点视频以及多视点深度各自编码量化参数QP之间的关系;

根据所述多视点视频编码率失真模型和多视点深度图编码率失真模型获得绘制所述虚拟视点所需要的编码码率和所述QP之间的关系;统计不同QP下所述虚拟视点的绘制失真和绘制所述虚拟视点所需要的编码码率以得到立体视频编码中虚拟视点的率失真模型。

2. 如权利要求1所述的立体视频编码率失真性能估计方法,其特征在于,其中,多视点视频编码采用基于视频编码标准H.264/AVC的多视点视频编码扩展版本来实现。

3. 如权利要求2所述的立体视频编码率失真性能估计方法,其特征在于,各视点内的每一个图像组的帧内编码帧、前向预测编码帧以及双向预测编码帧均采用相同的量化参数QP进行编码。

4. 如权利要求1所述的立体视频编码率失真性能估计方法,其特征在于,其中,在给定的多视点视频编码量化参数的条件下:

多视点视频编码的码率的估计方法为 $r_c = a/Q_{step}^2 + b/Q_{step} + c$,其中, $Q_{step} = 2^{(QP-4)/6}$,a为24.37,b为-0.303,和c为0.02;

多视点视频编码的失真的估计方法为 $PSNR_c = q_c \times QP + p_c$,其中, q_c 为-0.37,和 p_c 为49.30。

5. 如权利要求1所述的立体视频编码率失真性能估计方法,其特征在于,其中,在给定的多视点深度图编码量化参数的条件下:

多视点深度图编码的码率的估计方法为 $r_d = k/Q_{step} + t$,其中, $Q_{step} = 2^{(QP-4)/6}$,k为0.9996,和t为0.0040;

多视点视频编码的失真计算表达式为 $PSNR_d = q_d \times QP + p_d$,其中, q_d 为-0.65,和 p_d 为63.75。

6. 如权利要求1所述的立体视频编码率失真性能估计方法,其特征在于,所述获得多视点视频编码失真与虚拟视点绘制失真之间的关系进一步包括:

假设多视点深度图编码无失真,所述虚拟视点的失真仅由所述多视点视频编码失真引起,并选择绘制虚拟视点 C_v 所需的M个参考视点,则虚拟视点图像失真为 $E_{CW} = \sum_{i=1}^M E_{CW}^{ref_i}$,其

中, $E_{CW}^{ref_i} = \eta_{ref_i} w_{ref_i} E_C^{ref_i}$,其中, η_{ref_i} 为被遮挡的像素点的个数和参考视点整帧图像的像素点个数的比例, w_{ref_i} 为第i个参考视点视频编码失真对于虚拟视点图像绘制失真贡献的加权系数, $E_C^{ref_i}$ 为第i个参考视点 C_{ref_i} 的视频编码失真。

7. 如权利要求 1 所述的立体视频编码率失真性能估计方法,其特征在于,所述获得多视点深度图编码失真与虚拟视点绘制失真之间的关系进一步包括:

假设多视点视频编码无失真,所述虚拟视点的失真仅由所述多视点深度图编码失真引起,并选择虚拟视点 C_v 所需的 M 个参考视点,则虚拟视点图像失真为

$$E_{DW} = \sum_{i=1}^M \eta_{ref_i} \times w_{ref_i} \times E_{DW}^{ref_i},$$

其中, η_{ref_i} 为被遮挡的像素点的个数和参考视点整帧图像的像素点个数的比例, w_{ref_i} 为第 i 个参考视点视频编码失真对于虚拟视点图像绘制失真贡献的加权系数, $E_{DW}^{ref_i} = \sum_j e_{DW}^j$, e_{DW}^j 为虚拟视点绘制图像上与参考视点中 B_i^j 区域相对应的区域的失真。

8. 如权利要求 7 所述的立体视频编码率失真性能估计方法,其特征在于,其中,

$$e_{DW}^j = \|\Delta d_i\| \psi_C^j,$$

其中, ψ_C^j 的计算公式为 $\psi_C^j = \frac{1}{2(2\pi)^2} \left(\iint S_C^j(\omega) \omega_1^2 d\omega + \iint S_C^j(\omega) \omega_2^2 d\omega \right)$, S_C^j 为参考视点视频帧上 B_i^j 区域的亮度值或色度值所组成的矩阵进行傅立叶变换后得到的矩阵, $\omega = [\omega_1, \omega_2]$ 为 B_i^j 区域水平和垂直方向的角频率, $\|\Delta d_j\|$ 为视差的平均失真。

9. 如权利要求 1 所述的立体视频编码率失真性能估计方法,其特征在于,所述根据多视点视频编码失真与虚拟视点绘制失真之间的关系,以及所述多视点深度图编码失真与虚拟视点绘制失真之间的关系获得虚拟视点的失真与编码量化参数 QP 之间的关系进一步包括:

虚拟视点绘制图像的误差 E_T 为 $E_T = E_{CW} + E_{DW} + E_0$, 其中, E_0 为通过图像补绘来填补空洞所引起的误差, $E_0 = \sum_{i_0} e_{i_0}^o = \sum_{i_0} \sum_{y \neq y_{in}} P_{i_0}^o(y) \times (y - y_{in})$,

其中, i_0 为所述被遮挡像素点, $P_{i_0}^o$ 为所述被遮挡像素点 i_0 相邻的像素点的亮度或色度值分布, y 为所述像素点 i_0 的亮度或色度的取值范围, y_{in} 为所述被遮挡像素点 i_0 补绘时的亮度或色度值。

10. 如权利要求 1 所述的立体视频编码率失真性能估计方法,其特征在于,根据所述多视点视频编码率失真模型和多视点深度图编码率失真模型获得绘制所述虚拟视点所需要的编码码率和多视点视频编码码率以及多视点深度图编码码率之间的关系进一步包括:

立体视频编码所需的码率为 $r_t = r_c + r_d$, 其中, r_c 为多视点视频编码码率, r_d 为多视点深度图编码码率。

立体视频编码率失真性能估计方法

技术领域

[0001] 本发明涉及图像处理技术领域,特别涉及一种针对虚拟视点绘制的立体视频编码率失真性能估计方法。

背景技术

[0002] 随着多媒体技术的不断发展,传统的二维视频媒体已经不能满足人类视觉的需要,人们希望看到更具真实感和互动性的视频节目,也即具有高度立体沉浸感的三维视频。近年来逐渐兴起的三维视频技术,如多视点视频、自由视点视频以及立体视频等全新的视频媒体正逐渐成为多媒体技术重要的组成部分。所谓自由视点视频是指观看者通过自由选择观看视点来观看三维场景的视频媒体。

[0003] 自由视点视频发展经历了两个主要阶段,第一阶段被称作多视点视频,即观众可以从播放的多视点视频中选择自己需要的视点来观看,通过不同视点视频之间的视差来形成立体感。第二阶段被称为立体视频,即观看者可以自由选择观看视点观看具有立体感的视频。为了获得多视点视频,就需要采用多路摄像机来采集同一场景,并将采集到的多路视频传送到客户端,以及根据观众的需要来选择对应的多视点视频进行播放。而为了获得具有自由视点的立体视频,需要在多视点视频的基础上进一步引入场景的几何信息,通过多视点视频数据和几何信息来渲染得到观众随机选择的虚拟视点,从而让观众获得更强的观看立体感。

[0004] 区别于传统的单路二维视频数据,立体视频是由多路二维视频数据和场景几何信息组成,因此多视点视频的海量数据对传输带宽的需求远远大于传统的二维视频数据。此外,如何有效地压缩场景几何信息也成为立体视频编码需要面对的新挑战。因此,为了实现对立体视频的有效传输,需要针对立体视频特性来设计高效的编码技术。在早期,研究人员将传统的视频压缩技术应用于多视点视频压缩,采用传统的单路视频压缩技术将多视点视频中的每一个视点对应的视频单独进行压缩。该方案成为立体视频压缩最早的解决方案之一。

[0005] 然而,多视点视频序列之间存在较强的相关性,即使是采用最高效的 H. 264/AVC 编码技术,仍然无法有效地压缩不同视点之间的冗余。为此,研究人员设计了更加高效的多视点视频压缩方案。该方案将传统的单路视频编码技术拓展到多路视频上来,在时域预测编码的基础上加入视点间的预测编码,进一步压缩了多视点视频视点间的冗余,提高了多视点视频压缩率失真性能。此外,针对基于深度图的立体视频场景几何信息,即多视点深度图序列,研究人员同样采用多视点视频编码方案来实现对其高效压缩。随着立体视频技术的不断发展,新的立体视频压缩技术正在快速的发展。目前,实时高效的立体视频编码技术已经成为立体视频的应用走向市场的关键技术之一。

[0006] 在立体视频高效编码方案的基础上,为了实现对立体视频的有效传输,需要根据网络条件调整编码参数。在传统的视频编码中,当已知网络带宽的情况下,可以根据编码的率失真模型估计解码质量,从而更好地调整编码参数。因此,针对立体视频高效编码方案的

率失真模型分析是实现立体视频有效传输的重要环节。

[0007] 现有技术存在的缺点是,目前还没有针对虚拟视点绘制的立体视频编码率失真性能的估计方案。

[0008] 例如,在申请号为 200810163801,名称为一种网络自适应的立体视频编码方法的专利申请中,仅仅给出了立体视频不同的网络带宽下的自适应编码方案,没有针对立体视频编码的率失真性能进行分析。在申请号为 200810126528,名称为一种立体视频编解码方法、装置及系统的专利申请中,虽然公开了一整套完整的立体视频编码方案、装置和系统,但是同样没有给出立体视频编码率失真性能分析。在申请号为 200710164747,名称为一种面向多视点视频的码率控制方法的专利申请中,虽然公布了多视点视频编码的率失真模型,但是没有针对虚拟视点绘制的立体视频编码率失真性能。

发明内容

[0009] 本发明的目的旨在至少解决上述技术缺陷,提出了一种针对虚拟视点绘制的立体视频编码率失真性能估计方法。

[0010] 为达到上述目的,本发明一方面提出了一种立体视频编码率失真性能估计方法,包括以下步骤:获得多视点视频,并根据所述多视点视频获得对应的多视点深度图;根据所述多视点视频和所述多视点深度图分别获得多视点视频编码率失真模型和多视点深度图编码率失真模型;根据所述多视点视频和所述多视点深度图分别获得所述多视点视频编码失真与虚拟视点绘制失真之间的关系,以及所述多视点深度图编码失真与虚拟视点绘制失真之间的关系;根据所述多视点视频编码失真与虚拟视点绘制失真之间的关系,以及所述多视点深度图编码失真与虚拟视点绘制失真之间的关系获得虚拟视点绘制失真与多视点视频以及多视点深度图各自的编码量化参数 QP 之间的关系;根据所述多视点视频编码率失真模型和多视点深度图编码率失真模型获得绘制所述虚拟视点所需要的编码码率和所述 QP 之间的关系;并统计不同 QP 下所述虚拟视点的绘制失真和绘制所述虚拟视点所需要的编码码率以得到立体视频编码中虚拟视点的率失真模型。

[0011] 通过本发明提出的立体视频中虚拟视点绘制编码率失真分析方法,能够准确快速地估计立体视频中虚拟视点绘制编码率失真性能,从而为立体视频编码参数的选择与码率分配等问题提供了模型指导和解决方案,进一步提高了立体视频编码的效率。

[0012] 本发明附加的方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0013] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0014] 图 1 为本发明实施例提供的立体视频系统框图;

[0015] 图 2 为本发明实施例提供的多视点视频编码预测结构图;

[0016] 图 3 为本发明实施例的立体视频编码率失真性能估计方法流程图;

[0017] 图 4 为本发明实施例提供的虚拟视点绘制原理图;

[0018] 图 5 为本发明实施例提供的深度图四叉树分解示意图。

具体实施方式

[0019] 下面详细描述本发明的实施例,所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,仅用于解释本发明,而不能解释为对本发明的限制。

[0020] 由于立体视频中,虚拟视点是通过多视点视频和场景几何信息来绘制,因此虚拟视点绘制质量与多视点视频以及场景几何信息的编码质量相关。为此,本发明通过需要建立多视点视频与场景几何信息编码联合率失真模型来获得针对虚拟视点绘制视频的编码率失真模型,从而更好地指导立体视频编码参数的选择和码率分配等问题。

[0021] 本发明主要提出了一种实用的多视点视频与深度图编码联合率失真模型估计方法。该方法能够有效地估计多视点视频与深度图编码质量和虚拟视点绘制质量之间的关系,从而为给定网络带宽的条件下立体视频编码参数选择以及码率分配等问题提供了理论指导。

[0022] 本发明实施例的应用环境如下:本发明实施例应用的立体视频系统框图如图 1 所示。其中,用于立体视频编码的视频序列采用高清格式的名字为“Breakdancer”的标准测试视频序列;该高清格式视频序列的像素为 1024×768 ;解码器采用 H. 264/SVC(Multi-view Video Coding, 多视点视频扩展版本)标准的参考软件 JMVC(Joint Multi-view Video Coding, 多视点视频编码);编码器 GOP(Group of Pictures, 图像组)的帧数为 8;编码的时域预测编码采用 Hierarchical B(层次化双向预测编码帧,简称层次化 B 帧)的预测结构,编码预测结构图如图 2 所示。

[0023] 服务器端获取多视点视频,并根据多视点视频获得多视点深度图,并在立体视频编码中对多视点视频和多视点深度图进行编码并通过网络传输给客户端。客户端对接收到的多视点视频和多视点深度图编码码流进行解码,并根据观众选择的虚拟视点生成相应的虚拟视点视频并将其通过立体显示器显示给观众。

[0024] 在本实施样例中,虚拟视点绘制采用与虚拟视点相邻的两路视频和深度图来绘制。具体的,本实施样例采用“Breakdancer”序列的视点 4 和视点 6 这两路视频作为多视点视频输入序列。其中视点 4 称为左参考视点,视点 6 称为右参考视点。多视点视频和多视点深度图编码量化参数 QP 的取值范围为 0 到 51 之间的整数。生成的虚拟视点的参数与“Breakdancer”序列中视点 5 的参数相同。

[0025] 如图 3 所示,为本发明实施例的立体视频编码率失真性能估计方法流程图,在该实施例中,多视点视频编码采用基于传统视频编码标准 H. 264/AVC 的多视点视频编码扩展版本来实现。各视点内的每一个图像组的帧内编码帧(I 帧)、前向预测编码帧(P 帧)以及双向预测编码帧(B 帧)均采用相同的量化参数(QP)进行编码。该方法包括以下步骤:

[0026] 步骤 S301,建立多视点视频编码率失真模型。在给定多视点视频编码量化参数的情况下,估计多视点视频编码码率和失真。当给定编码量化参数 QP 的取值后,首先通过 $Q_{step} = 2^{(QP-4)/6}$ 得到量化参数 Q_{step} 的取值。例如,当 $QP = 28$ 时, $Q_{step} = 2^{(28-4)/6} = 16$ 。

[0027] 该步骤可进一步细分为以下两步:

[0028] 步骤 S101,多视点视频编码的码率 r_c 的计算表达式为 $r_c = a/Q_{step}^2 + b/Q_{step} + c$, 其中,参数 Q_{step} 的计算表达式为 $Q_{step} = 2^{(QP-4)/6}$,参数 a, b, c 需要根据不同视频序列具体设

置。具体地,在本发明的实施例中,对于“Breakdancer”序列,由于编码码率 r_c 和量化参数 Q_{step} 之间的关系表达式为 $r_c = a/Q_{step} + b/Q_{step}^2 + c$,因此可以通过线性回归的方法来拟合得到参数 a 为 24.37, b 为 -0.303 和 c 为 0.02,也即对于左右视点其编码码率 r_c 和量化参数 Q_{step} 间的关系表达式为: $r_c = 24.37/Q_{step}^2 - 0.303/Q_{step} + 0.02$ 。例如,当 $Q_{step} = 16$ 时, $r_c = 0.096$ bpp。

[0029] 步骤 S102,多视点视频编码的失真计算表达式为 $PSNR_c = q_c \times QP + p_c$,其中参数 q_c 和 p_c 需要根据不同视频序列具体设置。此时,在给定多视点视频编码量化参数 QP 的条件下,可以得到多视点视频编码码率 r_c 和失真 $PSNR_c$ 之间的对应关系。

[0030] 具体地,在本发明的实施例中,对于“Breakdancer”序列,由于多视点视频编码的失真计算表达式为 $PSNR_c = q_c \times QP + p_c$,通过线性回归的方法来拟合得到参数 q_c 为 -0.37 和 p_c 为 49.30,也即对于左右视点其解码质量 $PSNR_c$ 和编码量化参数 QP 之间的关系可以表示为 $PSNR_c = -0.37 \times QP + 49.30$ 。当 $QP = 28$ 时, $PSNR_c = -0.37 \times 28 + 49.30 = 38.94$ dB。

[0031] 步骤 S302,建立深度图编码率失真模型。在给定多视点深度图编码量化参数的情况下,估计多视点视频深度图编码码率和失真之间的对应关系。在该步骤中,多视点深度图也采用同步步骤 S301 中所述多视点视频编码进行编码。类似的,各视点内的每一个图像组的 I 帧、P 帧以及 B 帧也均采用相同的 QP 进行编码。同上述步骤,当编码量化参数 QP 的取值为 28 时 $Q_{step} = 16$ 。

[0032] 该步骤可以进一步细分为以下两步:

[0033] S221、多视点深度图编码的码率 r_d 的计算表达式为 $r_d = k/Q_{step} + t$,其中参数 Q_{step} 的计算表达式为 $Q_{step} = 2^{(QP-4)/6}$,参数 k, t 需要根据不同深度图序列具体设置。

[0034] 具体地,在本发明的实施例中,对于“Breakdancer”序列,由于编码码率 r_d 和量化参数 Q_{step} 之间的关系表达式为 $r_d = k/Q_{step} + t$,因此可以通过线性回归的方法来拟合得到参数 k 为 0.9996 和 t 为 0.0040,也即对于左右视点其编码码率 r_d 和量化参数 Q_{step} 之间的关系表达式为 $r_d = 0.9996/Q_{step} + 0.0040$ 。例如,当 $Q_{step} = 16$ 时, $r_d = 0.0665$ bpp。

[0035] S222、多视点视频编码的失真计算表达式为 $PSNR_d = q_d \times QP + p_d$,其中参数 q_d 和 p_d 需要根据不同深度图序列具体设置。此时,在给定多视点深度图编码量化参数 QP 的条件下,可以得到多视点深度图编码码率 r_d 和失真 $PSNR_d$ 之间的对应关系。

[0036] 具体地,在本发明的实施例中,对于“Breakdancer”序列,由于多视点深度图编码的失真计算表达式为 $PSNR_d = q_d \times QP + p_d$,通过线性回归的方法来拟合得到参数 q_d 为 -0.65 和 p_d 为 63.75,也即对于左右视点其解码质量 $PSNR_d$ 和编码量化参数 QP 之间的关系可以表示为 $PSNR_d = -0.65 \times QP + 63.75$ 。当 $QP = 28$ 时, $PSNR_d = -0.65 \times 28 + 63.75 = 45.55$ dB。

[0037] 接下来,本发明需要建立了虚拟视点失真与多视点视频编码失真以及深度图编码失真之间的关系模型。在立体视频中,虚拟视点绘制是通过参考视点的深度图计算得到参考视点和虚拟视点间对应像素点的位置关系,然后通过多个参考视点的视频序列中对应像素点的亮度值和色度值来加权计算虚拟视点对应像素点的亮度值和色度值。因此,虚拟视点的失真是由于多视点视频编码失真和多视点深度图编码失真共同引起。因此在本发明中,为了便于建立虚拟视点失真与多视点视频和多视点深度图编码参数之间的函数关系,将首先分析在多视点深度图无失真的条件下,虚拟视点失真与多视点视频编码失真之间的

关系。其次,分析在多视点视频无失真的条件下,虚拟视点失真与多视点深度图编码失真之间的关系。最后建立虚拟视点失真与多视点视频和多视点深度图编码失真的联合关系模型。

[0038] 步骤 S303,建立多视点视频编码失真与虚拟视点绘制失真之间的关系。在该步骤中,设多视点深度图无失真,此时虚拟视点的失真仅由多视点视频编码失真引起。此时不妨设生成虚拟视点 C_v 需要 M 个参考视点 $\{C_{ref_i} | ref_i = 1, 2, \dots, M\}$ 。虚拟视点绘制图像为 M 个参考视点通过深度图绘制的虚拟视点图像的加权平均。因此,当第 i 个参考视点 C_{ref_i} 的视频编码失真为 $E_C^{ref_i}$ 时,参考视点 C_{ref_i} 的视频编码失真对于虚拟视点绘制图像失真 E_{CW} 的贡献为 $E_{CW}^{ref_i} = w_{ref_i} E_C^{ref_i}$, 其中, w_{ref_i} 为第 i 个参考视点视频编码失真对于虚拟视点图像绘制失真贡献的加权系数。考虑到通过参考视点的视频和深度图来绘制虚拟视点视频时,会出现遮挡等现象,因此在计算虚拟视点失真 $E_{CW}^{ref_i}$ 时还需要将遮挡部分对应的误差减去。可以设通过参考视点 C_{ref_i} 来绘制虚拟视点的过程中被遮挡的像素点的个数和参考视点整帧图像的像素点个数的比例为 η_{ref_i} , 此时通过 C_{ref_i} 绘制虚拟视点所产生的失真 $E_{CW}^{ref_i}$ 的计算表达式应该修正为 $E_{CW}^{ref_i} = \eta_{ref_i} w_{ref_i} E_C^{ref_i}$ 。因此,本发明通过 M 个参考视点加权绘制得到的虚拟视点图像失真可以表示为 $E_{CW} = \sum_{i=1}^M E_{CW}^{ref_i}$ 。

[0039] 具体地,在本发明的实施例中,如图 4 所示,为绘制虚拟视点的原理图。在该步骤中设多视点深度图无失真,此时虚拟视点的失真仅由多视点视频编码失真引起。假设左右参考视点视频的编码失真分别为 E_C^L 和 E_C^R 。当左参考视点的视频编码参数为 $QP = 28$ 时,左参考视点视频的编码失真的计算方法为 $E_C^L = \frac{255^2}{10^{((-0.37 \times 28 + 49.30)/10)}} = 8.3$ 。左参考视点的视频编码失真对于虚拟视点绘制图像失真 E_{CW} 的贡献为 $E_{CW}^L = w_L E_C^L$ 。

[0040] 同理,当右参考视点的视频编码参数为 $QP = 28$ 时,右参考视点的视频编码失真 $E_C^R = \frac{255^2}{10^{((-0.37 \times 28 + 49.30)/10)}} = 8.3$ 。右参考视点的视频编码失真对于虚拟视点绘制图像失真 E_{CW} 的贡献为 $E_{CW}^R = w_R E_C^R$ 。

[0041] 由于虚拟视点到左右参考视点的距离相同,因此,加权系数 w_L 和 w_R 满足 $w_L = w_R = 0.5$ 。考虑到通过参考视点的视频和深度图来绘制虚拟视点的视频时,会出现遮挡等现象,因此在计算虚拟视点失真 $E_{CW}^{ref_i}$ 时还需要将遮挡部分对应的误差减去。通过实验分析可知,对于“Breakdancer”序列,通过视点 4 和视点 6 来绘制虚拟视点的过程中被遮挡的像素点的个数和参考视点整帧图像的像素点个数的平均比例为 0.15, 此时由于左右参考视点编码失真所引起的虚拟视点图像失真可以表示为:

[0042] $E_{CW} = E_{CW}^L + E_{CW}^R = 0.15 \times 0.5 \times (E_C^L + E_C^R) = 0.15 \times 0.5 \times (8.3 + 8.3) = 1.25$ 。在本发明的一个实施例中,例如当 QP 取值为 28 时, $E_{CW} = 1.25$ 。

[0043] 步骤 S304,建立深度图编码失真与虚拟视点绘制失真之间的关系。在该步骤中,假设多视点视频无失真,此时虚拟视点的失真仅由多视点深度图编码失真引起。根据虚拟视点图像绘制的原理,虚拟视点绘制是通过深度信息和相机参数获得虚拟视点中的像素与参

考视点中的对应像素间的映射关系。该过程等价于将求出虚拟视点中各个像素与参考视点中对应像素之间的视差。当深度图出现失真时,虚拟视点绘制过程中对应像素的位置发生偏移,也即视差出现失真。在此,可以假设深度值为 D_j 的深度图的失真为 ΔD_j ,那么视差的失真为 $\|\Delta \mathbf{d}_j\| = \frac{\alpha |\Delta D_j|}{D_j^2}$ 。其中, α 为与相机参数相关的一个常数。

[0044] 为了更好地衡量视差失真对虚拟视点质量带来的影响,首先需要对参考视点 C_{ref_i} 的视频帧图像进行分解,使得分解后得到的每一个区域 B_i^j 中的像素点对应的深度值方差不差过预先设定的门限。为了实现上述分解,在本发明中,可将 C_{ref_i} 的视频帧图像对应的深度图进行四叉树分解,使得分解后每个区域 B_j 内深度值方差不超过预先设定的门限 D_{th} 。对于 B_i^j ,根据给定的深度图编码量化步长可以得到该区域深度值的平均量化失真 ΔD_j 和平均深度 D_j 。然后通过 ΔD_j 和 D_j 来得到视差的平均失真 $\|\Delta \mathbf{d}_j\|$ 。在得到 $\|\Delta \mathbf{d}_j\|$ 后,由于视差失真所造成的虚拟视点绘制图像上与参考视点中 B_i^j 区域相对应的区域的失真记为 e_{DW}^j ,则 e_{DW}^j 的计算表达式为 $e_{DW}^j = \|\Delta \mathbf{d}_i\| \psi_C^j$ 。其中, ψ_C^j 的计算公式为 $\psi_C^j = \frac{1}{2(2\pi)^2} \left(\iint S_C^j(\omega) \omega_1^2 d\omega + \iint S_C^j(\omega) \omega_2^2 d\omega \right)$, S_C^j 为参考视点视频帧图像上 B_i^j 区域的亮度值或色度值所组成的矩阵对应的傅立叶变换矩阵, $\omega = [\omega_1, \omega_2]$ 为 B_i^j 区域水平和垂直方向的角频率。因此,对于参考视点 C_{ref_i} ,其深度图量化所引入的失真 $E_{DW}^{ref_i} = \sum_j e_{DW}^j$ 。

同所述步骤 S303,虚拟视点 C_v 是通过 M 个参考视点 $\{C_{ref_i} | ref_i = 1, 2, \dots, M\}$ 加权绘制得到。通过 C_{ref_i} 来绘制虚拟视点的过程中,没有被遮挡的点与整帧图像全部像素个数的比例为 η_{ref_i} ,在生成虚拟视点图像中, C_{ref_i} 的加权系数为 w_{ref_i} ,所以 C_{ref_i} 对应的深度图编码失真对由深度图编码失真所引起的虚拟视点失真 E_{DW} 的贡献为 $\eta_{ref_i} \times w_{ref_i} \times E_{DW}^{ref_i}$,也即

$$E_{DW} = \sum_{i=1}^M \eta_{ref_i} \times w_{ref_i} \times E_{DW}^{ref_i}。$$

[0045] 具体地,本发明假设多视点视频无失真,此时虚拟视点的失真仅由多视点深度图编码失真引起。当深度值为 D_j 的深度图的失真为 ΔD_j 时,视差的失真为 $\|\Delta \mathbf{d}_j\| = \frac{\alpha |\Delta D_j|}{D_j^2}$ 。其中,对于“Breakdancer”序列的视点 4 和视点 6,上式中与相机参数相关的常数 α 为 8.46。

为了更好地衡量视差失真对虚拟视点质量带来的影响,首先需要对左右参考视点的视频帧图像进行分解,使得分解后得到的每一个区域中的像素点对应的深度值方差不超过预先设定的门限。为了实现上述分解,本发明中,可将左右参考视点的视频帧图像对应的深度图进行四叉树分解,使得分解后每个区域 $\{B_i^j | i = 4, 6; j = 1, 2, \dots, N_R\}$ 内深度值方差不超过预先设定的门限 $D_{th} = 10$,其中, N_R 为每个深度图分解后的最大区域数目。如图 5 所示,为本发明实施例的对于深度图的四叉树分解结果。

[0046] 对于 B_i^j ,根据给定的深度图编码量化步长可以得到该区域深度值的平均量化失真 $\Delta D_j = 2$ 和平均深度 $D_j = 92$ 。然后,通过 ΔD_j 和 D_j 可得到视差的平均失真

$\|\Delta \mathbf{d}_j\| = \frac{8.46 \times 2}{92^2} = 0.002$ 。对于大小为 8×8 的区域 B_i^j ，以其亮度为例，设其亮度矩阵为：

$$[0047] \quad \begin{bmatrix} 62 & 63 & 62 & 61 & 63 & 58 & 62 & 63 \\ 61 & 62 & 62 & 63 & 61 & 64 & 65 & 64 \\ 62 & 61 & 57 & 61 & 63 & 59 & 63 & 63 \\ 67 & 61 & 58 & 62 & 61 & 61 & 66 & 64 \\ 63 & 62 & 62 & 58 & 60 & 62 & 63 & 62 \\ 61 & 62 & 61 & 61 & 61 & 63 & 60 & 58 \\ 62 & 62 & 61 & 62 & 62 & 62 & 60 & 62 \\ 64 & 63 & 61 & 62 & 62 & 62 & 60 & 61 \end{bmatrix}。$$

[0048] 其傅里叶变换所对应的矩阵为 $S_C^{L_j}(\omega)$ 。为了计算视差失真所造成的虚拟视点绘制图像上与 B_i^j 区域相对应区域的失真 $e_{DW}^{L_j}$ ，还需要计算 $\psi_C^{L_j}$ 的取值。由于 $\psi_C^{L_j}$ 的计算公式为 $\psi_C^{L_j} = \frac{1}{2(2\pi)^2} \left(\iint S_C^{L_j}(\omega) \omega_1^2 d\omega + \iint S_C^{L_j}(\omega) \omega_2^2 d\omega \right)$ ，而 B_i^j 区域垂直和水平方向的角频率为

$$\omega = [\omega_1, \omega_2] = \left[\frac{2\pi}{8}, \frac{2\pi}{8} \right]，因此，\psi_C^{L_j} = 62。那么，e_{DW}^{L_j} 的值为 e_{DW}^{L_j} = \|\Delta \mathbf{d}_i\| \psi_C^{L_j} = 0.124。$$

[0049] 同理，也可以得到其他区域上由于深度图量化误差所引入的虚拟视点图像绘制失真，即， $E_{DW}^L = \sum_j e_{DW}^{L_j} = 32.24$ 。同样，对于右参考视点中的每一个区域 B_R^j 区域，可以通过上述计算方法获得该区域的失真 $e_{DW}^{R_j}$ 。此时，由于右参考视点其深度图量化所引入的虚拟视点量化失真为 $E_{DW}^R = \sum_j e_{DW}^{R_j} = 32.16$ 。同步骤 S303 一样，通过视点 4 和视点 6 来绘制虚拟视点的过程中，没有被遮挡的点与整帧图像全部像素个数的平均比例为 0.15，左右参考视点的加权系数为 $w_L = w_R = 0.5$ ，所以左右参考视点的深度图编码失真对由深度图编码失真所引起的虚拟视点失真 E_{DW} 可以表示为：

$$[0050] \quad E_{DW} = 0.15 \times 0.5 \times (E_{DW}^L + E_{DW}^R) = 0.15 \times 0.5 \times (32.24 + 32.16) = 4.83。$$

[0051] 步骤 S305，联合上述关系建立立体视频虚拟视点编码率失真模型。为了建立虚拟视点的率失真模型，需要将上述四个步骤中得到的多视点视频和多视点深度图的编码码率和失真模型合并，并最终得到虚拟视点编码的率失真模型。在该步骤中，为了建立虚拟视点绘制的编码率失真模型，需要分为以下两个步骤来实现：

[0052] 步骤 S501，首先需要建立虚拟视点的联合失真与编码量化参数之间的关系。由于多视点视频失真主要造成虚拟视点图像上的像素点亮度及色度数值产生失真，而多视点深度图失真主要造成虚拟视点图像上的像素点位置发生偏移，因此多视点视频失真和多视点深度图失真对于虚拟视点图像分别所引起的失真相互独立，也即虚拟视点图像失真 E_T 可以表示为由多视点视频失真所导致的虚拟视点绘制失真 E_{CW} 和多视点深度图失真所导致的虚拟视点绘制失真 E_{DW} 的和的形式。

[0053] 进一步地，通过左右两个视点的视频和深度图来生成虚拟视点的过程中，由于深

度图的不连续性,在虚拟视点图像中仍然可能存在少量的空洞像素点,通过图像补绘来填补这些空洞所引起的误差用 E_0 来表示。由于图像具有一定的连续性,可以通过与被遮挡像素相邻的像素点获得被遮挡像素点 i_0 的亮度或色度值分布 $P_{i_0}^o$ 。该像素点由于补绘所带来的亮度或色度的误差可以表示为 $e_{i_0}^o = \sum_{y \neq y_{in}} P_{i_0}^o(y) \times (y - y_{in})$,其中 y 为像素点 i_0 的亮度或色度可能取值范围, y_{in} 为像素点 i_0 补绘时的亮度或色度取值。那么,图像补绘所带来的虚拟视点图像误差可以表示为 $E_0 = \sum_{i_0} e_{i_0}^o = \sum_{i_0} \sum_{y \neq y_{in}} P_{i_0}^o(y) \times (y - y_{in})$,其中, i_0 为被遮挡像素点, $P_{i_0}^o$ 为被遮挡像素点 i_0 的相邻像素点的亮度或色度值分布, y 为像素点 i_0 的亮度或色度的取值范围, y_{in} 为被遮挡像素点 i_0 补绘时的亮度或色度值。此时,虚拟视点绘制图像的误差 E_T 的计算表达式为 $E_T = E_{CW} + E_{DW} + E_0$ 。

[0054] 具体地,对于本样例中的虚拟视点, $E_0 = 0.13$ 。因此,虚拟视点绘制图像的误差 E_T 的计算表达式为 $E_T = E_{CW} + E_{DW} + E_0 = 1.25 + 4.83 + 0.13 = 6.21$,也即虚拟视点图像质量为 40.2dB。

[0055] 步骤 S502,建立虚拟视点绘制所需要的编码码率与编码量化参数之间的关系。由于在此前定义的立体视频系统中,多视点视频和多视点深度图是由多视点视频编码器分别进行编码,其编码量化参数也是独立选取的。因此,立体视频编码所需的码率 r_t 为多视点视频编码码率 r_c 和多视点深度图编码码率 r_d 之和,也即 $r_t = r_c + r_d$ 。

[0056] 具体地,在本发明实施例中,

[0057] 当 $QP = 28$ 时, $r_t = r_c + r_d = 0.096\text{bpp} + 0.0665\text{bpp} = 0.1625\text{bpp}$ 。

[0058] 最后通过统计不同 QP 下虚拟视点图像绘制的失真 E_T 和多视点视频以及多视点深度图编码所需的总码率 r_t 可以得到立体视频编码中虚拟视点图像绘制的率失真模型。

[0059] 通过本发明提出的立体视频中虚拟视点绘制编码率失真分析方法,能够准确快速地估计立体视频中虚拟视点绘制编码率失真性能,从而为立体视频编码参数的选择与码率分配等问题提供了模型指导和解决方案,进一步提高了立体视频编码的效率。

[0060] 尽管已经示出和描述了本发明的实施例,对于本领域的普通技术人员而言,可以理解在不脱离本发明的原理和精神的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由所附权利要求及其等同限定。

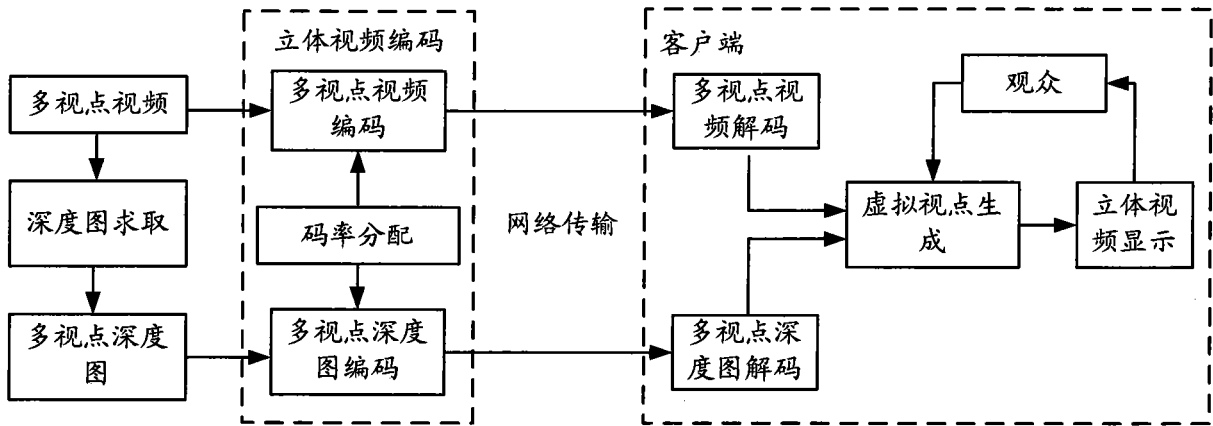


图 1

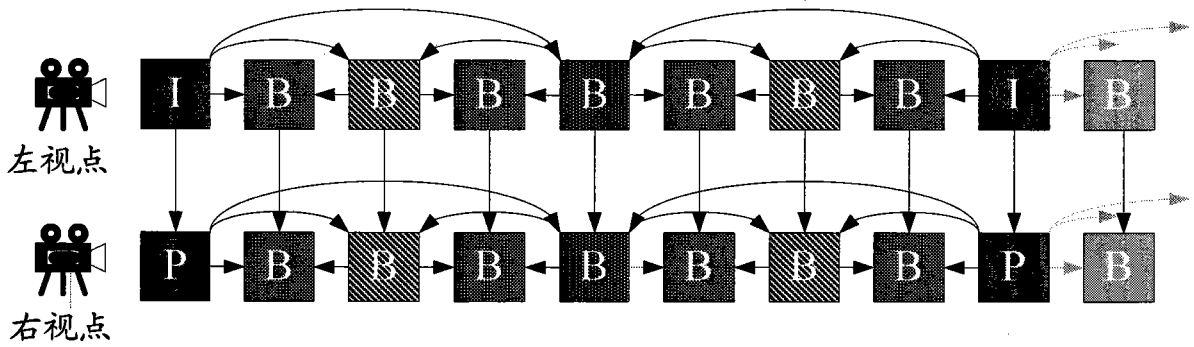


图 2

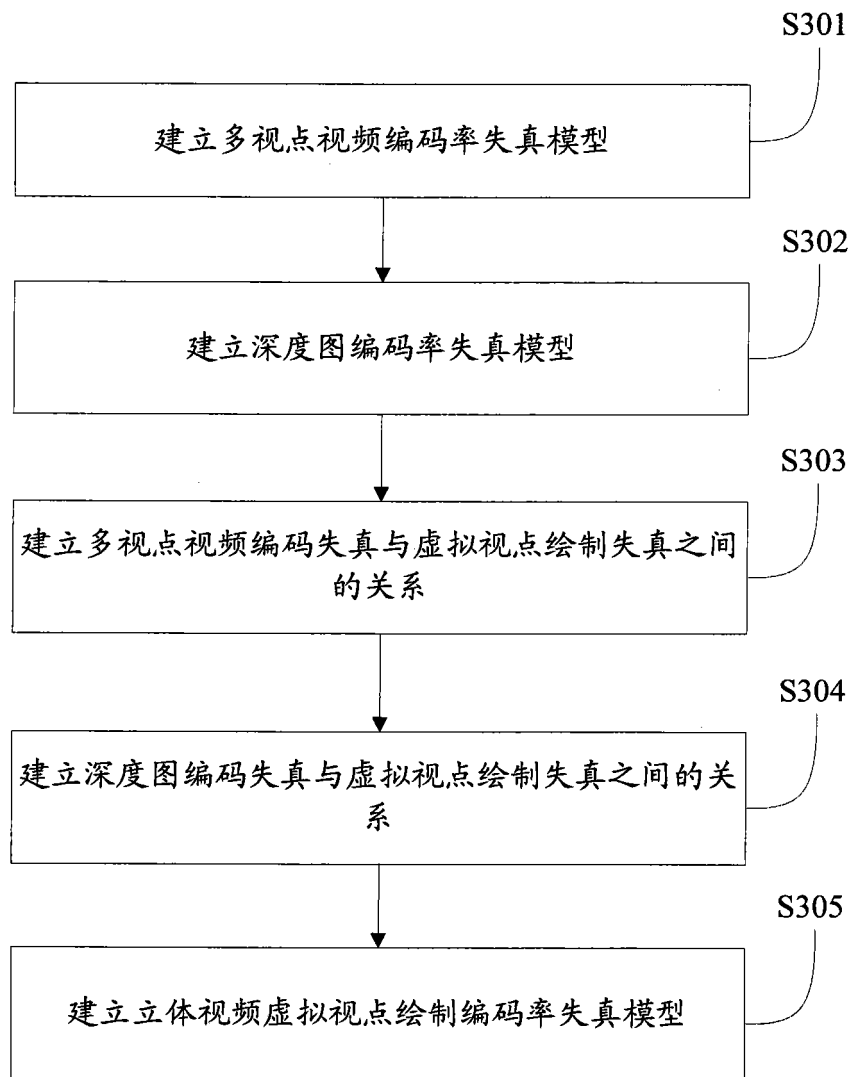


图 3

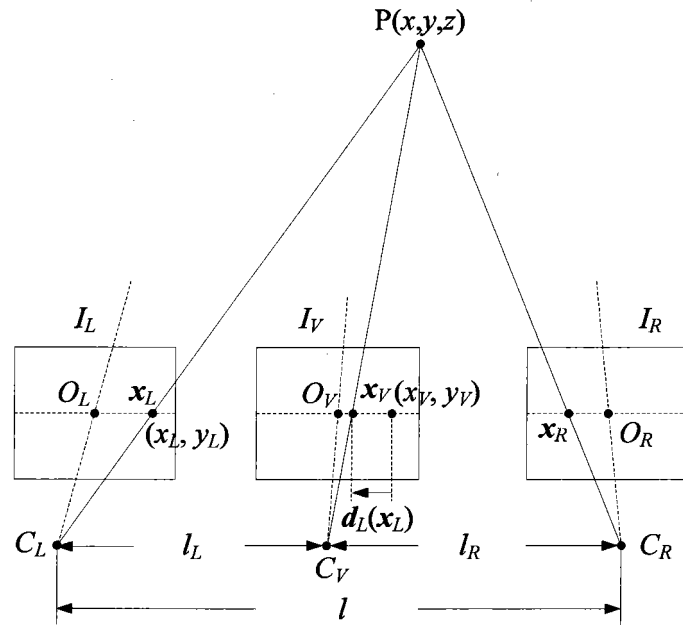


图 4

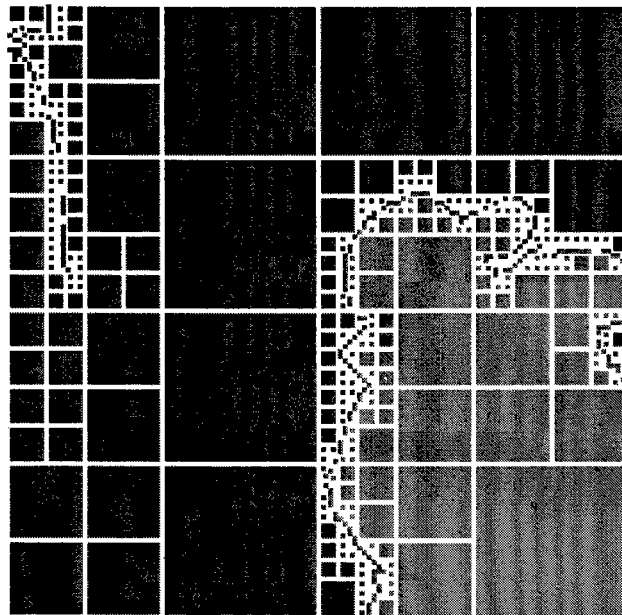


图 5