

引入拉格朗日算子的纹理深度联合比特分配方法

杨旭, 李兵兵, 常义林

(西安电子科技大学综合业务网理论及关键技术国家重点实验室, 710071, 西安)

摘要: 为了提高三维视频合成质量, 提出一种新的纹理视频和深度图编码比特分配方法. 该方法在比特分配中引入拉格朗日算子参量, 从而在限制比特下进一步降低虚拟合成误差. 首先依据统计实验并采用最小二乘拟合的方法, 对纹理视频和深度图分别建立以拉格朗日算子为参数的编码误差模型和编码比特模型, 然后利用编码误差和虚拟合成误差的相关性进一步推导出基于拉格朗日算子的虚拟合成误差的模型, 最后结合所推导的模型建立虚拟合成误差最小化目标函数, 利用无约束优化方法求解最优的纹理视频和深度图编码拉格朗日算子, 并依据此算子进行比特分配. 实验结果证明: 与固定比特算法和全搜索量化步长算法相比, 该方法在不同的编码条件下使合成视频质量平均提高 0.25 dB 和 0.68 dB, 且得到的编码总比特与目标比特最为接近, 平均误差仅为 0.45%.

关键词: 三维视频; 拉格朗日算子; 比特分配; 虚拟合成误差

中图分类号: TN911.7 **文献标志码:** A **文章编号:** 0253-987X(2012)12-0042-07

Joint Bit Allocation Method for Texture Videos and Depth Maps with Lagrange Multiplier

YANG Xu, LI Bingbing, CHANG Yilin

(State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China)

Abstract: A novel joint bit allocation method for the texture videos and depth maps coding is proposed to improve the synthesis quality of 3D video (3DV). The contribution of the proposed method is that the Lagrange multiplier is introduced into the bit allocation as a parameter to further reduce the virtual synthesis distortion under the constrained bit. Both distortion model and bit model for texture videos and depth maps are constructed through statistical experiments, and the Lagrange multiplier is treated as a parameter in both of the models. Then, a model of warping errors is deduced by a comprehensive analysis on the relationship between the coding distortion and warping error, and an objective function for minimizing the warping error is designed using this model. A constrained optimization method is used to obtain the optimal Lagrange multipliers of the texture video and depth map coding, and then the bit is allocated using the resulting Lagrange multipliers. Experimental results verify that compared with the full search based bit allocation algorithm and fixed bit algorithm, the proposed method achieves an average gain of 0.25 dB and 0.68 dB under different coding conditions, respectively. Meanwhile, the coding bit of the proposed method is closer to the constrained bit, and the average error is only 0.45%.

Keywords: 3D video; Lagrange multiplier; bit allocation; warping error

收稿日期: 2012-04-12. 作者简介: 杨旭(1984—), 女, 博士生; 常义林(通信作者), 男, 教授, 博士生导师. 基金项目: 国家自然科学基金资助项目(61001205, 60902052, 60902081, 60772134); 高等学校学科创新引智计划资助项目(B08038); 中央高校基本科研业务费资助项目(72105457); 西电-中兴联合研究资助项目.

网络出版时间: 2012-10-12

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20121012.1142.005.html>

三维视频(3DV)因其能为观众提供更真实的观感,已受到广泛的关注^[1-3],被应用到众多领域,如三维电视(3DTV)^[4]、任意视点电视(FTV)^[5]、计算机游戏和虚拟现实等。

在现有的研究中,3DV 有多种表示方式^[6-7],其中以多视点视频加深度(MVD)的表示方式^[7]最为简便且适合编码,因此被广泛应用。MVD 数据经过编码(如 H. 264/MVC^[8])压缩为码流,并传输到接收端。在接收端,解码器重建左右视点的 MVD 数据,并用基于深度图的描述方法(DIBR)^[9]对其进行合成得到虚拟视点视频。因此,虚拟视点视频的合成误差主要受纹理视频和深度图编码误差的影响,并且随着纹理视频和深度图的编码比特分配而变化。为了降低虚拟合成误差,已有大量文献对纹理和深度的比特分配问题进行了研究^[9-12]。Fehn 首先提出固定纹理和深度的比特分配方法^[9],该方法限定纹理和深度的编码比特之比为 5:1,但固定的比特分配并不适合所有序列和目标速率。Movarn 提出量化参数(QP)全搜索的算法^[10],该算法遍历所有纹理视频量化参数 Q_t 和深度图编码量化参数 Q_d ,在限定的目标比特率下寻找使虚拟合成误差最小的量化参数对(Q_t - Q_d),能够达到较优的虚拟视频质量,但复杂度较高。为了降低算法复杂度,Liu 和 Yuan 分别提出了 2 种虚拟合成误差模型,根据模型计算最优的 Q_t - Q_d 参数对^[11-12]。但是,以上纹理深度的联合比特分配方法仅考虑对 Q_t - Q_d 参数对的优化,而目前的视频编码标准(H. 264)的核心是率失真优化算法,且拉格朗日算子是率失真代价函数的重要参数,所以它对纹理视频和深度图编码比特分配亦会产生重要影响。

针对以上问题,本文提出一种新的纹理深度联合比特分配方法。该方法充分考虑到拉格朗日算子对虚拟合成误差的影响,推导出基于拉格朗日算子的虚拟合成误差模型;构建无约束虚拟误差最小化目标函数,优化纹理和深度编码拉格朗日算子对,以达到纹理和深度编码比特的最优分配。

1 虚拟视点合成及虚拟误差

目前广泛应用虚拟视点的合成技术是 DIBR 技术,该技术利用解码重建的左右视点纹理视频和深度图合成虚拟视图,其示意图如图 1 所示。

图中 M 点为空间点,其映射到解码重建的左右视点上的点为 m_L 和 m_R ,这 2 点所对应的亮度值分别为 $I_L(x_L, y_L)$ 和 $I_R(x_R, y_R)$ 则虚拟合成点 m_v 的

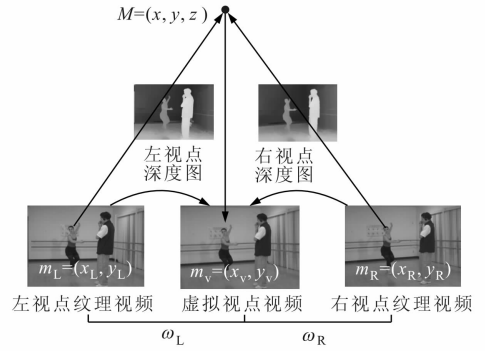


图 1 虚拟视点合成示意图

亮度值 $I'(x_v, y_v)$ 可由下式得到

$$I'(x_v, y_v) = \omega_L I_L(x_L, y_L) + \omega_R I_R(x_R, y_R) \quad (1)$$

式中: ω_L 和 ω_R 表示基于距离的合成权重^[11],且 $\omega_L + \omega_R = 1$ 。虚拟合成误差 D_v 是指合成的虚拟视图与原始的视点视图之间的均方误差,如下式所示

$$D_v = \frac{1}{WH} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} |I_v(x, y) - I'_v(x, y)|^2 \quad (2)$$

式中: $I_v(x, y)$ 和 $I'_v(x, y)$ 分别表示原始视点和合成视点在 (x, y) 处的亮度值; W 、 H 分别表示图像的宽和高。导致虚拟合成误差的因素除了深度图估计的准确性和摄像机参数影响外,更主要的是纹理视频编码和深度图的编码误差,所以在深度图估计和摄像机参数矫正准确的条件下, D_v 模型^[12] 可用纹理视频和深度图的编码误差 D_t 、 D_d 表示

$$D_v = AD_t(R_t) + BD_d(R_d) + C \quad (3)$$

式中: R_t 和 R_d 分别表示纹理视频和深度图的编码比特; A 、 B 、 C 表示模型参数。在总的限制比特 R_c 的条件下,最小化 D_v 的算法可演变为约束优化问题

$$\begin{aligned} & \min_{R_t, R_d} D_v \\ & \text{s. t. } R_t + R_d \leq R_c \end{aligned} \quad (4)$$

2 基于拉格朗日算子的 D_v 最小化问题

2.1 影响 D_v 的参数

由于 D_v 主要由纹理视频和深度图的编码误差所决定,所以影响编码误差的参数已成为最小化 D_v 问题的关键。目前 3DV 编码技术中纹理视频和深度图分别基于 H. 264/MVC 标准进行编码。H. 264 中核心算法为率失真优化算法,该算法可使在限制目标比特率下编码失真最小,其表达式如下

$$\begin{aligned} & \min J(\mathbf{S}, \mathbf{I} | \mathbf{Q}, \lambda) = \\ & \min(D(\mathbf{S}, \mathbf{I} | \mathbf{Q}) + \lambda R(\mathbf{S}, \mathbf{I} | \mathbf{Q})) \end{aligned} \quad (5)$$

式中: $J(\mathbf{S}, \mathbf{I} | \mathbf{Q}, \lambda)$ 表示拉格朗日代价函数; $D(\mathbf{S},$

$I|Q$ 和 $R(S, I|Q)$ 分别表示编码失真和编码比特; λ 和 Q 分别表示拉格朗日算子和量化步长; S 表示纹理视频或深度图所需的编码数据; I 表示编码模式信息。

在 H. 264 标准中,每个宏块均采用率失真优化算法,在给定 λ 和 Q 的条件下,选出使代价函数最小的 $D(S, I|Q)$ 和 $R(S, I|Q)$,因此 λ 和 Q 会直接影响编码失真和编码比特。随参数 Q 增大,编码保真度下降,从而导致编码失真的增加和编码比特的下降,反之亦然。参数 λ 会影响每个宏块最优模式的选择。此外, λ 对运动向量(MV)的选择也起着重要的作用。下式为最优 MV 选择的计算式

$$\{(M_v, R_{ef})_{opt} | I\} =$$

$$\operatorname{argmin}\{D_{fd}(M_v, R_{ef}) + \lambda^{1/2}R(M_v, R_{ef})\} \quad (6)$$

式中: R_{ef} 表示参考帧索引; M_v 表示 MV 信息; $D_{fd}(M_v, R_{ef})$ 表示运动补偿误差。从式(6)可看出, λ 将直接影响 MV 的选择,进而影响编码失真和比特的值。由此推得 D_v 受纹理视频和深度图编码的量化步长 Q_t 和 Q_d 以及拉格朗日算子 λ_t 和 λ_d 的影响。式(4)可重新定义为

$$\min_{Q_t, Q_d, \lambda_t, \lambda_d} D_v(Q_t, Q_d, \lambda_t, \lambda_d)$$

$$\text{s. t. } R_t(\lambda_t, Q_t) + R_d(\lambda_d, Q_d) < R_c \quad (7)$$

$$D_v = A D_t(Q_t, \lambda_t) + B D_d(Q_d, \lambda_d) + C \quad (8)$$

2.2 基于拉格朗日算子的虚拟失真模型

为了降低式(7)的复杂度,将其分 2 步计算:首先固定 λ_t 和 λ_d 的值,计算使得 D_v 最小的 Q_t - Q_d 参数对;其次基于该 Q_t - Q_d 参数对,对 λ_t 和 λ_d 进行计算,得到 λ_t - λ_d 最优组合。由于计算 Q_t - Q_d 参数对的方法已有大量文献^[10-13]提出,因此本文主要研究如何在最优 Q_t - Q_d 下,利用 λ_t - λ_d 组合进一步降低 D_v 。因此,式(7)可简化为

$$\min_{\lambda_t, \lambda_d} D_v(\lambda_t, \lambda_d | Q_t, Q_d)$$

$$\text{s. t. } R_t(\lambda_t | Q_t) + R_d(\lambda_d | Q_d) < R_c \quad (9)$$

为了定量研究在固定 Q_t 和 Q_d 下, λ_t - D_t 、 λ_d - D_d 、 λ_t - R_t 以及 λ_d - R_d 的模型,本文对 Balloons 序列 view3 的纹理视频和深度图分别进行实验。实验中,在每个固定的 Q_t 或 Q_d 下,分别用一组 λ_t 或 λ_d 对纹理视频和深度图进行编码,并得到对应的 D_t 和 R_t (D_d 和 R_d)。实验结果表明,对于每个固定的 Q_t (Q_d), λ_t 与 D_t (λ_d 与 D_d)呈指数关系,如图 2a~2f 所示; λ_t 与 R_t (λ_d 与 R_d)呈负幂次关系,如图 2g~2l 所示。

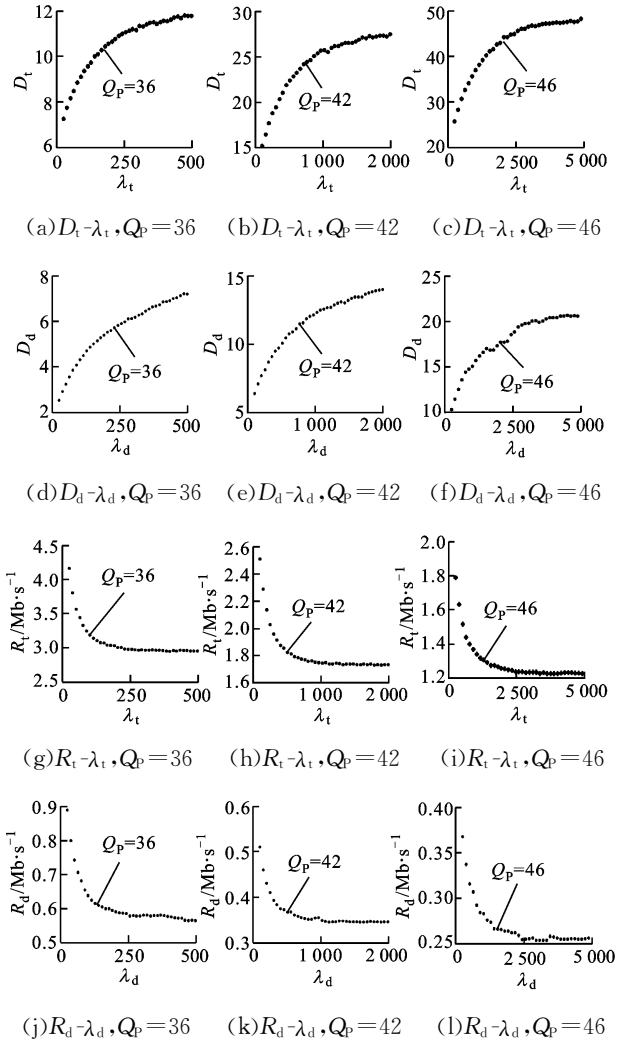


图2 不同 Q_p 下 D_t - λ_t 、 D_d - λ_d 、 R_t - λ_t 、 R_d - λ_d 模型的统计曲线

通过拟合可得 4 个模型曲线的幂次随 Q_t 和 Q_d 值的变化量很小,可近似认为定值。因此,本文首先用最小二乘拟合的方法分别得到 λ_t - D_t 和 λ_d - D_d 模型的幂次,并推出 λ_t - D_t 和 λ_d - D_d 模型

$$\begin{cases} D_t(\lambda_t) = a_t \lambda_t^{0.25} + b_t \\ D_d(\lambda_d) = a_d \lambda_d^{0.5} + b_d \end{cases} \quad (10)$$

式中: a_t 、 b_t 、 a_d 和 b_d 表示纹理视频和深度图失真模型参数。图 3 表示式(10)模型所计算的数据与真实数据的对比。实验采用序列 Balloons、Newspaper 和 Lovebird 的左视点纹理视频和深度图。从图中可看出,提出的模型能够很好地逼近真实数据。

同理,拟合出 λ_t - R_t 和 λ_d - R_d 的模型表示如下

$$\begin{cases} R_t(\lambda_t) = \alpha_t \lambda_t^{-1} + \beta_t \\ R_d(\lambda_d) = \alpha_d \lambda_d^{-0.5} + \beta_d \end{cases} \quad (11)$$

式中: α_t 、 β_t 、 α_d 和 β_d 分别表示纹理视频和深度图的编码比特模型参数。式(11)的拟合比特数据与真实

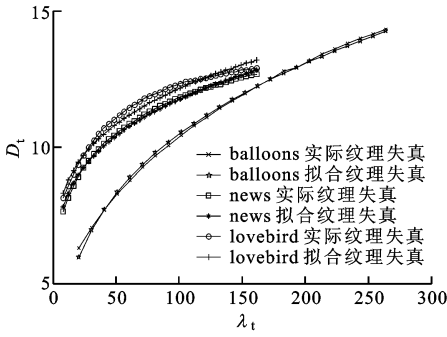
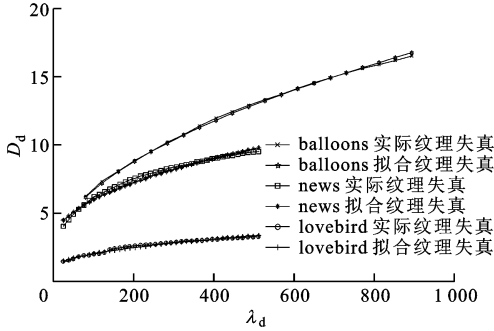
(a) $D_t - \lambda_t$ 模型与真实数据比较(b) $D_d - \lambda_d$ 模型真实数据比较

图 3 拟合失真数据与真实失真数据的比较

数据的比较如图 4 所示. 从图 4 中可看出, 模型 $\lambda_t - R_t$ 和 $\lambda_d - R_d$ 能够很好地逼近真实数据.

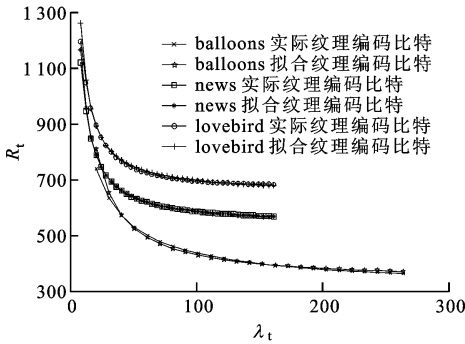
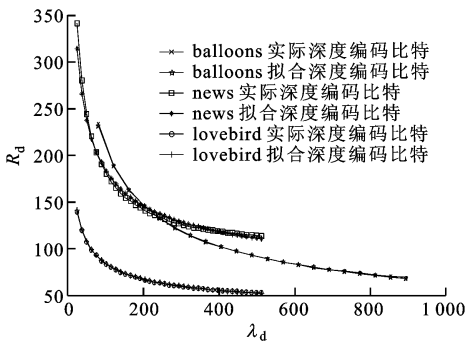
(a) $R_t - \lambda_t$ 模型与真实数据比较(b) $R_d - \lambda_d$ 模型与真实数据比较

图 4 拟合比特数据与真实比特数据的比较

因为 D_v 是由左右视点的纹理视频和深度图编码质量所决定的^[12], 所以式(7)可写成

$$D_v = \omega_L^2(D_t^L + \Theta D_d^L) + \omega_R^2(D_t^R + \Theta D_d^R) + C \quad (12)$$

式中: D_t^L 、 D_t^R 、 D_d^L 和 D_d^R 分别表示左右视点纹理视频和深度图的编码误差; Θ 表示深度图编码误差对合成虚拟视图误差的线性影响因子

$$\Theta = (f/l/Z)^2 \Psi \quad (13)$$

式中: f 和 l 分别表示摄像机的焦距和摄像机的间距; Ψ 表示由图像内容所决定的常数参量^[12]. 将式(10)、式(11)带入式(12), 可得到在固定 Q_t 、 Q_d 条件下, 基于拉格朗日算子的 D_v 模型如下

$$D_v = \omega_L^2(a_t^L(\lambda_t^L)^{0.25} + \Theta a_d^L(\lambda_d^L)^{0.5}) + \omega_R^2(a_t^R(\lambda_t^R)^{0.25} + \Theta a_d^R(\lambda_d^R)^{0.5}) + C \quad (14)$$

式中: a_t^L 、 a_t^R 和 a_d^L 、 a_d^R 分别表示左右视点纹理视频和深度图的模型参数; λ_t^L 、 λ_t^R 和 λ_d^L 、 λ_d^R 分别表示编码左右视点纹理视频和深度图的拉格朗日算子.

3 基于拉格朗日算子的比特分配

根据所推导的 D_v 模型, 提出一种新的纹理和深度联合的比特分配方法. 在最优 Q_t 、 Q_d 参数对条件下, 利用 D_v 模型求得最优的纹理和深度编码拉格朗日算子, 并对纹理视频和深度图编码比特进行再分配, 以达到降低虚拟合成误差的目的.

对 D_v 的最小化问题即为对式(14)的优化问题, 它可分解为对其第一项式和第二项式分别优化的问题. 首先对第一项式求最小值, 该项所表示的物理意义为左视点纹理视频和深度图所导致的虚拟合成误差, 其最小化目标函数如下

$$\begin{aligned} & \min_{\lambda_t^L, \lambda_d^L} (a_t^L(\lambda_t^L)^{0.25} + \Theta a_d^L(\lambda_d^L)^{0.5}) \\ \text{s. t. } & a_t^L(\lambda_t^L)^{-1} + \alpha_d^L(\lambda_d^L)^{-0.5} + \beta^L - R_c^L \leq 0 \end{aligned} \quad (15)$$

式中: R_c^L 表示左视点的限制比特; β^L 表示 β_d^L 与 β_t^L 之和. 式(15)可转换为无约束优化问题, 即

$$\begin{aligned} \min_{\lambda_t^L, \lambda_d^L} F^L = & \{a_t^L(\lambda_t^L)^{0.25} + \Theta a_d^L(\lambda_d^L)^{0.5} + m(a_t^L(\lambda_t^L)^{-1} + \alpha_d^L(\lambda_d^L)^{-0.5} + \beta^L)\} \end{aligned} \quad (16)$$

式中: m 表示约束因子. 本文中, Θ 取经验值 0.3, R_c^L 可通过 δR_c 得到, R_c 表示总的限制比特, δ 表示左视点编码比特占总比特的百分比, δ 值取多个序列测试得到的平均值 0.6. 对式(16)求偏导, 可得到最优的 λ_t^L 和 λ_d^L 值

$$\left. \begin{aligned} \frac{\partial F^L}{\partial \lambda_t^L} &= 0.25a_t^L(\lambda_t^L)^{-0.75} - m\Theta a_t^L(\lambda_t^L)^{-2} = 0 \\ \frac{\partial F^L}{\partial \lambda_d^L} &= 0.5a_d^L(\lambda_d^L)^{-0.5} - 0.5m\Theta a_d^L(\lambda_d^L)^{-1.5} = 0 \\ \frac{\partial F^L}{\partial m} &= \alpha_t^L(\lambda_t^L)^{-1} + \alpha_d^L(\lambda_d^L)^{-0.5} + \beta^L - R_c^L = 0 \end{aligned} \right\} \quad (17)$$

对式(14)的第二项式的最优化算法与第一项式相同,其物理意义是右视点纹理视频和深度图所导致的虚拟合成误差,通过对其优化可得到最优的 λ_t^R 和 λ_d^R 值。

本文所提的比特分配方法可归结为如下步骤:首先通过全搜索算法^[10]得到最优的 Q_t - Q_d 参数对,并计算出 R_c^L 和 R_c^R ;其次通过预编码方法求出左右视点纹理及深度图的模型参数,模型参数的计算方法以左视点纹理视频为例,分别用 $\lambda_{t,1}$ 和 $\lambda_{t,2}$ 对纹理视频进行编码,并得到相应的编码比特 $R_{t,1}$ 和 $R_{t,2}$,则 α_t^L 和 β^L 可由下式算出

$$\left. \begin{aligned} R_{t,1} &= \alpha_t^L(\lambda_{t,1})^{-1} + \beta^L \\ R_{t,2} &= \alpha_t^L(\lambda_{t,2})^{-1} + \beta^L \end{aligned} \right\} \quad (18)$$

那么,左视点剩余的模型参数 α_t^L 、 α_d^L 、 α_d^R 、 β_d^R ,以及右视点的模型参数 α_t^R 、 α_t^R 、 β_t^R 、 α_d^R 、 α_d^R 和 β_d^R 可利用类似的方法计算得出;最后,通过式(17)计算最优的 λ_t^L 、 λ_d^L 、 λ_t^R 和 λ_d^R ,并将它们的作为编码参数对左右视点纹理视频及深度图分别进行编码。图 5 给出本文方法的流程图。

4 实验结果及分析

为了验证本文比特分配方法的性能,本节用 3DV 序列 Balloons、Newspaper 和 Lovebird1 进行

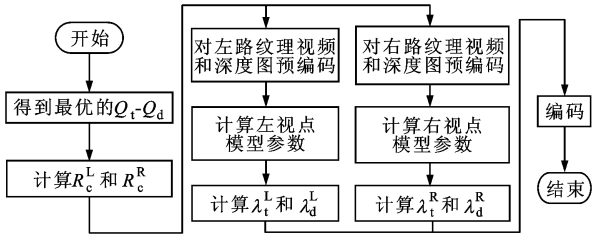


图 5 本文比特分配方法的流程图

实验,序列分辨率为 $1\,024\times 768$ 像素.实验所用的参考软件为 JMVC8.5,并在 2 种编码条件下进行.条件 1 为:分级 B 帧编码结构;全搜索运动估计;搜索范围为 96×96 像素;2 个参考帧;基于内容的自适应二进制编码(CABAC).条件 2 为:分级 B 帧编码结构;全搜索运动估计;搜索范围为 32×32 像素;基于内容的自适应可变速长编码(CAVLC).

在编码条件 1 和 2 下,分别对本文方法与固定比特分配方法^[9]以及目前效果最优的全搜索方法^[10]进行编码实验,结果如表 1 所示.表 1 中,峰值信噪比表示合成的虚拟视点的质量,比特率表示纹理视频和深度图编码比特之和, R_c 表示限制目标比特.其中虚拟视点合成的应用软件为 VS3.5.用 p_1 和 p_2 分别表示本文算法与全搜索算法和固定比特算法相比的加权平均峰值信噪比(BD PSNR)的增长值和减少值,采用 b_1 和 b_2 分别表示本文算法与全搜索算法和固定比特算法相比,其加权平均比特率(BD bitrate)的下降度和增长率. p_1 、 p_2 为正值且 b_1 、 b_2 为负值时表示本算法优于其他 2 种算法,反之亦然.计算 BD PSNR 和 BD bitrate 的算法由文献^[13]给出.实验计算的结果如表 2 所示.

表 1 3 种方法的编码实验结果

序列 (参数)	比特率/kb · s ⁻¹						峰值信噪比/dB						R _c / (kb · s ⁻¹)
	编码条件 1			编码条件 2			编码条件 1			编码条件 2			
	固定比 特算法	全搜索 算法	本文 方法	固定比 特算法	全搜索 算法	本文 方法	固定比 特算法	全搜索 算法	本文 方法	固定比 特算法	全搜索 算法	本文 方法	
Balloons (300 帧, 左视点 3, 右视点 5)	947.2	942.0	938.6	919.2	938.9	936.0	36.2	36.5	36.4	35.3	35.3	35.4	940
	581.2	609.2	599.9	607.4	608.1	597.8	33.8	34.2	34.2	32.2	32.5	32.7	600
	436.9	441.8	436.4	440.1	440.0	429.9	32.1	32.1	32.6	27.9	28.2	29.1	430
	333.4	325.7	320.5	330.1	328.1	322.4	28.0	28.2	28.9	26.5	26.8	27.4	320
Newspaper (300 帧, 左视点 4, 右视点 6)	1 339.9	1 287.6	1 299.4	1 310.8	1 305.3	1 296.2	35.0	35.1	35.1	33.5	34.5	34.7	1 300
	836.1	778.6	801.3	808.7	800.8	803.3	33.7	33.5	33.7	32.3	32.8	32.9	800
	536.1	521.0	525.4	534.3	529.2	523.6	32.0	32.0	32.1	30.5	30.6	30.8	525
	400.2	389.9	399.5	403.6	397.6	394.9	30.6	30.6	30.8	28.6	28.5	28.9	400
Lovebird1 (240 帧, 左视点 6, 右视点 8)	1 262.5	1 242.7	1 247.7	1 276.2	1 245.8	1 247.7	36.4	36.6	36.5	35.5	36.1	36.3	1 250
	716.3	745.3	753.8	766.8	745.2	744.3	34.7	34.8	35.0	34.2	34.3	34.5	750
	469.8	492.5	498.2	513.6	502.4	498.1	33.3	33.4	33.6	32.7	32.6	33.0	500
	376.8	368.2	371.3	361.6	379.3	373.3	32.2	32.2	32.4	30.9	31.1	31.5	375

表 2 3 种方法的性能比较

序列	p_1/dB		p_2/dB		$b_1/\%$		$b_2/\%$	
	编码条件 1	编码条件 2	编码条件 1	编码条件 2	编码条件 1	编码条件 2	编码条件 1	编码条件 2
Balloons	0.34	0.5	0.49	0.77	-6.66	-6.86	-6.76	-8.9
Newspaper	0.06	1.75	0.16	0.6	-1.48	-39.27	-3.08	-10.42
Lovebird1	0.18	0.08	0.25	0.41	-3.93	-2.22	-6.26	-9.85
平均	0.19	0.78	0.30	0.59	-4.02	-16.12	-5.36	-9.72

从表 2 可看出:本文方法的平均质量在编码条件 1 时分别比全搜索算法和固定比特算法提高了 0.19 dB 和 0.78 dB,在编码条件 2 时,分别提高了 0.30 dB 和 0.59 dB,2 种编码条件的平均值提高了 0.25 dB 和 0.68 dB;平均比特在编码条件 1 时分别比全搜索算法和固定比特算法降低了 4.02% 和 16.12%,在编码条件 2 时,平均比特分别降低 5.36% 和 9.72%。由此可看出,本文方法在不同的编码环境下均可有效地提高虚拟视点的合成质量和 3D 编码效率。

为了进一步验证所提方法的有效性,本文引入失配度(mismatch) E 来表示实际的编码比特 R_A 与目标编码比特 R_C 的匹配程度

$$E = \frac{R_A - R_C}{R_C} \times 100\%$$

(19)

并用 A_e 表示所有误差 E 的绝对值平均值。失配度能够验证比特分配方法的有效性, A_e 越低表明编码

比特越接近目标比特,方法有效性越高。表 3 列出了 3 种比特分配算法在不同编码条件下的 E 值和 A_e 。从表 3 中可看出,本文方法的总比特与目标比特最为接近, A_e 在编码条件 1 时为 0.36%,在编码条件 2 时为 0.53%,其平均值仅为 0.45%,比全搜索算法和固定比特算法的平均值低 0.75% 和 1.86%。

5 结 论

为了降低虚拟合成误差,本文分析了拉格朗日算子与纹理视频和深度图的编码质量和比特之间的关系,推导出基于拉格朗日算子的虚拟合成误差模型,并提出了一种新的纹理视频和深度图联合比特分配方法。该方法利用推导出的虚拟合成误差模型,结合无约束最优化算法对拉格朗日算子进行优化。实验结果表明,与全搜索算法和固定比特算法相比,本文提出的比特分配方法不仅能有效地减小虚拟合成误差值,而且具有较优的比特分配特性。

表 3 3 种方法的编码失配度对比

序列	$E/\%$					
	编码条件 1			编码条件 2		
	固定比特算法	全搜索算法	本文算法	固定比特算法	全搜索算法	本文算法
Balloons	0.76	0.22	-0.15	-0.10	-2.2	-0.42
	1.54	-3.13	-0.02	1.35	1.24	-0.36
	2.76	1.60	1.49	2.33	2.36	-0.01
	1.79	4.20	0.18	2.53	3.17	0.76
Lovebird1	-0.95	3.07	-0.05	-0.34	2.09	0.34
	-2.67	4.52	0.16	-0.64	2.25	0.64
	-0.75	2.13	0.07	0.49	2.72	-0.48
	-2.51	0.06	-0.12	1.16	-3.57	-1.15
Newspaper	-0.58	1.00	-0.18	0.41	0.84	-0.29
	-0.62	-4.49	0.51	0.10	1.22	0.41
	-1.49	-6.04	-0.34	0.80	1.77	-0.26
	-1.80	0.48	-0.97	-0.58	0.92	-1.26
$A_e/\%$	1.50	2.58	0.36	0.90	2.03	0.53

参考文献:

- [1] SMOLIC A, MUELLER K, STEFANOSKI N, et al. Coding algorithms for 3DTV: a survey [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(11): 1606-1621.
- [2] 刘晓仙,常义林,杨海涛,等. 任意视点电视系统中深度数据的自适应非均匀量化方法 [J]. 西安交通大学学报, 2010, 44(2): 82-87.
LIU Xiaoxian, CHANG Yilin, YANG Haitao, et al. Adaptive non-uniform quantization of depth data in free-viewpoint television system [J]. Journal of Xi'an Jiaotong University, 2010, 44(2): 82-87.
- [3] 邓智玘,贾克斌,陈锐霖,等. 三维视频编码技术的发展与挑战 [J]. 计算机应用, 2011, 31(9): 2543-2546.
DENG Zhipin, JIA Kebin, CHEN Ruilin, et al. New trend and challenges in 3D video coding [J]. Journal of Computer Applications, 2011, 31(9): 2543-2546.
- [4] ONURAL L, GOTCHEV A, OZAKTAS H M, et al. A survey of signal processing problems and tools in holographic 3-D television [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 17(11): 1631-1646.
- [5] TANIMOTO M, FUJII T. FTV: free viewpoint television [EB/OL]. (2002-07-22) [2011-03-15]. <http://wg11.sc29.org/m8595>.
- [6] THEOBALT C, ZIEGLER G, MAGNOR M, et al. Model-based free-viewpoint video acquisition, rendering and encoding [C]//Proceedings of Picture Coding Symposium. Piscataway, NJ, USA: IEEE, 2004: 1-6.
- [7] SMOLIC A, MULLER K, DIX K, et al. Intermediate view interpolation based on multi-view video plus depth for advanced 3D video systems [C]//Proceedings of International Conference on Image Processing. Piscataway, NJ, USA: IEEE, 2008: 2448-2451.
- [8] Telecommunication Standardization Sector. ITU-T recommendation H. 264: audio visual and multimedia systems infrastructure of audiovisual services: coding of moving video [S]. Geneva, Switzerland: Telecommunication Standardization Sector, 2010.
- [9] FEHN C. Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV [C]//Proceedings of International Conference on Stereoscopic Displays and Virtual Reality Systems XI. Bellingham, WA, USA: SPIE, 2004: 93-104.
- [10] MORVAN Y, FARIN D, DE WITH P H N. Joint depth/texture bit allocation for multi-view video compression [C]//Proceedings of Picture Coding Symposium. Glasgow, UK: EURASIP, 2007: 265-268.
- [11] LIU Yanwei, HUANG Qingming, MA Siwei, et al. Joint video/depth rate allocation for 3-D video coding based on view synthesis distortion model [J]. Signal Processing: Image Communication, 2009, 24(8): 666-681.
- [12] YUAN Hui, CHANG Yilin, HUO Junyan, et al. Model-based joint bit allocation between texture videos and depth maps for 3-D video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(4): 485-497.
- [13] BJONTEGAARD G. Calculation of average PSNR differences between RD-curves [EB/OL]. (2001-03-28) [2011-11-10]. http://wftp3.itu.int/av-arch/video-site/010_4_Aus/VCEG-M33.doc.

[本刊相关文献链接]

- 苏洪磊, 李兵兵, 杨付正. H. 264/AVC 网络视频编码失真评估的比特流层模型. 2012, 46(6): 36-41.
- 冯妮娜, 常义林, 李鹏, 卢朝阳. 适用于不等保护的自适应视频编码方法. 2012, 46(4): 94-99.
- 刘河潮, 杨付正, 常义林, 等. 考虑丢包特性的无参考网络视频质量评估模型. 2012, 46(2): 130-134.
- 王萍, 邓妍, 黄华. H. 264 中全零块检测的线性分类器算法. 2011, 45(10): 24-29.
- 陈士超, 武其松, 吴玉峰, 等. 一种大斜视双基地合成孔径雷达的频率变标成像算法. 2011, 45(10): 82-87.
- 李鹏, 常义林, 冯妮娜, 等. 用于视频传输的自适应分级正交幅度调制不等保护方法. 2011, 45(4): 72-76.
- 张志杰, 邹建华. 面向压缩域视频拷贝检测的主要边缘相对相位算法. 2010, 44(10): 8-12.
- 侯兴松, 薛端. 一种 H. 264 残差系数修剪算法. 2010, 44(10): 52-56.

(编辑 刘杨)