

# 基于自适应分段跳过规则的视点合成优化方法<sup>①</sup>

贾克斌<sup>②\*\*\*</sup> 刘 畅<sup>\*\*\*</sup> 窦 环<sup>\*\*\*</sup>

(\* 北京工业大学信息学部信息与通信工程学院 北京 100124)

(\*\* 先进信息网络北京实验室 北京 100124)

**摘 要** 对目前三维高效视频编码(3D-HEVC)标准的性能进行了分析,针对其深度图编码采用的合成视点失真变化(SVDC)计算复杂度高的问题,提出了一种基于自适应分段跳过规则的视点合成优化(VSO)方法。考虑到原始方法跳过了深度数据没有变化的模块或像素行,但忽略了可以在 VSO 中跳过的较小区域,该方法首先根据深度零失真准则和纹理平滑度准则判别出深度图中不会造成合成视点失真的区域,然后在深度图编码进行 SVDC 计算前基于自适应分段规则来提前跳过上述零失真区域,最后终止该区域的 SVDC 计算过程,有效地降低编码复杂度。对比实验结果证明了该方法的有效性,能在保持编码质量不变的前提下,平均减少 14.2% 的编码时间。

**关键词** 三维高效视频编码(3D-HEVC), 深度图编码, 合成视点失真变化(SVDC), 视点合成优化(VSO), 深度零失真准则, 纹理平滑度准则

## 0 引 言

近年来,随着多媒体技术的飞速发展,三维视频受到学术界、工业界,特别是电影业、家庭娱乐产业的重视。三维高效视频编码(three dimensional high efficiency video coding, 3D-HEVC)标准支持多视点视频加深度图(multi-view video plus depth, MVD)视频格式的编码<sup>[1]</sup>,解码端利用基于深度图像的绘制(depth image based rendering, DBIR)技术合成新的虚拟视点<sup>[2]</sup>。由于深度图用于虚拟视点的合成,所以编码深度图的失真会直接导致虚拟视点的失真。为了在编码深度图的同时考虑虚拟视点的质量,基于率失真优化(rate distortion optimization, RDO)方法的深度图编码决策过程还需在计算深度图自身编码失真的同时考虑虚拟视点的失真。因此 3D-HEVC 用合成视图失真代替原始深度图编码模式决

策过程中的率失真优化,这一技术被称作视点合成优化(view synthesis optimization, VSO)技术<sup>[3]</sup>。

目前,有很多国内外学者都将研究重点放在了如何减少 VSO 的复杂度上<sup>[4-8]</sup>。文献[4]提出了一种具有分段功能的允许深度失真模型,通过最小化给定比特率下的视图合成失真来建立模式决策和运动/视差估计的新率失真模型。此外,为了表征视图合成质量,文献[5]同时考虑了视频编码引起的失真和深度量化引起的失真和固有的几何失真。然而,这些算法是使用数学模型来度量合成视点的失真,它们的结果是近似值,并不能准确地表示合成视点的失真。为了精确地计算合成视点失真,Tech<sup>[9]</sup>等提出了合成视点失真变化(synthesized view distortion change, SVDC)方法来测量合成视图中的失真变化。但 SVDC 中引入的合成虚拟视点的过程增加了编码复杂度。为了能够加速 SVDC 模型的计算过

① 国家自然科学基金(61672064),北京市自然科学基金重点项目暨教委重点科技项目(KZ201610005007),中国博士后基金(2016T90022,2015M580029),北京市博士后工作经费资(2015ZZ-23),北京市自然科学基金面上项目(4172001)和北京市教委(040000546617002)资助项目。

② 男,1962年生,博士,博士生导师,教授;研究方向:视频编码和处理技术等;联系人,E-mail: kebinj@bjut.edu.cn (收稿日期:2017-06-27)

程,3D-HEVC 采用一种提前跳过(early skip, ES)方法<sup>[10]</sup>来提前终止某些不必要的 SVDC 计算。3D-HEVC 参考软件 HTM 采用的是基于块和基于行的 ES 方法。然而,在降低编码复杂度上仍有一定的空间。本文提出了一种基于自适应分段跳过规则的视点合成优化(VSO)方法,通过跳过一些小的像素片段避免复杂的合成视点失真变化(SVDC)计算过程。

### 1 SVDC 计算和提前跳过(ES)模式

SVDC 可用于精确地测量 VSO 的失真,它定义

为两个合成视点的失真之差,计算过程如图 1 所示。其中 V 表示由未编码的原始纹理图像帧 T 和未编码的原始深度图像帧 D 合成的虚拟视点图像帧,此外,由已编码的纹理图像帧 T' 分别与两幅部分编码的深度图像帧 D' 和 D'' 合成的两幅虚拟视点图像帧记作 V' 和 V'', 但 D'' 在 D' 的基础上还包含当前待决策深度图编码块中失真的深度信息,将 V' 和 V'' 分别与 V 计算的差值平方和记为 SSD' 和 SSD'', 最后计算 SSD'' 和 SSD' 的差值,得到 SVDC 的值。

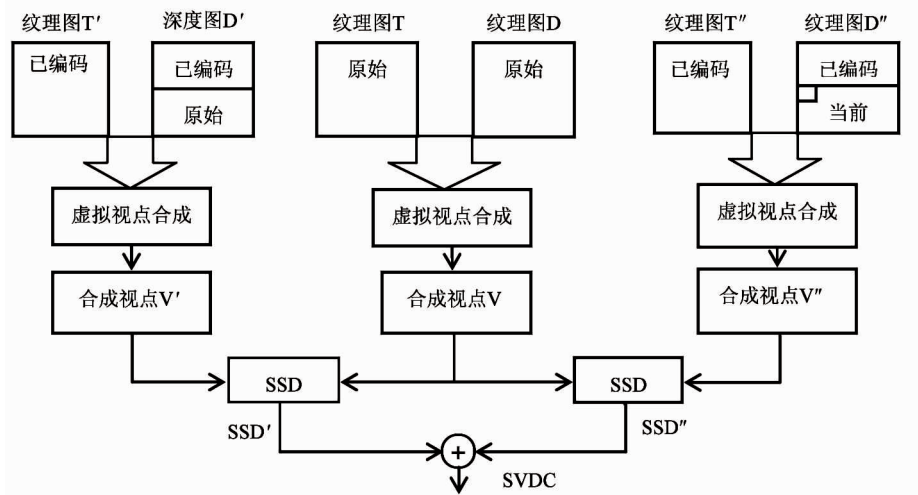


图 1 SVDC 计算过程框图

SVDC 的引入必然会带来高的复杂度,为了能够加速 SVDC 模型的计算过程,3D-HEVC 采用一种 ES 方法来提前终止某些不必要的 SVDC 计算过程。目前的 HTM-16.0 是采用基于块和基于行的 ES 决策模式来跳过不必要的编码块从而减少 SVDC 的计算时间。ES 方法的核心跳过条件为,如果当前深度值与原始深度值计算得到的视差矢量相同,则该深度值不会造成合成视点失真,可以直接跳过 SVDC 的计算。但在计算 SSD'' 时,每个决策候选都会由当前的待决策编码块编码得到不同的预测块和重建块,使得编码器对于每个候选都需要计算一次 SSD'', 这对于编码端而言仍然具有很高的计算复杂度。因此,虽然 ES 方法在一定程度上可以降低编码复杂度,但仍然有提高的空间。

## 2 自适应分段跳过规则

### 2.1 深度零失真准则

在视图合成过程中,来自深度图的数据可借助于相机参数转换为视差矢量。这些视差矢量用于将像素从原始视图转换为合成视图。根据双目视差法求深度的计算原理,视差矢量  $d$  的计算如下式所示:

$$d = \frac{f \times B}{z} \tag{1}$$

其中  $f$  和  $B$  分别表示相机焦距和两个视图间的基线距离, $z$  表示当前深度像素描述的物体在三维空间中的实际距离

$$z = 1 / \left[ \frac{1}{255} \times \left( \frac{1}{Z_{near}} - \frac{1}{Z_{far}} \right) + \frac{1}{Z_{far}} \right] \tag{2}$$

其中  $Z_{near}$  和  $Z_{far}$  分别表示相机中的最近和最远的深

度值;  $I$  是一个 8 位的量化深度值, 它的范围在 0 ~ 255 之间。

通过式(1)和(2), 可以得到视差向量  $d$  如下

$$d = \frac{f \times B \times I}{255} \times \left( \frac{1}{Z_{near}} - \frac{1}{Z_{far}} \right) + \frac{f \times B}{Z_{near}} \quad (3)$$

如果由编码的深度值获取的一段失真的视差矢

量与由原始深度值获取的视差矢量相同, 则可以忽略该深度段的失真, 并且可以跳过 SVDC 计算过程。

### 2.2 纹理平滑度准则

深度图中的像素失真并不是在合成视点中造成失真的充分必要条件, 如图 2 所示。



图 2 纹理平滑和不平滑像素的搬移过程

从图中可看出,  $B$  处于纹理平滑区域, 因此失真的虚拟视点像素位置  $B_2$  与正确的虚拟视点像素位置  $B_1$  具有相同的纹理内容。而对处于复杂纹理区域的  $A$  点来说, 失真后的像素位置明显有别于正确的虚拟视点像素位置。基于上述分析, 本文利用某一水平方向的像素对来定义纹理平滑度准则, 如下式所示:

$$|p_i - p_{i-1}| \leq T \quad (4)$$

其中, 阈值  $T$  的选择步骤如下:

(1) 在纹理图编码后, 记录所有选择帧内 DC 模式作为其最终编码模式的编码块。

(2) 对步骤(1)中的某一编码块求其平均水平方向的亮度差值。

(3) 计算步骤(2)中得到值的平均值, 并将其作为阈值  $T$ 。需要特别注意的是, 每编码一个  $I$  帧后都会更新一次阈值  $T$ 。

### 2.3 自适应分段跳过规则及其流程图

图 3 是整个基于自适应分段跳过规则的视点合成优化方法的流程图, 其中并没有展示出基于块和基于行的 ES 方法。首先判断一尺寸为  $N \times N$  的编码块是否满足基于块和基于行的 ES 方法, 假如该编码块满足基于块的 ES 条件, 则整个编码块的 SVDC 过程提前终止; 否则, 用基于行的 ES 条件来

判断该编码块中的所有像素行是否需要继续进行 SVDC 的计算过程。假如当前编码块中的某一像素满足了基于行的 ES 条件, 则终止该行的 SVDC 计算过程, 并转而判断下一像素行是否满足基于行的 ES 条件。如果某一像素行不满足基于行的 ES 条件, 则基于上述深度零失真准则和纹理平滑度准则, 可以找到不会造成合成视点失真的区域, 分别记为  $R_d$  和  $R_t$ , 记  $R_d$  的长度为  $L_d$ ,  $R_t$  的长度为  $L_t$ 。之后选取

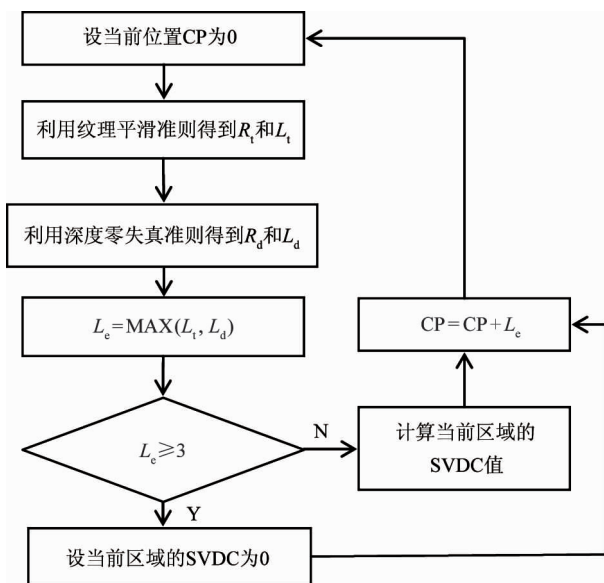


图 3 视点合成优化方法流程图

$L_d$  和  $L_l$  的最大值来确定最终的零失真区域,记为  $L_e$ 。在明确  $L_e$  后,再找到该最终零失真区域的下一个像素的位置,重复上述步骤直到此像素行的像素遍历完毕,这时只需对该像素行中除去所有零失真区域之外的其余区域计算合成视点失真。为了保证提出算法的有效性,当且仅当这些零失真区域的长度大于或者等于 3 时才会将其从当前像素行中除去。值得注意的是,本文提出的基于自适应分段跳过规则的视点合成优化方法是与基于块和基于行的 ES 方法共同工作的。

### 3 实验结果

为了评价本文提出方法的有效性,将 3D-HEVC 的参考模型 HTM-16.0 作为实验平台,测试序列均

为 JCT-3V 小组建议的标准序列。选取的标准测试序列为 Balloons(1024 × 768)、Kendo(1024 × 768)、Newspaper\_CC(1024 × 768)、Poznan\_Hall2(1920 × 1088)和 Poznan\_Street(1920 × 1088)。每个测试序列均编码 3 个视点和对应的 3 个深度图序列。量化参数设置为(25, 34)、(30, 39)、(35, 42)和(40, 45),它们分别用于编码纹理图和深度图。此外,将 GOP 长度设置为 8,帧内周期设置为 24,用于合成虚拟视点的量化步长设为 0.25,测试序列为 100 帧。所有实验在配置为 Intel(R) Core(TM) i5-3470 3.20GHz CPU,4.00GB RAM 的 PC 上独立执行。

依照文献[11]公布的 3DV 国际通用测试标准对 HTM 的原编码方法和本文提出的视点合成优化方法进行比较,编码时间结果比较如表 1 所示。其中  $\Delta T_1$  和  $\Delta T_2$  分别表示的是本文提出的编码方法

表 1 不同方法的编码时间比较

序列	QP	编码时间(s)			$\Delta T_1$ (%)	$\Delta T_2$ (%)
		原始方法	本文提出方法	文献[3]方法		
Balloons	25,34	6375.1380	5486.4430	6032.1560	87.8	94.8
	30,39	5039.5540	4348.6650	4758.9480		
	35,42	4413.0900	3865.3060	3901.8940		
	40,45	4075.8870	3722.8330	3798.5690		
Kendo	25,34	6813.8590	5403.4050	6589.2460	83.0	90.9
	30,39	5241.4240	4572.7730	4875.5620		
	35,42	4614.7770	3884.7390	4157.5840		
	40,45	4232.5140	3456.2730	3647.1240		
Newspaper_CC	25,34	5712.7590	4878.7820	5468.2780	81.4	91.3
	30,39	4958.4400	4001.6320	4357.1790		
	35,42	4362.0250	3579.3210	3894.7510		
	40,45	4264.2310	3309.8200	3579.4720		
Poznan_Hall2	25,34	14391.6000	12472.2180	13479.6040	89.1	71.2
	30,39	12086.2990	10870.0180	11479.2830		
	35,42	10994.9720	10066.6940	10136.4720		
	40,45	10990.5860	9706.5790	9956.1450		
Poznan_Street	25,34	14672.0560	14206.7610	14345.5840	88.1	95.6
	30,39	11832.2050	10785.4940	10987.4890		
	35,42	11414.7250	9901.5380	10997.7820		
	40,45	11949.8460	9422.1750	9875.4680		
均值					85.8	88.7

相较于原有的方法和文献[3]提出的方法所节省的编码时间百分比。从表1中可以看出,本文提出方法的编码时间比 HTM 原编码时间平均减少了14.2%的编码复杂度,比文献[3]的编码时间平均减少了11.3%的编码复杂度。因此本文提出方法能够比文献[3]提出的方法获得更高的编码效率,编

码复杂度更低。

除了对编码时间进行了对比,还对 BDBR 性能进行了比较。结果如表2所示。从表2中可看出,除了 Balloons 和 Poznan \_ Hall2 序列的合成视点的  $\Delta$ BDBR 较高外,本文提出的编码方法基本与原编码方法保持了相同的编码性能。

表2 本文提出方法与原始方法的 BDBR 性能比较

序列	原始方案与本文提出方法的比较(%)		
	video PSNR / video bitrate	video PSNR / total bitrate	synth PSNR / total bitrate
Balloons	0.1	0.1	1.8
Kendo	0.0	-0.1	1.1
Newspaper _ CC	-1.0	-4.3	0.0
Poznan _ Hall2	0.0	1.5	1.6
Poznan _ Street	0.0	-0.1	1.1
均值	-0.2	-0.7	1.1

## 4 结论

本文提出了一种基于自适应分段跳过规则的视点合成优化方法,在原有提前跳过(ES)方法的基础上以区域为单位灵活地计算合成视点失真,用区域计算代替整个模块或行的计算,主要是针对 3D-HEVC 中深度图编码采用的计算 SVDC 复杂度高的问题。本研究方法首先从理论的角度分析了传统的合成视点失真模型给编码端造成高计算复杂度的原因,之后根据深度零失真准则和纹理平滑准则判别出深度图中不会造成合成视点失真的区域,然后在深度图编码进行 SVDC 计算前基于自适应分段来提前跳过上述零失真区域,最后终止该区域的 SVDC 计算。实验结果表明,本文提出的视点合成优化方法能在保证率失真优化性能基本不变的前提下节省平均 14.2% 的编码时间。这一方法不仅可以用于纹理内容平坦的视频序列,还对运动快速的物体、视频内容复杂的视频序列适用。然而,关于 3D-HEVC 的纹理图和深度图编码优化问题仍有很多内容值得深入研究,例如,就本文来说,亮度差值差多少才能判断纹理平滑,此外,选择其它的帧内预测模式来代替本文中选择的帧内 DC 预测模式会不会有更好的效果。这些都是今后要进一步探讨的问题。

## 参考文献

- [1] Chen Y, Hannuksela M M, Suzuki T, et al. Overview of the MVC + D 3D video coding standard[J]. *Journal of Visual Communication and Image Representation*, 2014, 25(4): 679-388
- [2] Maugey T, Petrazzuoli G, Frossard P, et al. Reference view selection in DIBR-based multiview coding[J]. *IEEE Transactions on Image Processing*, 2016, 25(4): 1808-1819
- [3] Ma S W, Wang S Q, Gao W. Low complexity adaptive view synthesis optimization in HEVC based 3D video coding[J]. *IEEE Transactions on Multimedia*, 2014, 16(1): 266-271
- [4] Zhang Y, Kwong S, Hu S D, et al. Efficient multiview depth coding optimization based on allowable depth distortion in view synthesis[J]. *IEEE Transactions on Image Processing*, 2014, 23(11): 4879-4892
- [5] Liu Y W, Huang Q M, Ma S W, et al. Joint video/depth rate allocation for 3D video coding based on view synthesis distortion model[J]. *Signal Processing-Image Communication*, 2009, 24(8): 666-681
- [6] 张洪彬,伏长虹,苏卫民,等. 3D-HEVC 深度图像帧内编码单元划分快速算法[J]. *电子与信息学报*, 2016, 38(10): 2523-2533

- [ 7 ] Yang C, An P, Liu D Y, et al. Virtual view distortion estimation for depth map coding[C]. In: Proceedings of the IEEE Visual Communications and Image Processing (VCIP) Conference, Singapore, 2015. 1-4
- [ 8 ] Dou H, Chan Y L, Jia K B. Segment-based view synthesis optimization scheme in 3D-HEVC[J]. *Journal of Visual Communication and Image Representation*, 2017, 42: 104-111
- [ 9 ] Tech G, Schwarz H, Muller K, et al. 3D video coding using the synthesized view distortion change[C]. In: Proceedings of the 29th Picture Coding Symposium (PCS), Krakow, Poland, 2012. 25-28
- [ 10 ] Zhang Q W, Wu Q G, Wang X B. Early skip mode decision for three-dimensional high efficiency video coding using spatial and interview correlations [ J ]. *Journal of Electronic Imaging*, 2014, 23(5) :053017
- [ 11 ] Muller K, Vetro A. Common test conditions of 3DV core experiments[C]. In: Proceedings of the Joint Collaborative Team on 3D Video Coding Extension Development Meeting, Document JCT3V-G1100, San Jose, USA, 2014. 1-10

## Optimization of viewpoint synthesis based on adaptive segment skip rule

Jia Kebin<sup>\*\*\*</sup>, Liu Chang<sup>\*\*\*</sup>, Dou Huan<sup>\*\*\*</sup>

(\* College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124)

(\*\* Beijing Laboratory for Advanced Information Networks, Beijing 100124)

### Abstract

The performance of the standard of three dimensional high efficiency video coding (3D-HEVC) is analyzed, and aiming at its problem that the complexity of calculating synthesized view distortion change (SVDC) in depth coding is high, a view synthesized optimization (VSO) method based on adaptive segmentation skipping is proposed. Considering that the original method skips pixels of lines or blocks without the change of the depth data, but ignores the smaller areas that can also be skipped in VSO, the proposed method determines the region of the depth map that does not cause the synthesized view distortion according to the depth zero-distortion criterion and the texture smoothing criterion. Then, the zero-distortion region is skipped based on adaptive segmentation skip before the SVDC calculation of the depth map coding. Finally, the SVDC calculation process for the area is terminated. The experimental results show that the proposed method can reduce the coding time by 14.2% while keeping the coding quality unchanged.

**Key words:** three dimensional high efficiency video coding (3D-HEVC), depth coding, synthesized view distortion change (SVDC), view synthesized optimization (VSO), depth zero-distortion criterion, texture smoothness criterion