



主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2016
07
VOL.21

ISSN1006-8961
CN11-3758/TB



人脸图像处理 P865



第21卷第7期（总第243期）

2016年7月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 顾行发
编辑出版 《中国图象图形学报》编辑出版委员会
邮政信箱 北京9718信箱
邮 编 100101
电子信箱 jig@radi.ac.cn
电 话 010-64807995
网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱：北京399信箱 邮编：100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

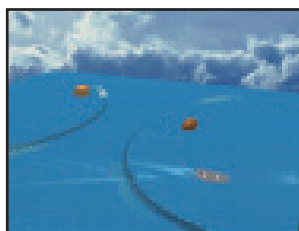
Editor-in-Chief Gu Xingfa
Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics
P.O.Box 9718, Beijing, P.R.China
Zip code 100101
E-mail jig@radi.ac.cn
Telephone 010-64807995
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China))
Printed by Beijing Kexin Printing Co., Ltd.

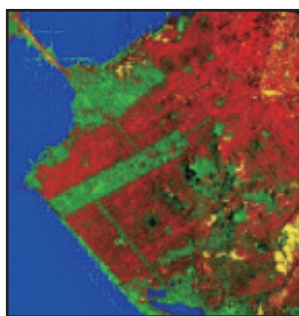
CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



大规模柔体的连续碰撞检测
算法(第901页)



大规模流体场景的真实感与
实时模拟(第913页)



星载SAR影像上船舶方位
向模糊去除算法(第951页)

图像处理和编码

低存储化压缩感知

王金铭, 叶时平, 徐振宇, 蒋燕君 835

3D-HEVC深度图像快速帧内编码方法

张洪彬, 伏长虹, 陈锐霖, 萧允治, 苏卫民 845

加权KNN的图文数据融合分类

康丽萍, 孙显, 许光奎 854

结合肤色分割与平滑的人脸图像快速美化

邱佳梁, 戴声奎 865

图像分析和识别

级联型P-RBM神经网络的人脸检测

叶学义, 陈雪婷, 陈华华, 顾亚风, 吕秋云 875

结合全局与局部信息活动轮廓的非同质图像分割

郝智慧, 郭满才, 宋杨杨 886

计算机图形学

带参数的多项式势函数与构造基于Metaball的过渡曲线

李军成, 宋来忠, 刘成志 893

大规模柔体的连续碰撞检测算法

周清玲, 刘艳, 程天翔 901

虚拟现实与增强现实

大规模流体场景的真实感与实时模拟

邵绪强, 杨晓丹, 李继荣, 阎蕾 913

医学图像处理

形状约束下活动轮廓模型冠脉血管图像多尺度分割

郭笑妍, 梅雪, 李振华, 曹佳松, 周宇 923

遥感图像处理

深度置信网络在极化SAR图像分类中的应用

邓磊, 付姗姗, 张儒侠 933

高光谱图像低秩表达与噪声水平估计

唐中奇, 付光远, 赵晓林, 陈进, 张利 942

星载SAR影像上船舶方位向模糊去除算法

邴磊, 邢前国, 邹娜娜, 李圳波, 吴樊 951

TDFA: 一种生成空间影像金字塔的方法

张云舟, 张陌, 王晋年, 张刚 959

无人机侦察视频超分辨率重建方法

张岩, 李建增, 李德良, 杜玉龙 967

CONTENTS

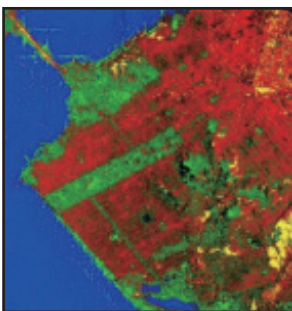
JOURNAL OF IMAGE AND GRAPHICS



Continuous collision detection algorithm for large-scale deformable objects (P901)



Realistic and real-time simulation of large-scale fluids (P913)



Application of deep belief network in polarimetric SAR image classification (P951)

Image Processing and Coding

- Reducing the storage space of the measurement matrix for compressive sensing
Wang jinming, Ye shiping, Xu zhenyu, Jiang yanjun 835
- Fast intra coding for depth map in 3D-HEVC
Zhang Hongbin, Fu Changhong, Chan Yuilam, Siu Wanchi, Su Weimin 845
- Fusion method via KNN with weight adjustment for the classification of image-text co-occurrence data
Kang Liping, Sun Xian, Xu Guangluan 854
- Fast facial beautification algorithm based on skin-color segmentation and smoothness
Qiu Jialiang, Dai Shengkui 865

Image Analysis and Recognition

- Cascaded probability state-restricted Boltzmann machine for face detection
Ye Xueyi, Chen Xueting, Chen Huahua, Gu Yafeng, Lyu Qiuyun 875
- Inhomogeneous image segmentation based on active contours with global and local information
Hao Zhihui, Guo Mancai, Song Yangyang 886

Computer Graphics

- The polynomial potential function with a parameter and transition curve based on Metaball technology
Li Juncheng, Song Laizhong, Liu Chengzhi 893
- Continuous collision detection algorithm for large-scale deformable objects
Zhou Qingling, Liu Yan, Cheng Tianxiang 901

Virtual Reality and Augmented Reality

- Realistic and real-time simulation of large-scale fluids
Shao Xuqiang, Yang Xiaodan, Li Jirong, Yan Lei 913

Medical Image Processing

- Multi-scale segmentation of coronary artery image based on active contour model with shape constraints
Guo Xiaoyan, Mei Xue, Li Zhenhua, Cao Jiasong, Zhou Yu 923

Remote Sensing Image Processing

- Application of deep belief network in polarimetric SAR image classification
Deng Lei, Fu Shanshan, Zhang Ruxia 933
- Low-rank representation for hyperspectral image noise level estimation
Tang Zhongqi, Fu Guangyuan, Zhao XiaoLin, Chen Jin, Zhang Li 942
- Azimuth ambiguity removal method for ship detection based on spaceborne SAR images
Bing Lei, Xing Qianguo, Zou Nana, Li Zhenbo, Wu Fan 951
- TDFA: a generation method of spatial image pyramid
Zhang Yunzhou, Zhang Mo, Wang Jinnian, Zhang Gang 959
- Super-resolution reconstruction for UAV video
Zhang Yan, Li Jianzeng, Li Deliang, Du Yulong 967

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2016)07-845-09

论文引用格式: Zhang H B, Fu C H, Chan Y L, Siu W C, Su W M. Fast intra coding for depth map in 3D-HEVC[J]. Journal of Image and Graphics, 2016, 21(7): 845-853. [张洪彬, 伏长虹, 陈锐霖, 萧允治, 苏卫民. 3D-HEVC 深度图像快速帧内编码方法[J]. 中国图象图形学报, 2016, 21(7): 845-853.] [DOI:10. 11834/jig. 20160702]

3D-HEVC 深度图像快速帧内编码方法

张洪彬^{1,2}, 伏长虹¹, 陈锐霖², 萧允治², 苏卫民¹

1. 南京理工大学电子工程与光电技术学院, 南京 210094; 2. 香港理工大学数字信号处理实验室, 香港

摘要: 目的 新一代 3D 视频采用了多视点加深度图像的格式, 其编码测试模型 3D-HEVC(3D 高效视频编码) 为深度图像引入了一些新的编码技术, 包括深度模型模式、分段直流(DC)残差编码模式和帧内跳过模式等, 但是同时也显著地增加了帧内编码器的复杂度, 严重制约了 3D 视频的实时应用, 因此本文提出了一种深度图像帧内模式选择方法。**方法** 首先, 分析在分层编码结构中, 父层帧内模式和子层帧内模式的关系, 并依据父层的帧内模式对子层的候选模式进行修剪; 其次, 通过分析粗略候选帧内模式, 把当前块分为平滑块、方向块和边界块, 并根据不同块进行候选帧内模式的选择。**结果** 与高效 3D 视频编码标准的测试模型相比, 本文的深度图像帧内模式选择方法可以节省约 44% 的编码时间。本文方法有效地减少了计算率失真代价的候选帧内模式数目, 从而降低了帧内编码器的复杂度。**结论** 该方法能在保证虚拟视点质量的前提下, 有效地降低深度图像帧内编码的复杂度。**关键词:** 高效 3 维视频编码; 深度图像; 帧内编码; 分层编码结构; 粗略模式决定

Fast intra coding for depth map in 3D-HEVC

Zhang Hongbin^{1,2}, Fu Changhong¹, Chan Yuilam², Siu Wanchi², Su Weimin¹

1. School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China;

2. Digital Signal Processing Laboratory, The Hong Kong Polytechnic University, Hong Kong, China

Abstract: **Objective** The new generation 3D video adopts multi-view video plus depth as the main format. To improve the coding performance and quality of synthesized views, several new techniques, including depth modeling mode, segment-wise DC coding, and skip intra mode, are introduced into the current 3D-HEVC test model for depth intra coding. However, these tools dramatically increase computational complexity, which obstructs the practical applications for the 3D video. Thus, in this study, a fast depth intra coding algorithm is proposed to alleviate the burden of the encoder. **Method** Considering that 3D-HEVC reuses the well-known hierarchical coding structure, the mode information of the parent coding units (CU) can be reused to predict the most probable modes for the current CU. Consequently, only a subset of the intra mode candidates is employed to calculate the rough rate-distortion cost. According to rough modes and their rough costs by rough mode decision, the current CU will be classified into three categories, namely, smooth, directional, and edge. For each class, a different set of rough mode candidates will be selected. **Result** Simulation results show that the proposed algorithm achieves a reduction of 44% encoding time for depth intra coding while maintaining the quality of synthesized views, compared with the original test model HTM-13.0. **Conclusion** Based on the mode information of parent CU and rough mode de-

收稿日期: 2015-09-08; 修回日期: 2016-02-24

基金项目: 国家自然科学基金项目(61301109)

第一作者简介: 张洪彬(1988—), 男, 2011 年南京理工大学信息与通信工程专业博士研究生, 主要研究方向为视频编码与分析。

E-mail: hongbinzhang0801@sina.cn

Supported by: National Natural Science Foundation of China (61301109)

cision, the proposed method can significantly reduce the number of intra mode candidates. Consequently, the proposed method can provide a remarkable time reduction of depth intra coding with a negligible BD rate increase.

Key words: three-dimension high efficiency video coding (3D-HEVC); depth map; intra coding; hierarchical coding structure; rough mode decision

0 引言

随着 3D 视频技术在电影工业中取得的巨大成功,潜在的家庭交互式 3D 视频应用引起了工业界和学术界的很大关注。起初,业界将多视点视频制定为第 1 代 3D 视频格式。由于多视点视频需要存储 20 到 50 个视点才可以实现自由视点视频,这无疑对目前的存储技术和网络带宽提出了很大的挑战。为了降低 3D 视频的数据量,运动图像专家组(MPEG)于 2007 年将多视点视频加深度图像(MVD)标准化为新一代 3D 视频格式^[1]。MVD 仅包含几个视点的纹理视频和对应的深度视频,其与传统 2D 视频相比仅需要增加 5% 到 20% 的比特率,因此在现有的网络带宽和存储能力下,MVD 是唯一可以实现自由视点视频的技术方案^[2]。在解码端,MVD 利用基于深度图像的渲染技术(DIBR)来合成虚拟视点,并将 3D 效果所需的视点呈现给多个用户^[3]。为了更好地推广新一代 3D 视频,联合视频开发组(JCT)在 2012 成立了 JCT-3V 工作组,专门负责 MVD 视频压缩的标准化工作。该工作组将 MVD 引入到新一代高效视频编码(HEVC)标准中,并将其压缩标准作为 HEVC 的一个重要扩展标准,称为 3D-HEVC^[4],对应测试模型的参考软件为 HTM。

HEVC 依然沿用传统的基于块的混合编码框架,通过采用更加灵活的分层四叉树编码结构来选择不同大小的编码单元,从而获得较高的压缩比^[5]。然而,编码器需要通过复杂的率失真计算来决定四叉树编码结构的最优分割,这导致了 HEVC 的计算复杂度远高于以往的编码器。虽然 3D-HEVC 继承了 HEVC 的编码框架,然而 HEVC 是为自然图像或彩色视频设计的编码方案,并不适用于深度图像编码。深度图像是一种特殊的灰度图像,其像素值表示物体到摄像机的量化距离。深度图像和自然图像的不同,主要表现为:1)深度图像大部分都是平坦的区域;2)这些平坦的区域被锐利的边缘所分割;3)深度图像并非用来观看,而是用来辅

助合成虚拟视点,所以编码器需要抑制深度图像边缘的失真,否则绘制的虚拟视点会产生振铃效应^[2],并且深度图像编码模式的选择不仅由深度图像本身的失真来衡量,而且还要考虑合成视图的失真。

针对深度图像的特点,3D-HEVC 引入了一些新技术来保留深度图像锐利的边缘,这些新技术包含帧内跳过模式、深度模型模式(DMM)、分段 DC 残差编码(SDC)和视角合成优化(VSO)等^[4]。与 HEVC 相比,深度图像的帧内候选模式数目多,包含 DMM 和 HEVC 的帧内模式,并且深度图像采用 VSO 来选择最优的帧内模式,而 VSO 的计算涉及高复杂度的视角合成算法,所以深度图像帧内编码的复杂度急剧增加,使得 3D-HEVC 的实时应用面临挑战^[4]。Muller 等人^[6]通过在双像素域搜索最优的 DMM,有效地减少了 DMM 的数目。而 Fu 等人^[7]将 DMM 按分割直线的方向分为 6 类,并通过分析当前块的方差来决定 DMM 的子搜索集。Gu 等人^[8]发现 DMM 并不适用于深度图像中的平滑区域,在此基础上,Park^[9]发现 DMM 也不适用于包含水平和垂直边界的编码单元,因此各自提出了相应的 DMM 跳过方法。Zhang 等人^[10]利用深度块的帧内邻近参考像素预测当前块的纹理方向,并依据此方向提出了一种帧内模式的修剪方法,有效地减少了 DMM 和 HEVC 帧内模式的数目。Gu 等人^[11]利用低复杂度的率失真代价判断是否需要进行 DMM 的模式选择。Tech 等人^[12]建立了一个简单深度图像率失真模型,通过重建纹理图像和深度图像失真的加权值来评价帧内模式,从而避免了视角合成的计算,节省了编码时间。此外,文献[13-16]利用深度图像和其纹理图像的相关性提出了一系列快速方法以降低 3D-HEVC 编码的计算复杂度。这些研究主要着眼于帧间和视点间的模式选择,提升了帧间编码的速度。

虽然 Muller^[6]、Gu^[8]和 Tech^[12]等人提出的方法已经加入到了当前的 3D-HEVC 的测试模型,但是在相同的测试条件下,一幅 I 帧深度图像的编码时间是纹理图像的 5 到 10 倍。因此,深度图像帧内编码已经成为制约 3D 视频技术发展的瓶颈,进一步降低其

计算复杂度对3D视频未来的实时应用意义重大。

本文主要研究分层四叉树编码结构中父层的帧内模式与子层帧内模式之间的关系,剔除一些在子层编码单元中出现概率极低的候选帧内模式,有效地减少子层编码单元需要计算率失真代价的帧内模式数目,同时还依据粗略帧内模式对编码单元进行分类,根据不同的编码单元类选择不同的候选帧内模式,从而进一步地降低编码复杂度。

1 深度图像的帧内编码

1.1 HEVC 的分层四叉树编码结构

与以往的编码标准相比,HEVC 采用了更加灵活的基于分层四叉树编码结构。首先,一幅待编码图像被分割为大小为 64×64 像素的最大编码单元(LCU)。根据率失真代价,每个 LCU 可以递归地分割为4个大小相同的子编码单元,直至大小为 8×8 像素的最小编码单元。图1为 LCU 最优分割的过程,其中编码单元中的数字为编码的序号。从图1可知,HEVC 的分层四叉树编码结构总是先计算父层编码单元的率失真代价得到最优帧内模式,然后计算子层编码单元的率失真代价得到最优模式,最终通过比较两个层的率失真代价来决定 LCU 的最优分割。

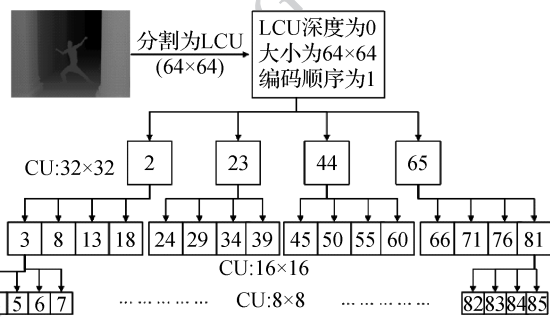


图1 HEVC 的分层四叉树编码结构

Fig.1 Hierarchical quadtree coding structure of HEVC

1.2 深度图像的帧内模式及其决定方法

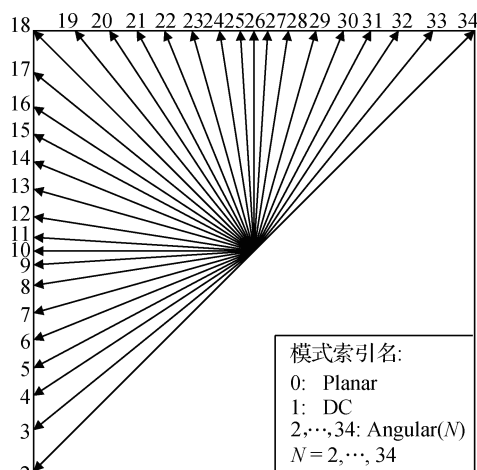
深度图像的帧内模式包含 HEVC 的帧内模式和新加入的深度模型模式。其中,HEVC 一共有35种帧内模式,如图2(a)所示模式2到模式34(Angular(N), $N = 2, \dots, 34$)分别对应了一个方向用于预测纹理,而 Planar 和 DC 模式用于预测平滑区域。图2(a)深度模型模式采用非矩形分割,如图2(b)所示,用一条直线 L_{SE} (S 为起点, E 为终点)将编码

单元分割为两个区域。表1列出了大小不同的编码单元对应的深度模型模式的数目。因为深度模型模式的候选模式数目众多,所以该模式可以很好地预测复杂多变的物体边界,但是也增加了模式选择的计算量。如图2所示,Planar 模式的索引为0,DC 模式的索引为1,33种 Angular 模式的索引为2~34,而深度模型模式的索引为35。原始的深度图像帧内模式选择方法步骤如下:

1) 计算35种 HEVC 帧内模式的粗略率失真代价

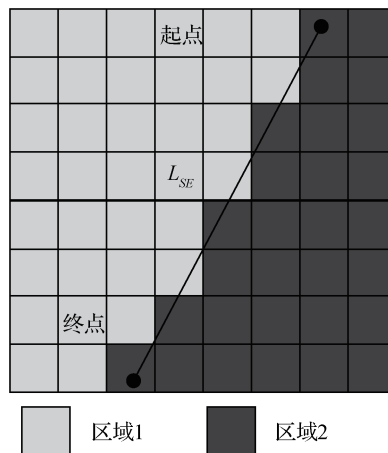
$$J_{vso} = D_{vso} + \lambda \times R(M) \quad (1)$$

式中, D_{vso} 为虚拟视点失真和当前块的哈达玛系数绝对值和的加权值,而 $R(M)$ 为候选模式的比特率, λ 为权衡因子。依据编码单元大小选择一定数目的粗略帧内模式 (64×64 像素和 32×32 像素的编码单元选择3个;其他编码单元选择8个),记为 RMD_i ,



(a) HEVC帧内模式

楔形分割



(b) 深度模型模式

图2 深度图像的帧内模式

Fig.2 Intra modes of depth map ((a) conventional HEVC intra mode; (b) depth modelling mode)

表1 不同大小的编码单元对应的深度模型模式数目

Table 1 The number of depth modelling mode

编码单元/像素	原始方法	快速方法 ^[6]
4 × 4	86	58
8 × 8	766	310
16 × 16	510	384
32 × 32	510	384

并和邻近块的帧内模式一起作为候选模式。 RMD_i 以按照粗略率失真代价的大小进行排序,即 RMD_0 为粗略率失真代价最小的帧内模式。

2) 计算表1中每个深度模型模式的失真代价,得到1个最优的深度模型模式,并作为候选模式。

3) 对第1)2)步选择的候选模式,分别计算每个模式的完整率失真代价,从而得到最优的帧内模式。

2 快速深度图像帧内编码方法

将现有3D-HEVC中的DMM部分停用以测试其对编码性能和时间的影响,并将相关结果在表2中列出。计算复杂度用 ΔT 来衡量,表示其占总编码时间的百分比。编码性能用虚拟视角的BDBR(BD-RATE)来衡量,由纹理和深度视频的总比特率和虚拟视点的峰值信噪比(PSNR)计算得到。从表2可以看出,不采用深度模型模式将导致BDBR

增加5.27%,同时深度模型模式的编码时间在总编码时间约占28%。可见深度模型模式是一种必不可少的编码工具,且需要进一步降低其计算复杂度。此外,从表2还可以看出,帧内编码主要的计算量来自于HEVC帧内模式(约72%),因为共有3个或8个HEVC帧内模式需要计算完整率失真代价,却仅有1个深度模型模式需要该计算。

表2 停用深度模型模式后编码性能和计算复杂度变化

Table 2 The influence on the coding performance when the depth modelling mode is disabled

测试序列	BDBR	ΔT /%
Kendo	4.40	-27.16
Balloons	4.03	-27.53
Newspaper	7.43	-29.77
Dancer	7.69	-26.54
Fly	3.81	-29.16
Hall2	5.91	-26.33
Street	2.10	-29.01
Shark	6.76	-28.61
平均值	5.27	-28.01

2.1 基于分层四叉树结构的帧内模式修剪方法

图3列出分层四叉树编码结构中,父层帧内模式 m 和子层帧内模式 n 之间的概率统计关系。其

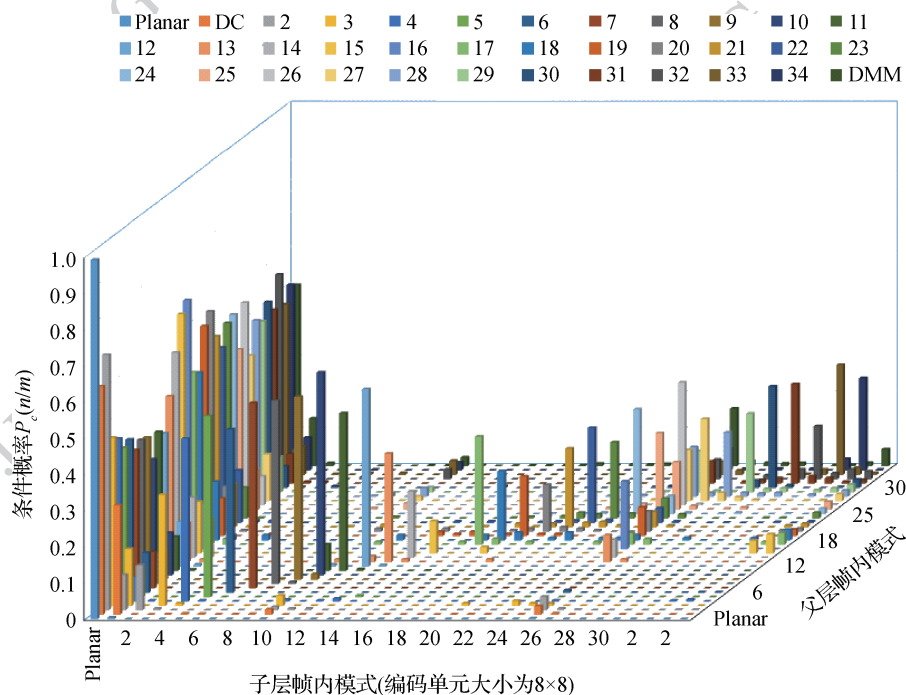


图3 父层帧内模式和子层帧内模式的统计关系

Fig. 3 The distribution of $P_c(n|m)$

中,横坐标表示大小为 8×8 像素的子层编码单元的帧内模式,纵坐标表示大小为 16×16 像素的父层编码单元的帧内模式, $P_c(n|m)$ 表示父层的帧内模式为 m 时,子层的帧内模式为 n 的条件概率。大量的实验证实其他大小的编码单元的 $P_c(n|m)$ 具有和图3相似分布。

父层帧内模式为该层编码单元空间信息的概括,而由于深度图像有很强的空间相关性^[17](详细论述参见文献[17]),所以父层的帧内模式可以为子层编码单元提供一种粗略的空间描述。因此如图3所示,考虑父层的帧内模式,子层的帧内模式的分布存在明显差异。更详细地讲,子层的 $P_c(n|m)$ 主要集中分布在 Planar 模式、DC 模式、水平模式(索引为 10)、垂直模式(索引为 26)和 DMM。此外,图3中剩余的主要概率值集中在主对角线附近。主对角线上的值表示子层继承父层帧内模式的条件概率。考虑一般的情况,如果父层的编码单元不包含边界,那么父层和子层都会选择 Planar 或是 DC 模式作为最优模式,因此 $P_c(n=0|m=0)$ 和 $P_c(n=$

$1|m=1)$ 的值都很大。如果父层的编码单元包含边界,即具有一定的方向信息,那么父层的编码单元将选择 DMM 或是 Angular 作为最优模式。在这种情况下,对于不包含边界的子编码单元仍然会选择 Planar 和 DC 模式;对于包含边界的子编码单元会选择父层帧内模式或是其方向邻近的模式作为最优模式,因此条件概率 $P_c(n=0,1,m|1 < m < 35)$, $P_c(n=35|m=35)$ 都具有相对较大的值。此外,由于水平和垂直方向广泛地存在于深度图像中,因此有一定的机率被选为最优帧内模式,故 $P_c(n=10|m)$ 和 $P_c(n=26|m)$ 的概率值也较大。综上所述,父层的帧内模式和子层的帧内模式有很大的关联性。

如 1.1 节所述,父层编码总是先于子层进行模式选择,因此可以利用父层的帧内模式对子层的候选帧内模式进行修剪。根据父层帧内模式,子层的候选帧内模式的修剪方法如图4所示。首先,对于给定的 LCU,通过 3D-HEVC 标准的帧内模式选择方法,得到 64×64 编码单元的最优模式,并存储下来。然后,对于其他大小的编码单元,先考虑其父层

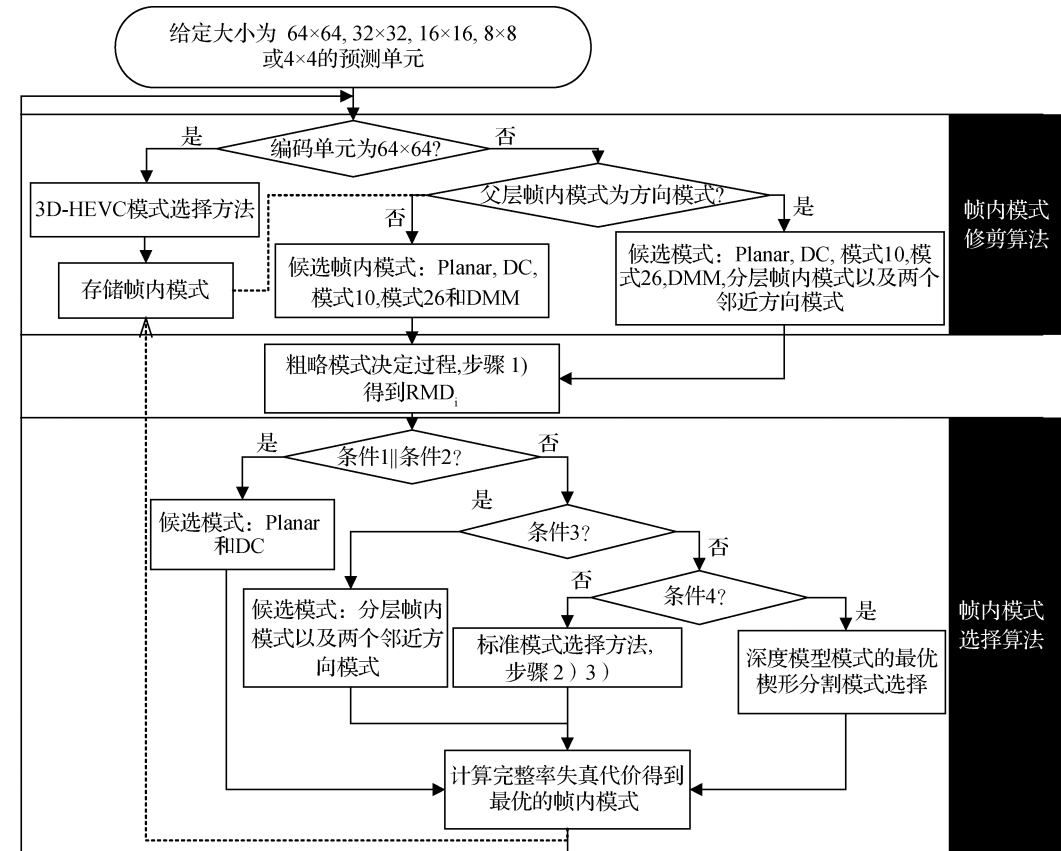


图4 本文方法流程图

Fig. 4 The flowchart of proposed algorithm

的帧内模式,如果该帧内模式不是 Angular 模式,则当前编码单元的候选帧内模式包括:Planar、DC、水平模式、垂直模式和 DMM,否则,除这 5 个模式外,候选帧内模式还包含父层的帧内模式 m 和该模式的两个邻近帧内模式 $m-1$ 和 $m+1$ 。

2.2 快速帧内模式选择方法

对帧内模式进行修剪后,根据式(1)计算每个候选帧内模式的粗略率失真代价,仍然可能有 3 个,5 个或是 8 个候选模式需要计算完整率失真代价。可见,进一步地减少候选帧内模式的数目是十分必要的,因此本文将同时考虑父层的帧内模式和粗略率失真代价来减少完整率失真代价的计算次数。因为不同的帧内模式适用不同的编码单元块,其中 Planar 和 DC 用于预测平滑区域;Angular 模式用于方向块的预测,而 DMM 是用于边界块的预测,所以本文根据父层的帧内模式和粗略率失真代价将编码单元分为平滑块、方向块和边界块。其分类的条件如下:

条件 1,父层帧内模式为 Planar 模式,并且子层的粗略率失真代价最小的两个模式为 Planar 模式和 DC 模式。

条件 2,父层的帧内模式为 DC 模式,并且子层的粗略率失真代价最小的两个模式为 DC 模式和 Planar 模式。满足条件 1 或是条件 2 的编码块为平滑块。

条件 3,父层帧内模式为 Angular 模式 m ,并且粗略率失真代价最小的 3 个模式为 $m, m+1$ 和 $m-1$,同时需要满足 RMD_0 的率失真代价小于某个阈值 T_1 。满足该条件的编码单元为方向块。

条件 4,父层的帧内模式为 DMM,而 RMD_0 的粗略率失真代价大于某个阈值 T_2 。满足该条件的编码块属于边界块。对于平滑块、父层和子层的编码单元都属于平滑区域, RMD_0 和 RMD_1 (Planar 模式和 DC 模式)可以很好地预测当前块,所以跳过 Angular 模式和深度模型模式,只有 Planar 和 DC 被选为候选帧内模式。方向块的编码单元有很强的方向性,故跳过 Planar 和 DC 模式,只考虑 Angular 模式。同时, RMD_0 的粗略率失真代价小于某个阈值 T_1 ,说明了 HEVC 的 Angular 模式可以很好地预测当前块,因此将不考虑深度模型模式。边界块包含锐利的边缘且 HEVC 帧内模式未能给出该类块一个可以接受的预测结果,因此对于此类编码块跳过所有的 HEVC 帧内模式,只考虑深度模型模式。

表 3 给出了本文提出的 4 个分类条件对应的候选帧内模式集和其预测的准确度,其中预测准确度为通过标准帧内模式选择方法得到的最优帧内模式包含在某个条件对应的候选帧内模式集的编码单元个数和满足该条件的总编码单元个数的百分比。首先预先对所有的帧内模式进行分类,而后通过粗略信息对编码单元进行合理的假设和分类,并为每一类编码块选择与之特征相符的模式类,因此如表 3 所示提出的 4 个分类条件的准确度的平均值超过了 92%。然而,Newspaper 序列的准确度仅为 85%,原因在于该序列的前景较多并且纹理复杂多变,难以预测,使得分类的准确性低。

表 3 本文提出的 4 个条件对应的候选模式和准确度

Table 3 The mode candidates accuracy rate of proposed conditions

测试序列	候选帧内模式				/%
	条件 1	条件 2	条件 3	条件 4	
	RMD_0 RMD_1	RMD_0 RMD_1	RMD_0 RMD_1 RMD_2	RMD_1 DMM	
Kendo	90.9	92.3	93.7	94.9	
Balloons	92.4	92.7	89.5	92.7	
Newspaper	89.3	82.6	81.4	89.7	
Dancer	95.6	94.2	93.1	94.3	
Fly	96.3	95.8	96.2	95.9	
Hall2	96.1	98.6	94.2	97.1	
Street	95.3	93.4	92.7	95.4	
Shark	92.3	94.3	97.1	94.9	
平均值	93.5	93.0	92.2	94.4	

条件 3 和条件 4 分别有一个阈值,其本质在于通过选取适当的最小率失真代价来控制 HEVC 帧内模式或深度模型模式是否作为候选模式。在计算粗略率失真代价时,式(1)中的 λ 与量化参数(QP)有关,即

$$\lambda = F_s \times 2^{(QP-12)/3} \quad (2)$$

式中, F_s 为伸缩因子, QP 为量化参数。其目的是为了将新的基于视角合成失真的方法整合到 HEVC 原有的率失真优化方案(RDO)。本文根据 λ 关于 QP 的分布来确定阈值 T_1 和 T_2 ,即

$$T_1 = 6 \times F_s \times 2^{(QP-12)/3} \quad (3)$$

$$T_2 = 12 \times F_s \times 2^{(QP-12)/3} \quad (4)$$

深度模型模式的候选模式数目多,意味着该模

式需要更多的比特来表示, 比特开销大。因此当编码系统对比特率提出更高的要求或 QP 增大时, 应该减少使用深度模型模式, 以节约比特数, 即需要放宽跳过深度模型模式的条件。由式(3)(4)可知, 阈值 T_1 和 T_2 的值可以根据量化参数自适应地调整, 这使得条件3和条件4可以根据编码器对失真和比特率的不同需要自适应地选择深度模型模式。例如, 当 QP 增大时, T_1 和 T_2 的值都将增大, 这意味着有更多的编码单元符合条件3, 同时符合条件4的编码单元数将减少, 最终只有少数的编码单元选择深度模型模式作为候选模式, 以满足编码器对比特率的要求, 同时降低了计算复杂度; 反之, 当 QP 减小时, T_1 和 T_2 可以通过条件3和条件4决定更多的编码单元选择深度模型模式, 以满足编码器对视频质量的要求, 同时也增加了模式选择的计算量。此外, 因为 HEVC 采用了基于分层的自适应量化参数调整技术, 可以在视频帧、片、CU 层对量化参数进行补偿, 所以阈值 T_1 和 T_2 也具有较好的分层自适应性, 可以很好地权衡方法复杂度和率失真。

基于以上的分析和论述, 快速帧内模式选择方法如图4所示, 首先计算经父层帧内模式修剪后的候选帧内模式的粗略率失真代价。其次, 同时考虑父层模式和粗略率失真代价及其对应模式 RMD_i , 判断当前块是否满足本文提出的4个条件。如果当前块满足4个条件之一, 则根据所满足的条件按表3选择对应的候选帧内模式; 否则, 将所有的 RMD_i 都视为候选帧内模式。如果当前块的候选模式中包含深度模型模式, 则根据原始方法第2)步得到最优的深度模型模式。最后, 计算所有的候选帧内模式的完整率失真代价, 从而得到最优的帧内模式。

3 实验结果与分析

为了评价本文方法的编码性能, 实验中在 3D-HEVC 的参考软件 HTM-13.0^[18] 实现本文方法以及对比方法 (Park's^[9])。最新的测试模型 HTM-13.0 与先前版本有很大改动, 因此本文方法仅和最新的文献 [9] 进行比较。值得注意的是文献 [6, 8, 12] 的方法已经被 HTM-13.0 采纳, 因此与 HTM-13.0 的比较结果等同与这3个方法的比较结果。实验测试序列包含8个标准测试序列, 分别为 Kendo、Bal-

loons、Newspaper、Dancer、GT_Fly、Hall2、Street 和 Shark。因为本文方法针对于深度图像帧内编码, 所以实验的配置条件为全 I 帧 (all_intra_main)。实验中纹理图像的量化参数设置为 25、30、35 和 40, 对应的深度图像的量化参数设置为 34、39、42 和 45。所有的实验均在配置为 CPU Intel Xeon(R) E3-1230 @ 3.3 GHz、RAM 16.0 GB 的 PC 机上进行。实验中采用与表2相同的 BDBR 定义来衡量各方法的率失真, 即由纹理和深度视频的总比特率和虚拟视点的 PSNR 计算得到虚拟合成视角的平均比特率。另外与原始方法相比, 深度图像编码时间减少的百分比 ΔT 被用来评价方法的复杂度。本文方法和对比方法分别和原始 HTM-13.0 比较, 得到 BDBR 的增加量和平均编码时间的减少量。

表4包含3个实验结果, 从表4可以看出, 本文的帧内模式修剪方法可以节省 24% 的帧内编码时间, 其主要的贡献来自于本文的方法仅选择 5 个或 8 个候选模式进行粗略率失真的计算, 同时对于较小的编码单元有时仅需对 5 个粗略帧内模式计算完整率失真代价, 而原始方法需要 8 个。本文的帧内模式选择可以减少 38% 的编码时间, 因为对编码块进行了分类, 可以根据当前块的分类结果选择 1 个、2 个或 3 个帧内候选模式, 同时为平滑块和方向块跳过最优深度模型模式的计算。本文方法总体可以减少 44% 的计算复杂度, 通过对比 3 个实验结果, 可以发现对于测试序列 Fly、Dancer、Hall2 帧内模式选择方法和总方法的编码性能相似, 这是因为这些测试序列中包含的纹理细节少, 大多数的编码块可以满足帧内模式快速选择方法的 4 个条件。然而, 对于其他序列, 帧内模式修剪方法利用父层的帧内模式对子层的帧内模式进行了修剪, 使得更多的子层编码块可以满足帧内模式选择方法所提的 4 个条件, 因此总方法能够更好地提高编码性能。

如表4所示, 对比方法^[9]仅节省 5% 的编码时间, 这是因为该方法需要在哈达玛变换域上对水平和垂直边界进行检测, 根据检测结果跳过深度模型本文方法, 而且 HTM-13.0 的快速方法^[8]与对比方法^[9]的思路相似, 对于平滑区域的处理两方法存在重叠, 所以对比方法^[9]在 HTM-13.0 上的编码效果提升并不明显。值得注意的是, 因为对比方法文献 [9] 需要计算哈达玛变换来检测水平和垂直边界并跳过深度模型模式, 而测试序列 Street 和 Shark 中包

含水平和垂直信息的边界块很少,这使得对比方法文献[9]对这两个序列的编码效率较低,无法获得增益。因为使用哈达玛变换进行的检测消耗了多余的时间,将对比方法应用到 HTM-13.0 中,在这两个序列上反而增加了 6% 和 3% 的编码时间。表 4 表明,本文方法有效降低了深度帧内模式的数目,减少了计算复杂度,提高了编码速度;本文方法对于高分

分辨率的视频序列效果更为显著,因为分辨率较大的视频序列包含更多的平滑区域,所以有较多的编码单元满足条件 1 或条件 2,使用本文方法可以减少候选帧内模式。本文方法和文献[6, 8, 12] (HTM-13.0)一起使用,仅平均增加 0.85% 的比特率,同时减少了 44% 的复杂度,因此本方法具有很好的加速效果和兼容性。

表 4 本文模式选择方法和文献[9]的编码性能比较结果

Table 4 The comparison between the proposed mode decision method and Park's algorithm^[9]

测试序列	分辨率/像素	文献[9]		本文的帧内模式修剪方法		本文的帧内模式选择方法		本文的总方法	
		BDBR	ΔT	BDBR	ΔT	BDBR	ΔT	BDBR	ΔT
Kendo	1 024 × 768	0.08	-10.2	0.65	-23.16	0.87	-31.29	0.94	-39.47
Balloons	1 024 × 768	0.03	-8.42	0.65	-23.48	0.69	-30.03	0.97	-40.6
Newspaper	1 024 × 768	0.16	-4.26	0.9	-23.15	1.05	-30.13	1.33	-37.72
Dancer	1 920 × 1 088	0.14	-5.73	0.28	-23.82	0.52	-42.01	0.71	-47.36
Fly	1 920 × 1 088	0.02	-6.10	0.38	-30.38	0.75	-51.14	0.79	-55.3
Hall2	1 920 × 1 088	0.47	-16.15	0.45	-24.87	0.64	-41.07	0.74	-44.94
Street	1 920 × 1 088	0.00	6.51	0.28	-23.73	0.46	-40.17	0.54	-43.22
Shark	1 920 × 1 088	0.00	3.25	0.51	-22.69	0.43	-37.02	0.78	-42.97
平均值	—	0.12	-5.14	0.51	-24.41	0.68	-37.86	0.85	-43.95

4 结 论

为了更好地编码深度图像,3D-HEVC 中引入了深度模型模式、分段 DC 残差编码等新技术,有效地提高了虚拟视点的视觉质量,但同时也增加了帧内编码的复杂度。本文提出了基于分层结构的深度帧内模式修剪方法,通过统计分层编码结构中父层帧内模式和子层的候选帧内模式的概率分布,利用父层帧内模式对候选帧内模式进行修剪,从而使得子层编码单元跳过一些非候选模式的粗略率失真计算。其次,本文通过粗略率失真代价对编码块进行分类,不同类的编码块采用不同的模式跳过策略,有效减少了需要计算率失真代价的模式数目。实验结果表明,与原始的方法相比,本文方法在几乎不降低合成视角质量的同时,节省了约 44% 的编码时间。

参考文献 (References)

[1] ISO/IEC. ISO/IEC 23002-3:2007 Information technology-MPEG

video technologies-Part 3: representation of auxiliary video and supplemental information[S]. Switzerland: ISO/IEC, 2007.

- [2] Müller K, Merkle P, Tech G, et al. 3D video formats and coding methods[C]//Proceedings of the 17th International Conference on Image Processing. Hong Kong, China: IEEE, 2010: 2389-2392. [DOI: 10.1109/ICIP.2010.5652030]
- [3] Kauff P, Atzpadin N, Fehn C, et al. Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability[J]. Signal Processing: Image Communication, 2007, 22(2): 217-234. [DOI: 10.1016/j.image.2006.11.013]
- [4] Müller K, Schwarz H, Marpe D, et al. 3D high-efficiency video coding for multi-view video and depth data[J]. IEEE Transactions on Image Processing, 2013, 22(9): 3366-3378. [DOI: 10.1109/TIP.2013.2264820]
- [5] Sullivan G J, Ohm J R, Han W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1649-1668. [DOI: 10.1109/TCSVT.2012.2221191]
- [6] Merkle P, Müller K, Zhao X, et al. JCT3V-B0039 Simplified wedgelet search for DMM modes 1 and 3, ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11[S]. Korea: ISO/IEC, 2014.
- [7] Fu C H, Zhang H B, Su W M, et al. Fast wedgelet pattern decision for DMM in 3D-HEVC[C]//Proceedings of IEEE Interna-

- tional Conference on Digital Signal Processing. Singapore: IEEE, 2015: 477-481. [DOI:10.1109/ICDSP.2015.7251918]
- [8] Gu Z Y, Zheng J H, Ling N, et al. Fast depth modeling mode selection for 3D HEVC depth intra coding[C]//Proceedings of IEEE International Conference on Multimedia and Expo Workshops. San Jose, California: IEEE, 2013: 1-4. [DOI: 10.1109/ICMEW.2013.6618267]
- [9] Park C S. Edge-based intra mode selection for depth-map coding in 3D-HEVC[J]. IEEE Transaction on Image Processing, 2015, 24(1): 155-162. [DOI: 10.1109/TIP.2014.2375653]
- [10] Zhang H B, Fu C H, Chan Y L, et al. Efficient depth intra mode decision by reference pixels classification in 3D-HEVC [C]//Proceedings of IEEE International Conference on Image Processing. Quebec, QC, Canada: IEEE, 2015: 961-965. [DOI: 10.1109/ICIP.2015.7350942]
- [11] Gu Z Y, Zheng J H, Ling N, et al. Fast bi-partition mode selection for 3D HEVC depth intra coding[C]//Proceedings of IEEE International Conference on Multimedia and Expo. Chengdu, Sichuan, China: IEEE, 2014: 1-6. [DOI: 10.1109/ICME.2014.6890324]
- [12] Tech G, Schwarz H, Müller K, et al. 3D video coding using the synthesized view distortion change[C]//Picture Coding Symposium. Krakow, Poland: IEEE, 2012: 25-28. [DOI: 10.1109/PCS.2012.6213277]
- [13] Lei J J, Li S, Zhu C, et al. Depth coding based on depth-texture motion and structure similarities[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(2): 275-286. [DOI: 10.1109/TCSVT.2014.2335471]
- [14] Lei J J, Sun J, Pan Z Q, et al. Fast mode decision using inter-view and inter-component correlations for multiview depth video coding[J]. IEEE Transactions on Industrial Informatics, 2015, 11(4): 978-986. [DOI: 10.1109/TII.2015.2446769]
- [15] Pan Z Q, Kwong S, Sun M T, et al. Early merge mode decision based on motion estimation and hierarchical depth correlation for HEVC[J]. IEEE Transactions on Broadcasting, 2014, 60(2): 405-412. [DOI: 10.1109/TBC.2014.2321682]
- [16] Shen L Q, An P, Liu Z, et al. Low complexity depth coding assisted by coding information from color video[J]. IEEE Transactions on Broadcasting, 2014, 60(1): 128-133. [DOI: 10.1109/TBC.2013.2289635]
- [17] Kang M K, Ho Y S. Depth video coding using adaptive geometry based intra prediction for 3-D video systems[J]. IEEE Transactions on Multimedia, 2012, 14(1): 121-128. [DOI: 10.1109/TMM.2011.2169238]
- [18] MPEG 3D-HTM Software V13.0. [CP]. [2015-09-08] <http://hevc.kw.bbc.co.uk/git/w/jetvc-3de.git>