

基于边缘检测的 3D-HEVC 深度图运动估计算法

谢晓燕^a, 辛晓斐^a, 朱 筠^b, 王飞龙^b, 刘 阳^a

(西安邮电大学 a. 计算机学院; b. 电子工程学院, 西安 710121)

摘 要: 针对 3D 高效视频编码(3D-HEVC)帧间预测未充分体现深度图特性、算法复杂度高、效率低的问题, 提出一种基于深度图边缘检测的运动估计算法。对深度图进行边缘检测预处理, 分别对边缘区域和平坦区域执行全搜索算法和六边形搜索算法, 通过快速搜索平坦区域以降低帧间预测中运动估计 SAD 的计算复杂度。在 3D-HEVC/HTM16.0 平台上的测试结果表明, 该算法减少了 6.7% 的深度图编码时间, 而合成视图的 BD-rate 指标仅为 0.146%, 有效提高了编码效率。

关键词: 高效视频编码; 深度图; 运动估计; 边缘检测; 六边形搜索

中文引用格式: 谢晓燕, 辛晓斐, 朱筠, 等. 基于边缘检测的 3D-HEVC 深度图运动估计算法[J]. 计算机工程, 2019, 45(7): 264-267.

英文引用格式: XIE Xiaoyan, XIN Xiaofei, ZHU Yun, et al. 3D-HEVC depth map motion estimation algorithm based on edge detection[J]. Computer Engineering, 2019, 45(7): 264-267.

3D-HEVC Depth Map Motion Estimation Algorithm Based on Edge Detection

XIE Xiaoyan^a, XIN Xiaofei^a, ZHU Yun^b, WANG Feilong^b, LIU Yang^a

(a. School of Computer Science and Technology; b. School of Electronic Engineering,
Xi'an University of Post and Telecommunications, Xi'an 710121, China)

[Abstract] To address the problem that the depth map characteristics are not fully reflected, and the algorithm has high complexity and low efficiency in 3D High Efficiency Video Coding (3D-HEVC) inter prediction, this paper proposes a motion estimation algorithm based on depth map edge detection. First, the edge detection preprocessing is performed on the depth map. Then, according to the results, the full search algorithm and the hexagon search algorithm are respectively performed on the edge region and the flat region. The fast search of flat areas can be implemented to reduce the computational complexity of motion estimation SAD in inter prediction. Test results on the 3D-HEVC/HTM16.0 platform indicate that the proposed algorithm reduces the coding time of the depth map by 6.7%, while the Band rate of the synthetic view is only 0.146%, and the coding efficiency is improved significantly.

[Key words] High Efficiency Video Coding (HEVC); depth map; motion estimation; edge detection; hexagon search
DOI: 10.19678/j.issn.1000-3428.0051505

0 概述

通过两个或多个摄像机在不同位置同时拍摄同一场景得到的 3D 视频, 其画面能提供所观察物体的深度信息, 给用户带来立体感的视觉体验。3D 视频技术的研究包括 3D 视频捕获、编码、压缩、传输等各个方面。在 3D 视频编码方面, 基于多视点视频加深度 (Multiview Video plus Depth, MVD) 的视频编码方法是当前的研究热点。但是

这些算法的开发着重于纹理属性, 而深度图像具有尖锐的边缘信息和大面积区域值相近的特性, 因此, 为了提高帧间预测总编码效率, 需要研究符合深度图特性的编码技术。

边缘信息是图像比较稳定的特征, 不会受外界的影响, 因此可以在运动估计中加入边缘信息来加快编码块的搜索。近年来, 国内外学者提出了一些基于边缘检测的深度图帧间预测的方法, 其中包括根据深度图边缘信息跳过一些模式的方法、新的边

基金项目: 国家自然科学基金 (61772417, 61602377, 61634004, 61272120); 陕西省科技统筹创新工程项目 (2016KTZDGY02-04-02); 陕西省重点研发计划 (2017GY-060)。

作者简介: 谢晓燕 (1972—), 女, 教授, 主研方向为图像处理、视频编解码算法; 辛晓斐, 硕士研究生; 朱 筠, 博士研究生; 王飞龙、刘 阳, 硕士研究生。

收稿日期: 2018-05-09 **修回日期:** 2018-06-10 **E-mail:** xxy@xupt.edu.cn

边缘检测算法和简化边缘检测器(Simplified Edge Detector, SED)的方法等。文献[1-3]根据物体边缘属性自适应地确定当前深度图编码单元(Coding Unit, CU),小分区适用于物体边缘区域的CU,而较大分区适合没有边缘区域的CU,提出可以跳过一些均匀区域中较少使用的预测模式,但是这种方法通过纹理图去预测深度图,对深度图的处理存在一定的偏差。文献[4]采用基于深度的前景分割,联合双边滤波插值的方法,实现一种融合彩色图的深度图增强方法,更好地保持边缘信息,但还存在其他非前景边缘不对称问题,其深度图质量需进一步提高。文献[5-6]采用Sobel算子边缘检测对CU内部的边缘信息进行提取,通过阈值来比较这些边缘值,预测CU的最佳预测模式,但是对于多噪声的图像,Sobel算子的处理效果不理想。文献[7]利用相应纹理图的边缘,使用一种新的边缘检测和深度图绘制算法来重建上采样中深度图边缘,提高了边缘信息的精度,提升了编码质量,但编码的计算复杂度降低不明显。文献[8-9]使用简化运动估计算法进行帧间预测,虽然提升了编码效率且降低了处理器的能耗,但需要为其设计专用硬件。

上述研究在一定程度上提高了编码效率,但这些方法都没有从运动估计的搜索算法角度对编码复杂度进行优化。在深度图中边缘区域比平坦区域更难编码,因此,边缘区域应使用更精细的搜索算法,以在不损失编码质量的同时提升编码效率。本文利用深度图的特性,提出一种运动估计算法,通过判断搜索区域是否为边缘区域而执行不同的搜索算法,以提取更精细的边缘区域。

1 深度图运动估计算法分析

1.1 深度图特点

3D高效视频编码(High Efficiency Video Coding, HEVC)采用MVD格式,除纹理外还使用深度信息,深度图与纹理帧具有一定的相关性。在实际应用中,通常利用红外传感器与摄像机的纹理帧相结合,获取深度图。这些贴图提供了物体与相机之间的几何信息,与纹理帧所呈现的复杂变化相比,深度图反映出具有大面积均匀区域和尖锐边缘的特征。图1(a)为测试序列Balloon的纹理图,图1(b)为对应的深度图,其中,1~4为提取的平坦区域,提取块中颜色均无变化,5~8为提取的边缘区域,其像素值有明显的边界变化。从图1可以看出,深度图含有大面积背景以及尖锐的边缘区域。针对这些平坦区域,可以使用快速搜索算法来降低运动估计算法的复杂度,从而提高编码效率。

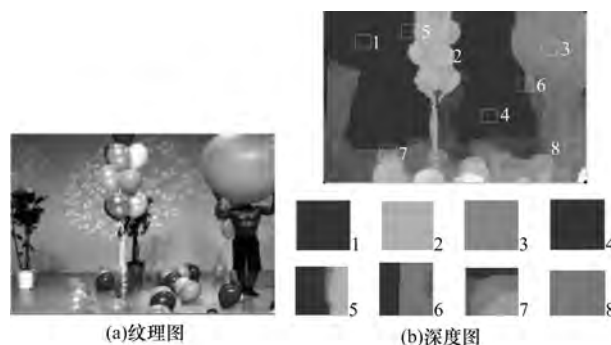


图1 Balloon序列纹理图和深度图

1.2 运动估计搜索算法分析

类似于纹理预测,深度图运动估计的目的是为当前帧中的每个块在参考帧中搜索到最佳匹配块。表1为在3D-HEVC/HTM12.0平台上统计视频编码各部分的时间占比^[10],其中,深度图编码占到总编码时间的47.50%,帧间预测占深度图编码时间的19.74%。

表1 视频编码各部分的时间占比 %

编码帧	功能	时间占比
纹理图 (52.50%)	帧间预测	66.01
	其他	33.99
深度图 (47.50%)	帧内预测	21.01
	DMM1	10.16
	DMM4	3.34
	帧间预测	19.74
	其他	45.75

3D-HEVC/HTM平台采用TZ搜索算法对纹理帧和深度图进行编码,在搜索过程中使用不同的搜索方法,包括菱形搜索和光栅搜索。它可以达到一个接近最优的搜索性能。尽管TZ搜索算法的复杂度仅为全搜索(Full Search, FS)算法的1/23,而且几乎没有编码损失^[11],但是考虑到深度图中存在的大面积平坦区域,可以使用比TZ搜索更简单的快速搜索算法来对平坦区域进行搜索^[12],从而降低运动估计算法的复杂度并保持良好的编码效率。

2 基于边缘检测的深度图运动估计

2.1 运动估计中SAD值分布

为了比较纹理图以及深度图的绝对误差和(Sum of Absolute Differences, SAD)分布,本文分析depth_balloon、depth_Newspaper、Poznan_Street 3种3D-HEVC测试序列,并给出其平均纹理图及对应深度图的SAD值热图,如图2所示。从图2可以看出,纹理图的SAD值分布比较复杂,并且存在局部最优值,快速算法很难找到最匹配块。然而,对应深度图的SAD值分布就比较规律,存在大面积的平滑区域,因此快速算法可以较快找到最优的SAD值,即使找不到最优值,误差也较小。

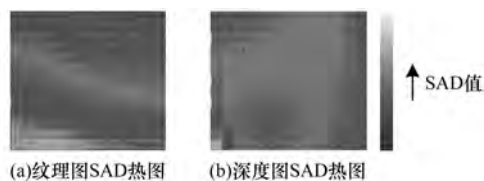


图 2 纹理图和深度图的 SAD 热图

在深度图测试序列中提取其边缘区域及平坦区域 SAD 值,如图 3 所示。图 3(a)为 Balloon 测试序列的第 30 帧深度图,其中,大的像素块为 80×80 ,小的像素块为 16×16 搜索块;图 3(b)为实线块的 SAD 热图;图 3(c)为虚线块的 SAD 热图。

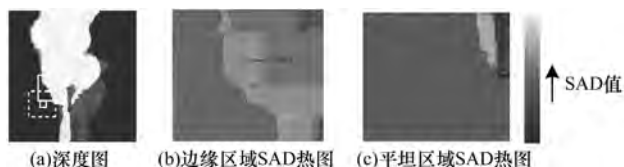


图 3 边缘、平坦区域 SAD 热图

从图 3 可以看出,平坦区域的深度图 SAD 值较低且集中分布,因此无需对此区域进行复杂的 TZ 搜索,快速搜索算法就能到最优的 SAD 值,从而达到提高图像质量的情况下减少搜索次数的目的。

2.2 算法步骤

基于深度图的特性,深度图中的平坦区域占比达到 85%,因此对于平坦区域进行计算量极大的率失真过程是冗余的^[13]。边缘区域 SAD 值分布情况较复杂,而平坦区域 SAD 值分布集中,本文对深度图进行边缘检测预处理,对不同的区域执行不同的搜索算法。

本文算法的过程为:

步骤 1 判断当前编码帧是否为深度图。如果是深度图,算法执行步骤 2;否则退出算法。

步骤 2 使用 Canny 算子对深度图进行边缘检测,采用双阈值判断其是否是边缘。其中,高、低阈值的选取采用直方图统计策略,经过多次实验分析,得到最终的低阈值为 0.46,高阈值为 0.82。

步骤 3 对边缘检测后的图像进行二值化处理,对每个 16×16 搜索块的像素进行查找搜索。

步骤 4 若搜索块中的像素全为 0,则判断该块为平坦区域,对该块进行六边形搜索 (HS) 算法;若搜索块中的像素不全为 0,则判断该块为边缘块,对该块执行 FS 算法。

3 实验与结果分析

3.1 搜索算法比较

本文采用 16×16 固定块进行块搜索,通过运动估计搜索算法中常用的 FS、四步搜索 (4SS)、菱形搜索 (DS)、HS 算法,对 7 种标准测试序列前 50 帧深度图进行了测试比较,3 种快速搜索算法结果如图 4

所示。其中, cost 表示平均每帧搜索的次数,平均搜索次数越低,算法的复杂度越低;平均峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 反映块匹配算法的精度,其值越高,搜索算法精度越高。因此,需要在保证 PSNR 的前提下,使用搜索次数较少的搜索算法来降低运动估计的复杂度。

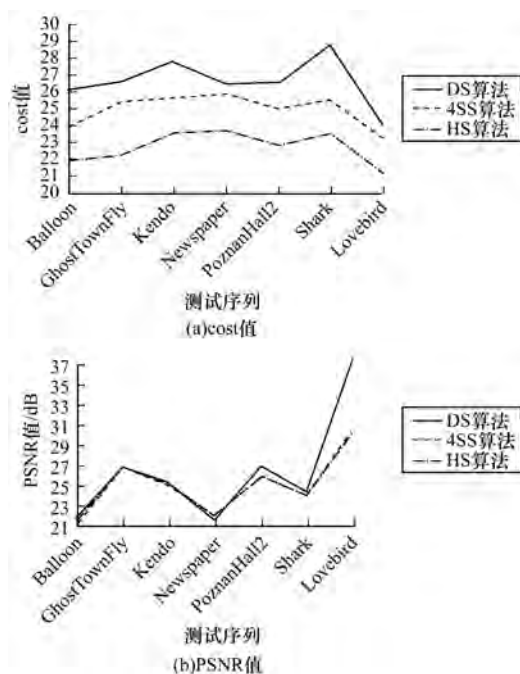


图 4 3 种序列的 cost,PSNR 值对比

从图 4 可以看出,DS 算法的平均搜索次数较高,4SS 算法其次,HS 算法最少;4SS 算法与 HS 算法的 PSNR 值接近,DS 算法的 PSNR 值较高。从测试序列分析,细节较多的序列其 PSNR 值较小,较平坦的序列其 PSNR 值较大。而在实验中,FS 算法搜索次数的范围为 800 ~ 900,远大于 3 种快速搜索算法。

3.2 有效性测试

本文在 3D-HEVC/HTM16.0 平台上进行算法有效性测试,实验参数的配置环境为标准测试环境^[14]。测试条件为:编码结构采用分层 B 帧, GOP 长度为 8, I 帧编码周期为 24,运动估计和视差估计范围为 64,LCU 的分割深度为 0 ~ 3,LCU 的大小为 64×64 ,编码帧数为 50,5 种测试序列、编码左视点、编码右视点和绘制虚拟视点的选择见表 2。

表 2 测试序列与参数选择

序列	左视点	右视点	虚拟视点
Balloon	1	5	3
Newspaper	2	6	4
Kendo	1	5	3
Shark	5	7	6
PoznanHall2	1	9	5

实验结果如表 3 所示。本文算法对深度图的编码时间减少了 6.7%,对于整个编码器来说,也就是纹理图和深度图减少了 2.9% 的编码时间,但这是以牺牲 BD-rate 为代价的。BD-rate 的值为 0.018% ~ 0.321%,其中,Balloon 序列的 BD-rate 最小,这是因为该视频运动比较缓慢,可以较好地区分平坦和边缘区域;Shark 序列的 BD-rate 最大,这是因为该视频中鲨鱼的运动较快且含有很多细节,使得边缘区域较难预测。

表 3 在 HTM16.0 平台上有效性测试结果 %

序列	BD-rate	节省的编码时间	
	合成视图	纹理图和深度图	深度图
Balloon	0.018	3.1	7.3
Newspaper	0.102	3.2	6.9
Kendo	0.021	3.1	7.5
Shark	0.321	2.6	5.7
PoznanHall2	0.268	2.8	6.3
平均	0.146	2.9	6.7

文献[15]使用 TZ 搜索算法和迭代小菱形搜索算法对深度图编码中帧间预测运动估计算法进行优化,表 4 是本文算法与文献[15]算法的对比数据,可以看出,本文算法对深度图节省的编码时间比文献[15]略低 0.2%,但是合成视图的 BD-rate 性能略有提升。实验结果表明,本文算法在一定程度上节省了 3D-HEVC 帧间预测中运动估计算法的编码时间。

表 4 对比实验结果 %

算法	BD-rate	节省的编码时间	
	合成视图	纹理图和深度图	深度图
文献[15]算法	0.148	3.2	6.9
本文算法	0.146	2.9	6.7

4 结束语

基于深度图边缘检测,本文提出一种 3D-HEVC 帧间预测中运动估计的优化算法,对不同的区域执行不同的搜索算法,在 BD-rate 值较小的情况下,实现平坦区域的快速搜索来降低帧间预测中运动估计 SAD 的计算复杂度。在 3D-HEVC/HTM 16.0 平台上对本文算法进行测试,结果表明,合成视图 BD-rate 平均值为 0.146%,纹理和深度图的编码时间平均减少 2.9%,深度图编码时间平均减少 6.7%。

参考文献

- [1] ZHANG Qiuwen, ZHANG Na, WEI Tao, et al. Fast depth map mode decision based on depth-texture correlation and edge classification for 3D-HEVC [J]. Journal of Visual Communication and Image Representation, 2017, 45: 170-180.
- [2] ZHANG Qiuwen, HUANG Kunqiang, WANG Xiao, et al. Efficient multiview video plus depth coding for

3D-HEVC based on complexity classification of the treeblock [J]. Journal of Real-Time Image Processing, 2017(4):1-18.

- [3] LEI Jianjun, DUAN Jinhui, WU Feng, et al. Fast mode decision based on grayscale similarity and inter-view correlation for depth map coding in 3D-HEVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(3):706-718.
- [4] 刘金荣,李淳芃,欧阳建权,等.基于联合双边滤波的深度图像增强算法 [J]. 计算机工程, 2014, 40(3): 249-252.
- [5] TSAI T H, SU Shengshuan, LEE T Y. Fast mode decision method based on edge feature for HEVC inter-prediction [J]. IET Image Processing, 2018, 12(5): 644-651.
- [6] ZHANG Qiuwen, CHEN Ming, HUANG Xinping, et al. Low-complexity depth map compression in HEVC-based 3D video coding [EB/OL]. [2018-04-27]. <https://link.springer.com/article/10.1186/s13640-015-0058-5>.
- [7] ZAMMIT L, DEBONO C J. Improved reconstruction of down sampled MV-HEVC depth video [C]// Proceedings of IEEE Eurocon 2017 International Conference on Smart Technologies. Washington D. C., USA: IEEE Press, 2017:176-180.
- [8] SANCHEZ G, SALDANHA M, PORTO M, et al. Real-time simplified edge detector architecture for 3D-HEVC depth maps coding [C]// Proceedings of IEEE International Conference on Electronics, Circuits and Systems. Washington D. C., USA: IEEE Press, 2017: 352-355.
- [9] 杨媛,陈福.图像深度估计硬件实现算法 [J]. 中国图象图形学报, 2018, 23(3):362-371.
- [10] SALDANHA M, SANCHEZ G, ZATT B, et al. Energy-aware scheme for the 3D-HEVC depth maps prediction [J]. Journal of Real-Time Image Processing, 2017, 13(1):55-69.
- [11] KUO Pinchen, LU Kuanhsing, HSU Y N, et al. Fast three-dimensional video coding encoding algorithms based on edge information of depth map [J]. IET Image Processing, 2015, 9(7):587-595.
- [12] SALDANHA M, SANCHEZ G, ZATT B, et al. Complexity reduction for the 3D-HEVC depth maps coding [C]// Proceedings of 2015 IEEE International Symposium on Circuits and Systems. Washington D. C., USA: IEEE Press, 2015:621-624.
- [13] ZHANG Qiuwen, CHEN Ming, ZHU Haodong, et al. An efficient depth map filtering based on spatial and texture features for 3D video coding [J]. Neurocomputing, 2016, 188:82-89.
- [14] SALDANHA M, SANCHEZ G, MARCON C, et al. Block-level fast coding scheme for depth maps in three-dimensional high efficiency video coding [EB/OL]. [2018-04-27]. <https://doi.org/10.1117/1/JEI.27.1.010502>.
- [15] SANCHEZ G, SALDANHA M, ZATT B, et al. Edge-aware depth motion estimation—a complexity reduction scheme for 3D-HEVC [C]// Proceedings of the 25th European Signal Processing Conference. Washington D. C., USA: IEEE Press, 2017:1524-1528.

编辑 刘盛龄