

基于区域划分的深度视频快速编码算法

田 涛, 彭宗举*

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

(*通信作者电子邮箱 pengzongju@nbu.edu.cn)

摘 要:多视点彩色视频与深度视频作为多媒体系统中主流 3D 场景表示方式,吸引了越来越多人的关注。深度视频反映场景的几何信息,如何对其进行快速编码尤为重要。提出了一种基于区域划分的深度视频快速编码算法。首先,根据深度视频的边缘和运动属性把深度视频分为四个区域;然后,深入分析不同区域内宏块模式分布比例以及参考帧选择特性,设计了不同的宏块模式选择和参考帧搜索策略来提高编码速度;最后,测试提出算法的编码时间、码率以及合成虚拟视点的质量。实验结果表明,提出算法在保证虚拟视点图像质量和编码码率基本不变的情况下,节约了 85.73%~91.06% 的编码时间。

关键词:三维;深度视频;视频编码;区域划分;快速编码算法

中图分类号:TN919.8 **文献标志码:**A

Fast depth video coding algorithm based on region division

TIAN Tao, PENG Zongju*

(Faculty of Information Science and Engineering, Ningbo University, Ningbo Zhejiang 315211, China)

Abstract: As the main scheme of 3D scene representation, multiview video plus depth attracts more and more attention. Depth video reflects the geometric information of the scene. It is important to design fast depth video encoding algorithm. A fast depth video coding algorithm based on region division was proposed. Firstly, the depth video was divided into four regions according to the features of edge and motion. Then, macroblock distribution proportion and multi-reference frame selection feature of different regions were analyzed. Consequently, different macroblock mode decision and reference frame selection methods were utilized to speedup depth video encoding. Finally, some experiments were conducted to estimate the proposed algorithm in terms of encoding time, bit rate and virtual view quality. Experimental results show that the proposed algorithm saves encoding time ranging from 85.73% to 91.06% while it maintains virtual view quality and bit rate.

Key words: 3D; depth video; video coding; region division; fast encoding algorithm

0 引言

随着三维显示及其相关技术的发展,三维视频能够提供全新的立体感体验,满足观赏者对深度感、交互式和立体感需求^[1-2]。多视点彩色视频与深度视频(Multiview Video plus Depth, MVD)作为三维场景信息的主要表示方式,可以完整地呈现出视角范围较广、深度层次丰富的场景信息,且能够利用基于深度图像绘制技术生成其他视点视频信号^[3-4]。然而,MVD需要编码的数据量随着视点数目的增加而成倍地增加^[5]。因此,为了解决多视点视频海量数据的传输和存储问题,需要对多视点视频数据进行高效压缩。两大国际标准化组织 ISO 和 ITU 组成的联合视频专家组提出了用于多视点视频编码的 JMVC 校验平台^[6]。JMVC 平台对每一个编码宏块采用可变块模式和多参考帧搜索技术。每个编码帧的宏块模式又分为帧内预测模式和帧间预测模式,帧内预测模式分为 Intra16×16、Intra8×8 和 Intra4×4。帧间预测模式分为 SKIP、Inter16×16、Inter16×8、Inter8×16、Inter8×8 和 Inter8×8Fext。每种帧间预测模式下,需要依次遍历三个搜索方向上的多个参考帧,选取率失真代价值最小的参考帧作为最优参考帧,然后,比较所有模式的率失真代价值,率失真代价值最

小的模式即为当前编码宏块的最优编码模式。这种编码技术可以获得较高的重建图像质量,但大大增加了编码的复杂度,是 MVD 进入实际应用系统的关键问题和难点。

近年来,有许多学者开展了视频编码快速算法的研究^[7-8]。这些算法在保证编码率失真性能前提下,提高了编码的速度。深度视频反映三维场景的几何信息,深度视频编码的目标是高压压缩效率和高性能的虚拟视点绘制。和普通彩色视频相比,深度视频信号比较平滑、纹理较少。因此这些算法不适合快速编码深度视频信号。文献[9]把深度图像分为连续区域和不连续区域,根据连续区域和不连续区域对绘制质量的影响程度,采用不同的编码策略。文献[10]根据深度值的大小,把深度图像分成近区域、中间区域和远区域,每个区域内又分为遮挡区域、静止区域和运动区域,通过减少每个区域的候选模式来减少编码的时间。文献[11]根据深度图像的平滑特性提前结束最优模式选择的过程。文献[12]将深度视频分为感兴趣区域和非感兴趣区域,两种区域采取不同的编码方法以达到减少编码时间的目的。这些算法在一定程度上减少了编码时间,但只从宏块模式选择的角度进行快速编码,没有进一步结合参考帧选择的特征。

本文结合深度视频平滑、纹理较少的特点,提出了一种基

收稿日期:2012-12-11;**修回日期:**2013-01-21。 **基金项目:**国家自然科学基金资助项目(61171163);宁波市自然科学基金资助项目(2012A610041);新一代移动互联网用户端软件科技创新团队资助项目(2010R50009)。

作者简介:田涛(1988-),男,安徽太和人,硕士研究生,主要研究方向:深度视频编码;彭宗举(1973-),男,四川南部人,副教授,博士,主要研究方向:三维视频信号的处理和编码。

于区域划分的深度视频快速编码算法。首先对深度视频进行区域划分,然后对不同区域的宏块模式和参考帧选择进行了深入分析,最后设计了快速编码算法。实验结果表明,本文算法不但大大提高了编码的速度,而且很好地保证了虚拟视点的绘制质量。

1 深度图像区域划分

本文根据运动和边缘属性,将深度视频划分为四个区域:静止平坦区域、运动平坦区域、静止边缘区域和运动边缘区域。

1.1 运动静止区域的划分

深度视频主要用于虚拟视点的绘制。研究表明,深度视频中的运动区域和对象的边缘区域对绘制图像质量的影响较大,静止区域对绘制图像质量的影响较小。考虑到深度视频和彩色视频中的对象具有运动一致性,彩色视频中运动的区域,在深度视频中极有可能也是运动区域;彩色视频中静止的区域,在深度视频中极有可能也是静止区域。根据此特性,分别求取彩色视频中当前帧的宏块和前向以及后向参考帧中对应宏块的平方差值和(Sum of Square Differences, SSD),其中较小的SSD值若小于等于设定的阈值 Th ,那么判定当前宏块属于静止区域,用 SR 表示。相反,则属于运动区域,用 MR 表示。以同一视点I、B和P三帧为例,根据编码预测结构,B帧在编码的过程中参考I帧和P帧。彩色图像分别记作 T_i 、 T_b 和 T_p ,对应深度图像分别记为 D_i 、 D_b 和 D_p 。计算 T_b 中的当前宏块和 T_i 以及 T_p 中对应宏块的SSD $_{bi}$ 和SSD $_{bp}$:

$$SSD_{bi} = \sum_{i=1}^{16} \sum_{j=1}^{16} |T_b(i, j) - T_i(i, j)|^2 \quad (1)$$

$$SSD_{bp} = \sum_{i=1}^{16} \sum_{j=1}^{16} |T_b(i, j) - T_p(i, j)|^2 \quad (2)$$

其中: $T_b(i, j)$ 表示彩色视频 T_b 中当前宏块内坐标为 (i, j) 的像素值, $T_i(i, j)$ 表示 T_i 中当前宏块内坐标为 (i, j) 的像素值, $T_p(i, j)$ 表示 T_p 中当前宏块内坐标为 (i, j) 的像素值。

SR 和 MR 通过如下公式划分:

$$MB \in \begin{cases} SR, & \min \{SSD_{bi}, SSD_{bp}\} \leq Th \\ MR, & \text{其他} \end{cases} \quad (3)$$

其中: Th 是通过实验确定的, $Th = 5500$; MB 表示当前宏块。

若 T_b 中当前宏块属于 SR ,那么判定 D_b 中对应的宏块也属于 SR ,其余的宏块则属于 MR 。图1(a)为Book Arrival序列的运动区域划分结果,白色区域表示 MR ,灰色区域表示 SR 。

1.2 边缘检测

由于深度视频中对象的边缘区域对虚拟视点的质量影响较大,故本文首先提取出深度视频中对象的边缘区域,对其进行保护,以保证虚拟视点的绘制质量。本文采用Canny算子求取边缘区域,具体步骤如下:

1) 用高斯滤波器对原始图像 $d(x, y)$ 进行平滑,去除图像噪声点,得到平滑图像 $f(x, y)$ 。高斯函数表示形式为:

$$G(x, y) = \sum_{x'=-m}^{x'=x+m} \sum_{y'=-m}^{y'=y+m} \exp \left| -\frac{x'^2 + y'^2}{2\sigma^2} \right| \quad (4)$$

$$m = \frac{n-1}{2} \quad (5)$$

$$f(x, y) = G(x, y) * d(x, y) \quad (6)$$

其中: n 为高斯滤波器窗口尺寸,“*”代表卷积。

2) 对于 $f(x, y)$ 中的每一个像素利用微分算子 G_x 和 G_y 求出沿两个方向的偏导数(f_x, f_y):

$$G_x = \frac{1}{2} \times \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \quad (7)$$

$$G_y = \frac{1}{2} \times \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \quad (8)$$

求出梯度 $|F|$ 大小和方向角 θ :

$$|F| = \sqrt{f_x^2 + f_y^2} \quad (9)$$

$$\theta = \arctan \frac{f_y}{f_x} \quad (10)$$

3) 对梯度幅值采用非极大值限制。仅得到全局的梯度并不能准确确定边缘,为此必须保留局部梯度最大值点,抑制非极大值点。

4) 阈值 Th_1 和 Th_2 ($Th_1 < Th_2$) 作用于非极大值抑制后的图像,得到两幅图像。 Th_2 作用后的图像去除了大部分噪声,但也丢失了部分边缘信息; Th_1 作用后的图像有较多的边缘信息,可补充 Th_2 作用后的图像,最终得到相对完整的边缘图。 Th_1 和 Th_2 分别为50和150。

图1(b)为Book Arrival序列的边缘图。在每个 16×16 块内,如果有至少3个像素点属于边缘,那么当前块即为边缘块。图1(c)为图1(b)边缘块化后的结果,黑色区域为边界区域,用 ER 表示,白色区域为平坦区域,用 NE 表示。

1.3 深度图像区域划分

基于运动对象提取和边缘检测的结果,利用式(11)对深度视频进行区域划分:

$$\begin{cases} A = SR \cap NE \\ B = MR \cap NE \\ C = SR \cap ER \\ D = MR \cap ER \end{cases} \quad (11)$$

其中: \cap 表示求重叠区域的操作,掩模 A 表示静止平坦区域, B 表示运动平坦区域, C 表示静止边界区域, D 表示运动边缘区域。图1(d)为Book Arrival的最终掩模图。其中,灰色、白色、灰白色和黑色的区域分别为 A 、 B 、 C 和 D 。

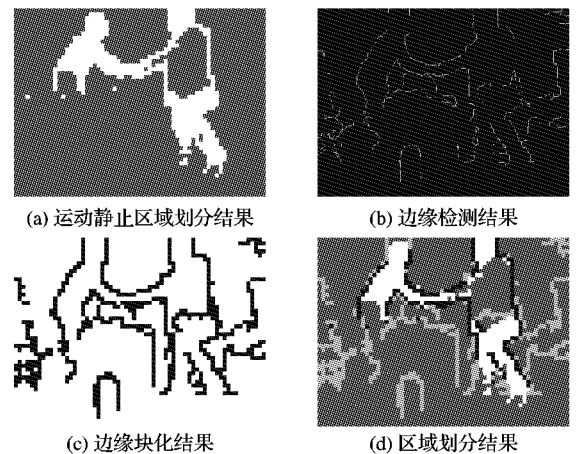


图1 Book Arrival序列区域划分结果

2 基于区域划分的深度视频快速编码算法

2.1 不同区域的最优宏块模式和参考帧选择特性分析

深度视频中不同区域的最优宏块模式分布比例存在区别。表1为Book Arrival序列不同区域的最优宏块模式分布比例。

从表1中可以看出,深度视频编码采用较多的SKIP模式作为最优编码模式,不同区域的最优编码模式有一定的规律性但比例极不均衡。 A 中,深度值比较稳定,最优编码模式主

要为 SKIP 模式,其余所有模式的比例总和不到 10%。B 中,相邻帧之间对应区域的深度值发生了改变,最优编码模式主要以 SKIP、Inter16×16 和 Intra 为主。C 是静止的边缘,时域相关性比较强,最优模式以 SKIP、Inter16×16 和 Intra 为主。D 是细节区域,和周围的非边缘区域深度值差别较大,各种模式作为最优模式的比例分布比较均匀,且存在一定比例的子宏块模式。

表 1 不同区域的宏块模式分布比例

区域	最优宏块模式比例
A	SKIP: 91.23%, 其余所有模式: 8.77%
B	SKIP: 36.28%, Inter16×16: 45.27%, Intra: 9.31%, 其余所有模式: 9.13%
C	SKIP: 74.25%, Inter16×16: 12.36%, Intra: 7.53%, 其余所有模式: 5.86%
D	SKIP: 58.88%, Inter16×16: 16.48%, Inter16×8: 6.78%, Inter8×16: 5.06%, Inter8×8: 3.56%, Intra: 9.23%

图 2 为 Lovebird1 序列和 Book Arrival 序列的最优模式分布图。白色和黑色边框的宏块分别采用 Inter 模式和 Intra 模式作为最优编码模式,其他所有宏块采用 SKIP 模式作为最优编码模式。从图 2 中可以看出, A 中最优模式为 Inter 模式和 Intra 模式的宏块主要分布在 B 或者 D 的周围,因为这些区域周围的深度值发生了跳变。

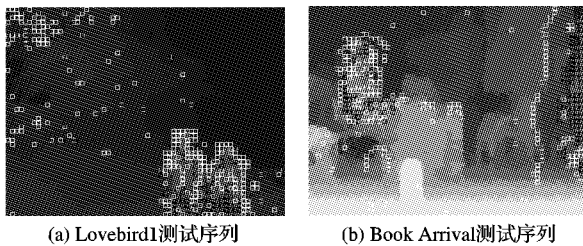


图 2 模式分布图

图 3 为 Book Arrival 序列和 Alt Moabit 序列中宏块最优搜索方向为双向搜索的分布情况,其中白色边框的宏块表示采用双向搜索作为最优搜索方向的宏块。由图 3 可以看出,选择双向搜索作为最优搜索方向的宏块所占比例小并且主要分布在 D 中,因为深度视频大部分区域都是平坦区域,只有少数深度跳变的边缘区域,而平坦区域的最优参考帧通常是前向或后向参考帧列表的第一参考帧。

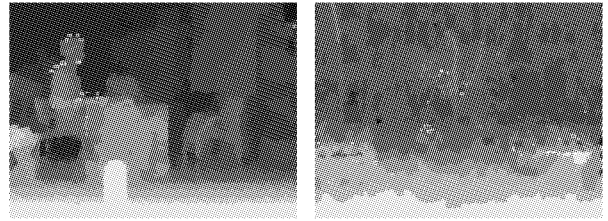


图 3 双向搜索分布图

2.2 深度视频快速编码算法

根据深度视频最优宏块模式和参考帧选择特性分析,本文提出了一种基于区域划分的深度视频快速编码算法,具体流程如图 4 所示。读入掩模后,如果当前宏块属于 A,判断与当前宏块相邻的已编码的宏块是否属于 B 或 D,如果是,则遍历 SKIP、Inter16×16 和 Intra 模式,搜索前向和后向参考帧列表中的第一参考帧;否则,则遍历 SKIP 和 Intra 模式。如果当前宏块属于 B 或 C,则遍历 SKIP、Inter16×16 和 Intra 模式并且不进行双向搜索。如果当前宏块属于 D,则遍历所有模式并进行双向搜索。

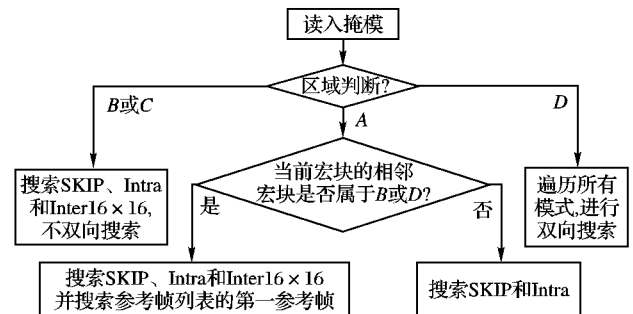


图 4 深度视频快速编码算法流程

3 实验结果

为了验证提出算法的性能,本实验在 JMVC8.3.1 校验平台上进行测试,选取的测试序列为 Book Arrival、Door Flowers、Alt Moabit、Lovebird1 和 Leave Laptop,利用深度估计软件 DERS 5.1 估计出这些序列的深度视频。本文算法从编码码率、编码时间(包括求取掩模的时间)和绘制虚拟视点的质量三个方面与 Yoon 的算法^[9]相比较,两种算法采用相同的测试条件。具体的测试序列和实验参数设置如表 2 所示。其中,QP 是编码过程中的量化参数。

表 2 测试序列和实验参数

序列	左右视点	虚拟视点	基本 QP	GOP 长度	编码帧数
Door Flowers, Leave Laptop	7, 10	8			
Alt Moabit	8, 10	9	22, 27,	15	61
Lovebird1	4, 6	5	32, 37		
Book Arrival	5, 7	6			

图 5 是 Yoon 的算法和本文算法编码时间节约百分比的比较结果。其中, ΔTS 表示编码时间节约的百分比,计算方法为:

$$\Delta TS_{\text{Proposed}} = \frac{Time_{\text{Original}} - Time_{\text{Proposed}}}{Time_{\text{Original}}} \times 100\% \quad (12)$$

$$\Delta TS_{\text{Yoon}} = \frac{Time_{\text{Original}} - Time_{\text{Yoon}}}{Time_{\text{Original}}} \times 100\% \quad (13)$$

式中, $Time_{\text{Original}}$ 、 $Time_{\text{Proposed}}$ 和 $Time_{\text{Yoon}}$ 分别表示原始 JMVC 算法、本文算法和 Yoon 算法编码深度视频的时间。实验结果表明本文算法节约了 85.73%~91.06% 的编码时间,而 Yoon 的算法节约了 44.86%~90.82% 的编码时间。因此,本文算法

的性能更好。对于 Lovebird1 序列,本文算法和 Yoon 的算法节省的时间都是最多的,但是本文算法的优势并不明显,主要因为 Lovebird1 序列的边缘部分比较少,有大量的平坦区域并且时域相关性很好。对于其他序列,因为这些序列的边缘区域比较多,本文算法把边界分为静止边缘和运动边缘分开来处理,所以相比较 Yoon 的算法,本文算法的优势更明显。

表 3 是 Yoon 的算法和本文算法码率变化比较的结果。其中, ΔBR 表示码率变化,计算方法为:

$$\Delta BR_{\text{Yoon}} = \frac{Bitrate_{\text{Yoon}} - Bitrate_{\text{Original}}}{Bitrate_{\text{Original}}} \times 100\% \quad (14)$$

$$\Delta BR_{\text{Proposed}} = \frac{\text{Bitrate}_{\text{Proposed}} - \text{Bitrate}_{\text{Original}}}{\text{Bitrate}_{\text{Original}}} \times 100\% \quad (15)$$

式中: $\text{Bitrate}_{\text{Original}}$ 、 $\text{Bitrate}_{\text{Yoon}}$ 和 $\text{Bitrate}_{\text{Proposed}}$ 分别表示原始 JMVC、提出算法和 Yoon 算法编码深度视频的码率。从表 3

可以看出, Yoon 的算法码率变化从 $-2.82\% \sim 2.76\%$, 平均增加了 0.37% ; 本文算法码率变化从 $-2.03\% \sim 3.30\%$, 平均增加了 0.29% 。两种算法在码率上的变化都可以忽略不计, 但是本文算法在码率上的增加稍微少些。

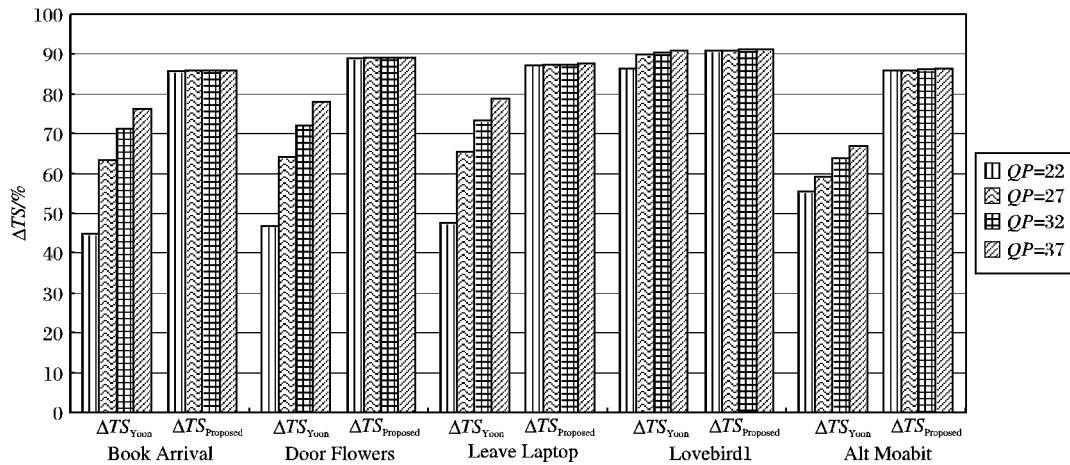


图5 Yoon 的算法和本文算法编码加速性能对比

表3 Yoon 的算法和本文算法码率变化对比

序列	$\Delta BR_{\text{Yoon}} / \%$				$\Delta BR_{\text{Proposed}} / \%$			
	QP = 22	QP = 27	QP = 32	QP = 37	QP = 22	QP = 27	QP = 32	QP = 37
Book Arrival	1.55	0.29	1.13	2.76	2.81	1.10	0.14	1.10
Door Flowers	1.10	-0.29	-0.96	-0.56	2.27	0.15	-1.41	-1.84
Leave Laptop	1.12	-0.71	-0.39	0.94	2.64	0.00	-0.97	-0.80
Alt Moabit	0.38	1.77	2.10	1.91	3.30	2.00	0.80	-0.40
Lovebird1	-2.82	-1.69	-0.33	0.17	0.35	-2.03	-2.31	-1.08
平均		0.37				0.29		

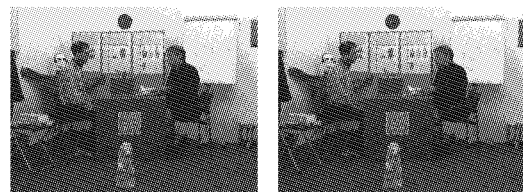
由于深度视频不是直接用于显示, 所以考虑深度视频的峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 没有多大的意义, 本文用合成的虚拟视点质量来评价算法的性能。表 4 是 Yoon 的算法和本文算法通过 MPEG 组织提供的 VSRS 3.5 软件绘制的虚拟视点的 PSNR 比较结果。其中, $\Delta PSNR$ 表示虚拟视点质量变化, 计算方法为:

$$\Delta PSNR_{\text{Yoon}} = PSNR_{\text{Yoon}} - PSNR_{\text{JMVC}} \quad (16)$$

$$\Delta PSNR_{\text{Proposed}} = PSNR_{\text{Proposed}} - PSNR_{\text{JMVC}} \quad (17)$$

式中: $PSNR_{\text{Yoon}}$ 、 $PSNR_{\text{JMVC}}$ 和 $PSNR_{\text{Proposed}}$ 分别为 Yoon 的算法处理后的深度视频、原始深度视频和本文提出的算法处理后的深度视频编码重建后合成的虚拟视点的 PSNR。与原始深度视频相比, 两种算法的重建的深度视频合成的虚拟视点的

PSNR 变化范围为 $-0.03 \text{ dB} \sim 0.02 \text{ dB}$, 基本保持不变。所以, 两种算法在节约编码时间的基础上, 都很好地保证了绘制的虚拟视点的质量。图 6 是原始算法和本文算法的重建的深度序列合成的虚拟视点图像, 两幅图像在主观上没有明显的不同。



(a) 原始算法 (b) 本文算法
图6 Leave Laptop 序列虚拟视点合成结果

表4 Yoon 的算法和本文算法绘制的虚拟视点的 PSNR 比较

序列	$\Delta PSNR_{\text{Yoon}} / \text{dB}$				$\Delta PSNR_{\text{Proposed}} / \text{dB}$			
	QP = 22	QP = 27	QP = 32	QP = 37	QP = 22	QP = 27	QP = 32	QP = 37
Book Arrival	0.00	0.00	-0.01	-0.02	0.00	0.00	-0.01	0.00
Door Flowers	0.01	0.01	-0.01	-0.01	0.01	0.00	-0.01	-0.01
Leave Laptop	0.00	0.01	-0.01	-0.01	0.01	-0.01	0.01	-0.03
Alt Moabit	0.00	0.01	0.01	0.00	0.02	0.02	0.01	-0.01
Lovebird1	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00

此外, 结构相似度 (Structural Similarity, SSIM) 是与主观感知非常接近的一种客观评价方法^[13], 更符合人类视觉系统, 可以更客观地评价本文算法的性能。表 5 是 Yoon 的算法处理后的深度视频、本文算法处理后的深度视频和原始深度视频编码重建后合成的虚拟视点图像 SSIM 值的比较结果。 $SSIM_{\text{Yoon}}$ 表示 Yoon 的算法对应的 SSIM 值, $SSIM_{\text{Proposed}}$ 表示本文算法的 SSIM 值, $SSIM_{\text{JMVC}}$ 表示原始深度视频的 SSIM 值。

从表中可以看出, Yoon 的算法和本文算法处理的深度视频与原始深度视频相比, 虚拟视点的质量基本一致。

4 结语

JMVC 采用可变块模式和多参考帧搜索技术提高了编码的压缩效率, 但却增加了巨大的编码复杂度。为了能够减少深度视频编码花费的时间, 本文首先把深度图像分为不同的

区域;其次,统计分析不同区域的最优模式分布比例和参考帧选择特性;最后,对不同的区域采取不同的编码方案。实验结果表明,在编码码率和虚拟视点的绘制质量基本不变的情况下,编码时间节约了85.73%~91.06%。

表5 三种算法的虚拟视点图像 SSIM 值的比较

算法	序列	QP			
		22	27	32	37
Yoon 算法	Book Arrival	0.959 0	0.959 0	0.959 0	0.958 8
	Door Flowers	0.958 4	0.958 4	0.958 5	0.957 9
	Leave Laptop	0.958 8	0.958 9	0.958 7	0.958 5
	Alt Moabit	0.955 5	0.956 1	0.956 8	0.957 4
	Lovebird1	0.953 9	0.953 8	0.953 8	0.953 5
本文算法	Book Arrival	0.959 0	0.959 0	0.959 0	0.958 9
	Door Flowers	0.958 4	0.958 5	0.958 5	0.958 0
	Leave Laptop	0.958 8	0.958 9	0.958 8	0.958 5
	Alt Moabit	0.955 5	0.956 1	0.956 9	0.957 4
	Lovebird1	0.953 9	0.953 9	0.953 7	0.953 5
原始深度视频	Book Arrival	0.959 0	0.959 0	0.959 1	0.958 9
	Door Flowers	0.958 3	0.958 5	0.958 5	0.958 0
	Leave Laptop	0.958 8	0.958 9	0.958 8	0.958 6
	Alt Moabit	0.955 5	0.956 0	0.956 8	0.957 3
	Lovebird1	0.953 9	0.953 9	0.953 8	0.953 5

参考文献:

- [1] VETRO A, TOURAPIS A M, MULLER K, *et al.* 3D-TV content storage and transmission[J]. IEEE Transactions on Broadcasting, 2011, 57(2): 384-394.
- [2] TANIMOTO M, TEHRANI M P, FUJII T, *et al.* FTV for 3-D spatial communication[J]. Proceedings of the IEEE, 2012, 100(4): 905-917.
- [3] MORI Y, FUKUSHIMA N, YENDO T, *et al.* View generation with 3D warping using depth information for FTV[J]. Signal Processing: Image Communication, 2009, 24(1/2): 65-72.
- [4] PARK Y K, JUNG K, OH Y, *et al.* Depth-image-based rendering for 3DTV service over T-DMB[J]. Signal Processing: Image Communication, 2009, 24(1/2): 122-136.
- [5] 邓智琰, 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): 2453-2456.
- [6] ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6. Draft reference software for MVC[S]. London: JVT-AE207, 2009.
- [7] 郁梅, 徐秋敏, 蒋刚毅, 等. 应用于多视点视频压缩的多参考B帧快速编码算法[J]. 电子与信息学报, 2008, 30(6): 1400-1404.
- [8] 杨春玲, 肖冬琴. 基于SSE和SSIM的H.264帧内预测模式选择改进算法[J]. 电子与信息学报, 2011, 33(2): 289-294.
- [9] YOON D H, HO Y S. Fast depth video coding method using adaptive edge classification[C]// Proceedings of 2011 IEEE Visual Communications and Image Processing. Piscataway, NJ: IEEE Press, 2011: 1-4.
- [10] LIN Y H, WU J L. A depth information based fast mode decision algorithm for color plus depth-map 3D videos[J], IEEE Transactions on Broadcasting, 2011, 57(2): 542-550.
- [11] WANG M H, JIN X, GOTO S. Difference detection based early mode termination for depth map coding in MVC[C]// PCS'10: Proceedings of Picture Coding Symposium. Piscataway, NJ: IEEE Press, 2010: 502-505.
- [12] WANG M H, LIU C, ZHANG T R, *et al.* Region of interest oriented fast mode decision for depth map coding in DIBR[C]// CSPA'11: Proceedings of IEEE 7th International Colloquium on Signal Processing and Its Applications. Piscataway, NJ: IEEE Press, 2011: 177-180.
- [13] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: From error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [14] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227.
- [15] 张春梅, 尹忠科, 肖明霞. 基于冗余字典的信号超完备表示与稀疏分解[J]. 科学通报, 2006, 51(6): 628-633.
- [16] AGGARWAL J K, CAI Q. Human motion analysis: a review[J]. Computer Vision and Image Understanding, 1999, 73(3): 428-440.
- [17] DUDA R O, HART P E, STORK D G. Pattern classification and scene analysis[M]. 2nd ed. New York: Wiley-Interscience, 1995.
- [18] JOHNSON W B, LINDENSTRAUSS J, SCHECHTMAN G. Extensions of Lipschitz maps into Banach spaces[J]. Israel Journal of Mathematics, 1986, 54(2): 129-138.
- [19] WANG Y, HUANG K, TAN T. Human activity recognition based on R transform[C]// IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2007: 1-8.
- [20] BORGEFORS G. Distance transformations in digital images[J]. Computer Vision, Graphics, and Image Processing, 1986, 34(3): 344-371.
- [21] 许秋燕. 二维泊松方程和扩散方程的一类显式并行算法[D]. 济南: 山东大学, 2010.
- [22] WEIZMANN[EB/OL]. [2012-08-20]. <http://www.wisdom.weizmann.ac.il/~vision/SpaceTimeActions.html>.

(上接第1705页)

- [2] 王亮, 胡卫明, 谭铁牛. 人运动的视觉分析综述[J]. 计算机学报, 2002, 25(3): 225-237.
- [3] 徐光祐, 曹媛媛. 动作识别与行为理解综述[J]. 中国图象图形学报, 2009, 14(2): 189-195.
- [4] 凌志刚, 赵春晖, 梁彦, 等. 基于视觉的人行为理解综述[J]. 计算机应用研究, 2008, 25(9): 2570-2578.
- [5] GORELICK L, BLANK M, SHECHTMAN E, *et al.* Actions as space-time shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(12): 2247-2253.
- [6] GAVRILA D M, DAVIS L S. 3-D model-based tracking of humans in action: a multi-view approach[C]// 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 1996: 73-80.
- [7] DAVIS J, BOBICK A. The representation and recognition of action using temporal templates[C]// 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 1997: 928-934.
- [8] BREGLER C. Learning and recognizing human dynamics in video sequences[C]// 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 1997: 568-574.
- [9] VINJE W E, GALLANT J L. Sparse coding and decorrelation in primary visual cortex during natural vision[J]. Science, 2000, 287(5456): 1273-1276.
- [10] WRIGHT J, YANG A Y, GANESH A, *et al.* Robust face