

Journal of Image  
and Graphics

# 中国图象图形学报



ISSN1006-8961  
CN11-3758/TB

2012 **6**  
Vol.17 No.

中国科学院遥感应用研究所  
中国图象图形学学会主办  
北京应用物理与计算数学研究所

# 中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年6月 第17卷 第6期(总第194期)

## 目次

### 综述

人脸图像的年龄估计技术研究 ..... 王先梅, 梁玲燕, 王志良, 胡四泉(603)

图像统计模型参数估计中的期望最大值算法 ..... 李旭超(619)

### 图像处理和编码

鲁棒的梯度驱动图像修复算法 ..... 叶学义, 王靖, 赵知劲, 陈华华(630)

快速去块效应的线性规划方法 ..... 金剑秋, 刘春晓, 王勋, 章志勇(636)

DWT-SVD 域全盲自嵌入鲁棒量化水印算法 ..... 叶天语(644)

Mean-shift 滤波和直方图增强的图像弱边缘提取 ..... 冀峰, 高新波, 谢松云(651)

### 图像分析和识别

梯度点对约束的结构化车道检测 ..... 王永忠, 王晓云, 文成林(657)

多水平外区抑制的轮廓检测 ..... 闫超, 张建州(664)

基于局部投影与块 LBP 特征的图像检索 ..... 邹彬, 潘志斌, 胡森(671)

### 图像理解和计算机视觉

视觉陌生度驱动的增量自主式视觉学习算法 ..... 瞿心昱, 姚明海, 顾勤龙(678)

霍夫空间中多足球机器人协作目标定位算法 ..... 许家铭, 解仑, 王志良, 倪善超(687)

### 计算机图形学

虚拟 CCD 线阵星载光学传感器内视场拼接 ..... 张过, 刘斌, 江万寿(696)

“第八届数字电视与无线多媒体通信国际论坛”会议专栏

追踪盗版者的抗共谋数字指纹 ..... 李晓强, 张煌, 赵洋洋, 汪晶晶(702)

LBP 与鉴别模式结合的热红外人脸识别 ..... 谢志华, 伍世虔, 方志军(707)

电影中复杂事件的检索与识别 ..... 杜吉祥, 郭一兰, 翟传敏(712)

方向特征和网格特征融合的离线签名鉴别 ..... 杨丹凤, 吕岳(717)

视觉关注度的立体图像质量评价 ..... 张艳, 安平, 张秋闻, 张兆杨(722)

低开销的异构数据交换 ..... 赵凯, 赵正德(726)

灵活的多视点视频编码预测结构 ..... 张艳, 蔡灿辉(730)

跨层反馈在无线视频自适应编码中的应用 ..... 王垚中, 郑世宝, 张重阳, 刘勃(736)

中国图象图形学报

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 6 期

2012 年 6 月 16 日出版

主管单位 中国科学院  
主 办 中国科学院遥感应用研究所  
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 李小文  
编辑出版 《中国图象图形学报》编辑出版委员会  
北京 9718 信箱 邮编 100101  
电子信箱:jig@irsa. ac. cn  
电话:010-64807995 010-82614429  
网 址:www. cijg. cn  
印刷装订 北京北林印刷厂  
广告经营许可证 京朝工商广字第 0346 号  
总 发 行 北京报刊发行局  
订 购 全国各地邮局  
国外发行 中国国际图书贸易总公司  
(中国国际书店)  
(北京 399 信箱 邮编 100044)

**Superintended by** Chinese Academy of Sciences  
**Sponsored by** Institute of Remote Sensing Application,  
CAS China Society of Image and Graphics  
Institute of Applied Physics and Computational  
Mathematics  
**Chief editor** LI Xiaowen  
**Editor, Publisher** Editorial and Publishing Board  
of Journal of Image and Graphics  
(P. O. Box 9718, Beijing 100101, China)  
E-mail:jig@irsa. ac. cn  
**Distributed by** Beijing Bureau for Distribution of Newspapers  
and Journals  
**Domestic** All Local Post Offices in China  
**Foreign** China International Book Trading Corporation  
(P. O. Box 399, Beijing 100044, China)  
**Printed by** Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

# Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 6 June 2012

## Contents

### Review

- Age estimation by facial image: a survey ..... Wang Xianmei, Liang Lingyan, Wang Zhiliang, Hu Siqian(603)  
Expectation maximization method for parameter estimation of image statistical model ..... Li Xuchao(619)

### Image Processing and Coding

- Robust gradient driving image inpainting method ..... Ye Xueyi, Wang Jing, Zhao Zhijing, Chen Huahua(630)  
Fast image de-blocking by linear programming ..... Jin Jianqiu, Liu Chunxiao, Wang Xun, Zhang Zhiyong(636)  
Perfectly blind self-embedding robust quantization-based watermarking scheme in DWT-SVD domain ..... Ye Tianyu(644)  
Weak edge detection using Mean-shift filtering and histogram enhancement ..... Ji Feng, Gao Xinbo, Xie Songyun(651)

### Image Analysis and Recognition

- Gradient-pair constraint for structure lane detection ..... Wang Yongzhong, Wang Xiaoyun, Wen Chenglin(657)  
Contour detection based on multilevel inhibition ..... Yan Chao, Zhang Jianzhou(664)  
Image retrieval method based on local projection and block LBP feature ..... Zou Bin, Pan Zhibin, Hu Sen(671)

### Image Understanding and Computer Vision

- Visual novelty driven incremental and autonomous visual learning algorithm ..... Qu Xinyu, Yao Minghai, Gu Qinlong(678)  
Target localization algorithm for cooperative multi-soccer robots based on Hough space .....  
..... Xu Jiaming, Xie Lun, Wang Zhiliang, Ni Shanchao(687)

### Computer Graphics

- Inner FOV stitching algorithm of spaceborne optical sensor based on the virtual CCD line .....  
..... Zhang Guo, Liu Bin, Jiang Wangshou(696)

### Special Issue of the IFIC' 2011

- Anti-collusion fingerprinting scheme capable of tracing pirate ..... Li Xiaoqiang, Zhang Huang, Zhao Yangyang, Wang Jingjing(702)  
Infrared face recognition using LBP and discrimination patterns ..... Xie Zhihua, Wu Shiqian, Fang Zhijun(707)  
Recognize and retrieval complex events in real movies ..... Du Jixiang, Guo Yilan, Zhai Chuanmin(712)  
Off-line signature verification based on combination of direction feature and grid feature ..... Yang Danfeng, Lv Yue(717)  
Stereo image quality assessment based on visual attention ..... Zhang Yan, An Ping, Zhang Qiuwen, Zhang Zhaoyang(722)  
Low overhead of heterogeneous data exchange ..... Zhao Kai, Zhao Zhengde(726)  
Flexible prediction structure for multi-view video coding ..... Zhang Yan, Cai Canhui(730)  
Cross-layer feedback based adaptive coding for wireless video transmission .....  
..... Wang Yaozhong, Zheng Shibao, Zhang Chongyang, Liu Bo(736)

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2012)06-0730-06

论文引用格式: 张艳, 蔡灿辉. 灵活的多视点视频编码预测结构[J]. 中国图象图形学报, 2012, 17(6): 730-735.

## 灵活的多视点视频编码预测结构

张艳, 蔡灿辉

华侨大学信息科学与工程学院, 厦门 361021

**摘要:** 多视点视频的时空相关性不仅随着序列的变化而改变, 而且视频序列内的相关性特征随着时间的推移也会发生变化, 固定的预测结构难以适应复杂情况下多视点视频编码的各种要求。在分析分层 B 帧预测结构和多视点视频图像中当前帧及其参考帧之间时间相关性和视点间相关性的基础上, 提出一种灵活的预测编码结构, 根据该多视点视频的时间及视点间相关关系和当前帧与其时间参考帧之间的距离来确定该帧是否采用耗时的视点间预测。实验结果表明, 与分层 B 帧的改进算法——可扩展性预测结构多视点编码算法相比, 本文算法在保证编码效率基本不变的前提下, 编码复杂度降低了约 45%, 随机访问性能提高 28%, 解码图像缓冲区要求降低 46% 左右。

**关键词:** 多视点视频编码; 预测结构; 时空相关性; 随机访问性能; 解码图像缓冲区

## Flexible prediction structure for multi-view video coding

Zhang Yan, Cai Canhui

School of Information Science and Technology, Huaqiao University, Fujian 361021, China

**Abstract:** The temporal and inter-view correlation of multi-view videos (MVV) changes not only from sequence to sequence but also from frame to frame. Therefore, the fixed prediction structure of the Joint Multi-View Video Coding (JMVC) has difficulties coping with various characteristics of different MVV sequences. In this paper, a flexible prediction structure for multi-view video coding (FPS\_MVC) is proposed based on the analysis of the hierarchical B picture prediction structure and the temporal correlation and inter-view correlation between the current frame and its references. Its temporal level (TL) and the temporal and inter-view correlation of the MVV sequence decide whether the inter-view prediction of the current encoded frame is adopted. Experimental results show that the proposed FPS\_MVC outperforms Scalable Prediction Structure (SPS\_MVC) on reducing computation complexity, improving random access (RA) ability, and decreasing the decode picture buffer (DPB) size by about 45%, 28%, and 46% on average, while maintaining almost the same coding efficiency.

**Key words:** multi-view video coding; prediction structure; temporal and inter-view correlation; random access; decode picture buffer

## 0 引言

多视点视频是传统单视点视频的 3 维扩展, 能

更加生动地再现自然场景, 给人身临其境的感觉, 可广泛应用于 3 维电视、自由视点电视和视频监控等领域<sup>[1]</sup>。由于多视点视频是由多个摄像机从不同角度对同一运动场景同步拍摄而得到的, 视点之间

收稿日期: 2011-09-26; 修回日期: 2012-03-06

基金项目: 国家自然科学基金项目(60772164); 福建省自然科学基金项目(2011J01354)

第一作者简介: 张艳(1984—), 女, 福建省华侨大学信号与信息处理专业硕士研究生, 主要研究方向为图像处理与视频通信。Email:

zy2009@hqu.edu.cn

存在着高度的时空相关性。与传统的单一视频相比,多视点视频数据量成倍增加,存储和传输问题限制了其应用,因此要求多视点视频编码(MVC)具有更高的压缩率、较低的编解码复杂度和更好的随机访问性能<sup>[2]</sup>。

现有的多视点视频编码算法基本都是基于块匹配的预测编码,因此预测结构显得尤为重要,它的设计已经在 MVC 研究中得到了广泛的关注,其中最具代表性的有顺序预测结构、GOGOP (picture of group of picture) 方案和分层 B 帧预测结构<sup>[3-5]</sup>。顺序预测结构<sup>[3]</sup>通过多参考帧的顺序预测来减少相邻视点和时间域的信息冗余,有较高的编码效率,但随机访问性能不佳,并且容易导致错误传递等问题。NTT 公司提出的 GOGOP 方案<sup>[4]</sup>使用多个帧内预测,提高了随机访问性能,却导致编码效率明显降低,而且所采用的 Inter GOP 多参考帧策略,增加了计算复杂度。由 HHI 提出的分层 B 帧预测结构<sup>[5]</sup>因其良好的编码性能而被 JVT 接受为 JMVC 标准的预测结构。对于时间冗余信息,分层 B 帧预测结构采用运动补偿预测(MCP)予以有效的消除;对于视点间冗余信息,则采用视差补偿预测(DCP)技术来消除<sup>[6]</sup>。分层 B 帧结构虽然能有效地利用时间和视点间的联合预测来提高编码性能,但是时间和视点间的联合预测提高了计算复杂度、降低了随机访问性能。为使多视点视频应用于实际场合,急需一种高效多视点视频编码方案,在保证高编码效率的前提下,尽可能降低计算复杂度,提高随机访问性能并降低解码图像缓冲区的要求。

本文在分析分层 B 帧预测结构和多视点视频序列相关性特征的基础上,提出一种改进的预测结构编码方案,以降低其编解码复杂度和提高随机访问性能。

## 1 分层 B 帧结构和序列相关性分析

### 1.1 分层 B 帧预测结构分析

分层 B 帧结构采用时间和视点间的联合预测实现良好的编码性能,作为基本的时间预测结构,能够利用时间相关性有效地消除时间冗余,并利用视点间预测进一步提高编码效率,但其高编码效率却是以高复杂度为代价获得的。因此,有必要进一步分析分层 B 帧预测结构,找出在基本不

影响编码效率的前提下,降低编码复杂度的方法。事实上,在分层 B 帧结构中,根据当前帧与其前向时间参考帧的时间间隔,非关键帧可以分成不同的时间层<sup>[7]</sup>(TL)。实验结果表明,视点间预测对不同时间层的图像编码性能有着不同的影响。在特定的时间层中,不同序列的时间相关性和视点间相关性的关系也有很大的差异。图 1 给出了 GOP 长度为 12 的分层 B 帧时间分级预测结构图,由图可知视点内的图像由关键帧(如 TLID = 0)和非关键帧(如 TLID = 1)组成,关键帧采用帧内编码或以已编码的关键帧作为参考进行帧间编码,非关键帧则采用分级预测结构进行编码。每一个时间层的图像数为

$$N_i = \begin{cases} 1 & i = 0 \\ L_{\text{GOP}} - (1 + \sum_{j=1}^{I_{\text{TLmax}}-1} 2^{j-1}) & i = I_{\text{TLmax}} \\ 2^{i-1} & \text{其他} \end{cases}$$

式中, $i$  是时间层索引, $N_i$  是第  $i$  个时间层的帧图像数, $L_{\text{GOP}}$  是 GOP 长度, $I_{\text{TLmax}}$  是当前 GOP 的最大时间层

$$I_{\text{TLmax}} = \begin{cases} \lfloor \log_2 L_{\text{GOP}} + 1 \rfloor & L_{\text{GOP}} = 12 \text{ 或 } 15 \\ \log_2 L_{\text{GOP}} & L_{\text{GOP}} = 2^n \end{cases}$$

式中, $\lfloor x \rfloor$  表示不大于  $x$  的最大整数。GOP 长度等于 12 时,TL 等于 0、1、2、3、4 的图像数分别是 1、1、2、4、4,当前帧与前向时间参考帧的距离分别为 4、3、2、1。处于第 3 层和第 4 层的帧占总帧数的 67%,即高时间层的图像数占了整体图像的绝大部分。

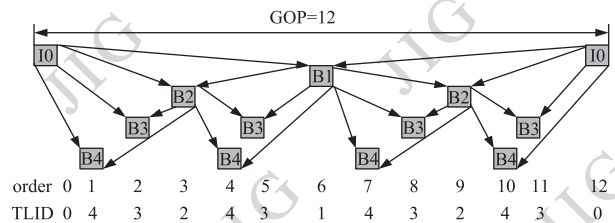


图 1 分层 B 帧的时间预测结构

Fig. 1 Hierarchical B picture structure for temporal prediction

### 1.2 不同时间层图像的时间相关性和视点间相关性关系分析

在分层 B 帧结构中,B 视点的图像帧均采用双向视点间参考,为了分析不同时间层图像的时间相关性和视点间相关性的关系,采用块匹配法<sup>[8]</sup>分析各层的时间相关性( $R_T$ )和视点间相关性

( $R_V$ ), 表 1 给出了一些典型序列在不同时间层的相关性分布情况。其中,  $R_T$  和  $R_V$  分别是经模式选择后采用时间预测和视点间预测占总宏块数的百分比, 显然  $R_T + R_V = 1$ 。从表 1 可以看出, 随着时间层索引的增大, 时间相关性逐渐增强, 这是因为随着 TL 的增大, 当前编码图像与时间参考图像的时间间隔逐渐减小, 而视点间的相关性由于摄像头间距的固定而保持不变。因此, 随着 TL 的增加, 序列的时间相关性增强, 时间预测的准确性提高, 而视点间预测对编码性能的贡献逐渐减小, 当 TL 达到最大时, 除出现场景切换外, 采用视点间预测进行编码的几率非常之小。

表 1 不同时间层的相关性分析

Table 1 Correlation analysis for frames in different TL

MVV 序列		/%			
		TL = 1	TL = 2	TL = 3	TL = 4
Vassar	$R_T$	95.4	96.7	97.6	97.9
	$R_V$	4.6	3.3	2.4	2.1
Exit	$R_T$	80.1	86.0	91.5	91.2
	$R_V$	19.9	14.0	8.5	8.2
Race1	$R_T$	71.9	79.5	90.3	93.9
	$R_V$	27.1	20.5	9.7	6.1
Ballroom	$R_T$	68.9	76.2	81.7	82.1
	$R_V$	31.1	23.8	17.3	17.9
Flamenco2	$R_T$	68.8	81.6	89.5	96.8
	$R_V$	31.2	18.4	10.5	3.2
Rena	$R_T$	65.5	76.3	86.7	92.6
	$R_V$	34.5	23.7	13.3	7.4
D_Xmas	$R_T$	12.6	17.5	20.4	28.3
	$R_V$	87.4	82.5	79.6	71.7

## 2 改进的分层 B 帧预测结构

### 2.1 灵活的多视点编码预测结构的设计

分层 B 帧结构需要对大量处于高时间层的图像进行耗时的视点间预测, 对某些序列而言, 这种预测对编码效率的贡献微乎其微, 而忽略对这些帧的视点间预测能够降低计算复杂度, 提高整体的编码性能。据此提出改进的编码方案, 使编码器能够根据序列各时间层的相关性特征灵活地选择最佳的预

测结构。

根据表 1 各时间层的相关性特征将多视点视频序列分为 4 类并设计 4 种对应的预测结构。第 1 类是时间相关性绝对占优的序列, 如 Vassar 序列, 此种序列由于对象运动缓慢且摄像机的间距较大, 导致序列的时间相关性占绝对优势, 忽略非关键帧的视点间参考虽会引起编码效益的微小损失, 但却能大大降低编码复杂度, 提高随机访问性能, 为此提出如图 2(a) 所示的时间相关性绝对占优预测结构 PS1; 第 2 类序列的时间相关性占主导但视点间相关性也占有相当大的比例, 如 Exit 序列, 此种序列的视点间相关性虽然不能完全忽略, 但是由于处于高时间层的图像帧采用视点间预测的概率也很小, 所以提出如图 2(b) 所示的部分时间相关性占优预测结构 PS2, 对最大和次大时间层的帧图像不进行视点间预测, 在对编码效率影响很小的前提下达到大大降低计算复杂度的目的; 第 3 类序列是时空相关性均衡的序列, 此类序列总的时空相关性基本相当, 但最高时间层图像的视点间预测效果仍然可以忽略, 对应的预测结构如图 2(c) 所示, 称之为时空相关性均衡的预测结构 PS3; 最后一类是视点间相关性占优的序列, 如 D\_Xmas 序列, 此序列由空间下采样获得, 时间相关性非常弱, 需要采用视点间预测才能获得较好的编码效果, 因此必须采用原始的双向 B 帧预测结构 PS4, 如图 2(d) 所示。编码时按照序列所属类型选择最佳预测结构。预测结构的选择策略为

$$PS = \begin{cases} FPS\_MVC\_PS1 & R_T > R_{T1} \\ FPS\_MVC\_PS2 & R_{T2} < R_T < R_{T1} \\ FPS\_MVC\_PS3 & R_{T3} < R_T < R_{T2} \\ FPS\_MVC\_PS4 & R_T < R_{T3} \end{cases}$$

式中,  $R_{Ti}$  ( $i = 1, 2, 3$ ) 是预先设置的阈值, 由实验测得。当  $R_{T1} = 90\%$ ,  $R_{T2} = 80\%$ ,  $R_{T3} = 60\%$  时达到最佳的编码效果。

### 2.2 灵活的多视点编码预测结构实现

灵活的多视点预测结构编码方案能够根据多视点视频序列的类型选择相应的预测结构进行编码, 降低计算复杂度, 但是必须预先获得  $R_T$  或  $R_V$ 。采用传统的块匹配方法进行计算需要 2 次编码, 不是一种可行的方案, 因此实现所提出的预测结构算法的关键在于是否能不计算  $R_T$  或  $R_V$  也能自适应选择最佳的预测结构。

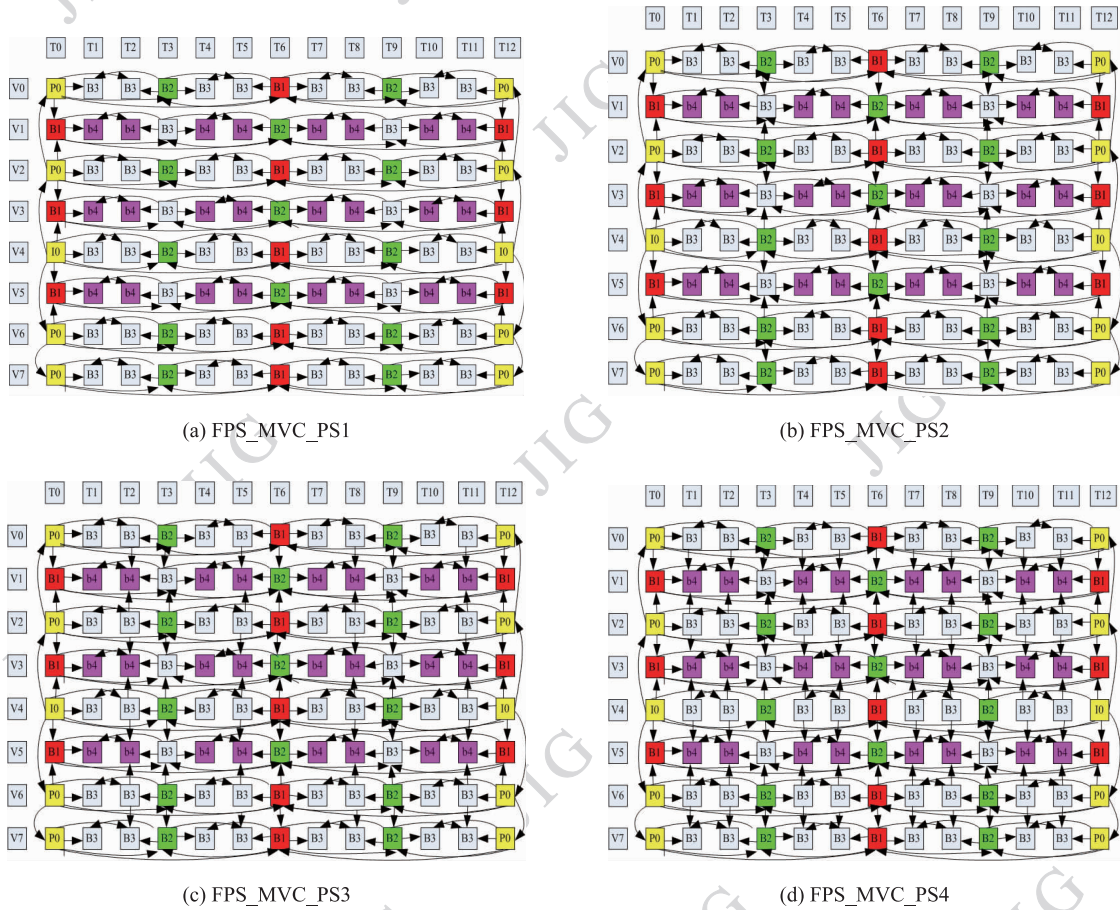


图 2 候选预测结构图

Fig.2 Four candidate prediction structures

进一步考察 4 种预测结构,不难发现其基本视点仅采用时间预测,而处于 0 时间层的关键帧仅采用视点间预测。因此基本视点中 B 帧的代价函数和关键帧中 B 帧的代价函数之比应该能够反映该 2 维 GOP 的时间相关和视点间相关关系。

令  $J_V$  为 2 维 GOP 关键 B 帧的平均拉格朗日代价值,  $J_T$  为基本视点 B 帧的平均拉格朗日代价值,定义相关性参数

$$\eta = \frac{J_T}{J_V}$$

则  $\eta$  也反映了该 2 维 GOP 各帧间相关关系,应该和  $R_T$  有一定的关系。

图 3 为相关性参数  $\eta$  和时间相关性  $R_T$  之间的对应关系,显然,  $\eta$  是  $R_T$  的单调增函数,可用  $\eta$  代替  $R_T$  来选择预测结构。通过 Matlab 数据拟合,可以得到其函数关系式为

$$\eta/\% = 32.72 + 4.0 \times \exp(100 \times R_T \times 0.03)$$

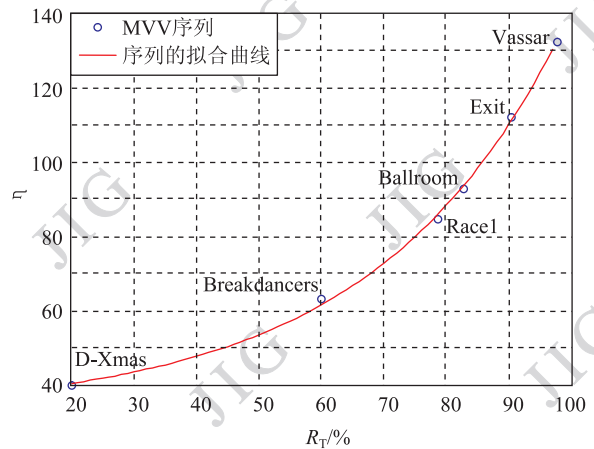


图 3 相关性特征  $\eta$  与时间相关性  $R_T$  的关系

Fig.3 The relationship between  $R_T$  and  $\eta$

由于不论采用图 2 中的哪一种预测结构,对基本视点和关键帧的预测结构都一样,而  $\eta$  是编码该 2 维 GOP 的基本视点和关键帧的副产品,不需要消耗额外的计算复杂度,因此预测结构的选择策略可

用表示为

$$PS = \begin{cases} \text{FPS\_MVC\_PS1} & \eta > TH_1 \\ \text{FPS\_MVC\_PS2} & TH_2 < \eta < TH_1 \\ \text{FPS\_MVC\_PS3} & TH_3 < \eta < TH_2 \\ \text{FPS\_MVC\_PS4} & \eta < TH_3 \end{cases}$$

式中,  $TH_1 = 95\%$ ,  $TH_2 = 75\%$ ,  $TH_3 = 55\%$ 。

### 3 实验结果

为评价所提出的编码方案的性能,以 JMVC 4.0 为测试平台,采用序列 Vassar、Race1、Exit 和 Ballroom 进行测试,表 2 给出了测试条件。表 3 是本文 FPS\_MVC 编码方案和 Huo 等人最近提出的双向 B 帧预测结构的一种改进算法——可扩展性预测结构 (SPS\_MVC) 编码方案<sup>[7]</sup>的率失真性能对比,其中 IMPV 代表本文所提出的方案相对于 SPS\_MVC 方案的改进程度。从表 3 中可以看出,两种方案的编码效率基本相同。多视点视频编码的复杂度通常用编码时间来考量,表 4 给出了两种方案的编码复杂度对比情况,从中可以看出 FPS-MVC 改进算法可以大幅降低计算复杂度。

表 2 多视点视频序列编码参数

Table 2 Coding parameters of MVV sequences

编码帧数	200	GOP	12
搜索范围	64	帧率	25
Based QP	28 30 32 34 36	搜索方式	FastSearch

表 3 编码效率比较

Table 3 Coding efficiency comparison

MVV 序列		FPS_MVC	SPS_MVC	IMPV
Vassar	Y/dB	34.95	34.96	-0.01
	Bitrate/(kbit/s)	177.27	172.96	2.49%
Exit	Y/dB	36.63	36.64	-0.01
	Bitrate/(kbit/s)	184.60	183.32	0.7%
Race1	Y/dB	35.51	35.53	-0.02
	Bitrate/(kbit/s)	311.52	314.69	-1.01%
Ballroom	Y/dB	34.38	34.35	0.03
	Bitrate/(kbit/s)	419.13	414.43	1.13%

表 4 计算复杂度比较

Table 4 Computational complexity comparison

MVV 序列	FPS_MVC/s	SPS_MVC/s	IMPV/%
Vassar	874.68	1 704.80	48.69
Race1	1 535.93	2 636.95	41.75
Exit	1 014.86	2 108.87	51.88
Ballroom	1 448.91	2 351.99	38.40
Aver MPV	1 218.60	2 200.65	45.18

良好的随机访问性能<sup>[9]</sup>有利于视点间的切换和任意访问视频序列的图像帧,随机访问性能由访问任一给定帧所需解码帧数的均值和最大值  $F_{AV}$  和  $F_{MAX}$  衡量。

$$F_{MAX} = \max \{ X_{j,i} | 0 < i < n, 0 < j < m \}$$

$$F_{AV} = \sum_{i=1}^n \sum_{j=1}^m X_{j,i} P_{j,i}$$

式中,  $X_{j,i}$  表示对第  $m$  个视点第  $n$  帧进行访问前需要解码的帧数,  $P_{j,i}$  表示用户访问第  $m$  个视点第  $n$  帧的概率。

解码图像缓冲区是用来衡量多视点视频编码对系统内存的需求指标,由于视点间预测需要额外的空间来存储解码后的图像,这对系统内存提出新的要求。表 5 和表 6 分别是两种编码方案的随机访问性能和对解码图像缓冲区的要求比较。从中可以看出 FPS-MVC 算法的随机访问性能提高约 28%,而解码图像缓冲区降低了 46% 左右,极大地优化了整体编码性能。

表 5 随机访问性能比较

Table 5 RA comparison

MVV 序列	SPS_MVC ( $F_{AV}/F_{MAX}$ )	FPS_MVC ( $F_{AV}/F_{MAX}$ )
Race1	10/18	8/16
Vassar	10/18	6.45/10
Exit	10/18	6.45/10
Ballroom	10/18	7.76/16
Aver	10/18	7.17/13

表6 解码图像缓冲区比较

Table 6 DPB size comparison

MVV 序列	SPS_MVC	FPS_MVC	IMPV/%
Vassar	24	10	58
Race1	24	16	33
Exit	24	10	58
Ballroom	24	16	33

## 4 结 论

在分析分层 B 帧预测结构及不同时间层上当前帧与参考帧的时间相关性和视点间相关性关系的基础上,对多视点视频编码的预测结构进行优化,提出一种灵活的多视点视频编码预测结构。编码器根据当前的 2 维 GOP 各图像帧所处的时间层和序列的时空相关性特征确定该 GOP 所属的类型,并选择最佳的预测结构进行编码,从而使得整体的编码性能达到最优。实验结果表明,相对于 Huo 等人最近提出的双向 B 帧预测结构的改进算法,本文提出的算法在保证编码效率几乎不变的前提下,使编码复杂度降低了 45%,随机访问性能提高了 28%,解码图像缓冲区大小降低了 46% 左右。实验结果表明了该算法的有效性和合理性。

## 参考文献 (References)

- [1] Tanimoto M. Proposal on Requirements for FTV. ISO/IEC MPEG & ITU-T VCEG, JVT-W127 [S]. California, USA: IEEE, 2007.
- [2] Smolic A, Mueller K. Coding algorithms for 3DTV-A survey [J]. IEEE Trans Circuits Syst. Video Technol, 2007, 17(11): 1606-1621.
- [3] Survey of Algorithms Used for MVC. ISO/IEC JTC1/SC29/WG11, N6909 [S]. Hong Kong, China: IEEE, 2005.
- [4] Kitahara M, Kimata H, Shimizu S. Multi-view video coding using view interpolation and reference picture selection [C]//Proceedings of IEEE International Conference. Multimedia & Expo. Toronto, Canada: IEEE, 2006: 97-100.
- [5] Merkle P, Smolic A, et al. Efficient prediction structures for multiview video coding [J]. IEEE Trans. Circuits and Systems for Video Technology, 2007, 17(11): 1461-1473.
- [6] Zhang Y, Kwong S, et al. Efficient multi-reference frames selection algorithm for hierarchical B pictures in multiview video coding [J]. IEEE Transactions on Broadcasting, 2010, 57(1): 15-23.
- [7] Huo J Y, Chang Y L, Li M. Scalable prediction structure for multiview video coding [C]//Proceedings of International Symposium on Circuits and Systems. Taipei, China: IEEE, 2009: 2593-2596.
- [8] Fecher U, Kaup A. Statistical analyses of multi-reference block matching for dynamic light field coding [C]//Proceedings of the 10th International Fall Workshop Vision, Modeling, and Visualization. Erlangen, Germany: Springer. 2005: 445-452.
- [9] Vetro A, Wiegand T, Sullivan G J. Overview of the stereo and multiview video coding extensions of the H. 264/MPEG-4 AVC standard [J]. Proceedings of the IEEE, 2011, 99(4): 626-642.