

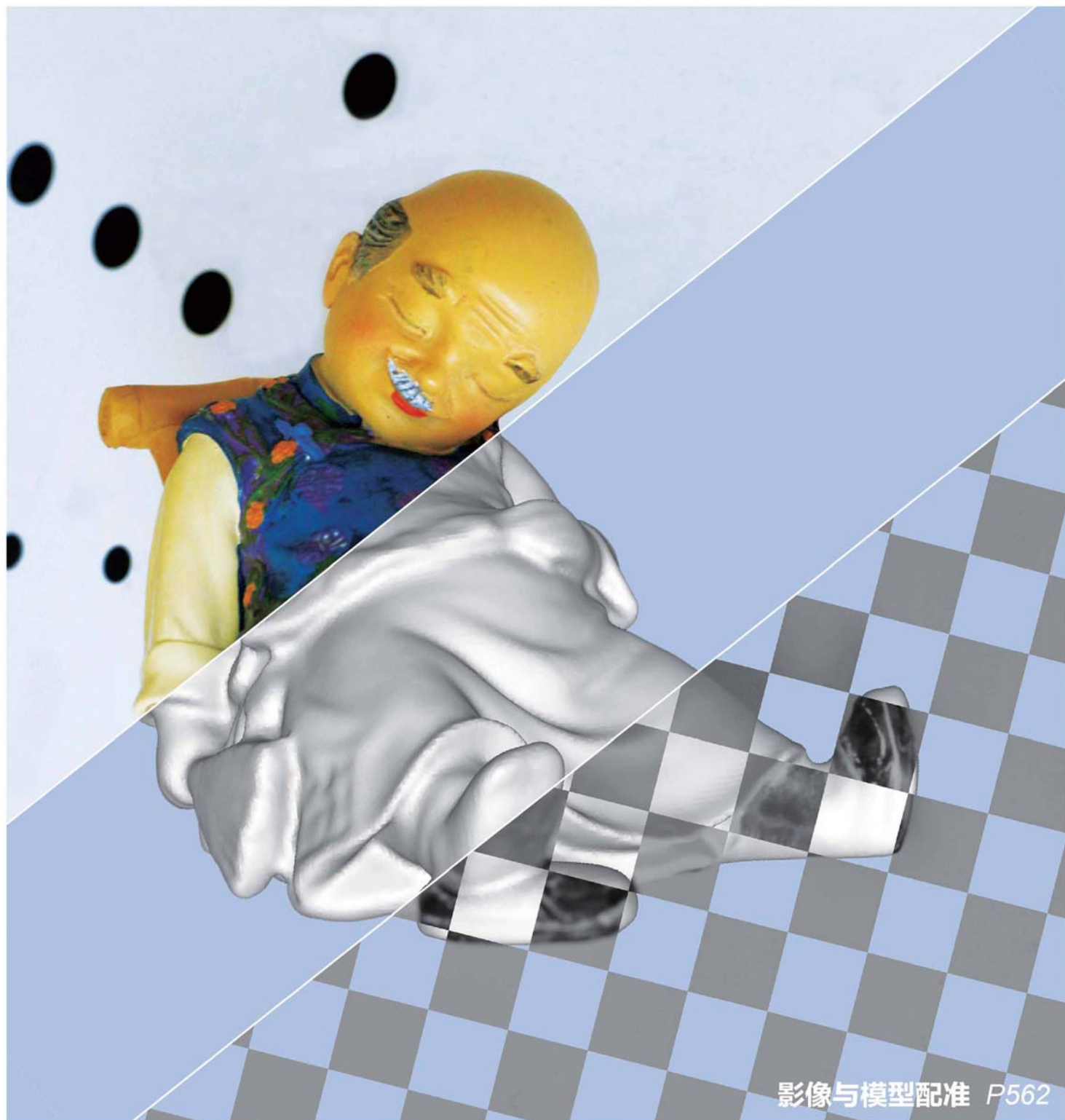


主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2014
04
VOL.19

ISSN1006-8961
CN11-3758/TB



影像与模型配准 P562



第19卷第4期（总第216期）

2014年4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

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 李小文

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@irsa.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 LI Xiaowen

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@irsa.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

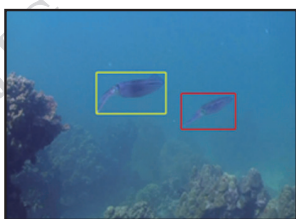
国外发行代号 M1406

国内邮发代号 82-831

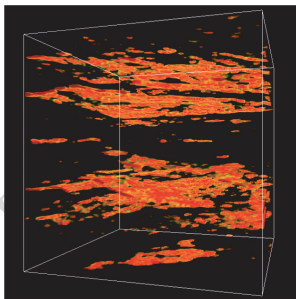
国内定价 50.00元



面泊松融合结合色彩变换的无缝纹理辐射处理(第512页)



融合视觉深度的特征计算与水下目标跟踪(第534页)



针对全空子数据体的GPU体绘制(第577页)

图像处理和编码

- 梯度矢量扩散控制实现边缘保持的彩色图像去噪
贾迪, 孟祥福, 张一飞, 董娜, 李思慧 0493
- 联合矩阵F范数的低秩图像去噪
刘新艳, 马杰, 张小美, 胡钊政 0502
- 面泊松融合结合色彩变换的无缝纹理辐射处理
周漾, 郑顺义, 黄荣永, 马电 0512
- 改进的局部方向模式纹理表示方法
刘海军, 常东超, 张凌宇 0520
- 图像插值空间大容量可逆数字水印算法
王继军 0527

图像分析和识别

- 融合视觉深度的特征计算与水下目标跟踪
王慧斌, 程勇, 陈哲 0534
- 融合颜色属性和空间信息的显著性物体检测
徐丹, 唐振民, 徐威 0541
- 采用人工鱼群的改进广义Hough变换目标定位
李志, 谢强 0549
- 非规则弯曲形变的不变量表示
李国祥, 夏国恩 0556

图像理解和计算机视觉

- 小物体3维重建中影像与几何模型的配准
郑顺义, 王晓南 0562
- 像素聚类改进二进制描述子鲁棒性
惠国保, 李东波 0569

计算机图形学

- 针对全空子数据体的GPU体绘制
李国和, 段忠祥, 吴卫江, 洪云峰, 刘智渊, 程远 0577

医学图像处理

- 基于无损水印的医学图像篡改检测和质量恢复
邓小鸿, 陈志刚, 毛伊敏 0583

遥感图像处理

- 特征点和不变矩结合的遥感图像飞机目标识别
曾接贤, 付俊, 符祥 0592
- 结合稀疏识别的自适应Wallis滤波在高分辨率影像控制点匹配中的应用
耿蕾蕾, 孙权森, 纪则轩, 刘佑鑫, 袁凯 0603

第9届和谐人机环境联合学术会议专栏

- 局部特征点的3维模型变换域水印算法
张建明, 周小梅, 王新宇, 詹永照 0613
- 足球视频搜索引擎中的用户偏好挖掘
胡雨成, 于俊清, 黄贤强, 何云峰, 管涛 0622
- 随机亮度差量化的二进制特征描述
庄东晔, 张冬明, 张勇东, 李锦涛 0630
- L_1 优化在网格去噪中的应用
王鹏, 王胜法, 曹俊杰, 李楠楠, 李波, 苏志勋 0637
- 多视点视频双向实时转码技术的设计与实现
孙立峰, 蔡飞飞, 杨士强 0645

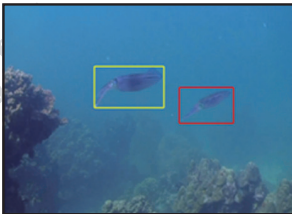
- 第五届数字地球高峰论坛征文通知 封三

CONTENTS

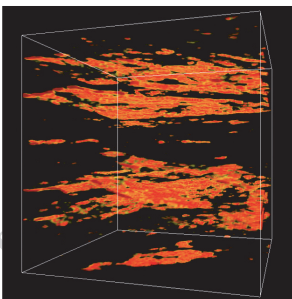
JOURNAL OF IMAGE AND GRAPHICS



Face-wise Poisson blending with color transfer in seamless texturing(P512)



Visual depth based feature calculation and underwater target tracking(P534)



GPU-based volume rendering for full-empty subdata blocks(P577)

Image Processing and Coding

Edge preserving denoising of color image through controlling the gradient vector Jia Di, Meng Xiangfu, Zhang Yifei, Dong Na, Li Sihui	0493
Image denoising of low-rank matrix recovery via joint Frobenius norm Liu Xinyan, Ma Jie, Zhang Xiaomei, Hu Zhaozheng	0502
Face-wise Poisson blending with color transfer in seamless texturing Zhou Yang, Zheng Shunyi, Huang Rongyong, Ma Dian	0512
Improved local directional pattern texture descriptor Liu Haijun, Chang Dongchao, Zhang Lingyu	0520
High capacity reversible watermarking for image interpolation space Wang Jijun	0527

Image Analysis and Recognition

Visual depth based feature calculation and underwater target tracking Wang Huibin, Cheng Yong, Chen Zhe	0534
Combining color names with spatial information for salient object detection Xu Dan, Tang Zhenmin, Xu Wei	0541
Improved generalized Hough transform using artificial fish swarm algorithm in target location Li Zhi, Xie Qiang	0549
The global descriptor for bending non-rigid shapes Li Guoxiang, Xia Guoen	0556

Image Understanding and Computer Vision

Images and geometric model registration in 3D reconstruction of small objects Zheng Shunyi, Wang Xiaohan	0562
Robust descriptor structured by pixel clustering Hui Guobao, Li Dongbo	0569

Computer Graphics

GPU-based volume rendering for full-empty subdata blocks Li Guohe, Duan Zhongxiang, Wu Weijiang, Hong Yunfeng, Liu Zhiyuan, Chen Yuan	0577
--	------

Medical Image Processing

Lossless watermarking algorithm for medical image's tamper detection and recovery with high quality Deng Xiaohong, Chen Zhigang, Mao Yimin	0583
---	------

Remote Sensing Image Processing

Aircraft target recognition in remote sensing images based on distribution of the feature points and invariant moments Zeng Jiexian, Fu Jun, Fu Xiang	0592
Adaptive enhancement via sparse recognition for automatic control point extraction in high-resolution images Geng Leilei, Sun Quansen, Ji Zexuan, Liu Jixin, Yuan Kai	0603

Special Issue on HHME 2013

Transform domain watermarking scheme of 3D models based on local feature points Zhang Jianming, Zhou Xiaomei, Wang Xinyu, Zhan Yongzhao	0613
User preference mining for soccer video search engine Hu Yucheng, Yu Junqing, Huang Xianqiang, He Yunfeng, Guan Tao	0622
Novel binary feature from intensity difference quantization Zhuang Dongye, Zhang Dongming, Zhang Yongdong, Li Jintao	0630
The application of l_1 -optimization in mesh denoising Wang Peng, Wang Shengfa, Cao Junjie, Li Nannan, Li Bo, Su Zhixun	0637
The design and implementation of real time MVC transcoding Sun Lifeng, Cai Feifei, Yang Shiqiang	0645

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2014)04-645-04

论文引用格式: 孙立峰, 蔡飞飞, 杨士强. 多视点视频双向实时转码技术的设计与实现[J]. 中国图象图形学报, 2014, 19(4): 645-648. [DOI: 10.11834/jig.20140420]

多视点视频双向实时转码技术的设计与实现

孙立峰, 蔡飞飞, 杨士强

清华大学计算机科学与技术系, 北京 100084

摘要: 目的 多视点视频编码标准 MVC 是国际标准组织针对 3D 视频制定的编码标准, 支持实时编码和兼容已有 H.264 标准是 MVC 应用推广迫切需要解决的问题。方法 基于拓扑排序, 提出一种 2 维预测结构到 1 维参考序列的映射变换机制, 利用已有的 H.264 编码单元技术, 实现了 MVC-H.264 双向实时转码。结果 该方法能够兼容现有的 H.264 编解码器, 支持高清 MVC 实时编码。通过 3D 视频直播实际应用验证了方法的有效性。结论 本文 MVC 与 H.264 标准间的预测参考结构映射和码流转换机制, 无需设计专用芯片就能同时兼容 H.264 和 MVC 标准。

关键词: H.264; 多视点视频编码; 转码; 并行编码

The design and implementation of real time MVC transcoding

Sun Lifeng, Cai Feifei, Yang Shiqiang

Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

Abstract: **Objective** The MVC standard, as an extension of H.264 standard, provides a standardized codec reference for 3D videos. Real time encoding and keeping compatible with H.264 are key problems for deploying MVC in 3D applications. **Method** In this paper, a transformation mechanism is proposed based on topological sort, which maps the 2D prediction structure of MVC to 1D reference sequence of H.264, and then uses the basic coding technologies of H.264 to get MVC bit stream with syntax reconstruction. **Result** The experiment results show the proposed method could achieve real time performance for MVC-H.264 transcoding and are compatible with existing H.264 codec. **Conclusion** The proposed method gives transformation mechanism of bit streams between MVC and H.264, which need no special designed SOC.

Key words: H.264; multiview video coding; transcoding; parallel encoding

0 引言

多媒体技术的发展日新月异, 3 维场景的体验与交互, 已经成为现实。随着 3D 视频拍摄、处理和显示技术的不断发展, 3D 视频成为新一代交互式媒体的重要应用。

视频编码专家组 (VCEG) 和动态图像专家组 (MPEG) 在现有的 H.264/AVC 标准基础上, 制定了多视点视频编码 (MVC) 的扩展标准。根据在标准测试序列上进行的相关测试报告^[1], 8 路视点的 MVC 编码, 码率平均可节省 20% 左右^[2], 常用 3D 视频的双路视点, 采用 MVC 编码可节省 20% ~ 30%^[3]。

收稿日期: 2013-10-24; 修回日期: 2013-11-28

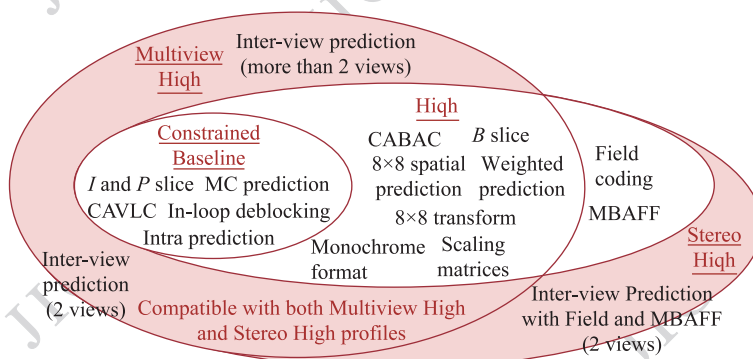
基金项目: 国家重点基础研究发展计划 (973) 基金项目 (2011CB302206); 国家自然科学基金项目 (61272231)

第一作者简介: 孙立峰 (1972—), 男, 副教授, 2000 年于国防科技大学获博士学位, 研究方向包括网络多媒体、视频编码、3D 视频处理与虚拟现实等, 在国际期刊会议上发表论文 60 余篇, 获 IEEE TCSVT 期刊 2010 年度最佳论文奖, ACM Multimedia 2012 最佳论文奖。

E-mail: sunlf@tsinghua.edu.cn

文献[4]从标准定义的档次角度,给出了 MVC 在标准扩展上的明确划分。由图 1 可以看到,标准在档次上明确规定了 MVC 扩展的内容,仅限于视点

间预测的新特性,对于 H. 264 已有的特性,如 I、P、B 帧,运动补偿预测、加权预测等等,都兼容已有的内容。

图 1 MVC 标准档次^[4]Fig. 1 MVC profiles^[4]

MVC 标准应用面临的现实问题是如何在支持 MVC 视频编码标准与已有标准兼容。现有编码设备和终端普遍采用 H. 264 标准,从工业界的角度看,标准的性能和实现效率一般需要数倍甚至是数量级的提升,才能推动软硬件的更新换代。从硬件上讲,编码器设备厂商在已有 H. 264 编码芯片的研发和制造工艺上投入了大量财力,不会轻易投入新的编码芯片的研制;内容提供商也不会单为 3D 视频而更换已有的 H. 264 编码器设备。从软件上讲,官方的 JMVC 软件只能用于测试验证,由于性能的原因,无法用于商业应用;而 x264 等多数自由软件和商业软件也未引入对 MVC 标准扩展的支持。

面对这种现状,编解码器如何利用标准的向下兼容,在已有的 H. 264 设备上改造实现对 MVC 标准的支持至关重要。最大化利用标准的兼容性,最小化对已有 H. 264 硬件设备的改造,是 MVC 编解码器实现的关键因素,本文给出一种 MVC-H. 264 转码的设计和实现。

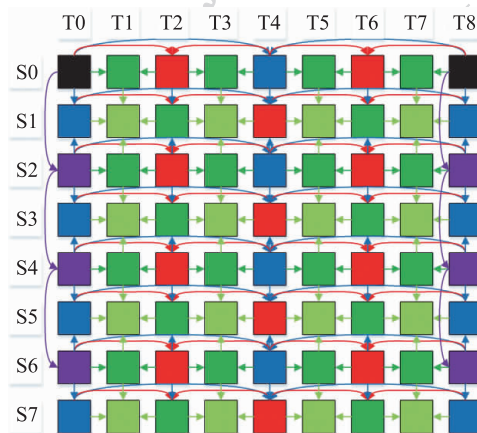
1 转码设计与系统实现

1.1 MVC-H. 264 转码

本文设计和实现的 MVC-H. 264 转码,目标是将符合 MVC 标准的多视点视频码流在不进行实际底层解码和编码的前提下转换为一路标准的 H. 264 视频码流。依照 MVC 标准文档,逐步说明 MVC-H. 264 转码的过程。

图 2 为一种典型的 MVC 编码预测结构,其中编

号 T0、T1、T2……表示按序的时间点,S0、S1、S2……表示不同的视点,多视点的视频序列按照时间序列和空间位置依次排列,形成 2 维的预测结构。

图 2 典型的 MVC 预测结构^[5]Fig. 2 Typical MVC prediction structure^[5]

MVC 的预测结构以帧为基本单元,而在帧以下的编码单元,如片或宏块,符合 H. 264 的语法标准定义。在帧的层级对 2 维的预测结构进行映射变换,将其变换为 1 维的预测结构,同时保持内部层级的编码结构。

首先考虑两种直接的映射变换:视点优先顺序与时间点优先顺序。在不考虑 B 帧的简单预测结构中,视点优先顺序依次编码 S0 视点内一个图片组的帧 T0、T1、…、T8, S1 视点的 T0、T1、…、T8, 至 S7 视点的 T0、T1、…、T8;时间点优先顺序依次编码 T0 时间点全部视点的帧 S0、S1、…、S7, T1 时间点的

$S_0, S_1, \dots, S_7$ 。参考文献[5]的分析,对于图2所示的预测结构,采用这两种映射变换,DPB(解码图片缓存)的最大帧数分别为

$$s \lceil n/2 \rceil + \lceil \lg(s) \rceil + 1 \quad (1)$$

$$n(\lceil \lg(s) \rceil + 1) + 2 \quad (2)$$

式中, s 为图片组大小, n 为视点数。

MVC向H.264码流转换的设计如下:

1) 预测结构双向映射。将多视点视频序列抽象成图的节点,将帧之间的预测参考关系抽象成节点之间的有向边,构造出一个在2维节点阵上的有向无环图,然后对有向无环图中的节点进行拓扑排序,利用原图的有向边计算新的1维节点序列之间的连边,将2维节点阵映射成为一个1维的节点序列。

2) 参考关系控制的码流生成。通过1维节点序列和2维节点序列之间的对应关系,计算出每个H.264视频帧码流对应的多路视频图片的位置,根据参考关系重新计算多视点视频码流中的参考序列控制指令(参考图像管理,又称为参考帧重排序过程)和内存控制指令(解码参考图像标记,又称为参考帧标记),将图片序列放入H.264的编码器中进行编码,同时通过对编码器的控制,使编码时每个图片只参考在1维节点序列中可以参考的图片,编码后得到一个H.264的码流。将H.264码流进行译码和重新组装并加上多视点视频的参考序列控制指令和内存控制指令得到符合MVC标准的码流。

以图2中的预测结构为例,使用拓扑排序的映射关系,可以得到如图3(a)所示的预测结构,按照图3(b)的顺序对视频帧依次排列。

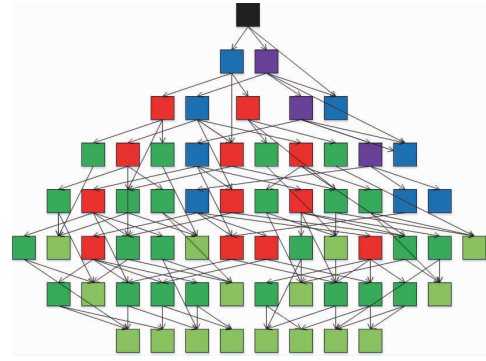
DPB的最大帧数为

$$(\lceil nv/2 \rceil + 1)(\lceil \lg(gs) \rceil + 1) + 2 \quad (3)$$

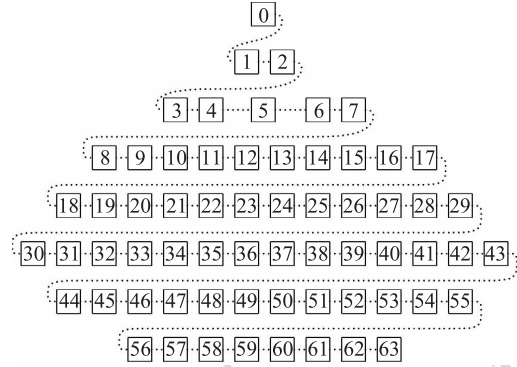
1.2 并行处理

当前CPU的发展方向,已经由单纯的提升最高主频,转为向多核的方向发展。因此一个系统的并行计算能力,将直接决定其系统的最高负载能力。利用实验室已有的研究成果^[6],将并行计算融合到编码器中,充分利用了高性能CPU的多核能力,并优化了内部调度机制和并行计算的通信机制。

对于图3中同一行的帧,相互之间没有预测参考关系,可以将其分配给任意的CPU核心,进行并行编码。对于相邻两行的帧,部分连续的帧上下行之间也没有预测参考关系,同样可以并行,如编号



(a) 预测结构



(b) 拓扑排序

图3 预测结构的拓扑排序

Fig. 3 Topological sort for prediction structure

6~10的帧。以图2中的预测结构为例,并行编码的时间为

$$T = \begin{cases} \left\lceil \frac{1}{P} \right\rceil + \left\lceil \frac{2}{P} \right\rceil + \left\lceil \frac{61}{P} \right\rceil & P \leq 5 \\ \left\lceil \frac{1}{P} \right\rceil + \left\lceil \frac{2}{P} \right\rceil + \left\lceil \frac{5}{P} \right\rceil + \left\lceil \frac{56}{P} \right\rceil & P > 5 \end{cases} \quad (4)$$

2 实验结果

以北京世纪鼎点软件公司的PoweSmart T42型号编码器作为硬件环境,进行了MVC的转码和编码测试。

表1中,第1组使用标准测试序列ballroom,8路视点,640×480像素分辨率;第2组使用欧洲杯半决赛的3D直播片源,2路视点,960×540像素分辨率。由于转码不涉及底层的编解码工作,只在码流的上层,如序列参数集(SPS)、图像参数集(PPS)、增强信息(SEI)、片头(slice header)进行语法解析,并对预测参考结构进行映射变换,所以计算复杂度和内存占用量都很小,两组测试显示,转码的CPU占用率低于5%,内存不超过1MB。

表 1 MVC-H. 264 转码测试

Table 1 MVC-H. 264 transcoding experiment

测试序列	码率/(bit/s)	CPU 占用率/%	内存/KB
ballroom	600K	3	700
欧洲杯直播	4M	5	800

将 MVC-H. 264 转码作为独立模块,加入 H. 264 编码器,再根据映射变换后的预测结构,对已有的 H. 264 编码器作少量的修改,即可实现 MVC 编码。

表 2 中,前两组是由微软研究院提供的测试序列^[7],8 路视点,1 024 × 768 像素分辨率,8 Mbit/s 码率,单路视点 1 Mbit/s,编码输出帧率达到400 帧/s,单路视点为 50 帧/s。对于 720P 规格的视频,基本可达到两路片源的实时编码。后 3 组使用标准测试序列,8 路视点,640 × 480 像素分辨率,4 Mbit/s 码率,单路视点 500 Kbit/s,编码输出帧率达到 900 帧/s,单路视点为 100 帧/s。可同时支持 4 路片源的实时编码。

表 2 MVC 编码测试

Table 1 MVC encoding experiment

测试序列	码率/(bit/s)	帧率/(帧/s)
Ballet	8 005.56	402.34
Breakdancers	8 003.03	393.91
ballroom	3 998.96	873.24
exit	4 002.83	970.89
vassar	4 002.60	918.93

3 结 论

针对 3DTV 系统面临的编码效率瓶颈和应用兼容性问题,利用 MVC 与 H. 264 在底层帧以下编码

单元技术的一致性,提出基于拓扑排序的 MVC-H. 264双向码流转换方法,能够实现实时转码并兼容 H. 264 现有编解码器。基于本文 MVC-H. 264 双向实时转码技术研制的多标准兼容 3D 视频编码器已在央视国际 2012 年的网络春晚和欧洲杯的 3D 直播中实际应用,是国内首次通过互联网开展的大规模 3D 视频直播应用。

参考文献 (References)

[1] Merkle P, Smolic A, Muller K, et al. Efficient prediction structures for multiview video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(11): 1461-1473. [DOI: 10.1109/TCSVT.2007.903665]

[2] Su Y, Vetro A, Smolic A. Common test conditions for multiview video coding. JVT-T207[R]. Klagenfurt, Austria: ITU-T Video Coding Experts Group (ITU-T SG16)2006.

[3] Chen T, Kashiwagi Y, Lim C S, et al. Coding performance of Stereo High Profile for movie sequences. ITU-T & ISO. IEC JTC1/SC29/WG11 Doc. JVT-AE022[R]. London, UK: ITU-T Video Coding Experts Group (ITU-T SG16)2009.

[4] 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. [DOI: 10.1109/JPROC.2010.2098830]

[5] Chen Y, Wang Y K, Ugur K, et al. The emerging MVC standard for 3D video services[J]. EURASIP Journal on Applied Signal Processing, 2009,(8):1-13. [DOI: 10.1155/2009/786015]

[6] Pang Y, Sun L F, Wen J T, et al. A framework for heuristic scheduling for parallel processing on multicore architecture: a case study with multiview video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(11): 1658-1666. [DOI: 10.1109/TCSVT.2009.2031463]

[7] Tech G, Wegner K, Chen Y. MV-HEVC Draft Text 4, Document JCV3V-D1004[R]. Incheon: ITU-T Video Coding Experts Group, 2013.