

【本文献信息】阮若林 胡瑞敏. 数字视频压缩编码技术标准现状与展望[J]. 电视技术 2014 38(3).

数字视频压缩编码技术标准现状与展望

阮若林¹ 胡瑞敏²

(1. 湖北科技学院 电子与信息工程学院 湖北 咸宁 437100; 2. 武汉大学 计算机学院 湖北 武汉 430072)

【摘要】信息技术和计算机互联网技术的飞速发展改变了人们的生活方式,以视频为核心的多媒体信息已经成为人们获取信息的主要来源。随着视频存储与传输的广泛应用,高效视频编码技术研究已经成为多媒体技术的研究热点。系统讨论了当前流行的视频编码技术标准及框架与采用的关键技术、标准的效率与复杂度的关系,并重点分析了未来视频编码技术的发展趋势。

【关键词】数字视频; 压缩编码; 技术标准; 感知编码

【中图分类号】TP391

【文献标志码】A

Status and Prospects of Digital Video Compression Coding Technology Standard

RUAN Ruolin¹, HU Ruimin²

(1. School of Electronic and Information Engineering, Hubei University of Science and Technology, Hubei Xianning 437100, China;

2. Computer School, Wuhan University, Wuhan 430072, China)

【Abstract】The rapid development of information technology and computer internet technology changes the people's way of life, the video has become a main source of information, and efficient video encoding technology has become the multimedia technology research hotspot. The current popular video encoding technology standards framework and its key technologies are introduced in this paper. And then the relationship between these standards of efficiency and complexity are discussed. Lastly the trends of the future video coding technology are analyzed in this paper.

【Key words】digital video; compression encoding; technology standard; perceptual coding

1 数字视频编解码技术标准

进入 21 世纪以来,随着信息技术、多媒体技术和计算机网络的飞速发展,人们的生活方式及获取信息的方式方法正在发生改变。处于其核心地位的多媒体信息已成为人类获取信息的最主要载体之一。特别是基于视频图像的编解码技术及网络通信的应用,已经成为当前多媒体技术研究的热点。当前,基于先进音视频编解码技术标注的数字高清电视、数字立体电视、网络流媒体、无线移动视频通信、高清视频监控等应用已经得到广泛应用。数字视频编解码技术标准成为这些应用的核心,高效视频编码技术的研究也成为多媒体技术广泛应用的关键。视频编码的根本目的是在降低编码码率的同时能够提供高质量的视频图像给终端用户。编码和传输人类视觉不能察觉到的或者不感兴趣的信息对于网络带宽和终端设备来说是一种浪费。从本质上来说,压缩机制和视觉系统面临着相似问题,即怎样以一种有效的方式来描述视觉对象以消除图像内部的视觉冗余。基于此种思想,从视频图像内容的角度出发,利用人眼视觉理论来指导视频压缩编码,在保持视频图像主观感知效果基本不变的条件下,尽可能地降低

编码码率成为未来视频压缩编码技术发展的趋势之一。因此,本文将从视频编解码采用的关键技术、当前主要的视频编解码标注制定机构及所制定的标注、各类标准的压缩效率与复杂度和未来视频编码技术的发展趋势等几个方面对视频编解码技术进行详细介绍。

2 视频编解码框架与关键技术

传统数字压缩编码技术是建立在 Shannon 信息论的基础上,经过大半个世纪的发展,已经取得了众多重要的研究成果。1948 年,Oliver Pierce 和 Shannon 等人在 Proceedings of IRE 上发表了 The philosophy of PCM^[1]一文,该文利用脉冲编码调制技术对模拟电视信号的数字化处理进行研究,这是现代数字视频压缩编码技术的起源。1969 年,在美国召开的第一届“Picture Coding Symposium”(图像编码会议)预示着数字图像视频编码已经以独立的学科身份跻身于学术界。

视频编码就是在限定失真条件下,通过特定的压缩编码技术消除视频信号间存在的空间冗余、时间冗余、统计冗余和视觉冗余等信息,以最低的编码码率来表示视频信号。

基金项目: 国家自然科学基金项目(61271256; 61231015); 国家“973”计划项目(2009CB320906); 国家科技重大项目(2010ZX03004-003); 湖北省教育厅科学技术研究项目(D20132802)

第一代数字视频压缩编码技术是以 Shannon 的信息论为基础,利用信源符号的统计概率模型,以像素或像素块为单位,采用传统信号分析方法消除视频图像及图像序列中存在的空间冗余、时间冗余和信息熵冗余,以达到压缩的目的。第一代数字视频压缩编码技术是基于像素和像素块的压缩方法,在编码过程中不关心视频图像的具体内容,基本不考虑人眼的视觉特性和视频信息的具体含义

及重要程度,在低码率时不能提供较好的视频质量。目前,国际视频编解码标准主要采用了第一代视频编码技术,都是以 Shannon 的信息论为基础,采用基于变换、量化、预测和熵编码等技术的混合编码框架,图 1 给出了基于变换/预测的混合视频编码框架图^[2],图 2 给出了从 1950 年到现在重要视音频编码技术的发展历程。

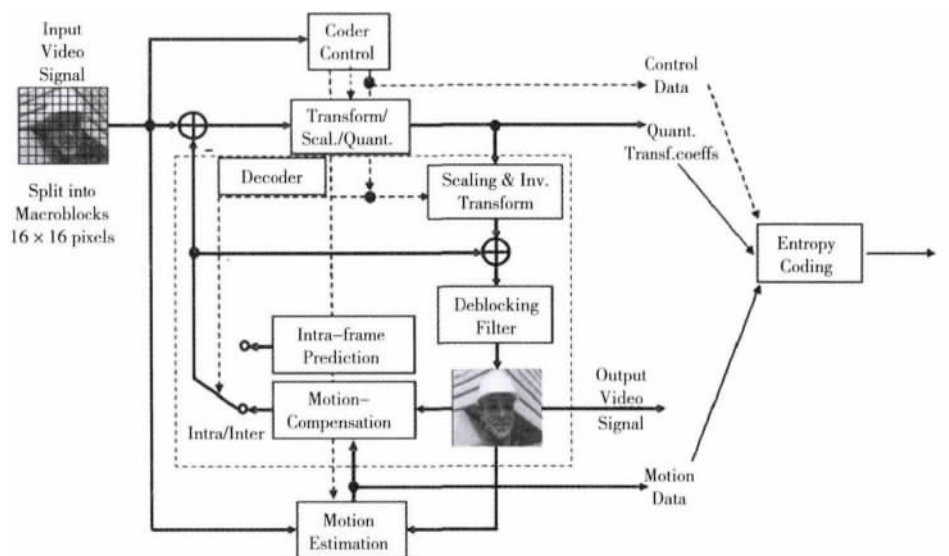


图1 基于变换/预测的混合视频编码框架图

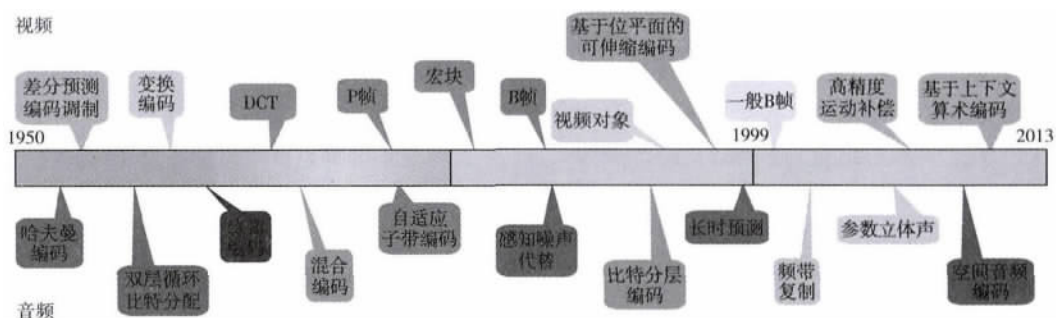


图2 重要视频、音频编码技术发展历程

3 主要视频编码标准介绍

1948年, C. E. Shannon 在 The Bell System Technical Journal 上发表的 A Mathematical Theory of Communication^[3]一文奠定了现代通信的理论基础,现有的视频编码框架都是以 Shannon 的信息论为基础,采用变换、量化、预测和熵编码等技术的混合编码方案,取得了许多重要的研究成果。20世纪90年代以后,在 ISO/IEC 和 ITU-T 等组织的协调下,对不同时期图像和视频编码的研究成果进行了收集、整理、综合、加工,制订了 JPEG^[4]、JPEG2000^[5]、MPEG-1^[6]、MPEG-2^[7]、MPEG-4^[8]、H.261^[9]、H.263^[10]和 H.264/AVC^[11]等三大系列图像视频编码标准。目前形成了以 JPEG2000、MPEG-4、H.264/AVC、

VC-1^[12](微软视频编解码标准 WMV9)和 AVS-P2^[13](我国自主知识产权音视频编解码标准)等为代表的多个先进视频编解码标准,得到了业界的广泛认可。这些标准融合了各种性能优良的视频图像编码方法,代表了不同时期视频图像编码的发展水平。

随着图像和视频编码与处理算法研究的不断发展,针对不同的应用需求,近几十年来相继形成了以 JPEG-x、MPEG-x、H.26x、VC-1 和 AVS 为主的五大系列图像和视频编码压缩标准,图 3 给出了这五大系列图像和视频编码国际标准的制定时间,大体上回顾了三十年来图像和视频压缩编码标准总的发展历程。

随着计算机网络技术和移动通信技术的迅速发展,目前最先进的视频编码标准 H.264/AVC 和 AVS-P2 都

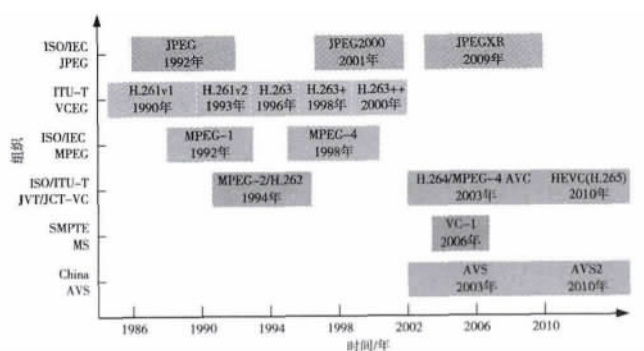


图3 重要图像、视频编码标准发展历程

将满足无线视频编码和传输应用作为自己的主要目标之一。H.264/AVC 以更高的编码效率和对各类网络的传输适应性的特点满足了无线传输中对带宽资源的限制要求,成为所有无线视频应用的一个最具吸引力的候选标准。2004年3月 AVS 工作组专门成立了专门的移动视频编码标准小组 AVS-M,制定移动视频编码标准 AVS-M 以满足移动无线环境下的各类视频编码传输应用需求^[14]。随着3G的全面应用以及4G标准的大力推进,目前一些国际知名公司的研究机构、各国科研机构 and 高等院校都将无线视频应用作为一个重要的领域展开相关研究,他们的积极参与极大地推动了新一代视频编码标准研究和发展的。

4 主要视频编码标准的压缩率与复杂度对比

第一代采用变换/预测的混合视频编码框架主要是以去除视频信号间的数据冗余来提升压缩编码的效率,较少考虑视觉上的冗余特性。近一二十年来,编码效率的提升主要依赖于压缩编码工具集的不断扩展。随着压缩编码工具集的不断扩展,编解码器的复杂度也不断大幅增加,图4给出了MPEG-x系列编码标准压缩率和编解码复杂度的关系曲线,以1993年推出的MPEG-1标准为基准,从图中可以看出当前先进视频编码标准的复杂度不断增加,压缩效率却提升很慢,同时,压缩效率的提高也显著增加了硬件实现的成本,阻碍了视频编码技术在实际标准中推广应用。

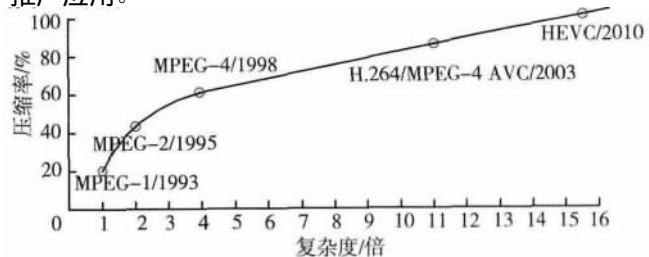


图4 视频编码标准压缩率与编解码复杂度关系曲线图

当前以JPEG-x, MPEG-x, H.26x, VC-1和AVS为主的五大系列图像和视频编码压缩标准融合了各种性能优良的视频图像编码方法,采用变换/预测的混合视频编码框架,虽然代表了不同时期视频图像编码的发展水平,也得到了业界的广泛认可,但是,该技术框架发展到现在遇到巨大瓶颈,主要表现在^[15]:

第一,该框架是以去除视频信号间的数据冗余来提升编码效率,较少考虑视觉上的冗余特性,随着编码工具集的不断扩展,数据冗余几乎被去除殆尽,再沿用传统的编码思路其编码效率已很难提升;

第二,传统框架的编码效率提升主要依赖于复杂度的大幅增加,要想进一步提升编码效率需要引入一些复杂度相当高的编码工具,这样就会显著增加硬件实现的成本,最终妨碍视频编码技术的实际应用。

因此,在传统数据冗余几乎被去除殆尽的情况下,再沿用传统的编码思路很难提升其编码效率,必须引入新的编码理论和方法。

5 未来视频编码技术的发展趋势

国际标准化组织ISO/IEC MPEG和ITU-T VCEG等非常重视这方面的工作,已于2010年成立专门协作团队JCT-VC(Joint Collaborative Team on Video Coding)开始了下一代视频编码国际标准H.265,即高效视频编码(High Efficiency Video HEVC)的研究和制定工作,主要包括高性能视频编码(High Performance Video Coder, HVC)和下一代视频编码(Next Generation Video Coding, NGVC)。并于2010年4月在德国召开了第一次会议,国际上许多著名的研究机构和公司均已开展相关的研究,一场新的技术竞争正在全球范围内展开。我国在多媒体编码和传输方面的研究历来薄弱,由于没有掌握MPEG-1, MPEG-2, MPEG-4和H.264/AVC等音视频编码标准的自主知识产权,导致DVD、VCD、数字电视等产业的发展受到很大制约,为了使移动多媒体业务不再重蹈覆辙,在新一轮的全球技术和标准竞争中占据有利地位,开展移动多媒体编码传输理论与技术的研究就显得非常必要和及时。2002年,在原信息产业部的大力支持下,我国音视频编解码技术标准组织AVS工作组(Audio Video coding Standard Workgroup of China, AVS)成立,并制定了相关视音频编解码标准AVS-P2和AVS-P3,现在也已经启动了AVS2的研究与制定工作^[17]。

近年来,随着计算机网络和通信技术的飞速发展,移动通信取得了商业运营的巨大成功,无线多媒体通信已经成为未来移动通信的发展趋势。未来通信必将以实现

“任何人可以在任何地点、任何时间与其他任何人以任意通信方式进行畅通无阻的多媒体通信”为目标^[18]。其中无线视频应用尤为重要,在移动终端上向用户提供视频通信服务正成为3G和4G等移动通信技术的发展方向。据原信息产业部预测,在3G时代,移动增值业务带来的收入将达到或超过运营商总收入的50%以上,各种基于高速移动通信网络的增值业务层出不穷,其中以移动视频点播、移动视频电话、移动视频监控、视频短消息、移动远程教育等为代表的移动多媒体业务最具发展前景,以无线视频通信为核心的这些多媒体业务将成为未来3G及4G业务的主导。

而在支撑移动多媒体业务的众多技术中,多媒体编码与传输理论和技术则是其中的核心,具有非常显著的地位。由于无线视频应用中存在着无线信道传输带宽波动、丢包率高、跨区切换频繁以及终端异构等一系列典型问题,它们将最终影响解码视频图像的质量。针对无线视频通信中高误码率问题,在先进视频编码的研究中,视频容错编码是其主要的手段之一^[19]。早在2005年,H.264/AVC及其扩展可伸缩视频编码(Scalable Video Coding, SVC)的需求文档JVT-N027^[20]中明确提出了视频编码的容错需求。我国数字音视频编解码技术标准组织AVS工作组已经完成了视频标准AVS-P2的制定工作,其制定的面向安防的视频编解码标准AVS-S^[21]也在其需求文档AVS-N1400^[22]中明确提出了容错的需求;同时,在2008年6月启动的AVS-P2中专门设立移动档次来研究无线视频编码的相关技术,其中对无线视频容错技术也提出了相应要求。但是,当前视频编码标准所采用的容错技术基本上还是靠增加冗余比特来增强视频压缩码流的容错性,容错导致了编码效率的降低,现有视频编码技术并没有很好解决容错性能和编码效率之间的矛盾。因此,为了满足有质量保证的视频业务在移动通信网络上传输的需求,必须突破传统视频编码思路的局限,引入新的编码思路和技术。

因此,下一代视频编码框架必须从编码理论上寻求突破,而自从Shannon创立信息论以来,视频编码技术与标准一直是沿着数字信号处理的思路发展。基于像素和像素块的视频处理系统与人眼视觉系统相比,其感知与认知能力相距甚远。因此,在探明人类视觉信息处理的基本生理、心理机理的基础上,建立更加高效的视觉信息编码及处理理论与方法已成为共识。

第二代数字视频压缩编码技术^[23]是建立在图像分析与合成、计算机图形学、计算机视觉等理论的基础上,与第一代视频压缩编码技术相比,进一步考虑到了去除视频图

像及图像序列中的结构冗余、知识冗余和视觉冗余,以获得更高的压缩效率。但是,第二代数字视频压缩编码中仍有许多编码技术尚需进一步的研究和探索。随着计算机技术、脑科学、神经心理学等学科的进一步发展,基于人眼视觉的编码理论有了较大发展,视觉编码理论认为人眼是视频信号的最终接收者,编码和传输人类视觉不能察觉到的或者不感兴趣的信息对于网络带宽和终端设备来说都是一种浪费,因此,编码过程中应充分考虑人眼视觉特性和视频图像内容本身对编码的影响。第二代数字视频压缩编码理论^[24]主要提出了基于对象的编码方法^[25]、基于模型的编码方法^[26]和基于分形的编码方法^[27]等几种编码技术,其中JCT-VC和AVS2目前正在展开基于人眼视觉理论的感知视频编码技术研究^[28],代表了下一代视频编码的研究方向,尚未达到成熟阶段,但其中超高清视频编码^[29]和立体视频编码技术^[30]有望在未来几年得到广泛应用。

6 结束语

当前主流视频编码技术仍然是第一代基于像素的编码方法,虽然能较好地消除图像的时空冗余和统计冗余信息,但主要是以较高的计算复杂度来换取较好的压缩效率,基本上没有有效去除其他冗余的机制。第二代数字视频压缩编码技术是建立在图像分析与合成、计算机图形学、计算机视觉等理论的基础上,进一步考虑到了去除视频图像及图像序列中的结构冗余、知识冗余和视觉冗余,以获得更高的压缩效率。相关标准制定组织正在大力开展下一代高性能视频编码技术标准的制定工作,相信在不久的将来一定会有性能优于当前先进编码方法的新一代图像视频编码技术。

参考文献:

- [1] OLIVER B, PIERCE J, SHANNON C. The philosophy of PCM[J]. Proceedings of the IRE, 1948, 36(11): 1324-1331.
- [2] WIEGAND T, SULLIVAN G, BJONTEGAARD G, et al. Overview of the H.264/AVC video coding standard[J]. IEEE Trans. Circuits and Systems for Video Technology, 2003, 13(7): 560-576.
- [3] SHANNON C. A mathematical theory of communication[J]. The Bell system Technical Journal, 1948, 27(6): 379-423.
- [4] ISO/IEC 10918. CCITT T. 81(JPEG), Grayscale and true color for static image compression international standards[S]. 1992.
- [5] ISO/IEC 15444-1: 2000(JPEG2000), Information technology - JPEG 2000 image coding system - part 1: core coding system[S]. 2001.
- [6] ISO/IEC JTC1 IS 11172(MPEG-1), Coding of moving picture and coding of continuous audio for digital storage media up to 1.5 Mbps[S]. 1992.

- [7] ISO/IEC JTC1 IS 13818(MPEG-2) ,Generic coding of moving pictures and associated audio[S]. 1994.
- [8] ISO/IEC JTC1 IS 14386(MPEG-4) ,Generic coding of moving pictures and associated audio[S]. 1998.
- [9] ITU-T Recommendation H. 261 ,Video codec for audiovisual services at px64 kbit/s[S]. 1993.
- [10] ITU-T Recommendation H. 263 ,Video coding for low bit rate communication[S]. 1996.
- [11] ITU-T Rec. H. 264/ISO/IEC 14496-10 AVC ,Draft ITU-T recommendation and final draft international standard of joint video specification [S]. 2003.
- [12] 韩云 陈祖爵. H. 264 ,VC-1 和 AVS 视频编码研究[J]. 电视技术 , 2007 ,31(2) :6-9.
- [13] 数字音视频编解码技术标准工作组(AVS) . 信息技术 先进音视频编码 第二部分: 视频(GB/T 20090.2) (AVS-P2) [S]. 2006.
- [14] 张新晨. 面向无线信道的视频编码与传输算法研究[D]. 武汉: 武汉大学 2006.
- [15] 阮若林. 基于视觉感知特性的无线视频容错编码技术研究[D]. 武汉: 武汉大学 2011.
- [16] 蔡晓霞 崔岩松 邓中亮 等. 下一代视频编码标准关键技术[J]. 电视技术 2012 ,36(2) :80-84.
- [17] AVS 工作组. 下一代 AVS 视频编码标准的技术需求 Version 2.0 [S]. 2009.
- [18] 章坚武. 移动通信[M]. 3 版. 西安: 西安电子科技大学出版社 2011.
- [19] 杜建超. 基于 H. 264/AVC 的视频容错技术研究[D]. 西安: 西安电子科技大学 2006.
- [20] SULLIVAN G ,WIEGAND T. Joint video team of ITU-T VCEG and ISO/IEC MPEG[S]. 2005.
- [21] 胡瑞敏 牟晓弦 李明. 面向视频监控的视频编解码技术[J]. 电视技术 2008 ,32(5) :68-71.
- [22] 数字音视频编解码技术标准工作组(AVS) . 面向监控应用的技术需求[S]. 2007.
- [23] KUNT M ,KONOMOPOULOS A ,KOCHER M. Second generation image coding techniques [J]. Proceedings of the IEEE ,1985 ,73(4) :549-575.
- [24] SIKORA T. Trends and perspectives in image and video coding [J]. Proceedings of the IEEE 2005 ,93(1) :6-17.
- [25] ISO/IEC JTC1/SC29 WG11/N3536 ,MPEG - 4 Overview. V. 15 [S]. 2000.
- [26] PEARSON D. Developments in model-based video coding [J]. Proceedings of the IEEE ,1995 ,83(6) :892-906.
- [27] HE C ,YANG S. Novel progressive decoding method for fractal image compression [J]. IEEE Proceedings Vision Image and Signal Processing 2004 ,151(3) :207-213.
- [28] 蒋刚毅 朱亚培 郁梅 等. 基于感知的视频编码方法综述[J]. 电子与信息学报 2013 ,35(2) :474-483.
- [29] 鲁业频 李素平. 立体视频技术的发展现状综述[J]. 电视技术 , 2012 ,36(12) :28-31.
- [30] 孙乐民 薛永林. 超高清数字电视关键技术研究[J]. 电视技术 , 2012 ,36(6) :17-20.

✉

作者简介:

阮若林(1974—) ,博士 副教授 硕士生导师 ,CCF、ACM 会员 主要研究方向为无线视频编解码与通信;

胡瑞敏(1964—) ,博士生导师 主要研究方向为多媒体信息处理、安防信息技术等。

责任编辑: 时 雯

收稿日期: 2013-04-01

(上接第 3 页)

- [3] CANCINO-SUAREZ S ,KLIMASZEWSKI K ,STANKIEWICZ O ,et al. Enhancement of stereoscopic depth estimation by the use of motion information [C]//Proc. 2012 44th Southeastern Symposium on System Theory (SSST) . Jacksonville ,FL: IEEE Press 2012: 94-98.
- [4] 武翠芳 王永芳 刘静 等. 一种改进的基于 BP 的立体匹配算法 [J] , 电视技术 2011 ,35(20) :28-31.
- [5] SCHARSTEIN D ,SZELISKI R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. International Journal of Computer Vision 2002(47) :7-42.
- [6] LI H ,CHEN G. Segment-based stereo matching using graph cuts [C]//Proc. CVPR 2004. [S. l.]: IEEE Press 2004: 74-81.
- [7] VEKSLER O. Efficient graph-based energy minimization methods [EB/OL]. [2013-02-02]. <http://www.csd.uwo.ca/faculty/olga/Papers/thesis.pdf>.
- [8] ANDREAS K ,MARIO S ,KONRAD K. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure [C]//Proc. 18th International Conference on Pattern Recognition ,ICPR 2006. Hong Kong: IEEE Press 2006: 15-18.
- [9] COMANICIU D ,MEER P. Mean shift: a robust approach toward feature space analysis [C]//Proc. PAMI 2002. [S. l.]: IEEE Press ,2004: 603-619.
- [10] BLEYER M ,GELAUTZ M. Graph-based surface reconstruction from stereo pairs using image segmentation [C]//Proc. Conference Videometrics VIII. [S. l.]: SPIE 2005: 288-299.

✉

作者简介:

董 欣(1988—) ,女 硕士生 主研 3D 视频高效深度提取技术;

王国中 博士 教授 博士生导师 研究方向为视频编解码与多媒体通信。

责任编辑: 时 雯

收稿日期: 2013-03-14