

基于不对称树结构的3D-SPIHT算法改进

吴德龄, 胡 栋

(南京邮电大学 江苏省图像处理与图像通信重点实验室, 江苏 南京 210003)

摘 要:对3D-SPIHT视频编码算法中时空方向树结构进行研究。虽然采用不对称树结构取代对称树结构,可以提高3D-SPIHT算法的压缩性能,但是在使用不对称树的条件下,如果每个GOF编码单元选取的比较小,会产生明显的边界效应,影响压缩性能。为此采用全样本对称周期延拓的方法进行边界延拓。实验结果表明,改进后不对称树结构的3D-SPIHT算法的压缩性能得到显著提升,从而证明经过边界延拓之后,基于不对称树结构的3D-SPIHT算法在不同的压缩率条件下信噪比均优越于对称树算法。

关键词:3D-SPIHT; 不对称树; 小波分解; 边界延拓

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2010)08-0079-04

An Improved Algorithm for 3D-SPIHT Based on Asymmetric Tree Structure

WU De-ling, HU Dong

(Jiangsu Key Laboratory of Image Processing and Image Communication, Nanjing University
of Posts and Telecommunications, Nanjing 210003, China)

Abstract: The structures of spatio-temporal orientation trees in 3D-SPIHT video coding system are investigated. Although asymmetric tree structure instead of symmetric tree structure can improve the compression ratio in 3D-SPIHT algorithm, under asymmetric tree structure, if there is a small size of GOF, it'll cause boundary effect. Therefore, full-sampled symmetric periodic extension method is adopted here in order to alleviate boundary effect. Experimental results show that the modified algorithm significantly improves the compression performance, which demonstrate that under different compression ratio, asymmetric tree structure outperforms symmetric tree structure for all the tested video sequences after boundary extension.

Key words: 3D-SPIHT; asymmetric tree; wavelet transform; boundary extension

0 引言

三维等级树集分割编码方法(3D-SPIHT)^[1]是B.J. Kim和W.A. Pearlman等人提出的视频编码方法,它通过三维小波变换来消除帧间冗余和帧内冗余,并将三维小波分解得到的系数在三维方向上构建三维时空方向树,通过三个集合来组织编码。这种算法复杂度低,并且产生嵌入式码流,因此有广泛的应用前景。然而, Kim和Pearlman提出的3D-SPIHT算法所构建的时空方向树仅仅是SPIHT算法中二维小波树在三维空间的扩展,即所谓的“对称树结构”,但视频序

列在时间和空间方向上不一定具有“对称”的统计特性,因此这种对称树结构对于3D-SPIHT算法来说并不是一种最好的选择。He Chao等人提出的“不对称树结构”^[2]是另一种组织时空方向树的方法。用这种不对称树结构来组织编码,可以使3D-SPIHT算法获得更好的压缩性能。但在实验中发现,用不对称树结构的3D-SPIHT进行编码时,如果取很少的帧来作为一个编码单元组(GOF),会出现明显的边界效应。针对这一情况,文中对不对称树结构的3D-SPIHT算法进行了改进,用全样本对称延拓的方法对每个GOF边界进行处理,从而有效地改善了边界效应,进一步提高了3D-SPIHT算法的信噪比。

1 不对称树结构的3D-SPIHT算法及改进

1.1 不对称树结构的3D-SPIHT分析

传统的3D-SPIHT算法通常运用三维二进小波

收稿日期:2009-11-27;修回日期:2010-01-29

基金项目:国家重大基础研究项目(2007CB310607-3);国家自然科学基金项目(GZ206005)

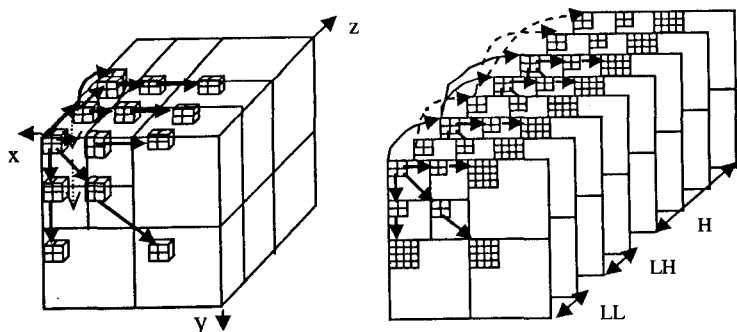
作者简介:吴德龄(1985-),女,江苏南京人,硕士研究生,研究方向为图像处理与多媒体通信;胡 栋,教授,硕士生导师,研究方向为图像处理与多媒体通信。

分解^[3~5],将空间和时间方向的小波分解交替进行,它是一种对称形式的三维小波分解方法。为了很清楚地理解这种分解方式,用 T_x, T_y 表示空间 X, Y 方向的小波分解, T_z 表示时间 z 方向的小波分解。在进行三维二进小波分解时,小波分解顺序为 $T_z T_x T_y T_z T_x T_y \dots$, 对称树结构就是针对三维二进小波分解系数而构建的。但是,这种小波分解方式在三维环境中并不是最佳的方式,因为视频序列在每个方向上并不具有对称的特性^[6],如果用平均标准偏差来衡量每个方向,可以发现,空间 X, Y 方向的平均标准偏差相近,而时间 z 方向的平均标准偏差比空间方向小的多。为了解决这个问题, KIM^[7] 等人应用三维小波包分解将时间和空间的小波分解独立开来。同样用 T_x, T_y 表示空间 X, Y 方向的小波分解, T_z 表示时间 z 方向的小波分解。在进行三维小波包小波分解时,小波分解的顺序为 $T_z T_z \dots T_z T_x T_y \dots T_x T_y$, 这样就可以将时间和空间做不同级数的小波分解,减少时间方向小波分解级数,而增加空间方向小波分解级数。

不对称树结构是针对三维小波包分解而产生的,是一种更好地组织三维小波时空方向树结构的方法。

图 1(a) 为不对称树结构立体图。它是以 2×2 个相邻系数而不是 $2 \times 2 \times 2$ 为一个系数单元。这种不对称树结构不但同样可以保持父代和子代小波系数的相关性,还可以使时空方向树构造的更长更窄。树更长可以聚集更多的零,从而令编码压缩效率更高。

图 1(b) 是不对称树结构逐帧细化图,其中以 8 帧图像为例,做两级小波包分解,其中时间轴方向第一次小波分解形成 L, H 子带,第二次对 L 子带分解形成 LL, LH 子带。



(a) 不对称树结构 3D-SPIHT

(b) 不对称树结构逐帧细化图

图 1 不对称树结构图

在不对称树结构中,视频序列的每一帧内树的结构基本和 2D-SPIHT 相同。在 2D-SPIHT 中,最低

频 LL 子带的每个系数单元中最左上角的系数没有孩子系数,它们不属于树结构的一部分,然而在不对称树结构中,但除了最高频时间子带以外,其余时间子带的空间最低频子带中每个系数单元最左上角的系数在其下一个时间子带相同的空间位置具有孩子系数,即存在一种帧间的联系。如图 1(b), LL 子带每帧的每个系数单元最左上角的系数在 LH 子带相应位置分别有 2×2 个孩子系数, LH 子带每帧的每个系数单元最左上角的系数在 H 子带每两帧的相应位置有 2 个 2×2 的孩子系数。如果将 2×2 个系数组成的系数单元看作一个整体,以这个整体左上角的系数坐标 (i, j, k) 来表示这个系数单元的坐标,那么每个系数单元的孩子系数单元的坐标可以表示为式(1)。

式(1)中, F 为时间最低频子带的帧数; R 为空间最低频子带的行数; C 为空间最低频子带的列数。

运用不对称树结构进行编码,在没有加入算术编码的情况下能够有效地提高 3D-SPIHT 算法的压缩性能^[2]。在运用算术编码的情况下,看到信噪比的提升效果没有不运用算术编码时明显,这是因为算术编码这种熵编码方法对压缩效率产生的提升优势覆盖了不对称树算法对压缩效率的提升优势,所以在没有加入算术编码的情况下比较两种算法会更加直观和清晰。不对称树算法还具有一定的灵活性,因为它适用的是小波包分解,当时间和空间方向采用不同的小波分解级数时^[8],对称树算法已不再适用,而此时不对称树算法不失为一种很好的选择。

1.2 改善边界效应的方法

在利用不对称树结构的 3D-SPIHT 算法对视频序列进行压缩的实验中发现,当采用不同长度的编码单元组(GOF)来进行编码时,如果将一个编码单元组(GOF)的帧数选取的很少,比如 4 帧, 8 帧, 16 帧,而时间轴进行小波变换时,在每个 GOF 的边界处, PSNR 会出现突然变小的现象,有时候甚至可以达到好几个 dB,造成视频序列的反馈抖动,这就是所谓的边界效应^[9,10]。但是,为了减少系数时延和所需的存储容量,时间轴一次选取的图像帧数不宜太多,通常选取每个 GOF 的长度在 16 帧以下,受小波滤波器的限制,这就必然导致边界效应。而且边界效应在处理运动幅度比较大的序列时会更加明显。产生边界效应的原因主要有两个

$$O(i, j, k) = \begin{cases} (i, j + R, k), (i, j, k + C), (i, j + R, k + C), (i + F, j, k) & i < F, j < R, k < C \\ (i, j + R, k), (i, j, k + C), (i, j + R, k + C), (2i, j, k), (2i + 1, j, k) & j < R, k < C \\ (i, 2j, 2k), (i, 2j + 2, 2k), (i, 2j, 2k + 2), (i, 2j + 2, 2k + 2) & \text{others} \end{cases} \quad (1)$$

方面,一是对于有限长度序列进行正反变换后没有满足完全重建,二是在上述情况下进行压缩后造成 GOP 间的差异。为了解决完全重建问题,需要在变换中使用合适的边界延拓。

文献[11]描述了几种常用的边界延拓方式,包括补零延拓、周期延拓、对称延拓、重复边界点延拓和双对称扩展延拓。对于对称延拓^[12],根据边界点像素不同的处理方式,分为半样本对称延拓、全样本对称延拓和半样本反对称延拓。针对不同类型的小波和信号,选取不同的边界延拓方式。如果采用正交小波进行小波变换,滤波器长度为偶数时可以采用单纯的周期延拓,而对称延拓适用于对称双正交小波变换。一般来说,使用双正交小波基可以获得比较好的图像质量,但是边界效应会变得更加明显,从而影响整个序列的平均压缩效果,进行边界延拓显得尤为重要。鉴于文中采用的是奇对称双正交小波进行小波变换,是一种具有线形相位的滤波器,在时间方向采用了全样本对称周期延拓法来完全重建。对于每个 GOF,在其信号头尾两端分别进行,如果延拓后长度依然小于滤波器的长度,则根据全样本对称延拓的结果周期性地延长信号,对该方法的示意图如图2所示。在小波变换前,先对输入信号进行全样本对称周期延拓,使其变成无限长信号,再通过滤波并隔点抽取以后,在不丢失信息的情况下,适当截取部分变换系数作为子带信号,从而使分解过程不增加信号的样本总数。采用该延拓方法不仅可以实现小波变换的完全重建,同时也不增加变换后的数据量。

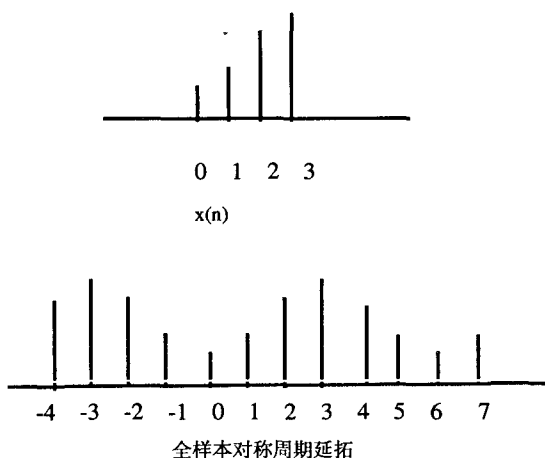


图2 边界延拓示意图

这里用 N 表示 GOF 的长度,其中从 $-(N-2)$ 到 $(N-1)$ 段全样本对称延拓所用的公式如下:

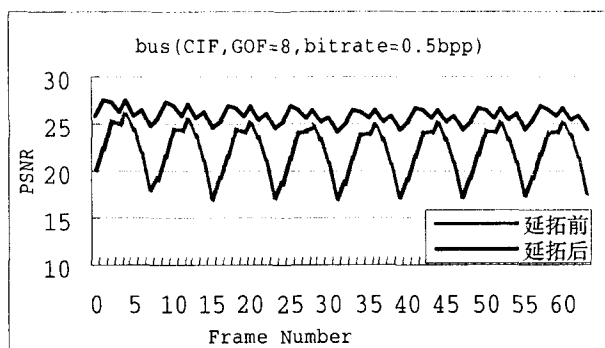
$$X(n) = \begin{cases} x(n) & 0 \leq n \leq N-1 \\ x(-n) & -(N-2) \leq n \leq -1 \end{cases} \quad (2)$$

在经过全样本对称周期延拓之后,之前存在的边界效应得到了明显改善,重建序列不但在边界处的

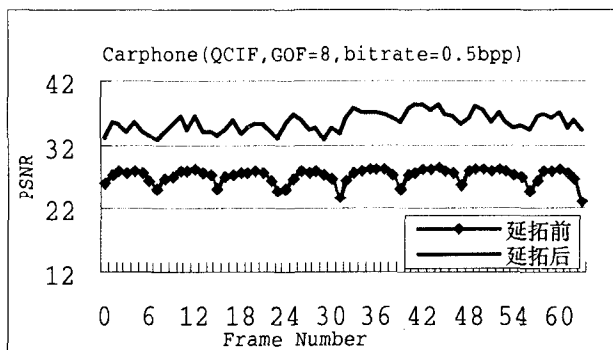
PSNR 值得到提高,整体的 PSNR 均有显著提升。实验中还发现,对一个视频序列进行边界延拓,压缩率越低,边界效应的提升效果越明显,在高压压缩率条件下,边界效应的提升效果变得淡化。这是因为对视频序列压缩程度高,使得量化步长变大,从而使小波变换后的系数产生的零变多,边界处的差异减小,导致边界延拓所起到的作用没有压缩率低时大。总而言之,通过对边界的处理有效地提高了基于不对称树的 3D-SPIHT 算法的信噪比。

2 实验结果及讨论

为了证实上述结论,选择 CDF9/7 小波进行三维小波变换,并选择了 bus、Carphone 两个测试序列。其中 bus 序列运动幅度比 Carphone 大,造成时间方向边缘抖动较大,边界效应较 Carphone 序列更为明显。在 Linux 下的 C 语言环境进行实验^[13],选取 GOF=8,压缩率为 0.5bpp,在图3中可以看到经过全样本对称周期延拓之后边界效应的和整体压缩性能的改善。



(a)

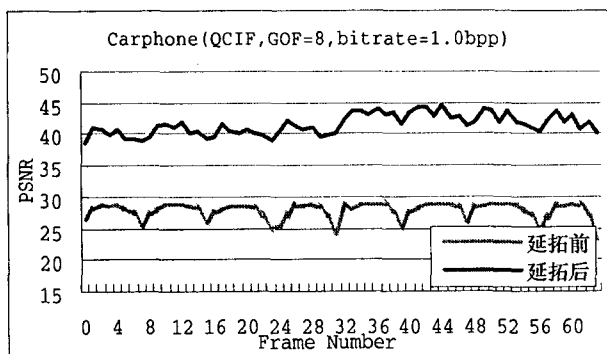


(b)

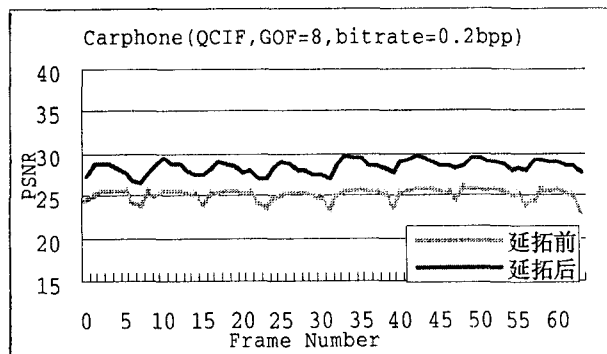
图3 边界延拓前后 bus 和 Carphone 序列逐帧 PSNR 对比图

图4是 carphone 序列在不同压缩率情况下的效果对比,包括 0.2bpp 和 1bpp 的情况,并结合图3中的 0.5bpp 的情况比较,可以看出在低压压缩率情况下边界效应的改善更明显。在 GOF=4 和 GOF=16 的情况下,边界延拓产生的效果如图5所示,边界效应同样得

到了改善。

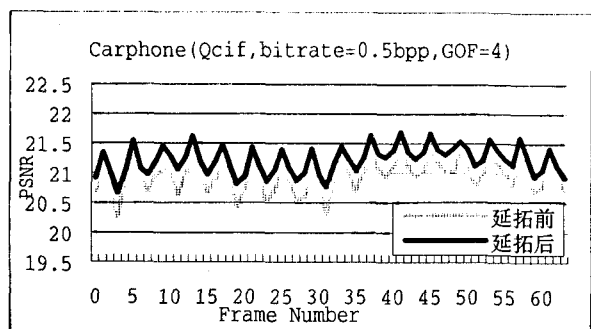


(a)

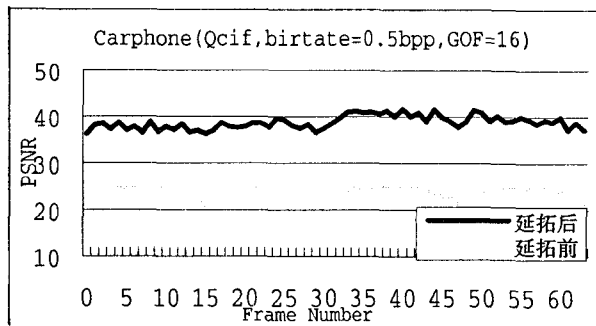


(b)

图 4 不同压缩率边界延拓前后逐帧 PSNR 对比图



(a)



(b)

图 5 不同 GOF 边界延拓前后逐帧 PSNR 对比图

该边界延拓的方法同样可以运用于传统的对称树结构的 3D-SPIHT, 同样也可以改善其边界效应, 提

高其 PSNR。图 6 显示了不应用算术编码的情况下, 在 0.5bpp 时, 不对称树和对称树算法都经过边界延拓之后其 PSNR 的比较。从而证明经过边界延拓之后, 基于不对称树结构的 3D-SPIHT 算法的压缩性能相对于传统的 3D-SPIHT 算法得到很大程度的提升。

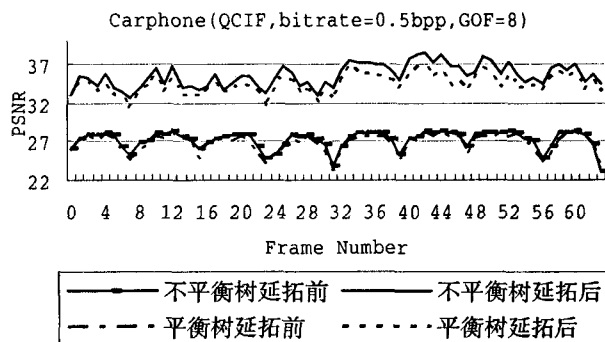


图 6 边界延拓前后对称树和不对称树算法逐帧 PSNR 对比图

3 结束语

对不对称树结构的 3D-SPIHT 编码算法进行了研究, 针对时间方向的边界效应问题, 通过采用合适的边界延拓方法加以改进, 并且通过实验对比验证了改进算法的性能。实验证明, 在经过边界延拓之后, 运用不对称树的 3D-SPIHT 算法的 PSNR 进一步得到了提高。

参考文献:

- [1] Kim B J, Pearlman W A. An Embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT)[C]//Proc. IEEE Data Compression Conf. [s.l.]: [s.n.], 1997:251-260.
- [2] He Chao, Dong Jianyu. Optimal 3-D Coefficient Tree Structure for 3-D Wavelet Video Coding[J]. IEEE transaction on circuits and systems for video technology, 2003(10): 961-972.
- [3] 粘永健, 吴乐华, 何世彪. 基于三维小波变换的视频图像编码研究[J]. 影像技术, 2006(2): 26-28.
- [4] 马 宣. 基于三维小波变换的视频图像压缩算法的研究[D]. 西安: 西安建筑科技大学, 2006.
- [5] 王 双. 基于三维小波变换的视频编码系统的研究与实现[D]. 北京: 北京邮电大学, 2006.
- [6] Burns T J, Rogers S K, Oxley M E, et al. A Wavelet Multiresolution Analysis for Spatio-Temporal Signals[J]. IEEE Trans. Aerosp. Electron. Syst., 1996, 32(4): 628-649.
- [7] Kim Beong-Jo, Xiong Zixiang, Pearlman W A. Low Bit-Rate Scalable Video Coding with 3-D Set Partitioning in Hierarchical Trees (3-D SPIHT)[J]. IEEE transaction on cir-

(下转第 86 页)

2.3.2 性能分析(对于情况(1)不再讨论)

定理 2:算法能保证得到 kmCDS;退出节点数目为 n 时,维护算法对近似系数的影响为 $O(n\Delta)$;维护算法最坏情况下的消息复杂度为 $O(\Delta^2)$,时间复杂度为 $O(\text{Diam})$,Diam 为该控制域内连通部分的直径。

证明:对于 m -控制性,只需要考虑 p 控制域内的边缘节点,在 p 退出后,这些节点能被原 kmCDS 中 $m-1$ 个节点控制,对于 a) 中,建立 CDS 后,CDS 外的节点又同时被 CDS 中至少一个节点控制,所以在 CDS 加入原 kmCDS 后,边缘节点至少被 m -控制;b) 中的 m -控制性显然满足。对于 k -连通性,由于 p 退出部分控制节点之间变为 $k-1$ -连通的,在执行算法后,a) 中这部分节点之间,通过新加入的 CDS 中的节点可以得到另一条连通路径,b) 中因为节点的退出不影响整个网络的 k -连通性,那么变为 $k-1$ -连通的节点之间一定存在第 k 条连通路径,而第 k 条路径的丢失是由于节点 p 的退出,可以在 p 的邻居集中找到一条路径来替换该条路径。所以 k -连通性得到了保证。

最坏的情况下,将 p 控制域里所有的边缘节点都加入 kmCDS,且假设所有退出的 n 个节点无公共的边缘节点。此时需要向 kmCDS 中加入 $n\Delta$ 个节点。所以对近似系数的影响为 $O(n\Delta)$ 。

情况 a) 中用 CDS-BD- $D^{[10]}$ 算法来构造 CDS,消息复杂度为 $O((\Delta+1)|V|)$,在这里 $|V|$ 决定于所有退出节点控制域的边缘节点数;情况 b) B_1 中节点周围的控制节点都要向 B_1 中发送消息确定其状态,需要 $O(\Delta^2)$ 的消息。所以最坏情况下消息复杂度为 $O(\Delta^2)$ 。

情况 a) 时间复杂度为 $O(\text{Diam})$;情况 b) B_1 直接加入 D 只需 $O(1)$ 时间,所以算法时间复杂度为 $O(\text{Diam})$ 。

3 结束语

针对无线网络终节点的移动情形做了分类分析,给出了一个移动情况下 kmCDS 的维护算法。该算法

针对不同的节点移动情形分别给出了不同的局部解决方法。并通过不同情形下的算法分析,证明了利用该算法在无线移动网络中可以很好地维护虚拟骨干网的 k -连通性与 m -控制性。

参考文献:

- [1] Dai Fei, Wu Jie. On Constructing k -Connected k -Dominating Set in Wireless ad hoc and Sensor Networks[J]. Journal of Parallel and Distributed Computing, 2006, 66(7): 947-958.
- [2] Shang Weiping, Yao F, Wan Pengjun, et al. Algorithms for Minimum m -Connected k -Dominating Set Problem[C]//COCOAS 2007. [s.l.]: [s.n.], 2007: 182-190.
- [3] Wu Yiwei, Wang Feng, Thai M T, et al. Constructing k -Connected m -Dominating Sets in Wireless Sensor Networks [C]//Military Communications Conference. Orlando, FL: [s.n.], 2007: 29-31.
- [4] Wu Yiwei, Li Yingshu. Construction Algorithms for k -Connected m -Dominating Sets in Wireless Sensor Networks [C]//Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing. Hong Kong, China: [s.n.], 2008: 83-90.
- [5] 凤旺森, 屈婉玲, 王捍贫, 等. 无线自组织网络中构造 2-连通 k -支配虚拟主干网的近似算法[J]. 计算机工程与科学, 2008, 30(10): 21-26.
- [6] 王雪瑜, 山炳强, 马 艳. 无线传感器网络中 2-连通 2-支配集的分布式构造算法[J]. 青岛大学学报: 工程技术版, 2008, 23(2): 22-26.
- [7] 孙立山, 张瑞宏, 武文斌. 2-连通 2-支配集的集中式构造[J]. 计算机工程与应用, 2009, 45(15): 107-110.
- [8] 施 伟. 移动 Ad Hoc 网络中连通支配集若干关键问题的研究[D]. 杭州: 浙江大学, 2007.
- [9] West D B. Introduction to Graph Theory[M]. 2nd ed. [s.l.]: Prentice-Hall, Inc, 2001.
- [10] Kim D, Wu Yiwei, Li Yingshu, et al. Constructing Minimum Connected Dominating Sets with Bounded Diameters in Wireless Networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2009, 20(2): 147-157.

(上接第 82 页)

- cuits and systems for video technology, 2000(8): 1374-1387.
- [8] Kim Beong-Jo, Pearlman W A. Fast Color - Embedded Video Coding with SPIHT[C]//Data Compression Conference. NY: [s.n.], 1998.
- [9] 孙延奎. 上波分析及其应用[M]. 北京: 机械工业出版社, 2005.
- [10] 陈武凡. 小波分析及其在图像处理中的应用[M]. 北京: 科学出版社, 2002.
- [11] Karlsson G, Vetterli M. Extension of Finite Length Signals for Subband Coding[J]. SP, 1989, 17: 161-168.
- [12] Li S P. Shape - Adaptive Discrete Wavelet Transforms for Arbitrarily Shaped Visual Object Coding[J]. IEEE transaction on circuits and systems for video technology, 2000(5): 725-743.
- [13] 杨志强. Linux 下的图像/视频编码算法开发平台的设计[D]. 天津: 天津大学, 2005.