

# 基于 H. 264 和双树小波的多描述视频编码

陈 婧 李 莉 蔡灿辉

(华侨大学, 信息科学与工程学院, 厦门, 361021)

**摘 要:** 本文针对互联网和无线信道等不可靠网络的视频传输问题, 提出一种基于 H. 264 和双树小波变换的多描述视频编码解决方案。采用分层的多描述视频编码框架, 实现 H. 264 和双树小波编码的有机结合。基本层用 H. 264 编码器对视频信号进行低码率编码后, 复制到各个描述; 增强层用三维双树小波变换对原视频和基本层重建视频的差值进行编码, 将产生的四棵三维小波树经噪声整形后两两组合, 编码送到不同描述。在解码端, 若能够接收到两个描述, 则通过中心解码器实现高质量的视频重建; 若丢失一个描述, 则通过边解码器解码仍可保证一定质量的视频重建。实验结果表明在相同码率下, 本算法的视频中心解码和边解码质量优于现有的多描述视频编码算法。

**关键词:** 多描述视频编码; 分层编码; H. 264; 双树小波变换

**中图分类号:** TN919.8    **文献标识码:** A    **文章编号:** 1003-0530(2011)08-1265-06

## H. 264 and Dual-tree Discrete Wavelet Transform based Multiple Description Video Coding

CHEN Jing LI Li CAI Can-hui

(School of Information Science and Engineering, Huaqiao University, Xiamen, China, 361021)

**Abstract:** A H. 264 and dual-tree discrete wavelet transform (DDWT) based multiple description video coding algorithm is proposed to solve the transmitting error or packet loss problem due to Internet or wireless network channel failure. Each description of the proposed multiple description coding scheme consists of a base layer and an enhancement layer. First, the input image sequence is coded by a standard H. 264 encoder in low bit rate to form the base layer, which is to be duplicated to each description. Then the error frames of the base layer and the input image sequence is coded by a 3D dual-tree wavelet encoder to produce four coefficient trees. After noise-shaped, these four trees are partitioned into two groups, forming the enhancement layer of related descriptions. If all descriptions are received, a high quality video can be reconstructed by a central decoder. If only one description is received, a side decoder can be used to guarantee an acceptable quality reconstructed video. The simulation results have shown that the quality of reconstructed video by the proposed algorithm outperforms the state-of-the-art of multiple description video coding methods.

**Key words:** multiple description video coding; layered coding; H. 264; dual-tree discrete wavelet transform

### 1 引言

互联网和无线通信网络传输中的数据错误、数据丢失和过度延时等现象会造成视频传输错误, 导致视频重建质量严重下降, 成为限制网络多媒体服务进一步发展的瓶颈。多描述编码(Multiple Description Coding, MDC)是针对不可靠信道传输问题提出的信源差错复原编码技术, 具有较好的抗突发干扰能力, 在实现差错掩盖和重建丢失分组方面有较广泛的适用范围, 为解决图像和视频经不可靠网络传输后质量严重下降问题提供了一种有效的解决方案, 成为近年来通信与

信号处理研究领域的一个热点<sup>[1]</sup>。

自 1993 年 Vainishampayan 发布第一个多描述编码器——量量化器<sup>[2]</sup>以来, 将多描述编码方法用于不可靠网络中图像和视频的传输有效性得到了越来越多的理论和实践验证。现有的多描述编码方法大致可以分为多描述量化编码、多描述变换编码、多描述子带编码(MD-SPIHT), 多相变换和选择量化算法(PTSQ), 基于非平衡保护的多描述编码、基于框架扩展的多描述编码和近几年提出的结合分层思想的多描述编码算法<sup>[3-4]</sup>等。

针对多描述视频编码方法的研究始于 1999 年。最直接的想法是将多描述图像编码方法直接扩展到多

描述视频编码,如基于量化的多描述视频编码<sup>[5]</sup>和基于变换的多描述视频编码<sup>[6]</sup>。这类方法的问题是由于传输错误所引起编解码端的参考帧不一致(失配)。消除失配的最有效方法是对每个描述分别进行运动预测和运动补偿<sup>[7]</sup>,但这不仅会降低编码效率,而且还会增加计算复杂度。因此如何解决失配问题是多描述视频编码的一个研究重点。近年来,为了提高 H. 264 编码标准的错误恢复性能,一些基于 H. 264 的多描述视频编码方案不断涌现。2004 年, Bernadini 和 Durigon 等人在图像处理国际会议(ICIP)上提出一个基于 H. 264 的多相空间亚采样多描述视频编码(Polyphase Spatial Subsampling Multiple Description Coding: PSS-MDC)<sup>[8]</sup>,对原始图像进行空间亚采样,得到的四个子图像序列。再采用 H. 264 编码器分别对他们进行编码,产生四路码流。这种方法具有很强的抗误码能力,即使在解码端有三路码流无法正常到达或是解码错误时,仍可以恢复出一定质量的信号。然而这种方法带有相当大的冗余量。文献[9]在 PSS-MDC 的基础上,先对子帧进行一定的处理以减少码率,再通过标准的 H. 264 编码器进行编码。文献[10]在 H. 264 的标准上对基于量化的多描述编码方案<sup>[2]</sup>进行了改进,使得更多的量化系数映射为 0 系数,以此提高压缩效率。文献[11]使用了 H. 264 的冗余片技术来产生多描述,但不适用于低码率传输。

本文在分层多描述编码框架的启发下,提出了一种基于 H. 264 和双树小波的分层多描述视频编码方案。基本层采用 H. 264 编码器对输入视频进行低码率编码产生两描述的公共信息分别拷贝到各个描述。增强层采用三维双树小波变换产生四组小波系数,再将码字分成两组编码后形成各描述的特有信息,分别送入两个描述。每个描述包含基本层和增强层信息,经不同信道分别传输。实验结果表明本方法可以获得较高的编码效率,使视频编码算法具有较强的抗信道干扰能力,且冗余度可调。

本文的其余部分安排如下:第二节介绍 H. 264 和双树小波变换,第三节讨论本文所提视频多描述编码方法的框架和算法实现,第四节列出实验结果并进行分析,第五节给出结论。

## 2 H. 264 和双树小波变换

### 2.1 H. 264

H. 264/AVC 是 ITU 和 ISO/IEC 联合制定的最新国际视频编码标准。和以前的视频编码标准一样,H. 264 采用传统的基于块的混合编码模式。因为 H. 264 运用了大量的新技术,如帧内预测编码,帧间预测编码,1/4 像素精度运动向量搜索,4×4 整数 DCT 变换,环内去块效应自适应滤波器,多帧预测,先进的熵编码等,使编

码性能比以前的视频编码标准有很大的提高。在同等的图像质量条件下,H. 264 的数据压缩比是 MPEG-2 的 2-3 倍,MPEG-4 的 1.5-2 倍。在差错控制方面,H. 264 虽然增加了参数集(Parameter Sets),灵活宏块排序(FMO),冗余片技术(RS)等差错复原工具,提高了网络适应性,但其抗击不可靠网络传输的能力仍无法达到令人满意的效果。

### 2.2 双树小波变换(DDWT)

双树小波变换的概念最初是由剑桥大学数字信号处理组主任 Nick Kingsbury 教授于 1998 年提出的。他设计了一对实数小波变换滤波器组,分别构成双树复小波变换的实部和虚部。实验验证了双树复小波变换具有近似的平移不变性、良好的方向选择性和有限的冗余性等性质<sup>[12]</sup>。然而,由于双树复小波的结构及复数系数的冗余性,其复系数的实部、虚部具有相同的方向选择性。为了降低双树复小波变换的冗余度,Sellesnick 提出了一种实系数双树小波变换(DDWT)<sup>[13]</sup>,成功地将系数冗余度降为原来的 1/2,且保留了原来的 6 个方向高频子带,如图 1 所示。

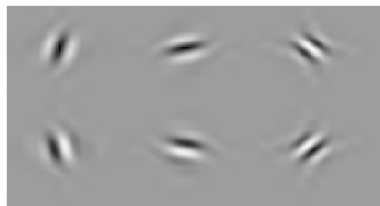


图 1 DDWT 的方向子带

Fig. 1 The directional subbands of DDWT

### 2.3 三维双树小波变换

三维双树小波变换可以利用四路三维离散小波变换(DWT)并行完成<sup>[14]</sup>,如图 2 所示。其中,\*表示卷积运算, $\downarrow 2$ 表示二下采样。 $x$ ,  $y$  和  $z$  分别表示视频序列的三个坐标轴方向(水平、竖直和时间)。 $h_0$  和  $h_1$  分别代表低通和高通滤波器,彼此构成 Hilbert 变换对,以使实数小波变换能够实现完美重构。为了方便描述,图 2 中的  $*xh_0$  表示视频信号与  $h_0$  低通滤波器在水平方向上进行卷积运算; $\downarrow_x 2$  表示信号在水平方向上进行二下采样,其余符号采用类似表示方法,不再一一说明。图 2 画出其中一路三维 DWT,共包含了 1 个低频子带(LLL)和 7 个高频子带(LLH, LHL, LHH, HLL, HLH, HHL, HHH)。构成三维 DDWT 需要四路类似的结构,因此三维 DDWT 有 4 个低频子带(LLL)和 28 个高频子带,具有很强的方向滤波作用。图 3 按行列出了三维 DDWT 生成的四棵小波树 4 个低频子带和 28 个高频子带的滤波方向图,并按各小波树在图 3 的出现顺序对其进行编号。

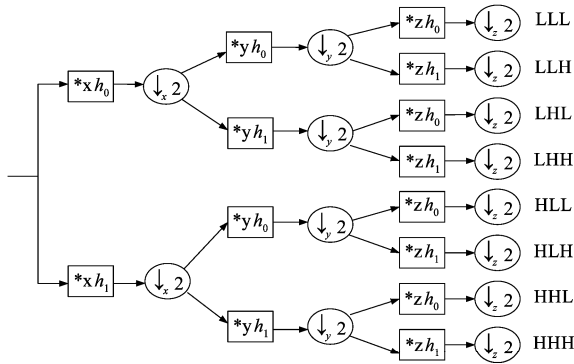


图 2 三维实数离散小波变换滤波器组示意图

Fig. 2 The filter band of 3D DDWT

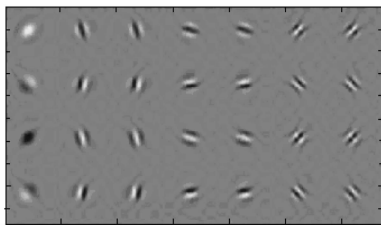


图 3 三维 DDWT 的 28 个高频方向子带

Fig. 3 The 28 directional high frequency subbands of 3D DDWT

### 2.4 双树小波系数的稀疏表示

然而,DDWT 的多方向选择性是通过提高变换的冗余度所换来的。一个视频序列经过三维 DDWT,输出的数据量是传统三维 DWT 的四倍。因此在对小波系数进行编码前,必须对三维 DDWT 所产生的四棵小波树进行稀疏化处理。Reeves 和 Kingsbury<sup>[15]</sup> 在研究了 DDWT 系数后发现,大多数的冗余系数的值都比较小,且较均匀地分布在各个子带中,类似于高斯白噪声。因此他们提出噪声整形(Noise shaping, NS)算法,在小波域和原始信号域进行信号迭代预测,挑选有效系数,摒弃无效系数,并对留下的重要系数进行误差修正,使其更加接近原始图像的能量,减少重建误差。文献[14]和[16]的视频编码方案也验证了噪声整形算法对小波系数进行稀疏表示的合理性和有效性。

## 3 基于 H. 264 和双树小波的分层多描述视频编码算法

### 3.1 分层多描述视频编码框架

所谓多描述编码就是将一个信源分解成两个或多个可独立编解码的码流,分别编码后通过不同的信道传输到接收端。对信源而言,每个码流就是一个描述。目前常用的多描述视频编码方案多数是讨论两个描述的情况。多描述编码具有下述特点:(1) 每一个描述都含有信源的基本信息。只要收到一个描述(码流),接收端就可以重建出具有一定质量的信源。(2) 每一个描述都包含一些别的描述所没有的信息。因此,接

收端接收到的描述数越多,重建信号的质量越高,当接收到所有描述时重建信号的质量最好。对于两描述的视频编解码而言,如果两个描述都接收到了,解码器可重建高质量的视频信号。如果只接收到一个描述,解码器也还可以利用该描述重建具有一定质量的视频信号。如果把每个描述的共有信息作为基本层,把每个描述专有信息的集合作为增强层,就构成了分层多描述编码。

分层多描述编码的优点是容易调整码率与冗余量,但由于分层多描述编码是把增强层分割到各个描述,为了避免失配,不能对增强层进行帧间预测,因此编码效率较低。为了提高编码效率,本文采用 H. 264 编码器对输入视频进行低码率编码作为基本层,用三维 DDWT 对基本层重建视频与输入视频的误差进行编码形成增强层。由于三维 DDWT 变换具有 28 个方向子带,能够有效地描述运动场景,因此可以对它直接编码,不必进行运动估计,这样既可以消除失配,又降低了增强层视频编码的计算复杂度。此外,用三维 DDWT 对增强层进行编码,还可以补偿 H. 264 编码产生的块效应。本文提出的分层多描述视频编码框图如图 4 所示。

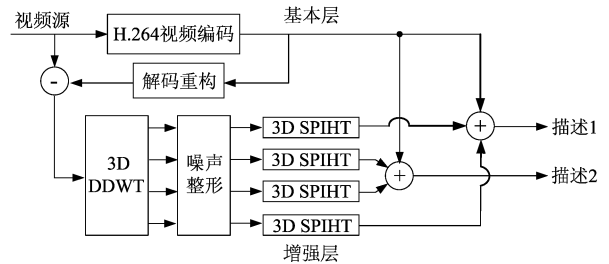


图 4 结合 H. 264 和双树小波的多描述视频编码框架

Fig. 4 The H. 264 and DDWT based multiple description video coding scheme

### 3.2 基本层的产生

由于基本层存放的是各个描述的公用信息,基本层决定了该多描述编码冗余的冗余量。因此,基本层的编码效率对整个多描述编码效率是至关重要的。因此,本文选择目前编码效率最高的 H. 264 编码器对输入视频进行低码率编码产生基本层信息,并复制到各个描述。H. 264 的多参考帧模式和 1/4 像素精度的运动估计保证了帧间预测的准确性,更精确的帧内预测编码进一步提高了其编码效率。

### 3.3 增强层的产生

对三维 DDWT 研究发现,一个视频序列经过三维 DDWT 产生的 4 棵三维小波树从不同方面反映了整个视频信息,它们之间具有一定的相关性。从图 3 容易看出,树 1 和树 3,树 2 和树 4 的方向具有较强的相关

性。如果把树1和树4分成一组,把树2和树3分成另一组,即使丢失其中一组小波树,整个视频序列的基本信息仍可以从未丢失的另一组小波树的低频子带中有所体现。因此,仍然可以通过另外未丢失的一组小波树进行一定质量的重建。

基于这个观察结果,本文首先对原始序列图像与基本层重建的序列图像间的误差经三维 DDWT 编码产生的四棵三维小波树。再把这四棵小波树送经三维噪声整形<sup>[16]</sup>,降低三维 DDWT 的冗余。然后对噪声整形后的四棵三维小波树分别经 3D SPIHT<sup>[17]</sup> 编码器进行编码形成增强层。把四棵三维小波树的码流两两分组分配给两个描述,树1和树4的码流作为一组分配给描述1;树2和树3的码流作为另一组分配给描述2。利用这四棵小波树之间的相关性,即使丢失其中两棵树的信息,仍然可以根据剩余的两棵树对增强层进行较高质量的重建。

基本层和树1、树4的码流构成描述1,基本层和树2、树3的码流构成描述2,两个描述通过不同信道进行传输。

### 3.4 解码算法

在解码端,如果能够完全接收到两个描述,则进行中心解码。先对基本层进行解码,然后再解码两个描述的四棵小波树,完整重建增强层信息。如果只收到其中一个描述,则进行边解码。首先对基本层进行重建,然后解码接收到的两棵小波树码流,并借助四棵小波树之间的相关性,重建增强层。基本层信息保证了视频的基本质量,接收到的部分增强层信息可以提供一定的误差补偿与纹理方向,保证较高质量的重建效果。

## 4 实验结果及分析

为了测试本文提出的多描述视频编码方案的性能,我们设计了多组实验,选取不同格式(cif, qcif)、不同运动速度类型序列(慢速,中速,快速),在不同帧率下进行测试。

第一组实验测试的是本算法与其他基于三维小波变换的多描述视频编码算法的性能比较。图5是本算法和 LMDVC<sup>[4]</sup> 在帧率为 15fps 时对 Foreman (qcif 格式) 序列的编码效果比较。图6给出本算法与 MCTD-MDSVC<sup>[7]</sup> 在帧率为 30fps 时对 Foreman 的测试结果。通过这两个比较图我们可以看到,本方法的中心解码和边解码效果在不同码率下均优于 LMDVC 和 MCTD-MDSVC 算法。特别是在速率较高时边解码的效果有明显的提高。图7给出的是本方法对 Foreman (cif 格式) 的重建效果图。由于本文基本层的编码效率高,而为了与文献[4]和文献[7]的实验结果进行比较,采用较高的冗余度,这

使得中心解码质量和边解码质量差距不大。

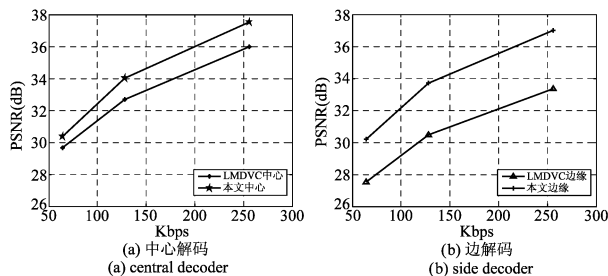


图5 本文方法和 LMDVC<sup>[4]</sup> 对 Foreman (qcif 格式, 15fps) 序列测试的结果比较

Fig. 5 The comparison of the proposed method with LMDVC<sup>[4]</sup> for Foreman (qcif, 15fps)

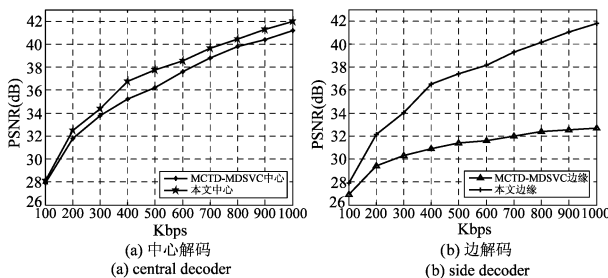


图6 本文方法和 MCTD-MDSVC<sup>[7]</sup> 对 Foreman (qcif 格式, 30fps) 序列测试的结果比较

Fig. 6 The comparison of the proposed method with MCTD-MDSVC<sup>[7]</sup> for Foreman (qcif, 30fps)

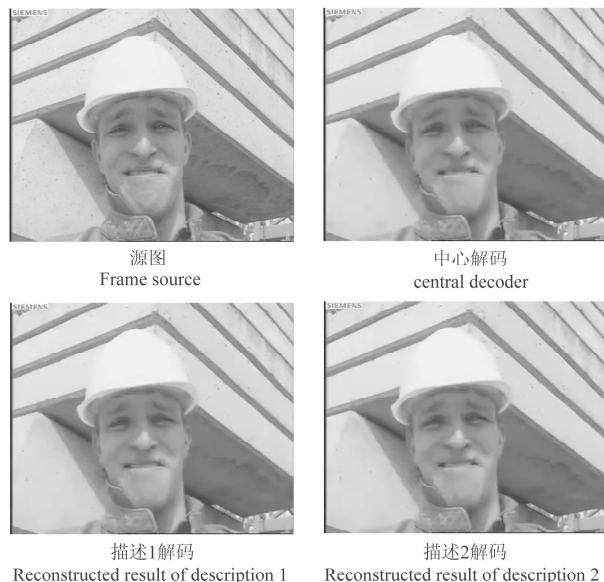


图7 Foreman 第34帧重构图像(cif, 256kbps, 30fps)

Fig. 7 The reconstructed result of the 34<sup>th</sup> frame of Foreman (cif, 256kbps, 30fps)

第二组实验是将本文方法与基于 H. 264 的多描述编码算法 PSS-MDC<sup>[8]</sup>、PSPT-MDC<sup>[9]</sup> 进行比较。图8是对不同测试序列的比较结果。对于慢速和中速的运动

序列,如 Miss-america 和 Foreman,本方法比上述两种基于 H. 264 的多描述编码效果要好得多,特别是当信道不可靠时,边解码的重建效果比上述两种有了很大的改善,这充分说明了对运动速度不是特别快的视频序列,双树小波编码作为增强层具有良好的效果。对于运动比较剧烈的序列,如 Football,本方法的中心解码效果明显优于 PSS-MDC,但稍低于 PSPT-MDC。主要原因是当运动比较激烈时,三维 DDWT 的 28 个方向子带不能有效地描述运动场景,因此直接用 3D SPIHT 对三维小波树进行编码的效果不如 H. 264。然而,当码率高于 450K/S 时本算法的边解码效果比 PSPT-MDC 好,这说明了本算法比 PSPT-MDC 有更强的差错恢复能力。

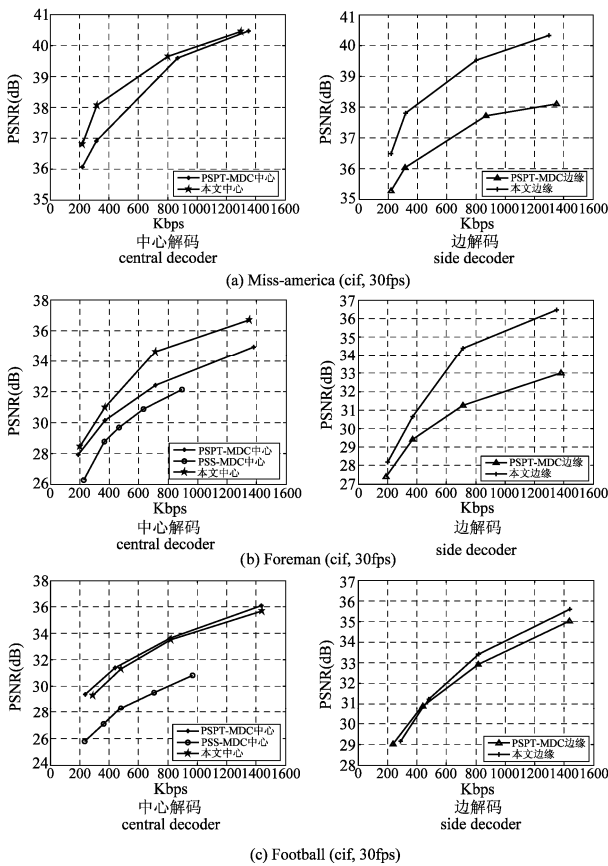


图 8 本方法和 PSS-MDC<sup>[8]</sup>、PSPT-MDC<sup>[9]</sup> 对不同序列的性能比较

Fig. 8 The comparison of proposed method with PSS-MDC<sup>[8]</sup> and PSPT-MDC<sup>[9]</sup> for different sequences

## 5 结论

本文将多描述编码与分层编码思想有机结合,提出并实现了一种新的高编码效率、冗余度可调的多描述视频编码方案。借助高效的 H. 264 视频编码器实现基本层编码,提高编码效率,降低冗余度。充分利用三

维 DDWT 产生的四棵小波树具有较强相关性这一特点,把经过稀疏表示后的四棵小波系数树根据方向特性两两分组构成不同描述的增强层信息。每个描述都由基本层信息和本描述的增强层信息组成。增强层的整帧双树小波变换改善了 H. 264 基于块的编码缺陷,三维 DDWT 的强方向滤波能力使之可以直接采用 3D SPIHT 编码器对小波树进行编码,既解决了多描述视频编码的失配问题,又降低了编码的计算复杂度。实验结果证明了本文算法的有效性。

## 参考文献

- [1] V. K. Goyal. Multiple description coding: compressing meets the network [J]. IEEE Signal Processing Magazine, 2001, 18(5): 74-93.
- [2] V. A. Vaishampayan. Design of multiple description scalar quantizer [J]. IEEE Transactions on Information Theory, 1993, 39(3): 821-834.
- [3] P. A. Chou, H. J. Wang, and V. N. Padmanabhan. Layered multiple description coding [C]. Proc. Packet Video Workshop, Apr. 2003.
- [4] J. Chen, C. Cai. Layered multiple description video coding [C]. Proc. IEEE Int. Conf. on Signal Processing, Beijing, China, 2008, 1296-1299.
- [5] V. A. Vaishampayan, S. John. Balanced interframe multiple description video compression [C]. Proc. IEEE Int. Conf. on Image Processing, 1999, 812-816.
- [6] R. A. Reibman, H. Jafarkhani, Y. Wang, M. T. Orchard, and R. Puri. Multiple description video coding using motion compensated temporal prediction [J]. IEEE Transactions on Circuits and System for Video Technology, 2002: 193-204.
- [7] C. Tillier, T. Petrisor, B. Pesquet-Popescu. A motion-compensated overcomplete temporal decomposition for multiple description scalable video coding [J]. EURASIP Journal on Image and Video Processing, 2007.
- [8] R. Bernardini, M. Durigon, R. Rinaldo, L. Celetto, and A. Vitali. Polyphase spatial subsampling multiple description coding of video streams with h.264 [C]. Proc. IEEE International Conference on Image Processing (ICIP), Singapore, Oct. 2004: 3213-3216.
- [9] Z. Wei, C. Cai, K. Ma. H. 264-based multiple description video coder and its DSP implementation [C]. Proc. IEEE International Conference on Image Processing (ICIP), Atlanta, GA, USA, Oct. 2006, 3253-3256.
- [10] O. Campana, R. Contiero, and G. A. Mian. An H. 264/AVC Video Coder Based on a Multiple Description Scalar Quantizer [C]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(2): 268-272.

- [11] T. Tillo, M. Grangetto, and G. Olmo. Redundant Slice Optimal Allocation for H. 264 Multiple Description Coding[C]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(1): 59-70.
- [12] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury. The dual-tree complex wavelet transform[J], IEEE Signal Processing Magazine, 2005, 22(6): 123-151.
- [13] I. W. Selesnick and K. Y. Li. Video denoising using 2D and 3D dual-tree complex wavelet transforms[C]. Proc. SPIE Conf. Wavelets X, CA, USA, Aug. 2003, vol. 5207: 607-618.
- [14] B. Wang, Y. Wang, I. Selesnick and A. Vetro. Video Coding Using 3D Dual-Tree Wavelet Transform [J]. EURASIP Journal on Image and Video Processing, 2007.
- [15] T. H. Reeves, N. G. Kingsbury. Overcomplete image coding using iterative projection-based noise shaping[C]. Proc. IEEE Int. Conf. on Image Proceeding, NY, USA, 2002, vol. 3, 597-600.
- [16] L. Li, C. Cai. Multiple description image coding using Dual-Tree Discrete Wavelet Transform[C]. International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), 2009, 655-658.
- [17] B. J. Kim, Z. Xiong, and W. A. Pearlman. Low Bit-Rate Scalable Video Coding with 3-D Set Partitioning in Hier-

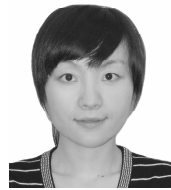
archical Trees (3-D SPIHT)[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(8): 1374-1387.

### 作者简介



陈 婧(1980-),女,福建厦门人,华侨大学信息科学与工程学院讲师,硕士学位,主要研究方向:图像和视频处理。

E-mail:chenjing8005@gmail.com



李 莉(1984-),女,河南开封人,华侨大学信息科学与工程学院硕士研究生,主要研究方向:图像和视频压缩编码。

E-mail:hababy@hqu.edu.cn



蔡灿辉(1954-),男,福建泉州人,华侨大学信息科学与工程学院教授,博士,副院长。IEEE 高级会员,信号处理学会委员,电子学会高级会员。近年主要研究方向:图像处理、数字视频、模式识别、多媒体通信等。E-mail:chcai@hqu.edu.cn