

基于预测式错误恢复机制的多描述视频编码研究

周文帅 白慧慧* 赵耀

(北京交通大学信息所 北京 100044)

摘要: 该文提出一种基于预测式错误恢复机制的多描述视频编码方法。在编码端通过预测单路解码可能产生的错误影响,为每个描述分配必要的冗余信息。考虑到视频编码压缩的效率问题,设计了不同的编码模式来处理冗余信息。解码端可以充分利用冗余信息,从而实现丢失视频帧的高质量恢复。实验表明,与传统时域采样方法相比,该方法具有更好的率失真性能。

关键词: 多描述视频编码; 标准视频编码器; 预测式错误恢复机制

中图分类号: TN919.81

文献标识码: A

文章编号: 1009-5896(2014)04-0817-06

DOI: 10.3724/SP.J.1146.2013.00858

Multiple Description Video Coding Research Based on Predictive Error Resilience

Zhou Wen-shuai Bai Hui-hui Zhao Yao

(Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China)

Abstract: A novel Multiple Description video coding scheme is proposed based on Predictive Error Resilience (MD-PER). At the encoder, the possible error caused by the single-channel reconstruction can be predicted firstly, and then the necessary redundancy information is inserted for each description. In view of the compression efficiency, different coding modes are designed to compress the generated redundant information. At the decoder, the redundancy information can be applied to achieve high quality recovery of the lost video frames. Experimental results demonstrate that compared with the traditional temporal sampling method the proposed scheme achieves better rate-distortion performance.

Key words: Multiple description video coding; Standard video encoder; Predictive error resilience

1 引言

随着因特网的爆炸式发展,视频传输应用越来越广泛。然而,网络拥塞,随机比特错误和数据包丢失会导致视频压缩数据的质量严重下降,这就给视频传输带来了极大挑战。因此,需要开发一种视频编解码方案,该方案需要有高的压缩效率,还要保证视频传输的鲁棒性。

多描述编码作为一种能在易错信道上提高传输鲁棒性的技术,吸引了越来越多的学者研究^[1,2]。当信源和信宿之间存在着很多信道时,一般情况下不会所有信道同时在一个视频帧出错,多描述编码就是基于这个思想提出来的。多描述编码在编码端信源产生多个比特流(称之为描述),每个比特流具有同样的优先级,在多个相互独立的信道上进行传输;在解码端,每一个描述都能被独立解码,重建用户可接受质量的视频序列,随着接收到的描述数量的

增加,重建视频序列的质量也随之提高。

多描述视频编码的基本思想是利用描述的冗余信息来提高传输的鲁棒性,因而,在多描述视频编码中需要一些必要的冗余,但这样会降低压缩效率,因此,多描述编解码方法的主要目标是在传输比特率和重建质量之间找到最佳折中。经典的多描述编码方法利用了变换和量化。根据多描述标量量化原则^[3],文献[4-6]中设计了多描述视频编码方法。多描述相关性变换在文献[7]中被用来设计运动补偿多描述视频编码。文献[8]提出了X树非平衡保护多描述编码方法。尽管上述方法得到了较好的实验结果,但这些方法不能兼容广泛应用的标准编解码器,如H.26x和MPEG系列;为了解决这个问题,文献[9]中多描述视频编码通过采用H.264/AVC标准中先进的视频编码工具来引入描述间的冗余信息。此外,很多多描述视频编码方法^[10,11]都针对某一种确定的标准编解码器,文献[11]针对H.264/AVC采用了片级和宏块级的多描述视频编码方法。另外,文献[12]提出的基于交织抽取与分块压缩感知策略的图像多描述编码方法压缩效率不高。文献[13]提出的适用于

2013-06-14收到,2013-08-12改回

国家自然科学基金(61272051)和教育部创新团队发展计划(IRT201206)资助课题

*通信作者:白慧慧 hlbai@bjtu.edu.cn

丢包信道的小波编码图像传输方法，利用了分层多描述编码来提高信源编码的容错性能，但该方法依赖于信源编码的结构。文献[14,15]提出的方法只针对特定图像。

为了解决已有的多描述视频编码方法只兼容特定标准编解码器的问题，以及进一步提高压缩效率和视频传输的鲁棒性，本文尝试提出一种更具广泛兼容性，压缩效率更高，视频传输更鲁棒的基于预测式错误恢复机制的多描述视频编码方法。该方法在编码端对原始视频序列进行时域抽样来产生初始视频子序列，以保证和当前标准编解码器的兼容性，然后在编码端预测解码重建可能出现的错误影响，为每个视频子序列分配必要的冗余信息。在解码端，这些冗余信息有助于更好地估计丢失信息从而获得较好的解码重建质量。

2 基于时域采样的多描述视频编码

图1是文献[2]提出的基于时域采样的多描述视频编码框图。首先，在编码端读入一组视频序列的前 N 帧，对其进行奇偶帧分离，得到一段奇数帧视频子序列和一段偶数帧视频子序列；然后，分别用标准视频编码器对两个子序列进行编码，再通过两个相互独立的信道传送到解码端。在解码端，如果两个描述均被收到，则进行中心路解码。由于收到了两路描述信息，可以用得到的两路信息进行奇偶帧排序得到中心路解码视频序列。单路解码时，单路描述根据对应解码器采用帧内插方式恢复出单路解码视频序列。本文采用的帧内插方式基于分段匀速运动假设，在重建中间帧的时候采用双向运动估计以提高预测性能，然而双向运动估计容易在所生

成的中间帧中产生重叠和孔洞。以下介绍帧内插方式具体过程。

用 f_k 和 f_{k+1} 分别表示视频序列的第 k 帧和第 $k+1$ 帧， f 表示 f_k 和 f_{k+1} 中间需要重建的帧， $MV(p)$ 表示像素运动矢量；首先，为了避免在重建帧 f 中出现孔洞，用式(1)将第 k 帧和第 $k+1$ 帧的像素值平均后生成一个初始重建帧，该初始重建帧虽然没有较高的重建质量，但可以有效避免孔洞的出现。

$$f(p) = \frac{1}{2}(f_k(p) + f_{k+1}(p)) \quad (1)$$

其中 $f(p)$ 表示重建帧中位置 p 处的像素值。然后，对第 k 帧 f_k 和第 $k+1$ 帧 f_{k+1} 分别通过后向预测和前向预测得到重建帧 f 。为了解决帧内插中出现重叠的问题，用式(2)取重叠部分的平均值作为运动补偿的最终值。最后，用所得平均值代替上一步中初始化生成的重建帧内对应像素值，从而进一步提高重建帧的质量。

$$f(p) = \frac{1}{2} \left(f_k \left(p + \frac{1}{2} MV_1(p) \right) + f_{k+1} \left(p + \frac{1}{2} MV_2(p) \right) \right) \quad (2)$$

3 基于预测式错误恢复机制的多描述视频编码方法

3.1 总体框图

图2是本文提出的基于预测式错误恢复机制的多描述视频编码方法的框图。首先读入一组视频序列，对其进行奇偶帧分离得到主信息奇数帧子序列 X_1 和偶数帧子序列 X_2 ，根据预测式错误恢复机制分别得到重建后的偶数帧序列 \tilde{X}_2 和重建后的奇数帧序列 \tilde{X}_1 ，经编码模式选择模块处理后得到奇数帧

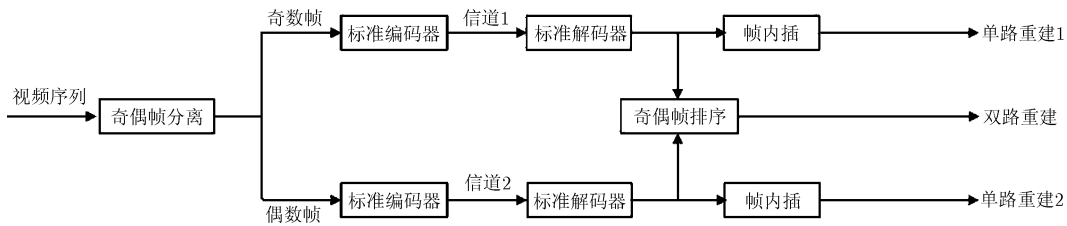


图1 基于时域采样的多描述视频编码框图

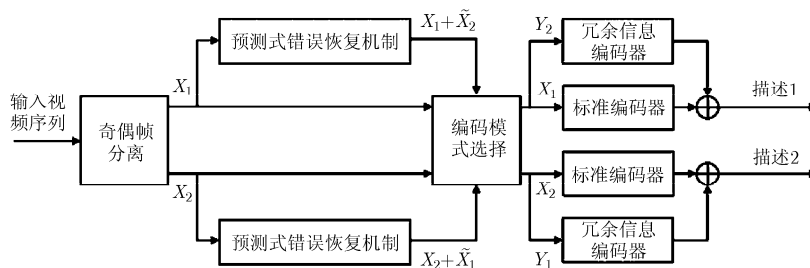


图2 基于预测式错误恢复机制的多描述视频编码框图

对应的冗余信息 Y_2 和偶数帧对应的冗余信息 Y_1 。然后对主信息 X_1, X_2 和冗余信息 Y_1, Y_2 分别采用标准编码器和冗余信息编码器进行编码。编码后, X_1 和 Y_2 得到描述 1, X_2 和 Y_1 得到描述 2, 经不同信道传输到解码端进行解码。

3.2 预测式错误恢复机制

在编码端, 对于奇数帧构成的视频子序列, 首先用奇数帧视频序列通过帧内插方式重建出偶数帧视频序列, 重建过程如图 3 所示。然后对重建得到的偶数帧进行分块, 块大小取 $a \times b$, 结合真实的偶数帧计算每一块的重建质量, 即 PSNR 值。

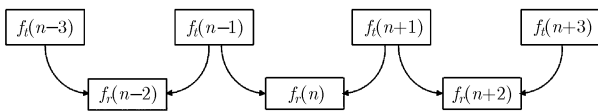


图 3 重建具体过程

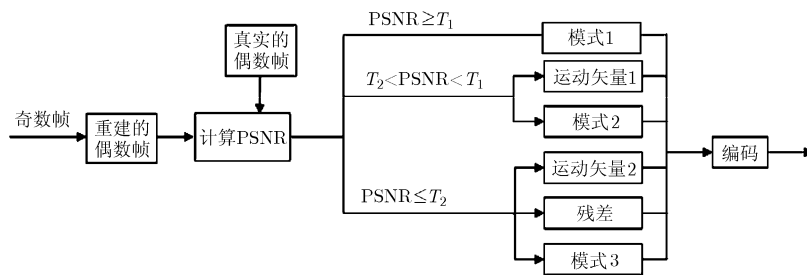


图 4 编码模式的选择过程

(2)模式 2, 当块的重建质量为 $T_2 < PSNR < T_1$ 时, 说明重建块需要一部分冗余信息来提高块的重建质量, 本文把运动矢量 MV 作为冗余信息传送到解码端, 即用 $f_i(n)$ 的当前块通过运动估计在 $f_i(n-1)$ 中寻找差异最小块, 记录其运动矢量 MV , 具体过程如图 5 所示。

(3)模式 3, 当块的重建质量为 $PSNR \leq T_2$ 时, 说明块的重建质量很差, 不仅需要传送模式 2 中的运动矢量 MV , 还需要传送残差信息。运动矢量

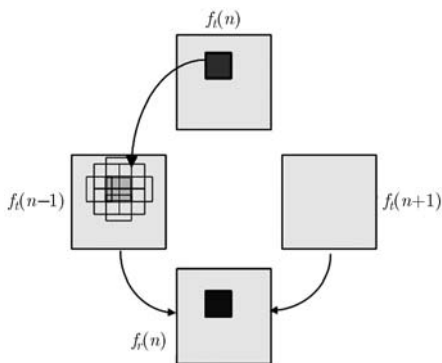


图 5 运动矢量和残差的获取

同样地, 对于偶数帧视频子序列, 用偶数帧序列通过帧内插方式重建出奇数帧序列, 然后对重建出的奇数帧进行 $a \times b$ 大小的分块, 并结合真实的奇数帧求得每一块的 PSNR 值。

3.3 编码模式选择

本节以奇数帧一路为例阐述编码模式的选择。根据上一模块计算得到的 PSNR 值设置两个阈值 T_1 和 T_2 ($T_1 > T_2$), 如果块的重建质量 $PSNR \geq T_1$, 则该模式定义为模式 1; 如果块的重建质量为 $T_2 < PSNR < T_1$, 则该模式定义为模式 2; 如果块的重建质量 $PSNR \leq T_2$, 则该模式定义为模式 3。选择过程如图 4 所示, 模式定义如下:

(1)模式 1, 当块的重建质量为 $PSNR \geq T_1$ 时, 说明只依靠主信息而不需要任何冗余信息就能够达到满意的重建质量, 因此采用灵活的 Skip 模式(不传送任何冗余信息), 降低了码率;

MV 的获取过程同模式 2。找到运动矢量 MV 后, 用当前块减去 $f_i(n-1)$ 中差异最小块得到另一部分冗余信息, 即残差数据。

由于运动矢量要无失真地传输到解码端, 因而直接对运动矢量信息进行算术编码。而对残差数据, 需要进行 DCT 变换并且量化取整, 而后进行算术编码。然后传送到解码端进行解码。

3.4 参数选择

由于阈值 T_1 和 T_2 ($T_1 > T_2$) 大致和冗余信息量呈正比关系, 当 T_1 和 T_2 取值太大时, 冗余信息就会很多; 而当 T_1 和 T_2 取值太小时, 冗余信息就会很少, 此时对视频序列的重建质量提高也不明显。因此, 可以根据信道带宽的大小或者用户要求的重建质量 (PSNR) 确定冗余量, 进而确定阈值 T_1 和 T_2 。

4 实验结果和分析

为了验证本文方法的有效性, 采用如表 1 所示的 4 个具有不同格式空间分辨率的标准视频序列来测试本系统的性能。将本文方法和文献[2]的方法在不同条件下的重建视频质量(Y 分量的 PSNR 平均

表1 输入的视频序列

序列名称	序列格式	序列分辨率	帧率(fps)
news	QCIF	176 × 144	30
mobile	CIF	352 × 288	30
soccer	4CIF	704 × 576	30
city	720P	1280 × 720	60

表 2 本文方法较文献[2]方法在复杂度和 PSNR 上的提高情况

序列名称	复杂度提高(%)	PSNR 提高(dB)
news	3.89	0.9322
mobile	4.10	0.6225
soccer	4.02	2.0489
city	4.37	2.2435

值)进行了比较。公平起见,在处理过程中,两种方法均选取各视频序列的前 100 帧,块大小都取 16×16。参数 T_1 和 T_2 的选取都是在仿真结果取得最好的性能下手动选取的,冗余信息量也由此确定。表 2 以奇数帧一路为例说明本文方法较文献[2]的方法在 QP=25 时复杂度和 PSNR 的提高情况(i5-3230 4 G 2.6 GHz PC)。为了保证取值的准确性,每个数据均测试 10 次,然后求其平均值。其中,主信息均通过 HEVC(HM-8.2, GOPsize 设置为 4, Max CUWidth 和 MaxCUHeight 设置为 64, IntraPeriod 设置为-1, TransformSkip 和 TransformSkipFast 设置为 0)进行编码,冗余信息采用算术编码(Arith06)方法进行编码。表 3 以接收到两个描述为例给出了本文方法和文献[2]的方法在不同丢包率下的重建视频质量比较。图 6 为各视频序列在收到一个描述和两个描述且丢包率为 0 时两种方法重建质量的比较。图 7 为 soccer 序列在不同丢包率的情况下本文方法和文献[2]的方法的比较。

由表 2 可以看出,本文方法相比于文献[2]的方法在复杂度提高很少的情况下,PSNR 值有了明显的提高。尤其在 720P 格式下,复杂度仅提高了 4.37%,但 PSNR 却有 2 dB 以上的提高。表 3 列出了两种方法在相同码率不同丢包率下视频序列的重建情况,从中可以看出随着丢包率的增加本文方法相比于文献[2]的方法的增益逐渐减小,这是因为随着丢包率的增加主要信息和冗余信息都丢失严重,因此重建视频质量将会随之下降。

如图 6(a)和图 6(b)所示,在丢包率为 0 的情况下,对于如 QCIF 和 CIF 格式的小分辨率视频序列,在低码率情况下,文献[2]的方法比本文方法有 0.5 dB 的增益。但在高码率情况下,本文方法比文献[2]的方法有 1 dB 以上的提高。另外,如图 6(c)和图 6(d)所示,对于大分辨率视频序列,在 4CIF 格式下平均可以提高 1~1.3 dB,在 720P 格式下可以提高 1~2 dB,其原因主要是由于高码率情况下本文方法的冗余分配能够起到更好的作用。由图 7 可以看出,

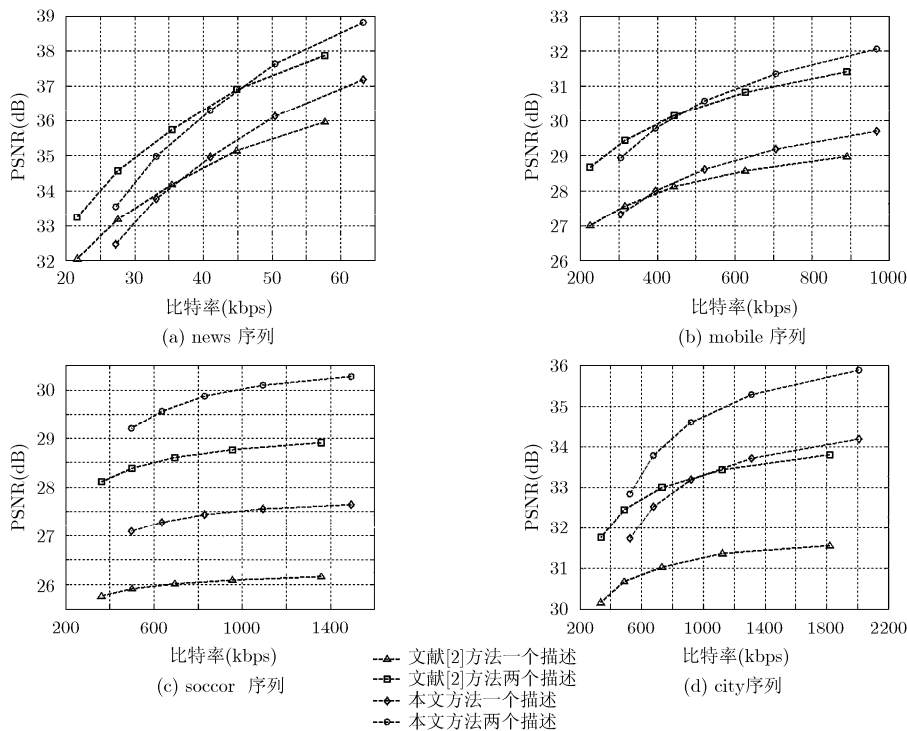


图6 本文方法和文献[2]方法在0丢包率下的PSNR

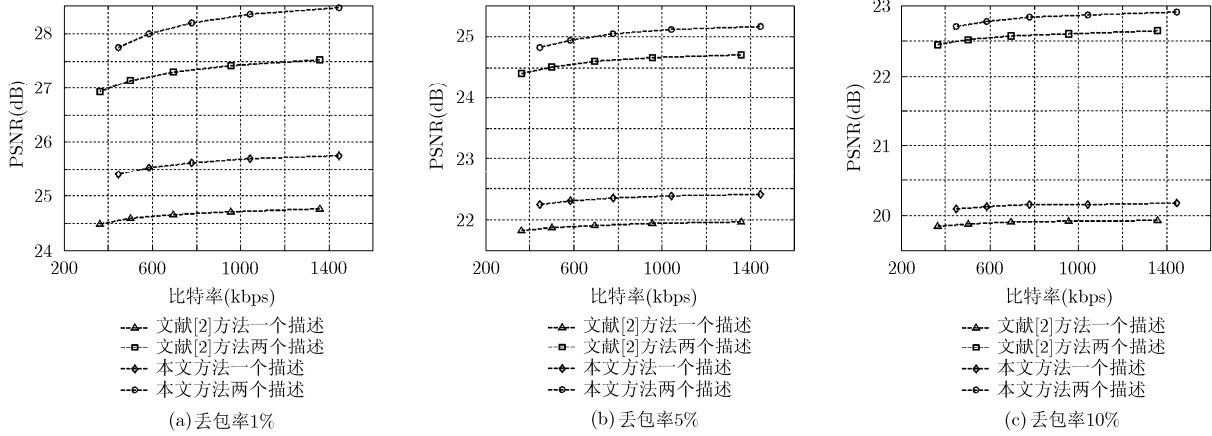


图7 soccer序列在不同丢包率下本文方法和文献[2]的方法的PSNR

表3 不同丢包率下两种方法的 Δ PSNR 比较

视频序列	码率(kbps)	1%			5%			10%		
		文献[2]方法	本文方法	Δ PSNR (dB)	文献[2]方法	本文方法	Δ PSNR (dB)	文献[2]方法	本文方法	Δ PSNR (dB)
news	20	29.3271	29.4593	0.1322	24.5140	24.5377	0.0237	21.7662	21.7629	-0.0033
	30	29.7428	29.9223	0.1795	24.6640	24.6947	0.0307	21.8323	21.8592	0.0269
	40	30.1460	30.3263	0.1803	24.7723	24.8089	0.0366	21.8849	21.9260	0.0411
	50	30.4430	30.6423	0.1993	24.8558	24.8872	0.0314	21.9272	21.9508	0.0236
	60	30.6341	30.7806	0.1465	24.9091	24.9254	0.0163	21.9464	21.9637	0.0173
mobile	250	26.9408	27.1018	0.1610	23.7170	23.7746	0.0576	21.4582	21.4705	0.0123
	350	27.4297	27.6777	0.2408	23.9343	24.0167	0.0824	21.5839	21.5970	0.0131
	500	27.8641	28.1450	0.2809	24.1249	24.2071	0.0820	21.6917	21.7077	0.0160
	650	28.2504	28.5441	0.2937	24.2770	24.3833	0.1063	21.7662	21.8048	0.0386
	900	28.5713	28.8880	0.3167	24.4043	24.4968	0.0925	21.8340	21.8718	0.0378
soccer	400	26.9265	27.7452	0.8187	24.4010	24.8269	0.4259	22.4540	22.7126	0.2586
	550	27.1346	27.9975	0.8629	24.5102	24.9480	0.4378	22.5215	22.7834	0.2619
	750	27.2971	28.1936	0.8965	24.6000	25.0484	0.4484	22.5796	22.8456	0.2660
	1000	27.4182	28.3423	0.9241	24.6650	25.1185	0.4535	22.6096	22.8756	0.2660
	1400	27.5195	28.4684	0.9489	24.7120	25.1693	0.4573	22.6449	22.9133	0.2684
city	400	29.5218	30.2042	0.6824	25.8615	26.1259	0.2644	23.4997	23.6390	0.1393
	550	29.9296	30.6594	0.7298	26.0255	26.2913	0.2658	23.5915	23.7298	0.1383
	800	30.2197	31.0121	0.7924	26.1396	26.4141	0.2745	23.6513	23.7909	0.1396
	1200	30.4397	31.2779	0.8382	26.2233	26.5034	0.2801	23.6955	23.8382	0.1427
	1900	30.6161	31.4933	0.8772	26.2874	26.5717	0.2843	23.7320	23.8752	0.1432

即使存在丢包的情况下本文方法依然优于文献[2]的方法。上述结果验证了本文方法的有效性。

5 结束语

本文提出了一种基于预测式错误恢复机制的多描述视频编码方法。该方法首先对原始视频序列进行奇偶帧分离得到主信息，并根据预测式错误恢复机制得到重建的奇偶数帧，而后结合真实的奇偶数帧在编码模式选择模块生成冗余信息。考虑到编码效率问题，本文设计了 3 种冗余信息产生模式，模

式 1 情况下，块的重建质量很高，因而不需要传送任何冗余信息；模式 2 情况下，块的重建质量差强人意，因而需要传送当前块的运动矢量；模式 3 情况下，块的重建质量很差，因而需要传送对应块的运动矢量和残差到解码端。由于采用了分类思想，本文方法能在保证恢复质量的前提下提高压缩效率。实验结果表明，本文方法在丢包率较低，大分辨率和运动相对平滑的情况下鲁棒性更好。实验结果也验证了本文算法的有效性。

参考文献

- [1] Wang Y, Reibman A R, and Lin S. Multiple description coding for video delivery[J]. *Proceedings of the IEEE*, 2005, 93(1): 57-69.
- [2] Goyal V K. Multiple description coding : compression meets the network[J]. *IEEE Signal Processing Magazine*, 2001, 18(5): 74-93.
- [3] Vaishampayan V A. Design of multiple description scalar quantizers[J]. *IEEE Transactions on Information Theory*, 1993, 39(3): 821-834.
- [4] Vaishampayan V A and John S. Balanced interframe multiple description video compression[C]. Proceedings of the IEEE International Conference on Image Processing (ICIP'99), Kobe, Japan, 1999, 3: 812-816.
- [5] 郑义, 史萍. 一种基于小波变换的多描述视频编码方法[J]. 电视技术, 2012, 36(23): 43-46.
Zheng Yi and Shi Ping. Wavelet-based multiple description video coding[J]. *Video Engineering*, 2012, 36(23): 43-46.
- [6] 铨欣. 基于多描述标量量化的分布式视频编码[D]. [硕士学位论文], 西安电子科技大学, 2012.
Xuan Xin. Distributed video coding based on multiple description scalar quantization[D]. [Master dissertation], Xidian University, 2012.
- [7] Reibman A R, Jafarkhani H, and Wang Y. Multiple description coding for video using motion compensated prediction[C]. Proceedings of the IEEE International Conference on Image Processing (ICIP'99), Kobe, Japan, 1999, 3: 837-841.
- [8] 陈婧, 蔡灿辉, 丁润涛. X 树非平衡保护多描述编码[J]. 电子与信息学报, 2005, 27(12): 1973-1977.
Chen Jing, Cai Can-hui, and Ding Run-cao. X-tree unequal protected multiple description coding[J]. *Journal of Electronics & Information Technology*, 2005, 27(12): 1973-1977.
- [9] Conci N and De Natale F G B. Multiple description video coding using coefficients ordering and interpolation[J]. *Signal Processing: Image Communication*, 2007, 22(3): 252-265.
- [10] Wen X, Au O C, and Xu J. A novel multiple description video coding based on H.264/AVC video coding standard[C]. IEEE International Symposium on Circuits and Systems, (ISCAS'09), Taipei, 2009: 1237-1240.
- [11] Peraldo L, Baccaglioni E, and Magli E. Slice-level rate-distortion optimized multiple description coding for H.264/AVC[C]. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP'10), Dallas, USA, 2010: 2330-2333.
- [12] 赵春晖, 刘巍. 基于交织抽取与分块压缩感知策略的图像多描述编码方法[J]. 电子与信息学报, 2011, 33(2): 461-465.
Zhao Chun-hui and Liu Wei. Image multiple description coding method based on interleaving extraction and block compressive sensing strategy[J]. *Journal of Electronics & Information Technology*, 2011, 33(2): 461-465.
- [13] 孙文珠, 王洪玉, 钱大兴, 等. 一种适用于丢包信道的小波编码图像传输方案[J]. 电子与信息学报, 2012, 34(10): 2342-2347.
Sun Wen-zhu, Wang Hong-yu, Qian Da-xing, et al. A transmission scheme for wavelet coding images over packet erasure channels[J]. *Journal of Electronics & Information Technology*, 2012, 34(10): 2342-2347.
- [14] 张培君, 王淑慧, 周开伦, 等. 融合全色度 LZMA 与色度子采样 HEVC 的屏幕图像编码[J]. 电子与信息学报, 2013, 35(1): 196-202.
Zhang Pei-jun, Wang Shu-hui, Zhou Kai-lun, et al. Screen content coding by combined full-chroma LZMA and subsampled-chroma HEVC[J]. *Journal of Electronics & Information Technology*, 2013, 35(1): 196-202.
- [15] Lin Tao, Zhang Pei-jun, Wang Shu-hui, et al. 4:4:4 screen content coding using dual-coder mixed chroma-sampling-rate (DMC) techniques[C]. JCT-VC-Meeting, JCTVC-I0272, Geneva, 2012, 4: 1-11.
- 周文帅: 男, 1989 年生, 硕士生, 研究方向为图像视频编码。
白慧慧: 女, 1979 年生, 博士, 副教授, 研究方向为图像视频编码。
赵 耀: 男, 1967 年生, 教授, 博士生导师, 研究方向为图像视频编码、数字水印与数字取证、图像视频检索与内容理解等。