

基于帧内区域的细粒度可扩展视频编码最优码率分配算法

文振¹, 张基宏¹, 孙立峰²

(1. 深圳大学信息工程学院, 广东深圳 518060, 2. 清华大学计算机系, 北京 100084)

摘要: MPEG-4 FGS 是 MPEG-4 中精细可扩展性编码方法, 由于采用帧 8*8 交织的位平面编码结构, 无法支持帧内区域级的动态码率分配, 不适合基于内容的增强和流化传输, 针对 MPEG-4 FGS 这编码限制, 本文根据 MPEG-7 的内容描述标准, 提出了一种基于帧内区域的细粒度可扩展 A-FGS 编码, 实现了实时的帧内区域流化编码, 并就该编码提出了一种最优码率分配算法. 利用帧内区域 8*8 宏块高位平面非零系数很少的特点, 提出了动态码率分配算法, 使在中等码率范围内编码效率提高了 0.5dB 左右.

关键词: FGS; 可扩展编码; 码率分配

中图分类号: TN915.43 **文献标识码:** A **文章编号:** 0372-2112 (2007) 04-0761-05

An Optimal Bit Rate Allocation Algorithm for the Intra-Frame FGS Video Coding

WEN Zhen kun¹, ZHANG Ji hong¹, SUN Li feng²

(1. College of Information Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China;

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

Abstract: MPEG-4 FGS is a fine grain scalable coding method for MPEG-4. The 8*8 frame interlaced bit plane coding structure used by the MPEG-4 FGS coding does not support intra frame dynamic bit rate allocation, and therefore is incapable of supporting content based streaming transport. Regarding the limitation of the MPEG-4 FGS coding, we propose an intra frame fine granular scalable coding method (A-FGS) based on the MPEG-7 content description standard to achieve the goal of real time intra frame streaming coding. An optimal bit rate allocation algorithm is also provided for the A-FGS coding method. By using the characteristic of having few non zero coefficient in the intra frame 8*8 macro block high bit plane, we apply a dynamic bit rate allocation method to improve the coding efficiency in mid bit rate range by about 0.5dB.

Key words: fine granularity scalable; scalable coding; rate allocation

1 引言

MPEG-4 于 1998 年开始征集细粒度可扩展的视频编码方案, 用于流式服务中的编码标准 (MPEG-4 Streaming Profile), 它的基本思想是将视频编码成一个可以单独解码的基本层码流 (Base Layer Bit-stream) 和一个可以在任何地点截断的增强层码流 (Enhancement Layer Bit-stream), 其中基本层码流能够适应最低的网络带宽, 提供一个基本的图像质量, 增强层码流能够覆盖一定网络带宽变化的动态范围, 提供从基本层图像到无损失图像质量的细粒度增强. 经过 MPEG 测试和评比, 从编码效率和复杂性两方面的考虑, 从众多的方案中最终采纳了 W. Li^[1,2] 等人提出的基于 DCT 变换系数的 FGS (Fine Granularity Scalable, 精细可伸缩性) 方案^[3], 并于 2001 年成为 MPEG-4 Streaming Profile 中的视频编码 MPEG-4 FGS

标准^[4].

FGS 应用非常灵活, 利用 8*8 块位平面编码可以对图像中感兴趣的前景部分优先编码和传输, 把图像中感兴趣区域的数据块以放大系数或 DCT 系数加权重的方式上移若干个位平面, 从而满足人眼对不同频率成分的敏感度, 位平面编码示意图如图 1 所示.

FGS 虽然具有优良的细粒度可扩展特性, 但它有两个致命弱点, 一是编码效率偏低, 二是虽然能够很好支持镜头级和帧级动态码率分配, 但由于采用交织的位平面编码结构, 很难支持帧内区域级的动态码率分配. 在 MPEG-4 FGS 中, 选择性的区域增强是在编码端通过位平面上移来实现的^[5], 一旦编码完成, 这种区域增强在流化传输服务器端固定不变, 所以 FGS 并不支持动态区域增强的码率分配, 不适合基于内容的流化传输^[6-8]. 针对 MPEG-4 FGS 这种编码限制, 本文提出了

一种基于帧内区域的细粒度可扩展 A-FGS (Area Fine Granularity Scalable 区域精细可伸缩性) 编码, 该编码支持帧内区域级的动态码率分配, 符合视频流化 MPEG-7 和基于内容的视频流化 CAVS (Content-Aware Video Streaming) 的区域级码率分配描述标准^[9-11], 本文同时提出了 A-FGS 编码的一种最优动态码率分配算法。

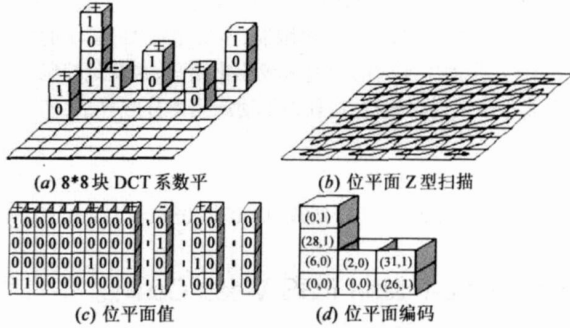


图1 8*8块位平面编码示意图

2 帧内区域的细粒度可扩展性视频编码

MPEG-4 FGS 编码和帧内区域细粒度可扩展 (A-FGS) 编码都是由基本层和增强层编码构成, 基本层用 MPEG-4 的单层编码, 增强层采用位平面编码, 它可以提供一个细粒度质量增强. 从解码端看, 解码器在接收到很少的数据 (基本层) 就能够解码, 随着所接收数据的增加, 解码图像质量逐步增强^[12,13]. MPEG-4 FGS 和 A-FGS 之间最大差别是增强层中的编码顺序不一样, 在 FGS 增强层编码中有两部分数据, 第一部分是帧头数据, 它包含了许多解码器必须用到的辅助信息, 如时间参考, 帧编码位平面个数等, 对于 FGS 和 A-FGS 是相同的; 第二部分是包含各个区域宏块编码数据的码流, FGS 是从一帧残差中最高位平面开始以整帧的位平面顺序进行 VLC 编码^[14-16], 而 A-FGS 是以帧中某一区域宏块按行扫描的顺序进行编码, 并且在区域宏块内是以该宏块位平面顺序从最高位到最低位进行编码. MPEG-4 FGS 编码和 A-FGS 编码不同主要有以下三个方面: (1) 编码结构; (2) 压缩编码码流的剪裁; (3) 区域宏块编码所用位平面层数的编码.

2.1 编码结构

FGS 是在同一个帧平面上以 8*8 块进行行扫描 VLC 编码, 如图 2(a) 所示, BP-0 是最高位平面, BP-N 是最低位平面, 图 3(a) 是 FGS 编码的码流结构, 每一个位平面的起始码字是一个 32 位比特同步码字, 它是字节对齐的, 表示一个位平面编码的开始, 而每个块的起始位并不是字节对齐的. A-FGS 是以 8*8 区域宏块为单位进行的 VLC 编码, 在每个区域宏块内部是从最高位平面到最低位平面进行的. 图 2(b) 是 A-FGS 的编码顺序示意图, 图 3(b) 是压缩编码后得到的码流结构, A-

FGS 每个区域宏块的起始位都是字节对齐的, 比 A-FGS 编码平均多用 4 个比特位, 主要是为了在流化传输服务器端快速分配码率和码流, 而在一个区域宏块内每个位平面起始位并不是字节对齐, 区域宏块的数据都是串行地组织在一起进行编码和传输.

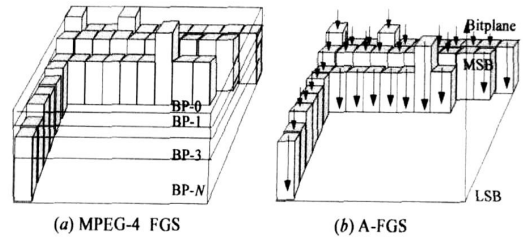


图2 两种 FGS 编码顺序比较

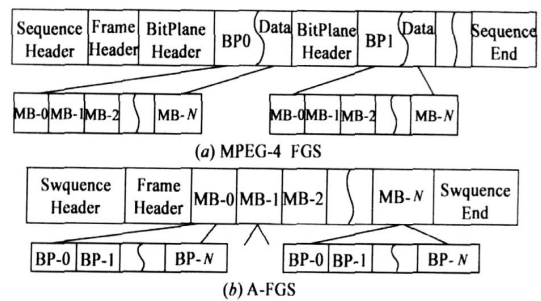


图3 压缩编码的码流结构比较

2.2 压缩编码码流的剪裁

MPEG-4 FGS 中, 码流由于它编码的特殊结构, 所以可以在码流的任意位置截断, 如果某一帧的编码码流在其中间某一宏块位置被截断, 解码器会由于 VLC 解码出错或者遇到下一帧帧头同步字而在这一帧截断位置附近停止解码, 根据帧头同步起始码字直接跳过这一帧到下一帧对应的区域位置处继续解码. 所以在 MPEG-4 FGS 解码的一帧图像中, 一般上半部分质量要比下半部分好.

在 A-FGS 中, 每个区域宏块的截断点是在每个位平面结束处, 并不是在区域宏块内部任意位置, 由于一个区域宏块的一个位平面编码并不需要非常多比特数, 尤其是在较高的位平面编码中出现没有编码信息时. 为了在解码端指示每个区域宏块的截断位置, 在区域宏块每层位平面编码结束处加了一位额外附加信息位, 这个标志位如果为 1 就表示结束, 解码器可以根据这一标志位的值就能知道当前区域宏块解码是否结束, 如果结束它将开始下一个区域宏块解码. 在视频流化传输服务器剪裁码流时, 只要将区域宏块所要截断的位平面处的标志位置为 1, 解码将丢弃该区域宏块剩余数据, 并与下一个区域宏块的解码数据连接起来.

2.3 区域宏块编码所用位平面层数的编码

在 MPEG-4 FGS 编码中, 位平面层数是以一个 5 位固定长度码字来表示的, 每个 8*8 块都有一个变长的

“CBP”码字来指示最高位平面(MSB)是否已经到达, 如果值为1就表示最高位平面编码还没有开始. 由于残差分布都趋近于高斯分布, 最高的前几个位平面非零系数很少, 所以 CBP 编码采用 VLC 编码, 而在后面的位平面中就采用每个块 1 比特的固定长度编码. 解码器根据这一帧的最高位平面层数和每个块 CBP 信息, 就可以知道帧内每个块的最高位平面层数.

在 A-FGS 编码中, 每个区域宏块可能具有不同的位平面层数, 从不同的位平面开始编码. 一帧最高位平面层数与区域宏块最高位平面层数的差采用 VLC 编码, 如果差值超过 2, 那么就输出一个“111”码字作为 Escape, 然后再输出一个 2 比特的固定长度编码表示超过 2 的数值, 为防止小概率使用过长码字, 因此采用固定长度编码.

3 A-FGS 码流剪裁中的最优码率分配

如果采用 A-FGS 编码, 并在增强区域内选择平均码率分配, 那么在一个很宽的解码范围内, 它要比相同码率下采用平均码率分配的 MPEG-4 FGS 要低 2dB 甚至更多, 其主要原因是 A-FGS 的平均码率分配并没有考虑到区域宏块残差的大小, 因为它是给每个区域宏块分配相同码率, 而 FGS 是从这一帧最高位平面开始按照位平面顺序从高到低进行码率分配, 所以对于残差大的区域块所得到的比特数就比残差小区域块要多. 从上面分析看, 尽管 A-FGS 平均码率分配非常简单, 但这种码率分配方法效率太低, 如果没有高效的码率分配方法, 虽然 A-FGS 有很好的可扩展特性, 但平均码率分配将损失掉由于在增强区域多分配码率所带来的增益, 限制了 A-FGS 在实际系统中的应用.

本文根据 JPEG-2000^[3] 中最优码率分配方法, 提出了 A-FGS 编码码流剪裁中的最优码率分配算法. A-FGS 编码的最优码率分配将根据每个区域宏块的率失真函数来进行码率分配, 是一个受约束的最优化问题, 最优码率分配目标是在一个给定码率下, 使解码图像的失真度达到最低.

3.1 最优码率分配描述

定义 R_i 表示第 i 个区域宏块分配的比特数, D_i 表示这个区域宏块在该比特数下对应的失真, 优化的码率分配问题就可以表示为:

$$\min(D_1(R_1) + D_2(R_2) + \dots + D_N(R_N)), \quad \text{s.t.} \quad \sum_{i=1}^N R_i = R_c \quad (1)$$

其中 N 是区域宏块总数, R_c 是分配给这一帧的总比特数. 利用 Lagrange 乘法, 可以将这样的优化问题转化为无约束的最优化问题, 引入 Lagrange 乘子 λ 优化的目标函数可以表示为:

$$J = \sum_i D_i(R_i) - \lambda \left(\sum_i R_i - R_c \right) \quad (2)$$

所要求的最优码率分配就是求解函数(2)的极值点, 由 $\partial J / \partial R_i = 0$, 得:

$$dD_i / dR_i = \lambda \quad (3)$$

如果能求解 λ , 那么就可以根据式(3)得到 R_i , 而由 $\partial J / \partial \lambda = 0$, 可以得到:

$$f(\lambda) = 0 \quad (4)$$

$f(\lambda)$ 由式(5)定义:

$$f(\lambda) = \sum_i R_i(\lambda) - R_c \quad (5)$$

如果 $R_i(\lambda)$ 是单调函数, 那么 Newton 迭代搜索算法就可以通过下面的迭代计算进行数值求解. 设 λ_k 为第 k 步计算得到参数值, 那么 $f(\lambda)$ 在 λ_k 附近行为可以用一阶近似来表示:

$$f(\lambda) \approx f(\lambda_k) + \left[\sum_{i=1}^N dR_i(\lambda_k) / d\lambda \right] (\lambda - \lambda_k) \quad (6)$$

将式(6)代入式(4)就可以解得 λ_{k+1} :

$$\lambda_{k+1} = \lambda_k - f(\lambda_k) / f'(\lambda_k) \quad (7)$$

$$\begin{aligned} f'(\lambda) &= \sum_i \frac{dR_i}{d\lambda} \\ &= \sum_i \frac{dR_i}{d(dD_i/dR_i)} = \sum_i \frac{1}{d(dD_i/dR_i)/dR_i} \\ &= \sum_i \frac{1}{d^2 D_i / dR_i^2} \end{aligned} \quad (8)$$

反复进行以上的迭代计算, 直到 $f(\lambda)$ 的绝对值小于某个设定阈值, 那么最优解 λ 也就求得, 每个区域宏块分配得到的比特数可以用下式表示:

$$R_i = R_i(\lambda), \quad i = 0, \dots, N$$

前面提到用 Newton 迭代计算求解的条件是 $R_i(\lambda)$ 是单调递增函数, 即每个区域宏块的可能截断点应该构成一个凸曲线, 否则需要从截断点集合中去除那些奇异点, 所以必须在码率分配之前对每个区域宏块截断点集合进行凸函数检查. 设集合 $TSet_i = \{1, 2, \dots, n\}$ 表示第 i 个区域宏块最初所有截断点集合, n 是所有可能截断点的总个数, p 是最新通过检查的截断点序号, k 是当前要检查的截断点序号.

3.2 截断点检查

Step 1: Set $p = 1$

Step 2: For $k = 2$ to n

If (k belongs to $TSet_i$) Then

Set $\Delta R_i^k = R_i^k - R_i^p$, $\Delta D_i^k = D_i^k - D_i^p$, $S_i^k = \Delta D_i^k /$

ΔR_i^k ;

If ($k \neq 2$ And $S_i^k < S_i^p$) Then

Remove p from $TSet_i$, goto Step 1.

Else

$p = k$;

End If

End If

End For

为了实现快速的最优码率分配,在 A-FGS 编码时将许多有关码率和失真数据记录下来,其中失真用 SSE(Sum of Square Error) 来度量的,并且在编码结束后对所有区域宏块截断点进行凸函数检查,将复合要求的截断点保存在辅助信息文件中,这样在码率分配时只要求解方程(4),然后根据每个区域宏块比特数对码流进行剪裁并装配.实现的最优码率分配算法中采用了快速 Newton 搜索方法求解(4),定义 k 表示当前的迭代步数, f_k 表示由(5)定义函数 $f(\lambda_k)$ 的值, $MaxNum$ 是预先设定的最大迭代步数, $MinDiff$ 是预先设定的码率分配误差阈值,如果迭代步数超过 $MaxNum$ 或者当前码率误差小于 $MinDiff$,那么就停止计算过程.

3.3 求解 λ 快速算法

Step 1: Initialization: $k = 0, \lambda_k = \lambda_0, f_k = f(\lambda_0), f'_k = f'(\lambda_0)$.

Step 2: Calculate λ_{temp} according to (6-7), then Calculate f_{temp} according to (6-5).

$$\lambda_{temp} = \lambda_k - f(\lambda_k) / f'(\lambda_k); f_{temp} = \sum_i R_i(\lambda_{temp})$$

- R_c ;

Step 3: If(Sign(f_i) and Sign(f_{temp}) is different) Then

$$\lambda_{k+1} = (\lambda_k + \lambda_{temp}) / 2; f_{k+1} = f(\lambda_{k+1});$$

$$\text{Else } \lambda_{k+1} = \lambda_{temp}; f_{k+1} = f_{temp};$$

End If

$$k = k + 1;$$

Step 4: If(($k < MaxNum$) and ($abs(f_k) > MinDiff$))

Then goto Step 2;

Else stop.

End If

Sign(f_k) 表示符号函数, 定义为

$$\text{Sign}(v) = \begin{cases} 1, & \text{if } v \geq 0 \\ -1, & \text{if } v < 0 \end{cases}$$

尽管最优码率分配比 MPEG-4 FGS 平均码率分配要复杂,但是由于采用了快速算法,在 P42.8 计算机上已经可以做到 16 倍的实时,因此 A-FGS 的最优码率分配对于有实时要求的视频流化传输应用是完全可行的.

4 实验结果

利用 MPEG-4 FGS 参考软件实现了本文提出的基于区域宏块的细粒度可扩展编码(A-FGS),实验用了 8 个 MPEG-4 标准测试视频序列在一定码率范围内进行解码后性能对比, CIF 分辨率的序列解码范围从 128

kbps 到 1 Mbps, 编码帧率是每秒 10 帧,序列长度是 10 秒,由于是测试 P 帧性能,所以除了第一帧是 I 帧外,其它都是 P 帧编码.编码时 CIF 序列的基本层编码码率是 128 kbps,在基本层都采用 TM5 码率扩展方法.

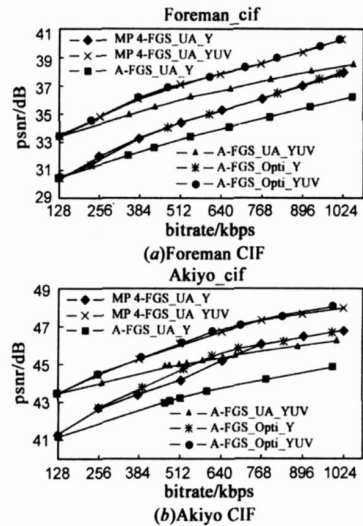


图 4 三种编码方法性能比较 (1) MPEG-4 FGS 平均码率分配, (2) A-FGS 平均码率分配, (3) A-FGS 最优码率分配

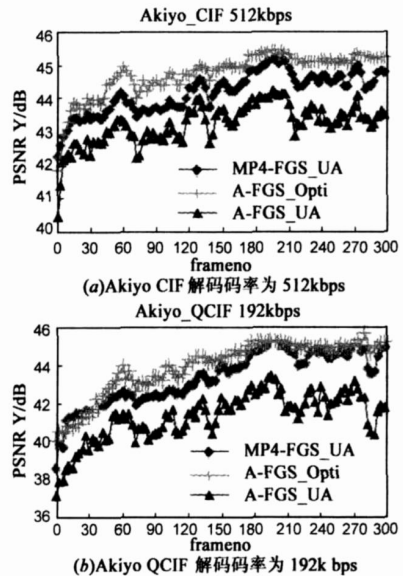


图 5 每帧 PSNR 性能比较

图 4 显示了 FGS、A-FGS 和 A-FGS 最优码率分配三种编码方法在不同解码码率下 PSNR 曲线,从图中可以知道,平均码率分配 A-FGS 比 FGS 差许多,其中 Y 分量的 PSNR 损失了 2.0dB 甚至更多,但是采用最优码率分配的 A-FGS 和 FGS 相当,在中等码率范围内还要好 1dB 左右.图 5 是 Akiyo CIF 和 Akiyo QCIF 两个序列在 512 kbps 和 192 kbps 解码码率下的每帧 Y 分量的 PSNR 比较,和其它两种方法相比,最优码率分配的 A-FGS 具有最好的编码效率.

5 结论

针对 MPEG-4 FGS 无法实现帧内区域动态增强的限制, 本文根据 MPEG-7 内容描述标准, 提出了一种基于帧内区域的细粒度可扩展 A-FGS (Area Fine Granularity Scalable 区域精细可伸缩性) 编码, 实现了实时的帧内区域流化编码, 并就该编码提出了一种最优码率分配算法. 利用帧内区域 8×8 宏块高位平面非零系数很少的特点, 采用动态码率分配方法, 使在中等码率范围内编码效率比 MPEG-4 FGS 提高了 0.5dB 左右.

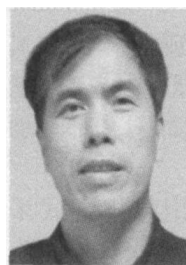
参考文献:

- [1] W Li. Streaming video profile in MPEG-4[J]. IEEE trans on Circuits and Systems for Video Technology, special issue on streaming video, 2001; 11(3): 301~ 317.
- [2] ISO/IEC TR 21000 1: 2001, Information Technology- Multimedia framework (MPEG 21) - Part 1: Vision[S]. Technologies and Strategy, ISO/IEC JTC 1/ SC29/ WG11/ N4040, Singapore: 2000.
- [3] MPEG. JPEG 2000 Part1 Final Committee Draft Version 1.0, ISO/IEC JTC 1/ SC29 WG1[S]. ISO/IEC JTC 1/ SC 29/ WG 1 (ITU- T SG8) N1646R, 2000. 195- 200.
- [4] 刘丹, 王相海. 基于小波子带统计特性的图像量化及质量可分级编码研究[J]. 计算机科学, 2006, 33(9): 211- 214. Liu Dan, Wang Xiang Hai. The research of the quantization scheme based on wavelet subband statistical characteristic and PSNR scalable image coding[J]. Computer Science, 2006, 33(9): 211- 214. (in Chinese)
- [5] M Schaar, Y T Lin. Content based selective enhancement for streaming video[J]. IEEE Transactions on Multimedia, March 2001. 977- 980.
- [6] ITU- T SG16/Q. 15, VCEG M73. Improved SP Frame Encoding[S]. 2001.
- [7] 张方, 吴成柯, 肖嵩, 王养利. 一种基于 MPEG-4 的感兴趣区域视频编码新方法[J]. 电子学报, 2005, 33(4): 634- 637. Zhang Fang, Wu Cheng ke, Xiao Song, Wang Yang li. A new regions of interest video coding method based on MPEG 4[J]. Acta Electronica Sinica, 2005, 33(4): 634- 637. (in Chinese)
- [8] 陈强, 卓力, 沈兰. 基于感兴趣区的 MPEG-4 FGS 增强层码率分配算法[J]. 电子与信息学报, 2005, 27(3): 402- 406. Chen Qiang, Zhuo Li, Shen Lan. ROF based rate allocation algorithm for MPEG-4 FGS[S]. Journal Of Electronics & Information Technology, 2005, 27(3): 402- 406. (in Chinese)
- [9] Y W He, R Yan, F Wu, S Li. H. 26L-based fine granularity

scalable video coding [A]. ISCAS 2002[C]. Phoenix, USA, 2002, 4: 548- 551.

- [10] PEG Homepage[OL]. <http://www.nist.gov/mpeg/>.
- [11] Wiegand T, et al. Overview of the H. 264/AVC video coding standard[J]. IEEE transactions on circuits and systems for video technology, 2003, 13(7): 560- 576.
- [12] 乔保军, 石峰, 王一拙. 空域精细可扩展编码预处理器的设计和实现[J]. 计算机应用, 2006, 26(4): 911- 913. Qiao Bao Jun, Shi Feng, Wang Yi Zhuo. Design and implementation of preprocessor for spatial fine granularity scalable coding[J]. Journal of Computer Applications, 2006, 26(4): 911- 913. (in Chinese)
- [13] 马社祥, 刘铁根, 刘贵忠. 图像的空间可分级压缩编码及其码率分配[J]. 光电子·激光, 2005, 16(12): 1500- 1505. Ma She xiang, Liu Tie gen, Liu Gui zhong. Spatial scalable coding of images and optimal bit allocation[J]. Journal of Optoelectronics Laser, 2005, 16(12): 1500- 1505. (in Chinese)
- [14] 朱红, 吴成柯, 方勇. 基于 H. 264 的自适应选择增强 FGS 视频编码[J]. 电子学报, 2005, 33(12): 2204- 2208. Zhu Hong, Wu Cheng ke, Fang Yong. Adaptive selective enhancement FGS video coding based on H. 264[J]. Acta Electronica Sinica, 2005, 33(12): 2204- 2208. (in Chinese)
- [15] Ohm J R. Advances in scalable video coding[J]. IEEE, 2005, 93(1): 42- 56.
- [16] Schwarz H, et al. SNR scalable extension of H. 264/AVC [J]. Singapore: Singapore, In Proc of IEEE International Conference on Image Processing (ICIP), 2004. 113- 116.

作者简介:



文振 男, 1962年3月出生于广东信宜, 清华大学计算机科学与技术专业毕业, 工学硕士, 现为深圳大学信息工程学院计算机系副教授. 主要研究方向为流媒体编码、基于内容视频检索. E-mail: wenzk@szu.edu.cn



张基宏 男, 1964年6月出生于江苏海安, 东南大学无线电专业毕业, 工学博士, 现为深圳大学教授. 主要研究方向有图像编码、矢量量化和模糊逻辑.