

基于 H. 264 的自适应选择增强 FGS 视频编码

朱 红, 吴成柯, 方 勇

(西安电子科技大学 ISN 国家重点实验室, 陕西西安 710071)

摘 要: 提出了一种新的选择增强 FGS 编码方法, 基本层仍采用 H. 264 编码方式, 增强层直接利用基本层的模式判决结果对不同运动分割宏块进行不等程度的增强. H. 264 采用了树型结构运动补偿, 较大的宏块分割尺寸适用于平坦区, 而较小的宏块分割尺寸可以刻画精细的运动细节; 分析表明, 对于特定类型的图像序列, 精细的运动细节又较为集中于视觉感兴趣区 (ROI). 本文利用这一特点, 对较小分割尺寸的宏块进行位平面提升以保证精细运动细节区的优先编码, 使它们更容易被包括进截断码流, 从而在不预先确定 ROI 位置的条件下仍能有效提高 ROI 的重建质量.

关键词: H. 264; 树型结构运动补偿; 感兴趣区 (ROI); 精细可分级编码 (FGS); 选择增强

中图分类号: TN919.81 **文献标识码:** A **文章编号:** 0372-2112 (2005) 12 2204-05

Adaptive Selective Enhancement FGS Video Coding Based on H. 264

ZHU Hong, WU Cheng ke, FANG Yong

(National Key Lab. on ISN, Xidian University, Xi'an, Shaanxi 710071, China)

Abstract: A new method of selective enhancement for FGS is proposed, in which the base layer is encoded with an H. 264 compliant encoder, while the enhancement layer, by using the outcome of mode decisions within the base layer, is encoded after performing unequal enhancement on motion compensated sub blocks of varying partition size. H. 264 uses tree structured motion compensation featuring a flexible hierarchy of partition sizes. In general, a large partition size is appropriate for homogeneous areas of the frame and a small one may be beneficial for motion detailed areas. Analysis shows that, for image sequences with particular characteristics, Region of Interest (ROI) may contain a significant part of areas with high motion detail. This feature can be used in FGS coding in which bit planes of macroblocks with small partition sizes are bit-shifted up prior to encoding in order to give them a higher priority and a higher probability of being included in a truncated bit stream, hence maximizing the quality of the potential ROI in the decoded sequence without the need to determine coordinates of the ROI in advance.

Key words: H. 264; tree structured motion compensation; region of interest; fine granularity scalability; selective enhancement

1 引言

网络异质造成的带宽波动是 Internet 视频流传输 (video streaming) 面临的挑战之一. FGS (精细可分级编码)^[1] 将视频压缩成基本层和增强层两个子码流. 基本层采用混合编码技术, 可以单独解码并提供一个粗糙质量; 增强层采用位平面编码技术, 编 3 解码依赖于基本层, 用于提供更优质量. 增强层码流允许任意截断, 重建质量同收到并解码的位平面数成正比, 所以 FGS 可以实现连续的增强层速率控制, 具有一定的网络带宽适应能力. 为了在低带宽下提高增强层的质量, FGS 引入了频率加权 (Frequency Weighting, FW)^[2] 和选择增强 (Selective Enhancement, SE) 机制. 前者考虑了人类视觉系统对各种频率成分敏感性的差异, 对不同频率成分采用不同的优先权, 使重要频率成分的比特优先进入编码码流. 后者考虑了不同空间区域对视觉效果的不同, 对不同的空间区域采用不同的优先权, 使一帧中的重要区域 (如 ROI, 视觉感兴趣区) 的比特优先进入编码码流. 文献 [3] 选择人脸作为 ROI 进行选择增强. 文献 [4] 提出了一种 B FGS 编码方案, 与基于帧的 FGS 编码不同, B FGS 对每个宏块进行独立的增强层编码, 允许区域增强同编码过程分

离, 支持区域级动态码率分配和内容感知视频流传输. 编码或传输时的 ROI 分割是文献 [3, 4] 方案的关键和难点.

为了降低视频传输对带宽的需求, 还需要更加高效的视频编码技术. 最新的视频编码标准 H. 264^[5] 虽然仍采用了混合编码框架, 但它吸收了原有标准 (如 H. 263, MPEG-4) 的优点并全面应用了视频通信的最新理论和技术成果, 对混合编码的诸多环节进行了改进并引入了一些新算法, 显著提高了编码效率^[6]. 目前的 H. 264 标准并不支持 FGS 编码, 但对该项工具的需求已列入 H. 264 下一步工作计划. 文献 [7] 提出了一种基于帧间小波变换的 H. 264 可分级编码方案, 对低运动复杂度图像序列比较有效. 文献 [8] 则提出了一种基于 PFGS (渐进 FGS) 的 H. 264 (H. 264 草案) 可分级编码方案, 将 MPEG-4 PFGS 的编码框架直接移植到 H. 264, 增强层编码并未利用 H. 264 特有的编码机制. 本文提出了一种新的基于 H. 264 的选择增强 FGS (SE FGS) 编码方法, 基本层编码仍采用 H. 264, 而增强层直接利用基本层的模式判决结果, 对较小分割尺寸的宏块进行位平面提升, 以保证精细运动细节区的优先编码, 使它们更容易被包括进截断码流, 从而在不预先分割 ROI 的条件下仍能有效提高潜在 ROI 的解码质量.

2 H. 264 运动分割规律分析

图像序列有时空间细节较多, 有时大面积平坦, 有时运动剧烈, 有时运动缓慢, 但图像中的不同区域对视觉上的重要程度并不相同. 视觉重要区域通常称为 ROI, 优先编码 ROI 可以提高图像的主观质量. 采用 ROI 编码的前提关键是对象分割. MPEG-4 的 VOP(视频对象平面) 概念就是根据人眼感兴趣的一些特性如形状、运动、纹理等, 将图像序列每一帧中的场景, 看成是由不同的 VOP 组成, 而同一对象连续的 VOP 称为 VO(视频对象). VO 可以是图像序列中的人物或具体的景物, 例如电视新闻中的播音员或者电视剧中一辆奔驰的汽车, 也可以是计算机生成的二维、三维图形. 实践表明, 对于特定类型的图像序列, 在运动特性和 ROI 之间存在一定的关联, 也就是说, 运动细节往往比较集中于 ROI 区, 因此通过对图像运动细节的分析能在一定程度上揭示出 ROI 的位置. H. 264 的树形结构运动补偿机制为图像运动细节的分析提供了便利条件, 这就引出了本文基于宏块分割模式的 SE-FGS 编码方案的基本思想, 如图 1 所示.

H. 264 允许对宏块的亮度分量采用 4 种分割模式 ($16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8$) 进行运动补偿, 如果采用 8×8 分割模式, 则每个 8×8 分割块又可进一步细分为 4 种分割模式 ($8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4$) 进行更精确的运动补偿, 多种尺寸块的组合可以给出大量的可能的宏块分割模式, 这种将宏块分割成不同尺寸运动补偿子块的方法称为树形结构运动补偿. 在 H. 264 中, 帧间编码宏块的模式集合为 $C: \{SKIP, 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4, Intra4 \times 4, Intra16 \times 16\}$, SKIP 是一种特殊的 16×16 模式, 其运动矢量和 DCT 信息不需要编码, 而 $Intra4 \times 4$ 和 $Intra16 \times 16$ 为帧内预测编码模式. 宏块的每种尺寸的分割块都需要做运动估计, 并返回一个独立的运动矢量, 宏块分割模式和各个运动矢量都需要编码和传输, 因此宏块分割尺寸的选择对压缩性能具有重要的影响. H. 264 参考软件 [9] 采用了率失真优化模式判决算法, 对每个宏块选择一个最佳的分割模式. 该分割模式使得在满足一定失真的条件下宏块编码所需的比特最少. 由于这种灵活而细微的宏块划分更切合图像中实际运动物体的形状, 因此能有效提高运动补偿的精度, 进而提高编码效率. 一般地, 对图像的平坦区采用较大分割尺寸的块, 对图像的运动细节区采用较小分割尺寸的块.

为了揭示运动细节同 ROI 的联系, 采用如表 1 所示的测试序列, 均为 QCIF@10Hz, Mobile 取 32 帧, 其它测试序列各取 100 帧进行实验. 对于不同的测试序列, ROI 定义的方法也不尽相同, 为便于阐述本文方案思想, ROI 定义如下: Akiyo 属头部肩部, 将头部肩部定义为 ROI; Hall 属监控图像, 将人体区定义为 ROI; Foreman 的脸部定义为 ROI; Silent 属哑语图像, 脸部表情和手势比较重要, 所以将脸部和上肢一起定义为 ROI. 由于包含了较多的视频对象, 所以 Container、News 和 Mobile 的 ROI 定义比较困难, 不失方案的一般性, 本文分别将 Container 的油轮、News 的两个头部肩部以及两个舞蹈人体、Mobile 的日历定义为 ROI. 采用 H. 264 参考软件 JM 8.1a, 编码帧类型 IPPP, QP=34, 5 个帧间预测参考帧, 允许率失真优化模式判决. 图 2 给出

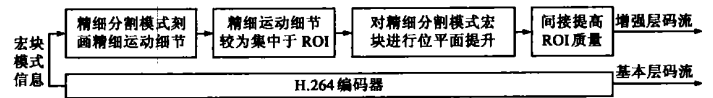


图 1 基于宏块分割模式的 SE-FGS 编码方案
了测试序列的一帧图像中宏块的分割模式. 从中可以直观看出, 对于 Akiyo、Container、Hall、News 和 Silent, ROI 中宏块的运动分割模式较为精细, 而对于 Foreman 和 Mobile, 各种宏块分割模式的分布都较为分散. 以下针对宏块模式同 ROI 的联系进行一些理论分析.

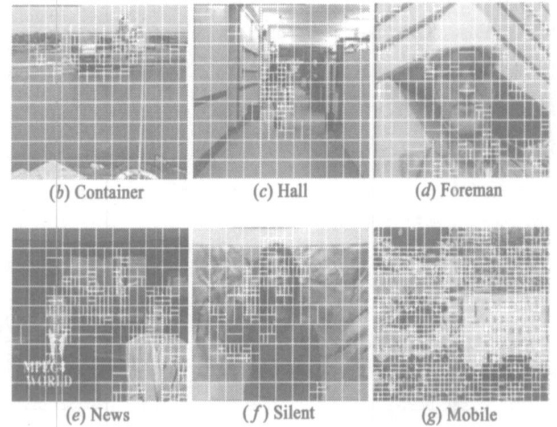


图 2 H.264 宏块的运动分割模式, (a)~(g) 分别为相应序列的第 27、78、22、3、62、7、9 帧编码图像

表 1 测试序列及其 ROI 定义

序列名称	序列特性	本文的 ROI 定义
Akiyo	A 类测试序列: 细节较少, 活动性较小	头部肩部
Container		油轮
Hall		人体
Foreman	B 类测试序列: 细节适中, 活动性较小; 或者细节较少, 活动性适中	脸部
News		头部肩部+ 舞蹈人体
Silent		脸部+ 上肢
Mobile	C 类测试序列: 细节较多, 活动性适中; 或者细节适中, 活动性较大	日历

按照运动分割模式精细程度的不同, 对集合 C 作如下划分: $C = C_a \cup C_b \cup C_c$, 各个子集为 $C_a = \{8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4\}$, $C_b = \{16 \times 8, 8 \times 16\}$, $C_c = \{SKIP, 16 \times 16, Intra4 \times 4, Intra16 \times 16\}$. 设 ROI 面积占整幅图像面积的比例为 $RatioS_f$, 下标 $1 \leq f \leq F$ 表示图像帧的序号. 设在第 f 帧中, C_a 和 C_b 模式宏块总数分别为 N_{a_f} 和 N_{b_f} , 而落在 ROI 中的 C_a 和 C_b 模式宏块的数目分别为 N'_{a_f} 和 N'_{b_f} , 则定义 $RatioN_{a_f} = N'_{a_f}/N_{a_f}$ 和 $RatioN_{b_f} = N'_{b_f}/N_{b_f}$, $RatioN_{a_f}$ 和 $RatioN_{b_f}$ 分别用于衡量 C_a 和 C_b 模式宏块集中于 ROI 的程度, 数值越大表示越集中. 定义 $El_{a_f} = RatioN_{a_f}/RatioS_f$ 和 $El_{b_f} = RatioN_{b_f}/RatioS_f$, El_{a_f} 和 El_{b_f} 的物理意义讨论如下: ① $RatioS_f$ 较小时, 对 ROI 所做的增强处理在时间或码率上的额外开销较小; ② $RatioN_{a_f}$ 或 $RatioN_{b_f}$ 较大时, C_a 或 C_b 模式宏块较为集中于 ROI, 即 C_a 或 C_b 模式宏块同

ROI 吻合程度较好; ③Ratio N_{af} 或 Ratio N_{bf} 越大, 同时 Ratio S_f 越小时, El_{af} 或 El_{bf} 的值越大, 表示越适合对 C_a 或 C_b 模式宏块进行选择增强处理; ④如果 $El_{af} > El_{bf}$, 则应优先对 C_a 模式宏块进行选择增强, 反之应优先对 C_b 模式宏块进行选择增强. 引入综合评价指标

$$GEI_a \text{ 和 } GEI_b: GEI_a = \frac{1}{F} \sum_{f=1}^F El_{af}, GEI_b = \frac{1}{F} \sum_{f=1}^F El_{bf}. GEI_a \text{ 或 } GEI_b \text{ 越大, 表示测试序列越}$$

适合采用“对 C_a 或 C_b 宏块选择增强以间接提高潜在 ROI 质量”的方法. 通过仿真实验, 可得 Akiyo、Container、Hall、Foreman、News、Silent、Mobile 的 Ratio S_f 平均值分别约为 0.15、0.13、0.15、0.21、0.19、0.17、0.15, 而各个测试序列的 El_{af} 、 El_{bf} 曲线及其指标如图 3 所示.

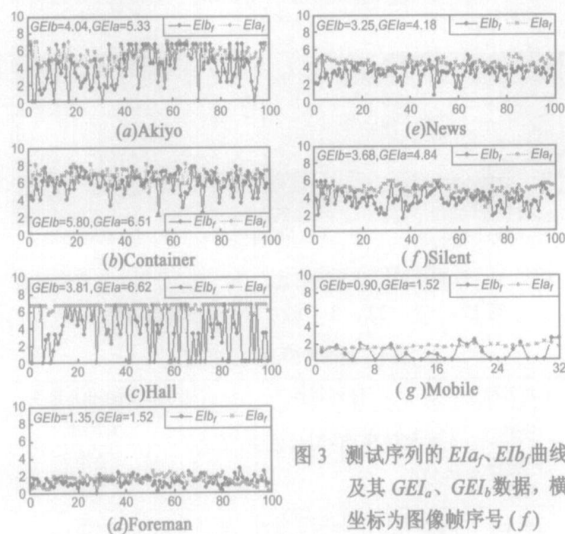


图 3 测试序列的 El_{af} 、 El_{bf} 曲线及其 GEI_a 、 GEI_b 数据, 横坐标为图像帧序号 (f)

考察 C_a 或 C_b 模式宏块同 ROI 的吻合程度, Foreman 和 Mobile 的 GEI_a 和 GEI_b 都明显低于其它序列的相应指标, 宏块模式同 ROI 的吻合程度较差, 因此并不适合采用本文的选择增强方案. 原因分析: Foreman 存在着较剧烈的全局运动, 使得 C_a 或 C_b 模式宏块在图像中的分布过于分散, 而不是比较集中于作为 ROI 的脸部, 这从图 2(d) 可直观看出; Mobile 中对象较多, 且运动比较复杂, 也使得 C_a 或 C_b 模式宏块在图像中的分布比较分散, 而不是相对集中在某个视频对象(例如日历)上, 这从图 2(g) 可直观看出. 而其它序列的 GEI_a 和 GEI_b 指标较大, $GEI_a \geq 4.18$, $GEI_b \geq 3.25$, C_a 或 C_b 模式宏块较为集中于 ROI 区, 这同图 2 的直观观察结果一致, 所以适合采用本文的选择增强方案. 可见, 本文方案比较适合具有如下特性的序列: 头肩部, 如 Akiyo; 背景固定或背景运动缓慢, 视频对象运动区域比较集中, 如 Container、Hall、News 和 Silent. 但不太适合具有如下特性的序列: 背景较复杂且存在剧烈全局运动, 如 Foreman; 视频对象较多且运动较复杂, 如 Mobile. 最后, 对所有测试序列都有 $GEI_a > GEI_b$, 表明总体上 C_a 模式宏块同 ROI 吻合得较好; 虽然在一小部分图像帧中, C_b 模式宏块同 ROI

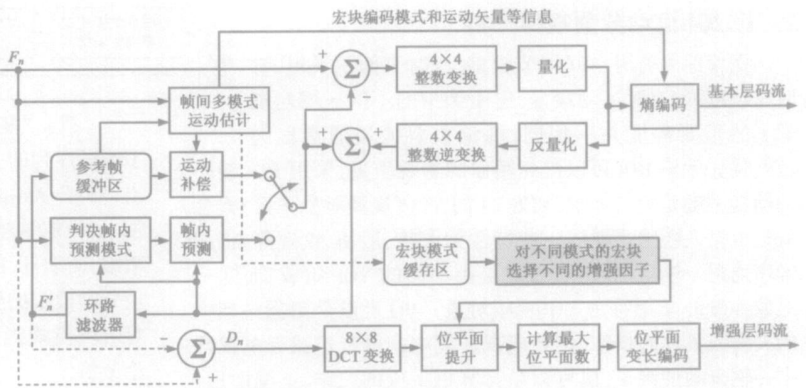


图 4 基于宏块模式的 SE-FGS 编码方案框图

的吻合性比 C_b 要好, 但从总体上讲, C_a 模式宏块的增强优先级应高于 C_b 模式宏块.

3 基于宏块模式的 SE-FGS 编码

通过对 H.264 运动分割规律的分析, 针对满足所述特性的序列, 可以总结出如下规律: ①两类运动分割模式子集 C_a 和 C_b 的宏块都比较集中于 ROI 区; ② C_a 宏块比 C_b 宏块更为集中于 ROI 区. 因此, 按一定优先级对 C_a 和 C_b 宏块进行选择增强可以间接提高 ROI 的质量. H.264 宏块的运动分割过程是自适应进行的, 因此 ROI 的增强也具有自适应性. 相比之下, 常规的图像分割方案有诸多不足之处. 实际上, 对包含自然场景的很多视频来说, ROI 的严格定义有时非常困难, 而且图像中还可能存在多个 ROI. 虽然基于对象分割的视频编码可以提高编码质量, 但对对象分割仍有一定的实现难度. 尽管 MPEG-4 框架已经制定, 但视频对象分割仍被认为是一个富有挑战性的难题^[10], 因此利用 H.264 宏块模式信息判别 ROI 的策略具有一定优势. 为此, 我们提出了如图 4 所示的 SE-FGS 编码方案, 核心思想是对不同模式的宏块选择不同的增强因子. 基本层仍采用 H.264 编码, 而增强层直接利用基本层的模式判决结果, 优先对较小分割尺寸的宏块进行位平面提升, 以保证精细运动细节区的优先编码, 从而在不必先分割 ROI 的条件下仍能有效提高 ROI 的解码质量.

增强层编码过程如下, 设 F_n 、 F'_n 和 D_n 分别表示输入信号、经环路滤波的基本层重建信号和残差, 则有 $D_n = F_n - F'_n$. 对 D_n 进行 8×8 块 DCT 变换和量化. 选择 8×8 块 DCT 变换的原因是, 它比 4×4 整数变换具有更好的能量集中特性^[11], 而且在位平面编码时可以直接使用 MPEG-4 FGS 的 VLC(变长码)码表. 与其它可分级编码方法不同, FGS 编码器对残差 DCT 量化系数采用了位平面编码方式. 同游程编码方式的最大区别在于, 位平面编码方式将每个量化的 DCT 系数看成是由若干比特构成的一个二进制整数, 而不是一个十进制整数. 因此采用位平面编码的任意编码比特都能重建 DCT 系数. 采用位平面编码, FGS 可获得对增强层的连续比特控制, 这是因为增强层比特流能在任意地方截断. 对每个块的 64 个 DCT 系数进行 Zigzag 扫描形成一维列表, 并计算残差帧的最大位平面数 N , $N = \max_k \{ \text{se}(k) + \lceil \log_2 \left\{ \max_{i,j} \{ c(i,j,k) \} \right\} + 1 \} \}$, 式中 $c(i,j,k)$ 为 DCT 量化系数, k, j, i 分别为宏块、块和系数的序号, $1 \leq i$

$\leq 64, 1 \leq j \leq 4, 1 \leq k \leq 99$ (QCIF 分辨率), $se(k)$ 为第 k 个宏块的选择增强因子, $0 \leq se(k) \leq 4$. $se(k)$ 对应于编码码流的语法元素 `shifted_bit_planes`; 若允许选择增强, 则语法元素 `fgs_vop_selective_enhancement_enable` 设置为 1, 否则设置为 0. 经选择增强后的新系数值 $c'(i, j, k) = 2^{se(k)} \cdot c(i, j, k)$, $c'(i, j, k)$ 等效为将 $c(i, j, k)$ 的全部比特位整体上移 $se(k)$ 个位平面. 从第 N 个位平面开始按降序进行, 依次对队列的 DCT 系数绝对值进行位平面编码. 每个位平面包含 64 比特, 各个比特依次取自所属 DCT 系数相同重要位置的比特位. 每个位平面都可以用一系列 (RUN, EOP) 符号表示. 其中, RUN 表示 1 之前的连 0 个数, 是否还有 1 存在则用 EOP 标识, 如果位平面全部为 0, 就用一个特殊符号 ALL-ZERO 标识. 最后, 将上述符号同 DCT 系数的符号位(对每个系数的符号位只编码一次)一起进行 VLC 编码后输出增强层比特流. 基本层的宏块模式判决结果被依次放于宏块模式缓存区, 以便于后续对不同模式的宏块进行不等程度的选择增强. 选择增强因子设置方法如下: 对于 C_a 模式宏块, $se(k) = 4$; 对于 C_b 模式宏块, $se(k) = 2$; 对于 C_c 模式宏块, $se(k) = 0$, 即不做增强处理.

4 实验结果和分析

本文对基于 H. 264 宏块模式的 SE-FGS 编码方案进行了仿真实验. 测试序列及其 ROI 的定义仍同表 1, 基本层编码采用 H. 264 参考软件 JM8. 1a, 图像分辨率 QCIF@10Hz, 编码帧类型 IPPP, Mobile 取 32 帧, 其它测试序列各取 100 帧进行实验, P 帧预测编码时使用 5 个参考帧, 允许率失真优化宏块模式判决, 基本层量化参数和熵编码方法分别为 QP=34 和 CABAC; 增强层编码时仅对 DCT 系数进行限幅 (clipping) 处理, 不作进一步量化. 并同基于 H. 264 的频率加权 FGS 方案 (FW-FGS) 进行了性能比较. 使用重建图像序列亮度分量的平均峰值信噪比 (Y-PSNR) 作为视频质量的客观评价标准. 本文并未同其它选择增强方案进行对比实验, 这是因为一般的 FGS 选择增强方案如文献 [3] 都需要预先知道 ROI 的位置, 该前提条件对本文方案来说则并不需要, 本文方案只是利用了宏块运动分割模式同潜在 ROI 的空间吻合关系, 通过对较小分割模式的宏块进行选择增强间接提高潜在 ROI 的重建质量.

频率加权采用一个 8×8 FW 矩阵对 DCT 系数进行加权, FW 矩阵每个元素的数值代表相应系数的位平面提升数, 提升一个位平面等价于 DCT 系数乘以权值 2. 通常情况下, 低频 DCT 系数比高频 DCT 系数更重要, 因此权值较大. 在编码码流中, 使用语法元素 `frequency_weighting_enable` 标识是否采用频率加权, 若是, 则使用另一个语法元素 `load_frequency_weighting_matrix` 标识 FW 矩阵是否已包含在码流中, 若是, 则使用 `frequency_weighting_matrix` 存储经 Zigzag 扫描排序的矩阵数据. 由于频率加权常常是优先对低频系数进行位平面提升, 因此位平面的统计特性可能发生变化. 为此, FGS 另外引入了两组 VLC 码表, VLC 码表的选择取决于直流 (DC) 系数的权值, 若 DC 的权值为 0 或 1, 就使用原始 VLC 码表; 若是 2 或 3, 就使用中端移位 (mediant shift) VLC 码表; 若是 4~7, 就使用高端移位 (high shift) VLC 码表. 对于 FW 矩阵的选择, 本文采用

文献 [12] 的方法, 对细节和活动性均较少的测试序列 Akiyo、Container 和 Hall 采用 FW1 矩阵, 而对细节和活动性均较多的其它测试序列使用 FW2 矩阵: $FW_1 = [4, 4, 4, 3, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 0, \dots]$, $FW_2 = [4, 3, 3, 3, 3, 3, 2, 2, 2, 2, 1, 1, 1, 1, 1, 0, \dots]$, 均采用 Zigzag 扫描方式. 表 2 给出了采用不同方案时测试序列的重建质量比较结果.

表 2 采用不同方案时测试序列的重建质量比较

测试序列	码率 (kbps)		平均重建质量 (Y-PSNR, dB)			
			FW-FGS		本文方案	
	基本层	截断码流	整幅图像	ROI	整幅图像	ROI
Akiyo	9.66	16	33.97	34.03	33.67	34.38
		32	34.38	34.55	33.98	35.41
Container	9.60	16	32.82	32.89	32.5	33.26
		32	33.29	33.45	32.86	34.47
Hall	13.23	16	33.80	33.84	33.49	34.17
		32	34.19	34.29	33.82	35.20
Foreman	31.86	32	32.36	32.38	32.12	32.49
		64	32.98	33.04	32.71	33.23
News	19.98	32	32.85	32.93	32.51	33.22
		64	33.45	33.59	33.07	34.40
Silent	20.55	32	32.25	32.32	31.92	32.59
		64	32.82	32.98	32.45	33.73
Mobile	134.79	200	28.61	28.58	28.48	28.81

对表 2 中实验数据的分析如下: ①针对整幅图像的平均重建质量, FW-FGS 优于本文方案, 对 A、B 和 C 类测试序列, 分别提高了约 0.36 dB、0.32 dB 和 0.11 dB; ②采用 FW-FGS 时, ROI 的平均质量同整幅图像的平均质量相差不大, 对 A、B 和 C 类测试序列, 差值分别约为 0.1 dB、0.09 dB 和 -0.03 dB; ③针对 ROI 的平均重建质量, 本文方案优于 FW-FGS, 对 Akiyo、Container、Hall、News 和 Silent 平均可提高 0.6 dB, 而对 Foreman 和 Mobile 平均可提高 0.18 dB; ④采用本文方案, ROI 的平均重建质量高于整幅图像的平均重建质量, 对 Akiyo、Container、Hall、News 和 Silent 平均可提高 1.06 dB, 而对 Foreman 和 Mobile 平均可提高 0.41 dB. 从中可以得出 3 点结论: ①采用 FW-FGS 比较有利于提高整幅图像的视频质量; ②本文方案则比较有利于提高 ROI 的视频质量; ③对测试序列 Foreman 和 Mobile, 本文方案的优势并不明显, 而对于其它测试序列, 本文的优势比较明显, 这说明本文方案具有一定的适用范围, 这同前文第 2 节中的分析结论是一致的. 第 ①、②结论的原因分析如下, FW-FGS 同 SE-FGS 是两种不同的 FGS 增强机制; FW-FGS 对所有图像区域的处理是平等的, 主要是突出 DCT 低频系数的重要性, 从而在截断码流中尽可能地保留图像的低频信息, 以保证整幅图像的重建质量, 同时可能降低细节区的重建质量; 而 SE-FGS 则是一种面向区域的增强方法, 主要是突出某块区域的重要性, 从而在截断码流中尽可能地保留该块区域的编码比特, 保证了该块区域的重建质量, 同时降低了其它区域的重建质量; 本文方案是一种 SE-FGS 方案, 由于对于特定的测试序列, 宏块的运动分割模式同 ROI 具有很强的吻合性, 因此采用本文方案可以有效提高 ROI 质量. 图 5 给出了采用不同方案时, 测试序列 Akiyo、Hall 和 News 的一帧重建图像.

从图 5 可以看出,采用本文方案时 ROI 的重建图像比较清晰,但背景区的重建图像比较模糊,这同上述客观质量分析的结果是相吻合的。

5 结论

对于特定类型的图像序列,精细的运动细节较为集中于视觉 ROI 区.本文利用这一特点,提出了一种新的 SE-FGS 编码方案,通过对较小分割尺寸的宏块进行位平面提升,保证了精细运动细节区的优先编码,使它们更容易被包括进截断码流,从而在不预先确定 ROI 位置的条件下仍能有效提高 ROI 的重建质量.由于基本层编码采用了率失真优化模式判决,宏块的运动分割模式随图像运动的变化而变化,而本文的增强过程直接利用了基本层的模式结果,因此本文的 SE-FGS 编码过程还具有对图像运动内容的自适应性。

参考文献:

- [1] Weiping Li. Overview of fine granularity scalability in MPEG-4 video standard[J]. IEEE transactions on circuits and systems for video technology, 2001, 11(3) : 301- 317.
- [2] Weiping Li. Frequency weighting for FGS[R]. ISO/IEC JTC1/ SC29/WG11, MPEG99/M5589, Dec, 1999.
- [3] Schaar M, et al. Content based selective enhancement for streaming video[C]. Thessaloniki, Greece: In Proc of IEEE International Conference on Image Processing (ICIP), 2001. 977- 980.
- [4] Yuwen He, et al. Block-based fine granularity scalable video coding for content aware streaming [C]. Rochester, New York: In Proc of IEEE International Conference on Image Processing (ICIP), USA, 2002. 22- 25.
- [5] Recommendation ITU-T H. 264. Advanced video coding for generic audiovisual services[S].
- [6] 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.
- [7] Schwarz H, et al. SNR scalable extension of H. 264/AVC [C]. Singapore: Singapore, In Proc of IEEE International Conference on Image Processing (ICIP), 2004. 113- 116.
- [8] Yuwen He, et al. H. 264 based fine granularity scalable video coding [C]. Scottsdale, USA: In Proc of IEEE International

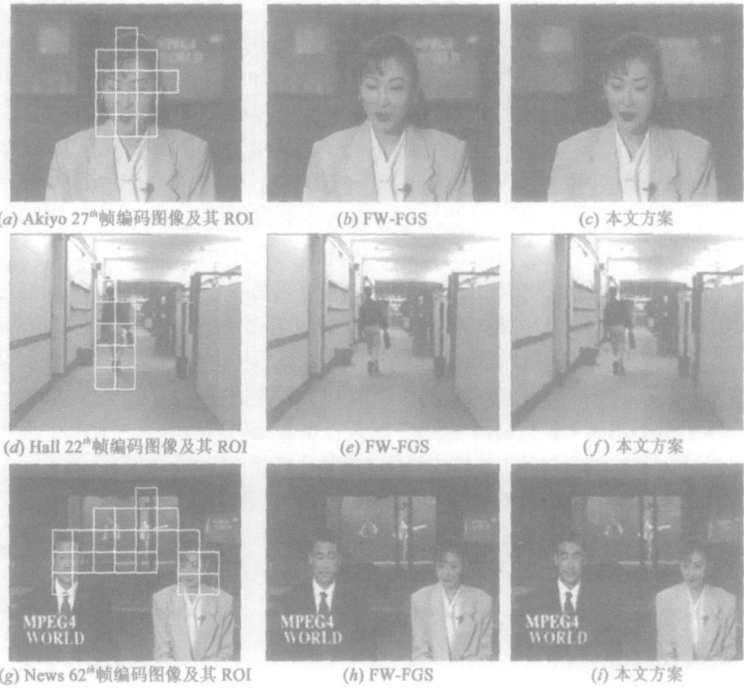


图 5 采用不同方案时的重建图像主观质量比较

Symposium on Circuits and Systems (ISCAS), 2002. 548- 551.

- [9] Download online: ftp://ftp.intc.files.org/jvt/experts/reference_software[Z].
- [10] ISO/IEC 14496-2 (MPEG-4 Video), Information technology generic coding of audiovisual objects-Part 2: Visual[S]. Dec. 1999.
- [11] Kalluri R, et al. Fine granular scalability for H. 264 based video streaming[C]. Auckland, New Zealand: In International Conference on Consumer Electronics (ICCE), 2002. 346- 347.
- [12] Sharon P and Schaar M. Scene dependent frequency weighting for subjective quality improvement of MPEG-4 fine granularity-scalability[C]. New York, USA: In International Conference on Image Processing (ICIP), Rochester, 2002, 697- 700.

作者简介:



朱红男, 1971年10月生于河北涿鹿, 1997年4月在中国空间技术研究院获得通信与电子系统硕士学位, 现为西安电子科技大学通信工程学院博士研究生, 专业为信号与信息处理, 主要研究方向为视频编码与无线/Internet视频传输。