

新的 H.264 降尺寸转码快速算法

吕卓逸, 贾克斌

(北京工业大学 电子信息与控制工程学院, 北京 100124)

摘要: 结合图像纹理特征和宏块运动特性, 提出一种新的 H.264 降尺寸转码算法, 充分利用输入编码码流中的编码信息来提高转码效率. 通过运动边缘检测划分出运动区域与背景区域, 针对不同区域内的宏块, 基于其运动特性, 采取不同的帧间模式选择策略. 提出一种基于感兴趣区域的运动矢量合成和修正算法. 利用运动矢量方差判断宏块运动的多样性, 基于宏块的多样性和重要性自适应地合成和修正运动矢量. 实验证明, 此算法具有较强的鲁棒性和广泛的适用性, 在转码后的视频质量和转码速度方面都取得了很好的效果.

关键词: H.264 降尺寸转码; 帧间模式选择; 运动矢量

中图分类号: TP 391

文献标志码: A

文章编号: 0254-0037(2011)12-1786-07

随着视频编码技术和网络技术的发展, 基于网络的多媒体通信已被广泛应用^[1]. 由于网络种类的多样性、终端设备性能的差异以及视频编码标准的不同, 因而提高视频流在不同应用环境之间的兼容性十分重要^[2]. 为解决网络带宽与接收终端显示材料的限制, 原始的高分辨率视频码流往往需要被转换为适合接收端显示的低分辨率码流之后进行传输, 由此引入了降低空间分辨率的视频转码技术.

同类视频转码, 尤其是基于 H.264 编码标准的视频转码成为近年来国内外研究的热点^[1]. H.264/AVC 是一个面向未来 IP 和无线环境下的视频压缩标准, 在视频压缩效率方面比目前所有的视频压缩标准都有显著提高^[3]. 由于运动估计和运动补偿需要占用 H.264 编码中 60% 的编码时间^[3], 如果在转码中重新对宏块进行运动矢量预测和修正, 计算量太大, 难以满足实时性要求^[4]. 为此, 利用输入码流中的信息来减少计算量, 是提高转码速度的有效途径. 在帧间模式选择算法的研究中, 文献[5]利用残差宏块和边缘块亮度值; 文献[6]利用 DCT 域内低频的 3 个变换系数来提前判决帧间预测模式; 文献[7]提出一种运动矢量重估计的体系结构 AMVR-DIM, 利用运动矢量的多样性和宏块的重要性来控制目标宏块的重编码模式. 解决宏块运动不一致造成的运动预测不精确问题, 需要新的运动矢量合成和修正方法来提高转码质量.

运动估计的准确性和高效性直接影响着视频转码后视频流的码率和解码后图像的质量, 使得模式选择和运动矢量合成成为视频转码中需要解决的关键问题. 本文利用图像纹理特征和宏块运动特性, 提出了一种基于感兴趣区域的帧间模式快速选择和运动矢量合成算法. 结合对宏块多样性和重要性的度量, 对运动矢量进行修正. 这种新的转码算法, 对各种运动程度的视频序列自适应地进行帧间模式选择和运动矢量重建, 既提高了转码后的视频质量, 又减少了运算量. 较之现有算法, 具有更高的鲁棒性和更广泛的适用性.

1 帧间模式快速选择算法

1.1 H.264 帧间模式选择

H.264 编码标准支持 7 种不同块大小的帧间编码模式. 每个宏块可按照 16×16 、 16×8 、 8×16 、 $P8 \times 8$

收稿日期: 2009-05-17.

基金项目: 国家自然科学基金资助项目(60672050); 北京市自然科学基金资助项目(4062005).

作者简介: 吕卓逸(1983—), 女, 北京人, 博士研究生.

进行分割, 如果选择 $P8 \times 8$ 模式, 还可按照 8×8 、 8×4 、 4×8 、 4×4 进行亚分割.

帧间编码采用直接拷贝模式 SKIP, 帧内预测模式 I4MB 和 I16MB. H. 264 的帧间模式选择算法流程如下:

1) 对当前宏块的模式 16×16 、 16×8 、 8×16 进行运动估计, 并计算率失真代价 (rate distortion optimal, RDO) 为

$$J(m) = D(m) + \lambda \cdot R(m) \quad (1)$$

式中, D 是宏块的失真; R 是需要编码的信息码率; λ 是用来调节失真与码率权重的系数; J 是率失真代价.

2) 对每个 $P8 \times 8$ 块的模式 8×8 、 8×4 、 4×8 、 4×4 进行运动估计, 并计算 J 值, 选择 J 值最小的模式作为 $P8 \times 8$ 块的最佳模式;

3) 重复执行 2), 直到计算完 4 个 $P8 \times 8$ 模块, 并得出 $P8 \times 8$ 模式的 J 值;

4) 为 SKIP 模式计算运动矢量和 J 值;

5) 计算帧内编码模式 I4MB 和 I16MB 的 J 值;

6) 从 16×16 、 16×8 、 8×16 、 $P8 \times 8$ 、SKIP、I4MB 和 I16MB 中选择 J 值最小的模式作为帧间宏块编码模式.

因需要对每种可能的宏块模式进行运动估计和计算 J , 计算量急剧增加, 编码速度下降.

1.2 快速帧间模式选择算法

为减少不必要的宏块模式搜索和 J 的计算, 提高转码速度, 基于感兴趣区域和宏块的运动特性, 本文提出了一种快速帧间模式选择算法来提前确定候选模式.

视频序列中分布均匀的区域, 例如背景区域, 通常采用大块分割模式 (16×16 、 16×8 、 8×16), 而分布不均匀的区域和边缘部分采用小块分割模式 (8×8 、 8×4 、 4×8 、 4×4). 同时, 视频序列中相邻图像间普遍存在着时间相关性, 即相邻图像的大部分区域的变化比较缓慢, 如背景区域中 SKIP 模式占相当大的比重, 而处于较长时间静止不动或者变化缓慢区域内的宏块通常采用大块分割模式. 本文分别对视频序列图像空间和时间两方面的特征进行研究, 在分析并提取其特征的基础上提出降尺寸转码新方法.

图像纹理特征的传统分析方法, 例如小波变换、统计学方法等虽可准确地获得运动边缘, 但是运算量较大. 利用输入编码码流中的残差值进行运动边缘检测, 划分出运动区域与背景区域, 是一种更为简单有效的方法. 实验结果表明, 简单地通过统计不为零的 DCT 变换系数便可以较准确地表现出宏块的空间特性. 图 1 所示是 CIF 格式的 Mother-daughter 序列和 Paris 序列的残差图像. 从中可以看出, 位于运动目标边缘以及运动目标内部的宏块由于变化较剧烈, 运动矢量预测不够准确, 产生的残差值较大. 相反, 位于背景区域的宏块运动比较缓慢, 经过运动估计后得到的最佳匹配块存在的残差相对较小. 因此, 通过对预测残差的计算可以判断宏块是否位于运动目标内部, 即通常被称作为感兴趣的区域 (region of interest, ROI).

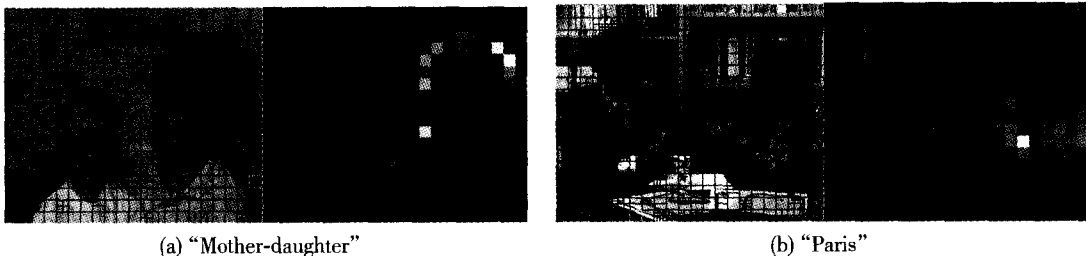


图 1 残差图像

Fig. 1 Residual image

首先, 获得运动边缘图像序列 (edge map), 将其定义为宏块重要性 I_n^t

$$I_n^k = \frac{\sum_{i=0}^3 C_i}{4} > T_l \quad (2)$$

式中, i 为原始尺寸视频的 4 个宏块; C_i 为宏块 i 不为零的 DCT 变换系数; T_l 为设定阈值.

按照划分出的感兴趣区域, 将图像分为前景和背景, 采用不同的帧间模式选择和运动矢量重建策略.

利用对宏块运动程度的计算来提取宏块运动特性. 在 H.264 编码中, 大块分割模式通常用于时间特性较稳定的区域, 而变化较剧烈的区域采用小块分割模式来减小运动估计后产生的残差, 此时运动矢量的值较大. 因此, 宏块的运动程度 A_n^k 定义为

$$A_n^k = \frac{1}{4 \times m} \sum_{i=0}^m (|v_{iv}| + |v_{ih}|) \quad (3)$$

式中, v_{ih} 代表宏块 MB_n^k 在水平方向上的运动矢量; v_{iv} 代表宏块 MB_n^k 在垂直方向上的运动矢量.

基于以上分析, 利用图像纹理特征得到感兴趣区域, 结合宏块的运动特性, 新的帧间模式选择算法流程如图 2 所示.

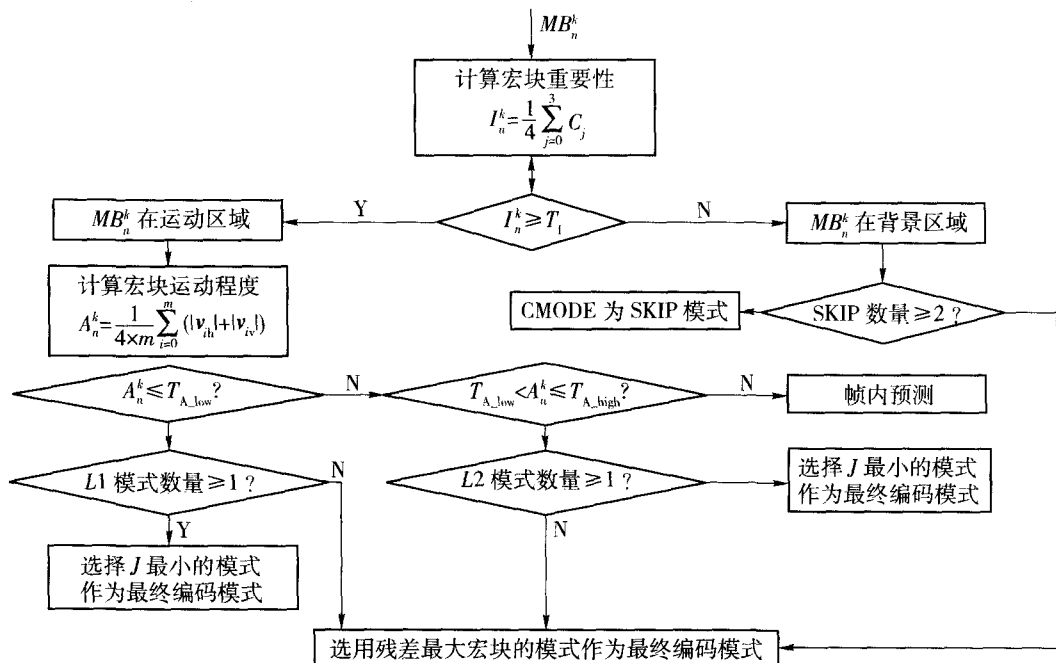


图 2 帧间模式选择算法流程

Fig. 2 Flowchart of inter mode decision

假设第 n 帧中的宏块 k 为 MB_n^k , 定义其编码模式为 M_n^k . 如果 M_n^k 为 $P8 \times 8$ 模式, 则每个 8×8 子块的亚分割模式为 S_n^k . 算法的具体实现步骤是:

1) 利用 DCT 变换系数划分出运动前景和背景区域之后, 计算位于运动目标区域内宏块的运动程度 A_n^k 为

$$A_n^k = \frac{1}{4 \times m} \sum_{i=0}^m (|v_{iv}| + |v_{ih}|) \quad (4)$$

2) 如果 $A_n^k \leq T_{A_low}$, 则从 4 个宏块的候选模式中选择大块模式 ($L1: 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8$) 计算其 J , 进行 5); 如果没有大块模式则进行 7);

3) 如果 $T_{A_low} < A_n^k \leq T_{A_high}$, 则从 4 个宏块的候选模式中选择小块模式 ($L2: 8 \times 8, 4 \times 8, 8 \times 4, 4 \times 4$) 计算其 J 后进行 5); 如果没有小块模式则进行 7);

4) 如果 $A_n^k > T_{A_high}$, 则采用帧内预测模式;

5) 选择 J 最小的模式作为最终编码模式;

- 6) 对于位于背景区域的宏块, 如果 4 个候选模式中有 2 个以上为 SKIP 模式, 则选择 SKIP 模式, 否则进行 7);
- 7) 选用残差最大宏块的模式作为最终编码模式.

2 运动矢量的合成与修正

鉴于运动估计和运动补偿在视频转码中的重要性, 本章首先介绍目前常用的几种运动矢量重建算法, 在分析这些算法特点的基础上, 提出一种基于宏块运动特征的运动矢量合成算法和基于感兴趣区域的运动矢量修正方法.

本文针对视频格式从 CIF 缩小为 QCIF 的转码进行研究, 如图 3 所示. 现有算法中, 平均加权法 (align-to-average weighting, AAW) 的计算简单, 但当宏块多样性较大时转码后图像质量下降较严重. 另 2 种运动矢量合成方法, 最佳加权法 (align-to-best weighting, ABW) 和最严重加权法 (align-to worst weighting, AWW) 基于对宏块预测残差值的计算, 分别选择不为零 DCT 变换系数最少和最多宏块的运动矢量. 上述算法统称为单一运动矢量重构法 (1-MV), 适用于宏块运动一致的情况. 针对运动不完全一致的情况 (如图 4 所示), 文献 [7] 提出了四运动矢量重构法 (4-MV) —— 自适应运动矢量重建算法 (adaptive motion vector re-sampling method, AMVR).

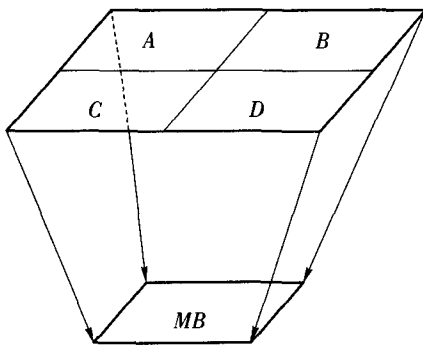


图 3 视频降尺寸 1/4

Fig. 3 Downscaling by 1/4

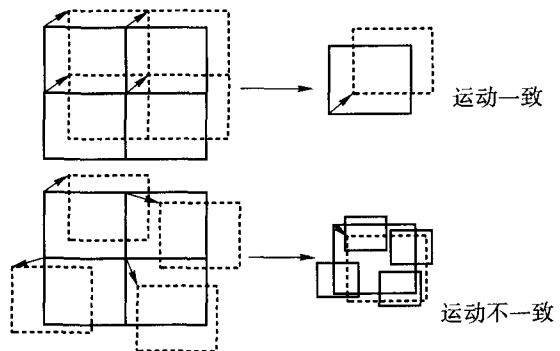


图 4 降尺寸视频运动矢量

Fig. 4 Down-sizing motion vector by two

文献 [6] 介绍了一种基于宏块多样性和重要性的运动矢量重建算法, 本文对其进行了研究和改进. 文献 [6] 中将宏块的多样性和重要性定义为:

$$D_j = \begin{cases} \text{high,} & \frac{d_j}{\frac{1}{N} \sum_{j=0}^{N-1} d_j} \geq 1 \\ \text{low,} & \frac{d_j}{\frac{1}{N} \sum_{j=0}^{N-1} d_j} < 1 \\ \text{zero,} & d_j = 0 \end{cases}, \quad d_j = \sum_{i=0}^3 (v_{ih} - \bar{v}_h + v_{iv} - \bar{v}_v) \quad (5)$$

$$I_j = \frac{m_j}{\frac{1}{N} \sum_{j=0}^{N-1} m_j} \geq 1, \quad m_j = \sum_{i=0}^3 C_i \quad (6)$$

式中, \bar{v}_h, \bar{v}_v 代表平均运动矢量; C_i 代表宏块不为零的 DCT 变换系数. 针对多样性和重要性的高低, 采用不同的动矢量重建和修正算法, 如表 1 所示.

表 1 运动矢量重建算法
Table 1 Motion vector re-construction

多样性	重要性	运动矢量重建
高	高	帧内预测 (Intra-refresh)
高	低	4-运动矢量重构法 (AMVR)
低	高	单一运动矢量重构法 (AWW) 大范围运动搜索 (±3 像素)
低	低	单一运动矢量重构法 (AWW) 小范围运动搜索 (±1 像素)
0 (运动一致)	高/低	单一运动矢量重构法 (AAW) 无运动搜索

与计算单个运动矢量偏离图像整体运动矢量均值的方法相比,利用方差更能准确地反映出 4 个运动矢量的差异. 因此,本文将宏块多样性定义为

$$V_j = \sum_{i=0}^3 v_i, \quad v_i = (v_{ih} - \bar{v}_h)^2 + (v_{iv} - \bar{v}_v)^2 \quad (7)$$

新的运动矢量重建和修正算法的具体实现步骤如下:

- 1) 利用式(2)计算宏块 MB_n^k 的重要性 I_n^k , 并划分出感兴趣区域;
- 2) 计算位于运动目标边缘及其内部宏块的多样性 D_n^k ;
- 3) 如果 D_n^k 大于设定阈值 T_D , 则表明原始视频 4 个宏块的运动不一致, 采用 4-MV 模式 (AMVR); 否则采用 1-MV 模式 (AWW);
- 4) 对于位于背景区域的宏块, 采用 1-MV 模式 (AAW).

在运动矢量修正方面, 为了避免直接进行 ±2 像素的运动搜索占用大量编码时间, 同时, 如果在运动搜索前可以获得理想的运动矢量, 就无需再进行修正, 基于对宏块重要性计算获得的感兴趣区域, 本文提出自适应的运动矢量修正算法. 即通过对宏块空间分布的预判, 对位于运动目标边缘和内部的宏块进行 ±2 像素的搜索; 对于位于背景区域的宏块不进行运动搜索.

基于感兴趣区域和宏块多样性、重要性计算的运动矢量重建和修正算法的流程如图 5 所示.

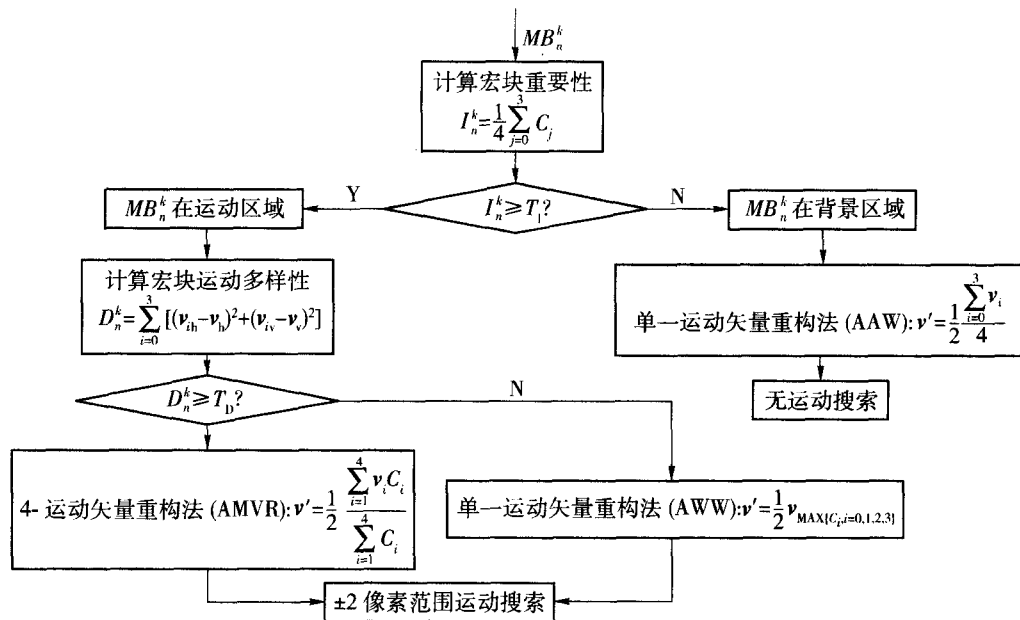


图 5 运动矢量重建流程

Fig. 5 Flowchart of ROIDI motion vector re-construction scheme

3 实验结果与分析

实验采用 H. 264/AVC 的测试模型 JM12. 2, 实验平台使用 Visual C + + 6. 0. 为了验证本文提出算法的适用性和高效性, 选取 4 个具有不同运动程度的 CIF 格式 (352 × 288) 的测试序列, 分别为 Silent, Foreman, Mother-daughter 和 Paris 序列, 将其缩小为 QCIF 格式序列 (176 × 144). 同时, 为了有效地说明算法的转码效果, 每个序列选取前 50 帧进行转码, 编码方式采用首帧为帧内编码 (intra-frame, I 帧), 后续帧为帧间编码 (inter-frame, P 帧).

在实验中, 本文提出的基于图像纹理特征和宏块运动特征的帧间模式选择和运动矢量合成修正算法 (proposed method), 与全搜索算法和 AWW 法在编码质量和编码效率两方面进行转码效果的比较.

由图 6 ~ 8 所示的 PSNR-Bitrate 曲线可看出, 与全搜索算法相比, AWW 的编码质量下降较严重. 简单地通过 ±2 像素范围的运动搜索可使编码质量有所提升, 但是对于运动变化较快和运动细节丰富的视频 (如 foreman 序列) 编码效果不理想. 而本文提出的算法, 对于不同运动程度的视频序列在 PSNR 方面比 AWW 算法有平均 0. 1 ~ 0. 2 dB 的提高. 因此本算法具有更广的适用性和更好的编码效果.

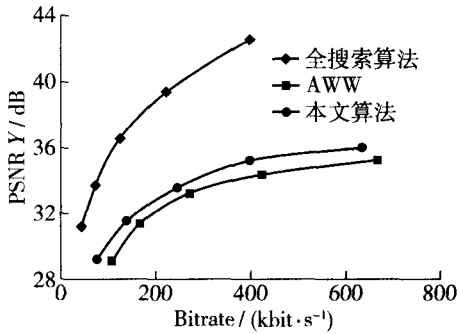


图 6 “Foreman” 序列
Fig. 6 Foreman

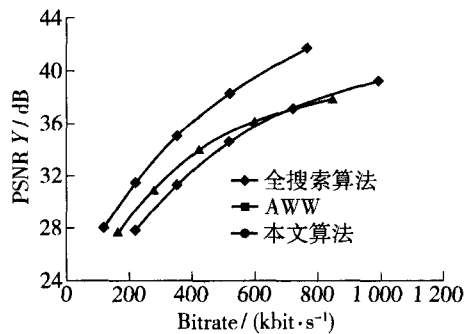


图 7 “Paris” 序列
Fig. 7 Paris

图 9 所示为运算复杂度的比较, 可以看出, 全搜索算法的运算复杂度非常大, 因此降低转码中运动估计占用的运算时间是解决转码实时性问题的关键. AWW 算法可将编码速度平均提高 97. 20%, 本文提出的算法编码速度平均提高 96. 73%, 而转码后的图像质量要比使用 AWW 算法平均提高 0. 34 dB. 综合考虑编码质量和编码速度双方面的因素, 本文提出的算法在有效解决降尺寸转码实时性问题的同时, 与现有算法相比, 取得了更优的编码效果和更广泛的适用性.

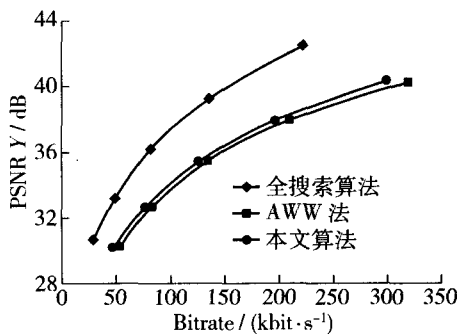


图 8 “Silent” 序列
Fig. 8 Silent

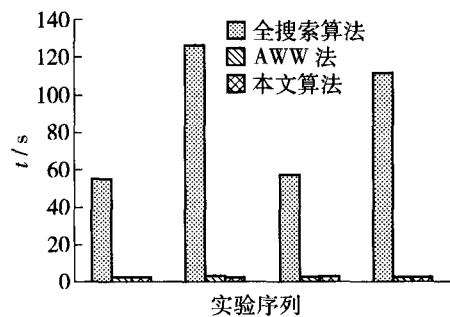


图 9 运算复杂度
Fig. 9 Computational complexity

4 结束语

重点分析了 H. 264 降尺寸转码中帧间模式选择和运动矢量合成修正算法, 针对目前已有算法的优缺点展开研究. 基于对图像空间纹理特征和宏块时间特性的分析, 提出了一种基于感兴趣区域和宏块运动特征的快速帧间模式选择算法以及改进的基于宏块多样性和重要性计算的运动矢量合成和修正算法. 对于不同运动程度的视频序列, 按照划分出的运动区域和背景区域, 自适应地进行帧间模式选择和运动矢量重建. 实验结果表明, 本文提出的算法与现有算法相比, 有效地提高了转码后的视频质量和转码速度, 具有更广泛的适用性.

参考文献:

- [1] WANG Jiao, YANG En-hui, YU Xiang. An efficient motion estimation method for H. 264-based video transcoding with spatial resolution conversion[C]//2007 IEEE International Conference on Multimedia and Expo. Beijing, 2007: 444-447.
- [2] 张宏波, 彭玉华, 刘兆广, 等. 一种改进的降尺寸视频转码算法[J]. 计算机工程, 2008, 34(16): 241-243.
ZHANG Hong-bo, PENG Yu-hua, LIU Zhao-guang, et al. Improved algorithm for down-sizing video transcoding[J]. Computer Engineering, 2008, 34(16): 241-243. (in Chinese)
- [3] 谢晶, 贾克斌. 一种基于二维直方图的 H. 264/AVC 快速帧内预测判决算法[J]. 电子与信息学报, 2005, 27(7): 1053-1057.
XIE Jing, JIA Ke-bin. A fast intra-frame prediction algorithm based on two-dimensional histogram for H. 264/AVC[J]. Journal of Electronics & Information Technology, 2005, 27(7): 1053-1057. (in Chinese)
- [4] FU Chang-hong, CHAN Yui-lan, SIU Wan-chi. Efficient motion estimation in H. 264 reverse transcoding[C]//IEEE International Conference on Image Processing. San Antonio, 2007: 317-320.
- [5] WU D, PAN F, LIM K P, et al. Fast inter mode decision in H. 264/AVC video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(6): 953-958.
- [6] TANG Qiang, NASIOPOULOS P, WARD R. Fast block size prediction for MPEG-2 to H. 264/AVC transcoding[C]//IEEE International Conference on Acoustics, Speech, and Signal Processing. Las Vegas, 2008: 1029-1032.
- [7] FUNG Kai-Tat, SIU Wan-Chi. Diversity and importance measures for video downscaling[C]//IEEE International Conference on Acoustics, Speech, and Signal Processing. Las Vegas, 2005: 1061-1064.

Novel Algorithm of Fast Video-downsizing Transcoding of H. 264

LÜ Zhuo-yi, JIA Ke-bin

(School of Electronic Information & Control Engineering, Beijing University of Technology, Beijing 100124, China)

Abstract: A novel fast variable block size selection method for an inter-MB in a video-downsizing transcoder for H. 264 is proposed, which uses the spatial homogeneity based on moving edge detection and temporal stationary characteristics by measuring activity magnitude. Based on the homogeneity and stationary of macro-blocks, only a few inter-modes are operated with the RDO process. In addition, adaptive motion vector re-construction and refinement scheme are proposed based on region of interest (ROI), besides of using variance and DCT coefficient for the definition of diversity and importance, respectively. Experimental results show that the new method is more robust and achieves good performance in both quality and complexity.

Key words: video down-sizing transcoder; inter mode decision; region of interest; motion vector re-construction; motion vector refinement

(责任编辑 张苏苹)