

H. 264 中基于时空相关性的帧间模式选择算法

何宜宝 毕笃彦 马时平 许悦雷

(空军工程大学工程学院信号与信息处理实验室, 西安 710038)

摘要 帧间块模式选择是 H. 264 视频编码器中比较耗时的模块之一, 为降低 H. 264 视频编码器帧间模式选择的复杂度, 提出了一种利用宏块的时空相关性实现帧间模式快速选择的方法。该算法通过对视频序列连续帧的时间相关性和相邻宏块的空间相关性进行统计、分析, 首先得到宏块模式转换的统计分布规律, 并根据此规律得到宏块的初始候选模式; 然后利用基于主宏块的宏块跟踪, 找到最相关宏块, 再由率失真门限筛选出最佳编码模式。仿真实验结果表明, 该快速模式选择算法可以在保证视频编码质量的同时, 很大程度上降低了编码器的计算复杂度, 平均节省编码时间约 64%, 平均信噪比下降约 0.31 dB。

关键词 H. 264 帧间模式选择 时空相关性 率失真优化 宏块跟踪

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2009)12-2456-07

An Inter-mode Decision Algorithm Based on Spatiotemporal Correlation in H. 264

HE Yi-bao, BI Du-yan, MA Shi-ping, XU Yue-lei

(Signal & Information Processing Laboratory, Engineering College, Air Force Engineering University, Xi'an 710038)

Abstract Inter-mode decision is one of the time-consuming units in H. 264 encoder. To reduce the complexity of the block mode decision process, an inter-mode decision algorithm based on spatiotemporal correlation of macro-block was proposed. First, the property of temporal correlation in successive frames and spatial correlation in neighboring macro-block is analyzed, and statistical results for the MB mode change is summarized. According to these characteristics, original candidate modes are chosen. Second, with macro-block tracking of the main macro-block, the most correlated macro-block is found. Then the best coding mode is chosen with a threshold of rate distortion. The results of simulation test show that the proposed fast mode decision strategy can highly improve the efficiency of real time H. 264 encoder while maintaining video coding quality. The method can save the encoding time up to 64% on average with -0.31dB performance degradation.

Keywords H. 264, Inter-mode decision, spatiotemporal correlation, rate distortion optimization, macro-block tracking

1 引言

高质量的视频导致所要处理的视频数据量激增, 而在有限的带宽内传输大量的数据又促进了高压缩率的视频压缩标准的发展。

H. 264^[1]是由 ISO/IEC(国际标准化组织/国际

电子技术委员会)的 MPEG(联合运动图像专家组)和 ITU-T(国际电信联盟)联合成立的 JVT(联合视频专家组)制定的一个新的国际视频压缩标准, 同时也是 MPEG-4 的第 10 部分(advance video coding)。

H. 264 与以往的视频编码标准相比具有更好的压缩性能和面向网络传输的“友好性”。H. 264 之

基金项目: 国家高技术研究发展计划(863)项目(2007AA701206)

收稿日期: 2008-07-26; 改回日期: 2008-11-06

第一作者简介: 何宜宝(1985 ~), 男, 空军工程大学工程学院信号与信息处理专业硕士研究生。主要研究方向为视频压缩编码算法。
E-mail: gudujiangboboo@yahoo.com.cn

所以具有优秀的压缩性能是由于多种新技术所产生效果的积累所致^[2-3]。这些新技术包括多模式帧内预测、可变尺寸块的运动补偿、多参考帧的运动补偿、4×4 整数变换、高效率的熵编码、环路滤波等。然而 H. 264 高效的压缩性能是以计算复杂度的大大增加为代价的,其中 H. 264 的运动估计和模式选择是编码器中最为复杂的编码模块之一。本文就 H. 264 中的帧间模式选择算法进行了较为深入的分析,提出了一种基于时空相关性的帧间编码模式选择算法。实验结果表明,本文提出的快速算法可以在保证视频编码质量的同时,有效地减少搜索模式数,从而可大幅度降低编码时间和计算量。

2 H. 264 标准中帧间多模式编码原理

H. 264 标准的帧间预测可以采用多模式运动估计、1/4 像素精度运动补偿和可供选择的多参考帧来对当前编码帧进行精确预测以提高压缩效率。在帧间预测编码时,每个亮度宏块可划分成形状不等的区域,作为运动描述区域,其划分方法如图 1 所示:当进行宏块级运动估计时,有 Skip, 16×16, 16×8, 8×16 等 4 种模式;当进行亚宏块级运动估计时,可以进一步划分成 8×8, 8×4, 4×8 和 4×4 等 4 种模式。H. 264 对于帧间宏块定义了 7 种编码模式,包括 Skip 模式、4 种帧间模式和 2 种帧内预测模式 (Intra16×16, Intra4×4)。

H. 264 标准中针对每个宏块由于可以提供多种编码模式,所以在编码端对编码的宏块就需要选取最佳的编码模式,以取得最佳的压缩效果。在 JM

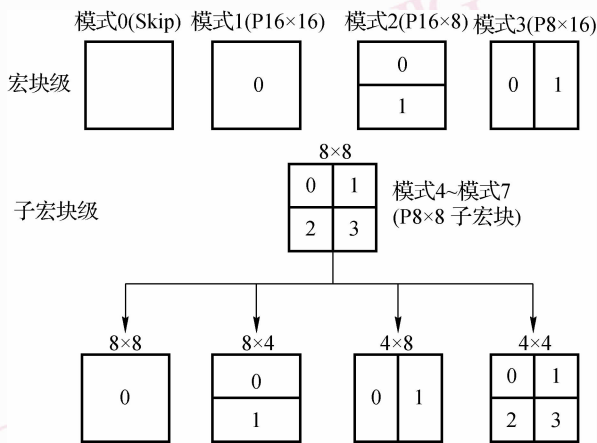


图 1 H. 264 运动补偿的宏块划分模式

Fig. 1 Division of the MB in motion compensation of H. 264

参考模型^[4]中提供了:基于失真的判决和基于率失真优化^[5] (rate distortion optimization, RDO)的判决两种最佳模式选择算法。

基于失真的判决在 JM 参考模型中的误差衡量函数为绝对误差和 (sum of absolute difference, SAD),可选取编码模式中绝对误差和最小的编码模式为最佳帧间编码模式,即

$$D_{SAD}(X_i, M) = \sum_{S \in X_i} |Y(S) - \hat{Y}_M(S)| \quad (1)$$

式中, $Y(S)$ 表示原始图像点 S 的像素值, $\hat{Y}_M(S)$ 表示在编码模式 M 下,重建图像点 S 的像素值。

基于率失真优化的判决开销函数为

$$J(\gamma^{\text{motion}}, \lambda_{\text{motion}}) = SAD(S, c(\gamma^{\text{motion}})) + \lambda_{\text{motion}} \times R(\gamma^{\text{motion}} - P) \quad (2)$$

其中, $\gamma^{\text{motion}} = (\gamma_x^{\text{motion}}, \gamma_y^{\text{motion}})^T$ 为模式 M 的运动矢量, $P = (p_x, p_y)^T$ 为模式 M 的运动矢量的预测值, λ_{motion} 为 Lagrange 乘数因子, SAD 函数同式 (1), $R(\gamma^{\text{motion}} - P)$ 为模式 M 所需的比特数。因此可在所有可供选择的编码模式中选择开销函数最小的编码模式作为最终编码模式。

在编码性能上,基于率失真优化的模式判决比基于失真的模式判决更为出色。本文主要是研究基于率失真优化的模式选择算法,同样适用于失真模式判决。

JM 参考模型的率失真优化的模式选择算法步骤如下:

- (1) 通过计算宏块级帧间编码模式的开销函数来选择最佳模式。其中可选模式有 Skip, 16×16, 16×8, 8×16;
- (2) 通过计算亚宏块级帧间编码模式的开销函数来选择最佳模式。其中可选模式有 8×8, 8×4, 4×8, 4×4;
- (3) 通过计算帧内编码模式的开销函数来选择最佳模式。其中可选模式有 Intra16×16, Intra4×4;
- (4) 根据率失真准则,在可供选择编码模式中选择率失真最小的编码模式为最终编码模式。

3 H. 264 中基于时空相关性的帧间模式选择快速算法

在 JM 参考模型中,由于帧间最佳模式的选择算法非常复杂、耗时,因此在实际应用中非常有必要对这个编码单元进行优化。众所周知,视频序列中

两个连续帧是高度相关的,本文依据视频序列相邻帧在时间上的高度相关性,认为对于序列中运动缓慢的目标,当前宏块的编码模式选择必然和前一帧对应位置的宏块编码模式有关;而对于运动剧烈的目标,由于相邻帧宏块的时间相关性相对较弱,因此可以利用邻近宏块的空间相关性,由已编码宏块的编码模式来得到当前宏块的模式信息。由以上获得的模式信息即可为模式选择快速算法提供依据。

3.1 初始候选模式选择

对于序列中运动缓慢的目标,连续帧的时间相关性可以为有效的模式搜索提供极为重要的信息。通常情况下,在视频图像中,有许多静止宏块或区域存在,如果可以较早地预测出这些宏块,则可以加速模式选择的搜索过程。

本文根据 JVT 所提供的测试代码 JM11,对不同视频序列的当前宏块所采用的最佳模式进行了概率统计,统计结果如表 1 所示。实验中用到了 6 个序列

表 1 基于时间相关性的宏块模式转换统计结果

Tab. 1 Statistical results for the MB mode change based on a temporal relationship (Unit: %)

| 模式转换 | 采用最佳模式的概率(%) | | | |
|-------------------|--------------|-------|-------|-------|
| | QP=24 | QP=28 | QP=32 | |
| Skip(模式 0) | Skip | 60.86 | 65.21 | 69.94 |
| | P16×16 | 20.03 | 17.49 | 16.54 |
| | P16×8 | 5.82 | 5.44 | 4.69 |
| | P8×16 | 5.62 | 6.00 | 5.33 |
| P8×8(子宏块模式) | P8×8 | 6.08 | 4.69 | 2.75 |
| | Skip | 22.04 | 24.59 | 30.03 |
| | P16×16 | 35.83 | 36.04 | 35.87 |
| | P16×8 | 10.39 | 10.28 | 10.51 |
| P16×16(模式 1) | P8×16 | 11.08 | 11.31 | 11.61 |
| | P8×8(子宏块模式) | 20.32 | 17.39 | 11.35 |
| | Skip | 17.59 | 20.82 | 25.97 |
| | P16×16 | 26.55 | 27.06 | 27.66 |
| P16×8(模式 2) | P16×8 | 18.60 | 19.36 | 19.55 |
| | P8×16 | 9.02 | 9.85 | 10.56 |
| | P8×8(子宏块模式) | 27.91 | 22.47 | 15.82 |
| | Skip | 15.14 | 19.14 | 24.63 |
| P8×16(模式 3) | P16×16 | 24.68 | 25.53 | 27.43 |
| | P16×8 | 10.43 | 9.36 | 9.89 |
| | P8×16 | 17.66 | 20.48 | 21.10 |
| | P8×8(子宏块模式) | 31.36 | 25.11 | 16.59 |
| P8×8(模式 4(子宏块模式)) | Skip | 7.62 | 10.16 | 11.36 |
| | P16×16 | 19.14 | 20.34 | 24.40 |
| | P16×8 | 8.41 | 9.62 | 11.86 |
| | P8×16 | 9.53 | 12.59 | 13.99 |
| P8×8(子宏块模式) | 54.96 | 46.81 | 38.08 | |

(mobile, flower garden, carphone, foreman, strfan, 和 salesman), 用各序列的前 100 帧进行实验分析。由表 1 可见,若当前一帧宏块编码模式为 Skip 或 16×16 时,即可得到对应的当前宏块候选模式的概率。统计实验结果表明:这两种宏块可以认为是静止区域部分。

当参考宏块的编码模式为 Skip 时,则当前宏块编码模式为 Skip 或 16×16 的概率大于 80%,而为 Skip, 16×16, 16×8 或 8×16 的概率则大于 91%。因此当前一帧宏块的编码模式为 Skip 时,则其对应的当前帧宏块的编码模式为 Skip, 16×16, 16×8 或 8×16 的可能性非常大。当参考宏块的编码模式为 16×16 时,其对应的当前帧宏块的编码模式为 Skip, 16×16, 16×8 或 8×16 的概率小于参考模式为 Skip 时的概率,但对于初始候选模式的选择来说,由 80% 左右的概率所提供的信息仍然是相当有价值的。当参考宏块的编码模式为矩形块(16×8 或 8×16)时,则当前帧宏块的编码模式为 Skip, 16×16, 16×8 或 8×16 的概率也将超过 70%。

依据以上分析得到的特性可知,当前宏块的初始候选模式可以由前一帧对应宏块的编码模式依次得到。例如,当参考模式为 16×16 时,则当前宏块的候选模式选为 {Skip, 16×16}; 当参考模式为 16×8 时,则当前宏块的候选模式选为 {Skip, 16×16, 16×8}。当参考模式为 8×8 时,则选择所有可能的模式 {Skip, 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4} 作为候选模式。

而对于运动剧烈的目标,则可以利用邻近宏块的空间相关性,由已编码宏块的编码模式来得到当前宏块的模式信息。与当前宏块相邻的已编码宏块共有 4 个(如图 2 所示)。如果其中有 3 个或 4 个有着相同的编码模式,则称该模式为主宏块模式。



图 2 邻近的已编码宏块

Fig. 2 Neighborhood MBs

通过对多种标准序列进行实验分析发现,存在主宏块模式的比例高达 42.6%,在主宏块存在的情形下,当前块编码模式也为主宏块模式的比例为

78.3%。因此,在选择初始候选模式时,应首先判断是否存在主宏块模式,若存在,则将该模式加入到初始候选模式中。

3.2 基于主宏块的宏块跟踪

目标跟踪就是在相邻帧中找到和当前目标最相关的目标区域的过程。为了找到感兴趣的目标区域,首先需要在时域定义一个跟踪标准,然后以此标准对目标进行搜索。在基于块匹配方法的视频编码中,本文算法以每个 16×16 的宏块作为感兴趣目标,用于最终的模式选择。如图 3 所示,一个 16×16 的宏块类型当用来在前一帧中找到和当前宏块最相关的目标区域时,这实际上就是对当前宏块的整像素进行运动估计。当经过整像素运动估计得到

最优运动矢量和最相关区域后,则和当前宏块最相关的宏块(前一帧中)定义如下:

$$MB_{\max} \{Area_A, Area_B, Area_C, Area_D\} \quad (3)$$

上式表示,从与匹配块相邻的 4 个宏块中选择一个主宏块作为和当前宏块最相关的块,该主宏块就是指与匹配块重叠面积最大的那个宏块。为了保证足够的相关,需定义一个相关系数门限 t 。当主宏块和当前宏块的相关系数大于这个门限时,则认为该主宏块即为最相关宏块。本文算法中选择相关系数门限 $t = 0.85 \sim 0.9$ 。在前一帧中找到最相关的宏块后,该宏块的率失真代价值即可用于判断模式选择过程是否自适应终止。

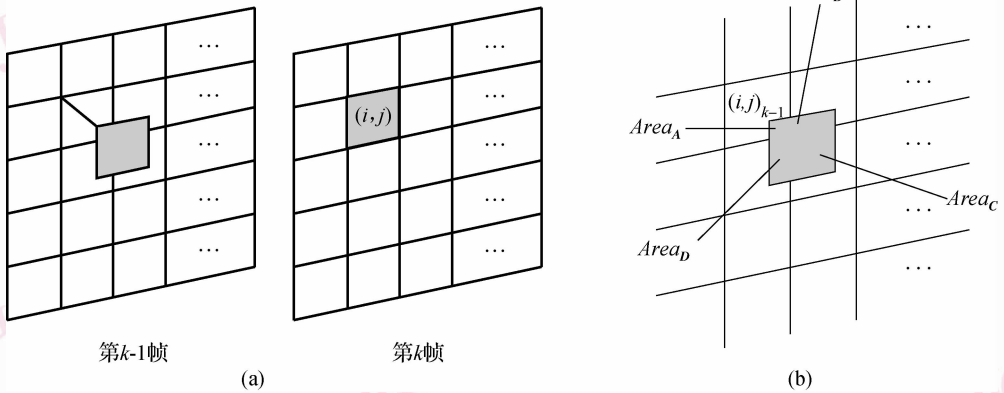


图 3 基于主宏块的宏块跟踪

Fig. 3 MB tracking scheme based on main MB

3.3 基于时空相关性的帧间编码模式选择快速算法

上文对 H. 264 编码器帧间编码模式选择中的宏块的初始候选编码模式进行了分析,并给出了基于时间相关性的宏块模式转换的统计分布规律。基于时空相关性的帧间模式选择快速算法是根据这些分布规律,结合宏块跟踪得到的最相关宏块的率失真代价值,采用门限控制的方法,依次对统计概率上可能是最佳的编码模式进行逐一筛选的方法。

基于时空相关性的帧间模式选择快速算法的具体步骤^[6-7]如下:

(1) 首先判断是否存在主宏块模式,若存在,则将该模式加入到以下得到的初始候选模式中。

(2) 若前一帧对应位置宏块的编码模式为 Skip,则先设当前宏块的初始候选模式为 {Skip}; 然后计算候选模式各自的率失真代价值,若其中最小

的率失真值小于参考宏块的率失真代价值,则该宏块的最终编码模式为率失真代价值最小者,并且不进行帧内模式搜索。否则,候选模式追加 $\{P16 \times 16, P16 \times 8, P8 \times 16\}$ 等模式,其中率失真代价值最小者即为当前宏块的最后编码模式。

(3) 若前一帧对应位置宏块的编码模式为 $P16 \times 16$ 模式,则先设当前宏块初始候选模式为 {Skip, $P16 \times 16$ }; 然后计算候选模式各自的率失真代价值,若其中最小的率失真值小于最相关宏块的率失真值,则该宏块的最终编码模式即为率失真代价值最小者,并且不进行帧内模式搜索。否则,候选模式追加 $\{P16 \times 16, P16 \times 8, P8 \times 16, P8 \times 8\}$ 等模式(追加 $P8 \times 8$ 模式是考虑到表 1 中候选模式为 $P8 \times 8$ 的概率比较大),其中率失真代价值最小者即为当前宏块的最后编码模式。

(4) 若前一帧对应位置宏块的编码模式为 $P16 \times 8$

模式,则先设当前宏块的初始候选模式为 $\{\text{Skip}, P16 \times 16, P16 \times 8\}$;然后计算候选模式各自的率失真代价,若其中最小的率失真值小于最相关宏块的率失真代价,则该宏块的最终编码模式为率失真代价最小者,并且不进行帧内模式搜索。否则,候选模式追加剩下的 $\{P8 \times 16, P8 \times 8\}$ 等模式,其中率失真代价最小者即为当前宏块的最后编码模式。

(5) 若前一帧对应位置的宏块编码模式为 $P8 \times 16$ 模式,则先设当前宏块的初始候选模式为 $\{\text{Skip}, P16 \times 16, P8 \times 16\}$;然后计算候选模式各自的率失真

代价,若其中最小的率失真值小于最相关宏块的率失真代价,则该宏块的最终编码模式为率失真代价最小者,并且不进行帧内模式搜索。否则,候选模式追加剩下的 $\{P16 \times 8, P8 \times 8\}$ 等模式,其中率失真代价最小者即为当前宏块的最后编码模式。

(6) 若前一帧对应位置的宏块编码模式为 $P8 \times 8$ 模式:则当前宏块的候选模式为以上提到的所有宏块编码模式,其中率失真代价最小者即为当前宏块的最后编码模式。

算法流程图如图 4 所示:

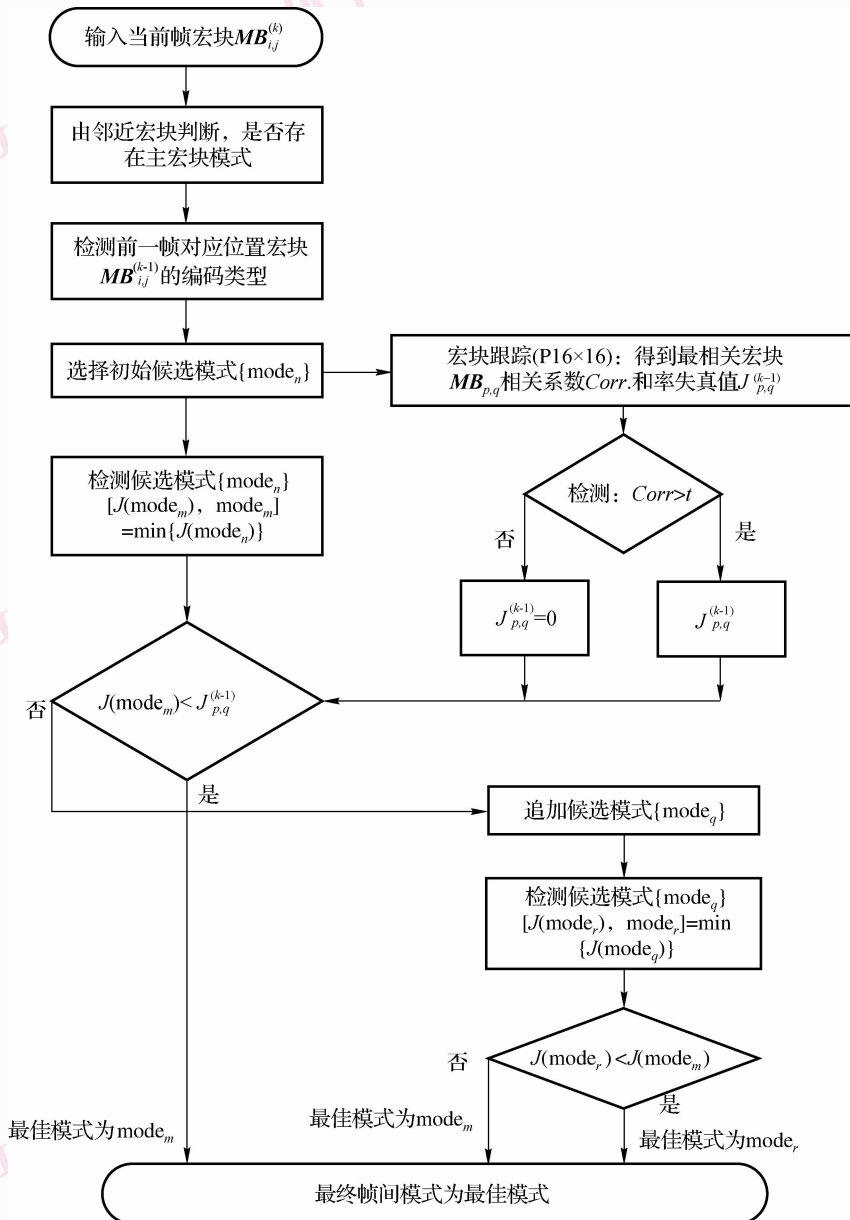


图 4 模式选择算法流程图

Fig. 4 Overall scheme of the proposed fast mode decision algorithm

4 实验结果与分析

在 H. 264 测试代码 JM11 基础上,加入基于时空相关性的帧间模式选择算法后,分别从图像质量、码率和编码速度 3 个方面和校验模型原始程序以及 Salgado's 算法^[8]和 Jeon's 算法^[9]进行了比较,以检验算法的有效性,Salgado's 算法和 Jeon's 算法都是比较经典的高效快速算法。本文采用 6 个测试序列的前 100 帧进行测试,这 6 个序列分别是“foreman”,“mobile”,“coastguard”,“flower”,“container”,“akiyo”;图像大小为 CIF;帧率为 30 fps;实验时,首先选用 Main Profile 编码,并选用一个参考帧,运动矢量搜索范围 $MV = \pm 16$;然后打开哈达码变换和率失真优化,进行 CAVLC 熵编码;最后统计每个序列在 QP = 28, 32, 36, 40 时的亮度的峰值信噪比(PSNR)值、码率和编码节省时间。

表 2 中列出了不同算法的编码性能比较结果, $\Delta PSNR$ 为编码前后视频 PSNR 值的变化, $\Delta bits$ 为编码前后视频的码率变化。负值的 $\Delta PSNR$ 和正值的 $\Delta bits$ 表示各算法的编码性能相对于测试代码 JM11 的编码性能下降,而节省时间则反映出各算法的效率高于校验模型 JM11。这是由于快速算法选用了较少的候选模式所造成的。从表 2 可以看出,与测试代码 JM11(参考模型)的编码性能相比,本文算法最差情况下的视频编码质量下降为 -0.417 dB,平均为 -0.314 dB,与 Jeon's 算法相比,虽不如 Jeon's 算法的编码性能,但优于 Salgado's 算法。这可以反映出,除了校验模型,Jeon's 算法是选用候选模式最多的,本文算法次之,Salgado's 算法最少。从人眼的视觉效果来看,当图像质量的下降在 -0.5 dB 以内时,人的主观感觉并不会太大的变化。由此可见,本文方法造成的图像质量下降,因其幅度很小,故基本上不会影响视频的主观视觉效果。

采用较少的候选模式能够减少编码器的复杂度和节省编码时间。表 2 比较了本文算法和其他算法节省的编码时间,从表 2 可以看出,与校验模型相比,本文算法最多可以节省约 69.54% 的时间,平均节省时间约 64%。可见本文算法节省时间明显高于 Jeon's 算法,和 Salgado's 算法接近,这与选用候选模式的多少是紧密相关的。

表 2 编码性能比较结果

Tab 2 Performance comparison of the proposed algorithm on the JM 11.0 reference encoder

| 序列(CIF) | 方法 | $\Delta PSNR$ (dB) | $\Delta bits$ (%) | 节省时间(%) |
|------------|--------------|--------------------|-------------------|---------|
| foreman | 本文算法 | -0.376 | 3.247 | 63.36 |
| | Salgado's 算法 | -0.476 | 3.924 | 69.08 |
| | Jeon's 算法 | -0.009 | -0.090 | 54.68 |
| mobile | 本文算法 | -0.390 | 2.321 | 61.06 |
| | Salgado's 算法 | -0.490 | 3.156 | 59.33 |
| | Jeon's 算法 | -0.010 | 0.040 | 34.59 |
| coastguard | 本文算法 | -0.287 | 3.310 | 64.02 |
| | Salgado's 算法 | -0.416 | 3.924 | 65.77 |
| | Jeon's 算法 | -0.009 | -0.130 | 21.05 |
| flower | 本文算法 | -0.417 | 5.147 | 62.28 |
| | Salgado's 算法 | -0.433 | 4.569 | 65.18 |
| | Jeon's 算法 | -0.398 | -0.013 | 44.01 |
| container | 本文算法 | -0.251 | 3.924 | 65.91 |
| | Salgado's 算法 | -0.516 | 4.103 | 76.47 |
| | Jeon's 算法 | -0.044 | -0.225 | 57.93 |
| akiyo | 本文算法 | -0.162 | 4.405 | 69.54 |
| | Salgado's 算法 | -0.176 | 3.952 | 79.06 |
| | Jeon's 算法 | -0.013 | 0.104 | 58.41 |

5 结论

本文提出了一种基于时空相关性和宏块跟踪的帧间模式选择快速算法。该算法之所以能产生较好编码效果的原因在于:它能够先通过对帧间编码的编码模式转换规律进行分析来得出其统计分布规律,并可由宏块跟踪得到最相关的宏块;然后就可根据所得的分布规律,依据门限控制得到最佳编码模式。仿真结果说明,本文算法应用在 H. 264 编码器中,可以在基本保证编码质量的同时,能很大程度地降低编码器的运算复杂度,同时候选模式的获取和门限的计算也比较简单,没有额外增加编码器的计算量。

参考文献(References)

- 1 ITU-T Rec. H. 264 | ISO/IEC 14496-10 AVC, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification[S]. In: Joint Video Team (JVT) of ISO/IEC MPEG

- &ITU-T VCEG 7th Meeting, Pattaya, Thailand, 7-14 March, 2003.
- 2 Wiegand T, Sullivan G J, Bjontegard G, *et al.* Overview of the H.264/AVC video coding standard [J]. *IEEE Transactions on Circuits and Systems Video Technology*, 2003, **13**(7): 560-576.
 - 3 Richardson E G. *H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia* [M]. New York, USA: John Wiley & Sons, 2003.
 - 4 H.264/AVC Reference Software JM11 [EB/OL]. <http://iphone.hhi.de/suehring/tml/download/>. (2005-03-15) [2008-07-26].
 - 5 Sullivan G J, Wiegand T. Rate-distortion optimization for video compression [J]. *IEEE Signal Processing Magazine*, 1998, **15**(6): 74-90.
 - 6 Kim Jung-ho, Kim Byung-gyu. Fast block mode decision algorithm in H.264/AVC video coding [J]. *Visual Communication Image Representation*, 2008, **19**(2): 175-183.
 - 7 Feng bin, Liu Wen-yu, Zhu Guang-xi. Fast adaptive mode decision algorithm for H.264 based on spatial correlation [J]. *Journal on Communications*, 2006, **27**(1): 75-80. [冯宾, 刘文予, 朱光喜. 基于空间相关性的 H.264 快速自适应模式选择算法 [J]. *通信学报*. 2006, **27**(1): 75-80.]
 - 8 Salgado L, Nieto M. Sequence independent very fast mode decision algorithm on H.264/AVC baseline profile [A], In: *Proceedings of International Conference Image Process* [C], Atlanta, GA, USA, 2006, **10**: 41-44.
 - 9 Choi I, Lee J, Jeon B. Fast coding mode selection with rate-distortion optimization for MPEG-4 part-10 AVC/H.264 [J], *IEEE Transactions on Circuits Systems Video Technology*, 2006, **16**(12): 1557-1561.