

H.264 中与码率控制相结合的快速运动估计

黄晓平, 沈未名, 郭晓云, 喻占武

(武汉大学测绘遥感信息工程国家重点实验室, 武汉 430079)

摘要: 根据 H.264 中整数变换、量化的特点, 可以证明在宏块预测 SAD 和实际 SAD 的误差小于某一门限时, 对编码质量不会产生影响。结合 H.264 码率控制算法中自适应 SAD 的预测, 提出一种根据 SAD 预测值来提前终止运动搜索的快速运动估计方法。实验结果表明, 该算法在图像质量稍有变化的情况下, 能有效提高运动估计速度, 最高可提高到 2.2 倍。

关键词: 整数变换; 运动估计; 提前终止; 预测 SAD; 码率控制

Fast Motion Estimation Combined with Rate Control in H.264

HUANG Xiao-ping, SHEN Wei-ming, GUO Xiao-xun, YU Zhan-wu

(State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, Wuhan 430079)

【Abstract】 Motion estimation is a most times-consuming module in H.264 encoder. Early terminating motion search is a best method for improving motion estimation. In H.264 encoder, rate control computes the predicted SAD which can be used as a decision condition for terminating motion search. It can be proved that image quality has no changes if the difference between predicted SAD and best SAD is small. Experimental results show the proposed early terminating motion search can effectively improve motion estimation speed and keep good image quality.

【Key words】 integer transform; motion estimation; early terminating; predicted SAD; rate control

1 概述

H.264^[1]是由ITU 和ISO/IEC的专家共同组成的联合视频小组JVT(Joint Video Team)制定的新一代视频编码国际标准。与以往H.263、MPEG-4 等视频编码标准类似, H.264 也采用基于块运动补偿和变换编码的混合编码框架, 但H.264 引入了许多新的编码技术, 如多模式帧内预测、多种块形状的运动估计、多参考帧、4×4 整数变换、自适应算术编码等。在相同的图像质量下, H.264 的压缩效率比H.263、MPEG-4 高50%以上^[2]。除了具有极高的压缩比, H.264 还具有良好的网络友好性, 能适用于高码率、低码率、有线和无线等各种领域, 具有极大的应用前景。但H.264 高效的压缩率是以高度复杂的计算为代价的, 其编码复杂度是H.263 的4~5倍^[2]。如何降低其计算复杂度, 提高编码效率, 成为H.264 能否取得成功的关键。

帧间预测是提高视频编码压缩比的关键技术, 通过运动估计找到最佳匹配块来实现。在 H.264 编码过程中, 运动估计部分计算复杂度时间占整个编码器 50%以上, 是优化和改进的关键。本文利用码率控制计算得到的预测 SAD 作为运动搜索提前退出的判决条件, 能在保持图像质量的前提下大大提高运动估计速度。首先根据 H.264 中 4×4 整数变换和量化的特点, 证明了在宏块预测 SAD 和实际 SAD 的误差小于某一门限时, 对编码质量不会产生影响, 从而提出了一种基于 SAD 预测的运动搜索提前终止准则, 可以有效提高运动估计速度, 因而提高了整个编码器的编码效率。

2 运动估计提前终止的条件

运动估计通过当前块与参考块对应像素差的绝对值和 SAD 的大小来判断是否为最佳匹配块。一般对搜索范围内所有块进行 SAD 计算, 取 SAD 最小的那个块作为最佳匹配块。对最佳匹配块与当前块对应像素的差(即残差块)进行整数变

换, 量化后进行变长编码得到码流。运动估计提前终止指没有搜索完区域内所有块时, 就已经得到一个合适的 SAD 值, 于是提前结束运动搜索, 减少计算量, 提高编码速度。运动估计提前终止的关键是确定这个合适的 SAD 值, 使编码器在保持图像质量的前提下提高速度。定义变量 SAD_{best} 为在运动搜索范围内的最小 SAD, SAD_{pred} 是对 SAD 的预测。定义变量 SAD_{thrh} 来表示提前终止的门限值:

$$SAD_{thrh} = SAD_{pred}(1 + \beta) \quad (1)$$

其中, β 为权系数, 通过调整 β , 可以在编码质量和搜索速度方面达到一种平衡。

在运动估计过程中, 当满足 $SAD_{best} \leq SAD_{thrh} \leq SAD_{exit}$ 时, 将停止搜索。很显然, 如果 $SAD_{thrh} \leq SAD_{best}$, 则不会提前终结。为了保证提前终结的策略不会带来图像质量的恶化, 定义门限 ΔTH , 并且假定当满足下面条件时, 本文的运动矢量已经足够精确, 因而不会导致图像质量的下降:

$$SAD_{exit} - SAD_{best} < \Delta TH \quad (2)$$

因为 $SAD_{exit} \geq SAD_{thrh}$, 所以式(2)等价于

$$SAD_{thrh} - SAD_{best} < \Delta TH \quad (3)$$

下面证明存在这样的 β 使得式(3)成立。因为 H.264 的 4×4 整数变换中残差值近似地满足零均值的拉普拉斯(Laplacian)分布^[3], 其协方差为

$$r(m, n) = \sigma_f^2 \rho^{|m|} \rho^{|n|} \quad (4)$$

其中, m, n 为 2 个像素间的水平与竖直距离; σ_f^2 为残差的方

基金项目: 国家“973”计划基金资助重点项目(2004CB318206)

作者简介: 黄晓平(1974 -), 女, 博士研究生, 主研方向: 图像与视频编解码, 视频通信; 沈未名, 教授、博士生导师; 郭晓云, 博士研究生; 喻占武, 教授、博士生导师

收稿日期: 2007-07-10 **E-mail:** hxp_cy@163.com

差； $\rho(|\rho| < 1)$ 是自相关系数，则变化可以表示为

$$F = AfA^T \quad (5)$$

其中， A 为变换矩阵； f 为残差矩阵。则第 (u,v) 个变换系数的方差 $\sigma_f^2[u,v]$ 可以写为

$$\sigma_f^2[u,v] = \sigma_f^2[ARA^T]_{u,u}[ARA^T] \quad (6)$$

$$\text{其中，} R = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}, \text{这里不妨取 } \rho=0.6(\text{Suitez 图像})$$

序列在不同帧率下的均值，对最终结果无影响)，则可推导出时域和空域方差的关系为

$$\sigma_f^2[u,v] = \sigma_f^2 \times M[u,v], u=0 \sim 4; v=0 \sim 4 \quad (7)$$

其中， M 矩阵为

$$M = \begin{bmatrix} 5.607424 & 2.125210 & 1.060864 & 0.678503 \\ 2.125210 & 0.805453 & 0.402067 & 0.257152 \\ 1.060864 & 0.402067 & 0.200704 & 0.128365 \\ 0.678503 & 0.257152 & 0.128365 & 0.082099 \end{bmatrix} \quad (8)$$

假定量化系数为 Q_{step} ，则第 (u,v) 个变换系数的2个不同值，如果满足下式，则对图像质量无影响：

$$\sigma_{F1}^2[u,v] - \sigma_{F2}^2[u,v] < Q_{step} \quad (9)$$

通过式(7)、式(9)，可得

$$(\sigma_{f1}^2 - \sigma_{f2}^2) \times M[u,v] < Q_{step}, u=0 \sim 3; v=0 \sim 3 \quad (10)$$

由于残差值近似的满足零均值的Laplacian分布，因此可计算时域的方差为

$$\sigma_f \approx \sqrt{2} \times SAD / B_{size} \quad (11)$$

其中 B_{size} 为当前块大小。如果认为 σ_{f1} 是 $SAD=SAD_{thr}$ 时时域的方差， σ_{f2} 是 $SAD=SAD_{best}$ 时时域的方差，则有

$$\frac{2}{(B_{size})^2} \times [SAD_{thr}^2 - SAD_{best}^2] < \frac{Q_{step}}{M[u,v]}, u=0 \sim 3, v=0 \sim 3 \quad (12)$$

因为 $M[0,0]$ 是 M 矩阵中最大的元素，并且 $SAD_{best} \approx SAD_{thr}$ ，所以式(12)可化简为

$$\frac{4 \times SAD_{thr}}{(B_{size})^2} \times [SAD_{thr} - SAD_{best}] < \frac{Q_{step}}{M[0,0]} \quad (13)$$

即

$$SAD_{thr} - SAD_{best} < \frac{Q_{step}}{M[0,0]} \times \frac{(B_{size})^2}{4 \times SAD_{thr}} \quad (14)$$

因此，门限值 ΔTH 为

$$\Delta TH = \frac{Q_{step}}{M[0,0]} \times \frac{(B_{size})^2}{4 \times SAD_{thr}} \quad (15)$$

从式(15)可以看出，存在一个门限 ΔTH ，当 $SAD_{thr} - SAD_{best} < \Delta TH$ 时，运动估计提前终止对最终的图像质量无影响。

3 运动估计提前终止算法

3.1 H.264的码率控制方法

通过对H.264码率控制方法的多个提案进行评估测试，JM参考模型最终采用了JVT-G012^[4]提出的方法，即用流量控制模型来分配每个基本单元目标比特数，并在宏块层编码采用二次率失真函数计算量化参数(QP)的算法。基本思想是根据图像和宏块复杂度来分配目标比特数，而计算QP在编码宏块之前，并不知道宏块的实际复杂度，于是采用预测宏块复杂度MAD的方法。使用线性预测模型通过上一帧同一位置宏块的MAD来预测当前帧当前宏块的MAD。分别用 MAD_{cb} 和

MAD_{pb} 表示当前帧当前宏块的MAD和上一帧同一位置宏块的MAD，则线性预测模型为： $MAD_{cb} = \alpha_1 \times MAD_{pb} + \alpha_2$ ，其中， α_1 和 α_2 是预测模型的2个系数，初始值分别为1,0，每编码完一个宏块更新一次。对模型系数进行更新时，为了提高模型的预测准确性，采用以下3个步骤：

(1)选择更新线性预测模型系数的数据点。数据点的数量和质量对于模型的精确性是十分关键的。一般来说，数据点的数量越多，模型的精度就越高，但同时计算复杂度也会提高。而对于数据点的质量评估，可以使用一些比较客观的指标(比如MAD)。考虑到这些因素，在数据点的选择上，采用了基于滑动窗口的数据选择机制。如果图像的复杂度有明显改变(例如剧烈的运动)，则使用只包含最近数据点的小窗口，否则使用大的窗口。通过滑动窗口机制，编码器能够自适应地选择那些有效的数据点。数据点的选择是基于当前对象(宏块，帧等)和前一编码对象的场景切换比较的。不过，基于滑动窗口的数据选择机制是尽量减小场景切换的影响，但并不能完全解决场景切换带来的问题。

(2)利用线性回归技术，计算模型系数 α_1 和 α_2 。

(3)排除不合理点，所谓不合理点就是那些由于场景切换造成的误差较大的点。在得到新的系数后，需要去掉这些不合理点，从而实现预测模型的改进和校准。具体为：如果预测结果和实际值的标准差大于 $K(K$ 是一个常数)，那么这个点就被认为是不合理点而被排除。但要注意最近的一个数据点一定要被选择。

3.2 运动估计提前终止策略

因为 $SAD_{thr} = SAD_{pred}(1 + \beta)$ ，所以要对图像质量无影响的运动估计提前终止，需要有一个准确预测的SAD。而在H.264中，码率控制算法已经在编码前就对MAD进行了较准确预测并且保存在相应变量中，而MAD就是SAD除以256得到。所以，在运动估计时，可以直接使用此SAD预测值，不需要重新计算，对于SAD的预测在运动估计部分不会带来额外的计算量。而码率控制部分采用基于线性预测模型和滑动窗口的自适应算法，MAD的预测精度较高，可以满足运动估计中的需要。而权系数 β 可以根据速度和精度的不同要求进行调整，从而实现效果较好的运动估计提前终止。具体就是在进行运动搜索时，如果某个运动矢量的SAD小于或等于 SAD_{thr} ，就停止运动搜索，这样可以减少许多不必要的运动搜索，从而提高运动估计速度。

4 实验及分析

为便于比较本文提出的把 SAD_{thr} ($SAD_{thr} = SAD_{pred}(1 + \beta)$)作为运动估计提前结束阈值的算法对编码速度和编码效率的影响，在H.264参考平台JM98的基础上，对不同的测试序列进行测试。本文选取具有代表性的3个不同运动大小的视频序列，其中，Claire是简单的头肩运动序列，Container是变化比较平缓的运动序列，而Foreman是变化较快的运动序列，大小都是 176×144 ，码率为96Kb/s，采用7种块搜索模式，全搜索且运动搜索范围为 ± 16 ，参考帧数为1，采用率失真优化，有码率控制且码率为96Kb/s，CAVLC变长编码，测试序列长度都为100帧，baseline，编码形式为IPPP等。测试结果见表1。

(下转第199页)