

一种 H. 264 帧内预测模式的快速选择算法

李文杰^{1),2)} 李刚¹⁾ 曹斌¹⁾ 吴功伟¹⁾

¹⁾(中国电信股份有限公司北京研究院,北京 100035) ²⁾(联想中国研究院,北京 100085)

摘要 H. 264/AVC 视频编码标准采用码率—失真度最优化(RDO)算法,以保证更高的压缩效率和更好的图像质量,但同时也大大增加了编码器的复杂度和运算负担。本文提出了一种新型的帧内预测模式的快速选择算法,该算法基于图像内部信息的变化剧烈程度进行判断,并通过方向矢量来减少候选帧内预测模式。试验结果表明,该算法可以将整个编码时间减少到原来采用 RDO 算法时的 40%~50%,而且和 RDO 算法比较,最终压缩后的视频图像质量基本没有下降。

关键词 H. 264 帧内预测 方向矢量 码率—失真度最优化(RDO)算法

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2009)06-1114-05

A Fast Intra-frame Prediction Mode Selection Algorithm for H. 264 Video Coding

LI Wen-jie^{1),2)}, LI Gang¹⁾, CAO Bin¹⁾, WU Gong-wei¹⁾

¹⁾(Beijing Research Institute, China Telecom Corporation Limited, Beijing 100035)

²⁾(Lenovo Corporate Research & Development, Beijing 100085)

Abstract In order to guarantee higher compression efficiency and video quality, a robust rate-distortion optimization (RDO) is employed in H. 264/AVC video coding standard, which is jointly developed by ITUT and MPEG. According to the specification of RDO, all of the prediction modes have to be tried exhaustively in order to find the best mode. As a result, the computational complexity of the mode decision is extremely high and is not suitable for the real time service, such as video conference. This paper proposes a fast intra prediction modes decision algorithm based on the discrepancy within the video picture. In addition, the Direction-Vector based preferred modes selection method is introduced in the proposed algorithm to reduce the candidate modes to be tested. Experimental results are given to demonstrate that our proposed algorithm achieves a reduction of processing time by 40%~50% as compared with the current RDO optimized mode decision with little quality degradation.

Keywords H. 264 intra prediction, directional vector, rate-distortion optimization (RDO)

1 引言

H. 264/AVC 是由 ITUT 和 MPEG 一起联合开发的新一代视频编解码标准,该标准可以在不牺牲图像质量的前提下提供更高的压缩效率。为了达到上述目的,H. 264/AVC 采用了许多新技术,例如帧内

编码的空间预测,可变块大小的运动补偿,多个参考帧以及自适应二进制编码等。同时也大大增加了编码器的运算复杂度。在这新的特性中,通过运用可变大小块和多个方向预测的技术,帧内预测模式被极大的丰富了。根据文献[1]的研究结果,借助于这些新特性 H. 264/AVC 帧内编码的性能超过了 JPEG-2000 的静态图像压缩标准。H. 264/AVC 中

收稿日期:2007-03-15;改回日期:2007-10-24

第一作者简介:李文杰(1970~),男,2000年于北京理工大学获计算机应用硕士学位。主要研究方向为计算机视觉、图像处理。

E-mail:liwj@ctbri.com.cn

采用 RDO 优化技术来进行帧内预测模式的选择来达到更高的编码效率。而 RDO 定义了所有的帧内预测模式都必须测试一遍,然后找到最佳的预测模式。因此,基于 RDO 技术的运算复杂度是非常高的,也就不适合一些需要实时应用的场合,例如视频会议等。

本文描述了一种高效的快速模式选择算法来进行 H. 264/AVC 中帧内预测模式的选择。该算法基于编码目标模块图像信息的变化剧烈程度和方向矢量,大致可以分为 2 个阶段。第 1 阶段,利用目标模块内部的变化剧烈程度信息来决定采用 16×16 或者 4×4 的模块进行亮度的编码预测。第 2 阶段,如果决定采用 4×4 模块,就利用目标模块的方向矢量来预测最佳编码帧内预测的候选模式。提出的算法可以先滤除掉大部分的候选模式,因此最终将减少到仅有 3~4 个模式,然后通过进一步的 RDO 运算从中选出最佳模式。

2 图像信息变化剧烈程度的计算

在 H. 264/AVC 中总共定义了 9 种 4×4 的亮度模块,4 种 16×16 的亮度模块和 4 种 8×8 色度模块的帧内预测模式^[2]。为了最终最佳的选择,H. 264/AVC 要求对每一种模块都进行一次 RDO 运算,即对亮度部分来说总共要进行 13 种模式的计算,对色度部分也要进行 4 种模式的计算。RDO 的运算是非常占用时间的,尽管它可以保证获得最佳的压缩效率和图像质量。需要注意的是在视频图像中的绝大部分图片内都包含了大量的冗余信息,特别是在那些图片内容变化不大的视频中^[3]。观察到对那些变化不是很剧烈的模块,如果采用 16×16 大小的模块进行预测,图像的质量不会有明显地下降。因此,可以省略掉那些 4×4 大小模块的 RDO 运算过程,从而明显地减少整个视频图像的编码时间。为了能够衡量一个宏块内部的变化剧烈程度的不同程度,采用了宏块内部信息的变化剧烈程度,如式(1):

$$Discrepancy = \sqrt{\sum_{x=0}^{15} \sum_{y=0}^{15} (L_{xy} - (\sum_{x=0}^{15} \sum_{y=0}^{15} (L_{xy})/256))^2} \quad (1)$$

式中, L_{xy} 表示的是一个宏块中的像素点的亮度值。如果用一个宏块中的像素点的亮度值构造一个矩阵的话,图像内部信息的变化剧烈程度 *Discrepancy* 就

是该矩阵的标准方差。在提出的算法中,用式(2)来代替式(1)表示一个宏块内部信息的变化剧烈程度,这样可以避免平方根的复杂运算。

$$D = \sum_{x=0}^{15} \sum_{y=0}^{15} \left(L_{xy} - \left(\sum_{x=0}^{15} \sum_{y=0}^{15} (L_{xy})/256 \right) \right)^2 \quad (2)$$

一个宏块内部信息的变化剧烈程度描述了宏块中的内部差异,可以用来作为一个门限值来决定采用 16×16 还是 4×4 的模块大小来进行预测。尽管如此,该门限值的大小可能依据视频图像特征的不同而变化。所以采用了一种自适应的方法来实时地根据不同视频图像的特点来调整该门限值的大小。算法的初始阶段,采用标准的 RDO 运算,每一个宏块内部信息的变化剧烈程度和对应的帧内预测模式将被记录下来。接下来,根据前面的统计结果和图像内部信息的变化剧烈程度与预测模式之间的关系,就可以来确定门限值的大小。这个过程将会在某一个固定的间隔内(默认可取 10 秒)重复地进行,以保证提出的算法能够及时适应视频图像的变化。一旦该门限值确定之后,每一个目标宏块内部信息的变化剧烈程度都将与该门限值进行比较,用来确定在下一步中将要采用的编码预测模块的大小。

3 方向矢量的确定

正如前面提到的,在 H. 264/AVC 中总共定义了 9 种 4×4 的亮度模块,4 种 16×16 的亮度模块和 4 种 8×8 色度模块的帧内预测模式,这些不同的预测模式可以按照不同的预测方向来进行划分。因此,如果可以估计出合适的预测方向,就可以选定某些预测模式来进行 RDO 运算,这样就省略掉了大部分不必要的预测模式的检测。实际上,像素点可以很好地由与它们的亮度值或色度值的梯度的垂直的方向上进行预测,而这个方向可以由各个像素点的亮度和色度的分布通过 Sobel 运算符得到^[4]。按照式(3)和式(4)定义一个像素点的方向矢量:

$$DV_x(x, y) = L(x-1, y+1) + 2 \times L(x, y+1) + L(x+1, y+1) - L(x-1, y-1) - 2 \times L(x, y-1) - L(x+1, y-1) \quad (3)$$

$$DV_y(x, y) = L(x+1, y-1) + 2 \times L(x+1, y) + L(x+1, y+1) - L(x-1, y-1) - 2 \times L(x-1, y) - L(x-1, y+1) \quad (4)$$

式中, $L(x, y)$ 为位于 (x, y) 像素点的亮度或者色度。

此外,方向矢量的方向角由式(5)确定。

$$\theta = \tan^{-1}(DV_x/DV_y) \quad (5)$$

因为 $-\pi/2 < \theta \leq \pi/2$, 而方向角的范围是从 $-\pi$ 到 π , 因此式(5)必须做如下的变动:

$$\theta = \begin{cases} \tan^{-1}(DV_x/DV_y) + \pi & DV_x < 0, DV_y > 0 \\ \tan^{-1}(DV_x/DV_y) - \pi & DV_x < 0, DV_y < 0 \\ \tan^{-1}(DV_x/DV_y) & \text{其他} \end{cases} \quad (6)$$

因为在 H. 264/AVC 中只有有限个帧内预测模式, 因此每一个像素点的原始方向矢量必须按照帧内预测模式的规则被重新设置, 具体设置方法如下:

$$\theta' = \begin{cases} \theta + \pi & -\pi < \theta < -3\pi/4 \\ \theta - \pi & 3\pi/4 < \theta < \pi \end{cases} \quad (7)$$

在这里 θ' 表示的是重新设置后的方向角, 必须注意的是方向矢量的幅度并没有发生变化。在重新设置过每一个像素点的方向角之后, 就可以通过矢量的加法运算计算出 4×4 模块的整体方向矢量。

$$D_x = \sum_{n=0}^{15} DV_x^n \quad D_y = \sum_{n=0}^{15} DV_y^n \quad (8)$$

式中, DV_x^n 和 DV_y^n 为重新设置过后的方向矢量的水平分量和垂直分量。这样, 平均方向矢量的方向角就可以按照式(9)计算。

$$\omega = \begin{cases} \tan^{-1}(D_x/D_y) - \pi & D_x < 0, D_y < 0 \\ \tan^{-1}(D_x/D_y) & \text{其他} \end{cases} \quad (9)$$

一旦 ω 确定之后, 就可以估计首选的预测方向和对应的帧内预测模式。把整个方向空间根据 8 个指定的预测方向(除了 DC 模式)分为 8 个不同的部分。因此, 首先每一个 4×4 block 的预测模式可以根据 ω 所在的方向空间确定出来, 详细的确定方法由式(10)描述:

$$Modes = \begin{cases} 1, 8 & 0^\circ < \omega \leq 45^\circ \\ 1, 6 & -26.6^\circ < \omega \leq 0^\circ \\ 6, 4 & -45^\circ < \omega \leq -26.6^\circ \\ 4, 5 & -63.4^\circ < \omega \leq -45^\circ \\ 5, 0 & -90^\circ < \omega \leq -63.4^\circ \\ 0, 7 & -116.6^\circ < \omega \leq -90^\circ \\ 7, 3 & -145^\circ < \omega \leq -116.6^\circ \end{cases} \quad (10)$$

式(10)中的模式号码的定义和文献[5]中的定义是一样的。需要注意的是 DC 模式, 模式号码为 2, 因为 DC 模式也是 H. 264/AVC 所指定的一种帧

内预测模式。而且观察到对各种不同的视频序列而言, 绝大部分的最佳帧内预测模式是 DC 模式。因此, 首选的帧内预测模式应该始终包含 DC 模式, 这样提出的快速算法最终对每一个 4×4 亮度模块有 3 个候选模式进行 RDO 运算, 比原先的 9 个候选模式减少了 2/3。

4 快速帧内模式选择算法

根据 JVT 提供的参考编解码模块 JM98, 一个完整的搜索必须对每一个宏块的亮度部分进行 13 种模式的 RDO 运算, 对色度部分进行 4 种模式的运算。从总体上来讲, 提出的快速算法可以分为 2 个阶段, 分别解决不同的问题。在第 1 个阶段, 利用每一个宏块内部信息的变化剧烈程度信息来决定在后面的亮度部分的预测编码中所采用的模块大小, 也就是说决定 4×4 还是 16×16 的亮度模块。变化剧烈程度越大, 表明一个图像内部变化的越剧烈, 此时采用 4×4 大小的模块可以更精确地反映原始图像的信息。需要注意的是判决门限值的大小是根据不同的视频序列进行动态调整的。在第 2 个阶段, 目标是基于方向矢量的信息尽可能地减少进行 RDO 运算的候选模式的数目。注意步骤如下所示, 这里只考虑了亮度部分的算法, 至于色度的和 16×16 亮度模块的快速算法相似:

(1) 根据视频序列的统计信息来更新门限值;

(2) 确定门限值之后, 每一个宏块内部信息的变化剧烈程度和该门限值进行比较。如果后者大于前者, 进入步骤(4), 否则进入步骤(3);

(3) 进行 16×16 亮度模块的快速选择算法, 然后重复步骤(2);

(4) 进行 4×4 亮度模块的快速选择算法, 然后重复步骤(2);

(5) 如果一定时间间隔的视频序列已经被编码完毕, 重复步骤(1)。

4.1 4×4 亮度模式快速选择算法

4×4 亮度模块帧内预测快速算法是基于 4×4 亮度模块的平均方向矢量。已经描述过了计算平均方向矢量的方法。尽管如此, 在实际应用中, 三角函数和反三角函数的计算也是比较消耗时间的, 而且这些计算也不是必需的, 可以利用垂直和水平分量的比值来代替具体的角度值。正如前面提到的, DC 模式始终是被选定作为一个首选模式参加 RDO 运

算的。综上所述,对一个 4×4 亮度模块的帧内预测编码,方向矢量可以指出两个首选的预测模式,加上 DC 模式,已经只需要进行 4 次 RDO 运算就可以达到目的,而且保证视频图像的质量没有明显下降。

4.2 16×16 亮度模式快速选择算法

如果一个目标宏块内部信息的变化剧烈程度小于指定的门限值,那么亮度部分将采用 16×16 模块进行帧内预测。根据前面相同的算法,首选的预测模式可以由 16×16 亮度模块的平均方向矢量决定。需要注意的是确定模式的规则和 4×4 亮度模块的规则是不一样的,如式(11)所示:

$$Modes = \begin{cases} 0 & 67.5^\circ < \omega \leq 112.5^\circ \text{ 或} \\ & -112.5^\circ < \omega \leq -67.5^\circ \\ 1 & |\omega| < 22.5^\circ \text{ 或 } 157.5^\circ < |\omega| < 180^\circ \\ 3 & 22.5^\circ < \omega \leq 67.5^\circ \text{ 或} \\ & -157.5^\circ < \omega \leq -112.5^\circ \end{cases} \quad (11)$$

同样地,DC 模式也必须作为一个候选模式。因此,对 16×16 亮度模块来讲,最终一共有两个候选模式需要进行 RDO 运算。

4.3 8×8 色度模式快速选择算法

对于色度模块,因为在 H. 264/AVC 中仅支持 8×8 大小的模块,所以就没有必要利用每一个宏块内部信息的变化剧烈程度来进行模块大小的选择。目标宏块的色度部分都是从先前编好的色度采样得到的,该色度采用可能来源于目标 8×8 色度模块的左侧或者上侧,或者左侧和上侧联合确定。也就是说色度模块中的 U 分量和 V 分量一直使用相同的预测模式。因此, 8×8 色度模块的首选预测模式一共有 2 个或 3 个,具体值可以根据先前编码的色度采用值得到。

综上所述,表 1 总结了基于快速模式选择算法的需要进行 RDO 运算的候选模式的数目。表 1 说明了基于本文算法的编码器只需要进行比标准 RDO 计算少得多的运算就可以实现基本相同的编码效果。因此,快速算法可以显著地减少编码时间。

表 1 RDO 运算的候选模式数目

Tab.1 The range of candidate modes

模块类型	H. 264	快速算法
4×4 Luma	9	3
16×16 Luma	4	2
8×8 Chroma	4	2 或 3

5 仿真结果

快速选择算法和 JVT 提供的 JM98 编解码模块整合在一起评估算法的性能。主要从编码时间、平均 PSNR 值和编码比特率 3 个方面来和标准的 RDO 模式选择算法进行比较。在仿真试验中采用的视频测试序列采用文献[6]中推荐的典型序列,并且测试视频序列中的帧都被设定为帧内预测方式进行编码,对应的量化参数设为 28。此外,设定测试序列的帧率为 30 帧/秒,两个编码的帧中间 5 帧不需要进行编码,测试序列总帧数为 300,所以最终进行帧内预测编码的帧数为 50。根据上述参数,平均编码比特率、平均 PSNR 和平均编码时间被记录下来,和标准的 RDO 算法进行比较。

此外,因为变化剧烈程度门限值的确定在快速算法中很关键,所以在仿真中使用了几个不同的门限值 D_0 ,以此来测试不同门限值对最终编码性能的影响。具体结果如表 2 所示。从表 2 可以看出, D_0 值设定的越大,编码所需的时间越短,但是 PSNR 值也会越来越低,编码比特率也会越来越高。因此,为了在减少编码时间的时候尽可能地保证较高的 PSNR 值,决定在仿真试验中设定 D_0 的值为 250。在表 3 和以后的仿真结果中的值都是基于该 D_0 值。

表 2 D_0 值对仿真结果的影响

Tab.2 Influence of D_0 on the simulation results

D_0	Δ Time(%)	Δ PSNR(dB)	Δ Bits(%)
100	-17.3	-0.13	0.0
200	-25.0	-0.25	0.9
250	-31.8	-0.29	1.2
300	-32.5	-0.33	1.5
400	-57.4	-0.51	2.9

表 3 不同测试序列的仿真结果

Tab.3 The simulation results of different test sequence

测试序列	Δ Time(%)	Δ PSNR(%)	Δ Bit-Rate(%)
News	-48.258	-0.265	3.214
Foreman	-64.356	-0.282	3.762
Bike	-59.462	-0.207	3.116
Mobile	-57.031	-0.214	3.147
Bus	-60.226	-0.203	3.608
Silent	-68.547	-0.197	3.342

从表 3 中可以看出,针对各个不同的视频测试序列,快速帧内模式选择算法平均可以比标准的 RDO 算法节约 55% 的编码时间,并且保证了可以忽略的 PSNR 的损失,与此同时最终的编码比特率的增长也不会超过 5%。因此,也证明了快速算法可以在实际应用中兼容各种不同的视频序列。

图 1、图 2 和图 3 分别显示了快速帧内预测模式选择算法的标准 RDO 算法的编码时间、平均 PSNR

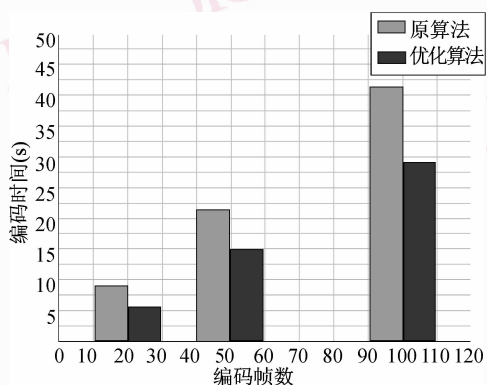


图 1 编码时间对比

Fig. 1 Comparison of encoding time

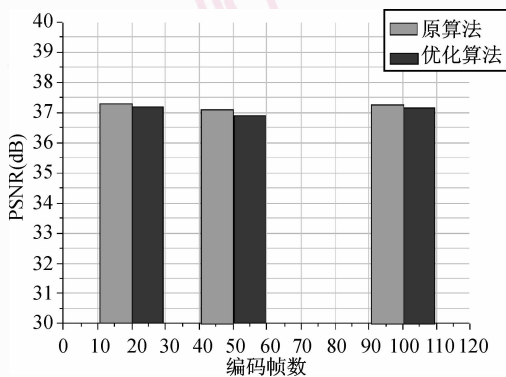


图 2 平均 PSNR 对比

Fig. 2 Comparison of the average PSNR

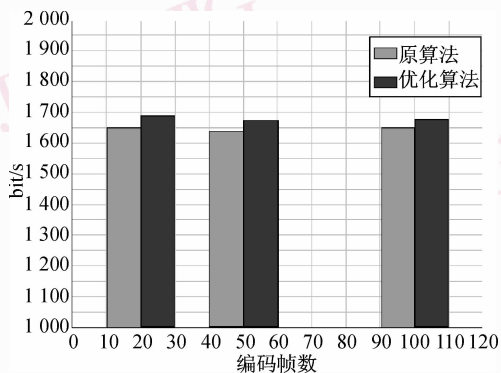


图 3 编码比特率对比

Fig. 3 Comparison of encoding bitrate

值和编码比特率的对比情况。选用的测试序列是 News。从中可以看出,快速算法和 RDO 标准算法基本达到了相同的 PSNR 值,但是却只用消耗 RDO 标准算法 40% ~ 50% 的编码时间。

6 结 论

提出了一种新型的帧内预测模式的快速选择算法。该算法在过程上分为 2 个部分。第 1 个阶段,基于目标宏块内部信息的变化剧烈程度来决定亮度部分编码模块的大小,这样就避免了一些不必要的候选模式。第 2 个阶段,算法使用方向矢量来滤除掉大部分的候选帧内预测模式,最终只剩下 2~3 个候选模式来进行 RDO 运算,从中确定最佳帧内预测模式进行编码。该算法被整合到 JM98 编码模块中,采用文献[6]中推荐的测试序列,和标准的 RDO 算法在某些性能指标上进行了对比,这些性能指标包括编码时间、平均 PSNR 值和编码比特率。仿真结果表明快速算法比标准的 RDO 算法在性能上改进很大。快速算法可以将 JVT 参考 JM98 编解码模块的运行时间在同一台主机上面减少到原来的 40% ~ 50%,而且并没有以损失视频图像质量为代价,仍然可以达到和原算法接近的 PSNR 值。最终的编码比特率也没有显著的上升。

参考文献 (References)

- 1 Wiegand T, Sullivan G J, Bjontegaard G, *et al.* Overview of the H. 264/AVC video coding standard [J]. *IEEE Transactions on Circuits and System on Video Technology*, 2003, **13**(7):1-19.
- 2 Kim Chang-sung, Li Qing, Kuo C C J. Fast Intra-prediction model selection for H. 264 codec[A]. In: *Proceedings of SPIE International Symposium ITCOM [C]*, Orlando, Florida, July, 2003, **5241**:99-110.
- 3 JVT-G013 Fast Mode Decision Algorithm for JVT Intra Prediction [S].
- 4 He Z, Zeng W, Chen C W. Low-pass filtering of rate-distortion functions for quality smoothing and bandwidth control in real-time video coding [J]. *IEEE Transactions on Circuits and System on Video Technology*, 2005, **15**(8):973-981.
- 5 ITU-T Recommendation H. 264 and ISO/IEC 14496-10 AVC, *Advanced Video Coding for Generic Audiovisual Services*[S].
- 6 ISO/IEC JTC/SC29/WG11 and ITU-T SG16 Q6, *Evaluation Sheet for Motion Estimation (Draft version 4)* [S].