

# H. 264 宏块模式的一种快速判决方案

朱 红, 吴成柯, 方 勇

(西安电子科技大学 ISN 国家重点实验室, 西安 710071)

**摘 要:** H. 264 采用了可变块运动估计和率失真优化模式判决, 极大地增加了编码器的复杂度. 本文利用 H. 264 参考软件对时间/空间相邻宏块模式之间的相关性进行了分析, 相应地提出了一种快速算法以降低模式判决的运算复杂度. 利用已编码相邻宏块的运动信息和模式类型比率信息预测当前宏块的候选模式类型, 然后采用部分率失真优化决定当前宏块的编码模式. 如果某些模式不在候选模式之列, 就不再对其做运动估计. 如果相邻宏块的平均失真超过了预定阈值, 当前宏块就采用常规的率失真优化模式判决. 实验结果表明, 对于 A、B 类测试图像序列, 采用本文方案可以明显降低编码复杂度, 而编码质量只有轻微的下降.

**关键词:** H. 264; 视频编码; 模式判决; 快速算法

**中图分类号:** TN919.81 **文献标识码:** A **文章编号:** 0372-2112 (2005) 09-1576-05

## A Novel Scheme for Fast Mode Decision within H. 264

ZHU Hong, WU Cheng-ke, FANG Yong

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

**Abstract:** H. 264 encoder complexity is remarkable due to mainly variable block size Motion Estimation (ME) and exhaustive RDO mode decision. In this paper, statistical characteristics of modes among adjacent macroblocks spatially and temporally are first studied using H. 264 reference software. Then a fast algorithm is proposed accordingly to reduce the computational load of the mode decision. Candidate modes for current macroblock are first inferred from given coded adjacent macroblocks by adopting motion information and ratios of defined mode sub-sets as inference measures. Then a partial RDO approach is applied to the candidate modes to make target mode selection. The ME is considered as a part of mode decision and no ME is required for particular modes that are eliminated in advance. The exhaustive RDO is invoked only if average distortion of the adjacent macroblocks goes beyond a predefined threshold. Simulation results demonstrate that, for image sequences of both class A and class B, the proposed algorithm can reduce encoder complexity significantly while coding quality is only slightly decreased.

**Key words:** H. 264; video coding; mode decision; fast algorithm

### 1 引言

H. 264<sup>[1,2]</sup>是适用于低速率传输的新一代视频编码国际标准,同先前的视频编码标准如 H. 263、MPEG-4 一样采用了混合编码框架,但 H. 264 吸收了原有标准的优点并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的算法(如多参考帧、多模式、1/4 像素精度的运动补偿、帧内预测编码、4 × 4 整数变换、环路滤波以及 U VLC、CABAC 熵编码等),显著提高了视频编码效率.

由于采用了相当复杂的编码技术和时间/空间预测模式, H. 264 的运算复杂度也远高于先前标准,其编码/解码复杂度分别约为 H. 263 的 3 和 2 倍. 这就限制了 H. 264 在网络中尤其是无线环境中的广泛应用,因为移动设备一般只具备很有限的运算能力. 因此,如何降低编码复杂度是 H. 264 具体实现时需要考虑的一个重要问题. 文献[3]提出了一种 UMHexagonS + CBFPS 运动估计算法,利用相邻(指时间/空间相邻,下同)宏块的相关性预测初始搜索点,并引入搜索过程早期中止机制(early termination),显著减少了匹配搜索点

数. 文献[4]提出了一种多参考帧运动估计算法,首先进行帧内预测并在前一帧中做运动估计,根据得到的结果决定是否需要在其他帧中继续做运动估计,同样采用了早期中止机制. 文献[5]根据视频对象的纹理特征和图像序列的时间平稳性决定宏块的可能模式. 文献[6]使用了 SKIP 模式优先(early SKIP mode decision)和选择性帧内模式判决(selective intra mode decision). 文献[7]基于率失真代价的不同对宏块模式进行了分类,并按一定的优先次序对各个模式类型进行测试,如果某个模式类型的编码代价小于阈值,就停止对剩余模式类型的测试. 文献[8]使用当前帧和前一帧中已编码的相邻宏块的模式预测当前宏块的模式,设当前宏块的空间坐标为  $(x, y)$ , 则其相邻宏块由当前帧中的  $(x-1, y)$ 、 $(x-1, y-1)$ 、 $(x, y-1)$  宏块以及前一帧中的  $(x+1, y+2)$  宏块组成,当前宏块的模式取代价为最小的相邻宏块的模式. 概括起来,文献中提出的运动估计或模式判决的快速算法主要体现了如下思想:基于图像序列信息(如运动矢量、编码模式、纹理特征等)的相关性,预测宏块的可能的编码模式,从而缩小搜索范围;采用早期中止机制减少不必要的运算;将模式判决和运动估计处理融为

收稿日期:2004-07-26;修回日期:2005-05-18

基金项目:国家自然科学基金(No. 60272050, 60372043)

一体,如果某些模式不在候选模式之列,就不再对其做运动估计,从而进一步降低模式判决的复杂度.类似的思想同样被用于帧内模式的快速判决<sup>[9-11]</sup>.

本文主要研究帧间编码宏块的模式判决问题.文献[7]并未考虑相邻宏块模式之间的相关性以及图像序列的运动特性.文献[8]在时间相邻宏块中仅选用了前一帧中的一个宏块,这在一般情况下显得不足.另外,基于相关预测所得的模式不一定全局最优,从而因模式误判可能降低视频编码质量,文献[8]对此并未给出解决措施.本文利用 H.264 参考软件对时间/空间相邻宏块模式之间的相关性进行了分析,相应地提出了一种新的快速算法以降低模式判决的运算复杂度.该算法利用了已编码相邻宏块组的运动度、模式类型比率信息来预测当前宏块的模式类型,然后利用部分率失真优化决定当前宏块的编码模式.

## 2 宏块模式的相关性分析

H.264 采用了树型结构运动补偿方式,宏块具有 4 种分割模式即  $16 \times 16$ 、 $16 \times 8$ 、 $8 \times 16$  和  $8 \times 8$ ,如果采用  $8 \times 8$  方式,则每个  $8 \times 8$  块又具有 4 种分割模式即  $8 \times 8$ 、 $8 \times 4$ 、 $4 \times 8$  和  $4 \times 4$ .在每种分割模式下,各个分割块都需要单独做运动估计,并返回各自的运动矢量.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,能有效提高运动估计的准确性.一般来讲,大尺寸分割块适合用于图像平坦区域,小尺寸分割块适合用于活动的图像细节区域.

H.264 采用了率失真优化 (Rate-Distortion Optimization, RDO)<sup>[12]</sup>的模式判决方法,计算并比较所有可选模式的率失真代价,代价最小的模式即为所选.率失真代价函数  $J$  定义如下:

$$J(ss, cc, \text{MODE}|QP, \text{MODE}) = \text{SSD}(ss, cc, \text{MODE}|QP) + \text{MODE} \cdot R(ss, cc, \text{MODE}|QP) \quad (1)$$

其中,  $QP$  和  $\text{MODE}$  分别是量化参数和 Lagrange 乘数.  $\text{SSD}(ss, cc, \text{MODE}|QP)$  为重建信号  $cc$  和原始信号  $ss$  的差值均方和.  $R(ss, cc, \text{MODE}|QP)$  为宏块的编码比特总数.帧间宏块模式  $\text{MODE}$  取自  $s = \{\text{SKIP}, 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4, \text{Intra}4 \times 4, \text{Intra}16 \times 16\}$ .  $\text{SKIP}$  模式是一种特殊的  $16 \times 16$  模式,其运动矢量和 DCT 信息不需要编码.  $\text{Intra}4 \times 4$  和  $\text{Intra}16 \times 16$  为帧内预测模式.计算每种模式的编码代价  $J$  时,包括运动补偿、整数变换、量化等编码过程以及相应的解码过程都需要重新进行.可想而知, RDO 模式判决的运算量非常可观.降低 RDO 模式判决运算量的途径之一是减少候选模式的数目,进而在候选模式中利用部分率失真优化 (Partial RDO, PRDO) 方法判决宏块的模式. PRDO 是一个相对的概念,它同 RDO 的区别在于候选模式的数目的不同.减少候选模式数目的措施之一是利用图像序列中相邻宏块信息之间的相关性,使用已编码宏块的信息推测当前宏块的可能模式范围.本文按照模式之间的相近程度,将模式集合划分成 4 种模式类型,  $s_1 \sim s_4$ , 如式(2)所示.  $s_1 = \{\text{SKIP}\}$ ,  $s_2 = \{16 \times 16, 16 \times 8, 8 \times 16\}$ ,  $s_3 = \{8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4\}$ ,  $s_4 = \{\text{Intra}4 \times 4, \text{Intra}16 \times 16\}$ .

$$s = s_1 \quad s_2 \quad s_3 \quad s_4 \quad (2)$$

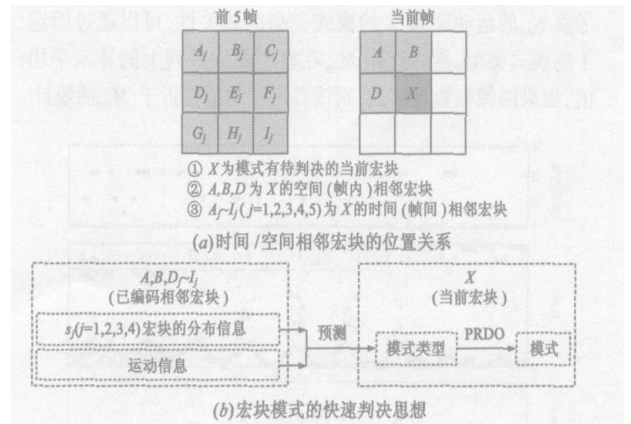


图 1 宏块编码模式的相关性分析模型

并使用如图 1 所示的宏块模式的相关性分析模型.在图 1(a)中,  $X$  表示模式有待判决的当前宏块,  $A, B, D$  表示  $X$  的已编码空间相邻宏块,  $A_j \sim I_j (j = 1, 2, 3, 4, 5)$  表示  $X$  的已编码时间相邻宏块,  $E_j$  的空间坐标同  $X$ .图 1(b)给出了  $X$  模式的快速判决设想,即根据已编码相邻宏块的运动信息和模式类型分布信息预测  $X$  的模式类型,进而对所得模式类型使用 PRDO 判决法以决定  $X$  的最终模式.如果  $X$  处于图像的中央区域,则所有相邻宏块均可用;否则某些相邻宏块可能位于图像之外而不可用,本文的处理方式是忽略这些宏块.根据  $X$  是否位于图像的中央区域还是不同的边缘区域,可以算出  $X$  的可用的时间/空间相邻宏块的数目取自  $\{20, 21, 23, 31, 33, 48\}$ .以下把可用相邻宏块的集合称为  $X$  的相邻宏块组,用表示,以下给出的“模式类型比率”和“运动度”的计算方法.  $S_{\text{ref}}$  中各模式类型宏块的比率

设  $N_{\text{ref}}$  表示  $S_{\text{ref}}$  中的宏块总数,  $N_{\text{ref},j}$  表示  $S_{\text{ref}}$  中模式类型为  $s_j$  的宏块总数,于是有式(3)成立,  $S_{\text{ref}}$  中各模式类型的比率  $R_{\text{ref},j}$  的计算如式(4).

$$N_{\text{ref}} = \sum_{j=1}^4 N_{\text{ref},j} \quad (3)$$

$$R_{\text{ref},j} = N_{\text{ref},j} / N_{\text{ref}}, j = 1, 2, 3, 4 \quad (4)$$

为了在统计意义上考察  $S_{\text{ref}}$  中各模式类型的比率同  $X$  的模式类型的相关性,可以通过固定  $X$  的模式类型,然后计算  $R_{\text{ref},j}$  在整个图像序列上的算术平均值,如果图像帧数足够大,则该算术平均值接近于  $R_{\text{ref},j}$  的统计平均值,用  $R(j, i)$  表示:

$$R(j, i) = E[R_{\text{ref},j} | \text{mode}(X) = s_i] \quad (5)$$

$S_{\text{ref}}$  的运动度

设  $\text{motion}(\text{MB})$  表示宏块 MB 的运动度,运动度越大表示运动越剧烈.规定当  $\text{mode}(\text{MB}) = s_1$  或  $\text{mode}(\text{MB}) = s_4$  时  $\text{motion}(\text{MB}) = 0$ ,而当  $\text{mode}(\text{MB}) = s_2$  或  $\text{mode}(\text{MB}) = s_3$  时,  $\text{motion}(\text{MB}) = \frac{1}{K} \sum_{k=1}^K \sqrt{MX_k^2 + MY_k^2}$ , 其中  $(MX_k, MY_k)$  为宏块的第  $k$  个分割子块的运动矢量,  $1 \leq k \leq K$ ,  $K$  为运动矢量总数.  $S_{\text{ref}}$  的运动度  $M_{\text{ref}}$  计算如下:

$$M_{\text{ref}} = \frac{1}{N_{\text{ref}}} \sum_{n=1}^{N_{\text{ref}}} \text{motion}(\text{MB}_{\text{ref},n}); \text{MB}_{\text{ref},n} \in S_{\text{ref}} \quad (6)$$

其中,下标  $n$  表示  $S_{\text{ref}}$  中宏块的序号.为了在统计意义上

考察  $S_{ref}$  的运动度同  $X$  的模式类型的相关性, 可以通过固定  $X$  的模式类型, 然后计算  $M_{ref}$  在整个图像序列上的算术平均值, 如果图像帧数足够大, 则该算术平均值接近于  $M_{ref}$  的统计

平均值, 用  $M(i)$  表示:

$$M(i) = E[M_{ref} | mode(X) = S_i] \quad (7)$$

实验采用的测试图像序列以及 H. 264 参考软件 JM 8. 1a<sup>[13]</sup> 的主要控制参数如表 1 所示, 对 H. 264 编码结果的分析数据由图 2、3 给出. 图 2 给出了 A 类序列 Akiyo、B 类序列 Coastguard 和 C 类序列 Mobile 的第 6 帧宏块 (作为当前宏块) 的模式类型同其  $S_{ref}$  的运动度、模式类型比率的关系. 图 3 则给出了各个图像序列的  $R(j, i)$  和  $M(i)$  数据.

表 1 实验采用的图像序列以及编码器主要控制参数

实验采用的图像序列		编码器主要控制参数
序列名称	序列特性	
Akiyo	A 类图像序列: 空间细节较少, 活动性较小	* 编码 150 帧, A、B 类图像序列的编码帧频设为 15 fps, QCIF 格式; C 类图像序列的编码帧频为 25 fps, SIF 格式; * IPPP 序列编码方式, QP 取 28; * 运动估计参考帧数目取 5, 运动估计最大搜索范围取 16, 允许所有 7 种宏块分割模式, 允许 RDO 模式判决; * 允许环路滤波, 允许哈达马变换, CABAC 熵编码;
Container		
Hall		
Mother		
Coastguard	B 类图像序列: 空间细节适中, 活动性较小; 或者空间细节较少, 活动性适中	
Foreman		
News		
Silent	C 类图像序列: 空间细节较多, 活动性适中; 或者空间细节适中, 活动性较大	
Football		
Mobile		

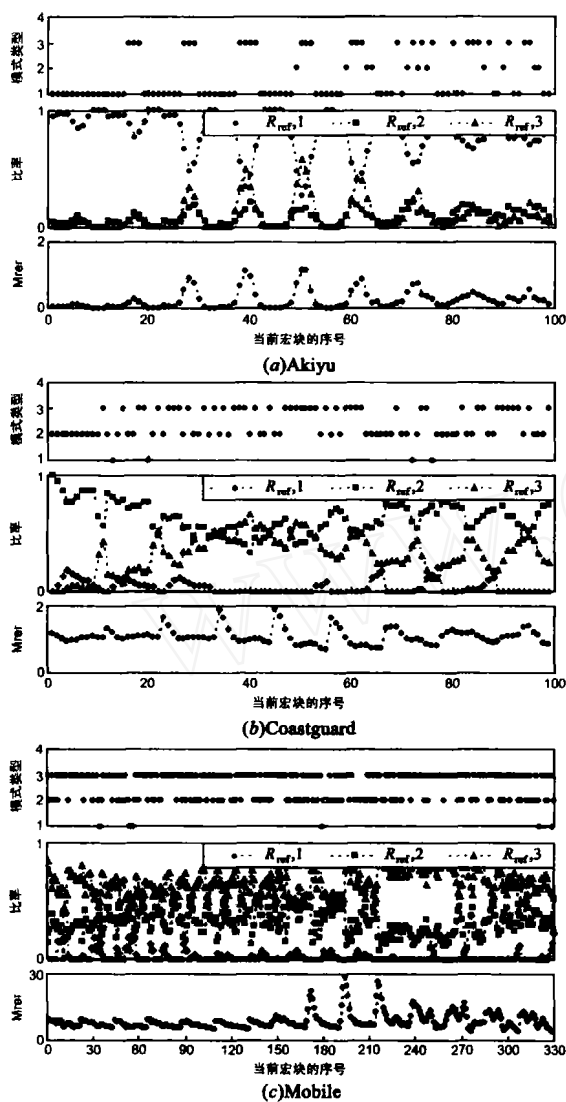


图 2 图像序列第 6 帧宏块的模式类型同  $R_{ref,j}$ 、 $M_{ref}$  的关系

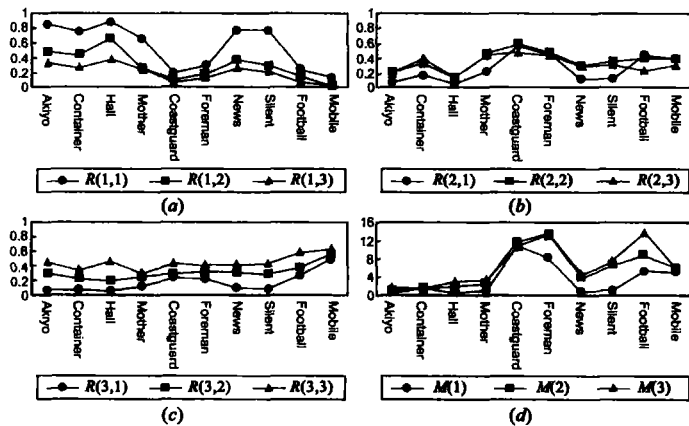


图 3 图像序列的  $R(j,i)$  和  $M(i)$  数据

从图 2(a) 可以看出, 对 Akiyo、Coastguard 和 Mobile, 均存在如下规律:  $R_{ref,j}$  较大时, 当前宏块的模式类型取  $S_j$  的几率也较大 ( $j=1, 2, 3$ ), 即当前宏块的模式类型取  $S_{ref}$  中比率占优势的模式类型的几率较大; 当  $M_{ref}$  较大时, 当前宏块的模式类型取  $s_2$  和  $s_3$  的几率较大, 即如果  $S_{ref}$  的运动度较大, 则当前宏块趋向于采用较小分割块模式. 图 3 进一步从统计意义上印证了上述规律的正确性.

从图 3(a) 可以看出, 对于所有测试序列都有  $R(1,1) > R(1,2)$  和  $R(1,1) > R(1,3)$ , 可见, 如果  $s_1$  为  $S_{ref}$  中比率占优势的模式类型, 则当前宏块的模式类型一般也取  $s_1$ . 从图 3(b) 可以看出, 对 Akiyo、Mother、Coastguard、Foreman、Silent、Mobile 有  $R(2,2) > R(2,1)$  和  $R(2,2) > R(2,3)$ . 可见, 如果  $s_2$  为  $S_{ref}$  中比率占优势的模式类型, 则当前宏块的模式类型一般也取  $s_2$ . 该规律对于 Football 而言是个例外, 对 Football 有  $R(2,2) < R(2,1)$  和  $R(2,2) > R(2,3)$ , 这同该序列运动较为剧烈从而相邻宏块模式类型之间的相关性较弱有关. 对 Container、Hall 和 News 而言, 都有  $R(2,2) > R(2,1)$ , 即如果  $s_2$  为  $S_{ref}$  中比率占优势的模式类型, 则当前宏块模式类型取  $s_1$  的可能较小, 但由于存在  $R(2,2) > R(2,3)$ , 所以当前宏块模式类型取  $s_2$  或  $s_3$  还存在不确定性. 从图 3(c) 可以看出, 对于所有测试序列都有  $R(3,3) > R(3,1)$  和  $R(3,3) > R(3,2)$ , 可见如果  $s_3$  为  $S_{ref}$  中比率占优势的模式类型, 则当前宏块的模式类型一般也取  $s_3$ . 从图 3(d) 可以看出, 对 Akiyo、Hall、Mother、News、Silent、Football、Mobile 有  $E(3) > E(2) > E(1)$ , 可见当  $S_{ref}$  的运动度较大时, 当前宏块的模

式类型趋向于采用  $s_3$ , 该规律严格上对 Container、Coastguard 和 Foreman 是个例外, 但对该 3 个序列都有  $E(2) > E(1)$  和  $E(2) > E(3)$ , 也即该 3 个序列的当前宏块的模式类型取  $s_2$  的可能性较大。

本文并未给出与模式类型  $s_4$  有关的相关性测试, 这是因为对于 A 类和 B 类序列的  $P$  帧, Intra (帧内编码) 宏块所占比例很小。例如, 对于 Akiyo、Container、Hall 和 Coastguard, 平均单个  $P$  帧中 Intra 宏块所占比例几乎为 0, 对于 Mother、Foreman、News 和 Silent, 该比例依次为 0.149%、0.014%、0.068% 和 0.129%。所以同 Inter (帧间编码) 宏块相比, Intra 宏块的运算量所占比例较小。C 类序列中的 Football 的 Intra 宏块的比例则较大, 达 11.09%, 这同 Football 的运动较为剧烈有关。而同为 C 类序列的 Mobile, Intra 宏块的比例约为 0.165%, 比 Football 小得多, 但仍大于 A、B 类序列, 这是因为 Mobile 的运动虽然比较平缓, 但运动方式较为复杂的缘故。当然, 采用 Intra 编码方式能有效地阻止视频质量的下降, 虽然编码比特会或多或少增加一些。本文提出的快速模式判决算法主要利用了宏块信息间的相关性。一般地讲, 细节越少或运动越小越简单, 宏块间的相关性越强, 快速算法的效果就越明显。鉴于此, 本文快速模式判决算法对 A 类序列将很明显, 对 B 类序列次之, 对 C 类序列较弱, 后文的实验结果也证实了这一点。

### 3 快速模式判决方法

通过上述对图 2 和图 3 的分析和讨论, 可以得出如下规律: 事件“ $R(i, i) > R(j, i); i=1, 2, 3; j=1, 2, 3; j \neq i$ ”出现的几率较大, 这个结果体现了相邻宏块模式类型之间存在一定的相关性, 即当前宏块的模式类型一般取  $S_{ref}$  中比率占优势的模式类型; 事件“ $M(3) > M(1); M(2) > M(1)$ ”出现的几率也较大, 这个结果体现了相邻宏块在运动度和模式类型之间也存在一定的相关性, 即如果  $S_{ref}$  的运动度较大, 则当前宏块的模式类型一般选择  $S_3$  或  $S_2$ 。从中, 受到如下启发:

利用已编码相邻宏块的模式类型比率信息可以预测当前宏块的模式类型。也即在判决当前宏块模式之前, 首先考察其  $S_{ref}$  中三种模式类型的比率 ( $R_{ref}; j=1, 2, 3$ ), 然后选择比率取最大值的模式类型作为当前宏块的候选模式类型, 从而减小模式判决的范围。

利用已编码相邻宏块的运动信息可以预测当前宏块的模式类型。即在判决当前宏块模式之前, 应考察其  $S_{ref}$  的运动度大小。虽然从统计数据上看,  $S_{ref}$  运动度的大小并不严格决定当前宏块的模式类型是取  $s_2$  还是  $s_3$ , 但是当运动度较大时, 采用  $s_3$  要比  $s_2$  更能保证运动估计的准确性。所以可做如下规定, 如果运动度较大, 则选择  $s_3$  作为当前宏块的候选模式类型, 如果运动度适中, 则选择  $s_2$  作为候选模式类型。这种运动度判决方法可作为上述模式类型比率判决方法的有益补充, 以提高模式判决的准确性。

利用宏块相关性的快速模式判决方法虽然能够减少 RDO 模式判决的候选模式数目, 但不一定能保证所选模式的全局最优, 所以存在模式误判的可能。在上述分析中, 并未考虑  $S_{ref}$  中模式类型  $s_4$  的比率对判决当前宏块模式类型的影

响, 因为对多数测试序列来说,  $s_4$  编码宏块所占比例都非常小。然而, 采用  $s_4$  模式类型的 Intra 编码具有较 Inter 编码高的视频质量。鉴于此, 可做如下规定, 如果  $S_{ref}$  的平均失真超出了预定阈值, 就仍采用常规的 RDO 方法判决当前宏块的模式, 此时无疑增加了模式类型  $s_4$  的选中几率, 从而可以避免视频编码质量的下降。

基于上述分析, 提出了一种流程如图 4 所示的快速模式判决方法。其中,  $D_{REF}$  表示  $S_{ref}$  的平均失真, 如式 (8) 定义。其中, SSD 和  $N_{ref}$  的含义分别同式 (1) 和 (3) 中的一致, 下标  $n$  表示  $S_{ref}$  中宏块的序号。  $D_{REF}$  的引入主要是为了复位模式判决算法的状态以避免因可能的模式误判而导致的视频质量下降, 即如果失真的积累导致超出了预定阈值  $D_{th}$ , 就执行进程跳转, 采用完全 RDO 方式来决定当前宏块的模式, 此时无疑会增加选择 Intra 编码模式的几率, 同时编码

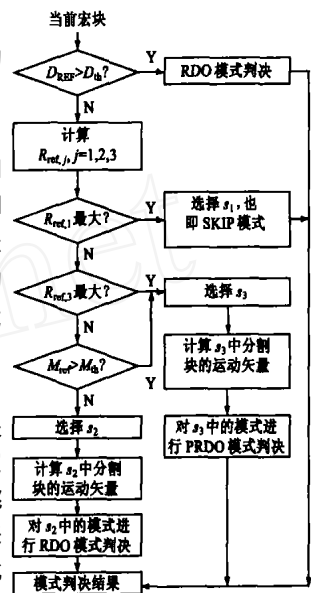


图 4 宏块模式的快速判决流程图

比特也会有所增加。测试表明, 对于本文使用 A、B 类测试序列而言, 上述进程跳转的发生次数很少, 所以在总的编码效率方面的损失也很小。在快速模式判决过程中, 运动度判决方法被用作一种辅助措施, 同模式类型比率判决方法一起, 用于在模式类型  $s_2$  和  $s_3$  之间作出选择。通过对运动度设置阈值  $M_{th}$  来实现, 即如果  $S_{ref}$  的运动度较大, 则选择模式类型  $s_3$  作为当前宏块的候选模式类型  $s_2$ , 否则  $s_3$  就选择模式类型  $s_2$ 。由于中的模式对宏块分割得更为精细, 所以当运动度较大时, 采用  $s_3$  要比  $s_2$  更能保证运动估计的准确性。在 H.264 参考软件中, 运动估计和模式判决两个环节是分开进行的, 参考文献 [7] 中将二者集成到了一起。也就是说, 如果某些模式已经被模式判决算法预先排除, 就不再对其进行运动估计, 于是可以显著减少模式判决的运算量。本文同样采用了这种处理方式。

$$D_{REF} = \frac{1}{N_{ref}} \sum_{n=1}^{N_{ref}} SSD(ss_n, cc_n, MODE_n | QP) \quad (8)$$

### 4 实验结果和分析

通过修改 H.264 JM 8.1a 参考软件实现本文提出的快速模式判决算法, 然后同参考软件进行性能比较。使用编码时间和亮度分量的平均峰值信噪比 (PSNR) 作为评价指标。计算机平台为 Intel Celeron 900MHz CPU, 256M RAM, Windows XP 操作系统。仍采用如表 1 所列测试图像序列和编码控制参数。对 A、B 和 C 类测试序列, 阈值 ( $D_{th}$ ,  $M_{th}$ ) 分别取 (35.22, 0.75)、(34.09, 2.81) 和 (33.31, 8.89), 均是通过实验方法确定的。编码时间、编码质量的比较结果如表 2 所示。

从中可以看出,同参考软件相比,采用本文方案可以明显降低编码器的运算复杂度.平均来讲,对 A、B、C 类序列分别可以减少 44.9%、33.3%、18.5% 的编码时间,同时编码质量分别有 0.07dB、0.12dB、0.4dB 的下降.可见,本文方案对于 A 类序列比较有效,而对 B 类序列次之,对 C 类序列则较弱.原因分析:本文的模式判决算法主要利用了宏块间编码信息的相关性,也即根据已编码相邻宏块组  $S_{ref}$  的运动度和模式类型比率的计算结果,预测当前宏块的可能的模式类型.一般来讲,图像序列的细节越少或运动越小越简单,宏块间的相关性就越强,预测效果就越好,从而快速算法的效果就越明显;反之,快速算法的效果就越差.

表 2 本文方案和 H.264 JM 8.1a 参考软件的实验结果比较

测试序列	编码时间(秒)				编码质量(PSNR, dB)			
	参考软件	本文方案	节省时间	平均节省时间	参考软件	本文方案	增益	平均增益
Akiyo	361.4	187.2	48.2%	44.9%	37.30	37.26	-0.04	-0.07
Container	378.2	208.8	44.8%		35.96	35.90	-0.06	
Hall	390.0	220.0	43.6%		37.21	37.14	-0.07	
Mother	371.2	211.2	43.1%		36.25	36.16	-0.09	
Coastguard	425.4	291.4	31.5%	33.3%	34.05	33.92	-0.13	-0.12
Foreman	383.9	258.0	32.8%		35.52	35.40	-0.12	
News	377.3	245.6	34.9%		36.68	36.57	-0.11	
Silent	381.7	251.5	34.1%		35.81	35.71	-0.10	
Football	1730.1	1366.8	21.0%	18.5%	33.37	33.04	-0.33	-0.40
Mobile	1750.8	1470.7	16.0%		33.25	32.77	-0.48	

为进一步验证本文方案的先进性,我们还同参考文献[8]中的模式判决方法进行了性能比较,仍采用了如表 1 所列图像序列和 H.264 编码控制参数.结果表明,采用<sup>[8]</sup>方案时的平均编码时间比本文方案还可进一步缩短 23%.而本文方案的重建视频质量比<sup>[8]</sup>方案平均提高了约 0.44~1.35dB.原因分析如下,文献[8]方案对当前宏块的模式进行预测时使用了较少的时间/空间相邻宏块,并使用了较简单的预测代价判决方法,所以运算复杂度较低,而模式判决的准确性也降低了;本文方案则使用了较多的时间/空间相邻宏块和较为复杂的运动度和模式类型比率预测法,所以运算复杂度较高,但模式判决的准确性也较高;本文方案考虑了因模式误判可能导致的质量下降,规定当  $S_{ref}$  的失真超过预定阈值时,就采用常规的 RDO 方法来判决当前宏块的模式,所以在提高模式判决速度的同时还可以保证视频编码质量,这种处理在文献[8]方案中并未考虑.

## 5 结论

本文采用 H.264 参考软件对时间/空间相邻宏块编码模式之间的相关性进行了分析,相应提出了一种新的模式判决快速算法.该算法利用已编码相邻宏块组的运动度和模式类型比率信息来预测当前宏块的模式类型,从而减小了模式判决的候选模式的数目,提高了模式判决的速度.利用宏块信息间的相关性进行模式预测是本文算法的核心,这种相关性越强,本文算法的优势越明显.但这种相关性受图像序列的空间细节、活动性、运动复杂程度等因素的影响较大,这也是本文

方案对于 A 类序列最有效、对 B 类序列次之、对 C 类序列较弱的主要原因.

## 参考文献:

- [1] ITU-T. Advanced video coding for generic audiovisual services [S]. Recommendation H.264, May 2003.
- [2] Thomas Wiegand, et al. Overview of the H.264 video coding standard [J]. IEEE Transactions on circuits and systems for video technology, 2003, 13(7): 1 - 19.
- [3] Zhibo Chen, et al. Fast motion estimation for JVT[A]. JVT-G016 Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), 7th Meeting: Pattaya II (C). Thailand: JVT-G016, March 2003.
- [4] Yur-Wen Huang, et al. Analysis and reduction of reference frames for motion estimation in MPEG-4 AVC/JVT/H.264 [A]. 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03) [C]. Hong Kong, China: ICASSP, 2003.
- [5] D Wu, et al. Block inter mode decision for fast encoding of H.264 [A]. ICASSP [C]. Montreal, Quebec, Canada: ICASSP, 2004.
- [6] Byeungwoo Jeon. Fast mode decision for H.264 [A]. JVT-J033, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6) [C]. 10th Meeting: Waikoloa, Hawaii, USA: JVT-J033, December 2003.
- [7] Jeyun Lee, Byeungwoo Jeon. Fast mode decision for H.264 with variable motion block sizes [C]. ISICIS 2003, Antalya, Turkey, 723 - 730.
- [8] A Ahmad, et al. Selection of variable block sizes in H.264 [A]. Quebec, Canada: ICASSP 2004, 319 - 321.
- [9] Bojun Meng, Oscar C. Au. Fast intra-prediction mode selection for 4x4 blocks in H.264 [A]. Hong Kong, China: ICASSP 2003, 389 - 392.
- [10] Bojun Meng, Oscar C. Au. Efficient intra-prediction algorithm in H.264 [A]. Spain: ICIP 2003, 837 - 840.
- [11] Changsung Kim, et al. Feature-based intra-prediction mode decision for H.264 [A]. Singapore: ICIP, 2004, 769 - 772.
- [12] Heiko Schwarz, Thomas Wiegand. An improved H.26L coder using Lagrangian coder control [A]. ITU-T/SG16/Q6/VCEG-D146 [C]. Porto Seguro, Brasil: ITU-T, 2001.
- [13] Download online: ftp://ftp.intc-files.org/jvt-experts/reference-software [Z].

## 作者简介:



朱红男, 1971 年 10 月出生于河北涿鹿, 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. E-mail: zhuehong@cast504.com.

吴成柯男, 1938 年 8 月出生于安徽黄山, 1962 年毕业于原西北电讯工程学院无线电通信工程系, 现为西安电子科技大学教授, 博士生导师, 1980 年至 1982 年在美国宾夕法尼亚大学做访问学者, 1990 年至 1991 年在法国南锡大学做客座教授, 目前主要从事计算机视觉、计算机图形学、图像处理和图像通信的教学和科研工作.