

一种适用于 H. 264/AVC 的自适应空域帧内预测算法

倪 伟 郭宝龙 杨 镠

(西安电子科技大学机电工程学院智能控制与图像工程研究所 西安 710071)

摘 要 提出了帧内预测算法中邻块间预测选项的相关性、预测选项和 SAD 相关性的概念,在此基础上提出了一种适用于 H. 264/AVC 的自适应空域帧内预测算法,该算法利用上述的两种相关性预测当前块的帧内预测选项;利用“半步停”技术自适应地中止预测.实验结果表明,该算法在保证重构图像质量和压缩比的同时,计算复杂度仅相当于原有算法的 7% 左右.

关键词 H. 264/AVC; 视频编码; 帧内预测; 预测模式; 半步停

中图法分类号 TP391

An Adaptive Intra Prediction Algorithm in Spatial Domain for H. 264/AVC

NI Wei GUO Bao-Long YANG Liu

(Institute of Intelligent Control and Image Engineering, School of Electromechanical Engineering, Xidian University, Xi'an 710071)

Abstract Intra prediction has been adopted in some video coding standards as an effective solution to increase the coding efficiency of I frame. In H. 264/AVC, the current block is predicted in spatial domain, which has a better performance in regard to the traditional prediction methods in frequency domain, but its computation complexity is much larger. By great number of experiments, authors found that the prediction modes of spatial adjacent blocks in H. 264/AVC are highly correlated, and this kind of correlation can be well represented by the SAD value. Based on these two characteristics, a novel fast intra prediction algorithm in spatial domain is presented in this paper. It utilizes the above characteristics to predict the intra prediction mode of the current block, and use the “half-way stop” technique to terminate the process of intra prediction adaptively. Experimental results show that with little degradation in the coding gain, the computational load of the proposed algorithm can achieve about 7% of original algorithm.

Keywords H. 264/AVC; video coding; intra prediction; prediction mode; half-way stop

1 引 言

近年来多媒体通信技术得到了广泛的应用与发展,国际电信联盟 (ITU) 和国际标准化组织 (ISO) 相继制定了一系列视频编码标准 (如 H. 261、H. 263^[1]、MPEG-1/2/4^[2]). 为了适应不断发展的视频流多媒体通信的需要,ISO/IEC 和 ITU 组成了联合视频组 (Joint Video Team) 共同开展了下一代视频编码标

准 H. 26L 的研究工作. 新标准于 2003 年 3 月获得正式通过,ITU 将该算法称为“H. 264 建议”^[3] (Recommendation H. 264),ISO 则称之为“ISO/IEC 14496 10 高级视频编码算法” (ISO/IEC 14496 10 AVC).

H. 264/AVC 标准仍然基于经典的运动补偿混合变换编码算法,在此基础上采用了一系列先进的编解码技术,如帧内预测、整数变换、高精度运动估计算法、基于上下文的自适应二进制算术编码 (CABAC)、多参考帧运动估计等,使得 H. 264/AVC 在所有码

收稿日期:2003-10-28;修改稿收到日期:2005-07-14. 本课题得到教育部优秀青年教师资助计划项目基金(2003)资助.倪 伟,男,1981年生,博士研究生,研究方向为视频编码、多媒体通信、图像处理. E-mail: weini@mail. xidian. edu. cn. 郭宝龙,男,1962年生,教授,博士生导师,主要研究领域为图像处理与图像通信、智能信息处理、神经网络等. 杨 镠,女,1981年生,博士研究生,研究方向为图像处理和虚拟现实.

率条件下都能够提供高质量的视频图像。H. 264 还采用了数据分区、数据掩盖和错误恢复等技术使之具备了在高误码率、丢包多发的信道中传输的能力。实验证明,与 H. 263+和 MPEG-4 简单类(simple profile)相比,在视频图像质量相同的情况下,H. 264/AVC 能够节省 50%左右的码率。

帧内预测技术作为提高 I 帧编码效率的有效手段,在 H. 263+[4]中被首次采纳,MPEG-4 的纹理编码中也使用了该项技术。H. 263+和 MPEG-4 的帧内预测基于 DCT 域,使用左方和上方子块的频域系数对当前子块进行预测,编码时只对预测值和实际值之间的残差进行量化编码,能降低编码后 I 帧所占用的比特数。H. 264/AVC 的帧内预测则基于子块在空间域的相关性进行预测,预测精度较传统的频域帧内预测算法有显著提高,但其庞大的运算量限制了在实时视频领域的应用。本文通过大量的实验和详细分析,得出了预测选项的空间相关性、预测选项和子块 SAD 值之间的关系,并结合“半步停”(half-way stop)思想,提出了一种新的适用于 H. 264/AVC 的自适应空域帧内预测算法。该算法能够有效地降低帧内预测的运算复杂度,同时仍然能够维持较好的图像质量和压缩性能。

本文第 2 节简要介绍了 H. 264 标准的帧内编码原理;第 3 节论述了 H. 264 的空域帧内预测算法,并分析其优缺点;第 4 节中首先证明了相邻子块的预测选项是高度相关的,并在此基础上提出了自适应帧内预测算法;最后给出实验结果,证明新算法的有效性。

2 H. 264/AVC 帧内编码原理

自然场景图像可被近似看作 Gauss-Markov 统计特性的信号,DCT 对此类图像具有近似最优的去相关能力,虽然在压缩效率上略逊于 K-L 变换,但是其可以做到的高效处理性却是 K-L 变换所无法比拟的,因而绝大多数视频编码标准都是采用基于 DCT 的混合变换编码算法,即帧内编码时通过 DCT 去除图像内部的空间冗余信息,帧间编码时通过运动估计/运动补偿技术去除帧与帧之间的时间冗余。

通常采用该种算法编码后,P 帧和 B 帧会得到有效的压缩。由于编码时将图像分隔成为固定尺寸(如 8×8 像素)且相互独立的子块作为编解码的最小单元,图像的全局性必然使得图像内部的子块和子块间存在较强的相关性。由于没有利用子块间的此种相关性,且不能利用帧间的相关性,因而仅使用

DCT 变换后量化编码,I 帧仍会占用较多的比特数。

为了提高 I 帧的压缩率,H. 263+首次采用了帧内预测技术,ISO 的 MPEG-4 标准也使用了称为自适应 DC/AC 预测的帧内预测算法,以上两种算法都是基于频域系数进行预测的。H. 264/AVC 基于图像在空域中的原始灰度值进行预测^[5,6]。H. 264/AVC 帧内编码原理如图 1 所示,从中可以看出,当图像信号输入编码器进行编码时,首先根据帧存储器中的已编码块进行帧内预测,编码时只对实际值和预测值之间的差值进行量化和编码。同时一路码流经过反量化和反变换之后存入帧存储器,作为以后宏块帧内预测的参考块。而所使用的预测模式也要作为必要的语法元素进行熵编码。解码是编码的逆过程,解码器首先根据输入的压缩码流的语法进行判断,如果为帧内编码,首先进行反量化和反 DCT 变换,得到当前帧的残差图像,再根据码流中的预测模式和帧存储器中的信息计算出当前子块的预测值,将残差和预测值相加,便可重构当前子块。

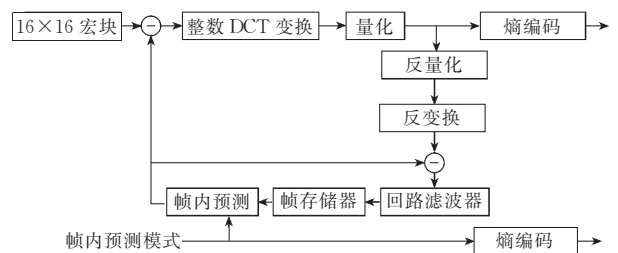


图 1 H. 264 帧内编码原理图

显然,帧内预测越准确意味着预测值和实际值之间的残差越小,编码后 I 帧占用的码字也就越少。如何设计最优的帧内预测算法,使得其预测值能够尽可能地逼近实际值是 H. 264/AVC 视频编码算法所需要解决的重要问题。

3 H. 264/AVC 帧内预测算法

自然场景图像中的前景和背景通常具有一定的纹理特性,按其方向性可以划分为水平纹理、垂直纹理和倾斜纹理等。在将图像分隔成为相互独立的子块后,图像中的物体一般位于相邻区域的数个或数十个子块中。正是由于物体对象的这种全局性,造成了当前块与其邻块的纹理方向往往是高度一致的,且各像素值之间相差不大。即不但子块内部的像素具有空间冗余特性,子块与子块间也存在空间冗余。H. 264/AVC 的帧内预测算法正是基于像素点在空间域的方向特性及子块像素间的相关性进行预测,去除子块间的空间冗余。预测时使用已经编码的子块(尤其是当前块上方和左方子块)的像素值来预

测当前块亮度分量和色度分量值. H. 264/AVC 对亮度分量进行帧内预测时包含两种预测模式: 基于 4×4 像素块和基于 16×16 像素块的帧内预测; 对色度分量只有一种预测选项. 每种预测模式设定了不同方向的预测选项 (prediction mode) 以尽可能地对应不同纹理特性的图像子块.

3.1 4×4 块亮度分量帧内预测模式

当图像区域中包含的细节较为丰富时, 此时子块内各像素点有一定差别, 将图像分割为通常的 8×8 像素块进行预测产生的误差比较大, 因此将图

像分为更小的子块进行预测是一种好的选择. H. 264/AVC 支持小至 4×4 像素块的预测, 具有更高的预测精度. 在该模式中, 预测以宏块中的 4×4 块为单位进行, 共有 9 种预测选项. Mode0 为 DC 预测, Mode1~Mode8 分别代表了 8 种不同方向的预测选项 (如图 2 所示), 块内同一方向的像素点具有相同的预测值, 以此来近似地逼近不同方向纹理特性的图像.

各预测选项的计算方法如表 1 所示, 块内各像素点的具体计算公式参见文献[3].

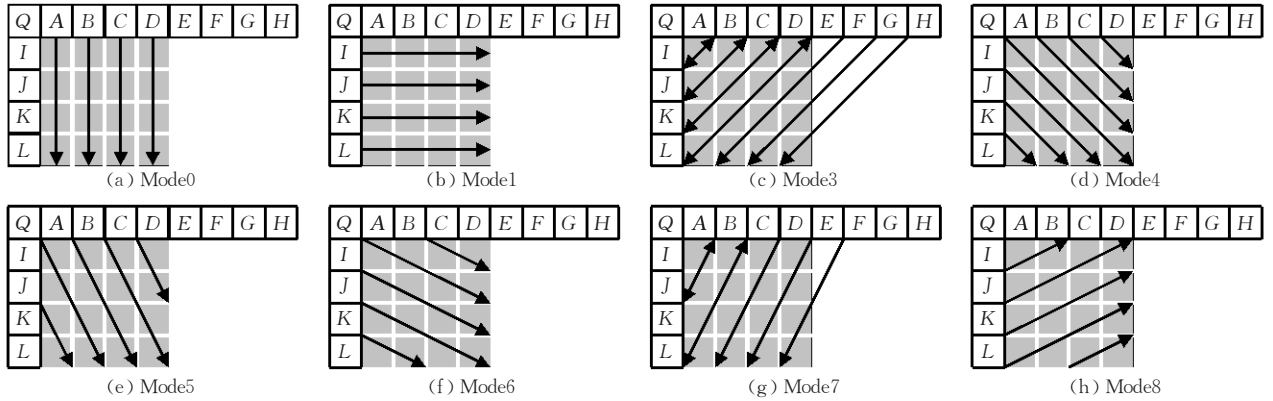


图 2 4×4 子块帧内预测选项方向示意图

表 1 4×4 块亮度分量帧内预测算法

| 序号 | 预测选项名称 | 预测算法 |
|--------|----------|---------------------------------|
| Mode 0 | 垂直预测 | 通过当前预测块正上方的相邻系数进行预测 |
| Mode 1 | 水平预测 | 通过当前预测块正左方的相邻系数进行预测 |
| Mode 2 | DC 预测 | $(A+B+C+D+E+F+G+H)/8$ |
| Mode 3 | 左对角预测 | 以 45° 角的方向预测当前块 |
| Mode 4 | 右对角预测 | 以 -45° 角的方向预测当前块 |
| Mode 5 | 垂直/右对角预测 | 以 $270^\circ + \alpha$ 的方向预测当前块 |
| Mode 6 | 水平/右对角预测 | 以 $-\alpha$ 的方向预测当前块 |
| Mode 7 | 垂直/左对角预测 | 以 $270^\circ - \alpha$ 的方向预测当前块 |
| Mode 8 | 水平/左对角预测 | 以 α 的方向预测当前块 |

当采用 4×4 块亮度分量帧内预测模式时, 依次使用上述的 9 种预测选项, 选取相应的相邻像素点进行预测. 其中 $\alpha = \text{tg}^{-1} 1/2$. 对于这 9 种预测选项, 通常使用绝对误差和准则 (SAD) 作为判决函数来比较预测选项的优劣. 编码时选取最接近当前子块实际值的预测结果即最优值, 并对残差和预测选项进行编码.

3.2 基于 16×16 块的帧内预测模式

在图像区域内部较为平坦、包含的图像细节不多的情况下, 可以对 16×16 的像素块直接进行帧内预测, 无需进一步分块, 从而进一步减少帧内预测的计算量. H. 264/AVC 针对 16×16 的帧内预测模式设定了 4 种预测选项, 分别为垂直预测、水平预测、DC 预测和平面 (Plane) 预测, 其预测原理如图 3 所示.

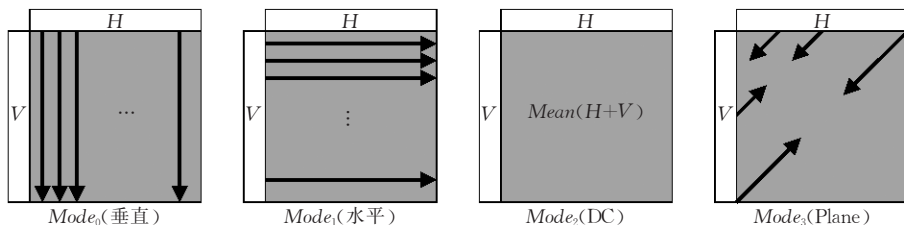


图 3 16×16 子块的帧内预测选项示意图

3.3 色度分量的帧内预测模式

除了上面所讨论的针对亮度分量所进行的帧内

预测外, H. 264/AVC 对于色度分量也进行单独的帧内预测. 人眼系统视觉特性 (HVS) 决定了人眼对

色度信号不如亮度信号敏感,因而在编码时色度分量一般采用较为粗糙的采样和量化方式.基于类似的思想,色度分量的帧内预测也采用了精度较低的方法.其帧内预测模式是以 8×8 块为单位进行的,只有一种预测模式.如图4所示,一个 8×8 像素的色度块被分为4个 4×4 像素的色度块A,B,C,D,S1,S2,S3,S4分别是与A,B,C,D相邻的4个像素色度值之和.根据当前 8×8 色度块在图像中的位置,在S1,S2,S3,S4中选取合适的样本进行预测.与亮度分量的预测模式不同,色度分量预测中 4×4 像素的色度块内部各采样点具有相同的预测值.

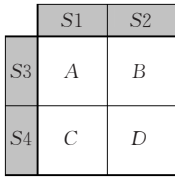


图4 色度分量的帧内预测模式

4 自适应空域帧内预测算法

H.264/AVC 设定了基于 4×4 像素块和基于 16×16 像素块的帧内预测,为了保证对不同纹理方向图像的预测精度,在各预测模式中又进一步定义了不同方向的预测选项,在 4×4 块的帧内预测模式中更是多达9种预测选项.对于每个 4×4 亮度子块,预测时都要依次使用9种预测选项进行预测,即必须进行 $4 \times 4 \times 9 = 144$ 种预测运算,得出相应的SAD值,再根据SAD值确定最优预测选项.此种帧内预测方式能有效地提高预测图像的精度,但是过多的预测选项也不可避免地使得帧内编码的运算量大大增加,从而降低了编解码的实时性.

事实上,ITU也注意到这个问题.在ITU给出的H.264/AVC验证模型JM中,其帧内预测算法在 4×4 块亮度分量帧内预测模式中并没有使用标

准草案中的全部9种预测选项,而是选取了其中使用频率最高、效果最好的5种预测选项,组成一种简化的帧内预测模式,我们称之为5阶帧内预测算法.5阶帧内预测算法包含DC预测、垂直/对角预测、垂直预测、对角预测、水平预测和水平/对角预测等5种预测选项,各选项方向如图5所示.JVC在报告中证明,使用简化的预测算法能够将运算量减小30%左右,同时预测的质量下降不多^[7,8].

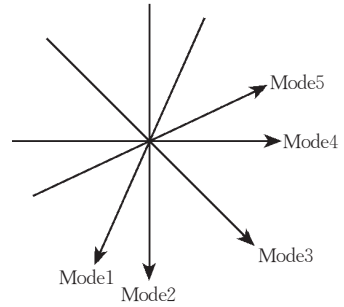


图5 JM中采用的5种方向预测选项

针对 4×4 块帧内预测计算量大的问题,我们在5阶帧内预测的基础上提出了自适应空域帧内预测算法.新算法首次提出了预测选项和SAD值的空间特性,在运算过程中使用了“半步停”技术(half-way stop),设定阈值灵活地中止预测,能够在保证图像质量的前提下进一步提高帧内预测的运算速度.

4.1 预测选项的空间相关性

通过对自然图像和H.264/AVC帧内预测算法结果的分析中我们发现这样一个规律,图像的相邻子块之间具有一定的相关性,由此造成进行帧内预测相邻图像子块之间的预测选项通常也具有相关性,即:当前子块的预测选项和其上方、左方、左上方和右上方子块的预测选项往往是高度一致的.图6(a)是CIF格式的标准图像序列Foreman中的一帧,图6(b)截取了(a)背景中 96×96 大小的矩形区域;(c)中给出了(b)中区域子块进行5阶帧内预测的预测选项结果.

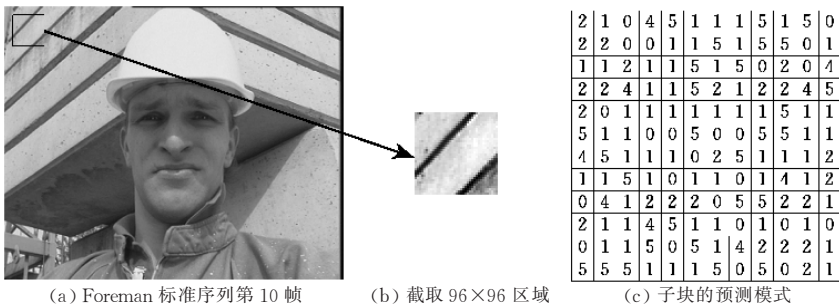


图6 相关性示意

为了更清楚地说明二者之间的关系,下面选取6个CIF格式的标准图像序列 Alex、Foreman、Miss

America、Tennis、Claire、Mobile,使用5阶帧内预测算法对以上序列的前100帧进行预测.表3中分别

列出了预测后子块的预测选项和左方、上方和右上方子块预测选项相同的数目及其所占百分比的平均值。在统计时,表 2 将当前块与其上方、左方和右上方的子块预测选项分离,结果中包含了当前块与一个以上的相邻块预测选项相同的情况。当前块与两个或三个

相邻块预测选项相同时,各邻块次数分别计算。表 3 中列出了帧内预测时忽略各子块间预测选项相同情况下的相关性结果,顺序为首先对上方块进行比较,如果当前块的预测选项与上方块相同,则转向下一子块的比较,否则继续比较左方块和右上方子块。

表 2 子块间的相关性(考虑邻块预测选项一致的情况)

| 图像序列 | 上方 | | 左方 | | 右上方 | | 不能预测的子块 | |
|--------------|-------------|--------------|-------------|--------------|-------------|--------------|-------------|--------------|
| | 次数 | 百分比 | 次数 | 百分比 | 次数 | 百分比 | 次数 | 百分比 |
| Alex | 3286 | 51.86 | 2823 | 44.55 | 2787 | 43.98 | 1806 | 28.50 |
| Foreman | 1828 | 30.37 | 1990 | 31.41 | 1792 | 28.29 | 2522 | 39.81 |
| Miss America | 2098 | 33.11 | 2054 | 32.43 | 1951 | 30.79 | 2397 | 37.84 |
| Tennis | 1946 | 30.71 | 2080 | 32.94 | 1786 | 28.19 | 2463 | 38.88 |
| Claire | 2023 | 31.93 | 1790 | 29.25 | 1718 | 27.11 | 2514 | 39.68 |
| Mobile | 2063 | 32.56 | 1911 | 30.15 | 1686 | 26.61 | 2460 | 38.82 |
| 平均 | 2202 | 34.92 | 2108 | 33.46 | 1953 | 30.83 | 2360 | 37.30 |

表 3 子块间的相关性(不考虑预测选项一致的情况)

| 图像序列 | 上方 | | 左方 | | 右上方 | | 不能预测的子块 | |
|--------------|-------------|--------------|-------------|--------------|------------|-------------|-------------|--------------|
| | 次数 | 百分比 | 次数 | 百分比 | 次数 | 百分比 | 次数 | 百分比 |
| Alex | 3286 | 51.86 | 843 | 13.31 | 401 | 6.33 | 1806 | 28.50 |
| Foreman | 1828 | 30.37 | 1146 | 19.04 | 649 | 10.78 | 2522 | 39.81 |
| Miss America | 2098 | 33.11 | 1169 | 18.46 | 668 | 10.59 | 2397 | 37.84 |
| Tennis | 1946 | 30.71 | 1253 | 19.78 | 674 | 10.63 | 2463 | 38.88 |
| Claire | 2023 | 31.93 | 1158 | 18.29 | 640 | 10.10 | 2514 | 39.68 |
| Mobile | 2063 | 32.56 | 1102 | 20.39 | 522 | 8.23 | 2460 | 38.82 |
| 平均 | 2207 | 35.09 | 1112 | 18.21 | 592 | 9.44 | 2360 | 37.30 |

从表 2 中可以看出,在一幅图像的 6336 个 4×4 子块中,平均有 34.92% 的子块的预测选项与其上方块一致;33.46% 子块的预测选项和左方子块相同;30.83% 子块的预测选项等于右上方子块的预测选项;只有 37.8% 的子块与相邻块不一致,无法通过相邻块的预测选项得出结果。统计结果证明了前述的结论,即一般的图像序列中子块间预测选项具有较强的相关性。按其邻块相关性从大到小排列分别为:上方子块、左方子块和右上方子块。表 3 说明按照上方、左方、右上方的顺序进行预测,平均有 62.17% 的子块可以这三步得到其预测选项。这种子块预测选项的相关性也是我们所提出快速帧内预测算法的基础。

4.2 SAD 值的空间相关性

我们首先给出 SAD 的定义。设当前子块各像素的实际值为 $f(i, j)$, 使用预测选项 $Mode_N$ ($N=0, 1, \dots, 5$) 做帧内预测所得到的预测值为 $f_N(i, j)$, 其中 i, j 分别代表块内像素点的横坐标和纵坐标。则 SAD 定义为

$$SAD = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |f_N(i, j) - f(i, j)| \quad (1)$$

设当前块为 X , 当前块的预测选项为 $Mode_X$, 当前块的某邻块 Y 编码的最优预测选项为 $Mode_Y$ 。SAD 在某种意义上反映的是块与块之间的相似度,

因此当前子块的 SAD 值与其相邻子块的 SAD 之间具有高阶相关性。事实上,如果当前子块的最优预测模式与其邻块的最优预测模式相同,即 $Mode_X = Mode_Y$, 那么使用两子块使用的 SAD 值通常相差不多。为了有效地对 SAD 进行比较,在此我们引入 SAD 相关性因子 λ 判断块 X 和其邻块 Y 的预测选项是否一致。

$$\lambda = \frac{SAD - SAD'}{SAD} \quad (2)$$

其中 SAD' 是子块 Y 的 SAD 值, SAD 代表当前块 X 使用选项 $Mode_Y$ 预测得到的 SAD 值。当 $\lambda \geq 0$ 时,其值越接近于 0, 说明当前块的 SAD 值与其邻块越接近,我们就认为预测选项和邻块的预测选项相同的几率就越大。当 $\lambda < 0$ 时,即 $SAD < SAD'$, 说明使用该预测选项对当前块进行预测的误差比其邻块还要小,此时就可以认为该邻块的预测选项即为当前块的预测选项。

我们对上述的 6 个图像序列的前 100 帧分别使用 5 阶帧内预测算法做预测,计算出子块与上方、左方或右上方的子块预测选项相同情况下的 λ 值,并得出其平均值。表 4 中列出了 λ 值的具体结果。

从表 4 的结果中可以看出,相邻子块的 SAD 值与预测选项之间存在数值上的相关关系,邻块的预测选项相等意味着 λ 的值处于一定范围之内,因此利用

λ 可以有效地反映出这种相关性. 例如, 在 Alex 序列中, 在当前块邻块预测选项相同的情况下, $\lambda = 0.2298$. 其他序列的 λ 值介于 $0.1209 \sim 0.3120$ 之间. 反之, 如果当前子块与其邻块的 λ 值满足上述条件, 我们可以近似地认为两邻块的预测选项相等.

表 4 各图像序列的 λ 值

| 图像序列 | λ | | | |
|--------------|-----------|--------|--------|--------|
| | 上方 | 左方 | 右上方 | 平均 |
| Alex | 0.1335 | 0.2581 | 0.2977 | 0.2298 |
| Foreman | 0.3322 | 0.3353 | 0.2685 | 0.3120 |
| Miss America | 0.0989 | 0.0979 | 0.1235 | 0.1068 |
| Tennis | 0.1336 | 0.1065 | 0.1727 | 0.1376 |
| Claire | 0.3482 | 0.3986 | 0.3767 | 0.3745 |
| Mobile | 0.1667 | 0.2174 | 0.1959 | 0.1209 |

对多个视频序列的统计分析说明: 图像序列中当前块与其上方、左方和右上方的子块存在某种比值关系. 使用 λ 对预测选项进行分析, 不受视频运动类型的影响, 具有较强的自适应性, 且能够较为准确地中止预测, 有助于使用较小的运动代价获得较好的预测效果. 设定阈值 $T \in [0.3, 0.5]$. 当满足条件 $\lambda \leq T$ 时, 认为当前子块与其邻块的预测选项一致. 此时预测的虚警概率和漏报概率能够被控制在合理的范围之内.

4.3 自适应帧内预测算法描述

基于子块间预测选项相关性, 并考虑预测选项与 SAD 值之间的关系, 我们构造了一种新的帧内预测算法, 称为自适应空域帧内预测算法. 在预测过程中自适应地选择 SAD 范围, 满足条件的子块直接跳出预测, 进一步提高预测速度, 体现了“半步停”

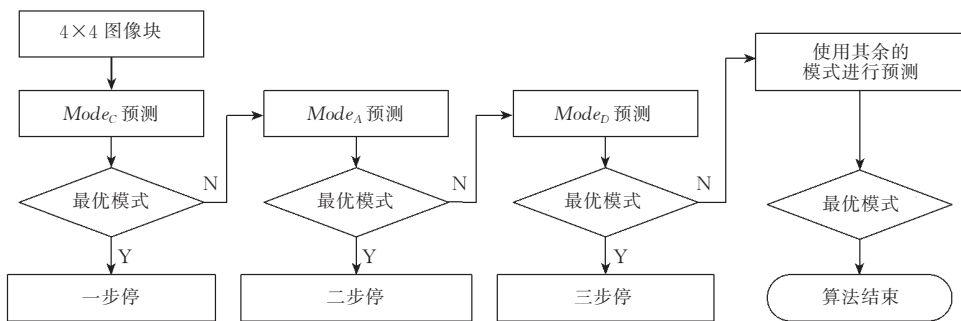


图 8 自适应帧内预测算法流程图

5 实验结果分析

在本节中, 我们仍然使用 Alex, Foreman, Miss America, Tennis, Claire, Mobile 这 6 个标准图像序列进行实验结果分析. 实验中选取每个序列的前 100 帧使用 H.264/AVC 草案中的预测算法(9 种预测选项)、5 阶帧内预测算法和本文提出的自适应空

(half-way stop) 的思想. 对于图 7 所示的各子块, 当前 4×4 块为 X, 块 A, B, C, D 分别为块 X 左方、左

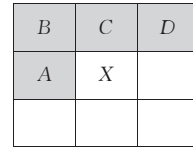


图 7 图像内部的子块划分方式

上方、上方和右上方的子块, 设定阈值为 T . 自适应帧内预测算法的具体步骤如下.

1. 令当前块的预测选项等于上方子块的预测选项, 即 $Mode_X = Mode_C$. 计算出预测结果 $f'(x, y)$ 和 SAE_X . 使用公式(2)得出块 X 和块 C 之间的相关性系数 λ . 当满足 $\lambda \leq T$, 则认为 $Mode_C$ 就是当前块的最优预测模式, $f'(x, y)$ 是最终的预测结果, 结束本块的预测; 如不满足上述条件, 则转入步骤 2.

2. 首先判断块 A 的预测选项是否与块 C 一致, 如果 $Mode_A = Mode_C$, 说明 $Mode_A$ 也非最选项, 此时跳转到步骤 3. 如果 $Mode_A \neq Mode_C$, 令当前块的预测选项等于左方子块的预测选项, 即 $Mode_X = Mode_A$. 计算出预测结果 $f'(x, y)$ 和 SAE_X , 得出此时的相似度系数 λ . 当满足 $\lambda \leq T$, 则认为 $Mode_A$ 就是当前块的最优预测模式, 结束本块的预测; 否则转入步骤 3.

3. 如果满足 $Mode_D \neq Mode_C$ 且 $Mode_D \neq Mode_A$, 令当前块的预测选项等于右上方邻块的预测选项, 即 $Mode_X = Mode_D$. 计算出预测结果 $f'(x, y)$ 和 SAE_X , 得出此时的相似度系数 λ . 当满足 $\lambda \leq T$, 则认为 $Mode_D$ 是最优预测模式, 结束本块的预测; 否则转入步骤 4.

4. 使用除 $Mode_A, Mode_B, Mode_C$ 之外剩余的预测选项进行预测, 在 5 种预测选项中选择 SAD 值最小的选项, 即为最优预测选项.

5. 结束本块预测. 整个算法的实现流程如图 8 所示.

域帧内预测算法分别进行预测, 得出各算法的图像质量和其运算量的情况并加以比较.

5.1 算法运算量比较

为了准确地比较各算法的运算量, 我们使用加、乘运算的次数作为衡量帧内预测算法的运算量的标准. 由于在帧内预测中各预测选项的乘法和加法运算都以 2 为单位, 在实际应用中尤其是使用 DSP 等硬件实现时, 通常使用移位运算来代替乘/除运算以

提高运算速度,则帧内预测的运算量可以更准确地用加法运算和移位运算的数目加以表达.表 6 给出了 H. 264/AVC 帧内预测算法(9 种预测选项)、5 阶

帧内预测算法和自适应空域帧内预测运算量的对比结果,并给出其每像素的平均运算量.

表 5 帧内预测算法计算量比较

(单位:次)

| 图像序列 | H. 264 帧内预测算法计算量 | | 5 阶帧内预测算法计算量 | | 自适应空域帧内预测算法计算量 | |
|--------------|------------------|---------|--------------|--------|----------------|--------|
| | 加法 | 移位 | 加法 | 移位 | 加法 | 移位 |
| Alex | | | | | 95589 | 110039 |
| Foreman | | | | | 136962 | 149315 |
| Miss America | 2211264 | 1919808 | 310464 | 386496 | 123107 | 146347 |
| Tennis | | | | | 135561 | 149593 |
| Claire | | | | | 136283 | 143227 |
| Mobile | | | | | 126875 | 144246 |
| 平均/子块 | 349 | 303 | 49 | 61 | 19.84 | 22.17 |

从表 5 中可以看出,相比较于 H. 264/AVC 中的帧内预测算法和 5 阶帧内预测算法,新算法的计算量有了很大的下降.对于一个 CIF 格式的图像,使用 H. 264/AVC 的帧内预测算法预测需进行 2211264 次加法和 1919808 次移位运算,5 阶预测算法需要 310464 次加法和 386496 次移位运算.由于此两种方法只使用全部预测选项进行预测,因此其总运算数目是恒定不变的.使用自适应帧内预测算法进行预测,预测时每个 4×4 子块需要的平均加法运算的数目从 349 次降低到 19.84 次,使用的移位运算数目从 303 次降低至 22.17 次,相当于 H. 264/AVC 帧内预测算法运算量的 5.68% 和 7.32%,说

明新算法是极为高效的.

5.2 重构图像的客观质量比较

我们使用 PSNR 作为图像客观质量的量度,比较三种算法的预测精度.PSNR 定义如下

$$PSNR = 10 \times \log \left[\frac{M \times N \times f_{\max}^2}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} [f'(x,y) - f(x,y)]^2} \right] \quad (3)$$

上式中, f_{\max} 为原图像中像素的最大灰度值, M, N 分别为子块的长和宽, $f(x,y)$ 和 $f'(x,y)$ 分别为原图像和重建图像中对应的像素灰度值. PSNR 越高,说明重构图像与原图像越接近,预测也越为准确.三种预测算法的 PSNR 对比结果如表 6 所示.

表 6 帧内预测算法 PSNR 比较

(单位: dB)

| 图像序列 | PSNR | | |
|--------------|-----------|-------------|---------------|
| | 5 阶帧内预测算法 | 自适应空域帧内预测算法 | $\Delta PSNR$ |
| Alex | 30.89 | 29.56 | -1.33 |
| Foreman | 26.04 | 24.62 | -1.42 |
| Miss America | 32.97 | 31.87 | -1.10 |
| Tennis | 25.04 | 24.34 | -0.70 |
| Claire | 31.89 | 31.07 | -0.82 |
| Mobile | 19.21 | 18.04 | -1.17 |
| 平均 | 27.67 | 26.58 | -1.09 |

从表 6 中可以看出,使用自适应空域帧内预测算法重构图像 PSNR 略有下降,根据图像序列不同,从 $-1.42\text{dB} \sim -0.7\text{dB}$ 不等.新算法 PSNR 平均值为 26.58dB,相比较于原有算法的 27.67dB 仅

下降 -1.09dB ,因此自适应空域帧内预测算法仍然具有较高的预测精度.

5.3 重构图像的主观质量比较

图 9 显示了 Foreman 序列和 Alex 序列使用 5



图 9 帧内预测算法重构图像主观质量比较

阶帧内预测算法和自适应帧内预测算法所得到的重构图像. 其中(a)和(d)是原始图像,(b)和(e)是使用5阶帧内预测算法得出的重构图像,(c)和(f)是使用本文快速算法得到的预测图像. 对比各图像可以看出,新算法的重构图像具有较好的主观质量,图像与使用5阶帧内预测算法得出的重构图像总体质量非常接近,仅有部分边缘细节的锯齿失真有轻微的增大.

6 结束语

H. 264/AVC 作为 ITU 和 ISO/IEC 制定的新一代国际视频编码标准,其优异的性能必将使其广泛用于可视电话、电视会议、远程教育、远程医疗等诸多领域,具有广阔的应用前景. 但总体上说,H. 264/AVC 性能的改进是以增加计算复杂度为代价获得的,而其编解码所需的庞大运算量是影响其实时应用的主要障碍. 目前国内外所作的相关工作主要集中在于使用低复杂度的整数变换以及快速运动估计算法方面,关于帧内预测算法的专门研究尚不多见.

本文依据 H. 264/AVC 帧内预测算法中各子块预测选项间的相关性,考虑预测模式与子块 SAD 之间的联系,提出了自适应空域帧内预测算法. 该算法依次使用当前子块上方、左方和右上方子块的预测选项进行预测,采用“半步停”技术设定阈值灵活地

判断满足要求的预测选项. 实验结果表明,相比于原有算法,该算法能有效降低帧内预测的计算复杂度,同时图像质量和原算法重构图像非常接近.

参 考 文 献

- 1 ITU-T. Recommendation H. 263: Video coding for low bit rate communication. ITU-T Recommendation H. 263 Draft, 1995
- 2 ISO/IEC/JTC1/SC29/WG11, ISO/IEC. Coding of Audio-Visual Objects-Part 2: Visual, ISO/IEC 14496-2 (MPEG-4 Visual Version 1), 1999
- 3 Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H. 264/ISO/IEC 14496-10 AVC), JVT-G050, 2003
- 4 ITU-T. Draft ITU-T Recommendation H. 263 version 2: video coding for low bit rate communication, 1997
- 5 Wenger S.. H. 264/AVC over IP. IEEE Transactions on Circuits and systems for video technology, 2003, 13(7): 645~656
- 6 ITU-T. H. 26L Test model long-term number 9 (TML-9) draft0, ITU Document VCEG-N83d1, 2001
- 7 Wiegand T., Sullivan G. J.. Overview of the H. 264/AVC video coding standard. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560~576
- 8 Ding Gui-Guang, Guo Bao-Long. H. 26L based fine granularity scalable video coding. Journal of China Institute of Communications, 2004, 25(9): 132~139(in Chinese)
(丁贵广,郭宝龙. 基于 H. 264 的精细度可伸缩视频编码. 通信学报, 2004, 25(9): 132~139)



NI Wei, born in 1981, Ph. D. candidate. His research interests include video coding, multimedia communication and image processing.

GUO Bao-Long, born in 1962, professor, Ph. D.. His main research interests include image processing and communication, intelligent information processing and neural networks.

YANG Liu, born in 1981, Ph. D. candidate. Her research interests include image processing, virtual reality.

Background

This work is a part of the project "Researches on Key Techniques in the JPEG2000 International Standard", supported by the Excellent Young Teachers Program of National Education Ministry of China (2003). Image/video coding is one of the key technologies in a wide range of applications. Recently, a merging video coding standard, known as H. 264/AVC is developed by the ITU-T and ISO. It can significantly improve the coding performance in both low and high bit rates as compared with previous coding standards. However the computational complexity of H. 264/AVC encoder is dramatically increased, which makes it difficult for practical ap-

plications. This is still an open problem to be solved. In the last few years, plenty of research work has been done on video and image coding by our research group, several novel coding algorithms have been proposed and used in practice, over 30 research papers have been published on domestic and overseas publications or conference proceedings. A spatial intra prediction algorithm was adopted in H. 264/AVC, lots of prediction modes must be used for various kinds of blocks, thus making it a highly time-consuming part. To reduce the computational complexity in intra prediction, authors have conducted the work presented in this paper.