

H.264 中 4×4 块的快速帧内预测算法

王启文, 黄东军

(中南大学信息科学与工程学院, 长沙 410083)

摘要: 视频压缩标准 H.264/AVC 的压缩率很高, 但其算法复杂, 编码时间较长。针对帧内 4×4 块提出基于方向预测模式度量的快速预测算法, 通过对当前 4×4 块帧内预测方向的度量, 利用相邻块和相邻方向预测模式的相关性进行帧内模式预测。实验结果表明, 与全搜索算法相比, 在采用全 I 帧编码的情况下, 该算法的编码时间减少 43%, 输出的视频比特率仅增加 0.8%, 且 PSNR 基本保持不变。

关键词: 帧内预测; 预测方向; 率失真优化

Fast Intra-frame Prediction Algorithm for 4×4 Block in H.264

WANG Qi-wen, HUANG Dong-jun

(Institute of Information Science and Engineering, Central South University, Changsha 410083)

【Abstract】 Video compression standard H.264/AVC has high compression ratio, but its algorithm is complex and the coding time is long. Aiming at the intra-frame 4×4 block, this paper proposes a fast prediction algorithm based on measurement of direction prediction mode. By measuring the intra-frame prediction direction of 4×4 block, it uses the correlation between neighboring blocks and neighboring direction prediction mode to predict the intra-frame mode. Experimental results show that compared with the full searching algorithm, this algorithm can decrease the coding time by 43% and only increase the output video bit rate by 0.8% while keeping PSNR almost unchanged.

【Key words】 intra-frame prediction; prediction direction; Rate Distortion Optimization(RDO)

1 概述

H.264/AVC 是新一代视频编码标准, 为了提高编码效率, 即在保证图像质量的前提下降低输出码率, H.264 引入了复杂的时间/空间编码模式, 并通过率失真优化(Rate Distortion Optimization, RDO)技术对各种模式进行计算, 以取得最优的编码模式, 并最大限度地降低输出码率。H.264 标准中采用 RDO 技术计算所有帧内和帧间预测模式, 导致编码时间过长, 不能应用到实时编码通信中。因此, 快速算法成为 H.264/AVC 的研究热点之一。

目前, 对帧内预测的研究已取得很多成果, 如快速三步法^[1]。该算法利用相邻预测方向的相关性, 选择性地计算帧内方向预测模式, 以达到减少率失真优化计算量的目的, 但它对编码时间的节约效果不理想。文献[2]利用子块边缘的变化方向从预选模式中选取最可能的预测方向, 该算法采用边缘梯度直方图法, 在几种最可能的预测方向中选取最佳预测方向, 在视频场景较复杂的情况下, 其视频压缩效果不佳。文献[3]对文献[2]算法进行改进, 但对预测正确率的提高效果不明显。文献[4]提出自适应阈值快速算法, 利用邻块相关性为当前块的率失真设置一个阈值, 提前判断当前块是否为帧内块, 但视频图像的动态变化性导致其阈值很难预测。鉴于此, 本文提出一种新的预测方向度量方法。

2 H.264/AVC 的帧内预测

H.264/AVC 的帧内预测应用于 I 帧和 p 帧。利用空间相关性, 用当前宏块的左边和上边宏块的相邻像素对当前宏块进行预测。对当前宏块与预测宏块的差进行变换、量化和熵编码, 从而对视频进行压缩。在 H.264/AVC 帧内预测编码中, 亮度块包括 4×4 块和 16×16 块, 在最新标准中加入了 8×8 块的预测, 其中, 4×4 块有 9 种预测模式; 16×16 块和 8×8 块

有 4 种预测模式。选择 RDO 计算代价最小的预测模式作为当前块的预测模式。RDO 代价函数如下^[5]:

$$RD_{cost} = SSD + \lambda_{mode} \times Rate$$

其中, SSD 为原始块与重建块像素值的平方差之和; λ_{mode} 为 QP 的指数函数; $Rate$ 表示采用该预测模式需要的比特数。对色度块只有一种预测类型, 即 8×8 块的预测, 它有 4 种预测模式, 与 16×16 块的 4 种预测模式相同。

图 1(a)中小写字母表示当前编码块, 大写字母表示上边和左边已经被编码和重建的像素。当前块的预测块 P 的像素按图 2 的方式根据样点(大写字母 A~M)计算得到。图 1(b)给出了 8 种预测方向。

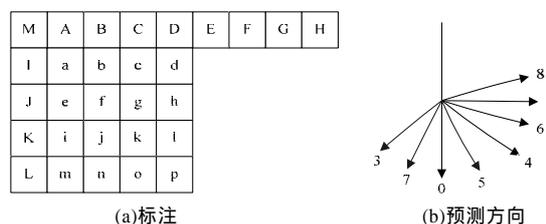


图 1 4×4 预测样点的标注和预测方向

图 2 给出了 4×4 的 9 种模式预测方式, 包括一种 DC 模式和 8 种方向预测模式, 图中的箭头表示每种模式的预测方向, 用 A~M 对当前块进行预测。其中, 模式 2 用 $\text{round}((A+D+I+L)/4)$ (round 表示取整)对像素进行预测。

基金项目: 国家自然科学基金资助项目(60873188)

作者简介: 王启文(1983 -), 男, 硕士研究生, 主研方向: 网络多媒体技术; 黄东军, 教授

收稿日期: 2009-07-24 **E-mail:** 619820762@qq.com

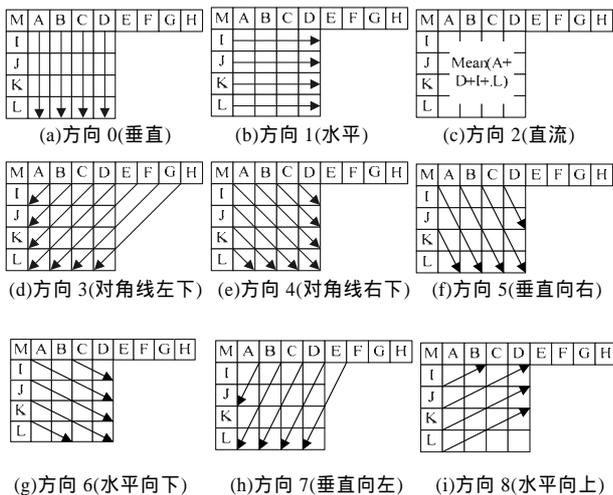


图2 4×4亮度预测模式

在 H.264 标准中,通过 RDO 技术,对所有帧内预测模式进行遍历,并选择一种最优预测模式。帧内预测块的预测模式不是直接进行熵编码,而是将其与最可能的模式相减后,进行熵编码。最可能模式的符号为 most-probable-mode^[4],它是由当前块的左边块(图 3 中的 A)和上边块(图 3 中的 B)的预测模式得到的。当前块的预测模式为 most-probable-mode 时,只要用一个比特来表示预测模式,与选用其他模式相比,其所用比特数最少。

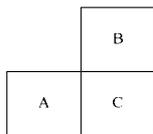


图3 当前块(C)与相邻块(A,B)

虽然 H.264 标准中的全搜索算法能保证取得最优的压缩率,但计算的时间复杂度较高。本文快速算法让 H.264 在编码性能与编码速度间取得较好的折中。

3 快速帧内预测算法

本文算法充分利用了 4×4 块预测模式的方向特性,以及相邻块和相邻方向预测模式的相关性。先对 4 个方向性强的方向预测模式进行度量,即垂直(模式 0)、水平(模式 1)、对角线左下(模式 3)、对角线右下(模式 4)。然后从度量结果中选出候选预测模式,并计算当前块是否具有方向性。如果具有方向性,则根据相邻方向预测模式相关性,将与候选模式相邻的预测模式作为候选预测模式。否则,把模式 2 选为候选模式。最后根据相邻块的相关性,把相邻块的预测模式作为候选模式。

3.1 帧内 4×4 块方向预测模式的度量方法

本算法先通过方向预测模式的方差值度量方向预测模式。方向预测模式的方差计算以图 2 中同一箭头指向像素为一列。然后计算每列的方差,最后每列的方差值相加。方差值越小,说明 4×4 块的纹理方向与该方向预测的方向相关度越高。以图 2 中模式 0,即垂直预测模式为例,说明基于方向预测模式的度量方法:

(1)计算每个方向的上像素的均值, $V1, V2, V3, V4$ 代表每列的均值,即

$$\begin{aligned} V1 &= (a+e+i+m)/4 \\ V2 &= (b+f+j+n)/4 \\ V3 &= (c+g+k+o)/4 \end{aligned}$$

$$V4 = (d+h+l+p)/4$$

(2)计算每个方向上的方差,为了便于计算,简化了方差的计算, $D1, D2, D3, D4$ 代表每列的方差值,即

$$\begin{aligned} D1 &= |V1-a|+|V1-e|+|V1-i|+|V1-m| \\ D2 &= |V2-b|+|V2-f|+|V2-j|+|V2-n| \\ D3 &= |V3-c|+|V3-g|+|V3-k|+|V3-o| \\ D4 &= |V4-d|+|V4-h|+|V4-l|+|V4-p| \end{aligned}$$

(3)对每个方向上的方差值进行累加, D 为总方差值,即 $D=D1+D2+D3+D4$

D 越小,相应模式作为候选模式的可能性越大。

3.2 帧内 4×4 快速算法描述

把 4×4 块的 9 种预测模式分为 4 组,即 $A1=\{0,1,3,4\}, A2=\{5,7,8,9\}, A3=\{2\}, A4=\{\text{upMode}, \text{leftMode}, \text{MostProbableMode}\}$ 。对 $A1$ 按 3.1 节方法进行计算。根据相邻方向预测模式相关性,从 $A2, A3$ 中选出相应的候选模式。根据相邻块预测模式相关性,从 $A4$ 中选出候选模式。具体算法描述如下:

(1)计算 $A1$ 中方向预测模式的方差。计算出的方差按升序排列,选出最优模式(D 最小)和次优模式(D 次小)。

(2)如果最优预测模式和次优预测模式在 $A1$ 中的预测模式中,方向上是相邻的(例如最优模式是 0,相邻的模式为 3,4),则表明当前预测块可能有很强的纹理方向性,根据相邻方向预测模式相关性,从 $A2$ 中把最优模式相邻的 2 个方向预测模式(例如,图 1(a)的最优模式是 0,相邻预测模式为 5,7)和 $A1$ 中的最优模式作为候选模式。反之,如果最优模式和次优模式在方向上不相邻,则表明当前预测块纹理方向性可能不是很强,因此,把 $A1$ 中的最优模式和模式 2 作为候选块。

(3)帧内预测模式与相邻块有很强的相关性,如果当前块具有很强的纹理方向性,即最优预测模式和次优预测模式在 $A1$ 的预测模式中方向上是相邻的,则可以推断相邻块也具有与当前块一样的纹理方向性,因此,把 most-probable-mode 作为候选预测方向。否则,把 2 个相邻块的预测模式都作为候选模式(图 3 中 A 和 B 块的预测模式)。

算法伪代码描述如下:

```

Int A1[4]; // 预测模式分组 A1
Int candidate[4]; // 候选模式
candidate[0]=GetMini(A1[4]); // 从 A1 中获取最优模式
candidate[1]=GetSecondMini(A1[4]); // 获取次优模式
if((candidate[0]-candidate[1])>1) // 具有方向性
{
candidate[1]=the right neighbor of candidate[0] in A2;
candidate[2]=the left neighbor of candidate[0] in A2;
candidate[3]=MostProbableMode;
}
else
{
candidate[1]=2; // 模式 2
candidate[2]=upMode;
candidate[3]=leftMode;
}

```

在本算法中,把 4×4 块的 9 种预测模式减少到 4 种,极大提高了编码效率。

4 实验结果

在参考软件 JM8.6 中实现本算法,并与 H.264 标准的全搜索算法、PAN Feng 的 sobel 算子算法进行实验对比。实验条件如表 1 所示。

表 1 JM 编码器参数设置

参数	值
Profile/Level IDC	66/30
Gop	IIII
RDO	Used
熵编码	CAVLC
编码帧数	30
量化参数	28, 36
Hadamard 变换	Used

分别对 container, brigde, news 的 QCIF 格式和 CIF 格式进行测试。表 2 和表 3 是针对各个测试序列进行实验得到的对比数据,其中,S 表示 PAN Feng 的算法;M 表示本文算法。由表 3 可见,PAN Feng 算法与全搜索算法相比,图像压缩比率增加了 5%左右,编码时间约减少 39%。本文算法与全搜索算法相比,视频图像压缩比率增加 0.8%左右,但编码时间约减少了 43%。

表 2 量化参数为 28 时 2 种算法与全搜索算法的比较

测试序列	算法	压缩比率/(%)	编码时间/(%)	亮度信噪比/dB
Container_cif	S	3.24	-39.74	-0.11
	M	0.71	-41.51	-0.06
Bridgefar_cif.yuv	S	3.95	-38.24	-0.07
	M	0.86	-41.54	-0.03
news_cif.yuv	S	8.17	-40.08	-0.13
	M	0.67	-42.16	-0.07
Container_qcif.yuv	S	2.85	-38.47	-0.14
	M	0.69	-41.26	-0.09
Bridgefar_qcif.yuv	S	4.44	-37.35	-0.09
	M	0.86	-42.04	-0.04
news_qcif.yuv	S	5.77	-39.37	-0.16
	M	0.92	-41.71	-0.09

表 3 量化参数为 36 时 2 种算法与全搜索算法的比较

测试序列	算法	压缩比率/(%)	编码时间/(%)	亮度信噪比/dB
Container_cif	S	5.23	-39.17	-0.12
	M	0.97	-41.47	-0.07
Bridgefar_cif	S	4.57	-36.91	-0.07
	M	0.61	-41.28	-0.03
news_cif	S	10.27	-38.62	-0.12
	M	0.48	-41.79	-0.08
container_qcif	S	4.52	-38.86	-0.12
	M	0.96	-41.89	-0.08
Bridgefar_qcif	S	6.65	-37.08	-0.05
	M	1.04	-42.05	-0.01
news_qcif	S	7.50	-38.08	-0.24
	M	0.50	-41.15	-0.09

实验结果表明,本文算法输出视频的大小小于 PAN Feng 的 sobel 算子算法。由于本文算法预测候选模式时考虑了预测样点的相关性,因此预测命中率很高。由压缩比率可知,

本文算法的输出视频大小与全搜索算法相当。由于本文算法是基于当前块整体计算的,因此与 PAN Feng 基于像素的计算相比,算法更简单且效率更高。表 2 和表 3 的亮度信噪比数据以及图 4 所示解码图像表明,本文算法能很好地保持图像质量。



(a)全搜索算法的解码图像 (b)本文算法的解码图像

图 4 解码图像

5 结束语

本文提出基于预测方向度量的帧内 4×4 块快速算法。由于在候选模式预测时考虑了预测样点,因此与文献[2]算法相比,本文算法的候选模式具有较高命中率,其压缩效果更好。

参考文献

- [1] Cheng Chao-Chung, Chang Tian-Sheuan. Fast Three Step Intra Prediction Algorithm for 4×4 Blocks in H.264[C]//Proc. of IEEE International Symposium on Circuits and Systems. Kobe, Japan: [s. n.], 2005.
- [2] Pan Feng, Lin Xiao. Fast Mode Decision for Intra Prediction[J]. IEEE Trans. on Circuits and Systems for Video Technology, 2005, 15(7): 813-822.
- [3] 李世平, 蒋刚毅, 郁梅. 快速帧内预测模式选择新方法[J]. 电子学报, 2006, 34(1): 141-146.
- [4] Kim B G. Fast Selective Intra-mode Search Algorithm Based on Adaptive Thresholding Scheme for H.264/AVC Encoding[J]. IEEE Trans. on Circuits and Systems for Video Technology, 2008, 18(1): 127-133.
- [5] Lim Keng-Pang. Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods[Z]. 2005.

编辑 陈 晖

(上接第 216 页)

通过实验发现,图像融合后,融合的边缘略有偏移,这主要与节点 A、节点 B 的镜头差异有关,包括径向畸变和切向畸变等。

5 结束语

视频传感器网络是一种新型视频信息获取途径,是对计算机视觉技术的扩展应用。与传统传感器网络技术相比,视频传感器网络以其巨大优势在军事、民用、商业领域中有广阔的发展前景。本文提出一种分层次的、基于动态注意力的多质量图像融合方法,利用节点的结构化部署和异构节点的角色划分,在降低网络带宽需求的同时有效提高了画面质量。

参考文献

- [1] 马华东, 陶丹. 多媒体传感器网络及其研究进展[J]. 软件学报,

2006, 17(9): 2013-2028.

- [2] Kulkarni P, Ganesan D, Shenoy P, et al. SensEye: A Multi-tier Camera Sensor Network[C]//Proc. of the 13th Annual ACM International Conference on Multimedia. New York, USA: ACM Press, 2005: 229-238.
- [3] Ma Huadong, Liu Yonghe. Correlation Based Video Processing in Video Sensor Networks[C]//Proc. of the IEEE Wireless Conf.. [S. l.]: IEEE Press, 2005: 987-992.
- [4] 杨垣, 杨万海, 裴继红. 基于小波分解的不同聚焦点图像融合方法[J]. 电子学报, 2001, 29(6): 846-848.
- [5] Ma Yufei, Zhang Hongjiang. A Model of Motion Attention for Video Skimming[C]//Proceedings of International Conference on Image Processing. New York, USA: [s. n.], 2002: 123-129.

编辑 陈 晖