

基于结构相似的 H.264 帧间预测改进算法

杨春玲¹⁾ 王华兴²⁾ 梁荣锬¹⁾

¹⁾(华南理工大学电信学院 广州 510640)

²⁾(爱立信(中国)通信有限公司 北京 100016)

摘 要 H.264/AVC 通过帧间多模式预测、率失真优化及其它新技术,大大提高了视频压缩效率.失真测度的选择对帧间预测模式及运动估计中匹配块的选择有很大影响.在 H.264 帧间预测过程中,失真测度选用的是当前块和候选块之间的绝对差和(SAD)及平方差和(SSD).结构相似度(SSIM)是新提出的一种图像质量评价方法,与其他评价方法相比,它更加符合人眼视觉系统(HVS)且计算量不大.文中将 SSIM 引入到 H.264 帧间预测作为失真测度,提出了基于结构相似的帧间预测方法(IPBSS).该方法是在作者前期工作积累上的进一步改进.实验结果显示,该算法可以在保持视觉质量不下降的前提下,平均节约 13%的码流.

关键词 H.264;帧间预测;失真测度;结构相似度;率失真优化

中图法分类号 TP319 **DOI 号**: 10.3724/SP.J.1016.2009.01603

Improved Inter Prediction Based on Structural Similarity for H.264

YANG Chun-Ling¹⁾ WANG Hua-Xing²⁾ LIANG Rong-Kun¹⁾

¹⁾(School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640)

²⁾(Ericsson(China) Communication Company Ltd., Beijing 100016)

Abstract H.264/AVC achieves higher compression efficiency by employing multiple modes inter prediction, rate-distortion(RD)optimal mechanism and other new techniques. Distortion metric plays an important role in video compression performance. Structural similarity(SSIM) is a new image quality assessment method, which is more consistent with Human Vision Systems(HVS). This paper proposes to adopt SSIM as the distortion metric in the inter prediction cost functions, named "improved inter prediction method based on SSIM"(IPBSS). It is an improved method of the authors' previous work on MEBSS(Motion Estimation Method Based on Structural Similarity). Simulation results show that the proposed IPBSS can averagely save bit rate more than 13% while maintaining almost the same video quality with $QP=10, 20$ and 30 . That is a better result than the authors' previous work MEBSS.

Keywords H.264; inter prediction; distortion metric; structural similarity(SSIM); rate-distortion optimization(RDO)

1 引 言

H.264/AVC 编码标准是由国际电信联盟

(ITU-T)和国际标准化组织(ISO/IEC)共同制定的最新视频压缩标准,同其他标准相比,在相同编码质量前提下,H.264 可以节约 50%的码率^[1].帧间运动估计多模式选择的引入极大提高了压缩效率和率

收稿日期:2007-03-14;最终修改稿收到日期:2009-07-03.本课题得到国家自然科学基金(60402015)、广东省自然科学基金(06025642)资助.杨春玲,女,1970年生,博士,教授,主要从事图像/视频压缩、图像质量评价方面的研究工作. E-mail: eeclyang@scut.edu.cn.王华兴,男,1981年生,硕士,研究方向为视频编码.梁荣锬,男,1984年生,硕士研究生,研究方向为视频编码.

失真(RD)性能. H.264 的帧间预测采用了 8 种不同的预测模式,分别为 Skip 模式、INTER 16×16 、INTER 16×8 、INTER 8×16 、INTER 8×8 ,其中 INTER 8×8 根据编码器参数设置以及优化原则决定是否需要进一步划分为 INTER 8×4 、INTER 4×8 或者 INTER 4×4 . 编码器在编码的过程中要计算所有模式的率失真代价(RDcost),选取具有最小 RDcost 的模式作为当前块的最优预测模式,这就是所谓的率失真优化算法.

率失真优化在视频压缩领域得到广泛的应用,其算法主要分为两类:第 1 类是静态 RDO,率失真函数是通过视频序列静态建模得到的,这一类率失真优化具有算法复杂度低、易于实现等特点,是得到最广泛应用的率失真优化算法.第 2 类是自适应 RDO,率失真函数根据要编码的残差块而自适应变化(例如,Chen 等人提出的基于 ρ 域的率失真优化^[2],Zhang 等人提出的基于上下文的率失真优化^[3]和最近 Li 等人提出的基于拉普拉斯分布率失真优化算法^[4-6]等).这两类的 RDO 都是基于峰值信噪比(PSNR)最优的率失真算法,然而,最近研究者发现 PSNR 并不能很好地反映主观的视觉感知,PSNR 较小的图像,其视觉误差并不一定很大.因此采用更加合理的评价标准有可能进一步改善 H.264 的压缩效率.

结构相似度是 Wang 等学者最先提出的一种图像质量评价方法^[7-8],他们在多年对图像处理、图像压缩、图像质量评价及人眼视觉系统的研究后指出,人眼在观看图像时是提取图像的结构信息,而且人眼能高度自适应实现这个目标,基于此设想提出了一种基于整体图像结构信息变化而非像素误差的结构相似度(Structural Similarity, SSIM)这一质量评价法.并通过大量的实验证明结构相似度(SSIM)比 MSE 更加符合人眼主观感觉,而且计算也比较简单.因此该方法一经提出,就引起了图像处理研究方面学者的广泛关注.已有学者在此结构信息理论的基础上提出多尺度模数相似度(Multi-scale Modular Similarity, M2S)评价法^[9],并且一些学者在研究图像质量评价方法时也会和该方法进行比较^[10].我们也对该方法进行了深入的研究,相关的研究结果发表在国内外学术刊物和国际会议上^[11-13].通过对 SSIM 的深入研究,我们发现结构相似度较大的两个图像块之间的差信号通常是一低频信号,而低频信号是较容易压缩的,且该方法的计算量虽比均方误差稍大,但和其它的基于人眼视觉系统的图

像质量评价方法相比要小很多.基于这些研究,我们提出把结构相似度用于视频编码的帧间预测中,用 SSIM 取代 SAD 和 SSD 作为块匹配及最优预测模式选择时的失真测度,称该方法为基于结构相似的帧间预测方法(IPBSS).

这一思想已在我们的早期工作 MEBSS 算法^[14]中有所体现,但本文的方法和文献^[14]中的处理方法很不相同. MEBSS 算法在计算当前块的 SSIM 时沿用了 H.264 中计算 SAD 和 SSD 的方法,即将当前编码块划分成 4×4 小块,分别计算这些 4×4 小块的 SSIM 的值,然后取平均作为当前块的 SSIM 值.这种计算方法对于 SAD 和 SSD 当然是合理的,因为它们是基于单像素误差的失真计算方法.但对 SSIM,当计算 16×16 、 16×8 、 8×16 、 8×8 、 8×4 、 4×8 这些大块的帧间预测时,这种方法并不合理,因为 SSIM 是基于块的失真测度方法而不是基于像素的,因而 4×4 分割的 SSIM 均值并不等于整块的 SSIM 值. MEBSS 在 QP 较大时实验效果较差原因也在于此.本文中改进了 SSIM 在帧间运动估计的计算方法,同时将 SSIM 引入到帧间预测模式选择中的率失真代价函数中.

本文第 2 节简单介绍 H.264 的帧间预测过程及 SSIM;第 3 节详细描述本文提出的基于结构相似的帧间预测方法(IPBSS);实验及结果分析在第 4 节;第 5 节是本文的结论.

2 H.264 帧间预测和 SSIM

2.1 H.264 帧间预测

多模式帧间预测是 H.264 视频压缩标准中的一个新技术,它的应用进一步提高了压缩效率.在 H.264 中,存在 8 种帧间预测模式,图 1 示出了其中的 7 种预测模式,另一种称为 Skip 模式.利用 Skip 模式编码的宏块,不传运动矢量和残差信号,在解码端利用当前块的上边和左边解码块的运动矢量估计

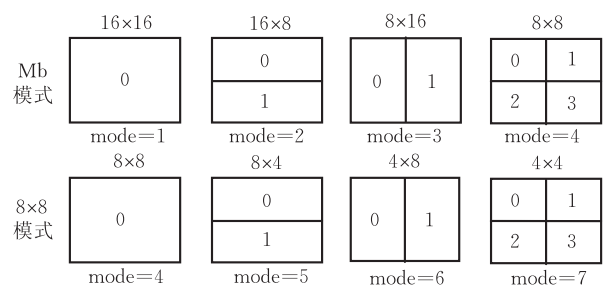


图 1 7 种帧间预测模式示意图

当前块的运动矢量,用匹配块作为当前块的估计块. Skip 模式常用于运动不大且有较好匹配块的宏块编码中.

帧间预测主要有两步,第 1 步是对于每种预测模式,在相应的运动估计搜索区域内寻找当前编码块的最佳匹配块,该过程采用拉格朗日代价函数为匹配准则:

$$MCOST(s, c) = SA(T)D(s, c) + \lambda_{MOTION} Bit(\Delta MV) \quad (1)$$

式(1)中, $SA(T)D$ 是失真测度, SAD 表示候选块 c 与当前编码块 s 之间残差信号的绝对值和,用于整数像素的运动估计中, $SATD$ 是候选块 c 与当前编码块 s 之间残差信号经 Hadamard 变换后的系数绝对值和,用于分数像素的运动估计. λ_{MOTION} 是拉格朗日乘子, ΔMV 是实际运动矢量和预测运动矢量的差值, $Bit(\Delta MV)$ 是对 ΔMV 编码所需的 Bit 数. 具有最小 $MCOST$ 的匹配块选为当前块的最佳匹配块.

第 2 步是选择最佳帧间预测模式,找出每种预测模式的最佳匹配块之后通过下式计算各模式对应的率失真代价.

$$J(s, c, MODE|QP) = D(s, c, MODE|QP) + \lambda_{MODE} R(s, c, MODE|QP) \quad (2)$$

其中 $D(s, c, MODE|QP)$ 是相应预测模式的重建块 c 相对于原图像块 s 的失真,此处用误差平均和 (SSD) 表示失真; QP 是量化参数, $MODE$ 表示预测模式, λ_{MODE} 是对应模式的拉各朗日乘子. 选择使率失真代价函数 $J(s, c, MODE|QP)$ 取值最小的预测模式作为当前宏块的最优预测模式,并按这一预测模式和相应的最佳匹配块对当前块编码.

2.2 结构相似度 SSIM

结构相似度采用图像结构信息的变化量来表征图像质量,与其他图像质量评价方法相比,具有计算量小,符合人眼主观视觉的优点. 它包括亮度、对比度和物体结构 3 个方面,定义为

$$SSIM(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y) \quad (3)$$

其中 x 和 y 表示原图像块和失真图像块, $SSIM(x, y)$

表示它们的结构相似度. $l(x, y)$ 是亮度比较,是原图像块和失真图像块的均值 μ_x 和 μ_y 的函数. $c(x, y)$ 是对比度比较,是原图像块和失真图像块的 δ_x 和 δ_y 的函数; $s(x, y)$ 是结构比较,用原图像块和失真图像块的归一化信号 $(x - \mu_x) / \delta_x$ 和 $(y - \mu_y) / \delta_y$ 的相关系数来表示. 这 3 个分量表示如下:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (4)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (5)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (6)$$

上式中 C_1 、 C_2 和 C_3 是为避免分母为零而加上的小常数. 可以看出 SSIM 最大值为 1, 最小值为 0. 且 SSIM 值越大表示待评估图像与参考图像越相似.

3 基于结构相似的帧间预测 (IPBSS)

在 2.1 节中,已对国际视频压缩标准 H.264 进行分析,该标准中帧间预测的代价函数的失真测度为 SAD 和 SSD ,这两种失真测度和 MSE 具有相同的物理意义,并不能很好地符合人眼视觉系统特性,从图 2 很容易看出这一点. 图 2 中的 $T1$ 是由原信号加上信号 $E1$ 得到的,而 $T2$ 是由原信号加上信号 $E2$ 得到的. $E1$ 和 $E2$ 是等幅值信号,但 $E1$ 中的每一个分量符号相同, $E2$ 的分量符号不同,可见 $E1$ 为一低频信号, $E2$ 为一高频信号. 很容易得出,用 SAD 或 SSD 作为失真测度来衡量 $T1$ 和 $T2$ 这两个信号时,它们相对于原信号的失真大小是相同的,但也很容易发现,信号 $T1$ 和原信号具有非常相似的波形,而 $T2$ 和原信号的波形差别很大,也就是说,在波形上 $T1$ 和原信号更加相似. 这种相似度是 SAD (或 SSD, MSE) 所不能体现的,而 $SSIM$ 就可表现这种失真程度. 文献[7-8]已利用大量的仿真实验证明 $SSIM$ 比 MSE 或 $PSNR$ 更加符合人眼视觉系统.

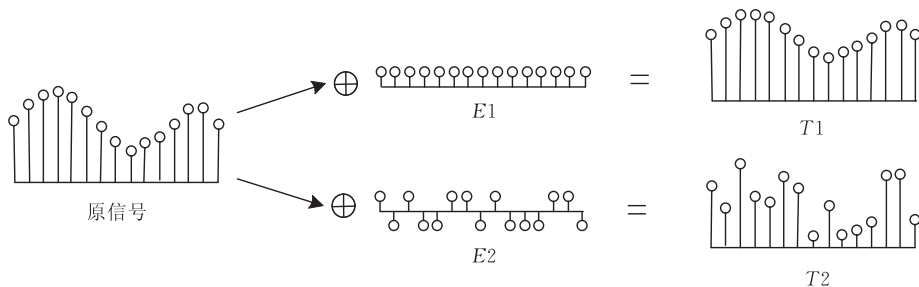


图 2 SAD 和 SSIM 的简单分析示例

另外,从图 2 可见, E_1 是 T_1 相对于原信号的残差信号, E_2 是 T_2 相对于原信号的残差信号,若失真信号和原信号在结构上非常相似,则残差信号在很大程度上是一个低频信号,而我们知道低频信号是容易压缩的. 因此我们提出,采用 $SSIM$ 代替 SAD 或 SSD 作为视频编码中帧间预测的失真测度,有望达到更高的压缩效率,称这种方法为基于结构相似的帧间预测方法,即 IPBSS.

根据 $SSIM$ 理论, $SSIM$ 值的取值范围是 $[0, 1]$, 且其值越大说明候选块与当前块越相似,而 SAD 和 SSD 却恰恰相反,因此用于代价函数中的失真测度取为 $SSIM$ 的函数,用 $D(SSIM)$ 表示:

$$D(SSIM) = 1 - SSIM \quad (7)$$

由于失真测度做了修改,对应的拉格朗日参数也要作相应的调整. 而 $SSIM$ 是一个小于 1 的浮点数,为了计算及硬件实现中的方便,对原拉格朗日公式修改如下:

$$MCOST(s, c) = K_1 \cdot (1 - SSIM(s, c)) + \lambda_{MOTION} Bit(\Delta MV) \quad (8)$$

用于模式选择的代价函数也修改如下:

$$J(s, c, MODE|QP) = K_2 \cdot (1 - SSIM(s, c)) + \lambda_{MODE} R(s, c, MODE|QP) \quad (9)$$

参数 K_1 、 K_2 是 QP 的函数,本算法通过对 K_1 、 K_2 的调整实现失真度和码率之间的平衡,其取值是通过大量的实验而得到的经验值. 上式中的拉格朗日乘子 λ_{MOTION} 、 λ_{MODE} 和 H.264 标准中取值相同.

在 IPBSS 中,帧间宏块的最优预测模式和最佳匹配块的选择主要分两个步骤,描述如下:

1. 对每一种预测模式,在搜索范围内找到它的最佳匹配块.

首先利用式(3)~(6)计算当前编码块与每一个候选块的结构相似度 $SSIM$,然后利用式(8)计算每一个匹配块的代价 $MCOST$,选择具有最小 $MCOST$ 的匹配块为当前编码块的最佳匹配块. 对于 Skip 模式, $MCOST$ 直接按照 $MCOST = K_1 \cdot (1 - SSIM(s, c))$ 计算,因为该模式不需要传送运动矢量和残差信号.

2. 选择最佳预测模式

对每一个预测模式的最佳匹配块,利用式(9)计算率失真代价 ($RDcost$),具有最小 $RDcost$ 的模式为该宏块的最佳预测模式.

本文中提出的 IPBSS 算法和我们在早期工作^[13]中提出的 MEBSS 算法主要有两点不同. (1) 当前预测块的 $SSIM$ 的计算方法不同,本文方法是直接计算当前编码块和匹配块的 $SSIM$ 值,而

文献[14]中的方法是先把当前块和匹配块划分为不重叠的 4×4 小块,计算对应小块的 $SSIM$ 值,然后取平均值作为当前块的 $SSIM$ 值. 这种算法是沿用了 H.264 中计算 SAD 和 SSD 的方法,是我们在该思想提出初期的一种简单实现,但是通过分析发现,这种方法虽然利于计算(可以重复利用 4×4 小块的计算结果来计算大块的 $SSIM$ 值),但并不合理,因为 $SSIM$ 是基于块而不是基于像素的. (2) 本文将 $SSIM$ 引入到最优模式选择的率失真代价函数中,用 $SSIM$ 取代了 SSD . 该算法取得了比 MEBSS 更好的视频压缩性能,性能的分析 and 比较在下一节给出.

4 实验及结果分析

4.1 实验环境

所有的实验结果是基于以下工作环境实现的:

(1) 本文提出的基于结构相似的帧间预测算法是在国际标准 H.264/AVC 的开放参考源代码 JM11.0^① 的基础上修改而实现的, H.264 的实验结果也是利用该参考软件得到的. 所有的实验中参考帧数选为 5 帧,且采用全搜索运动估计. 为了比较 IPBSS 算法和 H.264 原算法在帧间编码方面的性能,两种算法的 P 帧全部采用帧间编码模式.

(2) 实验中采用了 10 个常用的视频编码测试序列,这些序列的第 0 帧示于图 3. 每个序列编码 50 帧,帧率 30 帧/秒,所有实验第 0 帧都采用 $QP=10$ 的帧内编码,其余 49 帧为 P 帧,测试序列除了“waterfall”是 CIF 格式之外,其它序列都为 QCIF 格式.

(3) 仿真中用的计算机配置为 PD/2.8GHz, 512×2MB RAM, 操作系统为 Microsoft Windows XP SP2.

4.2 实验结果分析

通过大量的仿真实验,我们得到了式(8)和(9)中的参数 K_1 和 K_2 的经验值,如表 1 所示.

表 1 参数 K_1 和 K_2 的经验值

QP	K_1	K_2
10	200	80000
20	400	150000
30	1200	200000

$QP=10, 20, 30$ 条件下的实验结果示于表 2~4, 表中 IPBSS 表示本文算法, H.264 表示 H.264 视频

① <http://iphome.hhi.de/suehring/tml/download/>



图 3 实验中用到的序列第 0 帧图示

标准算法, 本文采用 $MSSIM$ 作为解码视频序列的客观质量评价准则, $MSSIM$ 由所有帧间编码序列的 $SSIM$ 平均得出. 表格中 $kbit/Pic$ 表示平均每帧的编码比特数, 由所有 P 帧编码比特数平均而得, 用它表示压缩率. 表格中分别给出了 H.264 和本文

所提算法 IPBSS 对这 10 个视频序列的压缩率 ($kbit/Pic$)、图像质量 ($MSSIM$) 和编码时间, 并给出了两种算法相对于这几个指标的比较值, 即 IPBSS 相对于 H.264 各指标的增加比例.

表 2 $QP=10$ 的实验结果比较

Sequence	H.264			IPBSS			结果比较		
	$kbit/Pic$	$MSSIM$	$Time/s$	$kbit/Pic$	$MSSIM$	$Time/s$	$\Delta(kbit/Pic)/\%$	$\Delta MSSIM/\%$	$\Delta Time/\%$
grandma	27.013	0.9948	5.136	24.068	0.9941	4.477	-10.90	-0.07	-12.83
carphone	39.141	0.9951	6.683	35.929	0.9946	9.301	-8.21	-0.05	39.17
coastguard	61.746	0.9961	8.396	59.400	0.9958	32.810	-3.80	-0.03	290.78
forman	51.110	0.9954	6.071	45.491	0.9937	14.572	-10.99	-0.17	140.03
apple	58.116	0.9935	7.134	57.150	0.9934	12.509	-1.66	-0.01	75.34
news	17.035	0.9962	4.064	12.211	0.9951	5.262	-28.31	-0.11	29.48
silent	20.763	0.9962	4.252	16.992	0.9956	8.607	-18.16	-0.06	102.42
trevor	39.077	0.9966	4.877	34.686	0.9959	12.653	-11.24	-0.07	159.44
akiyo	10.358	0.9965	3.279	6.798	0.996	2.989	-34.37	-0.05	-8.84
waterfall	214.880	0.9965	24.136	202.542	0.9961	79.465	-5.74	-0.04	229.24
Average							-13.34	-0.07	104.42

表 3 $QP=20$ 的实验结果比较

Sequence	H.264			IPBSS			结果比较		
	$kbit/Pic$	$MSSIM$	$Time/s$	$kbit/Pic$	$MSSIM$	$Time/s$	$\Delta(kbit/Pic)/\%$	$\Delta MSSIM/\%$	$\Delta Time/\%$
grandma	3.166	0.9881	3.406	2.420	0.9872	3.511	-23.56	-0.09	3.08
carphone	8.695	0.9848	5.140	7.249	0.9838	7.038	-16.63	-0.10	36.93
coastguard	21.591	0.9845	7.861	19.769	0.9821	32.018	-8.44	-0.24	307.30
forman	11.627	0.9816	6.068	9.464	0.9776	12.158	-18.60	-0.41	100.36
apple	6.467	0.9685	5.275	6.969	0.9689	6.475	7.76	0.04	22.75
news	4.853	0.9909	3.556	3.972	0.9897	4.834	-18.15	-0.12	35.94
silent	5.845	0.9902	3.485	5.250	0.9896	8.232	-10.18	-0.06	136.21
trevor	12.693	0.9879	4.632	10.782	0.9856	11.586	-15.06	-0.23	150.13
akiyo	2.419	0.9942	2.801	1.613	0.9933	2.718	-33.32	-0.09	-2.96
waterfall	39.084	0.9795	21.617	33.826	0.9787	49.719	-13.45	-0.08	130.00
Average							-14.96	-0.14	91.98

表 4 $QP=30$ 的实验结果比较

Sequence	H.264			IPBSS			结果比较		
	<i>kbit/Pic</i>	<i>MSSIM</i>	<i>Time/s</i>	<i>kbit/Pic</i>	<i>MSSIM</i>	<i>Time/s</i>	$\Delta(kbit/Pic)/\%$	$\Delta MSSIM/\%$	$\Delta Time/\%$
grandma	0.529	0.9818	2.011	0.406	0.9804	4.269	-23.25	-0.14	112.28
carphone	1.932	0.9669	3.718	1.325	0.9625	9.438	-31.42	-0.46	153.85
coastguard	4.357	0.9364	6.386	3.890	0.9303	37.393	-10.72	-0.65	485.55
forman	2.055	0.9599	4.326	1.579	0.9531	16.022	-23.16	-0.71	270.37
apple	0.937	0.9509	3.187	0.930	0.9504	7.602	-0.75	-0.05	138.53
news	1.347	0.9798	2.186	1.150	0.9762	5.846	-14.63	-0.37	167.43
silent	1.692	0.9754	2.586	1.536	0.9742	9.645	-9.22	-0.12	272.97
trevor	3.440	0.9631	4.058	2.659	0.9529	16.011	-22.70	-1.06	294.55
akiyo	0.498	0.9902	1.785	0.352	0.9888	3.096	-29.32	-0.14	73.45
waterfall	5.141	0.9496	17.134	4.383	0.9470	67.240	-14.74	-0.27	292.44
Average							-17.99	-0.40	226.14

从表中可以看出,在视频序列质量 *MSSIM* 几乎保持不变的情况下,对应于参数 $QP=10,20,30$, IPBSS 比 H.264 平均节约了 13.34%, 14.96%, 17.99% 的码流,最高节约了 34.37% 的码流. 为了

更好地说明解码序列的视觉质量,本文图 4~图 6 示出了 *MSSIM* 减少最多的 3 个序列最后一帧解码图像,从图中可看出, H.264 和 IPBSS 的解码图像从视觉上看不到差别.

图 4 $QP=30$ 时, coastguard 的第 49 帧解码图像图 5 $QP=30$ 时, trevor 的第 49 帧解码图像图 6 $QP=30$ 时, forman 的第 49 帧解码图像

另外从表中也可清楚地看出,对绝大多数序列,编码时间都有很大的上升,这是因为 *SSIM* 运算复

杂度大于 *SAD* 和 *SSD* 而引起的.但对于那些背景面积大且运动缓慢的视频序列,如“grandma”和

“akiyo”, 编码时间没有显著上升, 且在 $QP=10$ 的时候, 编码时间反而有少许下降. 这是因为, 在进行运动估计时, 程序中采用了优化算法, 如果当前匹配块的运动矢量的编码代价已大于已搜索过的所有匹配块的最小代价 M_{cost} , 则不再计算该匹配块和当前编码块的 $SSIM$, 而直接搜索下一个匹配块. 对于运动缓慢的图像序列, 就减少相当一部分匹配块的 $SSIM$ 的计算量, 整个序列的编码复杂度就降了下来. 由此可见, 利用 $SSIM$ 作为失真测度增加了运动搜索的计算量, 但若搜索点数减少, 增加的计算量就会相应减少, 因此若把该方法用到 H.264 的快速搜索算法中, 由于快速算法减少了搜索点数, 计算量方面的开销就会有很大改善.

从实验仿真结果可以分析出来 IPBSS 算法性能要比我们的前期工作 MEBSS 更好. 因为本文算法在 $QP=10, 20, 30$ 时, 都优于 H.264 原算法, 而 MEBSS 算法仅在 $QP=10$ 时, 比 H.264 压缩性能好, 所以就没有必要列出 IPBSS 算法与 MEBSS 算法性能比较.

5 结 论

本文提出了一种基于结构相似的帧间预测算法——IPBSS, 该算法是对我们前期的工作基于结构相似的运动估计(MEBSS)算法的改进. 在该算法中, 采用 $SSIM$ 作为帧间预测失真测度. 仿真结果显示在 $QP=10, 20, 30$ 时, 与 H.264 原算法相比, IPBSS 可以平均节约码率 13.34%, 14.96%, 17.99% 的码流, 而编码序列的视觉质量没有明显下降. 但由于引入 $SSIM$ 的计算量比原算法中的 SAD 和 SSD 的计算量大, 所以大部分序列的编码时间明显上升, 又由于程序采用了优化, 丢掉了一些匹配块 $SSIM$ 的计算, 运动不太大的视频序列的编码时间增加不大. 由此可见, 若把该方法用于 H.264 的快速算法中, 可望在计算量增加不多, 视觉质量基本不变的条件下, 有效地降低码率.

参 考 文 献

- [1] Wiegand T, Sullivan G J, Bjontegaard G, Luthra A. Overview of the H.264/AVC video coding standard. *IEEE Transactions on CAS for Video Technology*, 2003, 13(7): 560-576
- [2] Chen L, Garbacea I. Adaptive lambda estimation in Lagrangian rate-distortion optimization for video coding//*Proceedings of the Visual Commun Image Process (VCIP)*. San Jose, 2006, 6077(1): 1-8
- [3] Zhang J, Yi X, Ling N, Shang W. Chroma coding efficiency improvement with context adaptive Lagrange multiplier (CALM)//*Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*. New Orleans, 2007, 1(1): 293-296
- [4] Li X, Oertel N, Hutter A, Kaup A. Advanced Lagrange multiplier selection for hybrid video coding//*Proceedings of the IEEE International Conference on Multimedia & Expo (ICME)*. Beijing, China, 2007, 1(1): 364-367
- [5] Li X, Oertel N, Hutter A, Kaup A. Extended Lagrange multiplier selection for hybrid video coding using interframe correlation//*Proceedings of the Picture Coding Symposium (PCS)*. Lisbon, Portugal, 2007, 1(1): 1-4
- [6] Li X, Oertel N, Hutter A, Kaup A. Laplace distribution based Lagrangian rate distortion optimization for hybrid video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2009, 19(2): 193-205
- [7] Wang Z, Bovik A C, Sheikh H R, Simoncelli E P. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612
- [8] Wang Z, Lu L, Bovik A C. Video quality assessment using structural distortion measurement//*Proceedings of the IEEE International Conference on Image Processing*. Rochester, 2002: 65-68
- [9] Zhai G, Zhang W, Yang X, Xu Y. Image quality assessment metrics based on multi-scale edge presentation//*Proceedings of the IEEE Workshop on Signal Processing System Design and Impement*. Athens, 2005: 331-336
- [10] Chen Guan-Hao, Yang Chun-Ling, Xie Sheng-Li. Gradient-based structural similarity for image quality assessment//*Proceedings of the IEEE International Conference on Image Processing*. Atlanta, 2006: 2929-2932
- [11] Sheikh H R, Bovik A C. Image information and visual quality. *IEEE Transactions on Image Processing*, 2006, 15(2): 430-444
- [12] Yang Chun-Ling, Kuang Kai-Zhi, Chen Guan-Hao, Xie Sheng-Li. Gradient-based structural similarity for image quality assessment. *Journal of South China University of Technology*, 2006, 34(9): 22-25(in Chinese)
(杨春玲, 旷开智, 陈冠豪, 谢胜利. 基于梯度的结构相似度的图像质量评价方法. *华南理工大学学报*, 2006, 34(9): 22-25)
- [13] Chen Guan-Hao, Yang Chun-Ling, Po Lai-Man, Xie Sheng-Li. Edge-based structural similarity for image quality assessment//*Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. Toulouse, 2006: 933-936
- [14] Mai Z Y, Yang C L, Kuang K Z, Po L M. A novel motion estimation method based on structural similarity for H.264 inter prediction//*Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. Toulouse, 2006: 913-916



YANG Chun-Ling, born in 1970, Ph. D., professor. Her major research interests include image/video coding, image quality assessment.

WANG Hua-Xing, born in 1981, M. S.. His research interests focus on video coding.

LIANG Rong-Kun, born in 1984, M. S. candidate. His research interests focus on video coding.

Background

Rate-distortion optimization (RDO) plays a vital role in video compression applications, and it is widely researched in recent years. In literature, RDO for video compression can be classified into two categories. The first category computes the theoretical RD function based on a given statistic model for video data. And the second category uses an operational RD function, which is computed based on the data to be compressed. Such as ρ domain based RDO (proposed by Chen L et al), context based RDO (proposed by Zhang J et al) and Laplacian distribute based RDO (proposed by Li X et al). Both of the two categories above are based on the PSNR-Rate framework. However the challenge for designing a method under this framework is that PSNR do not correlate well with HVS. So, in this paper a new RDO method-IPBSS, which is based on structure similarity, is proposed. Simulation results

indicate that the IPBSS can average save bit rate more than 13%, while the video quality almost remain the same.

This work is supported in part by research project from National Natural Science Foundation of China (No. 60402015), and research project from Guangdong Natural Science Foundation of China (No. 06025642). These projects aim at construction methods of image quality assessment and its application in video coding and image compression.

The research interests of the group include image quality assessment, image/video coding. The group has proposed several kind of SSIM applications, such as SSIM based JPEG2000 encoding, and SSIM based H.264 inter frame encoding is presented in this paper.