

用于视频编码的维纳滤波器抽头系数表示新方法

李志斌, 常义林, 杨付正, 刘晓仙

(西安电子科技大学综合业务网理论与关键技术国家重点实验室, 710071, 西安)

摘要: 针对已有视频编码系统中自适应环路滤波器抽头系数编码方法存在冗余的问题,提出一种基于直流补偿的抽头系数编码表示方法(DCTCC). 首先,对解码恢复图像与原始图像的直流差值进行补偿;然后,以直流补偿后的恢复图像为输入信号、原始图像为期望值设计维纳滤波器,进一步补偿恢复图像的交流失真;最后,仅编码传送所设计滤波器的各非中心抽头系数,而根据提出的抽头系数之和等于1的特点则可计算得到中心抽头系数数值,从而提高滤波器参数传送效率. 实验数据表明,对于所用的各种格式视频测试序列,DCTCC的编码效率均优于已有自适应环路滤波方法,与 H. 264/AVC 标准相比,DCTCC 的编码效率最高可达 26.31%.

关键词: 自适应环路滤波器;维纳滤波器;抽头系数;直流补偿

中图分类号: TN919.8 **文献标志码:** A **文章编号:** 0253-987X(2010)10-0077-06

A New Description Algorithm for the Tap Coefficients of Wiener Filter in Video Coding

LI Zhibin, CHANG Yilin, YANG Fuzheng, LIU Xiaoxian

(State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China)

Abstract: A novel tap coefficients codec method based on DC offset (DCTCC) is proposed to overcome the redundant shortcoming in describing tap coefficients of adaptive loop filter in the video coding. Firstly, DCTCC performs compensating on the DC offset between the restored picture and the original picture. Then, the Wiener filter is designed in terms of the compensated picture as the filter input and the original picture as the desired response to further compensate AC distortion in restored picture. Finally, only the non-center tap coefficients are coded and transmitted, while the value of the center tap coefficient is derived by the feature that the sum of the tap coefficients equals 1, so that the transmission efficiency of the filter parameters is improved. Experimental results with test sequences in different formats show that DCTCC outperforms the existing algorithm with adaptive loop filter. Comparisons with H. 264/AVC show that the coding efficiency of DCTCC has an improvement up to 26.31%.

Keywords: adaptive loop filter; Wiener filter; tap coefficients; direct current offset

H. 264/AVC 是目前最新也是压缩性能最高的国际视频编码标准^[1-2],而近期涌现的几个关键技术则进一步提高了 H. 264/AVC 标准的编码性能,其中包括自适应环路滤波技术(Adaptive Loop Filtering, ALF)^[3-5]. 该技术利用维纳滤波器(Wiener Filter, WF)在信号处理领域突出的噪声抑制性

能^[6-7],将 WF 应用于视频预测编码环路中,不仅抑制了视频编码的量化噪声,提高了恢复图像质量,同时经过噪声补偿的恢复图像用作后续编码帧的参考图像,可提高帧间预测的准确性,从而达到提高视频编码效率的目的.

然而,现有 ALF 方法是将所有的 WF 抽头系

数依次全部量化编码传送到解码器,并没有考虑 WF 各抽头系数之间的相关性,导致 WF 抽头系数编码方法存在冗余,语法结构不严谨. 本文提出一种基于直流补偿的 WF 抽头系数编码方法(DCTCC),首先对解码恢复图像进行直流补偿,然后以补偿后的解码恢复图像为输入信号计算 WF,进一步补偿视频编码的交流失真,由于中心抽头系数可以根据 WF 各抽头系数之和等于 1 计算得出,DCTCC 方法只需对 WF 各非中心抽头系数进行编码,从而去除了 WF 参数编码表示方法的冗余,提高了视频编码语法的严谨性.

1 视频编码系统中的维纳滤波器

在视频编码系统中,WF 假设作为期望信号的原始图像 d 为平稳随机过程,并将视频编码失真看成是在原始图像上叠加的噪声 n ,视频编码系统的解码恢复图像 p 可表示为

$$p = d + n \quad (1)$$

最佳 WF 系数 w 的求解方程组^[7]如下

$$\begin{aligned} \mathbf{R}_{p,p} \mathbf{w} = & \\ \begin{bmatrix} R_{p,p}(0) & R_{p,p}(1) & \cdots & R_{p,p}(M-1) \\ R_{p,p}(1) & R_{p,p}(0) & \cdots & R_{p,p}(M-2) \\ \vdots & \vdots & & \vdots \\ R_{p,p}(M-1) & R_{p,p}(M-2) & \cdots & R_{p,p}(0) \end{bmatrix} \cdot & \\ \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{bmatrix} = & \begin{bmatrix} r_{p,d}(0) \\ r_{p,d}(1) \\ \vdots \\ r_{p,d}(M-1) \end{bmatrix} = \mathbf{r}_{p,d}^T \end{aligned} \quad (2)$$

式中: M 为抽头系数个数; $\mathbf{R}_{p,p}$ 是解码恢复图像的自相关函数矩阵; $\mathbf{r}_{p,d}$ 是解码恢复图像与原始图像的互相关函数矩阵.

采用最优 WF 补偿和最小化视频编码失真,可以提高解码恢复图像的质量. 维纳滤波器原理如图 1 所示.

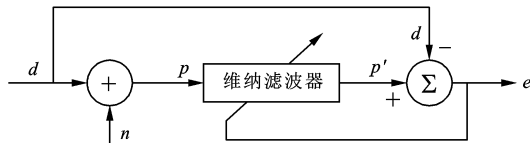


图 1 维纳滤波器原理图

从统计意义上讲,视频编码失真可以近似为白噪声,同时假设视频编码不会引起直流偏移,即编码失真的均值为 0,否则可以通过直流补偿使得补偿后的解码恢复图像与原始图像的均值保持一致. 编

码失真的方差和自相关函数分别为

$$D[n] = E[n^2] - E^2[n] = E[n^2] = \sigma^2 \quad (3)$$

$$R_{n,n}[\tau] = \begin{cases} \sigma^2, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases} \quad (4)$$

通常情况下,原始图像 d 与编码失真 n 相互独立,即下式成立

$$E[d n] = E[d] E[n] = 0 \quad (5)$$

分别化简 $\mathbf{r}_{p,d}$ 和 $\mathbf{R}_{p,p}$, WF 求解方程组(2)可变形为

$$\begin{bmatrix} R_0 + \sigma^2 & R_1 & \cdots & R_{M-1} \\ R_1 & R_0 + \sigma^2 & \cdots & R_{M-2} \\ \vdots & \vdots & & \vdots \\ R_{M-1} & R_{M-2} & \cdots & R_0 + \sigma^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{bmatrix} = \begin{bmatrix} R_0 \\ R_1 \\ \vdots \\ R_{M-1} \end{bmatrix} \quad (6)$$

式中: $R_\tau = R_{d,d}(\tau)$.

2 WF 各抽头系数之和的分析

首先,根据式(6)推导出系数和的表达式,其次基于平稳随机过程的假设,证明视频编码恢复图像信号的自相关矩阵为正定矩阵,然后基于实验统计给出图像信号自相关函数的特性,得出可接受视频编码失真条件下,失真功率远小于视频图像信号的自相关函数,最后得出用于补偿编码失真的 WF 各抽头系数之和近似等于 1 的结论.

2.1 抽头系数之和的推导

将 WF 求解方程组(6)的系数行列式记为 D ,根据克莱姆法则知,当 $D \neq 0$ 时,线性方程组存在唯一解,且方程组解的各项之和——最佳 WF 抽头系数之和可表示为

$$\sum_{i=1}^M w_i = \sum_{i=1}^M \frac{D_i}{D} = \frac{1}{D} \sum_{i=1}^M D_i \quad (7)$$

式中: D_i 是系数行列式 D 中的第 i 列元素换成常数项得到的行列式. 进一步化简可得

$$\sum_{i=1}^M w_i = 1 - \frac{\sigma^2}{R_0 + \sigma^2} \xi \quad (8)$$

式中

$$\xi = \frac{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \theta_1 & 1 & \cdots & \theta_{M-2} \\ \vdots & \vdots & & \vdots \\ \theta_{M-1} & \theta_{M-2} & \cdots & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \theta_1 & \cdots & \theta_{M-1} \\ \theta_1 & 1 & \cdots & \theta_{M-2} \\ \vdots & \vdots & & \vdots \\ \theta_{M-1} & \theta_{M-2} & \cdots & 1 \end{vmatrix}}, \quad \theta_i = \frac{R_i}{R_0 + \sigma^2} \quad (9)$$

2.2 抽头系数之和的分析

2.2.1. 求解方程组的系数行列式

引理 离散随机过程的自相关矩阵是正定或半正定矩阵^[7].

推论 叠加了独立的均值为 0 的白噪声的平稳随机过程的自相关矩阵是正定矩阵.

证明 如图 1 所示,假设 β 是任意 M 维非 0 复向量,定义 p' 为 β 与随机过程 p 的观测向量 p 的内积,则有

$$p' = \beta^H p \tag{10}$$

p' 的均方期望为

$$\begin{aligned} E[|p'|^2] &= E[\beta^H p p^H \beta] = \\ &\beta^H E[pp^H] \beta = \beta^H R_{p,p} \beta \end{aligned} \tag{11}$$

根据式(1)及式(5),下式成立

$$\begin{aligned} \beta^H R_{p,p} \beta &= \beta^H (R_{n,n} + R_{d,d}) \beta = \\ &\beta^H R_{n,n} \beta + \beta^H R_{d,d} \beta \end{aligned} \tag{12}$$

式中: $R_{d,d}$ 为正定或半正定矩阵,即二次型 $\beta^H R_{d,d} \beta \geq 0$; $R_{n,n}$ 为均值为 0 的白噪声的自相关矩阵

$$R_{n,n} = \begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix} \tag{13}$$

显然, $R_{n,n}$ 的各阶顺序主子式都大于 0, 那么 $R_{n,n}$ 为正定矩阵, 即有二次型 $\beta^H R_{n,n} \beta > 0$, 二次型 $\beta^H R_{p,p} \beta > 0$, 则根据二次型性质得出 $R_{p,p}$ 为正定矩阵, 推论得证.

由正定矩阵的性质知, WF 求解方程组的系数行列式大于 0.

2.2.2 图像信号的空间自相关函数 实平稳随机过程的自相关函数具有对称和有界的特点, 即式(14)和式(15)成立

$$R_{d,d}(\tau) = R_{d,d}(-\tau) \tag{14}$$

$$R_{d,d}(0) \geq R_{d,d}(\tau) \tag{15}$$

通常情况下, 图像信号可以看成是实平稳随机过程, 图像的空间自相关性用图像中当前位置 (x, y) 的像素与邻近位置 $(x + \Delta x, y + \Delta y)$ 的像素之积的均值 $R(\Delta x, \Delta y)$ 来刻画, 定义

$$R_{d,d}(\Delta x, \Delta y) = E[d(x, y) d(x + \Delta x, y + \Delta y)] \tag{16}$$

本文对大量图像信号的自相关函数进行了统计实验, 图 2 给出了具有代表性的图像 baboon(512 × 512 像素) 的归一化自相关统计数据.

从图 2 可以观察得出, 图像信号的空间自相关函数满足对称和有界的性质, 即在原点 $(0, 0)$ 处取得最大值, 并随着与原点距离的增大单调下降, 即呈现单峰形状, 同时下降也变得越来越缓慢. 为了便于准确描述空间自相关函数的变化趋势, 以水平方向上的自相关函数为例, 描述如下:

- (1) 当 $M \leq 5$ 时, $R_{d,d}(0, 0) > R_{d,d}(1, 0) > \dots > R_{d,d}(M - 1, 0)$, 如果记 θ'_i 为归一化自相关函数 $\frac{R_{d,d}(i, 0)}{R_{d,d}(0, 0)}$, 那么 $1 > \theta'_1 > \dots > \theta'_{M-1} > 0$;
- (2) 当 $M \leq 5$ 时, $\theta'_{M-1} > 0.8$.

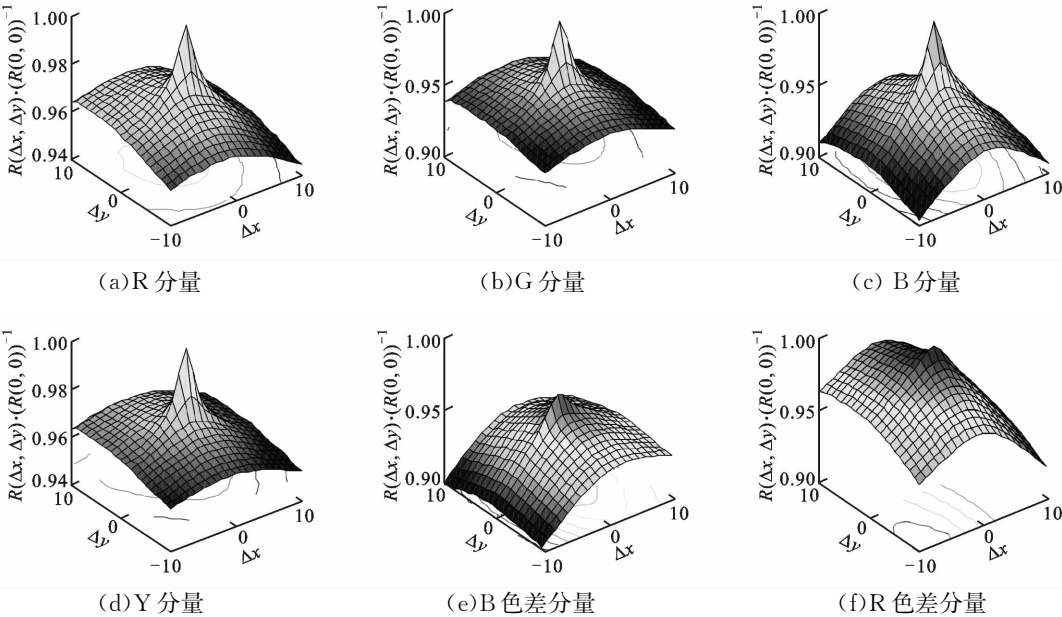


图 2 baboon 图像的各分量归一化自相关函数

2.2.3 视频编码失真功率 目前视频客观质量评估方法中,峰值信噪比(PSNR)被广泛用来度量视频编码中引起的失真.如果视频编码失真的均值为0,则视频编码的失真功率就是其方差 σ^2 ,将峰值信噪比记作 P ,则有

$$P = 10 \lg \frac{N^2}{\sigma^2} \quad (17)$$

式中: N^2 为视频信号的峰值功率.对于常见8 b/像素的视频图像信号, N 为255.根据实际经验,对于亮度分量,PSNR高于40 dB说明图像质量极好(即非常接近原始图像),在30~40 dB之间通常表示图像质量是好的(即失真可以察觉但可以接受),在20~30 dB之间说明图像质量差,PSNR低于20 dB则图像质量不可接受.因此,对于可接受的视频失真(PSNR大于30 dB)而言,根据式(17)有 $255^2/\sigma^2 > 10^3$.假设8 b图像信号在0~255范围内均匀分布, $R_{d,d}(0) = E[|d|^2] = 255^2/3$.因此,从统计角度讲视频图像自相关函数远大于可接受的视频失真功率,即 $\sigma^2 \ll R_{d,d}(0) = R_0$.

2.2.4 抽头系数之和近似等于1 根据上述结论,在可接受的视频失真情况下, $\sigma^2 \ll R_{d,d}(i) = R_i$,因此,式(9)中的 $\theta_i = \frac{R_i}{R_0 + \sigma^2}$ 和 $\frac{R_i}{R_0}$ 相同,也满足2.2.2节中论述的对称、有界且单调减的特性.基于此,根据不同的抽头系数数量 M 以及 θ_i 的最小值 θ_{M-1} 设计统计实验,通过实验结果证明式(8)中 ξ 在一个较小的区间内波动.具体步骤如下:在给定 M 和 θ_{M-1} 的条件下,首先在区间 $(1, \theta_{M-1})$ 产生 $M-2$ 个随机

数,从大到小分别记为 $\eta_1, \eta_2, \dots, \eta_{M-2}$;然后令 $\theta_1 = \eta_1, \dots, \theta_{M-2} = \eta_{M-2}$,使得式(9)中 θ_i 满足 $1 > \theta_1 > \dots > \theta_{M-1} > 0$ 的条件;最后计算 ξ 的值.每个实验统计50 000组随机数计算出结果,并根据计算得到的 ξ 绘制直方图.图3分别给出了抽头数为9个和7个, θ_{M-1} 分别为0.96、0.90、0.82和0.72的 ξ 的统计直方图.

根据图3中的统计实验结果,得到如下较为宽松的约束条件

$$|\xi| < 1 \quad (18)$$

同时,根据2.2.3节的结论 $\sigma^2 \ll R_0$,可得 $\frac{\sigma^2}{R_0 + \sigma^2} \xi \approx 0$,因此下式成立

$$\sum_{i=1}^M w_i = 1 - \frac{\sigma^2}{R_0 + \sigma^2} \xi \approx 1 \quad (19)$$

由以上分析可以得到结论:在可接受的视频编码失真条件下,用于补偿编码失真的WF各抽头系数之和近似等于1.

3 DCTCC 算法

首先计算视频编码引起的直流偏移并精确补偿,使得视频编码失真的均值为0,满足第2节结论成立的前提条件,然后以直流补偿后的解码恢复图像作为WF的输入信号,原始图像作为WF的期望信号,设计最佳滤波器,最后利用得到的WF补偿解码恢复图像编码失真,并量化编码各非中心抽头系数以及直流补偿值.

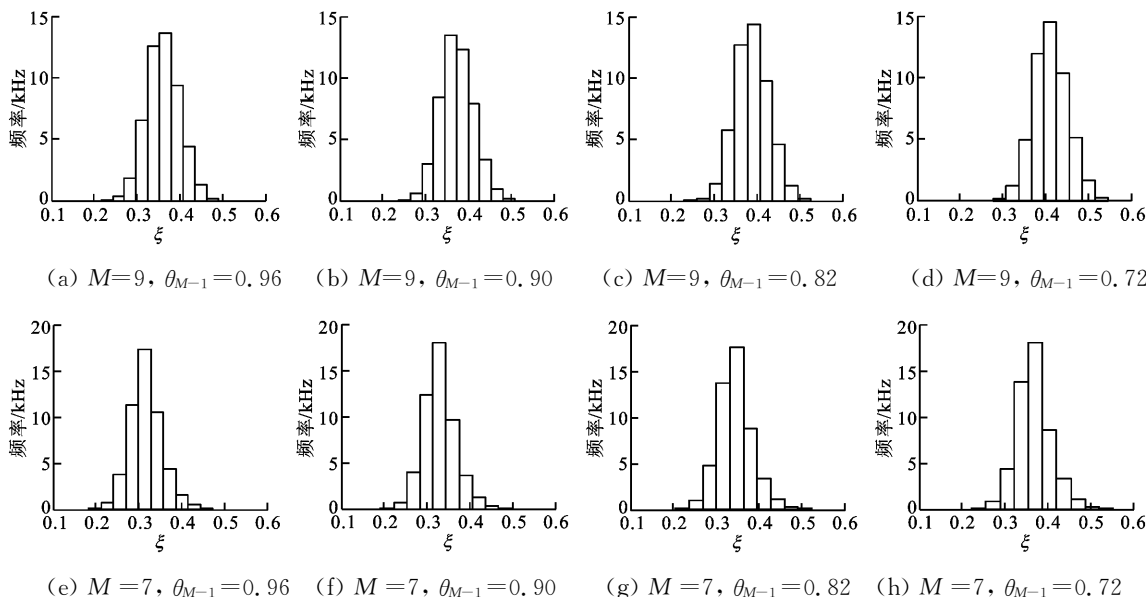


图3 ξ 的直方图

由于原始图像中当前位置像素与滤波器窗口中心位置对应的恢复图像像素的相关性最强,因此中心抽头系数远大于其他抽头系数,该系数编码开销最大.基于此,DCTCC 不对该系数进行量化编码,而是根据系数之和等于 1 的关系式,通过其他系数推导得到.同已有方法相比,DCTCC 方法的特点是对于编码开销最大的中心抽头系数不编码,在编码不引起直流偏移、视频编码失真可接受的情况下,利用 WF 各抽头系数之和等于 1 的分析结论,推导得到中心抽头系数.由于视频编码从统计角度讲不会引入直流偏移,DCTCC 引入了直流补偿值是为了防止可能出现的直流偏移,该补偿值的编码开销很小,并能精确修正直流误差.

4 实验数据

本文的实验采用 H. 264/AVC 标准参考代码 KTA2. 6^[8]作为仿真平台,并将本文方法 DCTCC 集成到 KTA2. 6 中.在现有的视频量化噪声补偿方法中,文献[4]的 BALF 方法由于提高编码效率显著,被 KTA 作为编码工具.为了全面比较本文方法和 H. 264/KTA2. 6 基准方法以及 BALF 方法,实验依照视频标准化组织推荐的测试条件^[9],选择了 12 个标准视频序列,其中 QCIF 格式 3 个,CIF 格式 4 个,720p 格式 5 个,测试序列的参数详见表 1.实验采用 IPPP 预测结构,分别在 Baseline 和 High Profile 这 2 个标准测试条件下进行比较.

表 1 测试序列的参数

序列名	分辨率/像素	帧率/帧·s ⁻¹	帧数
Container- qcif	176×144	15	100
Foreman- qcif	176×144	15	100
Silent- qcif	176×144	15	100
Foreman- cif	352×288	30	200
Mobile- cif	352×288	30	200
Paris- cif	352×288	15	150
Tempete- cif	352×288	30	200
BigShips	1 280×720	60	100
City	1 280×720	60	100
Crew	1 280×720	60	100
Night	1 280×720	60	100
ShuttleStart	1 280×720	60	100

为了定量地比较本文方法、BALF 和 KTA 定义的基准方法,实验采用 VCEG 组织推荐的由

Bjontegaard 提出的编码效率衡量方法^[10].该方法用 G 来衡量 2 条率失真曲线映射在 PSNR 上的平均增益,而用 B 来衡量 2 条率失真曲线映射在比特率上的平均节省量,从而可以有效表示改进算法相对于基准方法的提高程度.计算 G/B 的 4 组数据点所使用的 Q_P 参数同样采用文献[9]推荐的 4 个典型值,即 22、27、32 和 37.表 2 给出了 DCTCC 方法相对于 H. 264/KTA2. 6 参考代码的编码增益.为了便于比较,表 2 同时列出了 BALF 方法相对于 H. 264/KTA2. 6 参考代码的编码增益.从表 2 可以看出,在 Baseline 和 High Profile 的测试条件下,对于不同格式的测试序列,DCTCC 方法都可有效提高 H. 264/KTA2. 6 参考代码的编码效率,并优于 BALF 方法.

表 2 亮度分量的编码增益

测试序列	B/%			
	Baseline		High Profile	
	BALF ^[4]	DCTCC	BALF ^[4]	DCTCC
Container- qcif	11. 28	11. 52	3. 20	3. 43
Foreman- qcif	4. 02	4. 44	3. 17	3. 39
Silent- qcif	3. 61	3. 92	2. 96	9. 19
Foreman- cif	7. 89	8. 11	6. 11	6. 23
Mobile- cif	7. 92	8. 09	6. 89	7. 03
Paris- cif	2. 93	3. 05	2. 17	2. 27
Tempete- cif	4. 72	4. 84	3. 82	3. 91
BigShips	12. 14	12. 45	9. 64	9. 85
City	21. 96	22. 21	15. 96	16. 15
Crew	18. 09	18. 30	8. 29	8. 46
Night	6. 41	6. 75	5. 20	5. 43
ShuttleStart	25. 92	26. 31	16. 09	16. 31

表 3 给出了在相同的软、硬件条件下 DCTCC 和 BALF 方法的编、解码时间.表 3 数据表明,DCTCC 方法的计算量比 BALF 略有增加,这是由于 DCTCC 与 BALF 方法相比较,增加了一个直流偏移补偿模块,而 WF 方程组建立、抽头系数求解以及系数量化编码等其他算法模块的运算量和内存空间都没有明显的额外增加.相反地,编码器中 DCTCC 方法节省了一个抽头系数的量化编码,解码器中 DCTCC 方法也仅需要为恢复该抽头系数而增加系数求和的加法运算,这部分开销相对于算法中的其他模块而言可以忽略不计.

表 3 平均编解码时间比较

测试条件	平均编码时间/s		平均解码时间/s	
	BALF ^[4]	DCTCC	BALF ^[4]	DCTCC
Baseline	5 633.325	5 681.473	40.811	42.536
High Profile	10 836.235	10 998.383	39.496	41.326

5 结 论

本文分析了视频编码中 WF 的应用原理,化简了最佳滤波器抽头系数的求解方程组,推导出抽头系数之和的表达式,并通过对求解方程组系数行列式、视频编码失真特点和视频图像信号的空间自相关函数 3 个方面的分析,得出在可接受视频失真的情况下 WF 各抽头系数之和近似等于 1 的结论.利用该结论,针对现有 WF 抽头系数编码方法存在冗余信息的不足,提出了改进算法.实验结果表明本文所提方法可在保证视频恢复图像质量的同时提高视频编码效率,从而验证了本文对 WF 抽头系数之和的分析的合理性.

参考文献:

[1] Telecommunication Standardization Sector. ITU-T recommendation H. 264 advanced video coding for generic audiovisual services[S]. Geneva, Switzerland: Telecommunication Standardization Sector, 2009.

[2] WIEGAND T, SULLIVAN G. The H. 264/AVC video coding standard[J]. IEEE Signal Processing Magazine, 2007, 24(2):148-153.

[3] CHUJOH T, TANIZAWA A, YAMAKAGE T. Adaptive loop filter for improving coding efficiency[EB/OL]. (2008-04-23) [2009-03-27]. <http://www.itu.int/ITU-T/meetingdocs/index.asp>.

[4] WATANABE T, WADA N, YASUDA G, et al. In-loop filter using block-based filter control for video coding[C]//Proceedings of IEEE International Conference on Image Processing. Piscataway, NJ, USA: IEEE, 2009:1013-1016.

[5] CHUJOH T, YASUDA G, WADA N, et al. Improvement of block-based adaptive loop filter[EB/OL]. (2008-10-10) [2009-06-15]. http://ftp3.itu.int/av-arch/video-site/0810_San/VCEG-AJ13.zip.

[6] WIENER N. The extrapolation, interpolation and smoothing of stationary time series[M]. Cambridge,

USA: MIT Press, 1949.

[7] HAYKIN S. Adaptive filter theory [M]. 4th ed. New York, USA: Prentice Hall Inc. , 2002.

[8] Heinrich-Hertz-Institute. KTA reference software version JM11.0/KTA 2.6[CP/OL]. (2009-10-13) [2009-10-18]. <http://iphone.hhi.de/suehring/tml/download/KTA/jm11.0kta2.6.zip>.

[9] TAN T K, SULLIVAN G, WEDI T. Recommended simulation common conditions for coding efficiency experiments revision 1 [EB/OL]. (2007-01-14) [2008-07-08]. <http://ftp3.itu.int/av-arch/video-site/0701-Mar/VCEG-AE10.zip>.

[10] BJONTEGAARD G. Calculation of average PSNR differences between RD-curves[EB/OL]. (2001-03-28) [2006-02-05]. http://ftp3.itu.int/av-arch/video-site/0104_Aus/VCEG-M33.doc.

[本刊相关文献链接]

多扫描双向参考的快速帧内预测算法. 西安交通大学学报,2010,44(4):98-102.

一种方向提升小波变换和网格编码量化的图像编码算法. 西安交通大学学报,2010,44(2):67-71.

任意视点电视系统中深度数据的自适应非均匀量化方法. 西安交通大学学报,2010,44(2):82-87.

视频通信中的选择性重传技术. 西安交通大学学报,2009,43(6):82-87.

适合于小波视频编码的尺寸自适应图像组结构. 西安交通大学学报,2009,43(4):52-55.

采用帧延迟的实时编码协作方案. 西安交通大学学报,2009,43(2):63-66.

可分级视频编码覆盖网非一致性随机成员构建算法. 西安交通大学学报,2009,43(6):1-4.

利用 MPEG-4 视频流的运动目标检测算法. 西安交通大学学报,2008,42(4):454-457.

平板电视数字视频后处理芯片的设计与实现. 西安交通大学学报,2008,42(10):1285-1289.

一种八叉树编码加速的 3D 纹理体绘制算法. 西安交通大学学报,2008,42(12):1490-1494.

保证 H. 264 视频通信质量的实时传输协议载荷. 西安交通大学学报,2007,41(12):1455-1459.

基于视频的车辆瞬时停车延误检测. 西安交通大学学报,2007,41(6):692-695.

(编辑 刘杨)