

一种用于视频编码的低复杂度整数变换^{*}

周敬利 向东 陈加忠 余胜生

(华中科技大学计算机科学与技术学院 武汉 430074)

摘要 文中提出一种用于视频编码的低复杂度整数变换。与 H. 264 标准中所采用的整数变换相比较,本整数变换非常接近于 DCT 变换,而且无需乘法计算,只需要 16 位的加法和移位操作,大大降低了编码的计算复杂度,因此特别适合于低端的处理器。试验结果表明,该变换能够最大获得 1.56% 的比特率节省和 0.16 dB 的 PSNR 增益。

关键词 视频编码,整数变换,H. 264

A Low-Complexity Integer Transform for Video Coding

ZHOU Jing-Li XIANG Dong CHEN Jia-Zhong YU Sheng-Sheng

(College of Computer Science and Technology, Huazhong University of Science & Technology, Wuhan 430074)

Abstract This paper presents a new low-complexity integer transform for video coding. Unlike the 4×4 transform in H. 264, the new 4×4 transform matrix is much approximated to DCT matrix. Moreover, this transform can be calculated without multiplication, just additions and shifts, in 16-bit arithmetic, thus minimizing computational complexity, especially for low-end processor. Simulation results reveal a performance increasing up to 1.56% overall bit rate savings and 0.16 dB in peak signal-to-noise ratio.

Keywords Video coding, Integer transforms, H. 264

1 简介

离散余弦变换(DCT)被广泛用于图像和视频编码标准中。由于 DCT 反变换矩阵并没有用整数完全确定,从而增加了数据漂移(编码端和解码端解码数据的不一致)的可能^[1]。H. 264 视频标准采用了 4×4 整数变换,能够精确进行 2 进制计算,因此避免了编码端和解码端解码数据不一致的问题。除此之外,采用 4×4 变换能够减少振铃效应,提高重建视频图像的主观质量^[2]。

本文提出了一种新的整数变换,该变换无需乘法计算,只需要 16 位的加法和移位操作。与 H. 264 标准中的变换相比较,该整数变换更接近 DCT 变换。试验结果表明,该变换能够最大获得 1.56% 的比特率节省和 0.16 dB 的峰值信噪比(PSNR)增益。

2 整数变换的选择

对于大多数信号而言,DCT 变换的性能最接近统计最优的 K-L 变换^[3]。2 维 4×4 DCT 正变换可以表示为:

$$Y = AXA^T = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix} X \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix}$$

其中 X 为 4×4 的输入矩阵, A 是 DCT 变换矩阵, A^T 是矩阵 A 的转置,其中:

$$a = 0.5, b = \sqrt{0.5} \times \cos(\pi/8), c = \sqrt{0.5} \times \cos(3\pi/8)$$

对上式进行简单的运算,可得到:

$$Y = (CX C^T \otimes E_f) X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ d & 1 & -1 & -d \\ 1 & -1 & -1 & 1 \\ 1 & -d & d & 1 \end{bmatrix} X$$

$$\begin{bmatrix} 1 & d & 1 & 1 \\ 1 & 1 & -1 & -d \\ 1 & -1 & -1 & d \\ 1 & -d & 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} a^2 & ac & ba^2 & ac \\ ac & c^2 & ac & c^2 \\ a^2 & ac & a^2 & ac \\ ac & c^2 & ac & c^2 \end{bmatrix}$$

其中 $d = b/c$, \otimes 表示矩阵的对应元素相乘。矩阵 C 的行或列元素的范数不为 1,也就是 $CC^T \neq I$, I 为单位阵,所以需要在矩阵 E_f 中进行补偿。矩阵 E_f 中的元素为常数,因此可以把它们集成到量化过程中,这就降低了变换的计算复杂度。反向变换过程可以表示为:

$$X = C^T (Y \otimes E_f) C = \begin{bmatrix} 1 & d & 1 & 1 \\ 1 & 1 & -1 & -d \\ 1 & -1 & -1 & d \\ 1 & -d & 1 & -1 \end{bmatrix} \begin{bmatrix} a^2 & ac & a^2 a & ac \\ ac & c^2 & ac & c^2 \\ a^2 & ac & a^2 & ac \\ ac & c^2 & ac & c^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ d & 1 & -1 & -d \\ 1 & -1 & -1 & 1 \\ 1 & -d & d & -1 \end{bmatrix}$$

在 DCT 变换矩阵中,参数 $d = 1/(\sqrt{2}-1) = 2.4142\dots$,为无理数,因此不能精确地用二进制表示。我们可以通过整数变换来解决这个问题。为了得到接近于 DCT 变换的整数变换,参数 d 的值必须在 2.4 左右,参数 b, c 必须为整数,并且

^{*} 本课题得到国家“863”高技术研究发展计划(2002AA11901010)资助。周敬利 教授,博士生导师,研究方向为多媒体通信技术及其应用、高性能网络存储技术;向东 博士研究生,研究方向为视频压缩与通讯;陈加忠 博士,研究方向为小波理论和信号分析、图像与视频编码技术;余胜生 教授,博士生导师,研究方向为计算机网络与存储、多媒体处理与通信。

矩阵 C 必须为正交阵。然而,无论 d 为何值,矩阵 C 总为正交阵。同时,我们注意到,当参数 b, c 的值已确定时,相应的变换矩阵也已确定。为方便起见,在下文中,我们用变换 (b, c) 来表示由参数 b, c 所确定的变换矩阵。

在图像和视频编码中,变换矩阵通常用来去除输入信号之间的相关性。这种去相关性可以通过变换效率 η 来衡量^[4]: η 越大,去相关性能越好。设向量 X 为一维零均值、相关系数为 ρ 的马尔可夫过程的采样,其协方差矩阵 C_X 的第 (i, j) 个元素为 $\rho^{|i-j|}$ 。变换矩阵 T 的变换效率可以表示为:

$$Y = T \cdot X, C_Y = E(Y \cdot Y^T) = T C_X T^T$$

$$= \begin{bmatrix} s_{11} & \dots & \dots & s_{1n} \\ \vdots & & & \\ \vdots & & & \\ s_{m1} & \dots & \dots & s_{mn} \end{bmatrix}, \eta = \frac{\sum_{i=1}^n s_{ii}}{\sum_{p=1}^m \sum_{q=1}^n s_{pq}}$$

为了得到变换效率最高的矩阵,我们对 c 小于 8(3 比特)和 d 小于 5 条件下不同变换的 η 进行了计算,同时也比较了 DCT 变换。表 1 给出了当 $\rho=0.9$ 时的计算结果。从表中可以看出,当 d 约等于 2.4 时,去相关性最佳。为了决定最佳的 (b, c) 值,我们还必须考虑以下两个条件^[5]:

- 1) 为了减少计算复杂度, (b, c) 的值越小越好,这样加法和乘法操作的次数将会更少。
- 2) 为了提高去相关性能,整数变换矩阵每行元素的比应越接近 DCT 变换矩阵越好。

表 1 当 $\rho=0.9, c$ 分别等于 1, 2, ..., 7 时的最佳变换

变换效率 $\eta(\%)$	变换 (b, c)
95.24	(2, 1)
95.62	(5, 2)
95.89	(7, 3)
95.76	(9, 4)
95.78	(12, 5)
95.59	(13, 6)
95.82	(16, 7)
95.73	(17, 7)
95.75	DCT

在 H. 264 标准中,最先采用了变换 (17, 7),但它需要 32 位的数学运算,因此最终采用了变换 (2, 1)。但变换 (2, 1) 并不能很好地接近 DCT 变换矩阵,而变换 (5, 2) 比变换 (2, 1) 更接近 DCT 变换矩阵,其值也较小。因此我们选择变换 (5, 2) 作为上面两个条件的折衷,此时 $a=0.5, c=\sqrt{2/29}$ 。变换 (5, 2) 的正变换可以表示为:

$$Y = (C_f X C_f^T) \otimes E_f = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 5 & 2 & -2 & -5 \\ 1 & -1 & -1 & 1 \\ 2 & -5 & 5 & -2 \end{bmatrix} X$$

$$\begin{bmatrix} 1 & 5 & 1 & 2 \\ 1 & 2 & -1 & -5 \\ 1 & -2 & -1 & 5 \\ 1 & -5 & 1 & -2 \end{bmatrix} \otimes \begin{bmatrix} a^2 & ac/2 & a^2 & ac/2 \\ ac/2 & c^2/4 & ac/2 & c^2/4 \\ a^2 & ac/2 & a^2 & ac/2 \\ ac/2 & c^2/4 & ac/2 & c^2/4 \end{bmatrix}$$

对 8 比特的输入像素数据而言,预测误差需要 9 比特来表示。变换矩阵 C_f 的每一行绝对值之和的最大值为 14,因此变换后的变换系数需要 $9 + \log_2 14 \times 14 = 17$ 位来表示。由于许多处理器执行 32 位乘法运算的代价大于 16 位的乘法运算,因此我们希望把变换系数值的范围限制在 16 比特内。对

上面公式中矩阵 C_f^T 的第 2 列和第 4 列分别除 2,则正变换又可以表示为:

$$Y = (C_f X C_f^T) \otimes E_f = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 5 & 2 & -2 & -5 \\ 1 & -1 & -1 & 1 \\ 2 & -5 & 5 & -2 \end{bmatrix} X$$

$$\begin{bmatrix} 1 & 5/2 & 1 & 2 \\ 1 & 1 & -1 & -5/2 \\ 1 & -1 & -1 & 5/2 \\ 1 & -5/2 & 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} a^2 & ac & a^2 & ac \\ ac/2 & c^2/2 & ac/2 & c^2/2 \\ a^2 & ac & a^2 & ac \\ ac/2 & c^2/2 & ac/2 & c^2/2 \end{bmatrix}$$

上面矩阵变换后的变换系数需要 $9 + \log_2 14 \times 7 = 16$ 位来表示,我们可以用右移操作来代替除 2 的除法操作。当操作数为奇数时,右移操作会导入截断误差,但这些误差会被量化过程中的误差所覆盖^[6]。相应的反变换可以表示为:

$$X = C_f^T (Y \otimes E_f) C_f = \begin{bmatrix} 1 & 5/2 & 1 & 1 \\ 1 & 1 & -1 & -5/2 \\ 1 & -1 & -1 & 5/2 \\ 1 & -5/2 & 1 & -1 \end{bmatrix}$$

$$\left(Y \otimes \begin{bmatrix} a^2 & ac & a^2 & ac \\ ac & c^2 & ac & c^2 \\ a^2 & ac & a^2 & ac \\ ac & c^2 & ac & c^2 \end{bmatrix} \right) \begin{bmatrix} 1 & 1 & 1 & 1 \\ 5/2 & 1 & -1 & -5/2 \\ 1 & -1 & -1 & 1 \\ 1 & -5/2 & 5/2 & -1 \end{bmatrix}$$

为了降低计算复杂度,2 维变换可以先进行 1 行变换,再进行 1 列变换。图 1 给出了反变换的行变换和列变换的流程。我们采用左移和右移操作来代替乘法和除法操作,如乘 5 可以表示为 1 次左移 2 位操作和 1 次加法操作。因此,正变换和反变换无需乘法操作,只须加法和移位操作。但是,变换 (5, 2) 在进行行变换或列变换时,比变换 (2, 1) 多两次加法和移位操作。

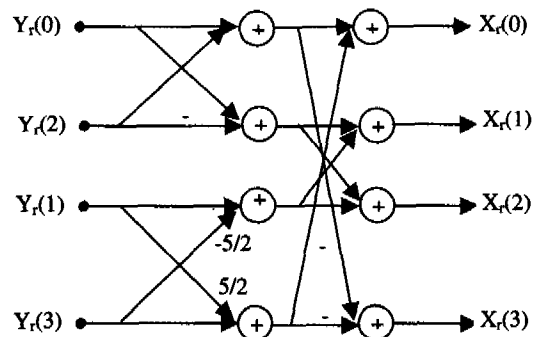


图 1 反变换的快速算法

3 量化和反量化

在有损压缩中,量化过程虽然损失了数据的精度,但换来了大的压缩比。变换 (5, 2) 的量化过程与 H. 264 标准中的量化过程相似。H. 264 标准采用了无需除法和浮点运算的标量量化,量化参数 (QP) 每增加 6,则量化步长约增加一倍。对于给定的 QP,量化过程可以表示为^[11]:

$$Y_q(i, j) = \text{Sign}(Y(i, j)) [(|Y(i, j)| \cdot A(Q_M, i, j) + f \cdot 2^{P+Q_E}) \gg (P+Q_E)]$$

其中 Q_E, Q_M 是 QP 除以 6 的商和余数; (i, j) 是行和列元素的索引值;参数 f 由编码器选择,其值一般在 0 到 1/2 之间。为了减小截断误差,参数 P 的值越大越好,但必须同时保证

(下转第 145 页)

5 Alonso G, Agrawal D, Abbadi A E, Kamath M, et al. Advanced Transaction Models in Workflow Contexts. In: Proc. of the Twelfth Intl Conf. on Data Engineering, 1996

6 Garcia-Molina H, Salem K. Sagas. In: Proc. 1987 SIGMOD Intl. Conf. on Management of Data, May 1987. 249~259

7 Endl R, Knolmayer G, Pfahrer M. Modeling Processes and Workflows by Business Rules. In: Proc. of the 1st European Workshop on Workflow and Process Management, Zurich, Switzerland, 1998. 47~56

8 WfMC. Workflow Management Coalition. Interface 1: Workflow Process Definition Interface-XML Process Definition Language. WfMC TC-1025, 2002

9 服务的商业流程执行语言. <http://dev2dev.bea.com/Web-services/BPEL4WS.jsp>

10 李红臣, 史美林. Petri 网在业务过程建模中的应用. 小型微型计算机系统, 2001, 22(1)

11 袁崇义. Petri 网原理与应用. 北京: 电子工业出版社, 2005

12 van der Aalst W M P. The Application of Petri Nets to Workflow Management. The Journal of Circuits, Systems and Computers, 1998, 8(1): 21~66

13 Jensen K. Coloured Petri Nets; Basic Concepts, Analysis Methods and Practical Use, Vol. 1: Basic Concepts, EATCS Monographs on Theoretical Computer Science, Springer, Berlin, 1992

(上接第 131 页)

矩阵 A 的元素值在 16 比特范围内, 我们选取 P=17. 反量化过程可以表示为:

$$Y_r(i, j) = [Y_q(i, j)B(Q_M, i, j) \gg (8 - Q_E)]$$

在量化和反量化过程中, 矩阵 A 和 B 的元素可以表示为: $A(Q_M, i, j) = M(Q_M, r)$, $B(Q_M, i, j) = S(Q_M, t)$. 其中, 当 $(i, j) = \{(0, 0), (0, 2), (2, 0), (2, 2)\}$ 时, $r=0, t=0$; 当 $(i, j) = \{(1, 1), (1, 3), (3, 1), (3, 3)\}$ 时, $r=1, t=1$; 当 $(i, j) = \{(0, 1), (0, 3), (2, 1), (2, 3)\}$ 时, $r=2, t=2$; 其他情况时, $r=3, t=2$. 矩阵 M 和 S 的元素分别表示为:

$$M = \begin{bmatrix} 52429 & 7231 & 27537 & 13768 \\ 47663 & 6574 & 25304 & 12517 \\ 40330 & 5563 & 21182 & 10591 \\ 37449 & 5165 & 19669 & 9835 \\ 32768 & 4520 & 17211 & 8605 \\ 29127 & 4018 & 15298 & 7649 \end{bmatrix}$$

$$S = \begin{bmatrix} 5120 & 1412 & 2689 \\ 5632 & 1554 & 2958 \\ 6656 & 1836 & 3496 \\ 7168 & 1977 & 3765 \\ 8192 & 2260 & 4303 \\ 9216 & 2542 & 4840 \end{bmatrix}$$

4 试验结果

我们通过修改 H. 264 参考软件 JM7. 3 来执行新的变换和量化方法. 我们对几个 QCIF 和 CIF 格式的视频序列进行了测试. 实验条件为 QP 等于 16, 20, 28 和 32, 两个参考帧, 运动向量的搜索范围为 32, 用 IPP...P 帧格式编码, 编码频率为 30Hz. 图 2 给出了变换(5, 2)和变换(2, 1)的率失真曲线比较, 表 2 给出了平均 PSNR 增益和平均比特率节省.

表 2 变换(5, 2)和变换(2, 1)的率失真性能比较

序列		平均 PSNR 增益 (dB)	平均比特率节省 (%)
QCIF	News	+0.063	+0.93
	FlowerGarden	+0.16	+1.56
	Foreman	+0.03	+0.76
	Silent	+0.085	+1.28
CIF	News	+0.036	+0.67
	Mobile	+0.03	+0.41
	Foreman	+0.033	+0.74
	Tempete	+0.058	+0.97
	Flowergarden	+0.06	+0.8
	Container	+0.002	+0.1
Average		+0.056	+0.82

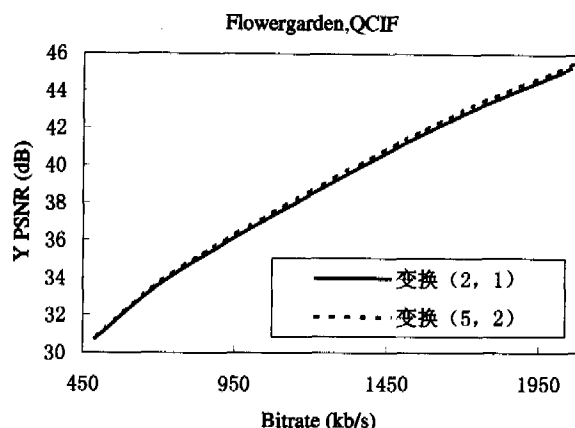


图 2 率失真曲线比较

从图 2 和表 2 中我们可以看到, 变换(5, 2)的率失真性能优于变换(2, 1), 其平均 PSNR 增益为 0.056dB. 变换(5, 2)的缺点是在正变换和反变换的列变换或行变换的计算需要多 2 次加法和移位, 但采用该变换带来的性能增益足以弥补计算量的增加, 因为加法和移位操作的计算非常简单.

结论 文中提出一种新的 4×4 整数变换矩阵, 该整数变换矩阵非常接近 DCT 变换, 并且无需乘法运算, 只需要 16 位的加法和移位操作, 因此大大降低了计算的复杂度. 试验结果表明: 该变换在率失真性能上优于变换(2, 1), 其平均 PSNR 增益为 0.056 dB.

参考文献

1 Malvar H S, Hallapuro A, Karczewicz M, et al. Low-complexity transform and quantization in H. 264/AVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 598~603

2 Wiegand T, Sullivan G J, Bjontegaard G, et al. Overview of the H. 264 / AVC video coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 1~19

3 Rao K R, Yip P. Discrete Cosine Transform: Algorithms, Advantages, Applications. Boston: Academic Press, 1990

4 Cham W. Development of integer cosine transforms by the principle of dyadic symmetry [J]. Proceedings of IEE Communications, Speech and Vision, 1989, 136: 276~282

5 Pei Soo-Chang, Ding Jian-Jiun. The integer transforms analogous to discrete trigonometric transforms [J]. IEEE Trans on Signal Processing, 2000, 48(12): 3345~3364

6 Wu F S, Cham W K. A comparison of error behaviour in the implementation of the DCT and the ICT [C]. IEEE Region 10 Conference on Computer and Communication Systems, Tencon, 1990