第 13 卷　第 6 期
2008 年 6 月
中国图象图形学报
Journal of Image and Graphics
Vol. 13, No. 6
June, 2008

整数 DCT 变换中变换基的通用生成算法

王中元　胡瑞敏　田　纲　李　明
(武汉大学国家多媒体软件工程技术研究中心,武汉　430072)

摘　要　H.264 是 ITU 与 ISO 联合共同开发的具有高编码效率、高压缩质量的视频新标准。整数变换是其提高压缩性能最主要的改进方法之一,基于同样的整数变换过程的变换基可以不唯一,因此在整数变换的理论确定后,寻找变换基是一件重要的工作,提出了一种通用的变换基生成算法。该方法通过分析整数 DCT 变换的原理,指出了整数变换矩阵应该满足的 4 个约束条件,以满足正交性约束为出发点,导出了整数矩阵元素之间的数量关系,并辅以另外 3 个约束条件,采用搜索的方法寻找变换基。实验结果表明,该算法在经过几十步的搜索后,就能找出所有可用的变换基,JVT 参考模型用到的变换基也在其中。

关键词　H.264　整数 DCT 变换　变换基
中图法分类号:TN919.8　　**文献标识码**:A　　**文章编号**:1006-8961(2008)06-1061-04

The Generic Generating Algorithm for Integer DCT Transform Radix

WANG Zhong-yuan, HU Rui-min, TIAN Gang, LI Ming
(*National Multimedia Software Engineering Research Center, Wuhan University, Wuhan* 430072)

Abstract　H.264 is the new video coding standard established by ITU and ISO, which has high coding efficiency and high compression quality. Integer DCT transform is one of the main renovation for compression efficiency improvement. For the integer DCT radix cannot be the unique within the same transform framework, finding the radix is another valuable research topic besides integer DCT theory. A generic generating algorithm for integer DCT transform radix is presented in this paper. Based on the mathematic analysis of integer DCT transform principle, four constraint conditions which shall be met by integer DCT transform matrix are given first. And then, starting from orthogonal constraint, the magnitude relation of matrix elements is formulized. The formulation, which combines the other three constraint conditions, results in a transform radix finding method by using search strategy. The simulation results show that, only through tens of seeking steps, the proposed method can find all valid radix including the one used by JVT reference model.

Keywords　H.264, integer DCT (discrete cosine transform), transform radix

1　引　言

新一代视频编码标准 H.264 是由国际电信联合会(international telecommunication union, ITU)与国际标准组织(international standards organization, ISO)的联合视频小组(joint video team, JVT)负责开发,它是面向实际应用的最新发展的标准。其目标是基于高的视频分辨率,提高图像质量,并能够覆盖所有低带宽和高带宽的应用。H.264 是在 ITU-T 增强型多媒体通信标准 H.26L 基础上推出的能够为 ITU-T 和 ISO/IEC (international electrotechnical commission)共同使用的新一代视频编码标准,并且在技术上同 MPEG(moving picture experts group)标准形成体系。

H.264 与先前的标准相似,对残差数据采用

基金项目:国家自然科学基金项目(60472040)
收稿日期:2006-08-18;**改回日期**:2007-01-16
第一作者简介:王中元(1972～　),男。2001 年获取武汉大学计算机应用专业硕士学位,现为武汉大学国家多媒体软件工程中心博士研究生。主要研究方向为视频编码、多媒体通信。E-mail:wzy_hope@163.com

基于块的变换编码，变换编码可以去除原始图像的空间冗余，使图像能量集中在一小部分系数上，这样可以提高压缩比增强抗干扰能力。而 H.264 采用的变换是近似 DCT（discrete cosine transform）的无乘法整数变换，在此称之为整数 DCT 变换。需要注意的是，此处的变换已经不是真正的 DCT，仍然称其为 DCT 变换只是为了说它是由 DCT 推导而来，且为了和另一个变换（Hadamard变换）相区别。H.264 的整数 DCT 变换中只有整数运算，消除了浮点运算，减少了运算量，并且精确的整数排除了编码器和解码器反变换之间的误差匹配问题[1]。

整数变换基不是唯一的，国际上一些大公司都提出了自己的变换基，并申请了专利保护，尽管这些变换基具有很好的变换性质，但一般都是通过试探的办法发现的，发现的过程带有偶然性，效率极低而且不全面。本文基于整数 DCT 变换的原理，提出了一种变换基的通用生成算法，从根本上解决了变换基的发现问题。该成果对于制定有自主知识产权的视频压缩标准具有积极意义。

2 整数 DCT 变换原理

为了达到正向变换和反向变换的可逆运算，整数 DCT 变换矩阵经过归一化后必须是正交矩阵[2]。若用 $\boldsymbol{I}$ 表示 8×8 整数 DCT 正向变换矩阵，$\boldsymbol{G}$ 表示归一化矩阵，$\boldsymbol{I}_0$ 为归一化后的正文矩阵，则归一化过程为(符号“ ⊗ ”表示点乘运算)

$$\boldsymbol{I}_0 = \boldsymbol{I} \otimes \boldsymbol{G}$$

$$\boldsymbol{G} = \begin{bmatrix} g_0 & g_0 & g_0 & \cdots & g_0 \\ g_1 & g_1 & g_1 & \cdots & g_1 \\ \vdots & \vdots & \vdots & & \vdots \\ g_7 & g_7 & g_7 & \cdots & g_7 \end{bmatrix}$$

$$g_i = \frac{1}{\sqrt{\sum_{j=0}^{7} I_{ij}^2}} \qquad i = 0,1,\cdots,7$$

若正向变换表示为

$$\boldsymbol{Y} = \boldsymbol{I}_0 \boldsymbol{X} \boldsymbol{I}_0^{\mathrm{T}}$$

则根据正交矩阵的性质，反向变换为

$$\boldsymbol{X} = \boldsymbol{I}_0^{\mathrm{T}} \boldsymbol{Y} \boldsymbol{I}_0$$

代入 $\boldsymbol{I}_0$，根据矩阵运算规律，容易证明

$$\boldsymbol{Y} = \boldsymbol{I}\boldsymbol{X}\boldsymbol{I}^{\mathrm{T}} \otimes \boldsymbol{S}$$
$$\boldsymbol{X} = \boldsymbol{I}^{\mathrm{T}}(\boldsymbol{Y} \otimes \boldsymbol{S})\boldsymbol{I}$$
$$\boldsymbol{S} = \boldsymbol{G} \otimes \boldsymbol{G}^{\mathrm{T}}$$

上式意味着 DCT 变换转化为整数变换后，点乘运算分别在量化和逆量化过程中完成，这也是整数变换没有精度匹配损失的原因。

下面给出上述两个结论的证明。先列出证明的依据，根据矩阵运算性质[3]，矩阵运算满足下述运算规律：

交换律：

$$\boldsymbol{A} \otimes \boldsymbol{B} = \boldsymbol{B} \otimes \boldsymbol{A}$$

结合律：

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$$
$$(\boldsymbol{A} \otimes \boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{B} \otimes \boldsymbol{C}) = \boldsymbol{A} \otimes (\boldsymbol{BC})$$
$$(\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C} = \boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C})$$

转置运算规律：

$$(\boldsymbol{A} \otimes \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} \otimes \boldsymbol{B}^{\mathrm{T}}$$

首先，证明结论 $\boldsymbol{Y} = \boldsymbol{I}\boldsymbol{X}\boldsymbol{I}^{\mathrm{T}} \otimes \boldsymbol{S}$。

证明：代入 $\boldsymbol{I}_0 = \boldsymbol{I} \otimes \boldsymbol{G}$ 到 $\boldsymbol{Y} = \boldsymbol{I}_0 \boldsymbol{X} \boldsymbol{I}_0^{\mathrm{T}}$ 中，得到

$$\begin{aligned}
\boldsymbol{Y} &= (\boldsymbol{I} \otimes \boldsymbol{G})\boldsymbol{X}(\boldsymbol{I} \otimes \boldsymbol{G})^{\mathrm{T}} & \\
&= (\boldsymbol{I} \otimes \boldsymbol{G})\boldsymbol{X}(\boldsymbol{I}^{\mathrm{T}} \otimes \boldsymbol{G}^{\mathrm{T}}) & \text{（转置运算）} \\
&= (\boldsymbol{G} \otimes \boldsymbol{I})\boldsymbol{X}(\boldsymbol{I}^{\mathrm{T}} \otimes \boldsymbol{G}^{\mathrm{T}}) & \text{（交换律）} \\
&= \boldsymbol{G} \otimes (\boldsymbol{IX})(\boldsymbol{I}^{\mathrm{T}} \otimes \boldsymbol{G}^{\mathrm{T}}) & \text{（结合律）} \\
&= \boldsymbol{G} \otimes (\boldsymbol{IXI}^{\mathrm{T}}) \otimes \boldsymbol{G}^{\mathrm{T}} & \text{（结合律）} \\
&= (\boldsymbol{IXI}^{\mathrm{T}}) \otimes \boldsymbol{G} \otimes \boldsymbol{G}^{\mathrm{T}} & \text{（交换律）} \\
&= (\boldsymbol{IXI}^{\mathrm{T}}) \otimes (\boldsymbol{G} \otimes \boldsymbol{G}^{\mathrm{T}}) & \text{（结合律）}
\end{aligned}$$

代入 $\boldsymbol{S} = \boldsymbol{G} \otimes \boldsymbol{G}^{\mathrm{T}}$，则 $\boldsymbol{Y} = \boldsymbol{IXI}^{\mathrm{T}} \otimes \boldsymbol{S}$，结论成立。

同理，可证 $\boldsymbol{X} = \boldsymbol{I}^{\mathrm{T}}(\boldsymbol{Y} \otimes \boldsymbol{S})\boldsymbol{I}$。

但问题是如何寻找整数变换矩阵 $\boldsymbol{I}$，整数变换矩阵需要满足以下几个特性[4]：

（1）能归一化成正交矩阵；

（2）变换运算可全部用移位运算实现；

（3）归一化后的正交矩阵同原始的 DCT 变换矩阵系数相近；

（4）整数变换矩阵系数不是很大，变换运算能用 16 位精度表示。

下面根据这几个原则，提出一种寻找整数变换矩阵的通用算法。

3 变换基生成算法

列出 8×8 DCT 变换系数矩阵[5]如下：

$$\begin{bmatrix} 0.353553 & 0.490393 & 0.461940 & 0.415735 & 0.353553 & 0.277785 & 0.191342 & 0.097545 \\ 0.353553 & 0.415735 & 0.191342 & -0.097545 & -0.353553 & -0.490393 & -0.461940 & -0.277785 \\ 0.353553 & 0.277785 & -0.191342 & -0.490393 & -0.353553 & 0.097545 & 0.461940 & 0.415735 \\ 0.353553 & 0.097545 & -0.461940 & -0.277785 & 0.353553 & 0.415735 & -0.191342 & -0.490393 \\ 0.353553 & -0.097545 & -0.461940 & 0.277785 & 0.353553 & -0.415735 & -0.191342 & 0.490393 \\ 0.353553 & -0.277785 & -0.191342 & 0.490393 & -0.353553 & -0.097545 & 0.461940 & -0.415735 \\ 0.353553 & -0.415735 & 0.191342 & 0.097545 & -0.353553 & 0.490393 & -0.461940 & 0.277785 \\ 0.353553 & -0.490393 & 0.461940 & -0.415735 & 0.353553 & -0.277785 & 0.191342 & -0.097545 \end{bmatrix}$$

系数共有 7 种取值：0.353 553、0.490 393、0.415 735、0.277 785、0.097 545、0.461 940、0.191 342，依次记为：c_0、c_1、c_2、c_3、c_4、c_5、c_6。令它们对应的整数 DCT 变换矩阵系数分别为 i_0、i_1、i_2、i_3、i_4、i_5、i_6，则整数矩阵 $\boldsymbol{I}$ 可写为

$$\begin{bmatrix} i_0 & i_0 & i_0 & i_0 & i_0 & i_0 & i_0 & i_0 \\ i_1 & i_2 & i_3 & i_4 & -i_4 & -i_3 & -i_2 & -i_1 \\ i_5 & i_6 & -i_6 & -i_5 & -i_5 & -i_6 & i_6 & i_5 \\ i_2 & -i_4 & -i_1 & -i_3 & i_3 & i_1 & i_4 & -i_2 \\ i_0 & -i_0 & -i_0 & i_0 & i_0 & -i_0 & -i_0 & i_0 \\ i_3 & -i_1 & i_4 & i_2 & -i_2 & -i_4 & i_1 & -i_3 \\ i_5 & -i_6 & i_6 & -i_5 & -i_5 & i_6 & -i_6 & i_5 \\ i_4 & -i_3 & i_2 & -i_1 & i_1 & -i_2 & i_3 & -i_4 \end{bmatrix}$$

为了满足正交性，考虑

$$\boldsymbol{II}^{\mathrm{T}} = \begin{bmatrix} a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b & 0 & -d & 0 & d & 0 & 0 \\ 0 & 0 & c & 0 & 0 & 0 & 0 & 0 \\ 0 & -d & 0 & b & 0 & 0 & 0 & d \\ 0 & 0 & 0 & 0 & a & 0 & 0 & 0 \\ 0 & d & 0 & 0 & 0 & b & 0 & d \\ 0 & 0 & 0 & 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & d & 0 & d & 0 & b \end{bmatrix}$$

式中，

$$a = 8i_0^2$$
$$b = 2(i_1^2 + i_2^2 + i_3^2 + i_4^2)$$
$$c = 4(i_5^2 + i_6^2)$$
$$d = i_3(i_1 + i_4) - i_2(i_1 - i_4)$$

性质(1)决定矩阵元素 d 必须为零，问题转化为如何寻找一组数 i_1、i_2、i_3、i_4 满足

$$i_3(i_1 + i_4) = i_2(i_1 - i_4)$$

的同时满足性质(2)～(4)。显然，可以用搜索的方法寻找，为了满足性质(2)，本文取搜速步长为 1/2，为满足性质(4)，取最大幅度为 15；为满足性质(3)，将系数 i_0、i_1、i_2、i_3、i_4、i_5、i_6 归一化得到标定系数 j_0、j_1、j_2、j_3、j_4、j_5、j_6，

$$j_k = \begin{cases} \dfrac{i_k}{\sqrt{a}} & k = 0 \\ \dfrac{i_k}{\sqrt{b}} & k = 1,2,3,4 \\ \dfrac{i_k}{\sqrt{c}} & k = 5,6 \end{cases}$$

因此，性质(3)可表述为

$$\sum_{k=0}^{6} |j_k - c_k| < T \quad k = 0,1,\cdots,6$$

式中，T 表示精度控制门限。考虑到大的系数对变换后的能量贡献较大，故加权，加权因子定义成

$$f_k = \frac{c_k}{\sum_{j=0}^{j=6} c_j} \quad k = 0,1,\cdots,6$$

所以，上式进一步写成

$$\sum_{k=0}^{6} |j_k - c_k| f_k < T_f \quad k = 0,1,\cdots,6$$

门限 T_f 取决于期望的候选变换基数目，本文取 0.03。

4　实验结果

以 i_1、i_2、i_3、i_4 为例，按照同原始 DCT 系数的偏差从小到大排列的一组变换基，如表 1 所示。其中 5 号基正好是 JVT 参考软件中用到的[6]。

4×4 整数变换的基也可以用类似的方法计算，DCT 系数 0.653 282、0.270 598 对应的整数变换系数如表 2 所示。其中 6 号基正好是 JVT 参考软件中用到的。需要说明的是，通过修改归一化矩阵 $\boldsymbol{G}$，变换基都可以尺度化到整数。

表 1　8×8 整数变换基
Tab. 1　8×8 integer DCT radix

序号	变换基				标定后基				偏差
	i_1	i_2	i_3	i_4	j_1	j_2	j_3	j_4	
0	12.5	10	7	2.5	0.492 760	0.413 919	0.275 946	0.098 552	0.001 971
1	12	10.5	7.5	2	0.478 471	0.418 662	0.299 045	0.079 745	0.011 475
2	13.5	11	7	3	0.502 244	0.409 236	0.260 423	0.111 610	0.011 478
3	7.5	6	4	1.5	0.504 505	0.403 604	0.269 069	0.100 901	0.011 481
4	10	9	6	2	0.475 651	0.428 086	0.285 391	0.095 130	0.011 481
5	12	10	6	3	0.499 134	0.415 945	0.249 567	0.124 784	0.011 604
6	7	6	4.5	1	0.480 196	0.411 597	0.308 697	0.068 599	0.014 149
7	6	5	2.5	2	0.502 625	0.418 854	0.209 427	0.167 542	0.025 839

表 2　4×4 整数变换基
Tab. 2　4×4 integer DCT radix

序号	变换基		标定后基		偏差
	DCT 0.653 282	DCT 0.270 598	DCT 0.653 282	DCT 0.270 598	
0	6	2.5	0.652 714	0.271 964	0.000 802
1	5	2	0.656 532	0.262 613	0.004 637
2	7	3	0.649 934	0.278 543	0.004 695
3	4.5	2	0.646 162	0.287 183	0.009 892
4	8	3	0.662 085	0.248 282	0.012 761
5	3	1	0.670 820	0.223 607	0.026 165
6	2	1	0.632 456	0.316 228	0.028 091

5　结　论

本文分析了 H.264 标准的整数变换和量化的原理,并在此基础上提出了一种通用的变换基生成算法。实验结果表明,应用该算法不仅可以找到 JVT 参考模型中使用的基,而且可以发现一组变换基,并按照它们同原始 DCT 系数的差距进行了对比排列。按照这个顺序,在设计新的视频压缩标准时可以挑选未用的变换基,从而避开国外公司已经进行了专利保护的基。这一发现对于制定有自主知识产权的本土视频压缩标准具有积极意义。

参考文献(References)

1 Hao Peng-wei, Shi Qing-yun. The inter implementation of reversible linear transforms [J]. China Science (Series E), 2000, **30**(2): 132～141.

2 Henrique S M, Antti H, Marta K, *et al*. Low-complexity transform and quantization in H.264/AVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, **13**(7): 598～602.

3 Mathematics Teaching and Research Section of Tongji University, Linear algebra[M]. Beijing: Higher Education Press, 1991.8. [同济大学数学教研室编,线性代数[M]北京:高等教育出版社,1991.8.]

4 ITU-T Recommendation H.264 & ISO/IEC 14496-10 AVC, Advanced Video Coding for Generic Audiovisual Services[S], 2003.

5 ITU-T Recommendation H.263, Video Coding for Low Bit Rate Communication[S], 1998.

6 JVT Reference Software Official Version JM61e [EB/OL], http://bs.hhi.de.html.