

一种类 DCT 的 8×8 整数变换

田晓华¹⁾ 朱光喜¹⁾ 王 曜²⁾

¹⁾(华中科技大学电信系宽带无线通信与多媒体系统研究中心, 武汉 430074)

²⁾(香港中文大学讯息工程系)

摘 要 整数变换因其高效的计算性能而被最新的国际视频编码标准 H. 264 采纳, 我国的第一个音视频编码标准 AVS 也将采用 8×8 整数变换。变换模块的性能直接影响编解码器的性能, 所以选择性能优异的变换基尤为重要。本文在 8×8 类 DCT 整数变换原理的基础上, 提出了一种结合能量集中率和去相关效率的变换基变换性能评价方法, 同时考虑计算复杂度, 在理论上选出了几组性能较好的整数变换基。通过分析视频序列测试结果, 讨论得出不同变换基的适用场合, 并且验证了理论评价结果的正确性。

关键词 整数变换 离散余弦变换 能量集中率 去相关效率 音视频编码标准

中图法分类号: TN911 **文献标识码:** A **文章编号:** 1006-8961(2004)08-0916-06

A DCT-like 8×8 Integer Transform

TIAN Xiao-hua¹⁾, ZHU Guang-xi¹⁾, WANG Yao²⁾

¹⁾(Broad Band Wireless Communication & Multimedia System R&D Centre, Department of Electronics & Information Engineering, Huazhong University of Science & Technology, Wuhan 430074)

²⁾(Department of Information Engineering, Chinese University of Hong Kong)

Abstract Integer transform has been adopted by the newest international video coding standard H. 264 for its high efficiency of computation. Our first national audio and video coding standard AVS will also adopt 8×8 integer transform for video compression. As transform module has strong influence on the efficiency of codec, it is important to select an excellent transform basis (transform matrix). In this paper, after introduction of basic theory of DCT-like 8×8 integer transform, the authors bring out three principles of transform basis evaluation; (1) orthogonality, (2) transform efficiency, and (3) computation complexity. First, the transform bases satisfied principle (1) can be found in a predefined data range. Principle (2) is embodied in a novel method combining evaluation of transform efficiency from two aspects; energy compaction and decorrelation efficiency in a proper range of correlation coefficients of input image prediction difference data. By using such evaluation method and taking principle (3), computation complexity of fast transform algorithm into consideration, several theoretical highly efficient transform bases such as (10, 9, 6, 2), (5, 6, 4, 1) and (4, 5, 3, 1) have been picked out and tested in the actual video sequences. Based on careful analysis of testing results, the authors find out the proper situation that each basis fits for and prove the validity of transform basis performance measurement that uses this theoretical evaluation method.

Keywords integer transform, DCT, energy compaction efficiency, decorrelation efficiency, AVS

1 引 言

离散余弦变换(DCT)^[1]自从1974年提出后,在图像和视频编码领域得到了广泛的应用^[2~4],其变换性能在所有次优变换中非常显著,能极大地去除图像元素在变换域中的相关性,为高效率的图像压缩奠定

了基础^[1]。但由于DCT的变换矩阵是用浮点数表示,浮点运算量较大,占用较多系统资源。为了提高变换效率,DCT又发展出用定点计算或者较大的整型变换逼近浮点运算,由于存在精度误差,即使没有经过量化,反变换后也不能完全恢复图像数据,也就是编码的可逆性不强。整数变换的产生,同时解决了计算精度误差大和编码效率低的问题,其特点是用整数变

换矩阵代替 DCT 的浮点数变换矩阵,这样变换过程完全是整数运算,不存在精度误差,保证了编码的可逆性;同时整数乘法可用加减法和位移代替,因此变换过程可以完全通过加减法和位移实现,运算量大幅度减少。近年来,在图像和视频处理领域有不少针对整数变换的研究^[5-7]。ITU 的 VCEG、还未完成的标准 H. 26L 均采用了整数变换^[8,9],并且最新的国际标准 H. 264/MPEG-4 Part 10^[10]也采用了整数变换,并取得了非常好的变换效果。

当前我国将要制定的第一个音视频编码标准(AVS)是面向高清晰电视(HDTV)的应用,但由于 H. 264 中采用的 4×4 整数变换的去相关性不足,需要再对变换后的直流系数做 Hadamard 变换。因此 AVS 决定采用去相关性更强的 8×8 整数 DCT 变换^[11]。

2 整数变换原理

8×8 离散余弦变换可以表示^[1]为

$$Y(u, v) = \frac{1}{4}C(u)C(v) \sum_{j=0}^7 \sum_{k=0}^7 (X(j, k) \times \cos\left(\pi u \frac{2j+1}{16}\right) \cos\left(\pi v \frac{2k+1}{16}\right)) \quad (1)$$

其中, $u, v \in [0, 7], C(0) = 1/\sqrt{2}, C(w) = 1, (w \in [1, 7])$ 。用矩阵的形式可以表示为

$$Y = P_0 X P_0^T$$

其中, X 为 8×8 的图像帧内帧间预测的残差矩阵, Y 为变换后的矩阵,

$$P_0 = \begin{bmatrix} a & a & a & a & a & a & a & a \\ b & d & e & g & -g & -e & -d & -b \\ c & f & -f & -c & -c & -f & f & c \\ d & -g & -b & -e & e & b & g & -d \\ a & -a & -a & a & a & -a & a & a \\ e & -b & g & d & -d & -g & b & -e \\ f & -c & c & -f & -f & c & -c & f \\ g & -e & d & -b & b & -d & e & -g \end{bmatrix}$$

$$a = \frac{1}{2\sqrt{2}}, b = \frac{1}{2} \cos\left(\frac{\pi}{16}\right), c = \frac{1}{2} \cos\left(\frac{2\pi}{16}\right), \\ d = \frac{1}{2} \cos\left(\frac{3\pi}{16}\right), e = \frac{1}{2} \cos\left(\frac{5\pi}{16}\right), f = \frac{1}{2} \cos\left(\frac{6\pi}{16}\right), \\ g = \frac{1}{2} \cos\left(\frac{7\pi}{16}\right).$$

根据 JVT 对 4×4 DCT 变换的改写过程^[12],可以改写 8×8 变换如下:从 P_0 的每一行提出一个公共系数,得到向量 $V_8 = [a, g_0, f, g_0, a, g_0, f, g_0]$,则

变换矩阵改写为

$$P_1 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ k_1 & k_2 & k_3 & k_4 & -k_4 & -k_3 & -k_2 & -k_1 \\ k_5 & 1 & -1 & -k_5 & -k_5 & -1 & 1 & k_5 \\ k_2 & -k_4 & -k_1 & -k_3 & k_3 & k_1 & k_4 & -k_2 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ k_3 & -k_1 & k_4 & k_2 & -k_2 & -k_4 & k_1 & -k_3 \\ 1 & -k_5 & k_5 & -1 & -1 & k_5 & -k_5 & 1 \\ k_4 & -k_3 & k_2 & -k_1 & k_1 & -k_2 & k_3 & -k_4 \end{bmatrix}$$

$$k_1 = b/g_0, k_2 = d/g_0, k_3 = e/g_0, k_4 = g/g_0, k_5 = c/f.$$

令矩阵 $E_8 = V_8^T V_8$, 则变换进一步改写为

$$Y = P_1 X P_1^T \otimes E_8 \quad (2)$$

其中, \otimes 表示 2 个矩阵的叉乘(即相同位置的元素对应相乘)。对于式(2)可以将与矩阵 E_8 的叉乘运算和量化操作放到一起,使变换得到简化。因此变换的重点在于 $P_1 X P_1^T$ 的计算,其中 X 是 8×8 的像素预测残差矩阵,为整数数据。如果 P_1 中的变量 k_1, k_2, k_3, k_4, k_5 都为整数,则变换将全部转化为整数运算。因此就需要确定 k_1, k_2, k_3, k_4, k_5 等 5 个参数的选值。通过大量实验证明,当选定 k_1, k_2, k_3, k_4 后, k_5 取 2 时的变换性能最好,文献[13]中也提出了相似的结论。因而将 k_5 固定取值为 2,仅研究其余 4 个参数的选择,并定义 (k_1, k_2, k_3, k_4) 为整数变换的变换基。

3 变换基的选择

3.1 候选变换基的确定

在研究浮点 DCT 变换特征的过程中发现,一个好的整数变换基需要满足以下几个原则:

(1)变换正交性 正交变换的特性保证变换仅是对坐标系的旋转,而图像的能量维持不变。为了保证变换的正交性,式(2)中的 P_1 必须满足条件

$$P_1 \cdot P_1^T \otimes E_8 = I \quad (3)$$

式(3)的条件可以等价于 $P_1 P_1^T$ 为对角阵,叉乘 E_8 后结果为单位阵。

(2)能量集中性 DCT 变换的目的在于去除元素之间的相关性,使变换后的能量尽可能集中在较少的几个系数上,以便提高量化后熵编码的压缩效率。整数变换基的选择同样也需要遵循这一原则。

(3)快速变换简单性 计算复杂度是衡量变换模块性能的重要指标,因此变换基的变换性能接近时,快速变换的运算次数越少,变换基的综合性能优

势越大。在实时性要求高的应用中,有可能牺牲一定的编码增益,以减小计算量的开销。

在文献[5]中要求待选基的 $k_1 \geq k_2 \geq k_3 \geq k_4$, 然而,这样的限制没有实质性的意义,同时漏选了部分优秀的整数变换基,因此没有采用这个限制。实际视频测试表明,数值较大的变换基的变换性能并没有显著提高,其快速算法的计算复杂度明显增大;而变换基数值较小的变换有可能设计出更加有效的快速算法,减少运算时间。另外,实验结果表明 k_4 取值较大时的变换性能并不好,可设定较小的取值范围。因此设定变换基的搜寻范围为 $k_1, k_2, k_3 \in [1, 10], k_4 \in [1, 4]$ 。

根据原则 1 的原理和原则 3 的限制范围,在设定的范围内循环搜索,得到满足正交条件的 56 组候选基。其中有大家已经较为熟悉的(10, 9, 6, 2)、(6, 6, 3, 2)、(3, 2, 1, 1),还有一些以前没有人提出来的(5, 6, 4, 1)、(4, 5, 3, 1)等。

3.2 变换基变换性能的度量

由于图像间的差异很大,单幅图像不能代表所有图像的特性,因此如果通过比较具体图像的变换域特性来衡量变换基的性能,很难对变换基性能做出科学评价。本文采用一种类似文献[5]、[9]、[14]的简单而又能较好表现图像像素点之间相关特性的一阶马尔科夫模型。由于式(2)中对图像残差的二维整数类 DCT 变换实质上是先后通过水平和垂直方向两次一维变换,使图像能量在两个方向上都得到集中,因此图像变换的性能取决于对一维向量变换的性能。

从图像的 8×8 残差块任取一行或者一列,得到长度为 8 的一维向量 $X_v = [x_1, x_2, \dots, x_8]^T$, 设相邻元素之间的互相关系数为 ρ , ($\rho \leq 1$), 则 x_i 和 x_j 元素之间的互相关系数为 $\rho^{|i-j|}$ 。因为残差的均值为零,所以该向量的第 i, j 元素之间的协方差为 $\rho^{|i-j|}$, 即协方差矩阵 $COV(X_v)$ 的第 i 行 j 列元素为 $\rho^{|i-j|}$ 。设变换矩阵 P 是对式(2)中的整数变换阵 P_1 归一化后得到的正交阵, P 对 X_v 做变换,得到 $Y_v = PX_v$, 易知,变换后的协方差矩阵为

$$COV(Y_v) = P \cdot COV(X_v) \cdot P^T \quad (4)$$

文献[14]讨论了式(4)中各元素的意义。其中,协方差矩阵的主对角线元素是一维向量元素的方差,而非主对角线元素是向量元素之间的协方差。因为变换的输入数据 x_i 为图像预测残差(-255~255),其均值为零,所以协方差矩阵对角线 $COV(X_v)_{(i,i)}$ 元素满足以下关系

$$COV(X_v)_{(i,i)} = \sigma_{x_i}^2 = E(x_i^2) - \bar{x}_i^2 = E(x_i^2)$$

即 x_i 的方差为其能量的均值。同理,变换后的协方差矩阵对角线元素 $COV(Y_v)_{(i,i)}$ 也表示 y_i 的能量均值。

3.2.1 能量集中率的比较

一些文献中通过变换的编码增益 G_{TC} (transform coding gain) 来衡量变换的能量集中 (energy compact) 性能^[8,15,16], G_{TC} 定义为变换域元素方差的算术平均值与几何平均值的比值。

因为式(4)中的正交变换 P 不改变向量 X_v 的能量,显然变换前后协方差矩阵主对角线元素的和不变。但正交变换却可以改变向量元素之间的能量分布,性能好的变换可以将大部分能量集中到个别元素中。因此可以用变换后能量集中的程度衡量变换的性能。由数学均值不等式的性质可知,能量分布越平均,元素能量的乘积越大;能量分布越集中,能量的乘积越小。所以对于正交变换 P , 变换的能量集中性能仅与变换域元素的方差的乘积或者几何平均值有关。

定义衡量变换能量集中性能的能量集中率为

$$\eta_E = \frac{1}{\sqrt[8]{\prod_{i=1}^8 COV(Y_v)_{(i,i)}}} \quad (5)$$

因为不同的相关系数 ρ 对应不同的输入数据协方差矩阵 $COV(X_v)$, 通过式(4)、(5)会得到不同的 η_E 。 ρ 的取值反映了图像残差特征, ρ 越大, 相邻像残差素点之间的相关性越强, 图像残差的细节信息较少, 冗余大, 可压缩性强; ρ 越小, 图像残差内的纹理越丰富, 细节信息量多, 冗余小。当 ρ 取值在 0.05~0.95 范围内, 所有 56 组候选变换基的 ρ - η_E 曲线如图版 I 图 1 所示。

由图版 I 图 1 可以看出, 能量集中率 η_E 随着 ρ 的增加而加速递增, 但各组变换基在相同 ρ 下的 η_E 较接近, 不易区分性能优劣, 为了使评价更合理, 采用类似罗马尼亚选择法^[17]的多指标综合评价方法。在相同 ρ 下, 采用式(6)对各组基的 η_E 标准化, 将 η_E 映射到 0~1 之间, 标准化后的 ρ - η_E 曲线如图版 I 图 2 所示, 从该图中比较容易区分各组基曲线的高低。

由于整数变换输入图像残差数据的互相关系数 ρ 并不是均匀分布在 0~1 之间的, 在 $\rho \in [0, 1]$ 的整个范围内评价 η_E 性能并不合理。采用文献[14]的方法对各类视频图像残差计算一维互相关系数 ρ , 得到 ρ 主要分布在 0.75~0.95 之间。 $\rho=0.95$ 时, 图像残差的细节信息较少; $\rho=0.75$ 时, 图像残差的细节信息较大。因此对变换基在 ρ 为 0.75~0.95 范围

内的 η_E 做整体评价可以反映变换基对各种视频图像的变换综合性能。

先计算所有候选基在 ρ 为 0.75、0.8、0.85、0.9、0.95 这 5 个采样点的 η_E ，再对相同的 ρ 下各组基的 η_E 标准化，得到第 i 组基的标准化结果为

$$Eval_{(i)} = \frac{\eta_{E(i)} - \text{Min}(\eta_{E(i)})}{\text{Max}(\eta_{E(i)}) - \text{Min}(\eta_{E(i)})} \quad (6)$$

再设定 5 点 ρ 值对应结果的权重，权重向量为 $w = [w_1, w_2, w_3, w_4, w_5]$ 。将各点标准化结果加权求和，得到第 i 组基最终的性能评价结果 $Eval_E$ 。因为图像相邻残差数据相关系数大的可能性高，对应的权重也较大。设定权重向量为 $w = [1/15, 2/15, 3/15, 4/15, 5/15]$ ，得到设定范围内性能最好的 10 组变换基。

3.2.2 去相关效率的比较

在文献[14]中，Clarke 用去相关效率来定义整数变换基的性能。他认为变换后的协方差能量越小，表明变换去相关的效率越高，则变换效果越好。例如最优的 KL 变换的非主对角线元素全为零，即协方差能量为零，去相关性的效率最高。因此，他通过计算变换前后协方差矩阵非主对角线元素(协方差)的改变，来衡量变换性能。文献[14]中定义的去相关效率为

$$\eta_C = 1 - \frac{\sum_{j \neq k} |COV(Y_V)_{(j,k)}|}{\sum_{j \neq k} |COV(X_V)_{(j,k)}|} \quad (7)$$

根据式(7)，计算各个候选变换基在相关系数 $[0.05, 0.95]$ 范围内的去相关效率，得到 ρ - η_C 曲线如图版 I 图 3 所示。由图版 I 图 3 可知去相关效率 η_C 随着相关系数 ρ 的增加而单调递增，但各组基 η_C 的差异随着 ρ 的增加而减小，在 $\rho \in [0.75, 0.95]$ 内已经很难区分。对各组候选基在相同 ρ 下的 η_C 标准化，得到标准化后的 ρ - η_C 曲线，如图版 I 图 4 所示。各组基在不同 ρ 下标准化的去相关效率容易区分。再运用上节的综合评价法，得到 $\rho \in [0.75, 0.95]$ 性能最好的 10 组基，其结果和能量集中度评价法选出的 10 组基完全相同，只是次序略有区别。

3.2.3 结合两种指标的综合评价

因为能量集中度 η_E 衡量的是式(4)中变换 P 对 $COV(Y_V)$ 主对角线元素能量的集中程度，而去相关效率 η_C 是衡量变换对 $COV(Y_V)$ 非主对角线元素能量的抑制能力，两者是从不同方面衡量变换基的性能。理想的变换基的去相关性能和能量集中性能都应该很优异，因此综合考虑 η_E 和 η_C ，设 $Eval_E$ 和 $Eval_C$ 分别是对能量集中性能和去相关效率的综合评价价值，

则通过加权平均得到对变换基的最终评价为

$$Eval = w_E \cdot Eval_E + w_C \cdot Eval_C \quad (8)$$

考虑到能量集中度直接影响变换后的压缩性能，相应的权重略大。这里取 $w_E = 0.6, w_C = 0.4$ ，得到评价结果最好的 10 组基，如表 1。理论上这几组基的能量集中率和去相关的性能都较好。

表 1 综合评价结果最好的 10 组基对 η_E 和 η_C 的评价结果和综合评价结果

变换基	$Eval_E$		$Eval_C$		$Eval$
	评价价值	排名	评价价值	排名	
10,9,6,2	0.9923	1	0.9762	1	0.9859
5,6,4,1	0.8901	2	0.8096	3	0.8579
6,6,3,2	0.8521	6	0.8322	2	0.8441
6,7,5,1	0.8804	3	0.7816	5	0.8409
4,5,3,1	0.8685	4	0.7595	7	0.8249
9,10,5,3	0.8363	9	0.7918	4	0.8185
7,8,6,1	0.8617	5	0.7530	8	0.8182
9,8,4,3	0.8416	7	0.7792	6	0.8167
8,9,7,1	0.8416	7	0.7316	9	0.7976
7,9,5,2	0.8307	10	0.7240	10	0.7880
DCT	1.0000		1.0000		1.0000

3.3 计算复杂度比较

整数变换用整数运算代替浮点运算，用加法和位移代替整数乘法，以尽量减少运算量。采用不同的变换基运算次数有很大区别，而相同的变换基采用不同的快速算法，运算次数也会改变。对一个 8×8 块做二维整数变换，可以分解成 16 次一维的 8 点变换，计算量为 一维变换的 16 倍。表 2 中列出了综合评价结果最好的 5 组变换基做一维 8 点变换的运算次数。反变换的计算次数与正变换相同。由表 2 可知，变换基(4, 5, 3, 1)的运算次数最少，(5, 6, 4, 1)其次。

表 2 综合评价结果较好的 5 组变换基一维变换运算次数和变换的动态范围

变换基	加法次数 (+/-)	位移次数 (<<)
10,9,6,2	36	10
5,6,4,1	32	6
6,6,3,2	36	10
6,7,5,1	32	10
4,5,3,1	28	6

4 视频序列测试结果

采用 JVT 的测试软件 JM5.0C 作为测试平台，变换部分强制采用 8×8 的整数变换，屏蔽块尺寸小于 8×8 的帧间搜索模式，量化采用类似 JVT 4×4

变换的量化策略。编码顺序为 IPBBPBB……。测试上述理论选择较好的 5 组变换基。每组基测试 AVS 提供的 7 个 CIF 序列 (bus_d、news、foreman、bus_f、container、garden、mobile, 其中, bus_d 和 bus_f 序列内容相似, 帧率不同), 每个序列分别在量化级数 QP 为 24、28、32、36 时编码 73 帧 (I_1, P_{24}, B_{48}), 得到码率和重建序列的亮度信噪比 SNR_Y 。采用 JVT 提供的曲线拟合软件 avsnr 比较变换基

的 RD 曲线的综合性能, avsnr 输入数据为 2 组变换基作测试时, 在上述 4 点 QP 的码率和亮度 SNR_Y 值, 通过比较 2 条拟合的 RD 曲线, 得到用 2 组基编码时 SNR_Y 和码率的平均差别, 从而确定 2 组基变换性能的差别。将理论选出最好的 5 组基中的 4 组与 (10, 9, 6, 2) 比较, 得到各组结果平均 SNR_Y 增加的 dB 和码率增加的百分比 (负号表示降低), 结果如表 3 所示。

表 3 4 组基的视频测试性能与 (10, 9, 6, 2) 的比较

CIF 序列	(5, 6, 4, 1)		(6, 6, 3, 2)		(4, 5, 3, 1)		(6, 7, 5, 1)	
	SNR_Y (dB)	Bits(%)	SNR_Y (dB)	Bits(%)	SNR_Y (dB)	Bits(%)	SNR_Y (dB)	Bits(%)
bus_d	0.0139	-0.4000	0.0012	-0.2082	0.0137	-0.4050	0.0069	-0.2688
news	-0.0536	0.9128	-0.0727	1.2857	-0.0918	1.6611	-0.0762	1.3155
foreman	-0.0044	0.1116	-0.0141	0.3490	-0.0124	0.3087	-0.0180	0.4439
bus_f	0.0244	-0.5447	0.0074	-0.2459	0.0235	-0.5312	0.0106	-0.2836
container	-0.0279	0.8550	-0.0141	0.4506	-0.0190	0.6097	-0.0262	0.8026
garden	0.1035	-1.9725	0.1243	-2.2857	0.0965	-1.8601	0.0995	-1.8629
mobile	0.0783	-1.5690	0.0881	-1.7672	0.0984	-1.9622	0.0626	-1.2461
平均	0.0191	-0.372	0.0171	-0.346	0.0156	-0.311	0.0085	-0.157

如果比较不同变换基在相同量化级数的性能, 从实验结果得到在 $QP=24$, 即在量化级数较低、码率较高时, (10, 9, 6, 2) 变换基的 RD 性能比其他几组基差, 同时在 $QP=36$ 时, 其性能也不是最优; 而在 QP 为 28 和 32 时其 RD 性能较好。因此 (10, 9, 6, 2) 变换基不适合在量化级数较大或较小的情况下工作, 在量化级数比较适中时其性能最好。(5, 6, 4, 1) 变换基在量化级数为 24 和 36 时的码率较小, SNR_Y 值较大, 综合性能最好; 在量化级数为 28、32 的性能没有明显优势。

观察试验采用的 7 个常用的标准视频测试序列可知, news、foreman、container 序列的场景变化较小, 只有人物头部或者主要物体的移动, 在此称之为慢场景; 而另外 4 个序列运动较为剧烈, 场景在不断变化, 称之为快场景。相应地, 慢场景编码输出码率较低, 而快场景所需要的编码比特较多。同时, 对于慢场景和快场景序列, 不同的变换基表现的变换性能差别较大。(10, 9, 6, 2) 变换基对慢场景序列的编码性能较好, 对快场景序列的性能较差; 而 (5, 6, 4, 1) 变换基的性能则恰好相反, 它对慢场景序列的编码性能仅次于 (10, 9, 6, 2) 变换基, 而对于快场景序列的性能最好, 如图 1、图 2 所示, 图中给出测试的 4 点 QP 下的 (10, 9, 6, 2) 变换基和 (5, 6, 4, 1) 变换基编码的率失真曲线。(4, 5, 3, 1) 变换基和 (6, 6, 3, 2) 变换基的性能略差于 (5, 6, 4, 1) 变换基, 率失真曲线和 (5, 6, 4, 1) 变换基非常接近。综合考虑计算的复杂

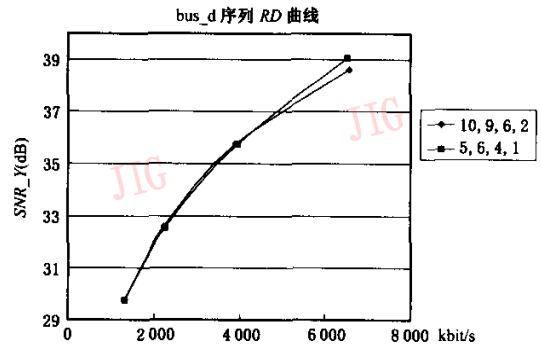


图 1 场景变化大的序列

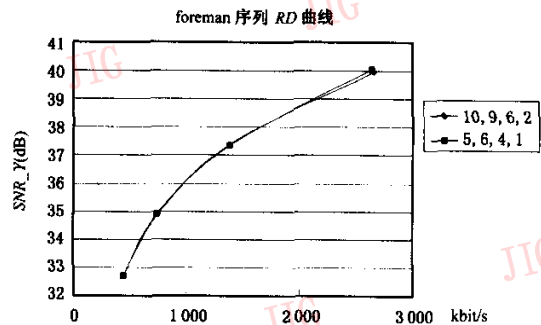


图 2 场景变化小的序列 (前 73 帧)

度和率失真特性, 在慢场景、低码率的场合使用 (10, 9, 6, 2) 变换基较好; 而在快场景、高码率的场合 (5, 6, 4, 1) 变换基是最佳选择, (4, 5, 3, 1) 变换基次之。

考虑 (6, 6, 3, 2) 这组基, 其去相关效率优于 (5, 6, 4, 1) 变换基, 实测性能比 (5, 6, 4, 1) 变换基差; 其

能量集中率劣于(6,7,5,1)变换基,而实测结果优于(6,7,5,1)变换基。单独从去相关效率或者能量集中率衡量变换性能,结论都不太准确。因此,实测结果验证了结合 η_E 和 η_C 的综合评价方式可以更加全面的比较整数变换基的性能。

5 结 论

提出了一种结合能量集中率和去相关效率的变换基性能评价标准。同时考虑计算复杂度和动态范围,在理论上选出了几组性能较好的整数变换基((10,9,6,2)、(5,6,4,1)、(6,6,3,2)、(6,7,5,1)、(4,5,3,1)),并通过实验验证了理论选择的正确性。通过实际序列测试,得出以下结论:

(1)在慢场景、低码率的场合使用(10,9,6,2)变换基较好;在快场景、高码率的场合(5,6,4,1)变换基是最佳选择,(4,5,3,1)变换基次之。

(2)从量化级数考虑,低量化级数编码时(5,6,4,1)变换基性能较好,量化级数适中(28~32)时(10,9,6,2)变换基性能较好。量化级数较高时(5,6,4,1)、(4,5,3,1)变换基的性能较好。

(3)如果忽略变换基之间小于 0.04dB 的编码增益的差别,从计算复杂度考虑,(4,5,3,1)变换基的运算量最小,是最佳选择。

通过理论和实测选择了几组率失真性能较好的 8×8 整数变换基,这几组基在特定方面的应用中具有较好优势,在 8×8 整数变换中有明显的应用前景。

参 考 文 献

- Ahmed N, Natarajan T, Rao K R. On image processing and a discrete cosine transform[J]. IEEE Transactions on Computers, 1974, C-23(1): 90~93.
- Sikora T. The MPEG-4 video standard verification model[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(1): 19~31.
- ITU Telecom. Standardization Sector of ITU, Video Coding for Low Bitrate Communication [S]. ITU-T Recommendation H. 263, Mar. 1996.
- Cote G, Erol B, Gallant M, et al. H. 263+: Video coding at low bit rates[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(7): 849~866.
- Giurcaneanu C D, Tabus I. Low-complexity transform coding with integer-to-integer transforms[A]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing[C], Salt Lake City, Utah, USA, 2001, 4: 2601~2604.
- Goyal V K. Transform coding with integer-to-integer transforms [J]. IEEE Transactions on Information Theory, 2000, 46(2): 465~47.
- Li Xin, Bo Tao, Orchard M T. On implementing transforms from integers to integers[A]. In: IEEE International Conference on Image Processing[C], New York, USA, 1998, 3: 881~885.
- Liang J, Tran T D. Fast multiplierless approximations of the DCT with the lifting scheme[J]. IEEE Transactions on Signal Processing, 2001, 49(12): 3032~3044.
- Malvar H. Low-complexity transform and quantization with 16-bit arithmetic for H. 26L [A]. In: Proceedings of IEEE International Conference on Image Processing[C], New York, USA, 2002, 2: II-489~II-492.
- Richardson, H. 264/MPEG-4 Part 10 "White Paper: Transform & Quantization"[EB/OL]. www.vcodex.fsnet.co.uk/h264-transform.pdf, 2003-4-30/2003-6-30.
- AVS 视频专题组. 编码技术测试通用条件[R]. AVS 决议 N1027, 2003.3.
- Hallapuro A, Karczewicz M, Malvar H. Low Complexity Transform and Quantization-Part I: Basic Implementation[S]. JVT-B038, 2002.
- Cham W K. Development of integer cosine transform by the principle of dyadic symmetry [J]. IEE Proceedings, 1989, 136 (4): 276~288.
- Clarke R J. Transform Coding of Images[M]. UK: Academic Press, 1985:19~31, 87~90.
- WEIDMANN Claudio. Image Transform Coding [EB/OL]. http://lcavwww.epfl.ch/~weidmann/SPclass/transform-codes/node10.html, 2000-02-01/2003-6-30.
- GirdBernd. Course Lecture: Image Communication I [EB/OL]. http://www.stanford.edu/class/ee398a/handouts/08-TransformCoding.pdf, 2003-2-25/2003-6-30.
- 谭跃进, 陈英武, 易进先. 系统工程原理[M]. 长沙: 国防科技大学出版社, 1999:191~193.

田晓华 1980年生。2004年于华中科技大学电子与信息工程系获硕士学位。研究方向为信源信道联合编码中的视频编码技术和率失真特性。
E-mail: tian1980xh@yahoo.com.cn



朱光喜 1945年生。教授、博士生导师。中国数字音视频编解码技术标准专家组成员。目前主要研究方向为多媒体通信、第四代移动通信等。



王 曦 1976年生。2003年于华中科技大学获信息与通信工程专业博士学位, 现在香港中文大学讯息工程系攻读博士后。主要研究方向为多媒体通信。



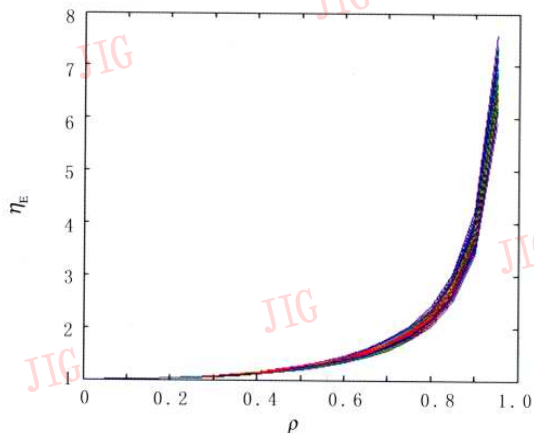


图1 56组候选变换基的 ρ - η_E 曲线

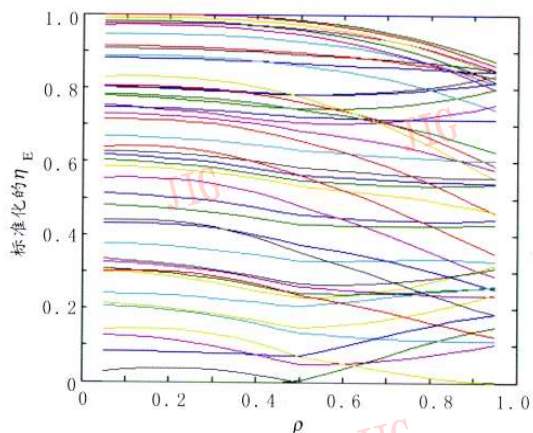


图2 56组候选变换基标准化后的 ρ - η_E 曲线

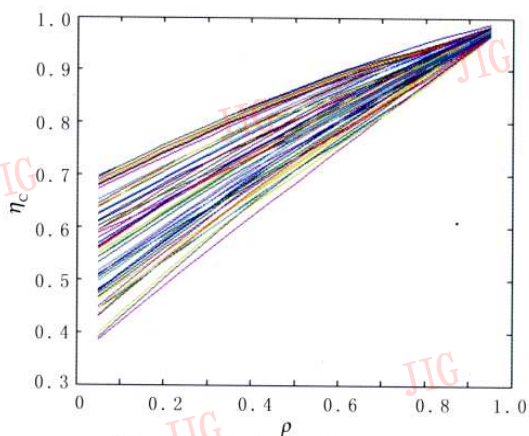


图3 56组候选变换基的 ρ - η_C 曲线

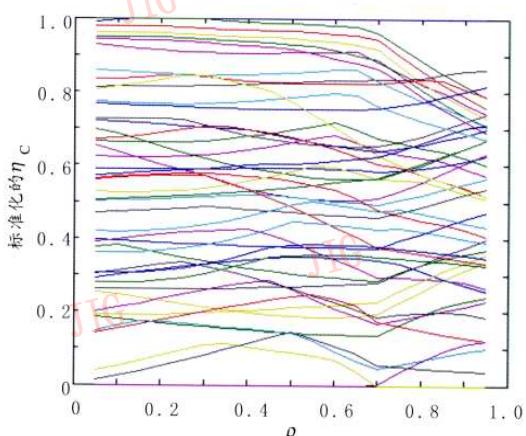


图4 56组候选变换基标准化后的 ρ - η_C 曲线

蔡 珣等：光照色调颜色恒常性算法研究



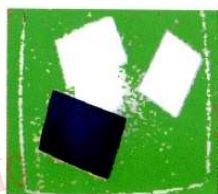
(a) 摄像采集的原图



(b) 直接RGB识别效果图



(c) 光照校正效果图



(d) 直接色调匹配效果图



(e) 光照色调校正效果图



(f) 线性回归颜色校正效果图

图1 各种识别算法效果图