

基于改进量化策略的 XYZ 视频编码

王福龙, 余英林

(华南理工大学 电子与通信系, 广东 广州 510641)

摘要: 对基于三维余弦变换(3D-DCT) 视频编码中 DCT 系数的量化问题展开了研究, 给出了一个三维量化矩阵. 实验表明, 基于此量化矩阵的 XYZ 视频编码, 具有很好的压缩效果及重建图象质量.

关键词: 离散余弦变换; 视频编码; 量化; 熵编码

中图分类号: TN919. 81

文献标识码: A

文章编号: 1007- 7162(2000)03-0060-04

随着通信技术的迅速发展以及人们对信息需求量不断增加, 大量的信息资源需要及时的存储、传送, 而图象信息量尤其是视频信息常因信息量巨大而难以得到有效及时的处理^[1]. 特别是在低码率信道更显突出. 因此人们正致力于研究新的方法来解决, 如结合小波变换、基于二维余弦变换(2-D DCT) 运动补偿及帧间预测、小波与分形相相合、三维余弦变换(3-D DCT) 等. 离散余弦变换是仅次于难以实现的 K-L 变换, 因其有快速算法得以在编码领域广泛应用, 如 JPEG、MPEG、H. 26x 等. 基于 3-D DCT 的视频编码^[2] 以其高压缩比及速度快等特点, 具有在低码率信道下实现视频图象实时编码的潜能, 对应用具有极高价值, 且其算法复杂度低. 本文对 DCT 变换的系数分布特点进行了深入研究, 结合人眼视觉特性, 通过实验得到一量化矩阵, 该矩阵可以有效量化变换系数, 从而得到很好的压缩效果同时又能保证恢复图象的质量.

1 XYZ 编码方案:

基于 3D-DCT 的 XYZ 编码方案如图 1:

将视频序列每 8 帧作为一个图象组(GOP), 从中依次取 8 × 8 × 8 立体块(video cube) 经 3D-DCT 变换后记为: $F(u, v,$

$w), u, v, w = 0, 1, \dots, 7$. 其中 3D-DCT 变换及逆变换为:

$F(u, v, w) =$

$$\frac{2c(u)c(v)c(w)}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \sum_{z=0}^{N-1} f(x, y, z) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right] \cos\left[\frac{\pi(2z+1)w}{2N}\right]$$

$$u, v, w = 0, 1, \dots, N-1.$$

(1)

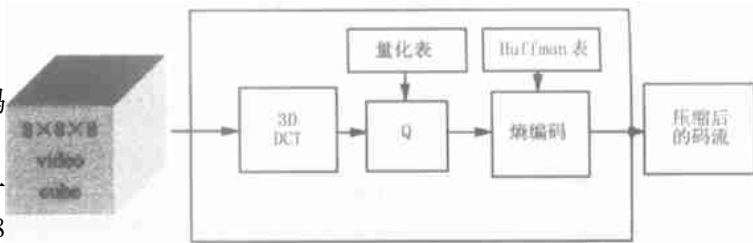


图 1 编码方案

收稿日期: 1999-10-26

基金项目: 国家自然科学基金资助项目(69772026), 广东省自然科学基金资助项目(970484)

作者简介: 王福龙(1968-), 男, 讲师, 现为华南理工大学在职博士; 主要研究方向: 分形理论, 图像处理与模式

$$f(x, y, z) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \sum_{z=0}^{N-1} c(u) c(v) c(w) F(u, v, w) \cos[\frac{\pi(2x+1)u}{2N}] \cos[\frac{\pi(2y+1)v}{2N}] \cos[\frac{\pi(2z+1)w}{2N}]$$

$$x, y, z = 0, 1, \dots, N-1.$$
 (2)

量化表记为: $Q(u, v, w)$, $u, v, w = 0, 1, \dots, 7$. 变换系数按下式量化:

$$F_q(u, v, w) = [\frac{F(u, v, w)}{Q(u, v, w)}]$$
 (3)

由于 DCT 变换系数能量大部都集中在直流及低频分量, 因此经量化后的变换系数 $F_q(u, v, w)$, 非零元素绝大部分都集中在以 $F_q(0, 0, 0)$ 为顶点的锥体顶部, 因此为了充分将非零系数排在一起, 从而使零系数都能连在一起, 得到高效的压缩编码, Zig-zag 按下列顺序进行: 用以 $F_q(0, 0, 0)$ 及 $F_q(7, 7, 7)$ 连线为法线的平面从 $F_q(0, 0, 0)$ 到 $F_q(7, 7, 7)$ 依次切割量化系数正立方体, 在平面上相交的系数依次逐帧排序.

最后, 将经 Zig-zag 排列后的 512 个系数做为一个单位进行熵编码: DC 分量减去前一个单位的 DC 分量, 对差值进行 Huffman 编码; AC 分量先进行 RLE 编码, 再进行 Huffman 编码.

2 量化策略的改进

我们知道在 JPEG 标准中适用于 2D-DCT 系数的量化矩阵, 该量化矩阵中对 DCT 变换系数低频区采用较小的量化系数, 而对高频区采用较大的量化系数. 二维 8×8 图象块经变换后其能量主要集中在低频区, 高频区的能量一般都很小, 这些系数经量化后大部分都为零, 再经基于零的行程编码及变长编码. 故其具有很好的量化效果及重建图象的质量. 这是因其充分考虑了人眼对低频比高频敏感的视觉特性.

对于基于 3D-DCT 的视频编码, 目前没有一个标准的高效的量化表, 所依据的主要是两个来源: 一个是通过一维量化表直接生成, 即:

$$1DQ[x] = (0, 1, 2, 3, 6, 11, 20, 25) \quad x = 0, \dots, 7 \text{ 则}$$

$$3DQ[x, y, z] = 5 + 1DQ[x] + 1DQ[y] + 1DQ[z] \quad x, y, z = 0, \dots, 7$$
 (4)

这种情况没有充分把视觉特性考虑进去, 实验表明效果不是很好.

另一个是直接使用二维 DCT 的量化表对每一变换后的帧块进行量化, 即:

这种情况直接照搬二维量化表到三维上去, 显然不是最优的. 为了探索更加有效的量化表, 我们从 3D-DCT 变换的特点进行了研究及大量的实验比较.

表1 二维 DCT 量化表

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

运动图象序列的每一帧可以看做是静止图象. 对运动图象序列进行 3-D DCT 变换, 可以视为先对每帧进行 2-D DCT 变换, 再对帧间方向进行 1-D DCT 变换. 每帧图象经 2-D DCT 变换后, 其能量大部分集中在直流及低频, 对帧间方向再进行 1D-DCT 变换, 即把后 7 帧的直流及低频的能量大部集中在第一帧的直流及低频区. 这样一来再考虑量化的时候, 既要保证首帧的直流及低频区能量有效并且较精确的编码, 又能将高频中能量较小的变换系数变为零, 而又使恢复图象具有较好的主观效果. 为此我们做了大量实验, 并取得了优于以上两种量化策略的量化表, 这里列出我们所使用的量化表:

表 2 改进的三维 DCT 系数量化表

40	42	46	50	56	66	74	80	42	44	48	52	58	68	76	82
42	44	48	52	58	68	76	82	44	46	50	54	60	70	78	84
46	48	52	56	62	72	80	86	48	50	54	58	64	74	82	88
50	52	56	60	66	76	84	90	52	54	58	62	68	78	86	92
56	58	62	66	72	82	90	96	58	60	64	68	74	84	92	98
66	68	72	76	82	92	100	106	68	70	74	78	84	94	102	108
74	76	80	84	90	100	108	114	76	78	82	86	92	102	110	116
80	82	86	90	96	106	114	120	82	84	88	92	98	108	116	112
46	48	52	56	62	72	80	86	50	52	56	60	66	76	84	90
48	50	54	58	64	74	82	88	52	54	58	62	68	78	86	92
52	54	58	62	68	78	86	92	56	58	62	66	72	82	90	96
56	58	62	66	72	82	90	96	60	62	66	70	76	86	94	100
62	64	68	72	78	88	96	102	66	68	72	76	82	92	100	106
72	74	78	82	88	98	106	112	76	78	82	86	92	102	110	116
80	82	86	90	96	106	114	120	84	86	90	94	100	110	118	124
86	88	92	96	102	112	120	126	90	92	96	100	106	116	124	130
56	58	62	66	72	82	90	96	66	68	72	76	82	92	100	106
58	60	64	68	74	84	92	98	68	70	74	78	84	94	102	108
62	64	68	72	78	88	96	102	72	74	78	82	88	98	106	112
66	68	72	76	82	92	100	106	76	78	82	86	92	102	110	116
72	74	78	82	88	98	106	112	82	84	88	92	98	108	116	122
82	84	88	92	98	108	116	122	92	94	98	102	108	118	126	132
90	92	96	100	106	116	124	130	100	102	106	110	116	126	134	140
96	98	102	106	112	122	130	136	106	108	112	116	122	132	140	146
74	76	80	84	90	100	108	114	80	82	86	90	96	106	114	120
76	78	82	86	92	102	110	116	82	84	88	92	98	108	116	122
80	82	86	90	96	106	114	120	86	88	92	96	102	112	120	126
84	86	90	94	100	110	118	124	90	92	96	100	106	116	124	130
90	92	96	100	106	116	124	130	96	98	102	106	112	122	130	136
100	102	106	110	116	126	134	140	106	108	112	116	122	132	140	146
108	110	114	118	124	134	142	148	114	116	120	124	130	140	148	154
114	116	120	124	130	140	148	154	120	122	126	130	136	146	154	160

3 实验结果

我们对 Miss America 序列及 Salesman 序列进行了实验, 并使用我们给出的量化表及前述两种量化表(分别称为量化策略 1 和 2) 分别作了实验, 实验结果比较如表 3 及表 4, 原图像及重建图像如图 2 所示.

表 3 Miss AMERICA 1- 24 帧实验结果比较

	量化策略 1		量化策略 2		本文量化策略		
平均压缩比	140. 17	160. 42	140. 17	160. 42	140. 17	160. 42	219. 53
平均信噪比	36. 32	35. 73	38. 52	38. 07	39. 31	38. 81	37. 73

表 4 Salesman 1- 24 帧实验结果比较

	量化策略 1		量化策略 2		本文量化策略		
平均压缩比	30.06	60.10	30.06	60.10	30.06	60.10	66.15
平均信噪比	35.13	31.63	35.70	32.36	36.35	32.84	32.42



从以上实验结果可以看出,采用新的量化策略能够得到更好的压缩效果及重建图象质量.这同时也说明了量化策略对编码效率的重要性.由于我们的量化表是在分析 3D- DCT 的特性基础上,经过大量实验得到的,相信还存在更好的、更优的量化策略.

参考文献:

[1] 余英林编著. 图像处理与模式识别[M] . 广州: 华南理工大学出版社, 1991. 108-120.

[2] Raymond Westwater, Borko Fuht. Real- Time Vidio Compression Teshniques and Algorithm[M] U. S. A. : Kluwer Academic Publishers. 1997. 35- 64.

(下转第 67 页)

$$\sum_{i=1}^k s_i = \sum_{i=1}^k r_i + \sum_{i=k+1}^n \min\{rk, r_i\} = \sum_{i=1}^k r_i + \sum_{i=k+1}^n r_i = \sum_{i=1}^n r_i = \sum_{i=1}^n s_i,$$

这与 $s_n > 0$ 矛盾($\mathbf{S}$ 是正向量), 因此等号不成立, 所以

$$\sum_{i=1}^k s_i < \sum_{i=1}^k r_i + \sum_{i=k+1}^n \min\{rk, r_i\}.$$

参考文献:

- [1] Brualdi R A, Michael T S. The class of matrices of zeros, ones and twos with prescribed row and column sums[J]. Linear Algebra Appl, 1989, 114/115: 181-198.
- [2] Brualdi R A. Matrices of zeros and ones with fixed row and column sum vectors[J]. Linear Algebra Appl, 1980, 33: 159-231.
- [3] Mirsky L. Combinatorial theorems and integral matrices[J]. J. Combin. Theory, 1968, 5: 30-44.
- [4] Anstee R P. Network flows approach for matrices with given row and column sums[J]. Discrete Math, 1983, 44: 125-138.

The Characteristics of the Nonempty Class $\mathcal{M}(\mathbf{R}, \mathbf{S})$

Hou Su-mei

(Dept. of Mathematics and Physics, GDUT, Guangzhou, 510090)

Abstract: Let $\mathcal{M}(\mathbf{R}, \mathbf{S})$ be the class of all $(0, 1, \dots, r)$ -matrices with a prescribed row sum vector $\mathbf{R}$ and column sum vector $\mathbf{S}$. We study the relationship between the vector $\mathbf{R}$ and the vector $\mathbf{S}$ when $\mathcal{M}(\mathbf{R}, \mathbf{S})$ is nonempty.

Key words: $(0, 1, \dots, r)$ -matrix; $\mathcal{M}(\mathbf{R}, \mathbf{S})$; irreducible matrix; structure matrix

(上接第 63 页)

XYZ Video Coding Based on Improved Quantization Policy

WANG Fu-long Yu Ying-lin

(South China Univ of Tech. Guangzhou, 510641 China)

Abstract: The quantizing policy of XYZ video coding has been researched, and a new quantization matrix on XYZ video coding is given. The simulation results show that this method can give high compression ratio and good reconstruct picture quality.

Key words: discrete cosine transform; video coding; quantization; entropy coding