维普资讯 http://www.cqvip.com

一种低码率快速视频编码方案

A Fast Video Coding Algorithm

王福龙　冯彦发　余英林
（华南理工大学电子与通信系　广州）

Abstract　A video code program based on part 3D-DCT method and modifred quantizing table is presented in this paper, which can have high compression ratio and good quanlity reconsftructed video picture result.

Keywords　Video coding, Discrete cosine transformation, Quantization

一、前言

图像信息尤其是视频信息其信息量巨大而难以得到有效及时的处理，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2-D DCT（二维离散余弦变换）运动补偿及帧间预测、小波与分形相结合、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，如结合2-D DCT 变换的运动估值/运动补偿、基于2D-DCT 变换域的帧间预测、基于3D-DCT 变换的视频编码[1]等。本文对 DCT（离散余弦变换）的特性进行深入研究，针对可视电话、会议电视的特点，提出基于部分3-D DCT 的视频编码方案，该方案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像实时编码的潜能，对应用具有极高价值，且其算法复杂度低。

二、视频编码方案

基于部分3-D DCT 编码方案如图1所示。

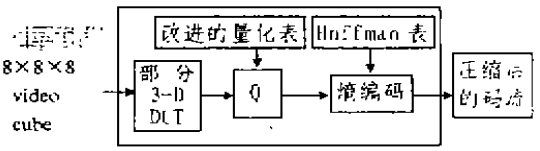

图1　基于部分3-D DCT 的视频编码方案

1. 部分3-D DCT 算法

自然图像在统计特性上，可近似为 Markov 过程，其协方差矩阵近似为 Toeplitz 矩阵的形式。而 DCT 的变换矩阵与 Toeplitz 矩阵的特征向量矩阵非常相似[2]，经过 DCT 后的自然图像的协方差矩阵，除了对角线以外，其它很多元素都近似为0，因此有很好的压缩效果。所以 DCT 变换被认为是准最佳变换，因其有快速算法，DCT 成为 JPEG 标准的基本算法。

运动图像序列的每一帧可以看做是静止图像，每帧图像经2-D DCT 变换后，其能量大部分集中在直流及低频，即集中在变换后矩阵次对角线的左上部，右下部经量化后绝大多数都是0。实验表明，对于帧间相关性很强的图像序列，如可视电话、视频会议等，每帧经2D-DCT 变换后，相邻帧对应位置的变换系数相关性很强。若帧间方向对这些非零系数进一步施加1D-DCT 变换，可以有效消除帧间 DCT 系数的相关性。因此我们提出帧内进行2D-DCT 变换后、帧间方向对次对角线左上部做1-D DCT 变换的部分3D-DCT 算法。

将视频序列每8帧做为一个处理对象，依次从中取出8×8×8视频单元。根据以上分析对每帧进行2-D DCT 变换，帧间进行部分1-D DCT 变换，即只对次对角线左上部（包括次对角线系数）进行变换，右下部不变换。8×8×8视频单元（video cube）经 DCT 变换后记为：F(u,v,w)，u,v,w=0,1,…,7。

2 量化表

对基于3D-DCT 的视频编码，目前没有一个标准的高效量化表，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)　x=0,…,7　则

3DQ[x,y,z]=5+1DQ[x]+1DQ[y]+1DQ[z]

x,y,z=0,…,7

王福龙、冯彦发　在职博士生（现在广东工业大学五山校区工作）。

维普资讯 http://www.cqvip.com

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响块进行量化，即：

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

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，

运动图像序列的每一帧可以看做是静止图像。对运动图像序列进行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，每帧图像经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地编码，又能将高频中能量较小的变换系数变为零，而又使恢复图像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1

40	42	46	50	56	65	74	80
42	44	48	52	58	68	76	55
46	48	52	56	62	72	69	56
50	52	56	60	66	87	80	62
56	58	62	66	68	109	103	77
66	68	72	64	81	104	113	92
74	76	78	87	103	121	120	101
80	92	95	98	112	100	103	99

表2

42	44	48	52	58	68	76	82
44	46	50	54	60	70	78	55
48	50	54	58	64	74	69	56
52	54	58	62	68	87	80	62
58	60	64	68	68	109	103	77
68	70	64	64	81	104	118	92
76	78	78	87	103	121	120	101
82	92	95	98	112	109	103	99

表3

46	48	52	56	62	72	80	86
48	50	54	58	64	74	82	55
52	54	58	62	68	78	69	56
56	58	62	66	72	87	80	62
62	64	68	72	68	109	103	77
72	74	78	64	81	104	113	92
80	82	78	87	103	121	120	101
86	92	95	98	112	100	103	99

表4

50	52	56	60	65	76	84	90
52	54	58	62	68	78	86	55
56	58	62	66	72	82	69	56
60	62	66	70	76	87	80	62
66	68	72	68	82	109	103	77
76	78	82	64	92	104	113	92
84	86	78	87	100	121	120	101
90	92	95	98	106	100	103	99

表5

56	58	62	66	72	82	90	96
58	60	64	68	74	84	92	55
62	64	68	72	78	88	69	56
66	68	72	76	82	87	80	62
72	74	78	82	88	109	103	77
82	84	88	92	98	104	113	92
90	92	78	100	106	121	120	101
96	98	95	106	112	100	103	99

表6

66	68	72	76	82	92	100	106
68	70	74	78	84	94	102	55
72	74	78	82	88	98	69	56
76	78	82	86	92	87	80	62
82	84	88	92	98	109	103	77
92	94	98	64	108	104	113	92
100	102	78	87	116	121	120	101
106	92	95	98	122	100	103	99

表7

74	76	80	84	90	100	108	114
76	78	82	86	92	102	110	55
80	82	86	90	96	106	69	56
84	86	90	94	100	87	80	62
90	92	96	100	106	109	103	77
100	102	106	64	116	104	113	92
108	110	78	87	124	121	120	101
114	92	95	98	130	100	103	99

表8

80	82	86	90	96	106	114	120
82	84	88	92	98	108	116	55
86	88	92	96	102	112	69	56
90	92	96	100	106	87	80	62
96	98	102	106	112	109	103	77
106	108	112	64	122	104	113	92
114	116	78	87	130	121	120	101
120	92	95	98	136	100	103	99

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

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

3. Zig-zag

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，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方体，

（下转第92页）

维普资讯 http://www.cqvip.com

的依赖性大；而基于贝叶斯统计学学习有机地综合了这两类信息，因此对实例数据集合的依赖性较低，其学习结果更精确。但贝叶斯统计学中的一些理论仍然存在许多争议，如：将参数 θ 看成随机变量是否妥当？参数的先验分布是否存在？在计算中如何选取？等。

在实际应用中，用于学习的实例数据往往是不完整的，这种情况下一般采用近似算法进行学习。常用的近似学习算法有：Gibbs 抽样算法、EM 算法和梯度上升算法。从算法的时间复杂度来看，Gibbs 抽样算法的效率是最低的，当选用相同的推理算法时，EM 算法和梯度上升算法的计算复杂性相同。另外，Gibbs 抽样算法只能用来对信度网的条件概率表进行批量更新，而 EM 算法和梯度上升算法既可以用于对条件概率表的批量更新，又可以用于条件概率表的顺序更新。

参 考 文 献

1 Spirtes P, et al Causation, Prediction, and Search. Spring-Verlag, New York 1993

2 Fisher R A. Theory of Statistical Estimation. In, Proc. of the Cambridge Philosophical Society, 1925, 22: 700～715

3 Dawid A P, Lauritzen S L Hyper Markov laws in the statistical analysis of decomposable graphical models Annals of Statistics, 1993, 21(3)

4 Dempster A, et al. Maximum likelihood form incomplete data via the EM algorithm. Journal of the Royal statistical Society, 1977, B39 1～38

5 Russel S, et al. Local learning in probabilistic network with hidden variables. In Proc. of the 14-th IJCAI, Montreal, Canada, Morgan Kaufmann, 1995. 1146～1152

6 Lauritzen S The EM algorithm for graphical association models with missing data Computational Statistical and Data Analyses, 1995, 19 191～201

7 Bayes T R An essay towards solving a problem in the doctrine of chance Phil. Trans sic , London, PP53, 370, 1763

8 Heckerman D. A Tutorial on Learning With Bayesian Networks [*Technial Report MSR-TR-95-06*]. 1995

9 Hryce) T. Gibbs Sampling in Bayesian Networks Artificial Intelligence, 1990, 46: 351～363

10 York J. Use of the Gibbs sampling in expert systems. Artificial Intelligence, 1992, 56, 115～130

11 Heckerman D, Geiger D. Learning Bayesian Networks: the combination of knowledge and statistical data; [Microsoft Technical Report MSR-TR-94-09]. 1994

12 Cooper G, Herskovits E A Bayesian method for induction of probabilistic networds from data Machine Learning, 1992(9): 309～347

13 Spiegelhalter D, Lauritaen S. Sequential updating of conditional probabilities on directed graphical structures. Networks, 1990, 20, 579～605

14 Thiesson B. Accelerated quantification of bayesian networks with incomplete data. In Proc of First Intl conf. on Knowledge Discovery and Data Mining, Montreal, QU, Morgan Kaufmann, 1995. 306～311

（上接第105页）

在平面上相交的系数依次逐帧排序。

4. 编码

将经 zig-zag 后的512个系数做为一个单位进行熵编码：DC（直流）分量减去前一个单位的 DC 分量，对差值进行 Huffman 编码；AC（交流）分量先进行 RLE 编码，再进行 Huffman 编码。

三、实验结果

我们对 Miss America, Chaire 人像等序列（灰度）分别进行了实验，每帧块进行2-D DCT 变换，帧方向只对每帧次对角线及其左上部进行1-D DCT 变换，右下部不变换并且采用前述量化矩阵进行量比，最后编码。

从实验结果可以看到，对于 Miss America, Claire 等图像序列，由于每帧帧内相关性强，且各帧变化区域小，每个帧块经过2-D DCT 变换后，次对角线右下部的变换系数经量化后，大部分都为零，并且在允许一定失真但保证图像质量的情况下，这些量化后的系数基本都是零。因此采用部分3-D DCT 算法及量化策略可以得到很高的压缩比及很好的解码恢复图像，同时由于采用部分3D-DCT 算法减少了近七分之一的 DCT 运算量，从而提高了编解码效率。

参 考 文 献

1 余英林编著. 图像处理与模式识别. 华南理工大学出版社

2 Westwater R, Furht B Real-Time Vidio Compression Techniques and Algorithm. U. S. A. Kluwer Academus Publishers. July 1997

3 N. 阿罕麦德, K R. 罗. 数字信号处理中的正交变换. 北京 人民邮电出版社, 1979