

变时长变速 XYZ 视频编码

王福龙 余英林

(广东工业大学 广州510643) (华南理工大学 广州510641)

摘要 将数字视频序列可以看作是从模拟视频信号中抽样得到,帧间相关性过强是由于采样过密造成的;而帧间相关性太弱是由于采样频率太小造成的。本文从采样定理的角度提出变时长变速 XYZ 编码方案,得到很好的编码效果。

关键词 离散余弦变换,视频编码,变速编码

Variant Rate XYZ Video Coding Method Based on Various Time Length

WANG Fu-Long YU Ying-Lin

(Guang Dong University of Technology, Guangzhou 510643) (South China University of Technology, Guangzhou 510641)

Abstract Digital video sequence can be regarded as sampling of analogue video signal, sampling too intensive can make the intraframe relationship higher, and sampling too sparse can make the intraframe relationship lower. Based on sampling theory, a variant rate XYZ video coding method is proposed in this thesis, which can improve the coding efficiency.

Keywords Discrete cosine transform, Video coding, Variant rate coding

1 引言

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地存储、传送,而图像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。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JPEG、MPEG、H. 26x等。基于3-D DCT^[1]的视频编码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。

本文从采样定理^[2~7]的角度提出变速XYZ视频编码。在数字视频编码中,数字视频序列可以看作是从模拟视频信号中抽样得到的,抽样频率一般为 $f_s=24$ 或 30 。而模拟视频信号的频率 f_c 却是变化的,对于运动较小的视频信号 f_c 相对较小,以 f_s 为采样频率能够将原始信号真实地反映出来,采样得到的视频序列相邻帧之间的相关性亦是较强的;而对运动较剧烈的视频信号,信号的变化频率 f_c 较大,这时若以不变的采样频率采样,必不能将实际信息完全真实地反映出来,因为此时的采样频率不一定能满足使频谱不发生混叠的最小抽样频率 $2f_c$ 。采样得到的视频序列相邻帧之间的相关性亦是较弱的。即使是同一信号序列,在不同时间段内信号频率也是不同的。依据采样定理,对信号频率大的视频信号相应采样频率亦应大些,对信号频率小的视频信号相应采样频率亦应小些。这样也保证了采样得到的视频序列具有较强的相关性。

对视频序列进行编码时,为了将信息压缩,就要有效消除空域和时域间的相关性。若采用XYZ视频编码方法,则要求帧间相关性强些。按照上段分析,对于运动缓慢的序列,信号

变化的频率小,此时用上述的不变的采样频率采样得到的视频序列帧间相关性显得过强,这是由于采样过密造成的,因此对于信号变化缓慢的视频信号采样频率应小些,这样同样能保证帧间的相关性及恢复信号的完整性,具体在编码操作时可以采用跳帧的方法。对于运动剧烈的序列,信号变化的频率大,此时用上述的不变的采样频率采样得到的视频序列帧间相关性显得过弱,这是由于采样周期过大造成的,因此对于信号变化剧烈的视频信号采样频率应大些,这样就能保证帧间的相关性及恢复信号的完整性,具体在编码操作时可以采用帧间插值的方法。

对帧间相关性太强的序列可以采用帧间跳帧的方法提高编码效率;而对帧间相关性太弱的视频序列可以采用帧间插帧的方法提高帧间相关性从而提高帧间编码效率。本文提出变时长变速XYZ编码方案,得到很好的编码效果。

2 变时长变速 XYZ 视频编码

首先,图像序列被分成一系列的时间窗(Time Window)W,每个时间窗所含的原始图像帧数固定。如果时间窗含有的帧数越多,那么相应变时长的3D DCT编码质量会越好,但存储空间也相应会增加。接下来,每个时间窗再分成空间大小为 $M \times N$ 大小的系列块。为了使每个3D编码块只含少量运动,引入一个景物变化检测器来检测每个系列块中景物变化,以便在变换之前将每个块 $M \times N \times W$ 检测分类出各个运动相对平稳区。设通过场景变化检测将长度为W的序列分成了P+1个不连续的部分 $t_{-1}(=0) < t_0 < t_1 < \dots < t_{P-1}(=W)$,则对应系列块分成了P+1个子序列块,记为 $F_j, j=0, 1, \dots, P, F_j$ 具有时间轴长度为 $L_j=t_j-t_{j-1}$,其具体形式为:

$$F_j = [I_{t_{j-1}}(x, y), I_{t_{j-1}+1}(x, y), \dots, I_{t_j}(x, y)] \quad (1)$$

P+1个子序列块的误差函数表达式定义如下:

$$E(P) = \sum_{j=0}^P \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} e[F_j(x,y)] \quad (2)$$

其中 $e[F_j(x,y)] = \sum_{i=0}^{L_j-1} [I_{i,j-1}(x,y) - \overline{I_j(x,y)}]^2$, 而 $\overline{I_j(x,y)}$ 表示 $I_{j-1}(x,y), I_{j-1+1}(x,y), \dots, I_{j-1}(x,y)$ 的均值。

对于每个视频序列块, 其分成的子序列块的个数受其所含运动的多少来决定。对于预先设定的阈值 E_0 , 不同的序列块要选择合适 P 使对应分解成的子序列构成的误差函数 $E(P)$ 相对变量 P 取得的最小值小于 E_0 。显然, 当块中含有的运动信息变化小时, P 的值相对较小, 当块中含有的运动信息多时, 相应 P 的值就会增加。从编码的角度来看, P 的值越小越能有效将对应块编码, 使其码率降低。使用自适应块滤波器 (Adaptive Block Filter, ABF)^[8], 可以有效求得最小误差函数表达式, 即将序列块按运动景物变化分解, 取得最优的子序列块。这种通过误差函数寻优分解序列块的方法虽然可以取得最优的子序列块, 但需要很大的计算量。

为了解决计算量过大的问题, 可以使用平均绝对误差

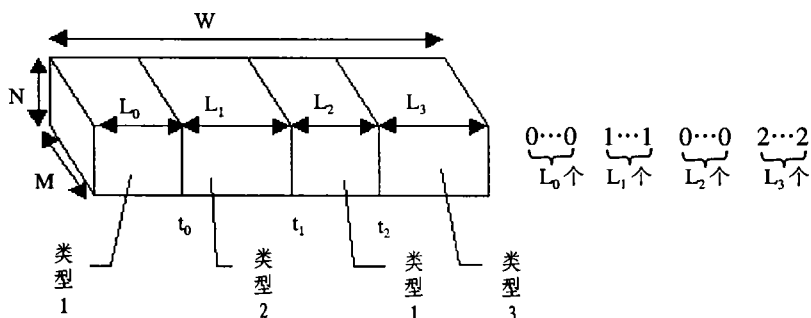


图1 时间窗长为 W 的序列块经运动检测分类图

(Mean Absolute Difference, MAD) 准则来检测帧间景物变化

$$MAD_i = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |I_{i+1}(x,y) - I_i(x,y)| \quad (3)$$

当 MAD_i 大于预先给定的阈值 T_0 时, 就认为帧间景物发生了变化, 此时需要在此另外分块。这种方法非常简单, 而且计算量非常小, 但其分块的精确程度不及 ABF 方法。

经过上述的方法将时长为 W 的序列块分成了 $P+1$ 个子序列块, 这些子序列块的时长是不等的, 其大小为 $L_j \times M \times N$, 对这些子序列变时长三维视频块, 进行进一步分类, 若子块内的误差值小于给定阈值 T_d , 则认为块内景物是不动区, 此时只编码第一帧二维块, 这样的子块中的每一帧标记为 0; 若子块内的误差值大于 T_d , 而小于 T_s , 则认为可跳帧区, 对相应子块隔帧进行变长 3DDCT 变换编码, 这样的子块中的每一帧标记为 1; 对于子块内的误差值大于 T_s , 进行变长 3DDCT 变换编码, 这样的子块中的每一帧标记为 2, 如图 1 所示。

应用与子块大小相匹配的 3D DCT 变换, 变为变换域, 设块大小为 $L \times M \times N$, 对应 3D DCT 定义^[9]如下:

$$X(u,v,w) = \frac{8}{L \times M \times N} \sum_{t=0}^{L-1} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} C(u)C(v)C(w)I_t(x,y) \cdot \cos \frac{\pi(2t+1)u}{2L} \cos \frac{\pi(2x+1)v}{2M} \cos \frac{\pi(2y+1)w}{2N} \quad (4)$$

其中, 当 $n=0$ 时, $C(n) = 1/\sqrt{2}$, $n \neq 0$ 时, $C(n) = 1$ 。

将每一序列块按上述方法分类编码, 并将序列块中每帧块的类型按顺序编号后采用算术编码, 编码流程图如图 2 所示。

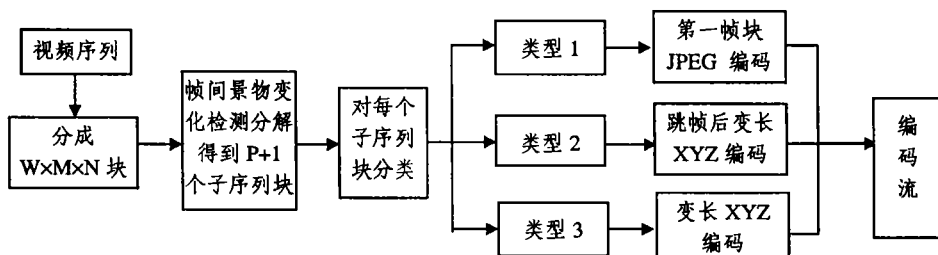


图2 变时长变长 XYZ 编码流程图

3 实验结果比较

我们对 Salesman (352×288) 和 Football (352×240) 序列 (30 帧/秒) 进行了实验, 并与 XYZ 编码方法、MPEG 方法及文 [9] (VLXYZ)、文 [10] (VLXYZ) 的结果进行了比较。实验中 XYZ 方法采用传统的 8×8×8 大视频块, 而 MPEG 方法采用 IPPPPPP 每 8 帧一个组, 运动估值采用块大小 16×16, 最大搜索位移为 8 的全搜索算法。对于本文提出的方法 (VSXYZ2), 采用时间窗大小 32, 即每 32 帧为一组进行运动景物变化检测, 将其分解成若干个子序列块。

通过检测图像景物变化, 将序列块有效分类为不动块, 相关块及变化块, 对应不同类型序列块采用不同的编码方法, 方法 VSXYZ1 对 Salesman 和 Football 序列都有很好的编码效果。但采用了定时长编码, 只限于 8×8×8 视频块, 当块内有景物变化较大时, 通过插值的方法不能有效提高相关性; 同时插值本身也增加了编码的信息量, 使得由于增加了相关性而提高的编码效率被部分抵消, 所以编码效果不及方法 VSXYZ2。

采用 VSXYZ2 方法通过检测时间窗内的 32 帧块景物变化, 将其分成若干个子块, 对每一子块进行分类为不动块、跳帧块或相关块, 按分类编码。可以将相关性较强的序列块分类

到一起,提高块内的相关性,经变长3DDCT变换后高频区很少会有较大系数,故可有效提高压缩;同时对不动块采用了单帧块编码,对跳帧块采用隔帧编码,这样既不必浪费计算时间,又可减少编码比特,能有效将编码比特用于运动变化区域,大大提高了编码效率。从实验结果可以看出,采用 ABF 方法能有效检测运动变化将时间窗分成相关性强的子块,效果好于 MAD 检测方法,但 ABF 方法需要很大的运算量。

实验结果如表1,2和图3,4所示。

表1 Football 序列实验结果比较

	bits per pixel	0.25	0.35	0.45	0.55
PSNR	XYZ	24.9	26	27.23	28.68
	MPEG	24.6	25.6	26.9	28.3
	VLXYZ	25.4	26.8	28	28.9
	VSXYZ1	25.1	26.51	27.73	28.7
	VSXYZ2	25.82	27.25	28.35	29.29

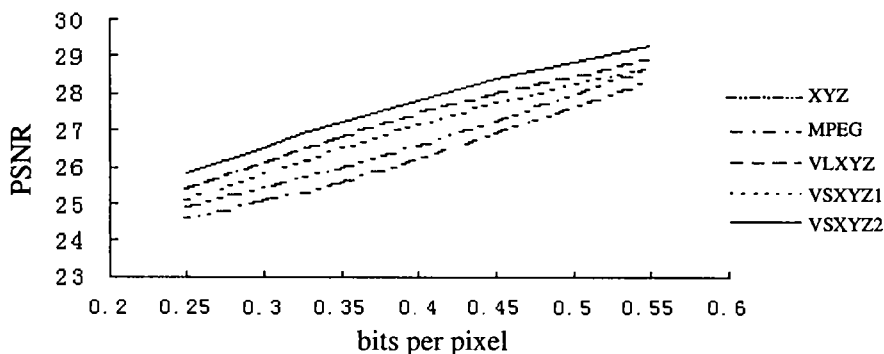


图3 Football 序列实验结果比较

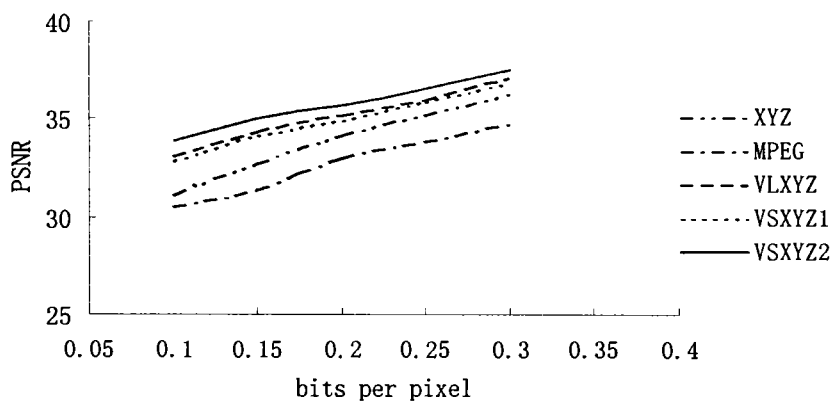


图4 Salesman 序列实验结果比较

表2 Salesman 序列实验结果比较

	bits per pixel	0.1	0.15	0.2	0.25	0.3
PSNR	XYZ	31.1	32.75	34.15	35.2	36.32
	MPEG	30.5	31.4	33	33.9	34.7
	VLXYZ	33.15	34.4	35.2	36	37.1
	VSXYZ1	32.8	34.2	34.9	35.85	36.8
	VSXYZ2	33.86	35.05	35.7	36.6	37.5

参考文献

- Westwater R, Furht B. Real-Time Video Compression Techniques and Algorithm. U. S. A. Kluwer Academic Publishers, July 1997
- 余英林编著. 图像处理与模式识别. 华南理工大学出版社
- 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

- 韦岗, 邱伟著. 现代信号处理理论与技术. 华南理工大学出版社, 1994
- 黎洪松著. 数字视频技术及其应用, 清华大学出版社, 1998
- 胡广书著. 数字信号处理—理论、算法与实现. 清华大学出版社, 1998
- 李振辉, 李仁和著. 探索图像文件的奥妙. 清华大学出版社, 1996
- Olstad B. Noise reduction in ultrasound images using multiple linear regression in a temporal context. Proc. SPIE, 1992, 1451: 269~281
- Chan Y L, Siu W C. Variable temporal-length 3-D discrete cosine transform coding. IEEE Trans. on Image Processing, 1997, 6(5): 758~763
- 王福龙. 定时长变速 XYZ 视频编码. 通信技术, 2002(1)