

一种新的自适应图像组结构算法^{*}

王员根^{1,2}, 梁 凡¹, 张秀丽³, 肖明明^{1,2}
(1 中山大学电子与通信工程系, 广东 广州 510275;
2 仲恺农业技术学院信息工程系, 广东 广州 510225;
3 荆门职业技术学院, 湖北 荆门 448000)

摘 要: 运动补偿时域滤波 (MCTF) 将若干连续的视频帧组成一个图像组进行时域小波分解以获得时域可分级性。与图像组大小固定的结构相比, 自适应图像组结构 (AGS) 提高了编码效率, 但实现复杂度也大大增加。提出了一种新的自适应图像组结构 (NAGS), 利用帧间预测判断视频序列的时域变化特性, 并根据视频序列的时域变化特性选择不同的图像组结构。新的自适应图像组结构大大降低了自适应图像组结构的计算复杂度, 并解决了自适应图像组结构时延的振荡性问题。实验结果表明, 新的自适应图像组结构的编码性能与自适应图像组结构基本一致。

关键词: SVG 运动补偿时域滤波; 提升; 图像组; 自适应图像组结构
中图分类号: TN919.81 **文献标识码:** A **文章编号:** 0529-6579 (2007) 02-0020-04

如何通过有效的时域变换技术来消除视频序列的帧间相关性是可分级视频编码 (scalable video coding SVC) 的关键, 也是当前 SVC 研究的重点。运动补偿时域滤波 (motion-compensated temporal filtering MCTF) 以其优越的性能越来越受到人们的重视。

MCTF 的思想首先由 Ohm^[1] 提出, 并由 Choi 和 Woods^[2] 进行了改进。与传统的闭环 (close loop) 预测编码把解码帧作为后续帧的运动补偿参考帧不同, MCTF 没有采用这种时域递归结构。MCTF 对若干连续的视频帧组成的图像组 (group of pictures GOP) 进行时域分解形成时域小波树, 这种在运动方向上进行的时域滤波是一个开环 (open loop) 过程, 滤波后的时域解相关信号再进行 2D 空域小波变换和嵌入式编码。在当前的可分级编码研究中, MCTF 是基于小波提升方法来实现的^[3]。

联合可分级视频编码模型 (joint scalable video model JSVM) 采用 MCTF 方法, 把 8 帧或 16 帧图像作为一个 GOP 进行时域滤波^[4], 当视频序列变化时, 这种固定图像组结构的 MCTF 分解并不能取得很好的编码效果。根据视频序列在时间轴上的变化特性, Park 等人提出了自适应图像组结构 (adaptive GOP structure AGS) 来进一步提高编码效

率^[5]。AGS 通过预编码来获得精确的图像分组结构, 大大增加了算法实现的复杂度和编码时延, 其复杂度相当于固定图像组结构的 4 倍。本文提出了一种新的自适应图像组结构 (novel AGS NAGS), NAGS 的实现复杂度与固定图像组结构相同, 并解决了 AGS 结构时延的振荡性问题。实验表明, NAGS 的编码效率与 AGS 基本一致。

1 自适应图像组结构

图 1 和图 2 分别是 GOP 为 8 帧率为每秒 15 帧的 Foreman 和 Football 序列 MCTF 分解结构。

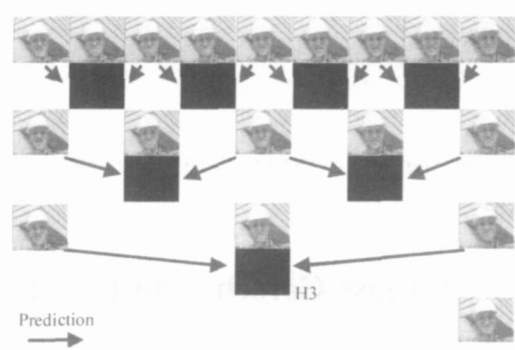


图 1 Foreman 序列的 MCTF 分解效果
Fig. 1 MCTF decomposition of Foreman

^{*} 收稿日期: 2006-06-14
基金项目: 广东省科技计划资助项目 (2006B21001010); 云浮—仲恺科研资助项目 (G2004012)
作者简介: 王员根 (1976 年生), 男, 硕士, 助教, 通讯联系人; 梁凡; E-mail: isslf@mail.sysu.edu.cn
(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

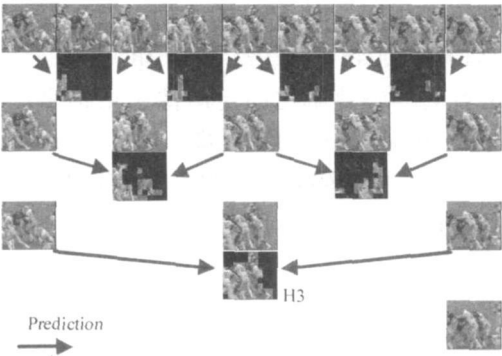


图 2 Football 序列的 MCTF 分解效果
Fig. 2 MCTF decomposition of Football

对 Foreman 序列来说，其图像特征变化相对缓慢，从图 1 中的黑块（帧间块）可以看出 MCTF 取得了良好的预测效果。然而，Football 序列的图像特征变化较大，用于预测的参考帧又相距较远，因此 MCTF 分解效果不理想。从图 2 可以看出，在 高通帧 H3 中存在很多的白块（帧内块）。

根据视频序列的时域变化特性，Pak 等人提出了改变固定图像组结构到适当的自适应亚图像组（sub-GOP）结构能提高编码效率^[5]。比如，将一个 GOP 为 16 的图像组分解为 [8 4 2 2] 的亚图像组结构，4 个亚图像组分别包括 8 帧、4 帧和 2 帧图像。实验结果表明，与固定图像组结构的 JSVM 1.0 相比^[4]，AGS 的编码效率最大能提高 0.62 dB^[6]。

AGS 虽然提高了 MCTF 的编码效率，但是它的实现复杂度过高，为了完成自适应选择图像分组，需要进行预编码：对每一个 GOP（如 GOP = 16）进行亚图像组分别为 16、8、4 和 2 的 MCTF 分解，生成的低通帧和高通帧再进行正常编码，为了提高压缩效率，在每一个亚图像组上将编码重建后的 MCTF 帧与初始的 MCTF 分解帧计算 MSE，根据最小的 MSE 而决定选择相应的图像组结构。AGS 问题如下：第一，在 MCTF 中，运动估值将是最耗时的一部分，AGS 进行预编码时，它的 MCTF 分解的复杂度和计算量是固定图像组结构的 4 倍，这里还不考虑需要的缓存容量；第二，由于亚图像组分组结构的变化，在图像组的边界需要不断地切换小波分解模式，也极大地增加了实现的复杂度；第三，亚图像组大小从 2 到 16 之间的选择带来了结构时延的振荡。

2 新的自适应图像组结构

我们通过研究发现：对于快速变化的序列若采用大尺寸的 GOP 将会产生高能量的高通帧，这对

视频压缩编码是极其不利的；而变化相对平缓的序列若采用小尺寸的 GOP 则无法很好地利用视频序列的长时间相关性，MCTF 分解得到的低通帧包含较高的能量，这也不理想。根据多数视频序列运动的缓变特性，即在短时间内（如一个 16 帧的 GOP 的时间）视频序列的运动剧烈程度相对不变，本文提出一种新的 AGS 算法 - NAGS，其思想如下：

- (1) 算法只采用两种亚图像组模式：[4 4 4 4] 和 [16]；
- (2) [4 4 4 4] 模式用于运动较剧烈的序列，而 [16] 模式对应运动相对平缓的序列；
- (3) 根据 MCTF 提升得到的 H3 帧（参见图 3）包含能量的多少来判决模式选择；

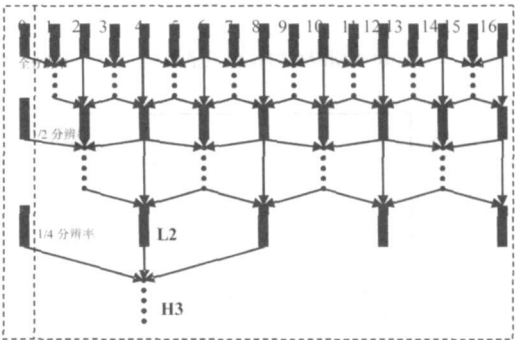


图 3 时域 2 步提升的 [16] 分组结构
Fig. 3 2 steps MCTF decomposition of [16]

- (4) H3 帧的能量满足 (1) 式将选择 [4 4 4 4] 模式，否则按 [16] 模式编码

$$\sum_{y=0}^{h-1} \sum_{x=1}^{w-1} H_3^2(x, y) \geq T \sum_{y=0}^{h-1} \sum_{x=1}^{w-1} L_2^2(x, y) \quad (1)$$

公式 (1) 中 w、h 分别是图像的宽和高，H 是高通子带，L 是低通子带，T 是判决阈值。

NAGS 大大降低了 AGS 的复杂度：开始执行 GOP = 16 的逐步提升，当提升到 H3 帧时与 L2 帧做比较，决定是否继续提升，还是选择 [4 4 4 4] 分组模式，它只比固定图像分组结构多一个低通帧 L2 的缓存和两个平方和累加计算，继承了固定图像组结构低复杂度的计算，又兼顾了 AGS 提高编码效率的思想。

3 实验结果及分析

我们对固定图像组结构、自适应图像组结构以及本文提出的新的自适应图像组结构的性能进行了实验研究。实验表明，公式 (1) 中阈值 T 取 0.25 时能取得较好的判决效果。图 4 给出了三种算法的编码性能比较。

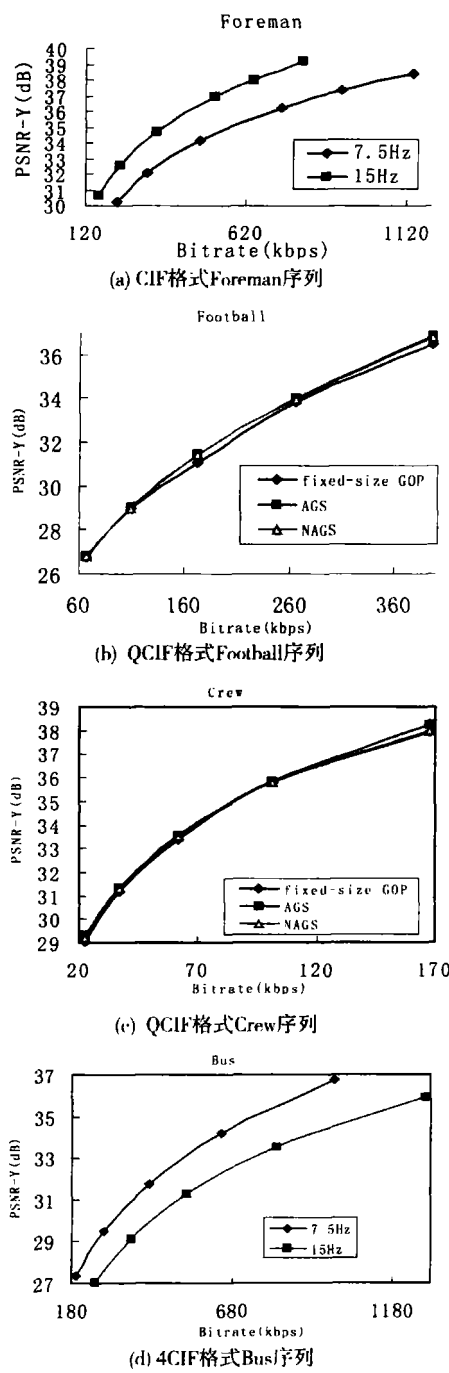


图 4 三种算法编码性能的比较

Fig. 4 Comparison of performance of three algorithms

对 Foreman 和 Bus 序列, 固定图像组结构、AGS和 NAGS选择了相同的 GOP结构, 因此其编

码性能完全相同。但 AGS的时间复杂度是固定图像组结构和 NAGS的 4 倍。Football序列的结果表明 NAGS的编码效率比固定图像组结构平均提高 0.29 dB 相对 AGS下降 0.02 dB。Crew 序列的结果表明 NAGS的编码效率比固定图像组结构平均提高 0.09 dB 相对 AGS下降了 0.06 dB。

5 总 结

本文提出了一种新的 AGS 算法 (NAGS)。NAGS算法实现复杂度与固定图像组结构相当。实验结果表明 NAGS 的编码性能与 AGS 基本一致。NAGS继承了固定图像组结构低复杂度的计算, 又兼顾了 AGS算法提高编码效率的思想; 在每一个 GOP内也有固定的提升小波, 进一步减少了 GOP内小波切换带来的计算复杂度, 使运算有规律便于硬件实现, 并且在结构时延的振荡性上也相对 AGS更稳定, 这对可分级视频编码器的应用将具有极其重要的现实意义。

参考文献:

- [1] OHM J - R. Three dimensional subband coding with motion compensation [J]. IEEE Trans Image Processing 1994 3(5): 559 - 571.
- [2] CHO I S - J WOODS J W. Motion compensated 3D subband coding of video [J]. IEEE Trans Image Processing 1999 8(2): 155 - 167.
- [3] SW ELDENS W. The lifting scheme: A custom - design construction of bi - orthogonal wavelets [J]. J Appl Comput Harmon Anal 1996 3(2): 186 - 200.
- [4] REICHEL J SCHWARZ H WIEN M. Joint scalable video model JSVM 1 [C]. JVT - N023 2005.
- [5] PARK G H, PARK M W, JEONG S et al Adaptive GOP structure for SVC [C]. ISO / IEC JTC1 SC29 WG11 M11563 2005.
- [6] PARK G H, PARK M W, JEONG S et al Improve SVC coding efficiency by adaptive GOP structure (SVC CE2) [C]. JVT - 0018 2005.

(下转第 26 页)

tems 2002 10(1): 95 – 103.

[6] YEE Leung WU Weizhi ZHANG Wenxiu Knowledge acquisition in incomplete information systems A rough set approach[J]. European Journal of Operational Research 2006 168 164 – 183.

[7] 魏大宽, 黄兵. 不完备模糊目标信息系统粗集与知识约简[J]. 计算机工程, 2006(8): 48 – 51.

[8] KRYSZKIEWICZ M. Rough set approach to incomplete information systems[J]. Information Sciences 1998 112 39 – 49.

Knowledge Reduction Based on Decision Attribute Support Degree
under Incomplete Information System

TAO Zhi SHANG Wei LI Weimin

(Civil Aviation University of China Science College Tianjin 300300 China)

Abstract A kind of attribute relative reduction for decision attribute support degree was proposed. States the significance of attributes in the incomplete decision table by introducing decision attribute support degree, which is then used as heuristic information for selecting attributes. This algorithm has a polynomial time complexity. Seeking the minimal relative reduction in a decision table is typically a NP-hard problem, and its complexity can be reduced by using the algorithm provided in this paper. An example shows this algorithm can achieve the minimal relative reduction of incomplete decision table.

Key words rough set; decision attribute support degree; incomplete decision table; attribute reduction

(上接第 22 页)

A Novel Adaptive Group of Pictures Structure

WANG Yuan-gen^{1, 2}, LIANG Fan¹, ZHANG Xirui³, XIAO Mingming^{1, 2}

(1. Department of Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou 510275, China;
2. Department of Information Engineering, Zhongkai Agrrotechnical College, Guangzhou 510225, China;
3. Jingnen Technical College, Jingnen 448000, China)

Abstract Motion-compensated temporal filtering (MCTF) decomposed a group of pictures (GOP) consisted of several successive frames to support temporal scalability by using subband/wavelet transform. Compared with fixed-size GOP structure, adaptive GOP structure (AGS) improves the coding performance, but the computational complexity of AGS is also increased significantly. A novel adaptive GOP structure (NAGS) is proposed to estimate the temporal characters of video sequences by inter-prediction. According to the temporal characters, it adopts different size of GOP. NAGS significantly reduces the computational complexity of AGS and fixes the problem of the jitter of temporal decomposition delay in AGS. The experimental results show that NAGS has almost the same coding performance as AGS.

Key words SVC; MCTF; lifting; GOP; AGS