

用于可分级视频编码的 MCTF 方法

王永玉¹, 袁超伟¹, 孙 衢²

(1. 北京邮电大学通信网络综合技术研究所, 北京 100876; 2. 四川大学电气信息学院, 成都 610065)

摘要: 提出用于可分级视频编码的 3 带小波提升与运动模型自适应结合的方法, 将 3 带小波的提升步与运动补偿自适应相结合来实现时间域的小波变换, 扩展了经典的运动补偿时域滤波, 实现了时间分级的灵活性, 同时又可以根据实际的视频序列选择适当的运动模型, 从而达到更好的编码效果。仿真结果验证了该方法的有效性和性能的优越性。

关键词: 小波变换; 提升; 可分级视频编码; 3 带时域分解; 运动补偿时域滤波

MCTF Method for Scalable Video Coding

WANG Yong-yu¹, YUAN Chao-wei¹, SUN Qu²

(1. School of Telecommunication and Network Technology, Beijing University of Posts and Telecommunications, Beijing 100876;

2. School of Electrical Engineering and Information, Sichuan University, Chengdu 610065)

【Abstract】 A new method for scalable video coding, using three-band lifting-based motion adaptive transform, is presented in this paper. A three-band motion-compensated lifting steps are used to implement the temporal wavelet transform. It extends the classical Motion-Compensated Temporal Filtering(MCTF) and offers higher temporal scalability flexibility, and it provides improved compression performance by using specific motion model according to real video sequences. The experimental results demonstrate the effectiveness of the framework.

【Key words】 wavelet transform; lifting; scalable video coding; three-band temporal decomposition; Motion-Compensated Temporal Filtering(MCTF)

1 概述

由于小波变换能够对信号进行多尺度表示, 并且基于小波系数的编码方法具有很好的嵌入式量化特性, 统一编码后可以在任意码流截断进行解码, 因此基于小波的视频编码非常容易实现编码的可分级。

文献[1]首先提出了在视频压缩中使用三维离散小波变换(3D-DWT)技术, 随后其他学者提出了多种编码算法, 如著名的 3D-SPIHT 算法^[2]。采用 3D-SPIHT 算法可以实现较好的编码效率, 然而缺乏运动补偿会导致时域方向的滤波在低通子带上形成明显的“鬼影效应”(ghosting artifacts)。因此, 如何将时域小波变换与运动补偿有效地结合是视频编码技术的关键, 具有重要的实际意义。文献[3]提出将三维小波变换与基于块的运动补偿相结合的方法; 文献[4]提出将三维提升小波变换与不规则网格运动补偿方法相结合的可分级视频编码^[4]。然而, 这些方法都是基于经典的 Haar 小波或更长的时域滤波器(如 5/3 滤波器)^[5]。这些经典的 2 带分解提供了时间可分级的 2 值因子, 如完全的时间分辨率为 30 帧/s, 还可提供 15 帧/s 和 7.5 帧/s 的时间分级。

为了适应更灵活的时间分级要求, 文献[6]提出了 3 带小波提升与基于块的运动补偿相结合的可分级视频编码方法, 实现了更灵活的时间可分级(例如, 可提供 30 帧/s、20 帧/s 和 10 帧/s 的时间分级)和改善的压缩性能, 然而在视频序列中存在较多的拉伸或收缩运动时, 基于块的运动补偿模型不能够达到很好的效果。

2 算法的构造与性能比较

有效的可分级视频编码的关键是在空间域、时域小波变

换中有效地结合运动补偿技术。这种有效的结合要满足 3 个要求: (1)保证在低频子带帧下实现高质量的重构, 特别是不能出现所谓的“鬼影效应”; (2)实现高的编码增益; (3)保证变换的可逆性。本文提出的 3 带小波提升与运动模型自适应结合的方法, 通过沿着真实场景的运动轨迹进行合适的小波滤波来实现前 2 点要求, 给出了其实现完美重构的条件, 保证其变换的可逆性。

2.1 类似于 Haar 的 3 带小波提升与运动模型相结合方法

在 3 带小波分解中, 2 个预测操作符(P^+ 和 P^-)产生 2 个细节子带; 2 个更新操作符(U^+ 和 U^-)计算近似子带。当 $P^+ = P^- = 1, U^+ = U^- = 1/4$ 时得到最简单 3 带小波提升格式, 即类似于 Haar 的 3 带小波^[6], 如图 1 所示。

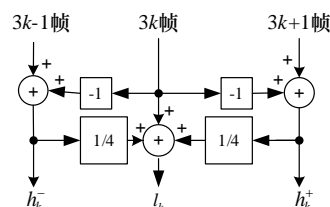


图 1 类似于 Haar 的 3 带小波提升格式

与 2 带 Haar 小波相比, 3 带小波除了增加时间的灵活性之外, 由于双向的更新操作, 将会获得效果更好的低频子带。

基金项目: 国家自然科学基金资助项目(60672132)

作者简介: 王永玉(1973-), 女, 博士研究生, 主研方向: 小波理论, 图像及视频编码技术; 袁超伟、孙 衢, 教授、博士生导师

收稿日期: 2007-08-07 **E-mail:** yongyuwang@sina.com

本文在提升步中加入运动补偿的统一结构如下: $x_k[n] \equiv x_k[n_1, n_2]$ 为原始视频序列第 k 帧的采样; $h_k[n] \equiv h_k[n_1, n_2]$ 为分解的细节(高频)子带; $l_k[n] \equiv l_k[n_1, n_2]$ 为分解的近似(低频)子带; 设帧 $k_1 \sim$ 帧 k_2 的运动补偿映射为 $W_{k_1 \rightarrow k_2}$, 使得 $W_{k_1 \rightarrow k_2}(x_{k_1})[n] \approx x_{k_2}[n]$, 此处并不指定特定的运动模型, 可以为基于块的运动模型或基于网格的运动模型等任何可逆运动模型。则加入运动模型后的提升格式为

$$h_{3k+1}^+[n] = x_{3k+1}[n] - W_{3k \rightarrow 3k+1}(x_{3k})[n] \quad (1)$$

$$h_{3k}^-[n] = x_{3k-1}[n] - W_{3k-1 \rightarrow 3k}(x_{3k-1})[n] \quad (2)$$

$$l_k[n] = x_{3k}[n] + \frac{1}{4}W_{3k+1 \rightarrow 3k}(h_{3k+1}^+[n]) + \frac{1}{4}W_{3k-1 \rightarrow 3k}(h_{3k}^-[n]) \quad (3)$$

其中, 2 个高频子带相当于运动补偿后的残差, 在运动模型精确的情况下应该接近于 0; 低频子带相当于沿着运动轨迹进行有效的低通滤波。如果运动模型设计得非常完美, 将会获得如图 2 所示的效果。

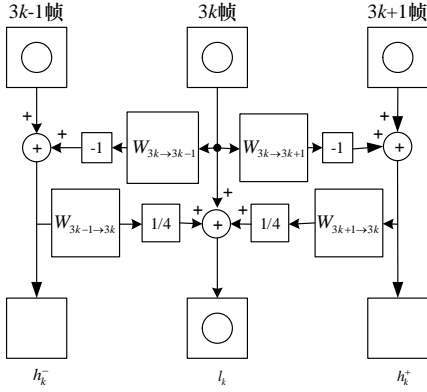


图 2 类似于 Haar 的 3 带小波提升与运动补偿相结合的格式

由图 2 可见, 沿着精确的运动轨迹做小波滤波可获得高质量的低频子带, 并且高频子带为 0, 获得了高的编码增益, 满足了有效可分级编码的前 2 个要求。对于第 3 点要求——变换的可逆性, 这种类似于 Haar 的 3 带小波可由其提升格式直接得出

$$x_{3k}[n] = l_k[n] - \frac{1}{4}W_{3k+1 \rightarrow 3k}(h_{3k+1}^+[n]) - \frac{1}{4}W_{3k-1 \rightarrow 3k}(h_{3k}^-[n]) \quad (4)$$

$$x_{3k+1}[n] = h_{3k+1}^+[n] + W_{3k \rightarrow 3k+1}(x_{3k})[n] \quad (5)$$

$$x_{3k-1}[n] = h_{3k}^-[n] + W_{3k-1 \rightarrow 3k}(x_{3k-1})[n] \quad (6)$$

2.2 类似于 5/3 的 3 带小波提升与运动模型相结合的方法

类似于 Haar 的 3 带小波采用了双向的更新操作。类似于 5/3 的 3 带小波还采用了双向的预测操作, 因此在视频编码中将会获得更好的效果^[6]。在提升步中加入运动补偿模型后, 其提升格式为

$$h_{3k+1}^+[n] = x_{3k+1}[n] - \beta W_{3k+2 \rightarrow 3k+1}x_{3k+2}[n] - (1-\beta)W_{3k \rightarrow 3k+1}x_{3k}[n] \quad (7)$$

$$h_{3k}^-[n] = x_{3k-1}[n] - \beta W_{3k-2 \rightarrow 3k-1}x_{3k-2}[n] - (1-\beta)W_{3k-1 \rightarrow 3k}x_{3k}[n] \quad (8)$$

$$l_k[n] = x_{3k}[n] + \frac{1}{4}W_{3k+1 \rightarrow 3k}h_{3k+1}^+[n] + \frac{1}{4}W_{3k-1 \rightarrow 3k}h_{3k}^-[n] \quad (9)$$

其中, $\beta \in [0, 1]$ 为权重因子, 可以根据实际应用中所选取的运动模型进行调节。当 $\beta = 0$ 时, 所得提升结构与类似于 Haar 的 3 带小波格式相同。图 3 给出了提升步中加入运动模型的效果。高频子带是进行了双向运动补偿预测后的参差帧, 低频帧是沿着运动轨迹进行低通滤波所得的。

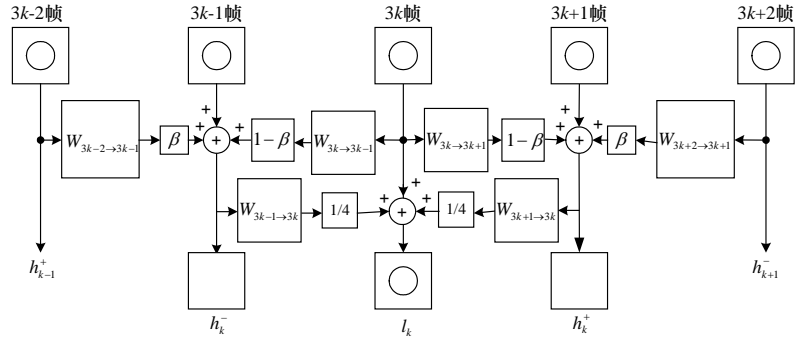


图 3 类似于 5/3 的 3 带小波提升与结合运动补偿相结合的格式

对于变换的可逆性, 由式(9)可直接解得

$$x_{3k}[n] = l_k[n] - \frac{1}{4}W_{3k+1 \rightarrow 3k}h_{3k+1}^+[n] - \frac{1}{4}W_{3k-1 \rightarrow 3k}h_{3k}^-[n] \quad (10)$$

设所选取的运动模型可逆, 即 $W_{k_1 \rightarrow k_2}W_{k_2 \rightarrow k_1} = 1$, 由式(7)和式(8)及图 3 可解得

$$x_{3k+1}[n] = \frac{1}{1-\beta^2}[\beta W_{3k+2 \rightarrow 3k+1}h_{3k+1}^-[n] + (1-\beta)W_{3k \rightarrow 3k+1}x_{3k}[n] + \beta(1-\beta)W_{3k+3 \rightarrow 3k+2}W_{3k+2 \rightarrow 3k+1}x_{3k+3}[n]] \quad (11)$$

$$x_{3k+2}[n] = \frac{1}{1-\beta^2}[\beta W_{3k+1 \rightarrow 3k+2}h_{3k+1}^+[n] + h_{3k+1}^-[n] + (1-\beta)W_{3k+3 \rightarrow 3k+2}x_{3k+3}[n] + \beta(1-\beta)W_{3k \rightarrow 3k+1}W_{3k+1 \rightarrow 3k+2}x_{3k}[n]] \quad (12)$$

可见, 在运动模型可逆的前提下, 此框架保证了可逆性。由于通常采用的大多数运动补偿模型都是可逆的, 或通过插值滤波等方法使其可逆, 因此此框架可实现有效的可分级视频编码。

3 仿真实例

要实现沿着运动轨迹进行有效的子带分解, 运动模型的设计非常重要。目前主要有 2 类运动模型: 一类是常用的基于块的运动模型; 另一类是基于可变形网格的运动模型。

(1) 在基于块的运动补偿模型中, 图像先分为 16×16 大小的块, 然后进行块的匹配搜索, 搜索范围为 8×8 , 搜索精度为 $1/4$ 。对于连接像素, 帧 k_1 到帧 k_2 的运动补偿映射 $W_{k_1 \rightarrow k_2}$ 为 $n_{k_2} = n_{k_1} - v$, 使得 $W_{k_1 \rightarrow k_2}(x_{k_1})[n] \approx x_{k_2}[n]$; 帧 k_2 到帧 k_1 的运动补偿映射 $W_{k_2 \rightarrow k_1}$ 为 $n_{k_1} = n_{k_2} + v$, v 为块平移矢量。对于无连接像素, $W_{k_1 \rightarrow k_2} \equiv W_{k_2 \rightarrow k_1} \equiv 0$ 。可见, 此运动模型是可逆的, 从而可获得有效的可分级视频编码。

(2) 在基于网格的运动补偿模型中, 采用规则的三角形网格划分方法: 图像先分为 16×16 大小的块, 然后沿每块的对角线分成规则的三角形。其中, 帧 k_1 到帧 k_2 的三角形网格节点的运动补偿映射 $W_{k_1 \rightarrow k_2}$ 为 $n_{k_2} = n_{k_1} - v$, 使得 $W_{k_1 \rightarrow k_2}(x_{k_1})[n] \approx x_{k_2}[n]$, v 为网格节点平移矢量, 块匹配方法, 搜索范围为 8×8 , 搜索精度为 $1/4$ 。帧 k_1 到帧 k_2 的每个三角形内部像素的运动补偿映射 $W_{k_1 \rightarrow k_2}$ 为 $n_{k_2} = An_{k_1} + B$, 使得 $W_{k_1 \rightarrow k_2}(x_{k_1})[n] \approx x_{k_2}[n]$, A, B 为仿射变换系数, 通过每个三角形网格节点的平移计算得出。可见, 此运动模型也是可逆的。

表 1 和表 2 给出了在本文所提出的框架下, 3 带小波提升分别与基于块的运动补偿和基于三角形网格的运动补偿模型相结合的测试结果, 以及 2 带小波提升结合三角形网格运动补偿的测试结果。其中, 3 带小波进行 3 级时域分解, 每个视频帧组包括 27 帧图像; 2 带小波进行 5 级时域分解, 每个视频帧组包括 32 帧图像; 时间分解子带都采用标准的双正

交 9/7 小波进行 4 级空间分解。

表 1 序列“MOBILE”，CIF 格式，30 帧/s 的测试结果(PSNR) dB

码率 (Kb·s ⁻¹)	3 带 Haar +块补偿	3 带 5/3 (β 取 0.19) +块补偿	3 带 Haar +网格	3 带 5/3 (β 取 0.20) +网格	2 带 Haar +网格	2 带 5/3 +网格
400	26.31	26.49	26.38	26.54	24.85	25.37
600	28.55	28.89	28.65	28.96	27.55	28.29
800	30.11	30.28	30.19	30.36	29.13	29.74
1 200	31.95	32.39	31.96	32.45	31.49	31.89

表 2 序列“FOREMAN”，CIF 格式，30 帧/s 的测试结果(PSNR) dB

码率 (Kb·s ⁻¹)	3 带 Haar +块补偿	3 带 5/3 (β 取 0.15) +块补偿	3 带 Haar +网格	3 带 5/3 (β 取 0.15) +网格	2 带 Haar +网格	2 带 5/3 +网格
400	31.39	31.45	31.91	31.95	30.82	31.05
600	32.87	33.05	33.22	33.53	32.68	32.98
800	34.20	34.45	34.75	34.99	33.87	34.22
1 200	35.86	36.25	36.29	36.54	35.73	36.12

从测试结果可见，对于序列中存在较多的扩张及收缩运动的“FOREMAN”序列，3 带小波提升与基于三角形网格的运动补偿模型相结合的性能明显优于 3 带小波提升与基于块的运动补偿模型相结合。两者都可以用统一框架来实现，只是在与运动模型相结合的过程中采用不同的运动映射。结果还显示，类似 5/3 的 3 带小波提升与运动补偿相结合性能优于类似 Haar 的 3 带小波提升与运动补偿相结合性能；而 3 带小波提升与运动补偿相结合的性能总体上又高于 2 带 MCTF 性能。

4 结束语

本文给出了用于分级视频编码的 3 带小波提升与运动模型自适应相结合的方法，与基于传统的 2 带小波的 MCTF 相比，它的主要优势在于实现了以因子为 3 的时域可分级，而

试验结果又显示了在相同条件下优于 2 带小波的性能；并且所述方法可以根据实际的视频序列选择适当的运动模型，从而获得性能更优越的视频编码系统。此外，还可在此框架基础上结合其他的 3 带小波(如构造类似 9/7 的 3 带小波)构造更适合应用的视频编码系统。

参考文献

- [1] Karlsson G, Vetterli M. Three-dimensional Subband Coding of Video[C]/Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing. New York, USA: [s. n.], 1988: 1100-1103.
- [2] Xiong K Z, Pearlman W A. Low Bit-rate Scalable Video Coding with 3D Set Partitioning in Hierarchical Trees(3D-SPIHT)[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(8): 1374-1387.
- [3] Ohm J. Three Dimensional Subband Coding with Motion Compensation[J]. IEEE Trans. on Image Proc., 1994, 3(5): 559-571.
- [4] Secker A, Taubman D. Highly Scalable Video Compression Using a Lifting-based 3D Wavelet Transform with Deformable Mesh Motion Compensation[C]/Proc. of the International Conference on Image Processing. Rochester, New York, USA: [s. n.], 2002: 749-752.
- [5] Sweldens W. The Lifting Scheme: A Custom-design Construction of Biorthogonal Wavelets[J]. Appl. Comput. Harmon. Anal., 1996, 3(2): 186-200.
- [6] Tillier C, Pesquet-Popescu B, Schaar M V D. 3-band Motion-compensated Temporal Structures for Scalable Video Coding[J]. IEEE Trans. on Image Proc., 2006, 15(9): 2545-2557.

(上接第 207 页)

50 城市和 75 城市的求解过程中，一些蚁群算法和鱼群算法都很难寻得全局最优解。但本文算法由于采用拥挤度渐变的办法，同时兼顾了两种算法的优点，在 50 城市 TSP 问题的求解中仍然可寻得全局最优解 427.855。而对于 75 城市的 TSP 问题，本文算法可寻得的最优解为 543.447，优于 Fogel^[6] 所得最优解 549.18。由此可见，本文所述混合优化算法确实具有较强的全局寻优能力。本文算法对 50 城市和 75 城市 TSP 问题求解所得最优解的形式如图 1、图 2 所示。

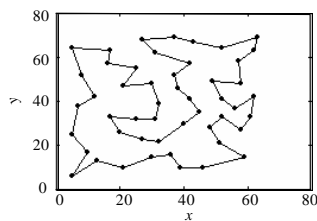


图 1 50 城市 TSP 问题最优解

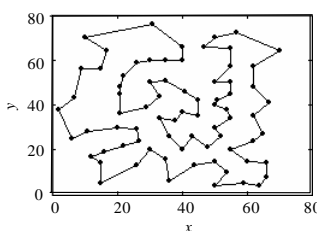


图 2 75 城市 TSP 问题所求结果

4 结束语

本文分析了蚁群和鱼群两种优化算法的寻优机理，通过类比揭示了两种优化算法中寻优个体在寻优过程中所遵循的运动规律的差异。并将鱼群算法中拥挤度的作用引入到蚁群算法中，提出了一种新的混合优化算法。在优化过程中拥挤度的指导作用逐渐减弱，既克服了蚁群算法初期可能出现的早熟等现象，增强了算法遍历寻优的能力，又使得算法具有较快的收敛速度，提高了算法的寻优性能。大量的仿真实验证明了该算法的有效性。

参考文献

- [1] 朱 玲, 施心陵, 刘亚杰, 等. 基于蚁群算法的甲状腺结节超声图像边缘检测法[J]. 计算机工程, 2006, 32(24): 178-179.
- [2] 陈传波, 罗增琦. 一种基于蚁群聚类的快速算法[J]. 计算机工程, 2007, 33(6): 206-207.
- [3] 俞 洋, 殷志锋, 田亚菲. 基于自适应人工鱼群算法的多用户检测器[J]. 电子与信息学报, 2007, 29(1): 121-124.
- [4] 卢辉斌, 范庆辉, 贾兴伟. 一种改进的自适应蚁群算法[J]. 计算机工程与设计, 2005, 26(11): 3065-3066.
- [5] 王 凌. 智能优化算法及其应用[M]. 北京: 清华大学出版社, 2001.
- [6] Fogel D B. Applying Evolutionary Programming to Selected Traveling Salesman Problems[J]. Cybernetics and System, 1993, 24(1): 27-36.