

基于冗余小波 DT 网格技术的单通道视频编码^{*}

高 韬¹, 刘正光¹, 于 明²

(1. 天津大学 电气与自动化工程学院, 天津 300072; 2. 河北工业大学 计算机科学与软件学院, 天津 300130)

摘 要: 提出了一种新型的基于冗余小波 DT(Delaunay triangle) 网格技术的单通道视频编码算法, 该算法通过冗余小波变换提取视频流中的运动信息, 结合 DT 和仿射变换算法对单通道视频流进行压缩。实验结果表明, 该算法在提高 PSNR 值的基础上减少了编码系统的复杂度, 具有一定的应用价值。

关键词: 视频编码; 冗余小波; Delaunay 三角网格

中图分类号: TN919.8 文献标志码: A 文章编号: 1001-3695(2008)11-3493-03

Single-channel video coding based on redundant wavelet and DT grid

GAO Tao¹, LIU Zheng-guang¹, YU Ming²

(1. College of Electrical Engineering & Automation, Tianjin University, Tianjin 300072, China; 2. College of Computer Science & Engineering, Hebei University of Technology, Tianjin 300130, China)

Abstract: This paper proposed a single-channel video coding method based on redundant wavelet and DT(Delaunay triangle) grid. The algorithm extracted the features and potential motion areas in the redundant wavelet transform domain, and then did the motion estimation in the time-domain. Experimental results show that the algorithm can increase the PSNR value and reduce the complexity. It has a certain practical value.

Key words: video coding; redundant wavelet; DT grid

利用人的视觉获取的信息称为视频信息, 视频信息的表现形式是视频信号。一般而言, 视频信号的信息量大, 传输网络所需的带宽相对较宽。例如一路高清新电视信号, 由于其信息量巨大, 不压缩需 1 Gbps。可见, 视频信息虽然具有直观性、确定性等优越特点, 但要传输包含视频信息的信号却需要较高的网络带宽, 所以视频信号在传输前要先进行压缩编码, 然后再在网络上进行传输。

目前, 视频编码的方式一般分为基于波形的编码和基于内容的编码^[1]。基于波形的编码采用了将预测编码与变换编码组合起来的基于块的混合编码方法。为了减少编码的复杂性, 使视频编码的操作易于执行, 在采用混合编码方法时, 首先将一幅图像分成固定大小的块, 然后对块进行压缩编码处理。基于内容的编码首先将视频帧分成对应于不同物体的区域, 然后分别对其进行编码。具体来说, 即对不同物体的形状¹ 运动和纹理进行编码。在最简单的情况下, 利用二维轮廓描述物体的形状, 利用运动矢量描述其运动状态, 而纹理则用颜色的波形进行描述。

针对基于内容和基于波形编码各自的特点, 本文提出一种以冗余小波变换为基础, 结合 DT 网格技术对单通道视频流进行压缩的视频压缩新算法。

1 视频压缩关键技术——运动估计

在视频压缩算法中, 运动估计是非常重要的部分, 也是视频压缩过程中耗时最长的一部分, 约占整个压缩过程运行时间的 50%^[2]。所以在保证估计的准确度基础上减少运动估计运行时间成为视频压缩的研究热点之一。

目前运动估计通常采用的搜索算法有全搜索算法、二维对数算法、“钻石”搜索算法。全搜索算法一定能够找到最优的匹配块, 但是计算复杂度很高, 硬件实现全搜索算法非常困难。二维对数和“钻石”搜索算法能够有效减少计算的复杂度, 但是搜索到的结果可能是局部最优点, 而不是全局最优点。这些算法都在时域进行, 而在小波变换域进行运动估计则引起了人们的关注。

文献[3, 4] 提出了冗余小波变换域运动估计方法, 对于搜索窗内的一个向量, 计算所有尺度下所有子带的绝对差, 然后全部相加, 结果称为交叉子带或交叉尺度形变。使交叉子带形变最小的矢量作为当前点的运动矢量(Δ_x, Δ_y) :

$$(\Delta_x, \Delta_y) = \arg \min_{-W \leq \Delta_x, \Delta_y \leq W} MAE(x - B/2, y - B/2, \Delta_x, \Delta_y)$$

其中 $MAE(x, y, \Delta_x, \Delta_y) = (1/B^2) \sum_{k=1}^B \sum_{l=1}^B AE(x + k, y + l, \Delta_x, \Delta_y)$

绝对差(AE) 为

$$AE(x, y, \Delta_x, \Delta_y) = \sum_{j=1}^J 2^{-j} \{ |V_j^c(x + \Delta_x, y + \Delta_y) - V_j^r(x, y)| + |H_j^c(x + \Delta_x, y + \Delta_y) - H_j^r(x, y)| + |D_j^c(x + \Delta_x, y + \Delta_y) - D_j^r(x, y)| + 2^{-J} |B_j^c(x + \Delta_x, y + \Delta_y) - B_j^r(x, y)| \}$$

其中: c, r 分别代表当前子带和参考子带; B_j, H_j, V_j 和 D_j 分别代表 j 尺度下近似子带分量、水平子带分量、垂直子带分量和对角线子带分量。搜索过程中, 运动矢量大小在搜索窗范围内, N 大小不固定。该算法利用小波变换的空间方向性提高搜索精度, 但是运算复杂度大、精度不稳定。文献[5] 提出的 EMRMC 方法是规则网格运动估计方法, 该方法利用运动图像极大的空间冗余性, 通过提取潜在运动区(PMA) 有效地缩小了搜索区域, 提高了效率。但是, 由于各分解子带大小不等, 需

收稿日期: 2007-12-06; 修回日期: 2008-04-07 基金项目: 国家自然科学基金资助项目(60302018)

作者简介: 高韬(1981-), 男, 博士研究生, 主要研究方向为模式识别、图像处理等(ggao0390@ yahoo. com. cn); 刘正光, 男, 教授, 博导, 主要研究方向为图像处理与模式识别等; 于明(1964-), 男, 教授, 院长, 主要研究方向为图像处理与模式识别。

要采用逐层扩散的方法得到潜在运动区, 运算量也比较大。

通过以上分析, 本文提出了一种冗余小波变换域提取特征点和潜在运动区域, 利用冗余小波变换各子带等大小的性质, 并结合时域搜索方法进行运动估计, 在给运算过程带来很大方便的同时提高了编码的质量。首先通过小波变换的方法提取特征点^[3], 得到 Delaunay 三角形网格; 再根据本文的算法在冗余小波变换域提取 PMA, 并考虑参考帧内 $N \times N$ 邻域的搜索区, 只对落在两区域的点进行运动估计。由于运动发生的局部性和连贯性以及小波变换的方向选择性, 使得潜在运动区域在高频子带通常只占整个图像的很小部分, 本算法只在很小的搜索区对有效的网格顶点进行运动估计, 从而提高了运行效率。

1.1 冗余小波变换理论基础

设 MRA 由满足如下条件的嵌套线性空间 $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$ 组成:

- a) 嵌套空间的并集在平方可积函数空间中稠密。
- b) 嵌套空间的交集只包含零向量。
- c) 如果 $f(t) \in V_k$, 那么 $f(2t) \in V_{k-1}$; 反之亦然。
- d) 存在函数(即尺度函数) $\phi(t)$, 使得 $\{\phi(t-k): k \in \mathbb{Z}\}$ 是 V_0 的基, 而 $\phi(t)$ 被定义为满足下列条件的函数:
 - (a) $\int_{-\infty}^{\infty} \phi(t) dt = 1$;
 - (b) $\|\phi(t)\|^2 = \int_{-\infty}^{\infty} |\phi(t)|^2 dt = 1$;
 - (c) $\langle \phi(t), \phi(t-n) \rangle = \delta(n)$ 。

则存在 $c(n), n=0, \pm 1, \pm 2, \cdots$, 使得 $\phi(2^{-k}t) = \sum_{n=-\infty}^{\infty} c(n) \phi(2^{-(k-1)}t-n)$, 说明 $V_k \subset V_{k-1}$, 即构成了一个 MRA。形象地说, 所有子空间的并集($V_{-\infty}$) 在 $L^2(R)$ 中稠密, 说明 $V_{-\infty}$ 能够逼近任意信号 $f(t) \in L^2(R)$ 。若令 $f_k(t)$ 为 $f(t)$ 在子空间 V_k 上的正交投影, 有 $f_k(t) = \sum_{n=-\infty}^{\infty} a(k, n) \phi(2^{-(k-1)}t-n)$, $k=0, \pm 1, \pm 2, \cdots$, 则可以得到 $a(k, n) = \sum_{m=-\infty}^{\infty} a(k-1, m) c(m-2n)/2$ 。假设 $\psi(t)$ 是 V_{-1} 中的函数, 且满足: $\int_{-\infty}^{\infty} \psi(t) dt = 0$; $\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$; $\langle \psi(t), \psi(t-n) \rangle = \delta(n)$; $\langle \psi(t), \phi(t-n) \rangle = 0$ 。则 $\psi(t)$ 就是与上述 MRA 对应的 DWT 需要的小波。令 W_0 是以 $\{\psi(t-n): n \in \mathbb{Z}\}$ 为基的线性向量空间, 则 $V_{k-1} = \bigoplus_{j=k}^{\infty} W_j$, 那么 $V_{-\infty} = \bigoplus_{j=-\infty}^{\infty} W_j$, 于是 $f_{-\infty}(t) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} b(k, l) \psi(2^{-k}t-l)$ 。同样, 可以导出 $b(k, n) = \sum_{m=-\infty}^{\infty} a(k-1, m) d(m-2n)/2$ 。令, $h(n) \equiv c(n)/2$, $g(n) \equiv d(n)/2$, $h(n) \equiv h(-n)$, $g(n) \equiv g(-n)$, 则 $a(k, n) = \sum_{m=-\infty}^{\infty} a(k-1, m) h(2n-m)$, $b(k, n) = \sum_{m=-\infty}^{\infty} a(k-1, m) g(2n-m)$ 。可以看做 $a(k, n)$ 、 $b(k, n)$ 是由 $a(k-1, m)$ 经过滤波器 h 滤波后抽样得到的。如果从频率域考虑, 令 $H(\omega) = \sum_n h(n) e^{-j\omega n}$, 则有 $|H(\omega)|^2 + |H(\omega+\pi)|^2 = 1$, 即 H 是一个正交镜像滤波器(QMF)。通过多分辨率分析可以将小波的求解转换为对数字滤波器的设计问题^[6,7]。

如果二维信号 $x(n, m) (n=1, 2, \cdots, N; m=1, 2, \cdots, N)$ 进行二维冗余离散小波变换后的低频分量、水平细节分量、垂直细节分量、对角细节分量分别为 $A_j(x, y)$ 、 $H_j(x, y)$ 、 $V_j(x, y)$ 和 $D_j(x, y)$, 小波滤波器组的分析滤波器和合成滤波器分别为 $\{h(k), g(k)\}$ 和 $\{h(k), g(k)\}$, 则二维信号分解过程为

$$A_j(x, y) = A_{j-1}(x, y) ([h]_{\uparrow 2^{j-1}}, [h]_{\uparrow 2^{j-1}})(-x, -y);$$

$$H_j(x, y) = H_{j-1}(x, y) ([h]_{\uparrow 2^{j-1}}, [g]_{\uparrow 2^{j-1}})(-x, -y);$$
$$V_j(x, y) = V_{j-1}(x, y) ([g]_{\uparrow 2^{j-1}}, [h]_{\uparrow 2^{j-1}})(-x, -y);$$
$$D_j(x, y) = D_{j-1}(x, y) ([g]_{\uparrow 2^{j-1}}, [g]_{\uparrow 2^{j-1}})(-x, -y)。$$

1.2 基于冗余小波变换的运动估计

1.2.1 冗余小波变换产生网格顶点

信号的奇异点在小波变换下, 随着尺度的增大而增大, 相邻尺度的小波系数直接乘可以增强信号的特征点。冗余小波具有平移不变性, 各子带图像大小等尺寸相同, 便于计算。因此算法采用相邻尺度子带系数相乘后求和的方法得到模板, 计算所有像素点的模板值, 如果大于阈值即确定该点为特征点。这样整个图像的特征点就确定下来, 从而生成一个基于图像内容的 Delaunay 三角形网格。模板如下:

$$\text{mask}(x, y) = |\prod_{j=J_0}^{J_1} HL^{(j)}(x, y)| + |\prod_{j=J_0}^{J_1} LH^{(j)}(x, y)| + |\prod_{j=J_0}^{J_1} HH^{(j)}(x, y)|$$

其中: J_0 和 J_1 代表开始和结束的尺度。因为冗余小波各子带的大小相同, 可以直接作算术运算。阈值 T 的提取操作可描述为将 $\text{mask}(x, y)$ 归一化后转换为 256 级灰度图像 mask , 通过迭代法自动确定阈值。大概步骤如下: 首先取图像灰度范围的中值作为初始阈值 T_0 (设共有 L 个灰度), 然后按下式进行迭代:

$$T_{i+1} = 1/2 \sum_{k=0}^{T_i} h_k \times k / \sum_{k=0}^{T_i} h_k + \sum_{k=T_i+1}^{L-1} h_k \times k / \sum_{k=T_i+1}^{L-1} h_k$$

其中: h_k 是灰度为 k 值的像素个数, 迭代一直到 $T_{i+1} = T_i$ 结束, 取结束时的 T_i 为分割阈值。

图 1 是通过冗余小波变换对 Akiyo 和 mother 序列一帧图像进行特征点提取形成的 Delaunay 三角形网格。



图 1 Delaunay 三角形网格特征点提取

1.2.2 结合 PMA 和网格顶点的运动估计

通过冗余小波变换提取特征点给后续的运动估计带来了方便和更高的效率。结合冗余小波变换的各子带系数高度相关、方向选择性、各子带与输入信号大小相等以及平移不变的性质, 提出了在冗余小波变换域进行的 PMA 模板提取, 并结合时域方法的运动估计和补偿。模板由多子带组成, 包括近似子带(LL)、低—高子带(LH)、高—低子带(HL)、高—高子带(HH)。提取的 PMA 模板如下:

$$\text{mask}(x, y) = \sum_{j=J_0}^{J_1} \left(|LL_C^{(j)}(x, y) - LL_R^{(j)}(x, y)| + |LH_C^{(j)}(x, y) - LH_R^{(j)}(x, y)| + |HL_C^{(j)}(x, y) - HL_R^{(j)}(x, y)| + |HH_C^{(j)}(x, y) - HH_R^{(j)}(x, y)| \right)$$
$$\text{PMA}(x, y) = \begin{cases} 1 & \text{mask}(x, y) \geq \Gamma_0 \\ 0 & \text{mask}(x, y) < \Gamma_0 \end{cases}$$

其中: J_0 和 J_1 代表开始和结束的尺度。因为冗余小波各子带的大小相同, 可以直接作算术运算。 X_C 、 X_R 分别代表当前图像和参考图像, Γ_0 为阈值。对 Akiyo 和 mother 序列相邻两帧图像进行 PMA 提取的结果如图 2 所示。



图 2 PMA 提取

网格顶点的运动估计是在经典的以该点为中心的宏块运动估计方法基础上的改进,即在当前帧以网格顶点为中心取得 $M \times M$ 的宏块,然后在参考帧以该点位置为中心 $L \times L$ (对于图像边缘的特征点的搜索,会随着到边界的距离而调整搜索窗)的范围内搜索匹配宏块,位移差即该顶点的运动矢量。本算法结合 PMA,在 PMA 中搜索匹配宏块,即判断当前帧和参考帧的网格顶点是否落在 PMA 内。不在 PMA 的当前帧和参考帧的网格顶点运动矢量规定为 0。值对落在 PMA 内的网格顶点进行运动估计。这样做的优点在于:由于连续两帧图像是非常相似的,大部分区域是相同的,PMA 常常分布稀疏且面积很小,运动矢量的搜索可以在局部高效完成。本文采用的最佳匹配准则为绝对帧间差和准则(SAD)^[8]。其数学表达式为

其中: $S_k(m, n)$ 为第 k 帧(m, n) 处的像素值; $S_{k-1}(m+i, n+j)$ 为 $k-1$ 帧在($m+i, n+j$) 处的像素值。获得三角形顶点的运动矢量后,三角形内部点的运动矢量可用六参数仿射变换求得:

其中: Δ_x 、 Δ_y 分别为 x 和 y 方向的运动矢量。由上式可确定该三角形区域的六个仿射变换系数,那么三角形区域内部各点的运动矢量也可求得。

2 单通道视频编码方案

本文提出的基于冗余小波的单通道视频编码方案框架如图 3 所示。视频序列第一帧作为 I 帧,可直接通过块 DCT 变换、量化、VLC 编码形成最终码流。后续帧可作为 P 帧,运动矢量由本文提出的算法得到并进行 VLC 编码,当前帧与预测帧的差值编码与 I 帧类似。从图 3 可以看出,本文提出的算法形成的最终码流具有兼容性,码流可在 MPEG 系列和 H.26x 系列的解码器上解码,无须另外设计解码器。

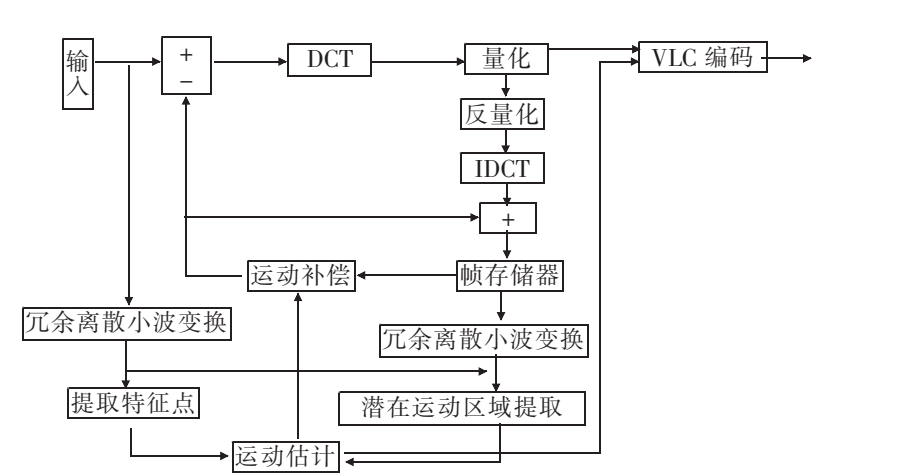


图 3 基于冗余小波的单通道视频编码框架

3 实验结果及讨论

本文对两个标准测试图像序列 Akiyo 和 mother 进行实验,将本文提出的方法与目前视频压缩算法 H.264 和 MPEG-4 进行比较。标准测试图像均为 CIF 格式,运动估计采用 Y-帧(亮度图像分量),包含了彩色视频图像几乎全部的运动信息。实验采用 Akiyo(352×288,15 帧)和 mother(352×288,15 帧);都采用第一帧作为 I 帧,后面帧作为 P 帧,根据前一帧的预测图像来估计当前帧。重建图像 PSNR 值如图 4、5 所示;对于视频图像数据的压缩率比较如表 1 所示。

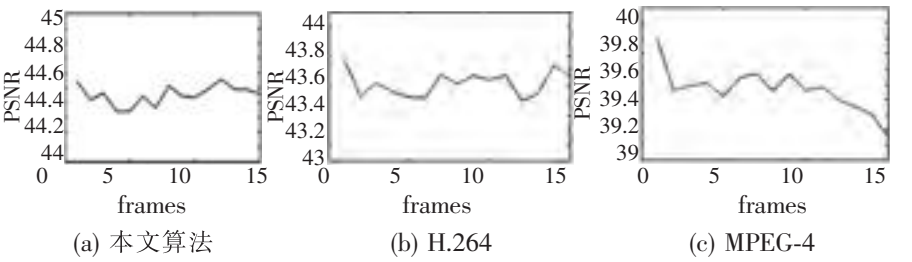


图 4 Akiyo 序列 PSNR 值比较

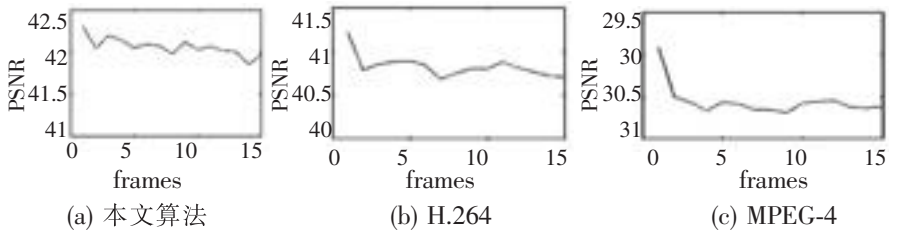


图 5 Mother 序列 PSNR 值比较

表 1 算法压缩率比较				
视频序列	原始大小 /MB	本文算法 /KB	H.264 /KB	MPEG-4 /KB
Akiyo sequence	2.17	30.9	22.8	24.8
mother sequence	2.17	17.6	10.6	8.29

从实验结果可得出,本文提出的算法得到的重建图像具有更好的主观视觉质量,PSNR 值要高于 H.264、MPEG-4 压缩算法,原因在于通过提取 PMA 和运动特征点提高了运动估计的准确性。从表 1 可以看出,本文提出的算法在压缩比方面要劣于 H.264 和 MPEG-4 算法,原因在于后续编码只是采用简单的 VLC 编码,如果结合 EZW 或者 SPIHT 编码,效果可能会有所提高。

4 结束语

本文提出了一种基于冗余小波变换提取 Delaunay 三角形网格特征点,结合 PMA 与三角仿射变换的单通道视频压缩方法。经实验结果证明,该方法能够在降低视频编码复杂性的同时提高 PSNR 值,主观视觉质量更好,具有一定的应用价值。今后研究重点在于提高算法的编码效率,减少码流的冗余度。

参考文献:

[1] 毕厚杰. 新一代视频压缩编码标准——H.264/AVC[M]. 北京:人民邮电出版社,2005.

[2] 李群迎,张晓林,刘荣科,等. 基于 TMS320C64x DSPs 的 MPEG-4 实时编码器设计与实现[J]. 电子技术应用,2005,31(7):43-45.

[3] CUI Su-xia, WANG Yong-hui, FOWLER J E. Mesh-based motion estimation and compensation in the wavelet domain using a redundant transform[C] //Proc of International Conference on Image Processing. Rochester: IEEE Press, 2002.

[4] CUI Su-xia, WANG Yong, FOWLER J E. Motion estimation and compensation in the redundant-wavelet domain using triangle meshes[J]. Signal Processing: Image Communication, 2006, 21(7):586-598.

[5] WEI Jie, LI Ze-nian. An enhancement to MRMC scheme in video compression[J]. IEEE Trans on Circuits and Systems for Video Technology, 1997, 7(3):564-568.

[6] DAUBECHIES I, SWELDENS W. Factoring wavelet transforms into lifting steps[J]. Journal of Fourier Analysis and Applications, 1998, 4(3):247-269.

[7] CALDERBANK A R, DAUBECHIES I, SWELDENS W, et al. Wavelet transforms that map integers to integers[J]. Journal of Applied and Computational Harmonic Analysis, 1998, 5(3):332-369.

[8] 尹黎明,王玲. 运动估计算法及其 DSP 优化[J]. 咸宁学院学报, 2005, 25(3):96-98.