一种新的小波域视频可分级运动估计方案

宋传鸣[1),3)] 王相海[1),2)]

[1)](辽宁师范大学计算机与信息技术学院 大连 116029)

[2)](北京大学视觉与听觉信息处理国家重点实验室 北京 100871)

[3)](南京大学计算机科学与技术系 南京 210093)

摘 要 首先提出了一种运动估计方法 MLBS(Modified Low Band Shift),并在此基础上提出了基于 MLBS 的小波域视频可分级运动估计方案 MLBSSME. 利用 MLBS 方法,提高了运动估计的准确度. 借助重叠块运动补偿方法,有效降低了最高分辨率下解码图像的块效应. 根据小波变换的多分辨率特性,利用缩小了的搜索窗口提高了搜索速度. 这种方法可有效应用于空间可分级和数率可分级的视频编码器中. 实验结果证明,对于多种类型的标准测试视频流,MLBSSME 算法始终能保持很高的估计精度. 利用该算法补偿得到的预测帧,其 *PSNR* 较之基于下层 LL 子带的分层运动估计方法和子带直接运动估计方法平均要高出 1～3dB,而对于空间细节较简单的视频,其 *PSNR* 比 LBS 方法提高了 0.5～1dB,并且算法的时空复杂度是 LBS 方法复杂度的 30%～40%.

关键词 运动估计;可分级编码;视频编码;小波域

中图法分类号 TP391

A New Scalable Video Motion Estimation Scheme in the Wavelet Domain

SONG Chuan-Ming[1),3)] WANG Xiang-Hai[1),2)]

[1)](*College of Computer and Information Technology*, *Liaoning Normal University*, *Dalian* 116029)

[2)](*National Laboratory on Machine Perception*, *Beijing University*, *Beijing* 100871)

[3)](*Department of Computer Science and Technology*, *Nanjing University*, *Nanjing* 210093)

Abstract This paper firstly proposes a new motion estimation method MLBS (Modified Low Band Shift), and then puts forward an MLBS-based scalable video motion estimation scheme in the wavelet domain (MLBSSME). The modified low band shift method improves the accuracy of the motion estimation. And the "blocking artifacts" at the finest resolution are effectively reduced by overlapped block motion compensation technique. According to the multi-resolution representation characteristics of wavelet transform, the estimation process is limited in a search window of reduced size so that the search speed is accelerated. The proposed scheme can be effectively applied to spatial scalable and rate scalable video codec. The simulation results show that the MLBSSME scheme can always reach high motion estimation accuracy for standard test video sequences which have different characteristics respectively. And the proposed scheme can achieve 1～3dB higher prediction quality on average than the next-finer-resolution LL subband based hierarchical motion estimation algorithm and the band-to-band motion estimation algorithms, and 0.5～1dB higher prediction quality than LBS method when the spatial detail of the source video is low. Moreover, the time and space complexity of the proposed scheme is only 30%～40% as much as the complexity of LBS.

Keywords motion estimation; scalable coding; video coding; wavelet domain

收稿日期:2005-08-29;修改稿收到日期:2006-06-28. 本课题得到国家自然科学基金(60372071)、辽宁省自然科学基金(20032105)和辽宁省高等学校优秀人才支持计划(RC-04-11)资助. **宋传鸣**,男,1980 年生,博士研究生,主要研究方向为图像、视频编码及多媒体水印技术. E-mail: chmsong@163.com. **王相海**,男,1965 年生,博士,教授,主要研究领域为计算机图形学及多媒体信息处理.

1　引　言

用单个编码器产生分层次的压缩码流，对不同层次的码流解码可获得不同的视觉质量，这种编码机制称为可分级编码(scalable coding)[1,2]. 国际多媒体编码标准 MPEG-2 和 MPEG-4 便体现了这一可分级编码思想，即支持空间可分级编码、时间可分级编码和数率(或信噪比)可分级编码[3,4]. 研究者们提出了多种视频可分级编码器[5~11]，其中，细粒度可分级编码 FGS(Fine Granularity Scalability)方法以其高度灵活的编码技术、精确的码率控制和较高的编码效率，而为 MPEG-4 所采纳.

目前的视频编码器利用运动估计和补偿技术来降低视频数据中存在的时间冗余，运动估计和补偿性能的优劣直接影响着编码器效率的高低. 而 FGS 采用传统的基于宏块的空间域运动估计和补偿方法，则存在着以下不足：(1)基于块匹配的运动估计往往会在相邻块间引入不连续的运动向量，尤其在低码率情况，导致“块效应”甚为常见；(2)若相邻块的运动向量不一致，当对帧差进行编码时，会产生大的高频系数，从而降低小波编码的效率；(3)在位平面编码时，由于没有考虑到不同频带对人眼系统不同的重要性，将所有频带同等处理，以致在相同的码率下，其解码的主观质量必低于基于小波的编/解码器的解码质量；(4)空间域估计方法不具有对运动的渐进的、细粒度的表示能力，直接限制了 FGS 编解码的性能. 因此，基于宏块的空间域运动估计不适合渐进、细粒度编码的需求.

本文首先深入分析了基于下层 LL 子带分层运动估计方法[2]的优势和不足，进而提出了一种运动估计方法 MLBS(Modified Low Band Shift)；在此基础上，给出了基于 MLBS 的小波域视频可分级运动估计方案. 文章最后对所提出的运动估计方案进行了实验，给出了实验结果，并与相应的算法进行了比较.

2　小波域运动估计的传统方法

小波变换的多分辨率表示能力使得它能够提供对对象的内容及其运动的渐进表示，很容易实现适合于人眼视觉系统特性的视频可分级编码和渐进传输的需求. 目前，小波域运动估计方法可以归纳为三类：克服平移可变性的运动估计技术[12]、子带直接运动估计技术[13,14]和子带间接运动估计技术[2,15]. 其中，子带间接估计技术中的分层运动估计方案尤适于视频可分级编码. 有关小波域运动估计的研究进展可参看文献[16].

2.1　基于下层 LL 子带的分层运动估计方法

多尺度运动估计思想最初由 Cheng 等于 1995 年提出[15]，文献[2]对其系统框架做了进一步明确和调整，并提出了基于下层 LL 子带的分层运动估计方法，其基本思想是：从最低分辨率开始，利用参考帧的最低频子带对目标帧的最低频子带进行基于块匹配的运动估计. 目标帧的高频子带经如下过程预测：首先用参考帧对目标帧的下一层较高分辨率下的 LL 子带进行运动估计，并利用重叠块补偿技术得到目标帧的下层 LL 子带的预测，然后对该预测子带进行一级小波分解，进而获得次低分辨率下三个上层高频子带的预测. 迭代进行上述步骤，即可得到各个分辨率下的高频预测子带.

该方法不仅有效地避免了“预测环损失(loss of prediction loop)[15]”，而且，由于未对高频子带进行直接运动估计，在一定程度上缓解了小波移变性[12]对运动估计的影响. 另外，通过引入重叠块补偿技术，降低了解码图像的块效应；通过利用快速块匹配运动估计方法和预测向量的中心偏置特性，提高了运动向量的搜索速度. 但是，移变性的影响仍未从根本上得到解决.

2.2　低频子带平移(LBS)运动估计方法

文献[12]在分析了小波移变性的特征和本质之后指出，信号在空间域中平移任意一个整数位移后分解得到的所有子带信号，完全能由在频率域中平移原始信号低频子带的方法产生出来，并提出了一种低频子带平移的运动估计方法. 其基本思想是：将参考帧的低频子带(最高分辨率下即为原始帧)分别沿着水平、竖直和对角线方向平移一个像素，它们与原始的低频子带构成了分别对应着空间域中位移为(1,0)，(0,1)，(1,1)和(0,0)的 4 个参考帧. 对目标帧和 4 个参考帧进行一阶小波变换，最后利用子带直接运动估计计算运动向量，其中，对于空间域中位移为(dx,dy)的目标宏块，选择 4 个参考帧中对应着位移为$(dx\%2,dy\%2)$的帧作为参考. 对于 i 级小波变换，可以对各参考帧的第$(i-1)$级分解生成的各低频子带重复上面的平移和分解操作. 变换第 i 层各子带中的目标宏块，若其空间域位移为(dx,dy)，则选择参考帧中处于相同分解层次上对应于

位移为$(dx\%2^i, dy\%2^i)$的子带作为参考进行运动估计.另外,文中还提出了"小波块"结构,并将其作为运动估计的最小单位.若设第 k 个小波块的第 i 层子带的起始坐标为$(x_{i,k}, y_{i,k})$,待匹配的运动向量为(dx, dy),则其匹配准则为

$$SAD_k(dx,dy)=\sum_{i=1}^{3}\sum_{x_i=x_{i,k}}^{x_{i,k}+M/2^i-1}\sum_{y_i=y_{i,k}}^{y_{i,k}+N/2^i-1}\left\{\left|HL_{\mathrm{cur}}^{(i)}(x_i,y_i)-HL_{\mathrm{ref}}^{(i)}\left(dx\%2^i,dy\%2^i,x_i+\left\lfloor\frac{dx}{2^i}\right\rfloor,y_i+\left\lfloor\frac{dy}{2^i}\right\rfloor\right)\right|+\right.$$
$$\left|LH_{\mathrm{cur}}^{(i)}(x_i,y_i)-LH_{\mathrm{ref}}^{(i)}\left(dx\%2^i,dy\%2^i,x_i+\left\lfloor\frac{dx}{2^i}\right\rfloor,y_i+\left\lfloor\frac{dy}{2^i}\right\rfloor\right)\right|+$$
$$\left.\left|HH_{\mathrm{cur}}^{(i)}(x_i,y_i)-HH_{\mathrm{ref}}^{(i)}\left(dx\%2^i,dy\%2^i,x_i+\left\lfloor\frac{dx}{2^i}\right\rfloor,y_i+\left\lfloor\frac{dy}{2^i}\right\rfloor\right)\right|\right\}+$$
$$\sum_{x_3=x_{3,k}}^{x_{3,k}+M/2^3-1}\sum_{y_3=y_{3,k}}^{y_{3,k}+N/2^3-1}\left\{\left|LL_{\mathrm{cur}}^{(3)}(x_3,y_3)-LL_{\mathrm{ref}}^{(3)}\left(dx\%2^3,dy\%2^3,x_3+\left\lfloor\frac{dx}{2^3}\right\rfloor,y_3+\left\lfloor\frac{dy}{2^3}\right\rfloor\right)\right|\right\} \quad (1)$$

其中,M 和 N 分别为小波块的宽度和高度,下标 cur 表示目标帧,ref 表示参考帧,上标 i 表示子带所处的分解层数.目标帧各子带含有两个参数,用来确定子带中某个系数的位置,而参考帧各子带含有 4 个参数,前两个用来选择参考帧,后两个用来确定系数在参考帧中的位置.

3 小波域可分级运动估计方案

3.1 基于 MLBS 的运动估计方法

由 2.1 节的讨论可知,基于下层 LL 子带的分层运动估计方法的瓶颈在于低频子带估计的准确程度.若低频子带的运动向量估计得较准确,则能够得到很好的补偿效果,降低帧差的能量;反之,不准确的向量不仅会影响相应分辨率下运动补偿的效果,而且会从低分辨率向高分辨率漂移,以致影响最终的估计质量.因此,解决小波移变性的负面影响,对于提高运动估计的性能是非常关键的.而低频子带平移 LBS(Low Band Shift)技术是解决这一问题的有效方法,它能够重构在小波分解过程中因下采样而丢失的奇数相位,这样就将空间域中的平移与频率域中的相移建立一一对应关系,进而准确地估计出小波域内各个小波块的运动向量.因此,将分层运动估计方法与 LBS 方法相结合是解决前者不足的有效途径.但是,这种结合并不是简单组合,原因在于 LBS 方法所采用的小波块结构不适合可分级运动估计的需要.首先,小波块包含了不同分辨率下处于同一空间位置的信息,利用它计算出的运动向量只有在最高分辨率下才能获得最优的解码质量,而在其它分辨率下并不是最优的,这不仅会显著增加计算的时间复杂度,而且会降低可分级编码器的效率;其次,各个参考帧所有分辨率下的所有子带系数必须存储在缓冲区中以进行运动估计,所需要的空间是相当大的,而且是不必要的.

综上分析,本文在 3 个方面作了进一步的工作.(1)放弃采用小波块结构,而利用符合 HVS 特性的大小变化的宏块结构,并为每个分辨率下的各宏块计算出与其对应的最优运动向量,最大程度地提高可分级编码器的效率;(2)采用层次运动估计结构,在对某一子带进行估计时,只以各参考帧的同一分辨率下的子带为参考,能够降低算法的时间复杂度,而且这样也将大大提高编码缓冲区的利用率,减少算法的空间开销;(3)运动估计时仅利用各分辨率下的低频子带,与高频子带相比,它是图像帧的更近似且更稳定的表示,有助于提高运动向量的准确度.进而,提出了一种改进的 LBS 方法 MLBS(Modified LBS).

设目标帧和参考帧分别施以 $L(L>0)$阶小波变换,则对其中第 $i(0<i\leq L)$层子带的宏块进行运动估计的具体过程如下:

1. 将目标帧进行小波逆变换,得到第 i 层 LL 子带;

2. 计算得出原始参考帧第 i 层 LL 子带,同时按照 2.2 节的平移方式,生成对应于不同空间域位移的参考帧.不同的是,对于所生成的各个参考帧,只保留其低频子带;

3. 将目标帧和参考帧的第 i 层 LL 子带划分成大小为$(p\times 2^{L-i})\times(q\times 2^{L-i})$的不重叠的宏块,$p\times q$ 为最高层子带宏块的大小;

4. 对于 LL 子带中的各个宏块,以$(i+1)$层 LL 子带相应位置宏块的运动向量为中心($i=L$ 除外),在大小为 $W\times W$ 的搜索窗口内,采用某种块匹配运动估计算法(如全搜索、菱形搜索法等),求出其运动向量.搜索过程所采用的匹配准则定义如下:设经过小波分解后,目标帧的第 i 层 LL 子带为 $LL_{\mathrm{cur}}^{(i)}$,参考帧的第 i 层 LL 子带为 $LL_{\mathrm{ref}}^{(i)}$,并设目标帧第 i 层 LL 子带内第 k 个宏块的起始坐标为$(x_{i,k}, y_{i,k})$,待匹配

的运动向量为(dx,dy)，则有

$$SAD_k(dx,dy)=\sum_{x_i=x_{i,k}}^{x_{i,k}+M/2^i-1}\sum_{y_i=y_{i,k}}^{y_{i,k}+N/2^i-1}\left\{\left|LL_{\mathrm{cur}}^{(i)}(x_i,y_i)-LL_{\mathrm{ref}}^{(i)}\left(dx\%2^i,dy\%2^i,x_i+\left\lfloor\frac{dx}{2^i}\right\rfloor,y_i+\left\lfloor\frac{dy}{2^i}\right\rfloor\right)\right|\right\}$$

并确定宏块 k 的最优的运动向量$(dx_{\min},dy_{\min})_k$为

$$(dx_{\min},dy_{\min})_k=\min_{(dx,dy)\in W}SAD_k(dx,dy),$$

其中，M 和 N 分别表示最高分辨率下宏块的宽度和高度，W 表示搜索窗口，其它符号约定与式(1)相同；

5. 对于参考帧第 i 层高频子带中的各个宏块，首先按照步 4 所述方法计算出参考帧第$(i-1)$层 LL 子带各宏块的运动向量，然后利用运动补偿技术得到该层 LL 子带的预测，最后将该预测 LL 子带进行一阶小波分解，进而获得参考帧第 i 层高频子带的预测.

3.2　基于 MLBS 的可分级运动估计方案

在 3.1 节提出的 MLBS 方法的基础上，本文提出的可分级运动估计方案 MLBSSME（Modified Low Band Shift based Scalable Motion Estimation）如图 1 所示.

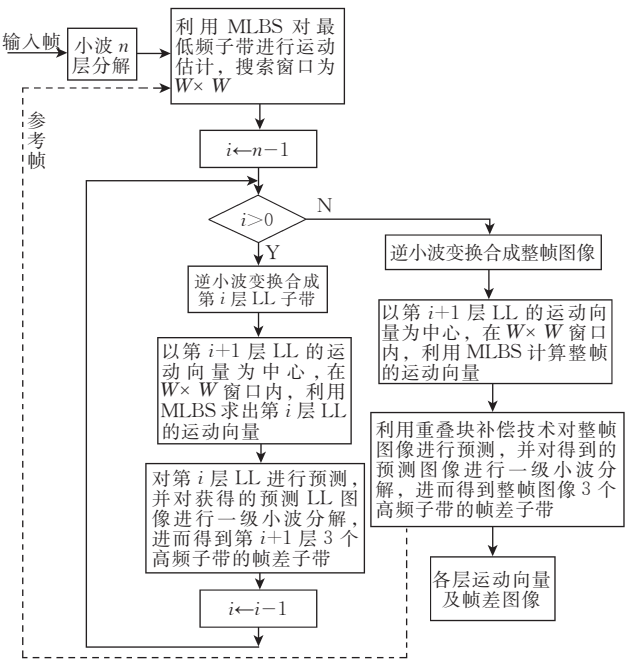


图 1　基于 MLBS 的可分级运动估计方案流程图

在该运动估计过程中，采用了 3 种方法进一步降低运动估计所耗费的时间和空间：(1)只在最高分辨率下的运动补偿采用重叠块技术，而较低分辨率下，由于小波变换及其逆变换本身就是一种全局滤波的过程，可以对各宏块的边缘进行平滑. 这样，既能有效地消除块效应，又有利于提高估计的速度；(2)利用小波变换的多分辨率特性以及人眼视觉系统的特点，在不同分辨率下，采用变块大小的运动估计(Variable Block-Size ME)，这种大小变化的宏块结构以及相同大小的搜索窗口，相对传统块匹配估计方法进一步节省了运算开销. 并且，开销的减少是在保证估计质量的前提下完成的，利用 MLBS 方法实现了低分辨率下运动向量的精确刻画，进而抑制了较高分辨率下错误运动向量的漂移；(3)高频子带由上层低频子带分解生成，省去了运动估计和补偿的过程，既能进一步节省时间和平移参考帧的空间需求，又能够解决高频子带估计不准确的问题. 因此，所提出的运动估计方案不仅能够显著提高基于下层 LL 子带分层运动估计方法的准确程度，也能有效地保证算法具有较低的时空复杂度.

4　实验结果与分析

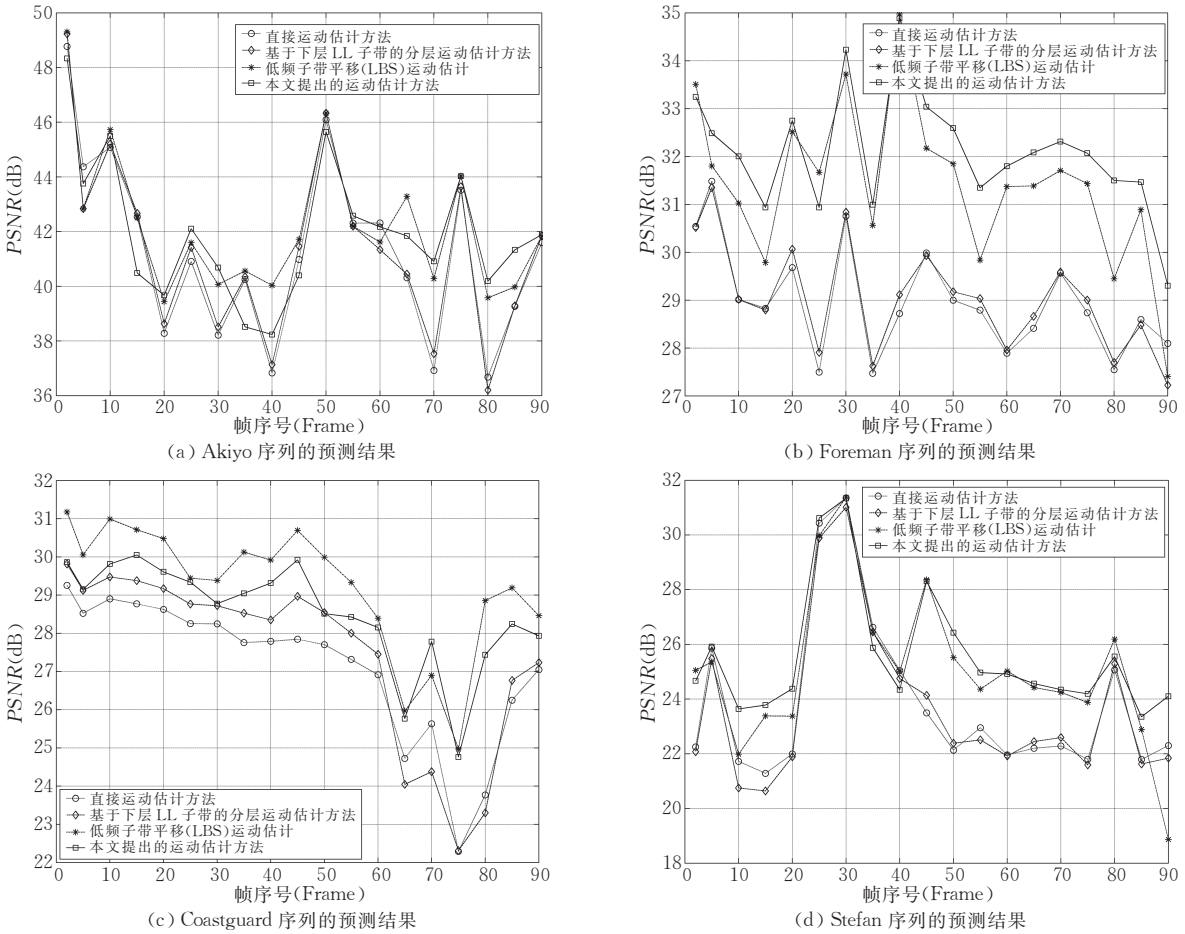
为了验证本文所提出的算法的有效性，分别以 CIF 格式的标准视频流 Akiyo、Foreman、Coastguard 和 Stefan 的 1～90 帧为例，采用 3 级提升方案 5/3 小波，进行了大量的实验，重构图像帧的质量用峰值信噪比进行衡量. 对于每一帧图像，将其前一帧作为参考帧，用 3.2 节的方法进行运动估计，并用参考帧中的最佳匹配块直接替代目标帧中的每个块，从而得到目标帧的预测.

为了验证算法的可分级特性，将 Akiyo 序列的第 1 帧作为参考帧，在空间分辨率分别为 1，1/2，1/4 和 1/8 时对第 2 帧进行了运动估计，得到的相应分辨率下的预测帧如图 2 所示，最高分辨率下的 $PSNR$ 为 48.33dB.

图 2　Akiyo CIF 序列的第 2 帧分别在其空间分辨率为 1，1/2，1/4 和 1/8 时的预测图像

为了验证所提出算法的运动估计性能，将实验结果与 3 种运动估计方法加以比较，其中第 1 种方

法是文献[13]提出的直接运动估计方法中最为精细的一种，首先计算最低分辨率下 4 个子带的运动向量，其它分辨率下各子带的运动向量以此为中心，进一步求精；第 2 种方法是文献[2]提出的基于下层 LL 子带的分层运动估计；第 3 种方法是文献[12]提出的低频子带平移(LBS)运动估计. 实验中，确定 $p=2$，$q=2$，$n=3$，$W=4$，在搜索窗口内对每个候选运动向量采用全搜索. 图 3(a)～(d)所示分别为 4 个测试序列使用上述 4 种运动估计方法在最高分辨率下得到的预测帧 $PSNR$ 的比较结果.



(a) Akiyo 序列的预测结果

(b) Foreman 序列的预测结果

(c) Coastguard 序列的预测结果

(d) Stefan 序列的预测结果

图 3 本文提出的运动估计方法与其它 3 种方法的预测结果的比较

在 MPEG-4 校验模型[17]的标准测试视频库中，Akiyo、Foreman、Coastguard 和 Stefan 这 4 个序列分别代表着具有不同空间和时间特性的视频流，其中，Akiyo 代表空间细节和运动复杂性都很低的一类视频；Foreman 和 Coastguard 代表空间细节复杂度中等而运动复杂性较低的一类视频. 但是相比较而言，后者的空间细节要比前者复杂，其复杂程度介于 Foreman 和 Stefan 之间；Stefan 代表空间细节复杂度较高而运动复杂性中等的一类视频. 对比研究图 3(a)～(d)就不难发现，尽管对于不同类型的视频流，各运动估计方法所获得的预测效果不尽相同，可是其性能优劣的比较仍是很明显的，本文提出的运动估计方法始终保持着很高的准确度，并且还可以看出：(1)基于下层 LL 子带的分层运动估计方法和本文提出的估计方法，尤其是后者，其估计效果均优于直接运动估计，并且随着运动复杂性的提高，这种优势则变得较为显著，这正是由于两个方法有效地缓解了小波移变性的缘故；(2)本文在基于下层 LL 子带的分层运动估计方法的基础上，利用 MLBS 方法和重叠块补偿技术提高了低频子带运动估计的准确度，预测帧的信噪比较之第 2 种方法有了进一步的提高，平均要高出 1～3dB；(3)在预测过程中，本文的方法能够获得最高的或者近似最高的 $PSNR$，说明利用该方法得到的补偿帧差的能量是最小的或者近似最小的，因此将它应用于可分级编码，并不会因兼顾可分级性而降低编码器的效率，甚至更加有利于提高编码效率；(4)本文方法与 LBS 方法相比，二者的平均性能相当. 对于空间细节较复杂的视频，LBS 方法性能往往略胜一筹，对于空间细节较简单的视频，用本文方法得到的预测帧的

PSNR 比 LBS 方法平均要高 0.5～1dB.

本文方法在时空复杂度方面也表现出了良好的性能，表 1 给出了所提出的方法与 3 种传统小波域运动估计方法在时空复杂度方面的比较结果. 这 4 种方法均运行在赛扬 1.7GHz CPU，256MB DDR RAM，Windows XP 操作系统环境下，实现语言为 Visual C++ 6.0. 由表 1 可以看到，本文算法在时空复杂度方面要较 LBS 方法好，但不及其它两种方法.

表 1　本文所提出的运动估计方法与 3 种传统小波域运动估计方法在时间开销和空间需求方面的比较

	时间开销(ms)	空间需求
直接运动估计方法	651	1 帧
基于下层 LL 子带的分层运动估计方法	381	1 帧
低频子带平移 LBS 方法	7261	10 帧
本文的 MLBSSME 方法	2343	4 帧

5　结　论

本文首先对基于下层 LL 子带的分层运动估计方法进行了深入的分析，针对其不足提出了一种新的运动估计方法 MLBS (Modified Low Band Shift)，并在此基础上，提出了一种基于 MLBS 方法的小波域视频可分级运动估计方案，在保持了原有方法优点的前提下，进一步提高了运动估计的准确度，其重构图像帧的 *PSNR* 平均提高了 1～3dB，并且预测质量明显优于直接运动估计方法. 与 LBS 方法相比，本文算法的预测质量与之相当，但空间复杂度仅是 LBS 的 40%左右，时间复杂度约是 LBS 方法的 30%左右. 对于各种类型的视频序列，该算法能够始终保持很高的估计准确度. 实验结果验证了方法的有效性.

参　考　文　献

1　Wang Xiang-Hai. Research on scalable image and video coding based on wavelet[Post-doctoral research report]. Nanjing University, Nanjing, 2001(in Chinese)
(王相海. 基于小波的图像和视频可分级编码研究[南京大学博士后研究工作报告]. 南京大学，南京，2001)

2　Wang Xiang-Hai. Research on spatial scalable video coding based on wavelet. Computer Science, 2002, 29(7): 115～117 (in Chinese)
(王相海. 基于小波的视频空间可分级编码研究. 计算机科学，2002，29(7)：115～117)

3　ISO/IEC JTC1/SC29/WG11. Information technology-generic coding of moving pictures and associated audio information, part 2: visual, ISO/IEC 13818-2, 1996

4　ISO/IEC JTC1/SC29/WG11 N3536. Overview of the MPEG-4 standard, 2000

5　Su Yih-Ching. Fine granularity video compression technique and its application to robust video transmission over wireless Internet[Ph. D. dissertation]. National Sun Yat-Sen University, Taibei, China, 2003

6　Zhang Qian, Guo Qu-Ji, Ni Qiang, Zhu Wen-Wu, Zhang Ya-Qin. Sender-adaptive and receiver-driven layered multicast for scalable video over the Internet. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(4): 482～495

7　Taubman D., Zakhor A.. Multi-rate 3-D subband coding of video. IEEE Transactions on Image Processing, 1994, 3(9): 572～588

8　Kim B., Xiong Z., Pearlman W. A.. Low bit-rate scalable video coding with 3D set partitioning in hierarchical trees (3D SPIHT). IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(12): 1374～1387

9　Li Wei-Ping. Overview of fine granularity scalability in MPEG-4 video standard. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(3): 301～317

10　Wu Feng, Li Shi-Peng, Zhang Ya-Qin. A framework for efficient progressive fine granularity scalable video coding. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(3): 332～344

11　Secker A., Taubman D.. Highly scalable video compression using a lifting-based 3D wavelet transform with deformable mesh motion compensation. In: Proceedings of the IEEE International Conference on Image Processing, Sydney, Australia, 2002, III: 749～752

12　Park Hyun-Wook, Kim H. S.. Motion estimation using low-band-shift method for wavelet-based moving-picture coding. IEEE Transactions on Image Processing, 2000, 9(4): 577～587

13　Zhang Ya-Qin, Zafar Sohail. Motion-compensated wavelet transform coding for color video compression. IEEE Transactions on Circuits and Systems for Video Technology, 1992, 2(3): 285～296

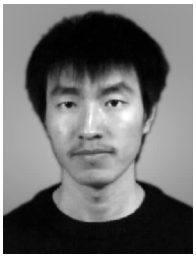
14　Kim S., Rhee S., Jeon J. G., Park K. T.. Interframe coding using two-stage variable block-size multiresolution motion estimation and wavelet decomposition. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(4): 399～410

15　Cheng Po-Y., Li J., Kuo C.-C. J.. Multiscale video compression using wavelet transform and motion compensation. In: Proceedings of the International Conference on Image Processing (ICIP'95), Washington, DC, 1995, 606～609

16　ISO/IEC JTC1/SC29/WG11 N3908. MPEG-4 video verification model, version 18.0, 2001

17　Song Chuan-Ming, Wang Xiang-Hai. Research progress in video motion estimation techniques in the wavelet domain. Chinese

Journal of Computers, 2005, 28(10): 1716～1727(in Chinese)
(宋传鸣,王相海.小波域视频运动估计研究进展.计算机学报, 2005, 28(10): 1716～1727)



SONG Chuan-Ming, born in 1980, Ph. D. candidate. His research interests include image and video compression, digital watermarking.

WANG Xiang-Hai, born in 1965, Ph. D., professor. His research interests include computer graphics and multimedia information processing.

Background

Fine granular scalable video coding is a difficulty and research hotspot in the state-of-art of video coding discipline. So far many effective schemes, such as FGS, PFGS, PFGST and FGSS, have been proposed. The authors are now presiding over the project "Research on Object-based Fine Granular Scalable Video Coding" supported by the National Natural Science Foundation of China under grant No. 60372071. The aim of the project is to propose a fine granular spatial scalable, rate scalable and hybrid spatial and rate scalable video object coding scheme based on the 2-D shape adaptive wavelet and its motion compensation so as to break the limitations of traditional hierarchical scalable coding scheme and the limitations imposed on the object-based video coding scheme.

Motion estimation in the wavelet domain is a key technique of the project. It influences the output video quality and determines the success of the scalable codec. The authors sum up the state-of-art techniques as three kinds, namely, band to band motion estimation techniques, indirect band motion estimation techniques and shift-variant resistant motion estimation techniques. The paper by the authors "Research Progress in Video Motion Estimation Techniques in the Wavelet Domain" has been published in "Chinese Journal of Computers". In this paper, combining the hierarchical structure and modified low band shift technique, an effective scalable motion estimation scheme is proposed.