第 28 卷 第 10 期
2005 年 10 月
计 算 机 学 报
CHINESE JOURNAL OF COMPUTERS
Vol. 28 No. 10
Oct. 2005

小波域视频运动估计研究进展

宋传鸣[1)] 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)](北京大学信息科学中心视觉与听觉信息处理国家重点实验室 北京 100871)

摘 要 小波域运动估计技术作为基于小波的视频编码器中最关键的技术之一,成为近年来研究的热点.文章首先对比论述了空间域运动估计技术的不足和小波域运动估计技术的优势,然后以小波域运动估计技术的发展为主线,将目前的研究方法归纳为三类:对子带直接运动估计技术、子带间接运动估计技术和克服平移可变性的运动估计技术,并对三类运动估计方法的发展进行了概括性阐述和比较性研究,对三类方法中各种算法的适用范围、优势和所存在的问题进行了深入的讨论,在此基础上对小波域运动估计技术的未来发展趋势进行了展望.

关键词 运动估计;小波域;视频编码;子带

中图法分类号 TP391

Research Progress in Video Motion Estimation Techniques in the Wavelet Domain

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

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

[2)](*National Laboratory on Machine Perception, Center of Information Science for Peking University, Beijing* 100871)

Abstract In recent decade, the fact, that wavelet transform has gained a great success in the field of still image coding, moreover, it has been accepted as a powerful tool by the JPEG2000 standard, has attracted numerous researchers' attention in video coding field, so that they devoted themselves one after another to the research of wavelet-based video coding techniques. As one of the key techniques of video encoder, motion estimation technique in the wavelet domain has become a research hotspot in recent years. This paper firstly discusses the deficiency of motion estimation techniques in the spatial domain and the advantages of the ones in the wavelet domain. And secondly takes the development of the motion estimation techniques in the wavelet domain as the thread to sum up the present techniques as three kinds, namely, band to band motion estimation techniques, indirect band motion estimation techniques and shift-variant resistant motion estimation techniques, and then make a general discussion and comparative study on the three kinds of methods, which carries on a deep discussion on the scope of application, advantages and disadvantages of every method of each kind of techniques. Finally, at the basis of above, the state of arts and possible new directions of motion estimation techniques in the wavelet domain are stated.

Keywords motion estimation; wavelet domain; video coding; subband

1 引 言

传统的视频编码标准MPEG-1、MPEG-2、ITU-T H.261、ITU-T H.263以及尚在不断完善中的MPEG-4标准[1],均采用基于块的空间域运动估计技术来减少视频帧的时间冗余.近十年来,在静态图像编码领域的应用中,小波变换异军突起,获得了巨大的成

收稿日期:2004-09-01;修改稿收到日期:2005-04-15.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资助.宋传鸣,男,1980 年生,硕士研究生,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.

功,并已经被纳入 JPEG2000 标准[2],取代了 DCT 变换,成为分析和处理自然图像信号的强有力的工具.这也吸引了视频编码领域众多专家学者的目光,他们纷纷致力于基于小波的视频编码技术的研究之中.作为视频编码最关键的环节之一,基于小波的运动估计技术则倍受人们的重视和关注,成为近年来研究的热点.

目前,基于小波的视频编码算法主要分为 3 类[3]:(1)在空间域进行运动估计和补偿,并对预测误差采用小波图像编码[4~8];(2)在小波域进行运动估计和补偿;(3)利用 3D 小波变换进行视频编码[9~13].Zhang 和 Zafar 在文献[14]中通过实验证明了第二类方法可以获得较之第一类方法更好的编解码效果,而第三类方法则具有编码延迟大、占用缓冲区大、对场景切换和快速运动较为敏感等诸多不足.因此,人们将大部分的精力都投入到了小波域运动估计技术的研究中,并相继提出了多种有效的估计方法.本文亦将重点集中于此,即以小波域运动估计技术的发展为主线,将目前的研究方法归纳为对子带直接运动估计技术、子带间接运动估计技术和克服平移可变性的运动估计技术三类方法,并对其发展进行了概括性阐述和比较性研究,同时对各类方法中各种算法的适用范围、优势和所存在的问题进行了深入的讨论.文中最后对小波域运动估计技术的未来发展趋势进行了展望.

文章第 2 节分别对空间域运动估计和小波域运动估计技术进行概述,并对前者的不足和后者的优势加以比较;第 3 节阐述了多种小波域运动估计方法的核心思想,并进行了相应的分析和讨论;第 4 节将对小波域运动估计技术的发展趋势进行分析和展望;最后总结全文.

2　空间域和小波域运动估计技术的比较

2.1　空间域运动估计技术概述

自视频编码技术出现以来,视频编码器均采用空间域内的运动估计和补偿技术减少帧间冗余,其中基于块匹配的运动估计技术因其在估计的精度和复杂度之间提供了很好的折衷,而应用得尤为广泛.在此之中,最简单的方法莫过于全搜索(FS).尽管它能够找到全局最佳匹配块,且易于实现,但是巨大的时间复杂度使得许多实时性要求较高的应用对之望而却步.因此,人们进行了积极的研究,提出了很多次最优快速搜索算法,诸如:二维对数法(TDL)[15]、三步法(TSS)[16]、对偶搜索法(CDS)[17]、交叉法(CSA)[18]、新三步法(NTSS)[19]、四步法(FSS)[20]、基于块的梯度下降搜索法(BBGDS)[21]、菱形搜索法(DS)[22]和边界辅助搜索法(EAS)[23]等.值得一提的是,新三步法首次提出了自然视频运动向量的中心偏置特性和中止判别技术,为后继各高效算法的提出奠定了很好的基础;菱形搜索法凭借其良好的搜索性能和整齐划一的处理步骤而为 MPEG-4 标准校验模型所接受.后来,人们进一步研究了相邻块间运动向量的关系,提出了 FS 的改进算法[24,25]等;针对 MPEG-4 基于对象的机制,又提出了任意形状视频对象的运动估计方法[26,27]等.算法虽种种,但其思想却基本相同,即放弃对不必要的匹配点的搜索,以尽可能降低搜索的时间复杂度,并尽可能减少因此导致的估计精度的下降.

2.2　小波域运动估计技术概述

小波域运动估计技术是近年来小波应用于编码领域的一个新兴的研究方向.自 1992 年 Zhang 和 Zafar 在文献[14]中首次提出该技术以来,经过数据编码、信号处理、应用数学等多领域的科研人员的不懈努力,出现了许多优秀的运动估计算法.我们将这些算法总体分为三类,即子带直接估计算法(Band-to-Band Motion Estimation)、子带间接估计算法(Indirect Band Motion Estimation)和克服平移可变性的运动估计算法(Shift-variant Resistant Motion Estimation).我们将在第 3 节对各类方法的发展情况进行分析和讨论.

2.3　空间域和小波域运动估计技术的比较

2.3.1　空间域运动估计技术的不足

尽管空间域运动估计技术被多个国际编码标准所采用,但它仍然存在着一些难以解决的问题:(1)估计的运算量大;(2)基于块匹配的运动估计方法往往在相邻块间引入不连续的运动向量,尤其在低码率的应用中,导致“块效应”甚为常见,从而引起人眼最为敏感的可视的降质,严重影响了编码的质量;(3)若相邻块的运动向量不一致,当对帧差进行小波图像编码时,就会产生大的高频系数,这将降低小波编码的效率;(4)Li 等人提出的基于空间域运动估计和 DCT 变换的细粒度编码方法 FGS[28,29],除了依然避免不了明显的“块效应”外,空间域估计方法亦不具有对运动的渐进的、细粒度的表示能力,这直接限制了 FGS 的细粒度编解码的性能.并且,在位平面编码时由于没有考虑到不同频带对于人眼

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. 因此，空间域运动估计不适合需求日益迫切的渐进、细粒度编码的需求.

2.3.2 小波域运动估计技术的优势

基于小波域的运动估计技术恰恰可以克服上述空间域运动估计技术的 4 点不足：(1)基于小波域的运动估计采用相对缩小了的搜索窗口和匹配块，有效降低了计算量；(2)小波变换是全局变换，它排除了“块效应”的困扰，为解码图像提供了更高的主观质量；(3)小波变换能够提供对信号的时频局部化的多分辨率和多频率表示，将一个非平稳信号分解成一组相对稳定的多尺度小波表示，从而易于编码；(4)小波的多分辨率表示能力使得它能够提供对对象和运动的渐进表示，因此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. 正是基于这些原因，小波域运动估计技术越来越受到重视.

3 小波域运动估计技术的研究进展

3.1 子带直接运动估计技术的研究进展

这类方法的出现最初源于多分辨率运动估计(Multi-Resolution Motion Estimation，MRME)技术，由 Metin Uz 等于 1991 年提出[30]. 其基本思想是认为典型的场景在任意尺度下都包含有运动，在较低的空间频率下，可以得到低精度的补偿帧差；在较高的频率下，则可以获得高分辨率的补偿帧差. 因此，采用一种由粗到细的策略，将图像帧表示成一个分辨率由上至下逐渐增大的金字塔，从最低分辨率图像开始，利用块匹配的运动估计方法，计算出粗略的运动向量；当达到较高分辨率时，再求精前一步中的运动向量，从而获得更为精细、准确的运动向量.

该方法首次将多分辨率分析引入运动估计中，符合人眼视觉特性，能够获得更高的主观质量，为小波域运动估计方法的提出奠定了基础. 并且，算法将块匹配运动估计的时间复杂度由平方级降至对数级，有效节省了运算量. 但是该项技术存在着以下不足：首先由于在各分辨率下均采用相同的分块大小，而导致忽略了较低分辨率下小对象的运动，甚至是大的运动，直接影响了估计的准确性(如图 1)；其次，处于金字塔中较低层的图像初始值是由高层图像的插值和求精得到的，容易产生较大的误差漂移；此外所采用的双线性插值的计算量也是一笔较大的开销.

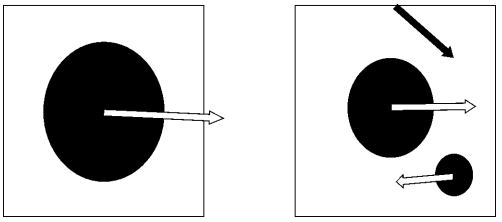


图 1 同样大小的宏块在不同分辨率下包含的物体(两个图中的大圆为同一个物体，其中右侧图中的黑色箭头表示宏块的运动向量。从图中可看出，小圆的运动向量因为大圆的运动而被忽略或改变了)

针对以上问题，Zhang 和 Zafar 于 1992 年提出了可变块大小的多分辨率运动估计算法[14]，其基本思想是：处于小波分解金字塔(如图 2)中不同层次上的运动是不同的，但它们却高度相关，它们实际上刻画了不同尺度和不同频率范围下的同一个运动结构. 首先将第 m 层的各个子带划分成大小为 $(p\cdot 2^{M-m})\times(q\cdot 2^{M-m})$ 的不重叠的宏块，其中 M 为小波分解的阶数，$p\times q$ 为最高层子带的宏块大小. 然后利用传统的块匹配运动估计算法(如 FS)计算最低分辨率下各宏块的运动向量；最后，较高分辨率下各宏块的运动向量，由处于低分辨率下同一空间位置的宏块的运动向量预测并求精. 约定用 $\boldsymbol{V}_{i,j}(x,y)$ 表示子带 $W_j^i(i=1,2,3,j=2,4,8)$ 中以 (x,y) 为中心的宏块的运动向量，则上述的估计算法可以表示为 $\boldsymbol{V}_{i,j}(x,y)=2\boldsymbol{V}_{i,2j}(x,y)+\boldsymbol{\Delta}(\delta x,\delta y)$，其中 $\boldsymbol{\Delta}(\delta x,\delta y)$ 为

$$\arg\min_{\delta x,\delta y\in\Omega}\left[\frac{1}{XY}\sum_{p=-X/2}^{X/2}\sum_{q=-Y/2}^{Y/2}\left|I_i(x_1+p,y_1+q)-I_{i-1}(x_1+p+x+\delta x,y_1+q+y+\delta y)\right|\right],$$

这里，Ω 表示搜索窗口，X 和 Y 分别表示宏块的长和宽，(x_1,y_1) 为宏块的中心坐标，$I_i(x,y)$ 表示第 i 帧中 (x,y) 点处的灰度值，图 3 为这一过程的图示.

在上述思想的基础上，文献[14]和[31]提出了 4 种运动估计方案：(1)仅计算最低频子带的运动向量，并将其扩大 2^{M-m} 倍即得到第 m 层子带的运动向量；(2)仅计算最低分辨率下 LL_M，LH_M，HL_M 和 HH_M 4 个子带的运动向量，第 m 层子带的运动向量由 M 层对应方向子带的运动向量扩大 2^{M-m} 倍确定；(3)计算最低频子带的运动向量，并将此向量扩大 2^{M-m} 倍，作为第 m 层子带运动向量的初始值，然后进一步求精，求出相应的 $\boldsymbol{\Delta}(\delta x,\delta y)$；(4)计算最低分辨率下 4 个子带的运动向量，第 m 层子带以 M 层对应方向子带的运动向量的 2^{M-m} 倍为初始值，然后进一步求精. 实验证明，方案(4)估计和补偿的效

果最好，方案(1)最差.

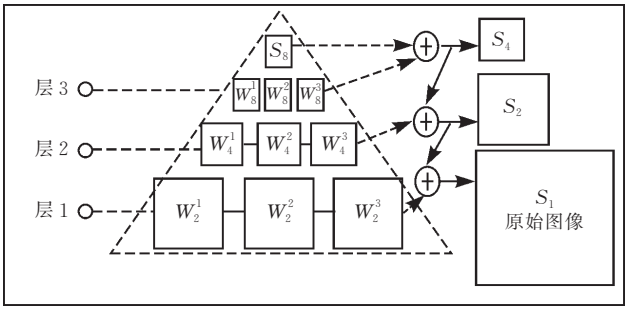


图 2　小波分解和重构的金字塔示意图

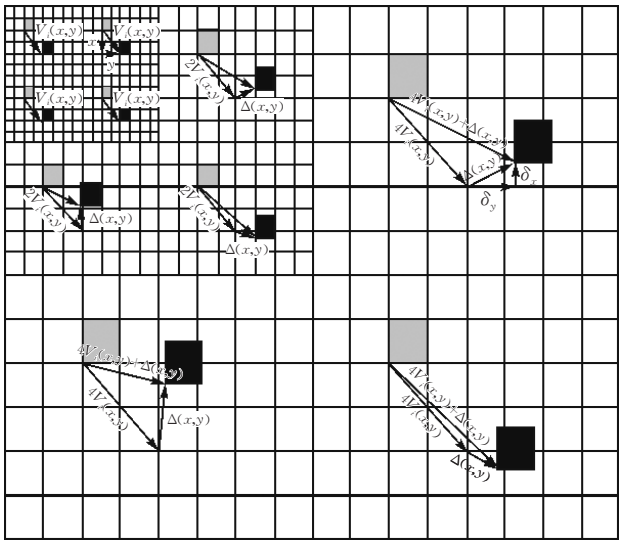


图 3　变块大小的多分辨率运动估计示意图

上述算法充分利用了小波分解良好的时频特性，使用大小变化的宏块结构，不仅所有子图中宏块的数量是一致的，也恰当地衡量了不同分辨率下运动的重要性，符合人类对不同分辨率下不同频率的感知特性，可以检测出最低分辨率下小对象的运动；此外，当求精每一分辨率下各宏块的运动向量时，不需要任何插值运算，大大减少了计算量. 并且，不同分辨率下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；还有该算法实现了低分辨率下运动的精确刻画，使得帧差较小，抑制了较高分辨率下错误运动向量的漂移.

然而，算法所存在的问题亦是显而易见的，表现在以下几个方面：(1)最低频子带的运动向量并不能保证其对于高分辨率子带亦是最优的，因为在运动估计时，忽略了高分辨率下各子带的以该向量为基础的运动向量进行运动补偿时可能产生的影响；(2)如果最低频子带的运动向量估计不准确，势必影响其它高频子带运动向量的估计，对重构帧造成不可逆转的破坏；(3)最低频子带的宏块过小，容易产生“孔径问题(aperture problem)”[32]，易得出错误的运动向量；(4)文献[31]中还尝试了小波域 1/2 像素和 1/4 像素精度的块匹配运动估计，但未获得比整数像素精度更高的重构质量.

针对上述问题，人们在以下几方面展开研究：

(1)针对第一方面的不足，文献[33]中提及了三种较上述四种方案更加精细的估计方法，但是，时间复杂度却较大；文献[34]则提出将小波变换后同一分辨率下各方向子图 HL_j，LH_j 和 HH_j 结合成全方向子图 $W_j^A(j=2,4,8)$，然后运动估计在目标帧和参考帧的各 W_j^A 之间进行，同分辨率下各子图具有相同的运动向量，据此，文中提出了快速估计方法(Fast MRME).

这种方法在运动估计时兼顾了各高频子图，能得到一个整体最优的运动向量，减少了搜索时间，也将编码每一帧所需要的运动向量的个数降至传统 MRME 的 40%左右. 但是，由于对高频子带的运动估计没有利用低频子带的估计结果，二者得出的向量有时并不一致，甚至是矛盾的. 据此，文献[35,36]中提出了一种充分利用低频子带的方法(New MRME)，每一分辨率下各高频子带的运动向量都以同分辨率的低频子带的估计结果作为初始值，再进一步求精，从而有效提高了重构帧的 PSNR.

文献[33～36]仅兼顾了同分辨率下的各子带，而文献[37]受文献[38]的启发，兼顾了所有分辨率下同方向子带的关系，提出了新的运动估计匹配准则 MSAD，基本思想是确定满足同一方向所有子带的 SAD 获得最小值时所对应的向量为运动向量. 这种方法在没有提高时间复杂度的情况下，获得了比传统 MRME 更高的 PSNR. 文献[39]则进一步发扬了这一思想，兼顾同一空间位置对应的所有分辨率的所有子带，在目标帧中以小波变换后形成的小波树为搜索对象，到参考帧中搜索与之最匹配的小波树，此时的向量即为小波树中所有系数的运动向量. 由于每棵树一个像素的平移相当于空域中 2^n (n 为小波变换的阶数)个像素的平移，故搜索区域只需限制在参考帧中少数几棵树上. 因此，这种方法简单、快速，尤适合于低码率、高实时性的应用. 由于估计过于粗糙，重构帧的 PSNR 不高.

(2)针对第二方面的不足，有效抑制低分辨率下错误向量的漂移，文献[40,41]提出了借助于中值滤波的运动估计算法(MV Prediction Through Median Filtering，MF)，其基本思想是：大部分自然场景

视频的运动都是平滑、连续的，一个错误的运动向量往往表现为与邻近块的运动向量不一致、不连续，这样，可以利用相邻块的运动向量来纠正错误的运动向量. 而纠正的方法就是将目标块的父窗口中各宏块的运动向量组成候选向量集合（如图 4 所示），对集合中各向量的水平和竖直分量分别进行中值滤波，最后将滤波后的两个结果重新组合成一个向量，适当扩大后即作为目标块的初始运动向量. 文献[41]中又提出了多候选向量的运动估计方法（Multi-Candidate MV Prediction，MC），即在父窗口中找出一个能获得最小帧差的向量，将其作为目标块的初始运动向量.

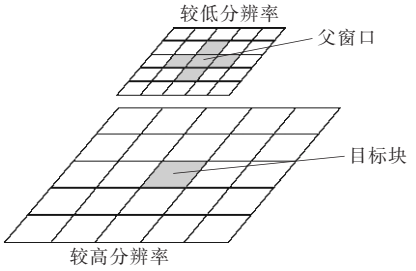


图 4 父窗口示意图

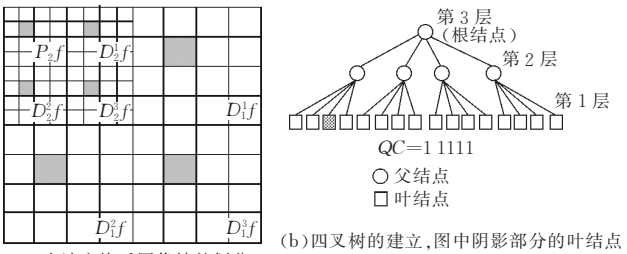
上述 MF 在未明显提高复杂度的情况下，能够克服错误的漂移，与传统 MRME 相比，能产生更加平滑的运动场，将平均帧差缩小 6%，达到更好的估计质量，提高了编码效率. 但是，中值滤波对高分辨率下高频子带非平滑的运动具有较大的负面影响，尤其是边缘块的高频子带，导致降低了估计的准确性，还会影响重构帧的主观质量. 而上述的 MC 方法除需要传输标识选择了某个向量的边信息，重构帧的质量亦劣于传统 MRME.

(3)针对第三方面的不足，文献[32]提出了两步估计法，具体思想是：在估计最低频子带的运动向量时，第 1 步，利用 4 个相邻宏块组成一个大宏块进行运动估计，以缓解“孔径问题”；第 2 步，每个小宏块在本身及相邻的 8 个宏块的运动向量中选择一个最优的作为运动向量，以减轻“块效应”，并以此作为其它高频子带运动向量的初始值. 算法通过提高对低频子带估计的准确度，有效地提高重构帧的 PSNR. 另外，在对高频子带的运动向量进行求精前，算法引入了误差阈值，只有当帧差大于阈值时，才进行求精操作. 这有助于降低算法的时间复杂度，却也相应增大了帧差.

(4)针对第四方面的不足，文献[42]对小波域系数的插值算法进行了理论推导和深入研究，提出了半像素精度的运动估计方法（Half-pixel MRME，HMRME），有效提高了运动估计的准确度.

此外，文献[43～46]在传统 MRME 的基础上提出了新的技术：(1)文献[43]指出，尽管相邻两视频帧很相似，但是其高频子带的相似性却较差，于是提出了用来估计高频子带运动向量的自适应阈值技术（adaptive thresholding MRME）. 其基本思想是：当目标块与参考帧中最佳匹配块的相似度低于一定阈值时，就不编码该块对应的运动向量，并在解码端将块初始化为零块（zero block）. AMRME 一方面可以减少编码运动向量和帧差的比特位，另一方面也有利于提高重构帧的主观质量. 但是，阈值的判断是在运动估计之后才进行的，因此额外增加了 MRME 的时间复杂度；(2)文献[44，45]借鉴了空间域快速估计算法的思想，利用连续两帧最低频子带中对应宏块的差与预设阈值进行比较，确定可能发生了运动的宏块，并将结果由上至下逐层传播，直到判断出所有分辨率下所有可能的运动宏块，最后只对找到的运动宏块进行运动估计. 算法显著降低了运算开销，是对传统 MRME 的非常有效的一种改进方法；(3)文献[46]在文献[14]的基础上，提出了可变块大小的两步多分辨率运动估计方法. 其基本思想是：用更精确的运动向量补偿细节区域，用较少的运动向量补偿平滑区域，以降低图像帧运动信息的总体数量，具体可以通过两步来实现：第 1 步，采用四叉树结构表示图像帧（如图 5），并对树的每个叶结点进行运动估计；第 2 步，分别估计每个叶结点及其父结点的熵值，若父结点的熵小于各子结点熵值的和，则将该子树合并成一个叶结点，这样保证了每棵子树的信息熵最小，此外，文中对文献[43]自适应阈值技术进行了改进，将零运动块的检测置于运动估计之前，并规定不对零运动块进行运动估计，免除了文献[43]中对零运动块不必要的运动估计过程. 该算法对自适应阈值技术进行了有意义的改进，加快了运动估计的速度；树的合并过程使得编码运动信息的比特能够根据帧的内容合理分配，提高了编码效率，保证了视频数据可以在更低的带宽下传输，并获得与传统 MRME 方法相媲美的质量；在运动估计时，文中提出用范数最小的运动向量代替“best_and_first match”的运动向量技术，能够取得更平滑的运动场. 但是，零运动块的使用在一定程度上增大了预测帧差的熵，另外，文中编码零运动块时，引入了一个特殊符号“NOT_CODED”，而在编码时过多的零块将使得“NOT_CODED”和“ZERO(表示零向量)”两个符号在所有运动向量的

分布上出现两个峰值,这都将降低系统熵编码器的性能.与算法的基本思想不一致的是,文中只提出用合并树的方法减少平滑区域的信息熵,却没有提出诸如分裂树等方法来增加细节区域的信息熵,因此,在较高带宽的应用中,该算法提供的视频质量将低于使用统一匹配准则的文献[14]中算法的质量;第 2 步的合并操作,尽管提高了编码效率,但是它所需要的计算量大约是第 1 步的 2.5 倍,使得编码器整体的时间复杂度比传统 MRME 还要高 2 倍左右.



(a)小波变换后图像帧的划分　　(b)四叉树的建立,图中阴影部分的叶结点对应(a)中的 7 个灰色的分块

图 5　图像帧的树结构

此外,文献[47]认为,边缘处能够很好地检测出运动,提出了基于边缘检测的快速运动估计方法,基本思想是:首先用 Sobel 算子对参考帧和目标帧的低频子带进行边缘检测,然后对边缘宏块进行运动估计.而中部宏块则是通过其周围各宏块和参考帧中相应位置的宏块的运动情况加以确定,其余各子带的运动向量以低频子带的运动向量为初始值进一步求精.该算法注意到了对象边缘在运动估计中的重要作用,这不仅有助于提高运动估计的速度,而且也有助于得到平滑的运动场,达到更好的运动估计质量.但是,边缘检测算法的性能是整个系统的瓶颈.

3.2　子带间接运动估计技术的研究进展

子带直接运动估计方法是将小波变换及其多分辨率特性运用于运动估计所产生的最朴素的方法,简单、快速.但是,在不断的应用中,人们发现,即使用最精细的方法估计、补偿的质量仍低于空域中的 FS,影响了基于小波的编码器的性能.1995 年,文献[48,49]为这种现象找到了合理的解释,即:小波变换的下采样操作破坏了原始信号频率的线性相位关系,导致频率的混叠.进而,使得信号在空间域中的平移与其小波系数在频率域中的相变不具有直接的线性关系(除平移的像素数是下采样周期的整数倍以外).我们又称这种现象为小波的"平移可变性".它将直接影响子带直接运动估计的准确度,尤其是高频子带的估计,严重降低了估计的质量.据此,人们提出了多种不同的方法以抑制混叠效应的发生,对参考帧的各子带进行了必要的处理后再进行运动估计,我们将这类方法称为"子带间接运动估计技术".

文献[48]提出了在空间域进行运动估计,在低分辨率下插值的方法,但是它仍旧是一种空域的运动估计方法.文献[49,50]则从移变性的根源出发,移去小波变换中的下采样过程,进行运动估计.其基本思想是:从最低分辨率开始,首先分别对参考帧和目标帧进行小波变换,前者变换过程中不进行下采样操作.然后,对最低频子带进行运动估计,得到的运动向量同时作为同一分辨率下高频子带的运动向量,并结合参考帧计算出目标帧最低分辨率下的 4 个预测子带.最后将 4 个预测子带经过一阶小波逆变换,得到次低分辨率下的最低频子带.重复上述各步,直到达到最高分辨率为止.该算法通过越过下采样的步骤直接进行运动估计,能够消除移变性所产生的影响,从而提高估计的准确度.随着分辨率的增大,运动估计和补偿按层次依次进行,这种分层结构尤利于视频的可分级编码.但是,算法仍然存在着不足:(1)算法采用的运动估计是不均衡的,即只消除了参考帧的频率混叠,而目标帧仍被混叠效应所影响,这时所进行的运动估计也必然是不准确的,定会引起估计质量的下降;(2)移去小波变换中的下采样过程,将运动估计所需要的存储空间增加了 1.5 倍;(3)算法采用的运动估计是基于像素的梯度下降法,进一步增加了时间复杂度,文中亦没有对此作必要的讨论和实验.文献[51]更加详细、生动地阐述了采样对运动估计所产生的影响,并提出用"algorithm àtrous"进行运动估计.与文献[49,50]相比较,算法对参考帧和目标帧进行小波变换的方式是相同的;不同的是,算法直接将目标帧的各个子带与参考帧的对应子带进行运动估计,从而较文献[49,50]更加准确地估计了高频子带的运动向量,也节省了文献[49,50]反复重构高分辨率子带的计算量.但是,算法仍旧没有解决文献[49,50]中估计不均衡、存储空间大和高计算量的问题.文献[52]提出了"预测环损失(loss of prediction loop)"的概念.所谓"预测环损失",是指用参考帧及运动向量进行补偿,得到的对目标帧的预测,经一阶小波分解所获得的低频子带,与先对参考帧及运动向量进行小波分解,然后再利用参考帧的低频子带进行补偿所获得的结果是不同的,即运动补偿和小波变换是不可换的操作.这意味着空间域中的补偿和小波域中的补偿将产生一定的偏差.显然,文献[49,50]中的算法由上至下的估计必将受到预测环损失的困扰.于是,文献[52]

兼顾考虑了移变性和预测环损失的影响，提出了一种适于多尺度视频编码的运动估计方法. 其基本思想是：从最低分辨率开始，首先利用参考帧的低频子带对目标帧的低频子带进行运动估计，同一分辨率下高频子带的运动向量由其下层低频子带的运动向量经一阶小波分解得出，并对该下层低频子带的预测进行一阶小波分解，同时得到 3 个上层高频子带的预测.

利用下层低频子带进行运动估计和补偿，保证了算法可以有效地避免预测环损失；由于未对高频子带进行直接运动估计，这样在一定程度上缓解了移变性对运动估计的影响. 可是，各低频子带的运动估计仍未消除移变性的影响，而这种影响不可避免地将漂移至整个运动估计的过程. 另外，算法为了追求多尺度编码，对已经得到预测的高频子带还要计算其运动向量，而该运动向量完全能在解码端利用其下层低频子带求得，故对高频子带运动向量的计算和编码是不必要的，额外地引入了冗余数据. 在此基础上，文献[53]对其系统框架进一步明确和调整，提出了基于下层 LL 子带的分层运动估计方法. 这是将小波域运动估计方法应用于空间可分级编码的具体框架和方法，亦是对可分级运动估计技术(参见第 4 节中的(2))的研究和探索. 通过在下层 LL 子带的运动估计中引入快速算法(如 TSS)并省略了各高频子带运动向量的计算，在时间复杂度方面较文献[52]有所提高. 可是，移变性的影响仍未得到根本解决. 文献[54]则提出了一种构造性的子带间接运动估计方法. 因为移变性所引起的运动估计误差在高频子带的估计过程中表现得尤为明显，所以算法提出，不对高频子带进行下采样操作，从而消除移变性对高频子带运动估计的影响. 这样，可以将低频子带估计得到的运动向量作为高频子带运动向量的初始值，再对各个高频子带进行求精即可. 该算法可以有效地避免移变性对高频子带运动估计的负面影响，提高了高频子带运动估计的准确性，而人眼对于视频中运动着的边缘(反映到小波域中即为高频子带)尤为敏感，故算法还可以提高重构帧的主观质量；算法的空间复杂度和时间复杂度均低于文献[49～51]. 但是，算法在编码时需要将未下采样的高频子带的运动向量及帧差全部编码，这意味着将比其它的算法多编码 2 倍左右的数据，而这些引进的冗余将直接降低编码器的效率. 另外，低频子带的运动估计仍然受移变性的影响.

文献[55,56]从理论上证明了混叠效应的发生，并据此推导出一种最优的插值滤波器，以达到尽量减少频率混叠的目的. 算法的基本思想是：每次对低频子带进行运动估计之前，都利用最优插值滤波器对参考帧和目标帧的低频子带进行滤波，然后采用与文献[53]几乎相同的方法逐层进行运动估计. 最优插值滤波器的构造，替代了小波变换的低通重构滤波器，最大程度地减少了低频子带的频率混叠效应，其时间和空间复杂度均明显低于文献[49～51,54]的复杂度. 与文献[49,50]相比，在保持算法复杂度不变的基础上，将有效地提高[49,50]算法的估计质量. 实验结果显示，重构帧的 PSNR 较 MPEG-2 和 H.263 平均高出 2db. 但是即便这样，也不能完全消除混叠效应的发生，估计出的运动向量仍然是不准确的.

3.3 克服平移可变性的运动估计研究进展

子带间接运动估计技术在一定程度上弥补了小波移变性所产生的影响，但是它所计算出的运动向量也并不准确. 在认真分析了移变性的特征和本质之后，文献[57]提出了一种能克服移变性的估计方法——低频子带平移技术(low band shift). 图 6 所示为一维低频子带平移方法的示意图，其中 $f(n)$ 为原始信号，$Lf(a)(b,na)$ 表示 $f(n)$ 在空间域中平移 b 个像素后经 a 阶小波变换所得到的低频信号，na 表示该平移信号在频率域内所产生的相应的平移，$Hf(a)(b,na)$ 表示同样情况下变换得到的高频信号.

由图 6 可以观察出：信号在空间域中平移任意一个整数位移 b 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的低频子带的方法产生出来. 在此基础上，提出了低频子带平移的运动估计方法(Low Band Shift Motion Estimation, LBSME). 基本思想是：将参考帧的低频子带分别沿着水平、竖直和对角线方向平移一个像素后，对目标帧和参考帧及其 3 个平移帧进行一阶小波变换，然后利用子带直接运动估计方法计算运动向量. 对于多级小波分解，可以对前一次分解生成的低频子带重复上面的平移和分解操作即可. 因每次都只平移低频子带，算法故而得名. 另外，文中还提出了一种称为“小波块(wavelet block)”的结构和一种新的匹配准则来辅助子带直接运动估计(如图 7 所示). 经过上述的操作，依次估计出每个小波块的运动向量，进而得到小波块内部不同分辨率下各个宏块的运动向量. 文献[58]在此基础上，提出一种基于 LBS 的直接子带估计算法. 与 LBS 不同，LBS 中首先以小

波块为整体进行估计,小波块内部各宏块的运动向量是由该小波块的向量经过适当放大后得到的,而后者则为每个内部宏块都计算出其最佳运动向量,可见后者能够更加精细地刻画各宏块的运动.

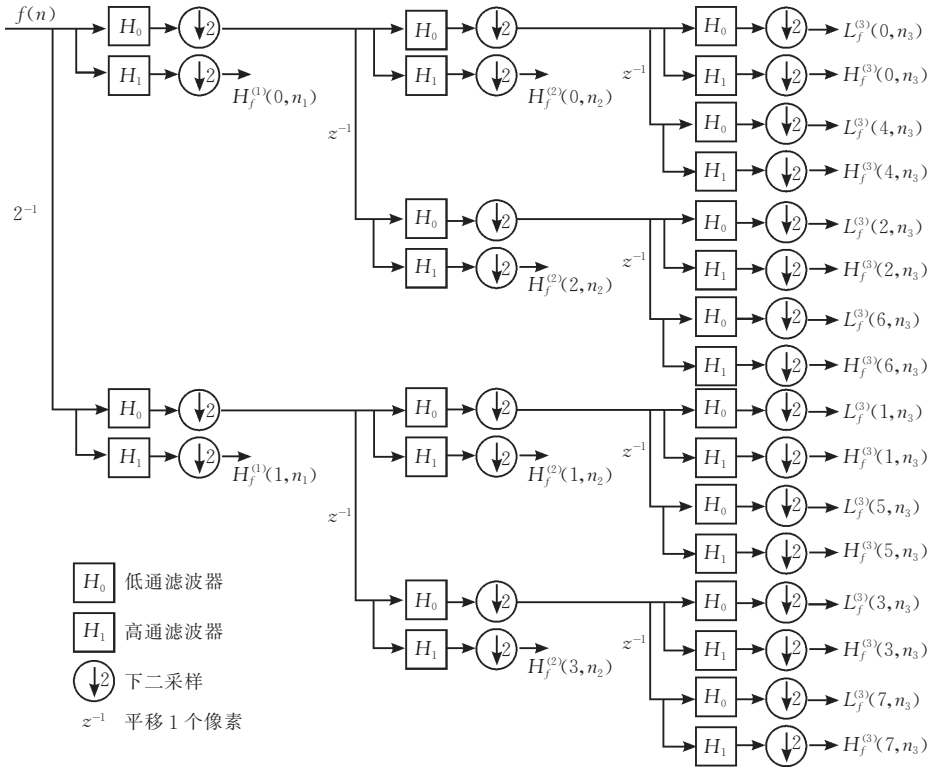


图 6　一维信号的低频子带平移方法示意图

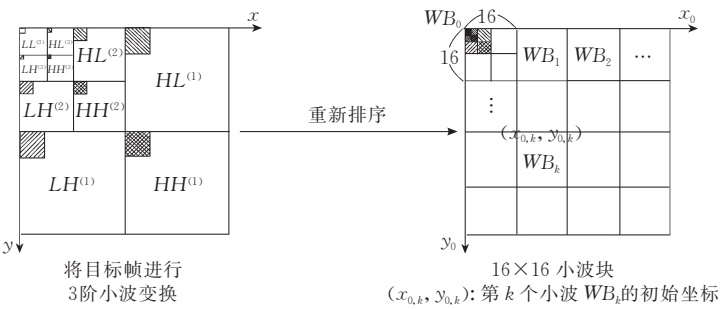


图 7　小波块形成示意图

算法克服了小波移变性,使得运动估计的准确度和主观质量明显高于空间域的 FS 和传统的子带直接运动估计方法,但是算法的时间复杂度和空间复杂度均很大,前者比 FS 时间复杂度大约高 10.3%,而后者则是空间域运动估计算法和传统子带直接运动估计算法空间复杂度的 10 倍.因此,从评价算法的角度看,该方法并不理想.另外,小波块结构本身不具有层次性,不适于视频可分级编码的需要.

边界宏块的运动估计历来都是一个棘手的问题,LBSME 同样未能免于边界问题的困扰.文献[59]在这个方面对 LBSME 技术作了进一步增强,提出了基于线性双正交小波的边界对称延拓(pad-

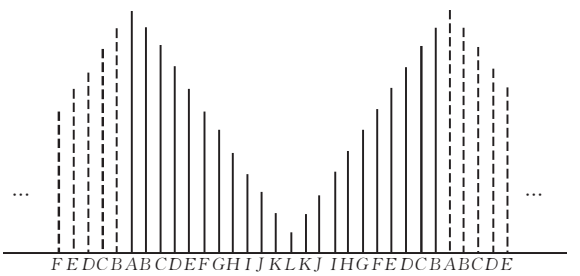


图 8　边界对称延拓技术示意图(其中 A～L 为原始信号样本)

ding)技术,其具体方法如图 8 所示.设信号 $f(n)$的长度为 N,则延拓后的信号长度为 $2N-2$,且具有 N 种唯一值,首尾两个信号样本只出现一次.在每一阶小波分解之前,都须按此规则对信号的低频子带进行延拓和系数调整,以保证信号的边界样本在

任何一次分解前后都具有对称特性.

这种延拓技术能够降低小波分解后边界宏块系数的幅值,并能够很好地支持不受限制的运动估计技术.实验证明,利用该方法进行运动估计,重构帧的 PSNR 比 LBSME 方法的平均高出 0.3～0.4db,从而进一步提高了 LBSME 方法估计的准确度和主观质量.不足的是,该方法会稍提高 LBSME 方法的计算量.

文献[57～59]中方法的复杂度均是很高的,文献[60]则从另一个角度出发,对一维阶跃信号 f 和平移一个单位后的信号 f_s 分别进行双正交小波分解,然后对两者高频系数的关系进行了理论上的推导.结果发现,对应系数的和的绝对值的和 SAS (Sum of Absolute Sum)比 SAD 要小,并且小波若使用 Biorthogonal(2,4)和 Biorthogonal(2,8)滤波器,则 SAS 为零.进而提出了一种适用于高频子带的运动估计匹配准则 SAS,并在此基础上提出了结合使用 SAS 和 SAD 的运动估计方法,即:使 SAS 和 SAD 获得二者中的最小值的向量作为目标块的运动向量.该算法在未提高传统子带直接运动估计方法的复杂度的情况下,能够在高频子带中对运动向量做出较为准确的估计,是实时应用和简单应用中的理想选择.不足的是,算法需对滤波器的选择做出一定的限制,即能够使 SAS 值为零或者接近零的滤波器才能较为准确地估计出运动向量,故影响了它的应用.另外,结合使用 SAS 和 SAD 的估计方法没有将 SAS 理论上的推导运用到实处,实际上会影响 SAS 准则的准确性.由于在高频子带中,正确的运动向量未必能使 SAS 获得比 SAD 更小的值,这样估计出的运动向量亦未必准确.

文献[61]将子带间接运动估计技术适用于可分级编码的层次结构与 LBSME 方法的高准确度相结合,提出了基于 LBS 的分层运动估计方法.其基本思想是:从最低分辨率开始,将目标帧的低频子带进行上采样和低通滤波,将参考帧经过 LBS 方法得到的 4 个低频子带进行与目标帧同样的操作,对目标帧的低频子带进行运动估计,然后将目标帧低频子带的预测、参考帧的次低频子带及其 3 个平移子带通过一定的方式结合成一个预测帧,对这个预测帧进行一阶小波分解,从而得到高频子带的预测.将当前分辨率下四个子带的预测进行一阶小波逆变换,得到次低分辨率下的低频子带.重复上述操作,直到达到最高分辨率为止.算法利用 LBS 克服了低频子带中移变性的影响,同时移去了目标帧低频子带的下采样操作,避免了不均衡的运动估计,进一步提高了文献[49,50]中低频子带以及整个系统运动估计的质量.但是其存储空间的需求是文献[49,50]的 4 倍,计算复杂度也有明显提高.可见,算法的代价也是我们所不得不考虑的.

除了上述用构造性的方法或结构克服移变性的算法以外,一些研究学者专门对小波进行了改进,提出了一些具有平移不变性的小波变换,如复小波(complex wavelet)、过完备小波(overcomplete wavelet)以及相应的运动估计技术.但是这些算法均具有计算量大、需要的存储空间大,甚至不利于编码等诸多不足.鉴于文章的篇幅,在此不再赘述,进一步的详情请参看文献[62～68].

4 小波域运动估计技术展望

随着人们对小波的进一步认识和多媒体应用领域的发展,对小波域运动估计技术的一些更高的要求将逐渐被提出来,我们认为未来小波域运动估计技术将在以下几个方面有更好的发展:

(1)更为高效的克服移变性的运动估计技术

小波域子带直接运动估计技术的瓶颈在于移变性,而子带间接运动估计技术和 LBS 技术的瓶颈又在于其空间和时间复杂度.因此,提出一种能有效克服或在一定程度上克服移变性,而又不明显增加空间、时间复杂度的算法,是至关重要的.随着基于小波的编码器的进一步研究和推广应用,对高质量、高效率的运动估计算法的需求将愈加迫切.

(2)适合于视频可分级编码的运动估计技术

伴随着 Internet 和无线网络技术的不断发展,图像和视频的可分级编码变得愈来愈重要,是近年来静态图像和视频编码领域的又一个研究热点,出现了很多可分级视频编码技术[28,29,36,42,53,70].而小波良好的多分辨率特性恰恰适合于视频可分级编码的要求,它可以对对象及其运动进行更细粒度的编码.因此,基于小波的细粒度可分级视频编码以及支持它的可分级运动估计技术将得以更广泛、更深入的发展.

(3)基于对象的小波域运动估计技术

"对象"是 MPEG-4 标准编/解码机制的基础.目前,已经出现了多种基于对象的空间域运动估计技术,如文献[26,27].但是基于对象的小波域运动估计技术却尚未提出.仅仅将对象直接引入到现有的小波域运动估计技术中来是远远不够的,也是不

科学的.如何利用形状信息来指导、加速和提高小波域运动估计技术的速度和质量,如何利用小波的特性给予对象及其运动更方便、直接和高效的表示和处理等等,仍是需要进一步深入研究和解决的问题.

(4)具有平移不变性小波变换的构造

目前具有平移不变性的小波,由于会引入冗余数据,均不适合于编码.而移变性则是小波域运动估计过程中必须考虑的因素,它直接影响着运动估计的精确程度.只有一种适于编码,而又满足移不变性的小波才能从根本上解决这个问题.因此,移不变小波的构造是必要的,也是重要的,更是值得深入研究的一个课题.

5 结束语

作为视频编码的核心技术之一,运动估计技术已经成为一个重要的研究领域.本文首先对比研究了空间域和小波域运动估计技术,然后以小波域运动估计技术的发展为主线,对三类小波域运动估计方法分别进行了概括性阐述,对不同方法的优势和不足进行了较为深入的研究和讨论,最后对未来小波域运动估计技术的发展进行了展望.

参 考 文 献

1 ISO/IEC JTC1/SC29/WG11, Information technology- generic coding of audio-visual objects, part 2: Visual, ISO/IEC14496-2, Final Draft of International Standard, 1998

2 ISO/IEC, JTC1/SC29/WG1, ISO/IEC FCD 15444-1: Information technology JPEG 2000 image coding system: core coding system [WG 1 N1646R], 2000

3 Zhong Min-Seng, Ma Zheng-Ming. Motion compensation based on wavelet coefficient blocks. Acta Automatica Sinica, 2004, 30(1): 64～69(in Chinese)
(钟敏生,马争鸣.基于小波系数块的运动补偿.自动化学报, 2004, 30(1): 64～69)

4 Katto J., Ohki J. I., Nogaki S., Ohta M.. A wavelet codec with overlapped motion compensation for very low bit-rate environment. IEEE Transactions on Circuits and Systems for Video Technology, 1994, 4(3): 328～338

5 Orchard M. T., Sullivan G. J.. Overlapped block motion compensation: An estimation-theoretic approach. IEEE Transactions on Image Processing, 1994, 3(5): 693～699

6 Martucci Stephen A., Sodagar I., Chiang T., Zhang Ya-Q.. A zerotree wavelet video coder. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(1): 109～118

7 Marpe D., Cycon H. L.. Very low bit-rate video coding using wavelet-based techniques. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(1): 85～94

8 Ekram Khan, Mohammed Ghanbari. An efficient and scalable low bit-rate video coding with virtual SPIHT. Signal Processing: Image Communication, 2004, 19(3): 267～283

9 Podilchuk C. I., Jayant N. S., Farvardin N.. Three-dimensional subband coding of video. IEEE Transactions on Image Processing, 1995, 4(2): 125～139

10 Chen Y., Pearlman W. A.. Three-dimensional subband coding of video using the zero-tree method. In: Proceedings of Visual Communications and Image Processing'96, SPIE 2727, Orlando, 1996, 1302～1312

11 Kim B. J., Pearlman W. A.. An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT). In: Proceedings of Data Compression Conference, Snowbird, Utah, 1997, 251～260

12 Xu J. Z., Xiong Z. X., Li S. P., Zhang Y. Q.. Memory-constrained 3-D wavelet transform for video coding without boundary effects. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(9): 812～818

13 Secker A., Taubman D.. Lifting-based invertible motion adaptive transform (LIMAT) framework for highly scalable video compression. IEEE Transactions on Image Processing, 2003, 12(12): 1530～1542

14 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

15 Jain J. R., Jain A. K.. Displacement measurement and its application in interframe image coding. IEEE Transactions on Communications, 1981, COMM-29(12): 1799～1808

16 Koga T., Iinuma K., Hirano A. and Ishiguro T.. Motion-compensated interframe coding for video conferencing. In: Procee. dings of National Telecommunications Conference (NTC81), New Orleans, LA, 1981, C9.6.1～9.6.5

17 Srinivasan R., Rao K. R.. Predictive coding based on efficient motion estimation. IEEE Transactions on Communications, 1985, COMM-33(8): 888～896

18 Ghanbari M.. The cross-search algorithm for motion estimation. IEEE Transactions on Communications, 1990, 38(7): 950～953

19 Li R., Zeng B., Liou M. L.. A new three-step search algorithm for block motion estimation. IEEE Transactions on Circuits and Systems for Video Technology, 1994, 4(4): 438～442

20 Po L. M., Ma W. C.. A novel four-step search algorithm for fast block motion estimation. IEEE Transactions on Circuits and Systems for Video Technology, 1996, 6(3): 313～317

21 Liu L. K., Feig E.. A block-based gradient descent search algorithm for block motion estimation in video coding. IEEE Transactions on Circuits and Systems for Video Technology, 1996, 6(4): 419～423

22 Zhu S., Ma K. K.. A new diamond search algorithm for fast block-matching motion estimation. IEEE Transactions on Image Processing, 2000, 9(2): 287～290

23 Chan Y. L., Siu W. C.. An efficient search strategy for block motion estimation using image features. IEEE Transactions on Image Processing, 2001, 10(8): 1223～1238

24 Brünig M., Niehsen W.. Fast full-search block matching. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(2): 241～247

25 Chung K. L., Chang L. C.. A new predictive search area approach for fast block motion estimation. IEEE Transactions on Image Processing, 2003, 12(6): 648～652

26 Yu D., Jang S. K., Ra J. B.. Fast motion estimation for shape coding in MPEG-4. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(4): 358～363

27 Hui K. C., Siu W. C., Chan Y. L.. Fast motion estimation of arbitrarily shaped video objects in MPEG-4. Signal Processing: Image Communication, 2003, 18(1): 33～50

28 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

29 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

30 Metin Uz K., Vetterli M., LeGall D. J.. Interpolative multiresolution coding of advanced television with compatible subchannels. IEEE Transactions on Circuits and Systems for Video Technology, 1991, 1(1): 86～99

31 Zafar S., Zhang Ya Q., Jabbari B.. Multiscale video representation using multiresolution motion compensation and wavelet decomposition. IEEE Journal on Selected Areas in Communications, 1993, 11(1): 24～34

32 Lee Jinhwan, Man Ro Y.. Hierarchical block matching algorithm in MRME. In: Proceedings of the 2000 International Conference on Image Processing (ICIP 2000), Vancouver, BC, Canada, 2000, 836～839

33 Shang Yan-Xia. Wavelet domain based video coding [M. S. dissertation]. Dalian University of Technology, Dalian, China, 2003(in Chinese)

(尚艳霞. 基于小波域的视频压缩编码[硕士学位论文]. 大连理工大学, 大连, 2003)

34 Panchanathan S., Chan E., Wang X.. Fast multiresolution motion estimation scheme for a wavelet transform video coder. In: Proceedings of SPIE Visual Communications and Image Processing, Chicago, 1994, 2308: 671～681

35 Chung M., Kang S., Choi J. H., Kwak H. S.. New multiresolution motion estimation for wavelet transform video coding. In: Proceedings of IEEE Asia Pacific Conference on Circuits and Systems, Seoul, Korea, 1996, 10～13

36 Wang Q., Ghanbari M.. Scalable coding of very high resolution video using the virtual zerotree. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(5): 719～726

37 Karlekar J., Desai U. B.. New multiresolution motion estimation and compensation scheme. In: Proceedings of IEEE International Symposium on Circuits and Systems, Orlando, USA, 1999, Ⅳ-459～462

38 Cosman P., Gray R. M., Vetterli M.. Vector quantization of image subband: A survey. IEEE Transactions on Image Processing, 1996, 5(2): 202～225

39 Joo S., Kikuchi H.. A new motion compensation on a wavelet transform domain. In: Proceedings of IEEE International Conference on Image Processing (ICIP 2000), Vancouver, Canada, 2000, 130～133

40 Zan Jin-Wen, Ahmad M. O., Swamy M. N. S.. Wavelet-based multiresolution motion estimation through median filtering. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02), Orlando, Florida, USA, 2002, IV-3273～3276

41 Zan Jin-Wen, Ahmad M. O., Swamy M. N. S.. New techniques for multi-resolution motion estimation. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(9): 793～802

42 Su Y.-C.. Fine granularity video compression technique and its application to robust video transmission over wireless Internet [Ph. D. dissertation]. National Sun Yat-Sen University, Taiwan, 2003

43 Mandal M. K., Chan E., Wang X., Panchanathan S.. Multiresolution motion estimation techniques for video compression. Optical Engineering, 1996, 35(1): 128～136

44 Wei Jie, Li Ze-Nian. An enhancement to mrmc scheme in video compression. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(3): 564～568

45 Zhao Ting R., Ohtsuki T.. A fast MRME scheme for video compression based on coefficients masking in the wavelet domain. In: Proceedings of the IEEE Region 10 Conference (TENCON), Cheju Island, South Korea, 1999, 127～130

46 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

47 Gleich D., Planinšic P., CucejŽarko. Low bitrate video coding using wavelet transform. In: Proceedings of the 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications (EC-VIP-MC), Zagreb, Croatia, 2003, 369～374

48 Tsunashima K., Stampleman J. B., Bove V. M.. A scalable motion-compensated subband image coder. IEEE Transactions on Communications, 1994, 42(2/3/4): 1894～1901

49 Nosratinia A., Orchard M. T.. A multi-resolution framework for backward motion compensation. In: Proceedings of SPIE Electronic Imaging, San Jose, CA, 1995, 190～200

50 Nosratinia A., Orchard M. T.. Multi-resolution backward video coding. In: Proceedings of IEEE International Conference on Image Processing, Washington, DC, 1995, 563～566

51 Skowronski J.. Pel recursive motion estimation and compensation in subbands. Signal Processing: Image Communication, 1999, 14(5): 389～ 396

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

53 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)

54 Meyer F. G., Averbuch A., Coifman R. R.. Motion compensation of wavelet coefficients for very low bit rate video coding. In: Proceedings of IEEE International Conference on Image Processing, Washington, DC, 1997, 638～641

55 Yang Xu-Guang, Ramchandran K.. Hierarchical backward motion compensation for wavelet video coding using optimized interpolation filters. In: Proceedings of International Conference on Image Processing (ICIP'97), Washington, DC, 1997, 85～88

56 Yang Xu-Guang, Ramchandran K.. Scalable wavelet video coding using aliasing-reduced hierarchical motion compensation. IEEE Transactions on Image Processing, 2000, 9(5): 778～791

57 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

58 Kim Hyung Sun, Park Hyun-Wook. Wavelet-based moving-picture coding using shift-invariant motion estimation in wavelet domain. Signal Processing: Image Communication, 2001, 16(7): 669～679

59 Fu Ming-Fai, Au O. C., Chan W. C.. Low band shift (LBS) motion estimation with symmetric padding in wavelet domain. In: Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS 2002), Scottsdale, Arizona, USA, 2002, III-13～III-16

60 van der Auwera G., Munteanu A., Lafruit G., Cornelis J.. Video coding based on motion estimation in the wavelet detail images. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'98), Seattle, USA, 1998, 2801～2804

61 Yuan Yu-Fei, Mandal M. K.. Low-band-shifted hierarchical backward motion estimation and compensation for wavelet-based video coding. In: Proceedings of the 3rd Indian Conference on Computer Vision, Graphics and Image Processing, India, 2002, 185～190

62 Magarey J. F. A.. Motion estimation using complex wavelets [Ph. D. dissertation]. Cambridge University Department of Engineering, London, 1997

63 Magarey Julian, Kingsbury N.. Motion estimation using a complex-valued wavelet transform. IEEE Transactions on Signal Processing, 1998, 46(4): 1069～1084

64 Xiong Hui-Lin, Zhang Tian-Xu, Moon Y. S.. A translation- and scale-invariant adaptive wavelet transform. IEEE Transactions on Image Processing, 2000, 9(12): 2100～2108

65 Kingsbury N.. Complex wavelets for shift invariant analysis and filtering of signals. Journal of Applied and Computational Harmonics, 2001, 10 (3): 234～253

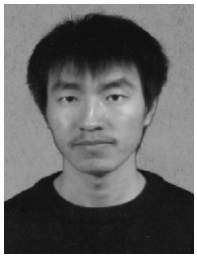
66 DeVore R., Petukhov A., Sharpley R.. Motion estimation with the redundant wavelet transform. In: Proceedings of International Workshop on Digital and Computational Video, Florida, USA, 2002, 53～59

67 Cui Su-Xia. Motion estimation and compensation in the redundant wavelet domain [Ph. D. dissertation]. Mississippi State University, Mississippi State, USA, 2003

68 Chen Jing, Cai C., Ding R.. Oversampled wavelet motion compensation and its hierarchical block-matching algorithm. In: Proceedings of IEEE International Conference on Neural Networks & Signal Processing, Nanjing, China, 2003, 1270～1273

69 Li Xin. Scalable video compression via overcomplete motion compensated wavelet coding. Signal Processing: Image Communication, 2004, 19(7): 637～651

70 van der Schaar M., Radha H.. Adaptive motion-compensation fine-granular-scalability (AMC-FGS) for wireless video. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(6): 360～371



SONG Chuan-Ming, born in 1980, M. S. 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

We're 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 project firstly studies the theories and methods of the object-based shape-adaptive wavelet transform and wavelet-based video object motion estimation techniques. And on the basis of the above, the final goal is to propose a fine granular spatial scalable, rate scalable and hybrid spatial and rate scalable video object coding scheme based on the 2D wavelet and its motion compensation. In the whole project, motion estimation in the wavelet domain is a key technique. Over the past year, we have been actively organizing the discussion in our research group once a week. So far, we have obtained some phased achievements. This paper is written on the basis of the discussion and gives an overview of the research progress of video motion estimation techniques in the wavelet domain.