

# 基于区域分割的小波分解运动补偿编码算法

刘红梅<sup>1)</sup> 黄金科<sup>2)</sup> 肖自美<sup>1)</sup> 郑慧诚<sup>3)</sup> 梁 凡<sup>1)</sup>

<sup>1)</sup>(中山大学电子与通信工程系, 广州 510275)

<sup>2)</sup>(广州石竹科技有限公司, 广州 510050)

<sup>3)</sup>(广东省电信公司广州市分公司数据分局, 广州 510055)

**摘要** 为了提高视频运动估值的速度和运动矢量编码的效率, 提出了一种基于区域分割的小波分解运动补偿的视频编码算法. 该算法的几个核心部分的特点如下: (1) 运动估值部分, 在基本的可变块多分辨率运动估值算法的基础上, 充分利用了小波分解后各子带所表征的运动结构相关性, 因而提高了运动矢量的编码效率; 同时提出了基于区域分割的多分辨率运动估值算法, 因而在有效减少运动估值时间消耗的同时, 进一步提高了运动矢量的编码效率; (2) 在量化部分, 提出了一种基于人的视觉系统的量化方法; (3) 采用改进的高效的零树小波编码算法对帧内和帧间信息进行编码. 实验结果表明, 所提出的算法是有效的.

**关键词** 视频编码 小波变换 零树小波 多分辨率运动估值 量化

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2002)03-0292-05

## A Region-Segmentation-Based Multiresolution Motion Compensation Wavelet Video coding Algorithm

LIU Hong-mei<sup>1)</sup>, HUANG Jin-ke<sup>2)</sup>, XIAO Zi-mei<sup>1)</sup>, ZHENG Hui-cheng<sup>3)</sup>, LIANG Fan<sup>1)</sup>

<sup>1)</sup>(School of Information Science and Technology, Zhongshan University, Guangzhou 510275)

<sup>2)</sup>(Carnation Technology Ltd., Guangzhou 510050)

<sup>3)</sup>(Guangzhou Telecommunication Corporation Data Communication Branch, Guangzhou 510055)

**Abstract** In this paper, a region-segmentation-based multiresolution motion compensation wavelet video coding algorithm is proposed. The features of the kernel parts of this coder are as follows: 1) In the part of motion estimation, based on the basic variable block multiresolution motion estimation algorithm, we take full advantage of the spatial and orientation correlations between the motion structures represented by the wavelet decomposition subbands and improve the coding efficiency of the motion vectors; a region-segmentation based multiresolution motion estimation algorithm is presented, under this algorithm, the computation complexity of motion estimation is reduced and the motion vector coding efficiency is improved farther. 2) In the part of quantization, a quantization scheme based on human vision system is proposed. 3) An improved version of zerotree wavelet coding algorithm with high performance is adopted to code the intraframe and interframes. The experimental results prove the effectiveness of the algorithm this paper presented.

**Keywords** Video coding, Wavelet transform, Zerotree wavelet, Multiresolution motion estimation, Quantization

## 0 引言

旨在实现有限带宽信道上视频传输的低码率视频编码是近年来迅速发展的一個研究领域. 目前, 在

此研究领域, 除了传统的基于分块 DCT 和运动补偿的视频编码方案已趋于成熟, 并已形成若干国际标准外, 又出现了许多新的思想和技术, 其中将变换编码与子带编码相结合的小波变换方案就是很有发展前景的技术之一<sup>[1]</sup>.

由于小波变换是以时(空)-频局部化函数所形成的小波基作为基底而展开的,因此它具有许多特殊和优异的性能,包括时(空)-频局部化特性、多分辨率分析特性等,并首先成功应用在静止图象的压缩编码上。目前,图象编码新标准——JPEG2000 就采用了基于小波变换的编码算法,其在性能上远远优于传统的基于 DCT 的图象压缩算法,但将小波变换应用于视频压缩编码,还需要增加消除时间冗余的机制,很多研究者在研究将 2D 小波方案进行扩展,以应用于视频编码的解决方法,这些方法归纳起来大致可以分为如下 3 类:第 1 类是将 2D 小波变换或子带编码方案扩展到三维子带编码<sup>[2]</sup>;第 2 类方法采用多分辨率运动估值/补偿(MRME/MRMC, Multi-Resolution Motion Estimation/Multi-Resolution Motion Compensation)算法来降低时间冗余<sup>[3~8]</sup>;第 3 类方法则采用时域内的块运动估值/重叠块运动补偿算法来降低时间冗余<sup>[9]</sup>。本文的方案是基于第 2 种思路,其理由是:一方面,这种方案在编码端减少了反小波变换运算,从而降低了编码端的运算复杂度;另一方面,在基于运动补偿算法的编码器中,主要运算开销往往在运动估值上,由于第 2 类方案在编码端采用多分辨率运动估值的方法,它可以充分利用小波分解后各子带间的相关性来降低运动估值所占用的时间,因而使这种方案在实时编码方面有优势。

可变块大小多分辨率运动估值算法是 Zhang 和 Zafar 将小波理论应用于视频压缩时提出的,它利用小波金字塔结构每层间的相关性,来降低运动估值过程的运算复杂性<sup>[3]</sup>。但文献<sup>[3]</sup>中的算法存在着运动矢量编码占用比特数太多及各子带采用全搜索方法进行运动估值消耗时间较多等不足。本文提出的基于区域分割的 MRME 算法,则在运动估值的速度和准确性方面均有提高。由于具有多分辨率分析特性的小波变换与人眼的视觉特性相符合,因此在编码器的量化策略上,本文提出了一种基于人眼视觉特性的量化策略。由于小波零树编码被证明是一种有效的对小波变换系数进行编码的方法,因此本文采用改进的零树小波编码算法——基于子集分割的零树编码(SDZW, Subset-Division-based Zerotree Wavelet)算法<sup>[10]</sup>来对帧内和帧间信息进行编码。

## 1 基于区域分割的小波分解运动补偿编码算法

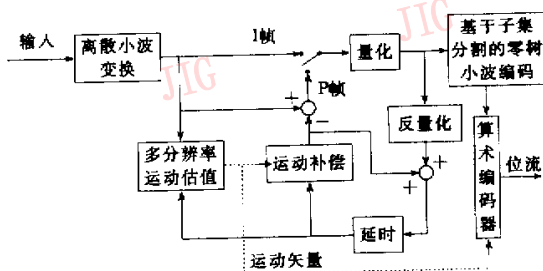
器框图。它包含以下 4 个部分:

(1) 离散小波变换部分 本文采用双正交 D9/7 滤波器进行小波分解与重建,其小波分解的级数根据编码序列的分辨率确定。

(2) 帧间冗余消除部分 提出了基于区域分割的多分辨率运动估值和运动补偿算法。

(3) 量化和反量化部分 提出了一种基于人眼视觉系统的量化策略。

(4) 帧内和帧间信息编码部分 采用改进的零树小波编码算法——基于子集分割的零树编码(SDZW, Subset-Division-based Zerotree Wavelet)及算术编码的算法。



### 1.1 基于区域分割的多分辨率运动估值算法

Zhang 在文献<sup>[3]</sup>中提出了基于小波变换域的多分辨率运动估值 MRME 算法,虽然该方法利用了小波分解域的子带相关性,从而减小了运动估值的计算复杂度,且易于实现可分级编码,但是 MRME 算法中运动矢量编码占用码率较高;针对这一不足,文献<sup>[4]</sup>提出了两阶段法,以改进 MRME 算法,这种方法通过合并相邻块的方法,以提高运动矢量的编码效率,但运算复杂度较高;文献<sup>[6]</sup>、<sup>[7]</sup>针对 MRME 算法使用全搜索法进行运动估值耗费时间多的问题,通过提取潜在运动区,使运动估值运算只在运动区域进行,从而使编码速度得到提高,但没有提高运动矢量的编码效率。

本文提出的基于区域分割的 MRME 算法,其基本思想是:

(1) 利用视频帧之间的时间冗余,以降低 MRMC 算法的计算开销,因为 DWT 是线性的、局部化很好的变换,其小波变换之后,时间冗余和局部特性将会保留在相邻帧的相应变换系数中,所以,如果运动估值之前和在小波子图中定位运动可能发生的区域,能使相应运动矢量的搜索仅在这些区域进行,则会提高 MRMC 算法的速度。本文通过运动检测,首先对小波变换的低通子带进行区域分割,将其

图 1 是基于多分辨率运动补偿算法的视频编码

分割成静止和运动区域;然后在每个子带,从粗到细进行有需要的局部范围内全搜索,这样就能将有限的运动时间分配到搜索补偿最有效的运动内容上。

(2) 由于 MRME 算法在本质上是一种分层的运动估值算法,且小波金字塔最低分辨率层上低通子带的运动矢量是其他所有子带进行运动估值的基础,它们的失真将对所有其他层的预测产生较大影响,并将导致在恢复图象中出现不可忽视的质量损伤,因此提高基础运动矢量的准确度,对 MRME 算法是很重要的. 本文的 MRME 算法在搜索最低分辨率低通子带的运动矢量时,由于是利用最低分辨率层所有子带提供的信息来获得作为其他子带估值基础的运动矢量<sup>[4]</sup>,从而有效地提高了 MRME 算法中基础运动矢量的准确性。

(3) 用区域分割结果来指导运动估值过程,以提高运动矢量的准确性、一致性以及降低运动矢量编码所需的比特数. 由于在视频编码与通信中,需要传输的数据中有很大大一部分是用于传送运动矢量和运动补偿后的预测残差,因此在低码率的视频通信中,用于运动矢量编码的比特数所占的比例不可忽视. 由于大多数基于运动补偿的视频编码方案中,对运动矢量的编码均采用差分编码(DPCM)方式,因此,如运动估值得到的运动矢量场越平滑,而且用于运动矢量编码越少,则可将节省下来的比特用于预测误差的编码,从而有利于提高复原图象的质量. 在进行运动估值时,本文首先利用分割结果,为每个待匹配块(也就是每个运动矢量),引入一个可靠性标志 reliable; 然后使用它指导运动估值过程<sup>[11]</sup>,如果该块所有的象素都运动或者都静止,则认为该块是可靠的;如果块内并非所有的象素都具有相同的状态,则可判断该块位于运动物体的边缘,而对于运动估值来说,则认为它是不可靠的. 运动估值的基本思想是:尽可能地利用预测值,整个估值算法分两次扫描象素,第1次只针对运动区域中可靠的块,那在较少预测信息的基础上进行估值;第2次再对运动区域中不可靠的块进行运动估值,此时由于可靠的块已全部进行了估值,因此可提供较多的预测信息,在此基础上,再提高估值的准确性和一致性<sup>[11]</sup>. 之后,用 DPCM 对运动矢量进行编码,就可较大地降低运动矢量所需的比特数。

### 1.2 基于视觉特性的量化方法

因为均匀门限量化器是 MSE 准则下的最佳标量熵约束量化器,而自适应死带又可用于码率控制,所以本文的视频编解码器采用自适应死带均匀门限量化器。

由于人眼视觉系统在压缩图象的最终视觉质量中扮演着重要的角色,而且通过对人眼视觉特性的分析表明,人眼对低频成分比对高频成分敏感,因此设计量化器时,可以随着分辨率的升高,将量化步长逐步加大. 本文的量化器采用了下面的步长分配策略:  $q(i) = Q_0 \cdot \lambda_r^{i-1}$ , 其中,  $r$  分别代表  $h$  (水平)、 $v$  (垂直)、 $t$  (斜方向);  $Q_0$  为基准步长,由码率决定;  $L$  为图象的小波分解级数;  $i$  表示当前待量子带所在的分解级,从最高分辨率层到最低分辨率层,分解级编号从 1 开始递增. 进行小波系数量化时,对于最低分辨率层的低通子带中的系数,取量化步长  $q = Q_0$ ,  $\lambda_r$  为不同分解级间量化步长的比例,  $\lambda_v, \lambda_h$  分别代表水平和垂直的不同分解级间量化步长的比例,取  $\lambda_v = \lambda_h = 1.3, \lambda_t = 1.5$ . 这与人眼对于斜方向细节不太敏感的特性也是一致的。

从低分辨率到高分辨率按比例逐级增加量化步长的一个附带好处是,使高分辨率子带中的系数在经过量化后出现更多的零,并且可进一步保证对应于同一局部空间的高分辨率子带中的小波系数比低分辨率子带中的小波系数小,从而更有利于零树编码。

量化器重建公式为:

$$R = \begin{cases} 0 & l = 0 \\ (l-1) \cdot q + \mu \cdot q + d & l > 0 \\ (l+1) \cdot q - \mu \cdot q - d & l < 0 \end{cases}$$

其中,  $l$  为量化电平,  $R$  为重建值,  $d$  为死带大小,死带大小由系数  $\alpha = 2d/q$  决定,对于帧内编码,本文取  $\alpha_{\text{intra}} = 1.5$ ; 对于帧间编码,取  $\alpha_{\text{inter}} = 2.2$ . 虽然在系数均匀分布的情况下,  $\mu$  可取 0.5,但实验表明,小波系数并非均匀分布,并且其分布特性与帧内或帧间编码模式有关,因此通过实验统计,本编码器取  $\mu_{\text{intra}} = 0.38, \mu_{\text{inter}} = 0.32$ .

### 1.3 基于子集分割的零树编码

小波零树编码被证明是对小波系数进行编码的有效方法. 与 Shapiro 所提出的传统零树编码<sup>[12]</sup>方法相比,本文所采用的基于子集分割的零树小波编码算法 (Subset-Division-based Zerotree Wavelet, SDZW), 在小波系数的树形结构、扫描系数的方式等方面进行了改进,并提出了幅值排序树的方法,以加快零树扫描判决速度<sup>[10]</sup>.

在零树编码中,最重要的是找出重要的系数和大块的“零值”系数. 而 SDZW 算法判断系数重要性的准则如下: ① 给定  $n, n$  是大于零的整数,对于系数  $C_{i,j}$ , 如果  $|C_{i,j}| \geq 2^n$ , 则称这个系数是重要的,否则为不重要的系数; ② 对于系数子集  $S_m$ , 如果

$\max(\{|C_{i,j}|\}) \geq 2^n, (i,j) \in S_m$ , 则称这个系数子集是重要的, 否则是非重要的; ③如果系数子集是重要的, 则可将  $S_m$  继续分成几个新子集, 并对这些新子集进行同样的判断, 如此一直重复下去, 就能找出所有的单个重要系数.

设  $c(i,j)$  为节点  $(i,j)$  所有子女的集合;  $g(i,j)$  为节点  $(i,j)$  后代的集合;  $d(i,j)$  为节点  $(i,j)$  孙辈的集合; 其中,  $g(i,j) = c(i,j) + d(i,j)$ . 图 2 显示了节点的对应关系. 文献[10]中子集分割小波系数的规则如下: ①系数  $(i,j)$  与阈值比较判断时, 如果  $g(i,j)$  具有重要的系数, 那么就将其分割为  $d(i,j)$  和 4 个元素  $(k,l) \in c(i,j)$ ; ②如果  $d(i,j)$  具有重要的系数, 那么就把  $d(i,j)$  分成 4 个子集  $g(k,l)$ , 其中,  $(k,l) \in c(i,j)$ .

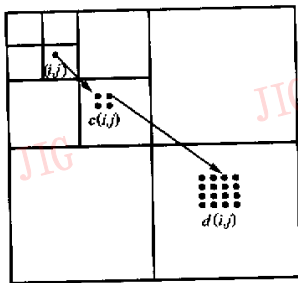


图 2 节点对应关系

在 Shapiro 零树编码中, 大量的比特开销在定位信息上, 而文献[10]提出将由小波系数构成的树形结构与上面的判断子集分割算法结合起来, 使解码器能根据判决的路径获得系数的位置, 从而减少用于定位信息的比特数, 实践证明, 其在限定码率的条件下, 能为高频中的大系数分配更多的比特, 使得恢复图象有更好的主客观质量.

## 2 实验结果

本文用 QCIF 格式序列 Carphone 和 CIF 格式序列 Hall, 对本文提出的基于区域分割的 MRME 算法和基本 MRME 算法的速度和运动矢量编码效率进行了比较实验, 结果见表 1. 其中, 方案 I 代表文献[3]中提出的基本的 MRME 算法; 方案 II 代表本文提出的基于区域分割的 MRME 算法.

表 1 多分辨率运动估值实验结果比较

序列	方案 I		方案 II	
	运动矢量编码效率	总耗时 (s)	运动矢量编码效率	总耗时 (s)
Carphone	397.26	10.37	255.38	8.42
Hall	269.41	30.37	164.51	6.79

表 1 给出了每个方案中运动矢量编码在码流中所占用的字节数, 以及在执行运动估值时所有帧耗费的总时间, 包括运动估值和运动补偿时间, 方案 II 则包括用运动检测器进行区域分割的时间. 实验是在 K6-266, 16×6M 内存的微机上的进行的. 实验时, 以 10 帧/s 的帧率, 对前 40 帧进行编码. 对于 QCIF 格式, 本文采用 2 层小波分解; 对于 CIF 格式, 采用 3 层小波分解. 从表 1 中可以看到:

(1) 方案 II 中, CIF 格式的 Hall 序列所耗的时间比 QCIF 格式 Carphone 序列所耗的时间还要少. 可见, 由于应用运动检测器来将视频帧分割为运动区域和静止区域, 编码器不再对静止区域进行运动估值和补偿, 因此所耗的时间从由图象大小决定变成主要由运动区域的大小来决定. 由于运动发生的局部性, 故可以有效地提高编码的速度.

(2) 由区域分割结果来指导运动估值过程, 可以有效地降低运动矢量编码所需的比特数, 进而提高运动矢量的编码效率.

采用本文提出的编码算法对 CIF 格式的测试序列 Miss America, Salesman 和 Claire 的前 20 帧亮度分量进行了编码(帧率为 10 帧/s), 并与其他采用 MRME 算法的编码器在相同条件下的编码结果进行了对比(如表 2 所示). 其中, MRME 代表文献[3]中编码器, AMRME 代表文献[6]中的编码器, JMRME 则代表文献[8]中的编码器.

表 2 编码效果比较

算法	Miss America		Salesman		Claire	
	Cr	PSNR(dB)	Cr	PSNR(dB)	Cr	PSNR(dB)
MRME	97	35.45	64	28.7	72	36.98
AMRME	97	36.55	64	29.7	72	37.80
JMRME	105	37.35	60	30.2	75	38.77
本文算法	97	38.05	64	30.71	72	39.20
	105	37.67	60	31.37	75	39.01

从表 2 中可以看出, 在码率相近或相同的情况下, 本文提出的编码器重建图象的 PSNR 值比其他算法均有提高.

### 3 结束语

实验结果表明,本文所提出的基于区域分割的多分辨率运动估值算法,在性能上优于基本的MRME算法,编码器的性能优于文献[3]、[6]、[8]中引用的其他采用多分辨率运动补偿方案的编码器。在编码性能和编码速度方面,本文提出的编码算法较已有方法均有不同程度的提高。今后将在此基础上,实现码流的多种可分级性。

### 参考文献

- 肖自美. 图像信息理论与压缩编码技术[M]. 广州:中山大学出版社,2000:282~304.
- Lewis S, Knowles G. Video compression using 3D wavelet transforms[J]. Electron. Lett., 1990,26(6):1184~1185.
- Zhang Y, Zafer S. Motion-compensated wavelet transform coding for color video compression[J]. IEEE Trans. Circuits and Systems for Video Technology, 1992,2(3):285~296.
- Seongman K, Seunghyeon R. Interframe coding using two-stage variable block-size multiresolution motion estimation and wavelet decomposition[J]. IEEE Trans. Circuits and Systems for Video Technology, 1998,8(4):399~409.
- Mandal M, Chen X Q. Multiresolution motion estimation technique for video compression[J]. Optical Engineering, 1996,35(1):128~135.
- Wei J, Li Z. An enhancement to MRMC scheme in video compression[J]. IEEE Trans. Circuits and Systems for Video Technology, 1997,7(3):564~568.
- 许洁斌,韦岗,布礼文. 一种基于形态变换的小波分解运动补偿编码新算法[J]. 中国图象图形学报,1998,3(9):727~733.
- Zhang X, Wang D. A video coding algorithm based on intraframe and interframe joint prediction of wavelet coefficients [J]. Journal of China Institute of Communications, 1998,19(5):11~17.
- Shen K, Delp E J. Wavelet based rate scalable video compression[J]. IEEE Trans. Circuits and Systems for Video Technology, 1999,9(1):109~122.
- 洪建林. 高分辨率图像小波压缩方法的研究[D]. 广州:中山大学电子与通信工程系,1997.
- 刘红梅,黄金科,肖自美等. 改进的多分辨率运动估值算法[J]. 中山大学学报(自然科学版),2001,(3):34~38.
- Shapiro J M. Embedded image coding using zerotrees of wavelet coefficients[J]. IEEE Trans. Signal Processing, 1993,41(12):3445~3462.

刘红梅 1969年生,中山大学电子与通信工程系讲师,1992年、1996年先后获清华大学计算机系学士学位和硕士学位,2000年获中山大学电子与通信工程系博士学位。主要研究方向为多媒体信息处理、视频压缩编码、图象/视频水印技术。

黄金科 1978年生,2000年获中山大学电子系理学学士学位,现为广州石竹科技有限公司助理工程师。主要工作方向是计算机软件、通信系统等。

肖自美 1938年生,现任中山大学电子系教授,中山大学信息与通信技术研究中心主任。主要从事图象压缩编码与处理、语音压缩与处理、多媒体技术与通信系统、数字传输与调制技术等领域的研究和教学工作。

郑慧诚 1974年生,1999年获中山大学电子系硕士学位,现为广东省电信公司广州市分公司数据通信分局助理工程师。主要工作方向为多媒体网络管理。

梁凡 1972年生,理学博士,中山大学电子系讲师。主要研究领域是图象压缩编码、无线移动网络可视多媒体通信系统等。