

# 序列图象编码系统中基于小波变换的运动估计与运动补偿方法

李 涛

俞斯乐

(中国航天工业总公司 8357 研究所, 天津 300141) (天津大学电视与图象信息研究所, 天津 300072)

**摘 要** 提出了一种基于小波变换的块大小可变的分辨率运动估计方法(VBMRME), 以减少图象间的时间冗余。首先估计最高层子图象的运动矢量, 低层子图象的运动矢量由高层子图象的运动矢量经细化导出, 子图象内每一块的大小随小波分解层数而变化, 这大幅度地减少搜索和匹配时间。

**关键词** 小波变换, 运动估计

## 1 引言

在图象编码领域里, 小波变换由于能够很好地表达非稳态图象信号及其与人眼的视觉特性相匹配, 近年来得到了人们越来越多的重视, 它将非稳态的信号分解成一系列相对平稳的子图象, 这有利于进行编码, 同时, 编码方案和参数可以依据每层子图象的特性进行调整, 因此编码每一稳态子图象要比编码整个非稳态图象来得容易。

序列图象在时间轴上具有较强的相关性, 运动估计(ME)及运动补偿(MC)技术可以有效地减小时间相关性, 因此该技术被广泛应用于各种视频压缩编码方案中。运动估计用来估计物体的位移, 得到运动矢量; 运动补偿根据得到的运动矢量, 对前一帧中由运动而产生的位移进行调整, 从而得到尽可能接近本帧的预测帧。由此可见, 运动估计算法越完善, 估计出的运动矢量越准确, 运动补偿的性能就越好, 从而使预测误差越小, 编码后需要传输的信息量也将随之大大减少, 整个系统的码率压缩比得到很大提高, 因此运动估计和补偿技术已成为序列图象编码系统中提高码率压缩比的重要技术。

## 2 小波理论简介

小波分解是一个给定信号  $f(x)$  在不同分辨率的信号空间的连续投影, 设  $V_j$  为分辨率  $2^{-j}$  的信号空间,  $A_j$  是信号在分辨率  $2^{-j}$  的空间的投影, 其中  $j$  是非负整数。在满足一定的条件下, 存在一个尺度函数  $\phi(x)$ , 则  $\{\sqrt{2^j}\phi_{j,n}(x)\}$  是  $V_j$  的正交基。  $A_j$  可以表示为:

$$A_j f(x) = \sum_{n=-\infty}^{+\infty} \langle f(u), \sqrt{2^j}\phi_{j,n}(u) \rangle \sqrt{2^j}\phi_{j,n}(x) \quad (1)$$

其中离散信号  $A_j^d f(u) = \langle f(u), \phi_{j,n}(u) \rangle$  称为  $f(x)$  在分辨率  $2^{-j}$  的离散逼近。由于  $V_j \subset V_{j-1}$ , 设  $O_j$  是  $V_{j-1}$  中  $V_j$  的正交补, 即

$$V_{j-1} = V_j \oplus O_j \quad (2)$$

在满足一定的条件下, 存在一个函数  $\psi(x)$ , 称为小波函数, 则  $\{\sqrt{2^j}\psi_{j,n}(x)\}$  构成  $O_j$  的正交基。于是, 信号  $f(x)$  在  $O_j$  的投影  $D_j$  为:

$$D_j f(x) = \sum_{n=-\infty}^{+\infty} \langle f(u), \sqrt{2^j}\psi_{j,n}(u) \rangle \sqrt{2^j}\psi_{j,n}(x) \quad (3)$$

而离散信号

\* 本课题获国家自然科学基金资助

收稿日期: 1996-12-16; 收到修改稿日期: 1997-10-05

$$D_j^d f(x) = \langle f(u), \psi_{j,n}(u) \rangle \quad (4)$$

称为  $f(x)$  在分辨率  $2^{-j}$  的离散细节信号, 由于  $\{\sqrt{2^j} \phi_{j,n}(x)\}$  和  $\{\sqrt{2^j} \psi_{j,n}(x)\}$  的正交性,  $A_j^d f$  和  $D_j^d f$  可直接从  $A_{j-1}^d f$  中导出。

$$A_j^d f(n) = \sum_{k=-\infty}^{+\infty} h(k-2n) A_{j-1}^d f(k) \quad (5)$$

$$D_j^d f(n) = \sum_{k=-\infty}^{+\infty} g(k-2n) A_{j-1}^d f(k) \quad (6)$$

上述的分解过程可以不断重复直至分辨率  $2^{-j}$ , 因此信号  $f(x) \in V_0$  可以连续分解

$$V_0 = O_1 \oplus O_2 \oplus \dots \oplus O_j \oplus V_j \quad (7)$$

相反, 由于  $V_{j-1} = V_j \oplus O_j$ , 可以由  $A_j^d f$  和  $D_j^d f$  恢复  $A_{j-1}^d f$ 。

$$A_{j-1}^d f(n) = 2 \sum_{k=-\infty}^{+\infty} h(n-2k) A_j^d f(k) + 2 \sum_{k=-\infty}^{+\infty} g(n-2k) D_j^d f(k) \quad (8)$$

二维情况下的小波分解如图 1 所示, 根据 Daubechies 算法可以用 FIR 滤波器实现 H 和 G。利用小波变换, 图象  $A_{j-1} f$  分解为相互正交的子图象  $A_j f, D_j^1 f, D_j^2 f$  和  $D_j^3 f$ ,  $A_j f$  对应  $A_{j-1} f$  的低频部分, 而  $D_j^1 f, D_j^2 f$  和  $D_j^3 f$  对应于  $A_{j-1} f$  和  $A_j f$  之间的差值信号, 在频率域它们表示相互独立且具有空间方向性的频带。

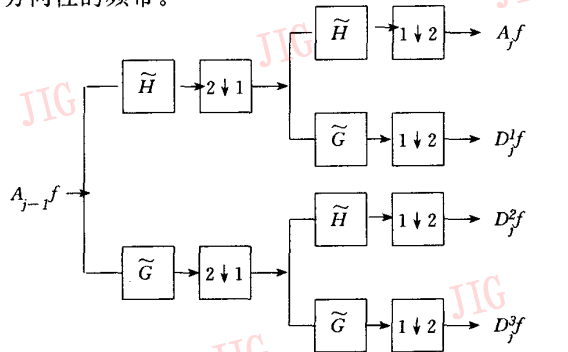


图1 二维小波分解

Fig. 1 Two-dimension wavelet decomposition

### 3 运动估计和运动补偿

根据小波多分辨率分解的特点, 一帧图象分解为若干层子图象, 分别具有不同的分辨率和频带范围及宽度, 因此, 运动形式在不同子图象是不同的, 但因为它们实质上表达的是同一物体的运动, 所以, 这些运动形式是高度相关的, 根据这个性质, 可采用

块大小可变的分辨率运动估计方案 (Variable Block-size Multiresolution Motion Estimation)。

在本方案中, 首先计算最高层低频子图象  $A_j f$  中每一宏块  $p \times p$  的全搜索运动矢量  $V_1(x, y)$ , 它的运动搜索范围  $\Omega$  为  $dm \times dm$ , 对应于整幅图象中  $2^j dm \times 2^j dm$  的搜索范围, 然后, 这个运动矢量乘以尺度作为其它层子图象内宏块的运动矢量的初始值, 各层之间的尺度大小为 2, 各层子图象内宏块的运动搜索范围仍为  $dm \times dm$ , 但各层子图象内宏块的大小随着层数的降低成倍地增加, 同时, 搜索范围相对减小了。

为实现多分辨率运动估计方案, 共有 4 种可选的方法:

(1) 求出子图象  $A_j f$  内宏块的运动矢量, 不再进行细化, 乘以尺度后直接作为其它层对应宏块的运动矢量。

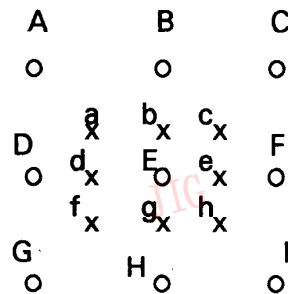
(2) 计算出子图象  $A_j f, D_j^1 f, D_j^2 f, D_j^3 f$  内各宏块的运动矢量, 乘以尺度后作为其它对应宏块的运动矢量。

(3) 计算出子图象  $A_j f$  内各宏块的运动矢量, 乘以尺度后, 作为其它子图象对应宏块的初始运动矢量并进行细化。

(4) 求出子图象  $A_j f, D_j^1 f, D_j^2 f, D_j^3 f$  内各宏块的运动矢量, 乘以尺度以后, 作为其它子图象对应宏块的初始运动矢量并进行细化。

以上 4 种方法中 (1) 和 (2) 计算比较简单, 但效果不太好, (3) 和 (4) 方法比较复杂, 但效果好, 经过比较筛选, 我们选用方法 (3)。

另外, 为了得到更精确的运动矢量, 在进行整像素的运动估计后, 需要采用半像素精度的运动估计。



$$a = (A+B+D+E)/4 \quad b = (B+E)/2 \quad d = (D+E)/2$$

图 2 半像素精度运动估计示意图

Fig. 2 Diagram of half-pixel precision motion estimation

### 4 系统方案和码流层结构

#### 4.1 系统方案

图象数字化后的数据量是非常大的,例如,对于 HDTV 图象,一帧大小为  $1440 \times 1152$ ,若每秒钟传输 25 帧,则所需的码流将高达 1Gb/s。按目前我国

专家建议的数字 HDTV 预选参数,需压缩到 21.5Mb/s,这就要求编码技术能尽量消除图象的冗余度。本方案主要是从时域和空域着手:运用运动估计 (ME) 和运动补偿 (MC) 去掉时间上的相关性;运用离散小波变换 (DWT), DCT, 视觉特性加权和熵编码消除空间的冗余度。图 2 是本系统方案的编码器框图。

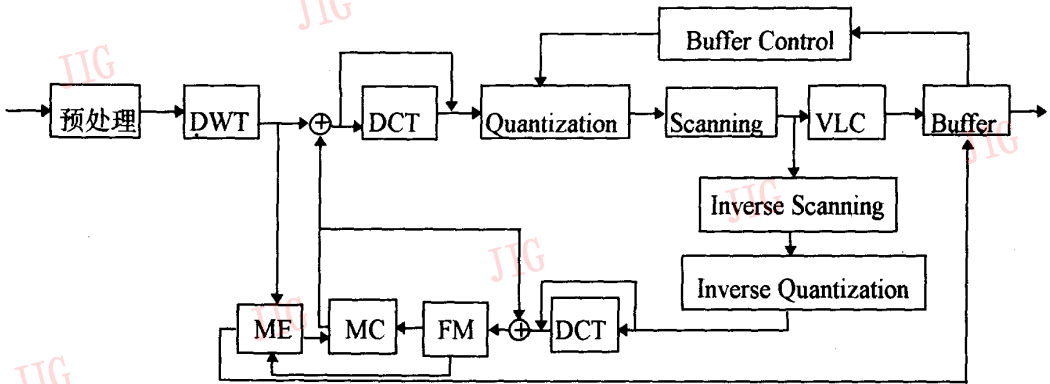


图 3 编码器框图

Fig. 3 Diagram of image sequence coding system

#### 4.2 码流层结构

为了便于图象编码处理,防止量化噪声的累积,需将一个序列图象分为以下 7 个层次:(1)序列图象层 (Sequence 层), (2)图象组层 (Group of Picture, GOP 层), (3)图象层 (Picture 层), (4)子图象层 (Sub-Picture 层), (5)条层 (Slice 层), (6)宏块层 (Macroblock 层) 和 (7)块层 (Block 层)。每一层都有自己单独的码头,关键参数和数据,层层镶嵌,是典型的一层套一层的树状结构。

在本方案中,将一个序列图象中的连续 12 帧划为一个 GOP,对于每帧图象有 3 种编码模式:I 帧 (帧内), P 帧 (前向预测) 和 B 帧 (双向预测)。

根据小波变换理论,只有当小波层数达到或超过 3 层时,它的效果才能超过传统的 DCT 变换,但层数越多,计算的复杂度也加大。因此,我们设计小波变换为 3 层,在图象内部,子图象的位置安排如图 4。

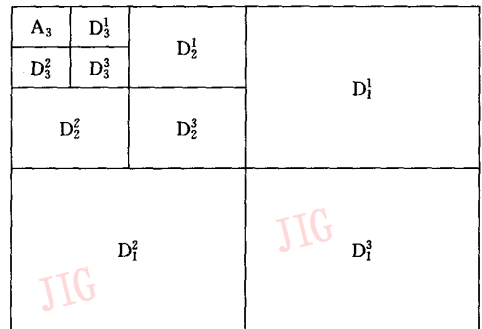


图 4 系统中各子图象的位置安排

Fig. 4 Disposition of sub-images after wavelet transform

### 5 模拟实验结果

为了验证基于小波变换的序列图象编码系统中运动估计的有效性,进行了模拟实验,并对解码恢复图象作了客观评价。本实验选用 2 个序列各 50 帧作为信源,包括 Renata 序列和 SKI 序列。客观测试是利用恢复图象和原始图象的峰-峰信噪比来衡量图象的好坏。图 5 和图 6 分别为 2 个序列图象的信噪比曲线 (亮度信号) 及每帧的实际比特数。由图可见,在压缩比约为 50~60 倍的情况下,信噪比都在 30dB 以上。

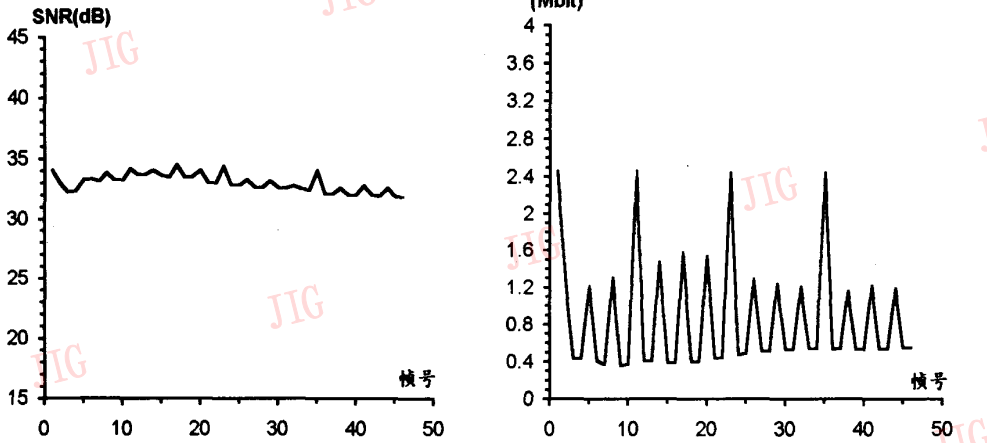


图5 Renata 的 Y 信号信噪比和比特数

Fig. 5 Luminance SNR and compressed bits for "Renata" Sequence

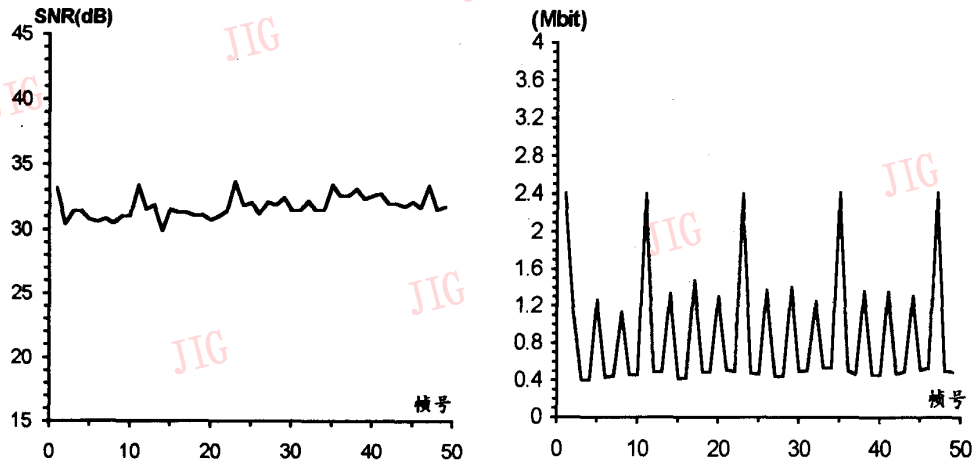


图6 Ski 的 Y 信号信噪比和比特数

Fig. 6 Luminance SNR and compressed bits for "SKI" Sequence

## 6 结束语

运动估计与补偿是为了消除序列图象的时间相关性,它是图象编码系统中最耗时的部分。至今,在 MPEG-2 图象编码标准中,这一部分仍属于进一步探讨的领域。根据目前的块匹配算法,当作逐点搜索时,系统负担大,硬件实现困难;而采用改进的搜索算法时,搜索精度又无法保证,很难找到最佳匹配点。本文提出的算法(VBMRME)可以克服以上的矛盾,在最高层子图象中搜索时,搜索范围为  $12 \times 12$ ,对于三层小波变换,这相当于整幅图象中  $96 \times 96$  的搜索范围,而且搜索时间短。为了提高搜索精

度,其它各层子图象在得到初始运动矢量之后,仍在  $12 \times 12$  范围内搜索,并且搜索精度为半像素,这样就能保证找到最佳的匹配点。

## 参考文献

- 1 Antonini M, Barland M, Mathieu P et al. Image Coding Using Wavelet Transform. IEEE Trans. on Image Processing, 1992, 1 (2): 205~220.
- 2 Cetin A, erek O, Ulukus S. Block Wavelet Transform for Image Coding. IEEE Trans. on Circuits and Systems for Video Technology, 1993, 3(6): 433~435.
- 3 Daubechies I. Orthonormal Bases of Compactly Supported Wavelets. Communication on Pure and Applied Mathematics, 1988, 41: 909~996.

- 4 Dvornik R, Jawerth B, Lucier B. Image Compression Through Wavelet Transform Coding. IEEE Trans. on IT. ,1992,38(2): 791~746.
- 5 Mallat S. A Theory for Multiresolution Signal Decomposition : The Wavelet Representation . IEEE Trans. on PAMI. , 1989, 11(7):674~693.
- 6 ISO/CD 13818-2 ITU-T Draft Recommendation H. 262 ,Information Technology-Generic Coding of Moving Pictures and Associated Audio Information ,1993.
- 7 Naveen T, Woods J. Motion Compensated Multiresolution Transmission of High Definition Video . IEEE Trans. on Circuits and Systems for Video Technology,1994,4(1):29~41.
- 8 Puri A, Aravind R, Haskell B. Adaptive Frame/Field Motion Compensated Video Coding. Signal Processing: Image Communication,1993,5(1~2):39~58.



李涛,1967年出生,1990年和1993年在天津大学计算机科学与工程系分别获软件专业学士学位和计算机组织与体系结构专业硕士学位,1996年获天津大学电子工程系信号处理专业博士学位,现任职于中国航天部8357研究所,主要从事图象处理,图象编码和小波变换等方面的研究。

## A New Method of Motion Estimation and Motion Compensation Based on Wavelet Transform in Image Sequence Coding

Li Tao

(1 The 8357 Research Institute of China Space Industry Corporation , Tianjin 300141)

Yu Sile

(2 Institute of Television and Image Information , Tianjin 300072)

**Abstract** A scheme of variable block-size Multiresolution Motion Estimation(VBMRME) is presented in this paper, which reduces temporal redundancy between pictures of the image sequence. Motion vectors in the highest level of the wavelet pyramid are first estimated. Then they are refined for the prediction of motion vectors in the lower levels. The size of a block in each sub-image is adapted to its level in the wavelet pyramid, so that the searching and matching time can be considerably reduced.

**Keywords** Wavelet transform , Motion estimation