

# 视频监控中的多尺度小波变换编码算法分析

王燕妮, 王慧琴, 周军妮

(西安建筑科技大学信息与控制工程学院, 陕西 西安 710055)

**摘要:**为了解决视频监控系统中信息快速、准确地传输问题,提出一种小波分解与重构的多尺度分析变换编码算法.由于相邻视频背景像素之间存在着很大的相关性,可采用预测匹配思想对运动变化幅度较小或无位移的背景环境进行编码;对于活动的、变化较快的人流区域,利用小波函数与尺度函数之间的正交性,通过多级递归逐渐改变视窗大小的方法,分步实行匹配,找到最优的运动矢量,得到尺度空间上的平滑逼近和小波空间的离散细节信号.与经典的预测匹配算法相比,仿真结果表明,该算法可取得较高的压缩比及质量较好的视频通信效果.

**关键词:**视频监控;多尺度分析;小波函数;变换编码

**中图分类号:** TN 919

**文献标志码:** A

**文章编号:** 1006-7930(2013)01-0141-06

视频信号编码技术随着数字通信、视频监控系统的发展受到了越来越多的关注.视频对象信息能够实时、准确高效地被传输是视频监控系统的核心,其编码技术也成为当前国内外视频安防事业的核心技术.

目前,在基于对象的视频编码技术中,运动估计编码<sup>[1-3]</sup>适用于对相关信源的各种编码要求,但由于高效、实时视频压缩的巨大计算量,其编码技术不能满足实时性和精度的要求;而各种多尺度分析<sup>[4-5]</sup>、形状及纹理编码技术<sup>[6-7]</sup>在提高压缩效率的同时,能很好地兼顾减少视频编码时间,但在传输速度很高时会出现边缘模糊效应,或者音视频不同步现象.对现有视频监控系统存储器的存储容量、通信干线的信道传输率而言,采用先进的视频压缩技术将视频信息进行压缩,非常必要.本文提出视频图像传输中的多尺度小波变换编码(MSWTC, Multi-scale wavelet transform coding)算法,采用多尺度分析分解、重构视频对象信号,通过仿真实验,达到既节省存储空间,又可以提高视频的传输质量及效率.

## 1 多尺度分析

多尺度分析是用小波函数的二进伸缩和平移表示函数这一思想的更加抽象和复杂的表现形式,它重点处理整个函数空间,而非侧重处理作为个体的函数.

将平方可积函数空间 $L^2(R)$ 用它的子空间 $V_j, W_j$  (其中 $j \in Z$ )表示,其中 $V_j, W_j$ 分别称为尺度空间和小波空间<sup>[8]</sup>.

尺度空间 $V_j$ 也称为逼近空间,具有递归嵌套关系:

$$\cdots \subset V_{+1} \subset V_0 \subset V_{-1} \subset \cdots \quad (1)$$

小波空间 $W_j$ 是 $V_j, V_{j-1}$ 之间的差,即

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

则 $W_j$ 是弥补 $V_j$ 逼近 $V_{j-1}$ 时丢失的信息.

将尺度空间做逐级二分分解产生一组逐级包含的子空间:

$$\cdots, V_0 = V_1 \oplus W_1, V_1 = V_2 \oplus W_2, \cdots, V_j = V_{j+1} \oplus W_{j+1}, \cdots \quad (3)$$

$j$ 是从 $-\infty$ 到 $+\infty$ 的整数, $j$ 值越小空间越大.

收稿日期:2012-04-11 修改稿日期:2013-01-10

基金项目:陕西省教育厅专项科研项目(11JK1023);陕西省教育厅产业化项目(2011JG12)

作者简介:王燕妮(1975-),女,陕西蒲城人,博士,副教授,主要从事视频压缩、图像处理等信号与信息处理方面的研究.

当  $j \rightarrow -\infty$ ,  $V_j \rightarrow L^2(R)$ , 包含整个平方可积的实变函数空间; 当  $j \rightarrow +\infty$ ,  $V_j \rightarrow 0$ , 即空间最终逼近到空集为止.

这样空间  $V_j$  与空间  $W_j$  正交, 各个  $W_j$  之间也正交, 即:

$$V_j \perp W_j; W_j \perp W_{j'}, j \neq j' \quad (4)$$

若  $f(t) \in L^2(R)$  是逐级逼近的极限, 每次逼近都是用一平滑函数  $\varphi(t)$  对  $f(t)$  做平滑的结果, 在逐级平滑时函数  $\varphi(t)$  也做逐级逼近, 并以此来逐级逼近待分析函数  $f(t)$ .

函数空间  $V_j$  具有以下特性<sup>[9-10]</sup>:

位移不变性. 函数的时移不改变其所属空间, 即

$$\text{如果 } f(t) \in V_j, \text{ 则 } f(t-k) \in V_j \quad (5)$$

二尺度伸缩性由下式表明:

$$\text{若 } f(t) \in V_j, \text{ 则 } f(t/2) \in V_{j+1}, f(2t) \in V_{j-1} \quad (6)$$

Riesz 基存在性. 存在  $\varphi(t) \in V_0$ , 使得  $\{\varphi(2^{-j}t-k), k \in Z\}$  构成  $V_j$  的 Riesz 基.

函数  $\varphi(t) \in L^2(R)$  为尺度函数, 若其经过整数平移  $k$  和尺度  $j$  上的伸缩, 得到一个尺度和位移均可变化的函数集合:

$$\varphi_{jk}(t) = 2^{-\frac{j}{2}} \varphi(2^{-j}t-k), j, k \in Z \quad (7)$$

称每一个尺度  $j$  上的平移系列  $\varphi_{jk}(t)$  所组成的空间  $V_j$  为尺度为  $j$  的尺度空间.

$$V_j = \text{span} \{ \overline{\varphi_{jk}(t)} \}, k \in Z \quad (8)$$

对于任意函数  $f(t) \in V_j$ , 有

$$f(t) = \sum_k a_k \varphi_{jk}(t) = 2^{-\frac{j}{2}} \sum_k a_k \varphi(2^{-j}t-k) \quad (9)$$

所以, 尺度函数在不同尺度下其平移系列组成了一系列的尺度空间.

$L^2(R)$  中  $W_j$  的一个标准正交基为:

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t-k), j, k \in Z \quad (10)$$

同一尺度下, 因为  $W_j \perp V_j$ , 所以小波函数和尺度函数之间是正交的, 即:

$$\int \psi_{jk}(t) \cdot \varphi_{jk}^* dt = 0 \quad (11)$$

若当前视频图像所选宏块的尺度为  $j$  (或  $2^j$ ), 所获得对象目标的信息是  $V_j$ , 当逼近对象时, 即尺度增加到  $j-1$  (或  $2^{j-1}$ ), 则所获得的对象信息为  $V_{j-1}$ , 应该比尺度  $j$  下获得的信息更加丰富, 即

$$V_j \subset V_{j-1} \quad (12)$$

则尺度越高, 距离越近; 反之, 则相反.

由于  $V_0 = V_1 \oplus W_1$ , 对于任意输入视频信号  $f(t) \in V_j$ , 有

$$A_0 f(t) = A_1 f(t) + B_1 f(t) \quad (13)$$

其中  $A_1$  表示  $V_j$  的权重系数;  $B_1$  表示  $W_j$  的权重系数. 因此, 输入信号函数  $f(t) \in V_j$  在  $V_j$  中尺度不断减小到  $j=0$  时, 得到平滑逼近.

## 2 视频监控中的多尺度小波编码算法

在视频图像中, 由于同一图像内部相邻像素之间运动矢量具有相同或相似的特征, 在进行匹配时, 可根据当前像素块左边块的运动矢量, 向右平移一个像素块的方法, 预测出当前像素块的运动矢量. 对于对象轮廓处的宏块像素则可采用可变宏块搜索法进行匹配, 这样就减少了搜索像素点数目及搜索时间.

经典的块匹配法因为块的尺寸是一定的, 所以选择适合大小的宏块很重要. 如果选择的宏块太小, 大范围的视频场景则不能预测其实际的位移; 相反, 宏块太大, 则可能无法找到精确的估计值. 采用可变宏块大小的多尺度小波分析编码算法, 通过分步实行块匹配, 每个步骤选取大小不同的宏块, 用一个趋于减小宏块尺寸的过程来实现运动估计, 并得到误差小, 且向量一致的矢量场.

由于  $\varphi_{j0}(t) = \frac{1}{2^{\frac{j-1}{2}}} \varphi(\frac{t}{2^j}) \in V_j$ ,  $V_{j-1} \subset V_j$ ,  $\varphi_{j-1,k}(t)$  是  $V_{j-1}$  中的正交基, 所以  $\varphi_{j0}(t)$  可表示为  $\varphi_{j-1,k}(t) = \frac{1}{2^{\frac{j-1}{2}}} \varphi(\frac{t}{2^{j-1}} - k)$  的线性组合,  $h_{0k}$  是权重系数, 即

$$\varphi_{j0}(t) = \sum_k h_{0k} \varphi_{j-1,k}(t) \quad (14)$$

$$\frac{1}{2^{\frac{j}{2}}} \varphi(\frac{t}{2^j}) = \sum_k h_{0k} \frac{1}{2^{\frac{j-1}{2}}} \varphi(\frac{t}{2^{j-1}} - k) \quad (15)$$

$$\varphi(\frac{t}{2^j}) = \sqrt{2} \sum_k h_{0k} \varphi(\frac{t}{2^{j-1}} - k) \quad (16)$$

根据多尺度分析理论, Mallat 提出了小波分解与重构的快速算法, 称为 Mallat 算法. 在此, 利用 Mallat 算法对输入的图像像素信号进行分解和重构.

由于  $A_j f(t) \in V_j$ , 所以

$$A_j f(t) = \sum_n x_n(j) \varphi_{jn}(t) \quad (17)$$

$$j=0 \text{ 时, } A_0 f(t) = \sum_n x_n(0) \varphi_{0n}(t) \quad (18)$$

$A_0 f(t)$  可表示信号  $f(t)$  在尺度函数方向上的逼近;  $x_n(0)$  是线性组合的权重.

$$x_n(0) = \langle A_0 f(t), \varphi_{0n}(t) \rangle \quad (19)$$

则  $x_n(0)$  表示在尺度  $j=0$  下小波函数的离散逼近.

$$x_n(1) = \langle A_1 f(t), \varphi_{1n}(t) \rangle = \langle A_0 f(t) - B_1 f(t), \varphi_{1n}(t) \rangle \quad (20)$$

因为  $B_1 f(t)$  与  $\varphi_{1n}(t)$  正交, 所以

$$x_n(1) = \langle A_0 f(t), \varphi_{1n}(t) \rangle = \langle \sum_n x_n(0) \varphi_{0n}(t), \varphi_{1n}(t) \rangle = \sum_n x_n(0) \langle \varphi_{0n}(t), \varphi_{1n}(t) \rangle \quad (21)$$

由于  $B_1 f(t)$  表示在小波函数  $W_1$  上的逼近, 有

$$B_1 f(t) = \sum_k d_k(1) \psi_{1k}(t) \quad (22)$$

$d_k(1)$  是线性组合的权重,

$$d_k(1) = \langle B_1 f(t), \psi_{1k}(t) \rangle = \langle A_0 f(t) - A_1 f(t), \psi_{1k}(t) \rangle \quad (23)$$

因为  $A_1 f(t)$  与  $\psi_{1k}(t)$  正交, 所以

$$d_k(1) = \langle A_0 f(t), \psi_{1k}(t) \rangle = \sum_k x_k(0) \langle \varphi_{0k}(t), \psi_{1k}(t) \rangle \quad (24)$$

这样,  $j \rightarrow 0$ , 分别得到  $V_j$  方向上的平滑逼近和  $W_j$  上的离散细节信号.

当进行信号重建时, 由  $V_{j-1} = V_j \oplus W_j$

$$A_{j-1} f(t) = A_j f(t) + B_j f(t) = \sum_k x_k(j) \varphi_{jk}(t) + \sum_k d_k(j) \psi_{jk}(t) \quad (25)$$

则  $j \rightarrow \infty$  时, 得到重建信号.

选择在参考帧中分割对象的对应宏块作为起点中心位置, 计算当前宏块的绝对误差值, 如果可以在预测帧中找到相同的宏块, 则直接编码传输. 对轮廓边界和轮廓外的宏块, 采用大宏块先进行绝对误差值比较, 如果没有找到匹配, 然后进行小宏块的绝对误差值比较.

采用分层多尺度块匹配搜索法, 在每级搜索层上采用不同尺寸的像素块, 通过递归法估计运动矢量. 则首先采用尺寸较大的视频窗, 估计出大范围运动. 当进入第二层搜索时, 视频窗尺寸变小, 可进行精确的估计. 转向最后一层时, 视频窗也收缩到最小, 可找到最优的运动矢量. 因此, 经过综合考虑, 采用倒置均值金字塔型结构能实现较好的运动估计.

对于已经分割的视频对象, 在其边缘轮廓区域, 取三层倒置金字塔结构模型, 第一层以  $8 \times 8$  像素块为单元进行分层, 第二层则以  $4 \times 4$  像素块为单元进行分层, 最后一层则为  $2 \times 2$ .

### 3 仿真分析

为了比较 MSWTC 算法的效果,采用视频录像序列对经典的全搜索算法(FS, Full Search)、菱形搜索算法(DS, Diamond Search)以及 MSWTC 算法在相同条件下进行计算机仿真。

#### 实验 1 重建帧的比较

选取一视频录像段的某一帧作为序列的第一帧参考帧,起始宏块的大小为  $8 \times 8$ 。分别用 DS 算法以及 MSWTC 算法得到第二帧的重建帧,从主观上分析,MSWTC 算法重建的视频图像非常接近原始帧,结果如图 1 所示。



图 1 重建帧的比较

Fig. 1 Comparison of rebuild frames

#### 实验 2 误差帧的比较

选取以上视频录像段的第五帧作为序列的参考帧,块的大小采用第一级为  $8 \times 8$ ,然后分级减小至  $2 \times 2$ 。分别用 DS 算法及 MSWTC 算法得到下一帧的误差帧均方误差(MSE, mean square error)。图 2 为 DS 算法和 MSWTC 算法产生的误差帧。

MSE 由下式决定:

$$MSE = \frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M [f(i, j) - g(i, j)]^2 \quad (26)$$

其中  $f(i, j)$  是视频图像序列的原始帧在  $(i, j)$  点处的像素灰度值,  $g(i, j)$  是利用运动矢量重建帧在  $(i, j)$  点处的像素灰度值。

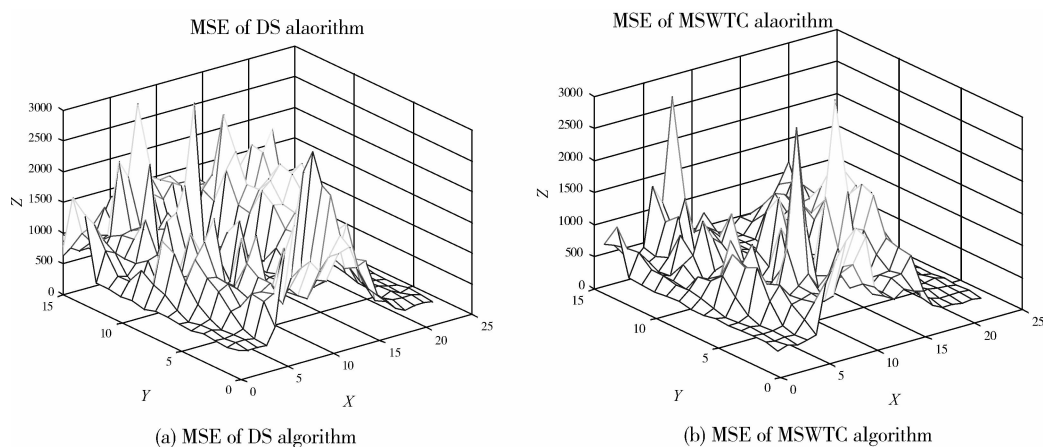


图 2 误差帧的比较

Fig. 2 Comparison of error frames

### 实验3 MSE和PSNR的比较

采用以上视频录像段的第一帧为参考帧,搜索范围  $P$  为 7. 分别用 FS 算法、DS 算法及 MSWTC 算法恢复下一帧,依次传输 100 帧视频,得到误差  $MSE$  和相应的峰值信噪比  $PSNR$  的曲线图,如图 3 所示.

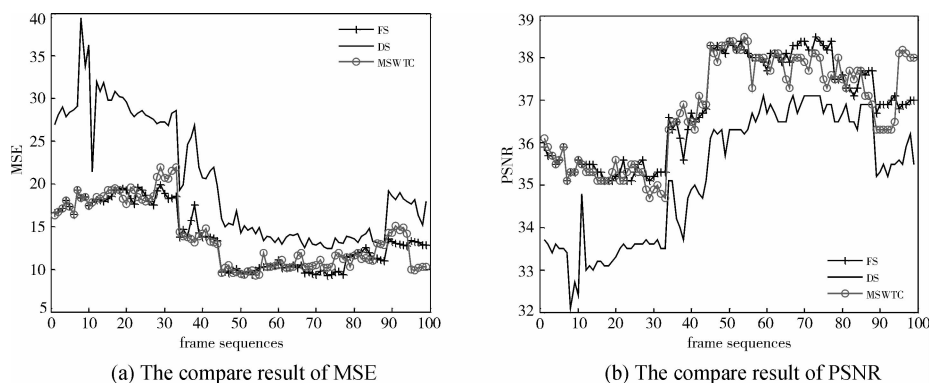


图3 各种算法对视频序列的性能比较

Fig. 3 Performance comparison of different algorithms

### 实验4 搜索匹配时间的比较

采用该视频序列的第五帧作为参考帧,块的大小采用  $8 \times 8$ . 在同样的环境下,分别用 DS 算法及 MSWTC 算法匹配得到恢复的第六帧,算法运行时间为表 1 所示:

在以上实验 1 中,采用视频录像序列的某一帧作为参考帧,用 MSWTC 算法重建其第二帧,从主观上看,MSWTC 算法重建的视频图像非常接近全搜索算法的重建图像. 在实验 2 中,通过误差帧的比较,可以看出对于任取的一段视频序列,MSWTC 算法比 DS 算法产生的误差平均降低了约 800. 在实验 3 中,通过误差和相应的峰值信噪比的比较,MSWTC 算法的性能曲线优于 FS 算法和 DS 算法. 在实验 4 中,用两种算法进行搜索,比较可得改进算法的匹配时间减少了约 68%. 同时对于该视频序列,进行了给定时间内存储信息量的比较,获得新算法的压缩比为 320 : 1,大大提高了监控系统的存储容量及传输效率.

表 1 视频序列的搜索匹配时间表

Tab. 1 The matching time of video sequences

| 搜索算法  | 算法运行时间/s |
|-------|----------|
| DS    | 0.003 2  |
| MSWTC | 0.001 9  |

## 4 结 论

视频图像编码算法是视频监控系统模块中的重要环节,其优劣决定系统整体性能的好坏. 提出多尺度小波变换编码算法并进行分析,利用不同尺度下视窗逐渐减小的方法,实现了对复杂背景下的视频监控录像准确而较快地编码与传输. 与经典算法相比,仿真结果表明,新算法恢复的视频压缩比高,预测质量好. 在实时视频通信中,算法可取得较好的效果.

### 参考文献 References

- [1] MEHRSERESHT N, TAUBMAN D. An efficient content adaptive motion compensated 3-Ddwt with enhanced special and temporal scalability [J]. IEEE Transactions on Image Processing, 2006, 15(6):1397-1412.
- [2] MA Kai-kuang, QIU Gang. Unequal-Arm Adaptive Rood Pattern Search for Fast Block-Matching Motion Estimation in The JVT/H. 26L[C]. IEEE International Conference on Image Processing (ICIP'2003), Barcelona, Spain, 2003(9):14-17.
- [3] ZHAO Fan, LIU Gui-zhong, NAN Nan. A high performance motion mode adaptive lifting motion compensation wavelet video codec [C]// Proceedings of 2007 IEEE International Conference on Multimedia and Expo. Piscataway, NJ, USA: IEEE, 2007:743-746.

- [4] LIU Yu, NGAN K N. Weighted adaptive lifting-based wavelet transform for image coding [J]. IEEE Transactions on Image Processing, 2008, 17(4): 500-511.
- [5] HOU Xing-song, LIU Gui-zhong, ZOU Yi-yang. SAR image data compression using wavelet packet transform and universal-trellis coded quantization [J]. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(11): 2632-2641.
- [6] CHANG Chuo-ling, GIROD B. Direction-adaptive discrete wavelet transfer for image compression [J]. IEEE Transactions on Image Processing, 2007, 16(5): 1289-1302.
- [7] 赵凡, 张娜, 任斐斐, 等. 适合于小波视频编码的尺寸自适应图像组结构 [J]. 西安交通大学学报, 2009, 43(4): 52-56.  
ZHAO Fan, ZHANG Na, REN Fei-fei, et al. An Adaptive Group of Pictures Structure for Wavelet-Based Video Coding [J]. Journal of Xi'an Jiaotong University, 2009, 43(4): 52-56.
- [8] Do M N, Vetterli M. The contourlet transform: an efficient directional multiresolution image representation [J]. IEEE Transactions on Image Processing, 2005, 14(12): 2091-2106.
- [9] Velisavljevic V, Beferull-Lozano B, Vetterli M, et al. Directionlets: anisotropic multi-directional representation with separable filtering [J]. IEEE Transactions on Image Processing, 2006, 15(7): 1916-1933.
- [10] 孙增寿, 范科举. 结构损伤识别的小波熵指标研究 [J]. 西安建筑科技大学学报: 自然科学版, 2009, 41(1): 18-24.  
SUN Zeng-shou, FAN Ke-ju. Research on the wavelet entropy index of structural damage identification [J]. J. Xi'an Univ. of Arch. & Tech.: Natural Science Edition, 2009, 41(1): 18-24.

## Analysis of multi-scale wavelet transform coding algorithm in video surveillance

WANG Yan-ni, WANG Hui-qin, ZHOU Jun-ni

(School of Information and Control Engineering, Xi'an Univ. of Arch. & Tech., Xi'an 710055, China)

**Abstract:** In order to solve the problem of information transmission quickly and accurately in video surveillance system, a multi-scale transform coding algorithm with wavelet decomposition and reconstruction is proposed. Owing to the considerable correlation in adjacent video background pixels, predictive matching can be used to cope with movement changed little or no displacement of the background. Using the Orthogonality between wavelet function and scaling function for the flow area with active and rapid changes, change the window size gradually through the method of multi-level recursion, carry the matching step by step, find the optimal motion vector, and get the smooth approximation on the scale-space and discrete detail signal on the wavelet space. The simulation results show that new algorithm can be wedte achieved higher compression ratio and better results for video communications as compared with the classical algorithm.

**Key words:** video surveillance; multi-scale analysis; wavelet function; transform coding