

文章编号 1004-924X(2012)03-0675-10

## 基于预搜索的高效双目分形视频编码

祝世平\*, 陈菊婧

(北京航空航天大学 仪器科学与光电工程学院 测控与信息技术系, 北京 100191)

**摘要:** 设计了基于预搜索的高效双目分形视频编码, 并成功应用于立体视频编码之中。对基本分形双目视频编码进行了改进。利用了树状划分准则, 起始帧采用块离散余弦变换(DCT)编码, 简化了块搜索范围并提前减少了重复运算。在双目立体视频编码中, 以左通道为基本层, 采用单独的运动补偿预测(MCP)方式进行编码, 充分利用了预搜索限制条件、改进的非对称十字形多层次六边形格点搜索算法、去方块环路滤波和分数像素块匹配算法; 以右通道为增强层, 采用 MCP 加视差补偿预测(DCP)方式进行编码, 选择误差最小的匹配块作为预测结果。在进行 DCP 编码时, 充分利用视差分布约束条件, 提出了快速的视差估计算法。实验结果表明, 提出的编码方法在保证一定的峰值信噪比(PSNR)前提下, 平均压缩时间是基本分形双目视频编码的 18%~23%, 压缩比提高了 15.13~47.49, 显著地改善了基本分形视频压缩算法的性能, 使分形视频压缩的应用具有更大的灵活性和实用性。

**关键词:** 双目分形视频编码; 预搜索; 运动估计; 视差估计; 环路滤波; 分数像素块匹配

**中图分类号:** TN919.81 **文献标识码:** A **doi:** 10.3788/OPE.20122003.0675

## Efficient stereo fractal video coding based on pre-searching

ZHU Shi-ping\*, CHEN Ju-qiang

(Department of Measurement Control and Information Technology, School of Instrumentation Science and Optoelectronics Engineering, Beihang University, Beijing 100191, China)

\* Corresponding author, E-mail: shiping.zhu@buaa.edu.cn

**Abstract:** An efficient stereo fractal video coding based on pre-searching was presented for the stereo video coding successfully. The basic stereo fractal video coding was improved by using tree structure macroblock partition, Discrete Cosine Transform(DCT) based encoding for original frame, reducing block searching range and computational repetition in advance. In stereo fractal video coding, the left view video was arranged as a main view and the Motion Compensation Prediction(MCP) was taken as coding structures by employing pre-searching limitation condition, improved UMHexagonS motion estimation algorithm, deblocking loop filter and sub-pixel block matching. The right view video was arranged as an enhanced view, the MCP and Disparity Compensation Prediction(DCP) were combined to be as coding structures, and the matching block with smallest matching error was chosen as the estimation results. In the DCP coding, a fast disparity estimation algorithm was proposed by making full

收稿日期: 2011-10-19; 修订日期: 2011-12-15.

基金项目: 国家自然科学基金资助项目(No. 61075011, No. 60675018); 教育部留学回国人员科研启动基金资助项目

use of disparity distribution restriction conditions. Experimental results indicate that the compression time by proposed stereo fractal video coding algorithm is 18%–23% that by the basic stereo fractal video coding algorithm, and the compression ratio has improved by 15.13–47.49 while a proper Peak Signal Noise Ratio(PSNR) is maintained. These results show the compression performance of the stereo fractal video coding algorithm and its applications are more flexible and practical.

**Key words:** stereo fractal video coding; pre-searching; motion estimation; disparity estimation; de-blocking loop filter; sub-pixel block matching

## 1 引言

分形图像编码方法是利用图像的局部自相似性来消除图像数据中存在的跨尺度冗余,对图像进行压缩的一种新颖压缩方法。它是目前公认的三种最有前途的新一代图像编码技术之一。Barnsley 于 1988 年首次提出迭代函数系统理论(Iterated Function System, IFS)并将其成功应用于图像压缩领域<sup>[1]</sup>。Jacquin 提出了全自动的分形图像压缩方法<sup>[2]</sup>,该方法采用基于图像分块的方式以局部的仿射变换代替全局的仿射变换。之后,Fisher 利用二叉树改进了这一方法<sup>[3]</sup>,使得分形压缩具有面向对象的自适应性,大大提高了编码效率,并成为目前分形图像编码中的主流方法。

与普通的单通道视频相比,立体视频要处理的数据量很大,因此如何解决立体视频数据的压缩传输问题显得尤其重要。在双目视频中,左右目视频由 2 个相机在不同时刻获得,这 2 个相机间的位置参数和它们的性能决定了左右目视频的特点。运动估计和视差估计<sup>[4-5]</sup>是立体视频图像压缩编码中的关键技术。运动估计是从时间上去除图像之间的冗余,提高编码效率。相对应地,视差估计可以从空间上去除图像之间的冗余,进一步提高编码效率。与运动估计相比,视差估计<sup>[6-7]</sup>要复杂得多。在运动估计中,移动的像素点通常只有小部分,而在视差估计中几乎所有像素的视差都不为 0,除非物体表面离成像平面很远,视差才可以近似看成 0。因此,对于立体视频的编码,研究高效的运动估计和视差估计算法相当重要。

总之,分形编码具有高压缩比,分辨率无关性和解码速度快等优点,但由于计算量巨大、压缩时

间过长,编码时间和解码的质量也有待提高,使得分形编码距离实用还有很大距离,需要进一步改进和提高。文献[8]提出了一种混合式的基于立方体和基于帧的自适应划分分形视频压缩算法。文献[9]提出了一种在分布式计算环境中并行实现的分形视频压缩算法,但是两种算法在压缩比和压缩时间上还有待进一步提高。

文献[10]虽然在峰值信噪比(PSNR)和压缩比基本保持不变的前提下,能够将运算复杂度降低为全搜索算法的 0.028~0.029 倍,但是压缩时间减少的幅度较低,没有从根本上解决加快编码速度的问题。因此,本文对基本的分形编码进行了改进,提出了一种基于预搜索的高效双目分形视频编码,并充分利用了分形视频编码和视频运动的不同特征。

## 2 分形编码的理论基础

下面介绍分形图像压缩的数学基础—迭代函数系统(IFS)理论。设  $D$  是  $\mathbf{R}^n$  欧氏空间的子集, $\omega$  为  $D \rightarrow D$  的映射,如果存在一个实数  $C, 0 \leq C < 1$ ,使得对于  $\mathbf{R}^n$  上的度量  $d$  满足对任意  $x, y \in D$ ,有  $d(\omega(x), \omega(y)) \leq C(d(x, y))$ ,则称  $\omega$  为压缩映射,实数  $C$  称为  $\omega$  的压缩因子。

完备的度量空间  $(X, d)$  以及  $n$  个压缩映射  $\omega_i: X \rightarrow X$  (其压缩因子分别为  $C_1, C_2, \dots, C_n$ ) 一起,就组成一个迭代函数系统 IFS,记作  $\{X: \omega_1, \omega_2, \dots, \omega_n\}$ 。 $C = \max(C_1, C_2, \dots, C_n)$  称为 IFS 的压缩因子。因此  $\{\mathbf{R}^2: \omega_1, \omega_2, \omega_3\}$  就是一个 IFS。

分形图像压缩中,一般的匹配准则是 RMS,即:

$$\text{RMS} = \sum_{i=1}^n (s \cdot a_i + o - b_i)^2, \quad (1)$$

其中,  $s$  和  $o$  分别为尺度因子和偏移因子;能够通

过选择合适的  $s$  和  $o$  来使匹配误差最小,此时有:

$$\begin{cases} s = \frac{[n \sum_{i=1}^n a_i b_i - \sum_{i=1}^n a_i \sum_{i=1}^n b_i]}{[n \sum_{i=1}^n a_i^2 - (\sum_{i=1}^n a_i)^2]}, \\ o = \frac{1}{n} [\sum_{i=1}^n b_i - s \sum_{i=1}^n a_i] \end{cases} \quad (2)$$

其中:  $b_i$  是子块 (Range block) 的像素值,  $a_i$  是父块 (Domain block) 的像素值,  $N$  为块的像素个数,  $s$  为尺度因子,  $o$  为偏移因子。对应的匹配误差为:

$$\text{RMS} = \frac{1}{N} \left[ \sum_{i=1}^N b_i^2 + s \left( s \sum_{i=1}^N a_i^2 - 2 \sum_{i=1}^N b_i a_i + 2o \sum_{i=1}^N a_i^2 \right) + o(N \cdot o - 2 \sum_{i=1}^N b_i) \right]. \quad (3)$$

在此基础之上,一些学者和研究人员把分形图像压缩的方法应用到视频序列的压缩,最为经典和影响较大的是应用循环预测编码和非收敛帧间预测编码 (CPM/NCIM) 相结合的方法<sup>[11]</sup>。类似于标准视频编码方法所采用的运动估计/补偿技术,该方法利用了相邻帧之间的时间强相关性,对视频序列压缩取得了较好的效果。在 CPM 和 NCIM 中,每个子块都由来自相邻帧相同大小的父块通过运动补偿得到。CPM 和 NCIM 两者间最大的不同在于 CPM 在解码的过程中需要具备收敛性,而 NCIM 不需要。

### 3 基于预搜索的双目分形视频编码

目前基本的分形图像和视频压缩算法的运算量很大,编码速度较慢,并且解码的质量有待提高,使得分形图像和视频压缩算法还需要进一步改进和提高。而且由于左右目所拍摄的图像存在空间信息冗余,所以需要开发高效的双目视频编

码算法。本文针对这些问题,从不同方面提出了以下的改进方法。

#### 3.1 对分形编码的基本改进

为了改进分形立体视频编码,本文首先从分形编码本身进行了如下基本改进<sup>[12-13]</sup>:

(1) 树状划分准则: 本文采用类似于 H. 264 标准的块划分方法,先把图像划分为由若干固定大小尺寸(一般为 16 pixel×16 pixel)的宏块组成的棋盘格式,再对每个宏块进行编码。

(2) 基于 DCT 变换的 I 帧编码: 对于起始参考帧,采用类似于 JPEG 静止图像压缩标准的方法,进行基于 DCT 变换的处理。

(3) 减少重复运算: 计算子块时,只计算与给定子块有关的量,包括子块的像素和与像素平方和等;计算父块时,只计算与搜索到的父块有关的量,包括父块的像素和与像素平方和等。

(4) D 块搜索策略和搜索范围: 本文中不会对子块和父块进行分类和翻转,并且由于搜索范围太大不会对结果有明显提升,所以把搜索范围限定在 7~15 pixel。

相比于基本分形视频编码,改进后的方法能保持解码图像质量基本不变,同时一定程度上减少了压缩时间和存储空间。

#### 3.2 预搜索限定条件

在分形编码过程中,通过设定合适的尺度因子  $s$  和偏移因子  $o$ ,可以使得仿射变换后的  $a_i$  值有来自  $b_i$  值的最小平方距离。当 RMS 关于  $s$  和  $o$  的偏微分为 0 时,得到 RMS 的最小值,此时对应的  $s$  和  $o$  如公式(2)所示。

因此,将公式(2)重新带入公式(1)时,可得到不同于基本分形编码的 RMS 形式如下:

$$\begin{aligned} \text{RMS} &= \sum_{i=1}^n (s \cdot a_i + o - b_i)^2 = \sum_{i=1}^n \left\{ s \cdot a_i + \frac{1}{n} \left[ \sum_{i=1}^n b_i - s \sum_{i=1}^n a_i \right] - b_i \right\}^2 = \\ &= \sum_{i=1}^n \left( \left( a_i - \frac{\sum_{i=1}^n a_i}{n} \right) \cdot \frac{[n \sum_{i=1}^n a_i b_i - \sum_{i=1}^n a_i \sum_{i=1}^n b_i]}{[n \sum_{i=1}^n a_i^2 - (\sum_{i=1}^n a_i)^2]} + \frac{\sum_{i=1}^n b_i}{n} - b_i \right)^2 = \\ &= \sum_{i=1}^n \left( (a_i - \bar{a}) \cdot \frac{[\sum_{i=1}^n a_i b_i - n \bar{a} \bar{b}]}{[\sum_{i=1}^n a_i^2 - n \bar{a}^2]} + \bar{b} - b_i \right)^2 = \end{aligned}$$

$$\sum_{i=1}^n ((a_i - \bar{a}) \cdot \frac{\sum_{i=1}^n (b_i - \bar{b})(a_i - \bar{a})}{\|a_i - \bar{a}\|^2} + \bar{b} - b_i)^2 =$$

$$\|b_i - \bar{b}\|^2 \sum_{i=1}^n \left( \frac{(a_i - \bar{a})}{\|a_i - \bar{a}\|} \cdot \frac{\sum_{i=1}^n (b_i - \bar{b})(a_i - \bar{a})}{\|b_i - \bar{b}\| \|a_i - \bar{a}\|} - \frac{b_i - \bar{b}}{\|b_i - \bar{b}\|} \right)^2, \quad (4)$$

让  $\hat{a} = \frac{(a_i - \bar{a})}{\|a_i - \bar{a}\|}$ ,  $\hat{b} = \frac{(b_i - \bar{b})}{\|b_i - \bar{b}\|}$ , 且可知  $\|\hat{a}\|^2 = 1$ ,  $\|\hat{b}\|^2 = 1$ , 则式(4)可进一步推导如下:

$$\text{RMS} = \|b_i - \bar{b}\|^2 \sum_{i=1}^n (\hat{a} \cdot \sum_{i=1}^n \hat{b} \hat{a} - \hat{b})^2 = \|b_i - \bar{b}\|^2 (1 - (\sum_{i=1}^n \hat{b} \hat{a})^2), \quad (5)$$

其中对于每个确定的子块来说,  $\|b_i - \bar{b}\|^2$  是已知的, 因此为了得到最小匹配误差 RMS,  $m = 1 - (\sum_{i=1}^n \hat{b} \hat{a})^2$  的取值比较重要, 要求它越小越好, 可以推导出一个预搜索限制条件, 从而提前结束搜索过程, 减少编码时间, 提高编码效率。

本文算法中, 在每个子块的匹配过程中, 对于  $m$  值进行判断, 如果  $0.9 < m < 1$ , 才进行接下来的匹配误差 RMS 的大小判断; 否则说明 RMS 已经足够小, 不需要下一步的比较, 直接保存结果, 进行下一个子块的匹配。

预搜索限制条件使得搜索过程提前结束, 并保证一定的解码质量, 这样存储空间和压缩时间大大减少了。当匹配误差 RMS 足够小时, 进一步的子块划分是多余的, 所以及时结束搜索过程能大大减小需要存储的压缩信息。

### 3.3 改进的运动估计算法

非均匀多层次六边形格点整像素运动搜索方法(UMHexagonS)是一种整像素的快速运动估计算法, 是针对 H. 264 的多预测模式、多参考帧的特点提出的, 并已被 H. 264 的测试软件所采用。由于该搜索方法是针对 H. 264 编码方法提出的, 本文对其进行了下述改进, 使其应用于基于分形的视频编码方法中:

#### 3.3.1 起始点预测

基于分形的视频编码算法没有涉及到多参考帧, 并且宏块和子块具有不同的大小, 故利用如图 1 所示的三种方式进行起始点预测:

(a) 空域中值预测: 取当前子块的左、上、右相邻块的运动矢量中间值为预测运动矢量;

(b) 原点预测: 根据时间相关性, 令运动矢量取值为(0, 0);

(c) 相邻参考帧预测: 利用前一参考帧中对应位置块的 MV 按比例进行预测。

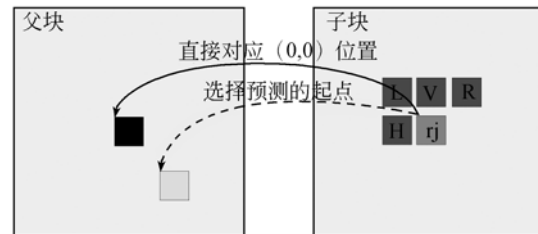


图 1 起始点预测

Fig. 1 Prediction for starting point

对于每个子块, 采用这 3 种方式进行预测, 选择对应 RMS 最小的预测方式, 进行后续的搜索。

#### 3.3.2 非对称十字模板搜索时的阈值转移条件

根据分形编码的特点, 按照块的不同大小选择不同的阈值。其中, 子块和父块的误差匹配准则 RMS 为式(2)、(3)。非对称十字模板搜索完毕之后, 选择最佳匹配点作为新的起始点进行后续模板的匹配。本文提出的阈值转移条件能及时结束搜索过程, 而且保证满足条件的匹配误差, 因此使得编码速度提高。本文方法及时去掉不需要的子块划分, 减少了需要存储的信息, 从而提高了压缩比。

#### 3.3.3 提前终止条件

根据分形编码算法的特点将提前终止分为两种情况: 一是在非均匀多层次六边形格点整像素运动搜索过程中, 除了该算法本身的提前终止条件外, 为了减少搜索复杂度, 如果最优点位于六边形中心, 可停止搜索; 二是基于分形的视频编码算法采用树状划分结构, 如图 2 所示, 首先按照模式 1 进行非均匀多层次六边形格点整像素运动搜索方法, 如果满足阈值条件, 那么终止该宏块的编

码,进行下一宏块的编码,否则按照模式 2 将模式 1 的宏块进行划分,对每个子块进行非均匀多层次六边形格点整像素运动搜索方法的运动估计,并以此类推。

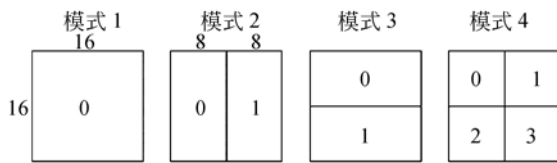


图 2 树状划分  
Fig. 2 Tree structure

本文所提出的基于非均匀多层次六边形格点整像素运动搜索的快速运动估计算法能大大地加快编码过程,同时保持可以接受的解码质量和压缩比,这对于分形视频编码来说非常重要。

### 3.4 快速的视差估计算法

对于立体视频的右目视频,除了利用时间上的冗余外,还利用空间上的冗余,即采用运动估计和视差估计联合预测的方式,选择误差最小的一种作为预测结果。本系统所利用的视差分布约束条件如下:

(1)外极线约束;对于左目图像上的一点,通过相对定向参数找出右目图像上与其对应的极线<sup>[14]</sup>,其对应点就在此极线上搜索。对于平行系统的视差搜索,只需沿扫描线,进行 X 方向的搜索即可。

(2)方向性约束;对于同一个景物,其透视投影左图像相对于右图像局部地向右移动。因此在搜索时只需沿一个方向搜索即可。例如,当寻找左图像块在右图像中的匹配块时,可优先向右进行搜索。

(3)空域相关性与时域相关性:连续变化视差场中的视差矢量有很强的相关性,即同一帧内视差矢量之间存在相关性。对于相邻两帧图像,仅有少数像素发生了运动,多数像素的位置并没有变化。对于位置不变的像素来说,其视差基本不变。所以在进行视差估计时,可以用前一帧图像的对应视差矢量作为搜索起始点进行小范围内的搜索,从而快速找到实际视差矢量。

视差估计能大大提高编码效率,它充分利用左右视点间的相关性,能够更快地找到最小的匹

配误差。除了提高编码效率之外,本文算法还能减少编码的复杂度,充分利用了双目分形视频编码的特点。

### 3.5 去方块环路滤波

在视频编码器中,加入去方块滤波器的方法有两种:后置滤波器和环路滤波器<sup>[15]</sup>。后置滤波器只处理编码环路外显示缓冲器中的数据,而环路滤波器处理编码环路中的数据。在编码器中,被滤波的图像帧作为后续编码帧运动补偿的参考帧,在解码器中,滤波后的图像输出显示。

在去除滤波方块效应时,应该先判断该边界是图像的真实边界还是方块效应所形成的边界(假边界),如图 3 所示。对真实边界不进行滤波处理,而对假边界则要根据周围图像块的性质和编码方法采用不同强度的滤波。

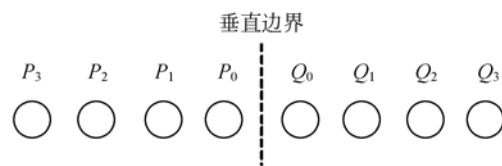


图 3 边界示意图  
Fig. 3 Sketch map of the boundary

通常来讲,真实边界两侧像素值的梯度差要比仿射变换误差造成的虚假边界两侧像素梯度差大得多。因此我们在对边界滤波的时候要考虑边界的真实性,我们在对边界像素滤波操作之前,首先对每一个位于  $4 \times 4$  亮度块之间的块边界定义参数,块边缘强度 BS,在滤波之前首先确定边缘强度,针对不同强度的块边缘,选择的滤波器和所需要滤波的像素点数也不一样,定义边界强度的判决树如图 4 所示。

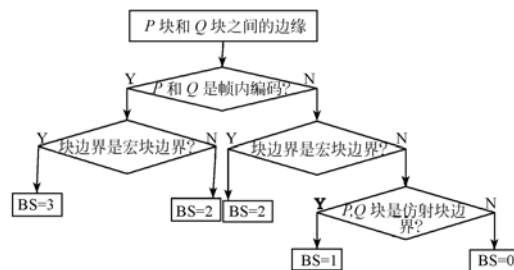


图 4 块边缘强度判决树  
Fig. 4 Decision tree for block edge strength

滤波是以宏块为单位的,先对垂直边界进行水

平滤波,再对水平边界进行垂直滤波,对宏块的两个方向上的滤波都完成后,才能进行后面宏块的滤波,对图像中宏块的滤波按光栅扫描方式进行。

去方块环路滤波使得参考帧有更高的质量,所以重构得到的图像和原图像更加相似。根据块和其周围块的编码方法,选择不同的滤波器。滤波器的选择形式如下,其中,  $P_0', Q_0', P_1', Q_1'$  表示滤波后的像素值,  $P_0, P_1, Q_0, Q_1$  表示原始的像素值:

BS=3 时,需要进行强滤波,滤波器表示为:

$$\begin{cases} P_0' = (P_1 + P_0 + Q_0) / 3 \\ Q_0' = (P_0 + Q_0 + Q_1) / 3 \\ P_1' = (2 \cdot P_1 + P_0') / 3 \\ Q_1' = (2 \cdot Q_1 + Q_0') / 3 \end{cases} \quad (6)$$

BS=2 时,滤波器表示为:

$$\begin{cases} P_0' = (P_1 + 2 \cdot P_0 + Q_0) / 4 \\ Q_0' = (P_0 + 2 \cdot Q_0 + Q_1) / 4 \end{cases} \quad (7)$$

BS=1 时,滤波器表示为:

$$\begin{cases} P_0' = (P_1 + 3 \cdot P_0 + Q_0) / 5 \\ Q_0' = (P_0 + 3 \cdot Q_0 + Q_1) / 5 \end{cases} \quad (8)$$

当 BS=0 时,不进行滤波。

### 3.6 分数像素块匹配

帧间编码宏块中的每个块和亚宏块分割区域都是根据参考帧中同尺寸的区域经仿射变换得到的,它们之间的关系用运动矢量表示。假设最佳的匹配发生在与当前块偏离整数个像素的位置,但是由于自然物体运动具有连续性,相邻两帧之间的块的运动矢量不总是以整数像素为基本单位的。

对于很多块,在一个内插到半像素精度的区域进行搜索可能会找到更好的匹配,为了获得更精确的运动向量和更高的压缩比,本文在分形编码算法中引入了分数像素运动估计技术,分数像素运动估计大致分为运动向量预测,整数像素运动估计搜索和分数像素运动估计搜索三个主要步骤,主要过程如下:

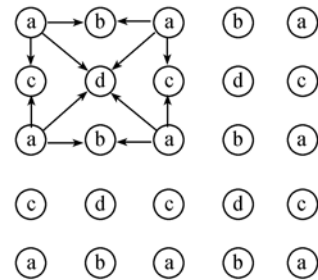
(1)对参考帧中搜索区域内的像素进行内插形成一个更高分辨率的区域;

(2)在内插区域进行整数像素和半像素位置搜索找到最佳匹配;

(3)具有最小匹配误差的相应像素保存为压缩信息。

图 5 显示了半像素内插结果,原来的整数像素位置在块中用 a 标记,像素 b 和 c 是通过对一

对整数像素进行线性内插得到的,像素 d 是通过周围的四个整数像素进行线性内插得到的。



a: 原来的整像素 b,c,d: 内插像素箭头表示内插方向

图 5 半像素内插

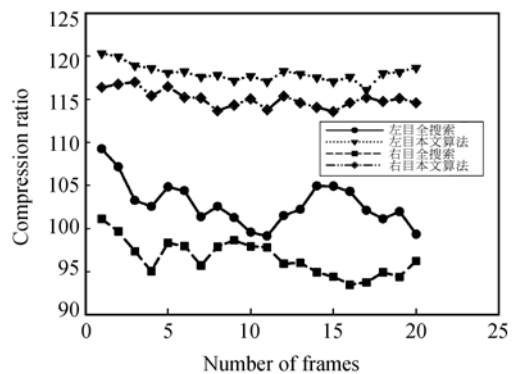
Fig. 5 Sub-pixel interpolation

分数像素块匹配能获得更准确的匹配像素,对于高质量解码图像的需求是必须的。尽管分数像素的计算增加了复杂度,但是分数像素快速运动估计和补偿的性能要好于整数像素运动估计和补偿的性能。

## 4 实验结果与分析

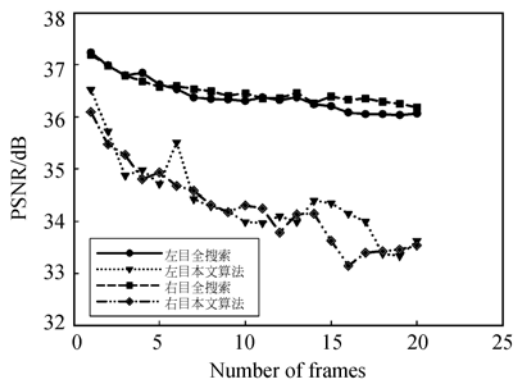
为了验证本文所改进的分形算法,分别利用基本的分形双目视频编码算法和本文改进的算法对标准测试立体视频序列“exit.yuv”的 1~20 帧进行实验。其中,基本分形算法的运动和视差估计都是基于全搜索的。计算机 CPU 为 Intel Core™2 Duo 处理器,主频 1.98 GHz,内存 2 G,实验平台为 Visual C++ 6.0。

实验结果如图 6 所示,本文改进方法所用的平均压缩时间仅为全搜索方法的 0.23 倍,压缩比

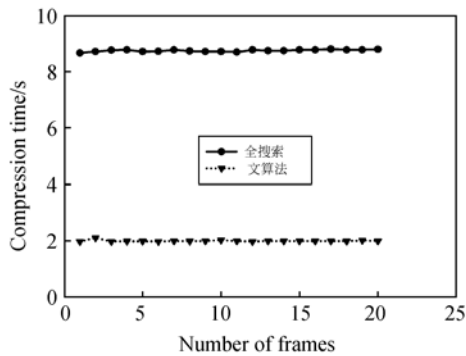


(a) 压缩比对比结果

(a) Comparison of compression ratio



(b) PSNR 对比结果  
(b) Comparison of PSNR



(c) 压缩时间对比结果  
(c) Comparison of compression time

图 6 “exit”立体视频序列 1~20 帧分别采用全搜索和本文提出的算法得到的实验结果对比

Fig. 6 Experimental comparison of the 1-20 frames of stereo video sequences “exit” by full search and the proposed algorithm

提高了 15.13~18.47, 而 PSNR 值仅减少了 1.98~2.21 dB, 这从主观上是很难直接分辨出来的。

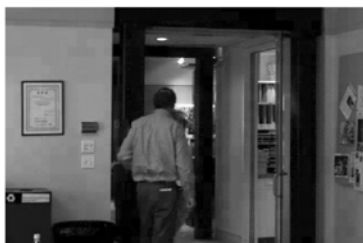
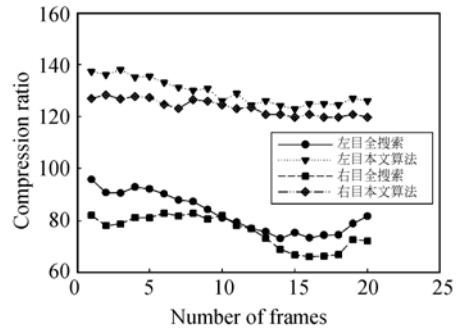


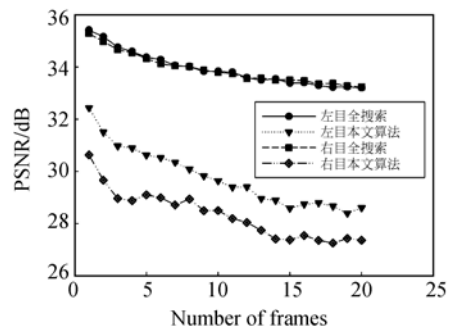
图 7 本文算法对立体视频序列“exit”右目第 3 帧的解码图像 (PSNR=35.27 dB, CR=117.01)

Fig. 7 Decoded image for the third frame on right view of stereo video sequences “exit” by proposed algorithm (PSNR=35.27 dB, CR=117.01)

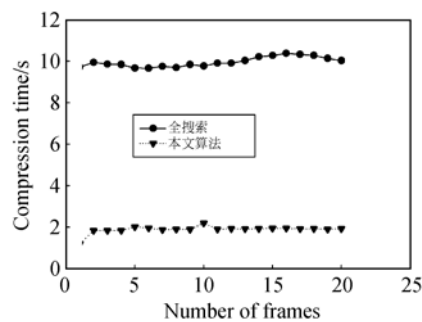
对立体视频序列“exit.yuv”利用本文算法的第 3 帧解码图像如图 7 所示。本文算法充分利用分形编码和视频运动的不同特点, 针对提高解码图像质量采用了去方块环路滤波和分数像素块匹配, 针对提高编码速度和压缩比采用了预搜索限制条件、快速的运动和视差估计算法。



(a) 压缩比对比结果  
(a) Comparison of the compression ratio



(b) PSNR 对比结果  
(b) Comparison of PSNR



(c) 压缩时间对比结果  
(c) Comparison of the compression time

图 8 “ballroom”立体视频序列 1~20 帧分别采用全搜索和本文提出的算法得到的实验结果对比

Fig. 8 Experimental comparison of the 1-20 frames of stereo video sequences “ballroom ” by full search and the proposed algorithm

此外,为了验证本文算法对于不同视频的适用性,对标准测试立体视频序列“ballroom.yuv”<sup>[16]</sup>的 1~20 帧进行了实验。所得结果如图 8 所示,本文改进方法所用的平均压缩时间仅为全搜索方法的 0.18 倍左右,压缩比提高了 46.41~47.49,而 PSNR 值仅减少了 4.1~5.5 dB。本文算法的第 3 帧解码图像如图 9 所示,可以看到主观上仍有不错的效果。



图 9 本文算法对立体视频序列“ballroom”右目第 3 帧的解码图像 (PSNR = 28.96 dB, CR = 126.73)

Fig. 9 Decoded image for the third frame on right view of stereo video sequences “ballroom” by proposed algorithm (PSNR = 28.96 dB, CR = 126.73)

在文献[8]中,使用了 4 段“videoconference”视频(256 pixel×256 pixel)和 4 段“movie”视频(720 pixel×576 pixel)。其中针对 4 段“videoconference”视频解压缩后的平均 PSNR 为 35.05 dB,平均压缩比为 13;针对 4 段 movie 视频解压缩后的平均 PSNR 为 30.59 dB,平均压缩比为 7.61。在文献[9]中,针对“videoconference”视频

其中 4 帧解压缩后的平均 PSNR 为 36.8 dB,压缩比为 6.72,16 帧的压缩时间为 337 s,平均 21s / 帧;针对“movie”视频其中 3 帧解压缩后的平均 PSNR 为 32.28dB,压缩比为 3.65。同时,相比于文献[10]中的方法,本文算法针对提高压缩比和编码速度,提出了预搜索限制条件、非均匀多层次六边形格点整像素运动搜索方法、去环路方块滤波方法和分数像素块匹配方法;并针对该文献中已经提出的视差估计方法,提出了一种更加快速简单的视差估计方法。对于同一个双目视频序列“exit.yuv”的 1~20 帧,文献[10]中平均性能如下:左目中,PSNR = 36.4 dB,CR = 102.89;右目中,PSNR = 36.47 dB,CR = 97.42;平均压缩时间为 5.69 s。而本文提出的系统的平均性能如下:左目中,PSNR = 34.43 dB,CR = 118.03;右目中,PSNR = 34.28 dB,CR = 115.06;平均压缩时间为 2 s。具体数据如下表 1~表 3 所示。其中表 1 表示左目的性能对比,表 2 表示右目的性能对比,表 3 表示压缩时间对比,这里只列举出了部分帧的数据。可以看到信噪比仅仅下降了 1.97~2.1 dB,但并不影响主观的观看,同时压缩比提高了 15.14~17.64,大大减少了所需要的存储空间,压缩时间平均减少了 3.69 s,本文改进方法所用的平均压缩时间仅为文献[10]中方法的 0.35 倍。由此分析得到,本文改进的算法比文献[10]中改进的分形算法在解码图像质量处于可接受范围内时,压缩比有很大提高,编码速度也大大提高了。

表 1 左目的压缩结果对比

Tab. 1 Comparison of compression results for left view

类型/dB	帧号								
	1	3	5	7	9	11	13	15	17
PSNR(文献[10])	37.23	36.79	36.62	36.37	36.33	36.37	36.37	36.2	36.05
PSNR(本文)	36.52	34.87	34.71	34.43	34.19	33.97	33.99	34.35	34
压缩比(文献[10])	109.25	103.25	104.82	101.36	101.27	99.14	102.22	104.92	102.13
压缩比(本文)	120.31	118.89	118.06	117.58	117.16	117.04	117.94	117.04	116.07

表2 右目的压缩结果对比

Tab. 2 Comparison of compression results for right view

类型/dB	帧号								
	1	3	5	7	9	11	13	15	17
PSNR(文献[10])	37.18	36.76	36.57	36.52	36.42	36.33	36.44	36.38	36.28
PSNR(本文)	36.09	35.27	34.93	34.58	34.18	34.25	34.14	33.63	33.4
压缩比(文献[10])	102.42	98.48	99.12	96.2	99.35	99.46	97.34	95.01	94.64
压缩比(本文)	115.9	116.01	114.34	112.5	113.11	112.91	114.43	110.93	113.14

表3 右目的压缩时间对比

Tab. 3 Comparison of compression time for right view

类型/dB	帧号								
	1	3	5	7	9	11	13	15	17
压缩时间(文献[10])	5.69	5.73	5.67	5.7	5.67	5.69	5.7	5.69	5.69
压缩时间(本文)	1.98	1.98	1.989	2	2	2	2	2	2

## 5 结 论

本文设计了基于预搜索的高效双目分形视频编码,对于基本的分形视频压缩方法从不同方面进行了改进和提高,充分利用了分形视频编码和视频运动的不同特征。实验结果表明,相比于基本的分

形全搜索算法,本系统在保证可接受的峰值信噪比(PSNR)前提下,平均压缩时间仅为全搜索方法的18%~23%,压缩比提高了15.13~47.49,大大改善了编码性能。由于分形编码自身的特点,未来的研究可能会从分形原理上进行改进,将基于对象的思想引入分形编码等,这将进一步提高分形立体视频编码的灵活性和应用范围。

## 参考文献:

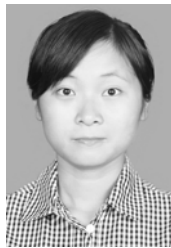
- [1] BARNESLEY M F, SLOAN A D. A better way to compress images[J]. *Byte*, 1988(1):215-233.
- [2] JACQUIN A E. Image coding based on a fractal theory of iterated contractive image transformations [J]. *IEEE Transactions on Image Processing*, 1992, 1(1): 18-30.
- [3] JACOBS E W, FISHER Y, BOSS R D. Image compression: A study of the iterated transform method [J]. *Signal Processing*, 1992, 29(3): 251-263.
- [4] 李桂菊, 刘刚, 梁静秋, 等. H. 264 快速运动估计算法的改进[J]. *光学精密工程*, 2010, 18(11): 2489-2496.
- LI G J, LIU G, LIANG J Q, *et al.*. Improvement of fast motion estimation algorithm used in H. 264 [J]. *Opt. Precision Eng.*, 2010, 18(11): 2489-2496. (in Chinese)
- [5] 祝世平, 申晓东. 十字交叉六边形块运动的估计搜索[J]. *光学精密工程*, 2009, 17(12): 3069-3076.
- ZHU SH P, SHEN X D. Cross-hexagon search algorithm for fast block-matching motion estimation [J]. *Opt. Precision Eng.*, 2009, 17(12): 3069-3076. (in Chinese)
- [6] 王爱红, 王琼华, 李大海, 等. 立体显示中立体深度与视差图获取的关系[J]. *光学精密工程*, 2009, 17(2): 433-438.
- WANG A G, WANG Q H, LI D H, *et al.*. Relationship between stereo depth and parallax image captured in stereoscopic display [J]. *Opt. Precision Eng.*, 2009, 17(2): 433-438. (in Chinese)

- [7] 纪华, 张叶, 韩广良, 等. 结合 MSA 与方向编码的快速景象匹配[J]. 光学精密工程, 2010, 18(6): 1404-1412.  
JI H, ZHANG Y, HAN G L, *et al.*. Fast scene matching combining MSA and direction code [J]. *Opt. Precision Eng.*, 2010, 18(6): 1404-1412. (in Chinese)
- [8] WANG M Q, LIU R, LAI C H. Adaptive partition and hybrid method in fractal video compression [J]. *Computers & Mathematics with Applications*, 2006, 51(11): 1715-1726.
- [9] WANG M Q, HUANG Z H, LAI C H. Matching search in fractal video compression and its parallel implementation in distributed computing environments [J]. *Applied Mathematical Modelling*, 2006, 30(8): 677-687.
- [10] 祝世平, 侯仰拴. 快速分形立体视频编码系统的设计[J]. 光学精密工程, 2010, 18(11): 2505-2512.  
ZHU SH P, HOU Y SH. Design of fast fractal stereo video coding system [J]. *Opt. Precision Eng.*, 2010, 18(11): 2505-2512. (in Chinese)
- [11] KIM C S, KIM R C, LEE S U. Fractal coding of video sequence using circular prediction mapping and noncontractive interframe mapping [J]. *IEEE Transactions on Image Processing*, 1998, 7(4): 601-605.
- [12] 祝世平, 侯仰拴, 陈菊婧, 等. 一种基于分形的双目立体视频压缩编解码方法, 中国: 201010522161.9[P].  
ZHU SH P, HOU Y SH, CHEN J Q, *et al.*. A fractal-based stereo video coding method, China: 201010522161.9[P]. (in Chinese)
- [13] 祝世平, 侯仰拴, 陈菊婧, 等. 一种基于对象和分形的双目立体视频压缩编解码方法, 中国: 201010522152.X[P].  
ZHU S P, HOU Y S, CHEN J Q, *et al.*. An object and fractal-based stereo video coding method, China: 201010522152.X[P]. (in Chinese)
- [14] 张广军. 机器视觉[M]. 北京: 科学出版社, 2005: 108-112.  
ZHANG G J. *Machine Vision* [M]. Beijing: Science Press, 2005: 108-112. (in Chinese)
- [15] 毕厚杰. 新一代视频压缩编码标准——H. 264/AVC [M]. 北京: 人民邮电出版社, 2008: 138-144.  
BI H J. *The H. 264 Advanced Video Compression Standard* [M]. Beijing: Posts & Telecom Press, 2008: 138-144. (in Chinese)

#### 作者简介:



祝世平(1970—),男,河北深泽人,副教授,1991年、1994年于西安理工大学分别获得学士、硕士学位,1997年于哈尔滨工业大学获得博士学位,主要从事图像处理和视频压缩、计算机视觉、机器视觉在精密测量中的应用等方面的研究工作。E-mail: shiping.zhu@buaa.edu.cn



陈菊婧(1988—),女,湖北洪湖人,2009年毕业于重庆大学获得学士学位,2009年进入北京航空航天大学攻读硕士学位,主要从事图像处理、视频编码、DSP应用方面的研究。E-mail: chenjuqiang927@163.com

(本栏目编辑:李自乐)