

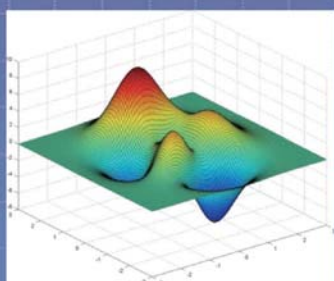


主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

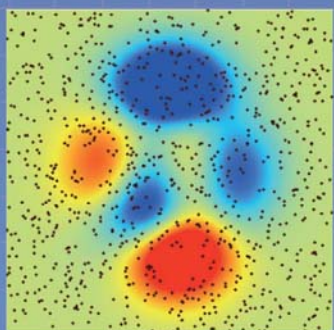
中国图象图形学报

2014
02
VOL.19

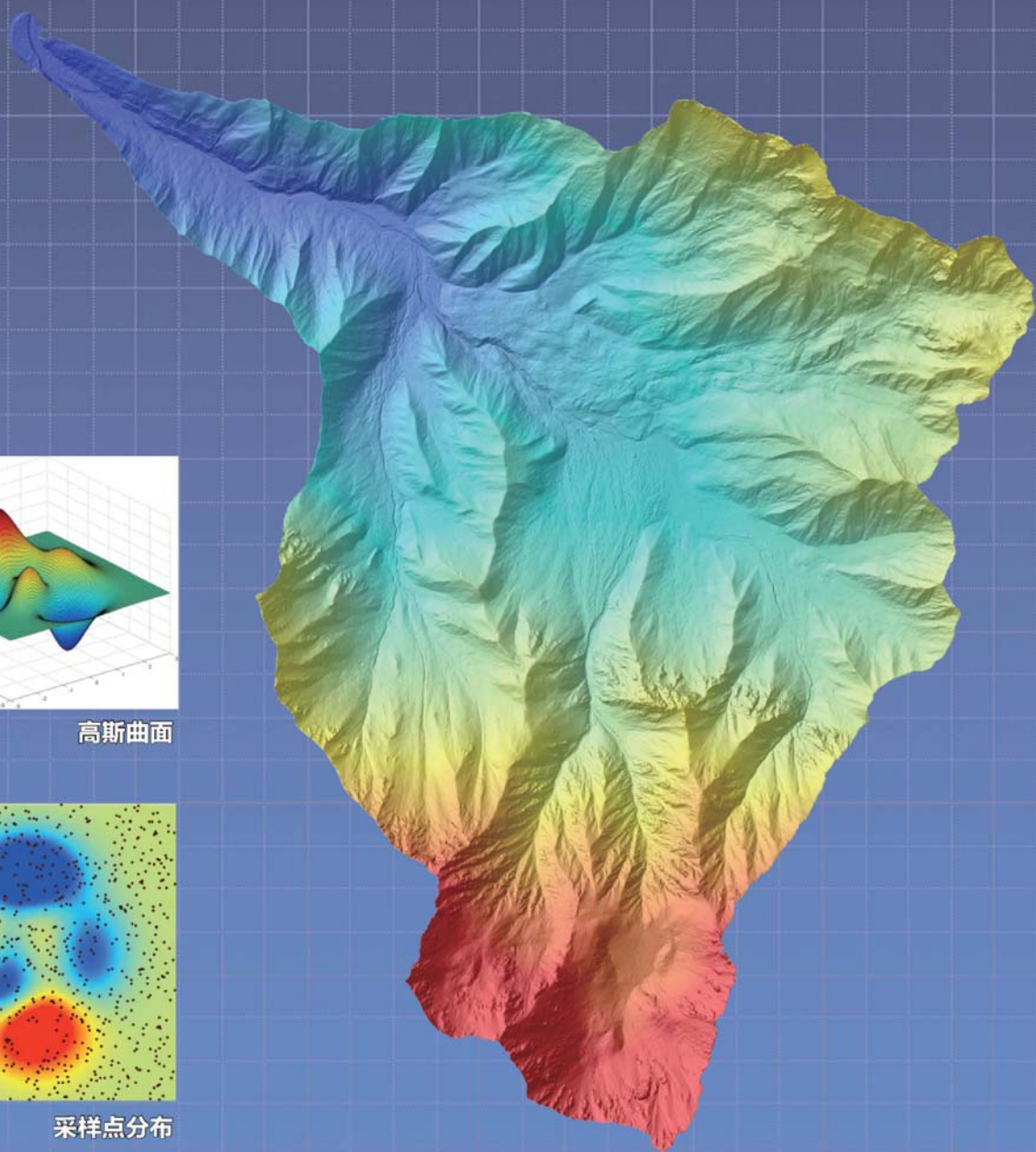
ISSN1006-8961
CN11-3758/TB



高斯曲面



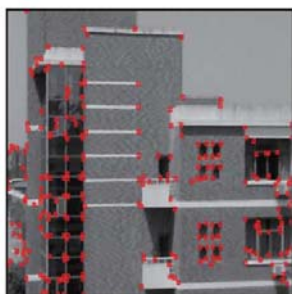
采样点分布



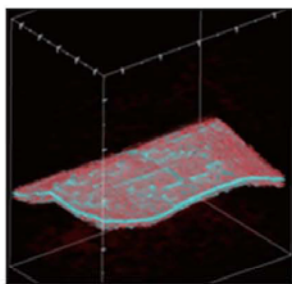
HASM地形模拟 P290



邻域最短距离法寻找最佳拼接缝(第227页)



曲率尺度空间与链码方向统计的角点检测(第234页)



目标特征指导的多分辨率体绘制算法(第283页)

计算机视觉前沿论坛

行为理解的认知推理方法

陶霖密, 杨卓宁, 王国建 0167

深度学习及其在目标和行为识别中的新进展

郑胤, 陈权崎, 章毓晋 0175

图像处理 and 编码

分布式视频编码中基于改进FCM聚类的相关噪声模型估计

杨春玲, 吴娟 0185

结合一阶自回归滑动平均和压缩感知的视频模型

王教余, 徐小红, 沈仁明, 廖重阳, 杨勋 0194

信息量加权的梯度显著度图像质量评价

徐少平, 杨荣昌, 刘小平 0201

优化加权TV的复合正则化压缩感知图像重建

费选, 韦志辉, 肖亮, 李星秀 0211

消除图像伪轮廓的各向异性自适应滤波

倪婧, 王朔中, 廖纯, 曾兴 0219

邻域最短距离法寻找最佳拼接缝

郑悦, 程红, 孙文邦 0227

图像分析和识别

曲率尺度空间与链码方向统计的角点检测

曾接贤, 李炜烨 0234

基于方向特征的手掌静脉识别

周宇佳, 刘娅琴, 杨丰, 黄靖 0243

复杂环境下高效物体跟踪级联分类器

江伟坚, 郭躬德 0253

图像理解和计算机视觉

均值规范化对比度的局部特征描述符

颜雪军, 赵春霞, 袁夏, 徐丹, 刘凡 0266

计算机图形学

三角域上Said-Ball基的推广渐近迭代逼近

张莉, 李园园, 杨燕, 檀结庆 0275

目标特征指导的多分辨率体绘制算法

郭思奇, 鲁才, 聂小燕 0283

地理信息技术

高精度曲面建模优化方案

赵明伟, 岳天祥, 赵娜 0290

医学图像处理

应用于人体关节缺损修复的建模与可视化

邢慧君, 杨健, 李勤 0297

小邻域统计信息核磁共振医学图像分割模型

张建伟, 方林, 陈允杰, 詹天明 0305

遥感图像处理

结合暗通道原理和双边滤波的遥感图像增强

周雨薇, 陈强, 孙权森, 胡宝鹏 0313

改进NSCT和IHS变换相结合的遥感影像融合

刘慧, 周可法, 王金林, 王珊珊 0322

分段2维主成分分析的超光谱图像波段选择

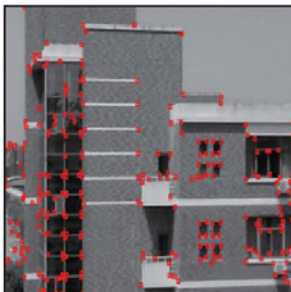
张婧, 孙俊喜, 阮光诗, 刘红喜 0328

CONTENTS

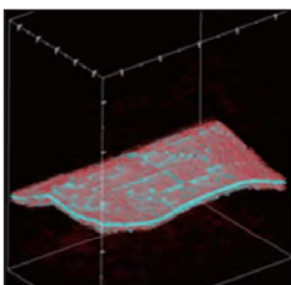
JOURNAL OF IMAGE AND GRAPHICS



Finding an optimal seam-line through the shortest distance in the neighborhood(P227)



Corner detection based on curvature scale space and chain code direction statistics(P234)



Target feature based multi-resolution volume rendering(P283)

Forum: Forefront of Computer Vision

Cognitive reasoning method for behavior understanding

Tao Linmi, Yang Zhuoning, Wang Guojian 0167

Deep learning and its new progress in object and behavior recognition

Zheng Yin, Chen Quanqi, Zhang Yujin 0175

Image Processing and Coding

Correlation noise modeling based on improved Fuzzy C-Means clustering in distributed video coding

Yang Chunling, Wu Juan 0185

Compressed sensing video model based on a first-order autoregressive moving average model

Wang Jiaoyu, Xu Xiaohong, Shen Renming, Liao Chongyang, Yang Xun 0194

Information content weighted gradient saliency structural similarity index for image quality assessment

Xu Shaoping, Yang Rongchang, Liu Xiaoping 0201

Compound regularized compressed sensing image reconstruction based on optimal reweighted TV

Fei Xuan, Wei Zhihui, Xiao Liang, Li Xingxiu 0211

False contour suppression with anisotropic adaptive filtering

Ni Jing, Wang Shuozhong, Liao Chun, Zeng Xing 0219

Finding an optimal seam-line through the shortest distance in the neighborhood

Zheng Yue, Cheng Hong, Sun Wenbang 0227

Image Analysis and Recognition

Corner detection based on curvature scale space and chain code direction statistics

Zeng Jiexian, Li Weiye 0234

Palm-vein recognition based on oriented features

Zhou Yujia, Liu Yaqin, Yang Feng, Huang Jing 0243

Efficient cascade classifier for object tracking in complex conditions

Jiang Weijian, Guo Gongde 0253

Image Understanding and Computer Vision

Local feature descriptor based on mean normalized contrast

Yan Xuejun, Zhao Chunxia, Yuan Xia, Xu Dan, Liu Fan 0266

Computer Graphics

Generalized progressive iterative approximation for Said-Ball based on triangular domains

Zhang Li, Li Yuanyuan, Yang Yan, Tan Jieqing 0275

Target feature based multi-resolution volume rendering

Guo Siqi, Lu Cai, Nie Xiaoyan 0283

Geoinformatics

HASM optimization based on the improved difference scheme

Zhao Mingwei, Yue Tianxiang, Zhao Na 0290

Medical Image Processing

Method of modeling and visualization for repairing of osteoarticular defect

Xing Huijun, Yang Jian, Li Qin 0297

Magnetic resonance medical images segmentation based on a local statistical information model

Zhang Jianwei, Fang Lin, Chen Yunjie, Zhan Tianming 0305

Remote Sensing Image Processing

Remote sensing image enhancement based on dark channel prior and bilateral filtering

Zhou Yuwei, Chen Qiang, Sun Quansen, Hu Baopeng 0313

Remote sensing image fusion based on an improved NSCT and IHS transformation

Liu Hui, Zhou Kefa, Wang Jinlin, Wang Shanshan 0322

Segmented 2DPCA algorithm for band selection of hyperspectral image

Zhang Jing, Sun Junxi, Ruan Guangshi, Liu Hongxi 0328

中图法分类号: TN919.8 文献标识码: A 文章编号: 1006-8961(2014)02-0194-07

论文引用格式: 王教余, 徐小红, 沈仁明, 廖重阳, 杨勋. 结合一阶自回归滑动平均和压缩感知的视频模型[J]. 中国图象图形学报, 2014, 19(2): 194-200. [DOI:10.11834/jig.20140204]

结合一阶自回归滑动平均和压缩感知的视频模型

王教余, 徐小红, 沈仁明, 廖重阳, 杨勋

合肥工业大学计算机与信息学院, 合肥 230009

摘要: **目的** 针对视频处理中面临的采样数据量大及采样时间长的问题, 把视频状态空间一阶自回归滑动平均模型和压缩感知模型相结合, 提出了一种基于一阶自回归滑动平均的视频压缩感知模型。 **方法** 主要思想是在压缩感知理论框架下, 充分利用视频帧内稀疏性和帧间相关性, 把视频分割成动态部分和静态部分同时采样但分别处理, 利用凸优化等方法得到视频状态空间一阶自回归滑动平均模型的关键参数。 **结果** 多组真实场景下的实验结果表明, 该模型较大程度上降低了帧间冗余度和数据采集量, 视频采集压缩比为 100 ~ 200 时, 仍然能取得较好的重建效果。 **结论** 结合压缩感知和线性预测技术, 提出了一种新的视频获取模型, 对视频的静态部分和动态部分分别处理, 并给出了该模型使用的条件。实验结果表明, 该模型对帧间变化不大的视频, 具有良好的压缩效果。

关键词: 视频处理; 压缩感知; 稀疏表示; 自回归滑动平均模型

Compressed sensing video model based on a first-order auto regressive moving average model

Wang Jiaoyu, Xu Xiaohong, Shen Renming, Liao Chongyang, Yang Xun

School of Computer and Information, Hefei University of Technology, Hefei 230009, China

Abstract: **Objective** In order to reduce the large data volumes in video processing, we combine the first-order Auto Regressive Moving Average (ARMA) model video model with compressed sensing theory, and propose a compressed sensing video model, which is based on the first-order ARMA. **Method** The main idea is making full use of video sparsity and frame coherence under the theoretical framework of compressed sensing, and dividing the video into a static part and a dynamic part. The new model gets the key parameters through simultaneous sampling and separate processing. Moreover, we discuss the construction conditions of the model and provide concrete guidelines on how to use this new model with provable performance. **Result** We present experimental evidence that, within our framework, the data volume can be reduced largely and reconstructed video shows a robust result even with compression rates at a ratio of 100 to 200. **Conclusion** Combining with the compressed sensing and linear prediction technology, we propose a new video acquisition model, which static and dynamic parts of video can process respectively. Additionally, we give the conditions of using this model. Experiments show that the model has a well compression effect while facing smooth video.

Key words: video processing; compressed sensing; sparse representation; auto regressive moving average model

收稿日期: 2013-07-23; 修回日期: 2013-09-08

基金项目: 安徽省自然科学基金项目(128085MF91)

第一作者简介: 王教余(1988—), 男, 合肥工业大学信号与信息处理方向在读硕士研究生, 主要研究方向为图像处理、模式识别。

E-mail: wangjiaoyu0701@163.com

0 引言

由 Donoho、Candes 等人^[1-2]提出的压缩感知(CS)理论,作为一种新的信号采集方法,突破了原有的香农采样定理。当信号满足 RIP (restricted isometry property) 准则^[2]时,其采样和压缩可以同时进行,这可以大大减小数据采集量,同时通过解一个优化问题就可以从压缩的信号重建出原始信号。压缩感知主要由一个欠定方程 $\mathbf{z} = \Phi \mathbf{y} + \mathbf{e}$, 通过观测结果 $\mathbf{z}$ 和观测矩阵 Φ 来重建目标信号 $\mathbf{y}$, 式中 $\mathbf{y} \in \mathbf{R}^N$, $\Phi \in \mathbf{R}^{M \times N}$, $M \ll N$, $\mathbf{e}$ 是测量误差。该理论成立的一个重要的前提是信号 $\mathbf{y}$ 在某个基 Ψ 下是稀疏的, 即, $\mathbf{y} = \Psi \boldsymbol{\theta}$, 式中 $\boldsymbol{\theta}$ 有 k 个非零值, $K \ll N$ 。经 $M = O(k \log(N/K))$ 次测量后, 信号 $\mathbf{y}$ 可以通过解式(1)求得, 即

$$\min \|\boldsymbol{\theta}\|_1 \quad \text{s. t.} \quad \|\mathbf{z} - \Phi \Psi \boldsymbol{\theta}\|_2 \leq \varepsilon \quad (1)$$

式中, ε 是测量误差的上限。解决问题式(1)的算法有很多, 如贪婪算法、门限算法和基于凸优化压缩感知框架算法等。

利用 CS 理论, 可大大减小了数据采集量、采样时间、存储及传输负担。因此, 该理论提出后, 受到一直被大数据困扰的视频处理领域的高度关注^[3]。如何利用 CS 理论有效地降低视频采样数据, 具有很强的现实意义。视频压缩感知就是利用测量矩阵得到观测值并重建出原始视频的过程, 即

$$\mathbf{z}_t = \Phi_t \mathbf{y}_t \quad (2)$$

式中, $\Phi_t \in \mathbf{R}^{M \times N}$ 是 t 时刻的测量矩阵, $M \ll N$, $\mathbf{y}_t$ 是 t 时刻的视频帧, $\mathbf{z}_t$ 是 t 时刻的观测值。目的是通过已知的 $\mathbf{z}_{1:T} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_T\}$ 和 Φ_t , 恢复视频帧 $\mathbf{y}_{1:T} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T\}$, $1:T$ 表示从第 1 帧到第 T 帧。

目前实现视频压缩感知的模型主要有 Wakin 的 3D 小波稀疏基模型^[4]、Wakin 和 Park 的 coarse-to-fine 统计模型^[5]、Cevher 基于静态背景的压缩感知模型^[6]、Kim 的 iterative self-similar 模型^[7]、Chang 的 4D MRA (four dimensional model reference adaptive) 模型^[8]。其中 3D 小波稀疏基模型把视频帧分成许多块, 块的数量为 1D, 然后对每个图像块进行 2D 上的小波压缩, 故称为 3D 小波稀疏基, 它实现的理论基础是对于每个视频帧, 其帧内图像块之间是具有空间上的相关性的, 在每个图像块压缩时, 考虑到块之间的空间相关性。这种方法容易出现视频块效应。coarse-to-fine 模型较 3D 小波稀疏

基模型, 引入了视频相邻帧间块相似度统计信息。但是块效应仍不能很好地消除。iterative self-similar 模型对每帧使用压缩感知方法, 利用压缩感知手段重建。它只是简单地把图像压缩感知方法应用到视频领域, 没有利用视频结构的特殊性做任何的改变, 这种方法实用性不强, 另外重建时间复杂度较大。基于静态背景的压缩感知模型, 只适合二元背景的视频场景, 用于遥感和地质成像, 对一般的视频不具有实用性。4D MAR 模型, 对视频帧内进行分块离散小波变换(DWT), 同时用 2 维的方位信息去表征块之间的相关性。方法上和 3D 小波稀疏基模型相似, 相比较, 4D MAR 模型分块更小, 视频重建鲁棒性更好。因而一定程度上消除了 3D 小波稀疏基模型带来的明显的块效应问题。然而上述模型中只有 coarse-to-fine 统计模型利用到了视频相邻帧的内容相关性, 其余模型均把每个视频帧看成一个独立的视频帧来处理, 只用到帧内稀疏性, 还是传统的图像压缩感知的思想, 并没有利用视频内容帧间相关性去进一步降低视频数据采集量。因此, 将视频压缩感知和视频状态空间自回归滑动平均模型 (ARMA model)^[9] 相结合, 提出了一种基于一阶 ARMA 的视频压缩感知模型 (CS-ARMA 模型), 同时利用视频帧内和帧间的稀疏性, 把视频分割成动态部分和静态部分, 利用凸优化和线性优化方法, 得到视频 ARMA 模型的关键系数, 实现视频信号重建。

1 ARMA 视频模型

视频处理中, ARMA 模型是最重要的模型之一, 在目标识别^[10]、交通监控^[11]、人物活动场景^[12] 中都有广泛的应用。ARMA 描述视频时, 把视频看成具有时间相关性和有着相同特点的图像序列, 例如波浪、烟雾、植物、旋风和车辆行驶等目标的视频都可以用 ARMA 模型来表述。该模型描述视频有着诸多优点, 首先它是线性模型, 能很好的处理局部信息丢失问题, 结构较为简单, 只经过少量计算步骤即可确定模型的参数, 这些参数描述的是整个视频的特征而不是局部特征, 这利于分析视频的主要结构及其时间上的关联性。其次 ARMA 模型系统参数计算量较小, 能快速求解, 减轻了在最小方差下进行线性预测和系统控制的计算复杂度, 提高视频分析和处理的实时性。

视频状态空间 ARMA 模型^[13] 中视频帧 $\mathbf{y}_t$ 描

述为

$$y_t = \sum_{j=0}^q c_j x_{t-j} \quad (3)$$

$$x_t = \sum_{i=1}^p a_i x_{t-i} + e_t \quad (4)$$

由式(3)(4)及文献[14]知,其一阶状态空间方程为

$$y_t = Cx_t + \varepsilon_t \quad (5)$$

$$x_{t+1} = Ax_t + \eta_t \quad (6)$$

式(5)(6)所描述的即是视频状态空间一阶 ARMA 模型(后称 ARMA 模型),式(3)~(6)中 $x_t \in \mathbf{R}^d$ 是 t 时刻的状态向量, d 是状态空间的维度, $a_i, A \in \mathbf{R}^{d \times d}$ 是状态转移矩阵, $c_i, C \in \mathbf{R}^{N \times d}$ 是一滤波器组, $d \ll N$, ε_t, η_t 及 e_t 是噪声, p 和 q 是 ARMA 模型的阶数。由式(5)(6)可知 ARMA 视频模型由矩阵对 (C, A) 描述的,滤波器组 C 和状态转移矩阵 A 互不相关。滤波器组 C 的选择要充分考虑到视频的内容和计算效率,其他参数的求解就是线性系统的系统识别问题。文献[9]给出了该模型参数具体的求解过程和系统误差分析及其修正方法。

2 CS-ARMA 视频模型

为了在视频压缩感知理论框架下充分利用视频帧内稀疏性和帧间相关性,将压缩感知和一阶 ARMA 相结合,提出基于一阶 ARMA 的视频压缩感知模型,即由式(2)(5)(6)可知

$$z_t = \begin{pmatrix} \tilde{z}_t \\ \tilde{z}_t \end{pmatrix} = \begin{bmatrix} \tilde{\Phi} \\ \tilde{\Phi}_t \end{bmatrix} y_t = \begin{bmatrix} \tilde{\Phi} y_t \\ \tilde{\Phi}_t y_t \end{bmatrix} = \begin{bmatrix} \tilde{\Phi} C x_t + \varepsilon_t \\ \tilde{\Phi}_t C x_t + \eta_t \end{bmatrix} \quad (7)$$

$$x_{t+1} = Ax_t + \eta_t \quad (8)$$

式中, $\tilde{z}_t \in \mathbf{R}^{\tilde{M}}$ 、 $\tilde{z}_t \in \mathbf{R}^{\tilde{M}}$, 每帧测量的总次数为 $M = \tilde{M} + \tilde{M}$, $\tilde{M}$ 为动态测量次数、 $\tilde{M}$ 为静态测量次数。

式(7)由两个部分构成:动态量观测矩阵 $\tilde{\Phi}$ 和常量观测矩阵 $\tilde{\Phi}_t$, 称 $\tilde{z}_t$ 为动态测量值(即视频动态部分)、 $\tilde{z}_t$ 为测量值常项(即视频静态部分)。对于式(7),只知道测量结果序列 z_t 和观测矩阵 Φ_t , 这是常见的双线性问题,分两步来求解该系统的参数。首先,通过测量值常项 $\tilde{z}_{1:T}$ 估计状态序列 $\hat{x}_{1:T}$,然后再用动态测量值 $\tilde{z}_t$ 和状态序列 $\hat{x}_{1:T}$ 去计算滤波器组 C , 图 1 是 CS-ARMA 模型的求解流程图。

$\tilde{z}_{1:T}$ 可通过 N4SID (subspace state space system

identification)^[15]、PCA-ID (principal component analysis and identification)^[16] 和最大期望算法 (EM)^[17] 来求解。N4SID 算法通过识别状态子空间逼近系统参数,这种算法最大问题是计算空间复杂度太大。EM 算法收敛效果不好,对噪声比较敏感。使用一种适度的算法——PCA-ID 算法,假设滤波器组 C 在空间和时间上是不相关的,利用主成分分析 (PCA) 的方法去近似逼近系统参数,可以快速有效地求解 $\tilde{z}_{1:T}$ 。

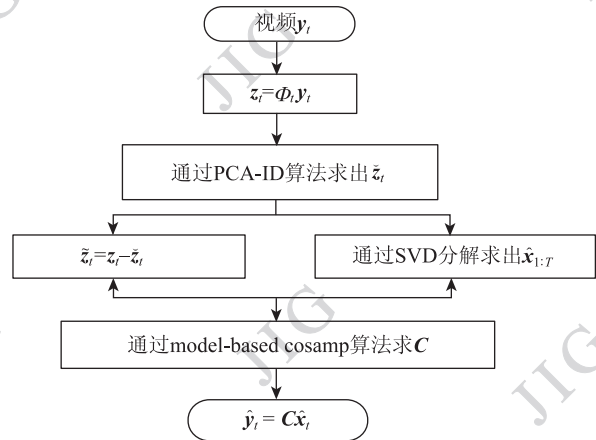


图 1 视频 CS-ARMA 求解流程图

Fig. 1 Flow chart for CS-ARMA

2.1 状态序列 $\hat{x}_{1:T}$ 估计

2.1.1 CS-ARMA 模型构造条件

由于常量观测矩阵 $\tilde{\Phi}$ 是不变的,由式(4)(5)可知 $(\tilde{\Phi} C, A)$ 描述的是线性系统。对于 CS-ARMA 系统,输入 $y_0, y_1, \dots, y_{d-1}$ 以及在有限采样周期内测量到的有限输出 $z_0, z_1, \dots, z_{d-1}$,若能唯一确定任意初始状态 x_0 ,则称该系统是完全可观测的,简称系统是可测的,详见文献[18],该模型的构造条件即定理 1。

定理 1 在 CS-ARMA 模型中,当且仅当系统矩阵

$Y(\tilde{\Phi} C, A) = [\tilde{\Phi} C, \tilde{\Phi} CA, \dots, \tilde{\Phi} CA^{d-1}]$ 为满秩时,系统是可测的,这时候初始状态值 x_0 是可求的,即模型是可有效构造的。

证明

由式(7)(8)可得 $\begin{cases} \tilde{z}_t = \tilde{\Phi} C x_t + \varepsilon_t \\ x_{t+1} = Ax_t + \eta_t \end{cases}$, 其中 ε_t, η_t

为系统误差,为了讨论可测性,可设系统误差为零,初始化输入为 $x_0 = [x_{0,0}, x_{0,1}, \dots, x_{0,(d-2)}, x_{0,(d-1)}]^T$ 。上式动态方程为

$$\begin{cases} \tilde{\mathbf{z}}_0 = \tilde{\Phi} \mathbf{C} \mathbf{x}_0 \\ \tilde{\mathbf{z}}_1 = \tilde{\Phi} \mathbf{C} \mathbf{A} \mathbf{x}_0 \\ \vdots \\ \tilde{\mathbf{z}}_{d-1} = \tilde{\Phi} \mathbf{C} \mathbf{A}^{d-1} \mathbf{x}_0 \end{cases} \quad (9)$$

其向量-矩阵形式为

$$\begin{bmatrix} \tilde{\Phi} \mathbf{C} \\ \tilde{\Phi} \mathbf{C} \mathbf{A} \\ \vdots \\ \tilde{\Phi} \mathbf{C} \mathbf{A}^{d-1} \end{bmatrix} \begin{bmatrix} x_{0,0} \\ x_{0,1} \\ \vdots \\ x_{0,(d-1)} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{z}}_0 \\ \tilde{\mathbf{z}}_1 \\ \vdots \\ \tilde{\mathbf{z}}_{d-1} \end{bmatrix}$$

$$\text{令 } \mathbf{Y}(\tilde{\Phi} \mathbf{C}, \mathbf{A}) = \begin{bmatrix} \tilde{\Phi} \mathbf{C} \\ \tilde{\Phi} \mathbf{C} \mathbf{A} \\ \vdots \\ \tilde{\Phi} \mathbf{C} \mathbf{A}^{d-1} \end{bmatrix}, \text{ 称 } \mathbf{Y}(\tilde{\Phi} \mathbf{C}, \mathbf{A}) \text{ 为 CS-}$$

ARMA 系统矩阵。

式(9)展开后有 $d\tilde{M}$ 个方程,若其中有 d 个独立方程,便可唯一确定一组 $[x_{0,0}, x_{0,1}, \dots, x_{0,(d-1)}]$ 。故 CS-ARMA 系统可有效构造的充分必要条件是 $\text{rank}(\mathbf{Y}(\tilde{\Phi} \mathbf{C}, \mathbf{A})) = d$, 证毕。

2.1.2 $\hat{\mathbf{x}}_{1:T}$ 的求解

由式(7)(8)可知,测量值常项 $\tilde{\mathbf{z}}_{1:T}$ 是由 $(\tilde{\Phi} \mathbf{C}, \mathbf{A})$ 决定的,当定理 1 条件得到满足时,求状态序列 $\hat{\mathbf{x}}_{1:T}$ 问题就转化为系统识别问题。对于 CS-ARMA 系统的卜循环矩阵 $\mathbf{Hankel}(\tilde{\mathbf{z}}_{1:T}, d)$ 有

$$\mathbf{Hankel}(\tilde{\mathbf{z}}_{1:T}, d) = \mathbf{Y}(\tilde{\Phi} \mathbf{C}, \mathbf{A}) [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{T-d+1}] \quad (10)$$

式中

$$\mathbf{Hankel}(\tilde{\mathbf{z}}_{1:T}, d) = \begin{bmatrix} \tilde{\mathbf{z}}_1 & \tilde{\mathbf{z}}_2 & \cdots & \tilde{\mathbf{z}}_{T-d+1} \\ \tilde{\mathbf{z}}_2 & \cdots & \tilde{\mathbf{z}}_{T-d+2} \\ \vdots & \vdots & \vdots \\ \tilde{\mathbf{z}}_d & \cdots & \tilde{\mathbf{z}}_{T-1} & \tilde{\mathbf{z}}_T \end{bmatrix}$$

由文献[9]可知对 $\mathbf{Hankel}(\tilde{\mathbf{z}}_{1:T}, d)$ 奇异值分解,便有 $[\hat{\mathbf{x}}_{1:T-d+1}] = \mathbf{\Sigma}_d \mathbf{V}_d^T$, $\mathbf{\Sigma}_d$ 是包含 $\mathbf{\Sigma}$ 矩阵中前 d 个最大的奇异值的对角矩阵, $\mathbf{V}_d$ 是对应的酉矩阵向量。

2.2 滤波器组 $\mathbf{C}$ 的估计

若已求出状态序列 $\hat{\mathbf{x}}_{1:T}$, 由 $\tilde{\mathbf{z}}_t = \tilde{\Phi}_t \mathbf{C} \mathbf{x}_t$, 可知滤波器组 $\mathbf{C}$ 和 $\tilde{\mathbf{z}}_t$ 是线性关系。另外本文模型中 $\mathbf{C}$ 是时不变的,可以通过积累一段时间的 $\tilde{\mathbf{z}}_t$ 序列去估计 $\mathbf{C}$, 这样每帧的动态部分测量次数 $\tilde{M}$ 就大大减小。这对 CS-ARMA 模型有着重要的意义。

在求解 $\mathbf{C}$ 的时候,要先验证其稀疏性,对绝大多数视频来说,滤波器组 $\mathbf{C}$ 得到的是视频的主要结

构,当视频帧间在空间上具有相关性,那么滤波器组 $\mathbf{C}$ 在小波基或 DCT 基下都具有稀疏性^[14]。因此可以通过解凸优化问题来求 $\mathbf{C}$, 即

$$\text{已知 } \hat{\mathbf{x}}_{1:T} \text{ 有 } \min \sum_{i=1}^N \|\mathbf{s}_i\|_1$$

$$\text{s. t. } \mathbf{C} = \mathbf{\Psi} \mathbf{S}, \forall t, \|\tilde{\mathbf{z}}_t - \tilde{\Phi}_t \hat{\mathbf{C}} \hat{\mathbf{x}}_t\|_2 \leq \varepsilon \quad (11)$$

式中, $\mathbf{s}_i$ 是 $\mathbf{S} = \mathbf{\Psi}^T \mathbf{C}$ 的第 i 列, $\mathbf{\Psi}$ 是 $\mathbf{C}$ 的稀疏基, $\hat{\mathbf{C}}$ 是 $\mathbf{C}$ 的估计值。式(11)的求解过程是一个混合 L_1 - L_2 范数的问题^[19]。SPG-L1^[20] 和 model-based cosamp^[19] 算法都可求解式(11), 本文选择 model-based cosamp, 它能快速收敛到较为准确的值, 计算过程也比较直观, 易于理解。

3 实验

通过多组实验去检测 CS-ARMA 模型的鲁棒性。引入了压缩比的概念即, $R = \frac{N}{M}$, 其中 N 是单个视频帧 Nyquist 采样需要的采样数, M 是本文方法所需的单个视频帧采样数, 通过重建信噪比来评价恢复的视频的鲁棒性, 重建信噪比 RE_{SNR} 为(单位 dB)

$$RE_{\text{SNR}} = 10 \lg \left(\frac{\sum_{t=1}^T \|\mathbf{y}_t\|_2^2}{\sum_{t=1}^T \|\mathbf{y}_t - \hat{\mathbf{y}}_t\|_2^2} \right) \quad (12)$$

输入信噪比 IN_{SNR} 为(单位 dB)

$$IN_{\text{SNR}} = 10 \lg \left(\frac{\sum_{t=1}^T \|\mathbf{y}_t\|_2^2}{T\sigma^2} \right) \quad (13)$$

式中, $\mathbf{y}_t$ 和 $\hat{\mathbf{y}}_t$ 分别为原始视频和重建后视频, σ 为噪声标准方差, T 为视频的帧数。

3.1 动态部分测量次数 $\tilde{M}$ 值的选择

为有效恢复滤波器组 $\mathbf{C}$, 至少需要 $O(dK \times \log(N/K))$ 次测量^[1-2], K 为常数。经过大量的实验, 给出一个较佳的经验值 $4dK \log(N/K)$ 。在 T 帧内需要测量 $\tilde{M}T$ 次, 即

$$\tilde{M}T = 4dK \log(N/K) \Rightarrow \tilde{M} = 4 \frac{dK}{T} \log(N/K) \quad (14)$$

式(14)表明, 当 T 较大的时候(如高速视频采样), $\tilde{M}$ 将变小。令 $T = \tau f_s$, f_s 是采样频率, τ 是视频时长。因此, 这个模型的复杂度往往取决于 τ 的大小。对于固定的 τ , 当 f_s 增加时, 测量次数 $\tilde{M}$ 也相应地减小。

3.2 不同静态测量次数 $\tilde{M}$ 对重建效果的影响

在定理 1 的约束条件下,通过实验分析在不同 IN_{SNR} 下,不同测量次数 $\tilde{M}$ 对 IN_{SNR} 的影响。实验采用 Wakin^[5] 提供的蜡烛视频,视频转为 352×288 的灰度图像,其中 $\tilde{M} \geq 1$,基础空间向量维度 $d = 10$, $T = 500$ (即有 500 帧),当 IN_{SNR} 比较高时,即使 $\tilde{M}$ 较小时,也能得到较好的重建效果,详细结果见图 2,其中,不同曲线代表不同的 IN_{SNR} ,实验结果表明 $\tilde{M}$ 大于 10 时,视频可以较好地重建,由于视频是稀疏的,当 $\tilde{M}$ 增加到一定值时,再增加静态测量次数,重建效果改善并不明显。

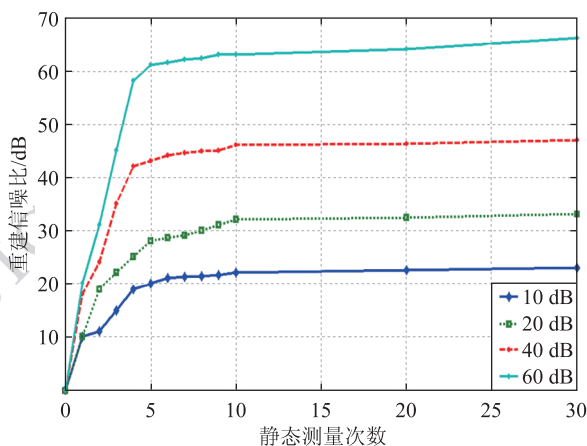


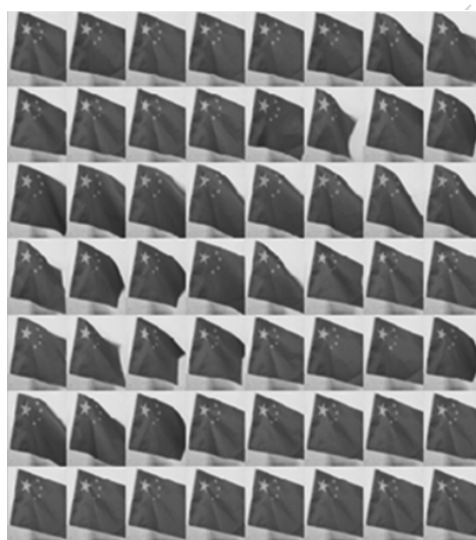
图 2 不同 $\tilde{M}$ 值对 RE_{SNR} 的影响

Fig. 2 The affection of different $\tilde{M}$ value for RE_{SNR}

3.3 实际场景实验

实验中两组视频均来自实际场景中的视频。为了增强实验结果的视觉对比,视频均转为 352×288 的灰度图。第 1 组 Flag 视频是用佳能 IXUC220HS 摄像机拍摄的 30 s 视频,用本文方法重建视频,结果见图 3,压缩比 R 为 200,每帧测量次数 M 为 506, RE_{SNR} 为 27.95 dB,由于红旗视频具有较强帧间相关性和良好的稀疏性,采用 30 s 的视频为一组求解 ARMA 系统参数,而对一些帧间变化较为剧烈的视频,为了保证重建效果,视频分组时间可以更短。由于不同观测基 Ψ 会影响 C 的重建精度,故也对该组视频在不同观测基 Ψ 下重建效果做了对比,结果见图 4。图 5 是为了跟文献[5-6,8-9]实验结果对比,采用 Wakin^[5] 提供的蜡烛视频,视频时长为 25 s,选了一组重建信噪比均在 28 dB 左右的重建视频,其中图 5(b)一(f)所示重建方法对应的压缩比 R 分别为 200、3、70、50 及 55。图 6 为该对比实验的全部数据,描述了在不同模型下,重建信噪比和压缩比的关系,

可以看出在相同压缩比下,由于本文模型同时利用了帧间相关性和帧内稀疏性,故具有更强的鲁棒性。



(a) 原始视频



(b) CS-ARMA模型重建视频

图 3 国旗视频

Fig. 3 Flag video

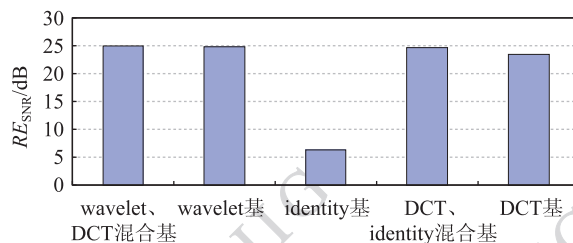


图 4 不同观测基重建信噪比

Fig. 4 The affection of different observation matrix for RE_{SNR}

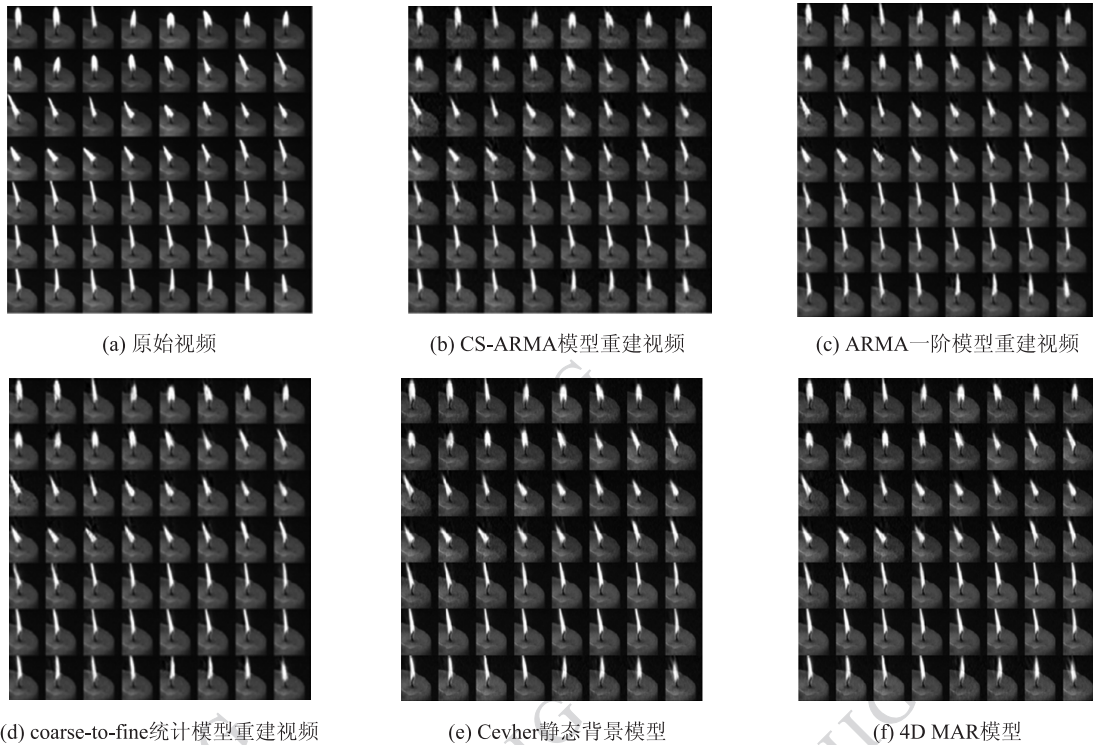


图 5 蜡烛视频

Fig. 5 Candle video

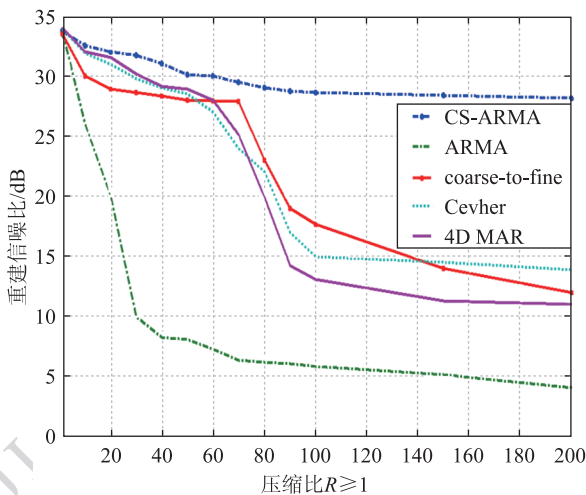


图 6 不同模型重建信噪比和压缩比的关系

Fig. 6 The affection of different models for R and RE_{SNR}

4 结 论

利用视频帧内稀疏性和帧间相关性, 结合了视频压缩感知和视频 ARMA 模型, 提出了一阶 ARMA 的视频压缩感知模型, 并分析得出了其构造条件。该模型对于帧间相关性较高和帧内满足 RIP

稀疏度准则的视频来说具有较强的实用性, 实验表明当视频压缩比达 100 ~ 200 时, 仍然能较好的重建。同时也和经典 ARMA 视频模型、coarse-to-fine 视频模型、Cevher 基于静态背景的压缩感知模型及 Chang 的 4D MAR 模型做了对比, 在相同重建信噪比情况下, CS-ARMA 模型的压缩比更高。这对森林航拍、交通监控有着重要的实用意义。下一步研究重点主要是将视频 CS 理论和高阶的 ARMA 视频模型结合, 进一步提高模型精度, 以增强视频重建质量, 同时希望借鉴美国 Rice 大学 Baraniuk 教授等人研制单像素相机^[21]和可编程像素相机^[22]原理, 实现硬件方面的突破。

参考文献 (References)

- [1] Donoho D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [2] Candès E J. The restricted isometry property and its implications for compressed sensing [J]. Comptes Rendus Mathématique, 2008, 346(9): 589-592.
- [3] Stankovic V, Stankovic L, Cheng S. Scalable compressive video [C]//Proceedings of the 18th IEEE International Conference on Image Processing. TX, USA: IEEE, 2011: 921-924.

- [4] Wakin M, Laska J, Duarte M, et al. Compressive imaging for video representation and coding [C] // Picture Coding Symposium. Houston, USA: Rice University, 2006: 303-308.
- [5] Park J Y, Wakin M B. A multiscale framework for compressive sensing of video [C] // Picture Coding Symposium. Colorado, USA: IEEE, 2009: 1-4.
- [6] Cevher V, Sankaranarayanan A, Duarte M F. Compressive Sensing for Background Subtraction [M] // Computer Vision-ECCV 2008. Berlin Heidelberg: Springer, 2008: 155-168.
- [7] Kim Y, Oh H, Bilgin A. Video compressed sensing using iterative self-similarity modeling and residual reconstruction [J]. Journal of Electronic Imaging, 2013, 22(2): 021005-021005.
- [8] Chang T, Nadar M S, Ghring J, et al. Non-Contrast-Enhanced 4D MRA Using Compressed Sensing Reconstruction: USA. Patent 20,130,121,550 [P]. 2013-5-16.
- [9] Doretto G, Chiuso A, Wu Y N, et al. Dynamic textures [J]. International Journal of Computer Vision, 2003, 51(2): 91-109.
- [10] Hadid A, Pietikinen M, Li S Z. Learning Personal Specific Facial Dynamics for Face Recognition from Videos [C] // Analysis and Modeling of Faces and Gestures. Berlin Heidelberg: Springer, 2007, 4778: 1-15.
- [11] Chan A B, Vasconcelos N. Probabilistic kernels for the classification of auto-regressive visual processes [C] // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. London, UK: IEEE, 2005, 1: 846-851.
- [12] Turaga P, Veeraraghavan A, Chellappa R. Unsupervised view and rate invariant clustering of video sequences [J]. Computer Vision and Image Understanding, 2009, 113(3): 353-371.
- [13] Simon J, Peter J. Digital Audio Restoration—a Statistical Model Based Approach [M]. Cambridge, USA: MIT Press, 1998: 122-126.
- [14] Aggarwal G, Chowdhury A K R, Chellappa R. A system identification approach for video-based face recognition [C] // Proceedings of the 17th International Conference on Pattern Recognition. Barcelona, Spain: IEEE, 2004, 4: 175-178.
- [15] Van Overschee P, De Moor B. N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems [J]. Automatica, 1994, 30(1): 75-93.
- [16] Candès E J, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information [J]. IEEE Transactions on Information Theory, 2006, 52(2): 489-509.
- [17] Van Den Berg E, Friedlander M P. Probing the Pareto frontier for basis pursuit solutions [J]. SIAM Journal on Scientific Computing, 2008, 31(2): 890-912.
- [18] Curtain R F, Pritchard A J. Infinite Dimensional Linear Systems Theory [M]. Berlin: Springer Verlag, 1978: 90-94.
- [19] Baraniuk R G, Cevher V, Duarte M F, et al. Model-based compressive sensing [J]. IEEE Transactions on Information Theory, 2010, 56(4): 1982-2001.
- [20] Van Den Berg E, Friedlander M P. SPGL1: A solver for large-scale sparse reconstruction [EB/OL]. [2013-06-25]. <http://www.cs.ubc.ca/labs/sci/spgl1>, 2007.
- [21] Duarte M F, Davenport M A, Takhar D, et al. Single-pixel imaging via compressive sampling [J]. IEEE Signal Processing Magazine, 2008, 25(2): 83-91.
- [22] Sánchez N, Freshwater D W. Vegetative, reproductive and molecular characterization of *Gloiocladia microspora* (Faucheaecae, Rhod.) [J]. European Journal of Phycology, 2010, 45(1): 63-78.