

基于压缩感知和改进自适应正交匹配的 稀疏信号重构

张宗福

(广东江门职业技术学院, 广东 江门 529090)

摘要: 针对传统香农-奈奎斯特采样定理指出在保证原始信号重构精度的前提下, 采样频率必须为原始信号频率的 2 倍, 提出了一种基于压缩感知理论和改进的自适应正交匹配追踪算法的稀疏信号重构方法; 首先引入了压缩感知模型和信号重构目标函数, 然后在经典正交匹配追踪类算法进行分析和总结的基础上, 为克服其不足, 设计了一种二次筛选支配原子集的方法, 即通过计算信号的 QR 分解并计算具有最大势能的原子从而得到能量候选原子集, 通过计算余量与原子的相关性选出相关性最大的原子从而得到相关候选原子集, 并将能量候选原子集和相关候选原子集的交集作为最终支配原子集; 最后定义了具体的采用自适应正交匹配算法实现信号重构的算法; 在 Matlab 仿真环境下试验, 结果表明: 文章方法能有效地进行稀疏信号重构, 具有较小的重构误差, 且与其它方法相比, 具有收敛速度快和重构效果好的优点。

关键词: 信号重构; 压缩感知; 正交匹配; 噪声

Sparse Signal Reconstruction Based on Compressed Sensing and Adaptive Orthogonal Matching

Zhang Zongfu

(Jiangmen Polytechnic, Jiangmen 529090, China)

Abstract: Aiming at the traditional Shannon-Nyquist sample theorem argued that under the guarantee of signal reconstruction accuracy, the sampling frequency should be two times more than the primitive signal sampling frequency, a sparse signal construction method based on compressed sensing and improved adaptive orthogonal matching was proposed. Firstly, the compressed sensing and signal construction function was set, then the classic orthogonal matching pursuit algorithm was summarized, and the atom re-selection method was designed, namely, by computing the QR decomposition of signal and energy selective set, and by computing the relation between remain value and atom to get the relative selective atom set, and the intersection between energy selective atom set and relative selective set was built as the final supporting atom set. Finally, the algorithm for signal construction was defined. The simulation experiment was operated in Matlab environment, and the result shows the method can realize sparse signal construction with less construction error, and compared with the other methods, it has the advantage of rapid convergence and good construction effect.

Key words: signal reconstruction; compressed sensing; orthogonal matching; noise

0 引言

为有效和快速地对规模巨大的数据量进行处理, 需要首先对信号进行压缩, 然后再基于压缩的信号进行各类操作^[1]。

香农-奈奎斯特采样定理指出: 对于任何一个连续带宽的有限信号, 只有当信号采样频率大于或等于信号带宽的 2 倍时, 才能保证的无损重构^[2-3], 但随着待处理信号的带宽增大时, 香农-奈奎斯特采样定理主要面临着下面两个问题^[4]:

(1) 存储资源、传输资源和计算资源等硬件成本增大而信号获取效率减少;

(2) 当获取数据后, 先压缩再进行存储或传输的处理方式导致了资源利用率的降低。

因此, 如何建立一个有效的信号处理框架, 能以远小于香农-奈奎斯特采样定理中要求的采样频率对原始信号进行无损地采样, 同时保证信号能精确地恢复成原始信号, 是信号处理领域的一个重要研究方向^[5]。

现有的基于压缩感知理论 (compressed sensing, CS)^[6] 进行的信号重构工作主要有: 文献 [7] 基于压缩采样匹配追踪 (CoSaMP) 算法建立了对压缩感知信道进行估计的新方法。文献 [8] 对原子匹配迭代过程进行改进, 其缺点是较传统的正交匹配算法仍然具有较高的时间复杂度。文献 [9] 利用雷达目标信号在多普勒域内的稀疏特征, 由于其针对多普勒域, 因此, 具有通用性不强的缺点。

上述工作均基于压缩感知的信号重构, 具有重要意义, 但仍然存在如效率低、误差大以及时间复杂度大的问题, 因此, 本文在上述工作的基础上, 设计了一种基于改进压缩感知和改进正交匹配算法的稀疏信号重构方法, 具有重构精度高和时间复杂度低的优点, 实验证明了文中方法的有效性。

1 压缩感知模型

当信号为稀疏的或可压缩时, 原始数据信号可以通过少量信号精确恢复来重建, 假设存在一个长度为 N 的离散信号, 则该信号可以通过与一组标准正交基的线性组合表示, 即:

$$\mathbf{X} = \sum_{i=1}^N \alpha_i \boldsymbol{\varphi}_i \quad (1)$$

式 (1) 中, $\boldsymbol{\varphi}_i$ 为列向量, 是 $N \times N$ 矩阵 $\boldsymbol{\Psi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_N]$ 中的

收稿日期: 2014-01-20; 修回日期: 2014-02-20。

作者简介: 张宗福 (1981-), 男, 广东南雄人, CCF 会员, 硕士, 讲师, 主要从事无线传感器网络和计算机应用方向的研究。

列向量, α_i 是原始信号 X 在域 Ψ 上的等价表示, 当 α_i 中大部分的元素为 0, 而只有少部分元素不为 0 时, 则 α_i 是稀疏的, 当 α_i 中仅有 K 个元素不为 0 且 $K \leq N$, 则称信号 X 为 K 稀疏的。

由于信号 X 的稀疏性, 因此, 可以构造一个与变换基 Ψ 不相关的测量矩阵, 其满足有限等距性质 (Restricted Isometric Property, RIP 准则), 即:

$$\begin{cases} (1 - \delta_k) \|z\|_2^2 \leq \|Ab\|_2^2 \leq (1 + \delta_k) \|z\|_2^2 \\ \delta_k \in (0, 1) \end{cases} \quad (2)$$

选取 $M(M \leq N)$ 个观测值构成 $M \times 1$ 的向量 Y , 并在接收端通过观察向量 Y 来精确重建原始信号, 上述过程可以描述为:

$$Y = \Phi X = \Phi \Psi \alpha \quad (3)$$

式中, Φ 为 $M \times N$ 的测量矩阵。

信号 X 的重构可以描述为利用 $M \times 1$ 的观测值向量 Y 来重构具有 N 个元素的信号 X , 由于 $M \leq N$, 因此, 式 (2) 是欠定的, 即具有多个满足条件的解, 为了精确地对原始信号 X 进行重构, 可以通过求解最小 l_0 范数来根据测量值重构信号:

$$\begin{cases} \alpha = \operatorname{argmin} \|\alpha\|_0 \\ s. t. Y = \Phi X = \Phi \Psi \alpha \end{cases} \quad (4)$$

式中, $\|\alpha\|_0$ 为向量 α 中非零元素的个数, 因此求解最小 l_0 范数是一个 NP-hard 问题, 根据最小 l_0 范数在一定程度上与最小 l_1 范数具有等价性, 将 (4) 可以修改为:

$$\begin{cases} \alpha = \operatorname{argmin} \|\alpha\|_1 \\ s. t. Y = \Phi X = \Phi \Psi \alpha \end{cases} \quad (5)$$

从上述可知, 压缩感知理论主要包含 3 方面的问题即信号的稀疏表示、设计满足 RIP 条件约束的测量矩阵以及重构算法的设计。而其中重构算法就是根据观测值和测量矩阵从低维信号中最大程度地回复出原始的高维数据信号, 因此, 重构算法是压缩感知模型最重要和最关键部分。

2 匹配追踪算法

匹配追踪算法是一类贪婪追踪算法, 主要有: MP (Matching Pursuit, MP) 算法^[10], 正交匹配 OMP 算法 (Orthogonal Matching Pursuit, OMP) 算法^[11] 和正则化正交匹配追踪算法 (Regularized Orthogonal Matching Pursuit, ROMP)^[12]。

MP 算法的基本原理为在算法的每轮迭代中进行原子选择和余量计算更新, 选择测量矩阵 Φ 中与余量内积绝对值最大的原子作为匹配原子, 实现原始信号的稀疏逼近并重新计算余量, 重复上述过程直到原始信号重构完毕。MP 算法能在迭代过程中保证余量与选择的匹配原子的正交, 但不能保证余量与原来选择的所有匹配原子正交, 因此, 需要较多的迭代次数才能保证重构精度。

OMP 算法在 MP 算法基础上对其进行改进, 即在选取了最匹配原子后, 采用 Gram-Schmidt 正交方法对原子进行正交化, 然后将数据投影到所有已选正交原子上, 得到数据信号的分量和余量, 其它过程均与 MP 算法相同, 但由于其不会重复选择原子, 在较少的迭代次数后就能有效重构原始图像, 因此, 运行效率高于 MP 算法。

ROMP 算法从测量矩阵中挑选多个原子作为候选集, 然后从候选集中按照正则化原则挑选出匹配原子放入最终的支撑集, 对于稀疏度为 K 的信号, 当余量与测量矩阵中原子的相关系数为 $\rho \in \mathbf{R}^{M \times 1}$, 将其中相关系数最大的 K 的原子作为候选集 S , 并分为若干组, 对于 S 中的任意一个分组 F , 若 F

中任意的两个原子索引 $i, j \in (1, m)$, 都满足:

$$|\rho_i| \leq 2 |\rho_j| \quad (6)$$

则选择 S 中具有最大能量的原子集 F 作为最终的候选集 Φ_F , 并采用最小二乘法计算信号 X 的逼近 \bar{X} 以及对应的余量 r , 如下所示:

$$\begin{cases} \bar{X} = \operatorname{argmin} \|y - \Phi_F X\|_2 \\ r = y - \Phi_F \bar{X} \end{cases} \quad (7)$$

ROMP 算法能保证最终支撑集中的原子能量大于其余未选中的测量矩阵中的原子能量, 在经过 K 次迭代就能较为精确地对信号进行重构, 其缺点是需要信号的稀疏度, 否则会影响重构质量。

3 改进的自适应正交匹配追踪算法

3.1 改进的自适应正交匹配追踪算法原理

为了克服上述匹配追踪算法的不足, 文中设计了一种改进的自适应正交匹配追踪算法, 其主要改进之处主要包含选择能量最大且满足稀疏性的原子构成能量候选原子集、选择与余量相关性最大的一些原子构成相关性候选原子集, 将候选原子集和相关性候选原子集的交集作为最重的支配集。

3.2 QR 分解选择能量候选原子集

采用 QR 分解根据原子能量选择能量候选原子集, 借鉴 ROMP 算法的正则化思想, 采用 QR 分解法来增大测量矩阵的奇异值, 从测量矩阵中选出能量最大且满足稀疏性的原子, 原信号 X , 假设测量矩阵为 $M \times N$ 的 Φ , 则原始信号 X 在 N 维空间运用 QR 分解法进行分解, 如下所示:

$$\Phi = R^T Q^T \quad (8)$$

式中, 矩阵 R 中位于主对角线上的元素值远大于其它值, 因此, 保留主对角线上的元素, 将其他元素值设置为 0, 得到新的矩阵, 如下所示:

$$\Phi' = R'^T Q^T \quad (9)$$

式中, R' 为对角矩阵, 当前的原子索引集为 F , 且其主对角线上每个元素值 R'_{ii} 可以通过下式求取:

$$R'_{ii} = \frac{1}{\sqrt{\Phi'^T (\Phi \Phi')^{-1} \Phi'}} \quad (10)$$

因此, 在满足式 (2) 所示的 RIP 准则的前提下, 则信号在 l_2 范数下的精确稀疏解可以表示为:

$$X' = \Phi'^T (\Phi \Phi')^{-1} \Phi' \quad (11)$$

计算原始信号在各原子上势能:

$$\Delta E_i = E(X_1, \dots, X_i, \dots, X_m) - E(X_1, \dots, X_i', \dots, X_m) \approx \lambda |X_i - X_i'| \quad (12)$$

按照下式依次选择具有最大势能的原子 φ_i 作为该轮选中的原子, 直到本轮选择的原子数达到 d 倍原子总数。

$$i = \operatorname{arg} \max_i \Delta E_i \quad (13)$$

将每次选择的 φ_i 加入到能量候选原子集中, 即最终的能量候选原子索引集 F_e 和能量候选原子集 Φ_{F_e} 为:

$$\begin{cases} F_e = F_e \cup \{i\} \\ \Phi_{F_e} = \Phi_{F_e} \cup \{\varphi_i\} \end{cases} \quad (14)$$

3.3 匹配相关候选原子集

相关候选原子集即选择与余量相关性较大的原子获得最终的支撑集, 通过设计匹配策略, 即:

$$|[r, \Phi_{F_e}]| \geq d * \max |[r, \Phi] \quad (15)$$

式中, Φ_{F_e} 为当前选择的原子集, d 为匹配因子, 在迭代过程中

行自适应地随着余量的变化而进行调整:

当余量收敛程度低于某一预设阈值 th_1 时, 使 d 随迭代次数增加非线性下降, 以增加最终支配集中原子的个数, 如下所示:

$$d_{l+1} = d_l * e^l \quad (16)$$

当与余量差的收敛程度高于某阈值时 th_2 , 为了防止原子选择陷入局部最优值, 使 d 随迭代次数增加非线性增加, 以防止最终支配集中原子个数增长过快:

$$d_{l+1} = (1 + d_l) * e^l \quad (17)$$

当选择出 Φ_{F_ρ} 后, Φ_{F_ρ} 就作为当前的相关性候选原子集。

3.4 支配原子集

当能量候选原子集和相关候选原子集都计算完毕后, 就将两者的交集作为最终的原子支配集, 然后通过不断迭代, 即采用最小二乘法计算信号的估计值, 在此基础上更新余量, 直到余量满足迭代终止条件。

文中改进算法的总体框架可以描述为图 1 所示。

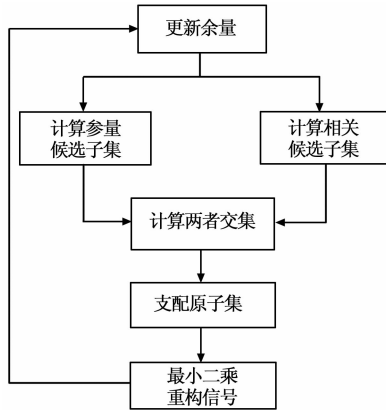


图 1 改进算法总体框架

4 算法描述

文中基于两次选择候选原子集的改进正交匹配算法可以定义为:

算法 改进的自适应正交匹配追踪算法

初始化: 余量初值 $r_0 = \mathbf{X}$, $M \times N$ 的测量矩阵 Φ , 余量当前值 $r = r_0 = \mathbf{X}$, 匹配因子 d 的初值 d_0 ($1 < d_0 < 1$), 能量候选原子索引集 $F_e = \varphi$, 能量候选原子集 $\Phi_{F_e} = \varphi$, 相关候选原子索引集 $F_\rho = \varphi$, 相关候选原子集 $\Phi_{F_\rho} = \varphi$, 支配原子索引集 F , 支配原子集 Φ_F , 当前迭代次数 $l = 1$, 余量收敛性阈值 th , 余量收敛程度阈值 th_1 和 th_2 ;

(1) 计算能量候选原子集: 根据式 (11) 计算信号在 QR 分解下的稀疏表示, 根据式 (12) 计算使该信号具有最大势能的原子, 并以 d 为选择原子比例选择一定数目的原子, 将其索引加入到 F_e^l , 从而能量候选原子集 $\Phi_{F_e^l}$;

(2) 计算相关候选原子索引集: 计算信号与各原子的内积, 并根据式 (15) 根据参数 d 选择相关性较大的若干原子, 将这些原子的索引加入到集合 F_ρ^l , 从而得到相关候选原子集 $\Phi_{F_\rho^l}$;

(3) 取能量候选原子集和相关候选原子索引集的交集作为支配原子索引集 F^l :

$$F = F_e \cap F_\rho \quad (18)$$

根据下式计算支配原子集 Φ_F :

$$\Phi_{F^l} = \Phi_{F_e^l} \cap \Phi_{F_\rho^l} \quad (19)$$

(4) 根据式 (7) 所示的最小二乘法计算原始信号的重构值 $\bar{\mathbf{X}}_l$ 和对余量 r_l 进行更新;

(5) 判断余量 r 是否满足收敛性阈值 th , 即判断:

$$\|r\|_2 \leq th \quad (20)$$

当上式成立时, 重构信号已经满足要求, 算法结束, 输出支配原子集 Φ_{F^l} 和重构信号 $\bar{\mathbf{X}}_l$;

否则根据下式计算余量的收敛程度:

$$u = \left| \frac{\|r_{l+1}\|_2 - \|r_l\|_2}{\|r_{l+1}\|_2} \right|$$

当 $u \geq th_2$, 采用式 (16) 对匹配因子 d 进行更新;

反之, 采用式 (17) 对匹配因子 d 进行更新;

(6) 当前迭代次数 $l = l + 1$, 并返回 (2) 继续迭代。

5 仿真实验

为了验证文中改进的自适应正交匹配追踪算法的优越性, 在 Matlab 仿真环境^[13]下进行实验, 文中算法参数设置如下: 信号压缩比为 15, 稀疏度设置为 $K = 15$, 匹配因子 d 的初值 $d_0 = 0.1$, 余量收敛性阈值 $th = 0.05$, 余量收敛程度阈值 $th_1 = 0.34$ 和 $th_2 = 0.21$, 采用文中方法对长度 $N = 200$ 的一组语音信号进行重构, 原始信号和重构误差分别如图 2 和图 3 所示。

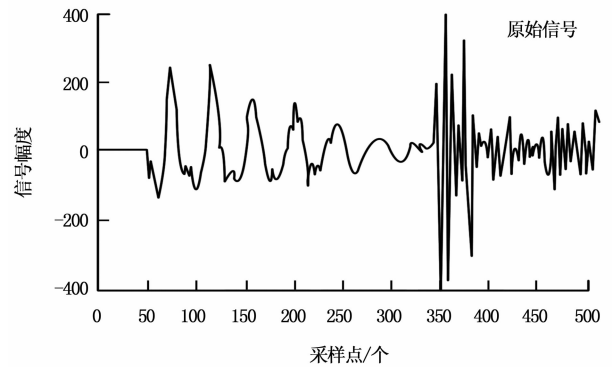


图 2 语音原始信号

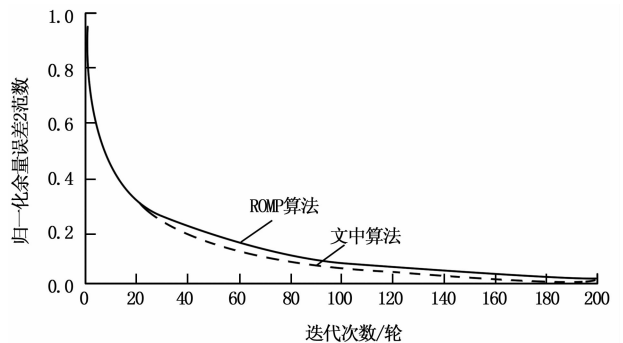


图 3 余量误差 2 范数比较

从图 2 和图 3 可以看出, 文中方法不仅具有较快的收敛速度, 同时具有较小的余量误差。ROMP 算法中求得的余量误差在迭代次数为 20 次之后, 就一直低于文中方法。这是由于文中算法在支配原子集的选择过程中, 选择与余量相关大同时能量最大的原子, 因此, 能较为精确地对原始信号进行重构, 具有较少的重构误差, 同时由于克服了 ROMP 算法的易于先入局部最优的缺点, 因此, 收敛效果也更好。

将 OMP 算法、ROMP、文献 [7]、文献 [8] 和文中方法

对应的峰值信噪比、余量误差 2 范数和运行时间进行比较，结果如表 1 所示。

表 1 不同算法重构结果比较

算法名称	峰值信噪比(dB)	余量误差 2 范数	运行时间(s)
OMP	25.653	0.142 1	31.24
ROMP	26.784	0.094 3	13.535
文献[7]	27.674	0.063 4	10.564
文献[8]	29.454	0.056 3	24.344
文中方法	30.124	0.043 2	14.241

从表 1 中可以看出，文中方法的重构性能最好，这表现在文中方法具有最大的峰值信噪比和最小的余量误差 2 范数，由于文中方法求取了两次候选子集，因此执行时间略长于 ROMP 算法和文献 [7]，综合上述，文中算法在具有较少运行时间的同时能有最小的余量误差和最大的峰值信噪比，因此，文中方法更优。

6 结论

为了有效和快速地对规模巨大的数据量进行处理，需要在对信号进行压缩的基础上实现信号重构，为此，设计了一种基于压缩感知理论和改进的自适应正交匹配算法的信号重构方法。通过在信号重构阶段对候选原子集进行两次筛选，即选择具有最大能量的候选原子集和具有最大相关性的原子集，并取两者的交集作为最终的支配原子集，在此基础上通过重复迭代更新余量实现信号的精确重构，仿真实验证明了文中方法能实现信号的精确重构，具有较小的余量误差，较其他方法具有较大优越性。

参考文献：

[1] Baraniuk R G. Compressive sensing [J]. IEEE Signal Processing Magazine, 2007, 24 (4): 118 - 121.

(上接第 1542 页)

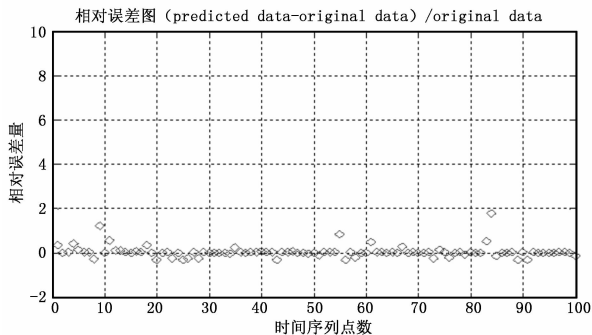


图 5 海杂波预测数据相对误差图 (PSO-SVM)

5 结束语

海杂波具有混沌特性，而支持向量机算法能够有效地对混沌序列进行回归预测。本文采用互信息法和改进的伪邻近点法从雷达实测数据中提取海杂波混沌特性的延迟时间和嵌入维数，再利用相空间重构得到 SVM 训练样本，应用改进的 PSO 算法优化 SVM 的参数，最终建立回归模型，并进行仿真实验。仿真实验结果表明：回归模型预测海杂波是可行的，而 PSO-SVM 法预测海杂波能进一步提高预测精度。

[2] Donoho D L, Tsaig Y, Drori I, etc. Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit [J]. IEEE Transactions on Information Theory, 2012, 58 (2): 1094 - 1121.

[3] 杨真真, 杨震, 孙林慧. 信号压缩重构的正交匹配追踪类算法综述 [J]. 2013, 29 (4): 486 - 489.

[4] Karahanoglu N B, Erdogan H. A orthogonal matching pursuit: Best-first search for compressed sensing signal recovery [J]. Digital signal processing: A review journal, 2012, 22 (4): 555 - 568.

[5] 路锦正, 张启衡, 徐智勇, 等. 超完备稀疏表示的图像超分辨率重构方法 [J]. 系统工程与电子技术, 2012, 34 (2): 403 - 408.

[6] 马春晖. 压缩感知重构算法研究 [D]. 杭州: 杭州电子科技大学, 2012.

[7] 陈艳, 王彪. 基于压缩感知的稀疏信道估计方法 [J]. 计算机工程与设计, 2013, 34 (7): 2338 - 2341.

[8] 白凌云, 梁志毅, 徐志军. 基于压缩感知信号重建的自适应正交多匹配追踪算法 [J]. 计算机应用研究, 2011, 28 (11): 4061 - 4063.

[9] 刘寅, 吴顺君, 张怀根, 等. 一种快速的基于压缩感知的多普勒高分辨方法 [J]. 西安电子科技大学学报 (自然科学版), 2011, 38 (2): 82 - 87.

[10] Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries [J]. IEEE Transactions on Signal processing, 1993, 41 (12): 3397 - 3415.

[11] Tropp J. Greed is good: algorithmic results for sparse approximation [J]. IEEE Trans. Inf. Theory, 2004, 50 (10): 2231 - 2242.

[12] Donoho D L, Tsaig Y, Drori I, et al. Sparse solution of underdetermined linear equations by stage orthogonal matching pursuit [R]. Technical Report, 2006.

[13] 左来. 基于小波包分析和 LS-SVM 的柴油机故障诊断方法 [J]. 计算机测量与控制, 2009, 17 (11): 2150 - 2152.

参考文献：

[1] 张云展. 混沌学与海杂波 [J]. 火控雷达技术, 2008, 37 (4): 1 - 6.

[2] 张天文, 袁山增. 基于分形理论的目标航路预测 [J]. 指挥控制与仿真, 2008, 30 (2): 21 - 23.

[3] 马晓光, 胡非. 利用支撑向量机预报大气污染物浓度 [J]. 自然科学进展, 2004, (3): 349 - 353.

[4] 张小花, 陈玮. 基于相空间重构的制造系统混沌研究 [J]. 机械科学与技术, 2013, 32 (2): 309 - 312.

[5] 李中群, 裴承鸣. 一种新的多嵌入延迟相空间重构方法 [J]. 计算机工程与应用, 2009, 45 (33): 120 - 123.

[6] 混沌-支持向量机回归在流量预测中的应用研究 [J]. 计算机科学, 2009, 36 (7): 243 - 245.

[7] 刘楠, 黄金全. 应用改进粒子群算法的涡轮发动机性能寻优 [J]. 南京航空航天大学学报, 2013, 45 (3): 303 - 308.

[8] 邵信光, 杨慧中, 陈刚. 基于粒子群优化算法的支持向量参数选择及其应用 [J]. 控制理论与应用, 2006, 23 (5): 740 - 748.

[9] 尹波, 夏静波, 付凯, 等. 基于 IPSO 混沌支持向量机的网络流量预测研究 [J]. 计算机应用研究, 2012, 29 (11): 4293 - 4299.

[10] 田海累, 李洪儒, 许葆华. 基于改进人工鱼群算法的支持向量机预测 [J]. 计算机工程, 2013, 39 (4): 222 - 225.

[11] 姜斌, 王宏强, 黎湘. S 波段雷达实测数据海杂波混沌分型特性分析 [J]. 电子与信息学报, 2007, 29 (8): 1809 - 1812.

[12] 宋绍楼, 陈龙虎, 陈晓菊, 等. 基于粒子群多峰值 MPPT 算法的光伏系统研究 [J]. 计算机测量与控制, 2012, 20 (5): 1354 - 1356.