

用于 AIS 实时信号稀疏表示的追踪算法研究

怀率恒, 张淑芳*

(大连海事大学 信息科学技术学院, 辽宁 大连 116026)

摘要: 为解决非周期船舶自动识别系统(AIS)信号的载波跟踪问题,利用信号稀疏表示方法对 AIS 实时信号进行重构,以获得 AIS 信号的信道信息,从而为实现定位信息的测量打下基础.从 AIS 信号特点和 AIS 实时信号处理的要求出发,对基于 K-SVD 算法所构造的自适应冗余字典,从信号处理时间、稀疏表示精度、误码率和受噪声的影响几方面对比了基追踪(BP)和正交匹配追踪(OMP)两种不同的追踪算法.实验结果表明,两种算法都能有效精确地重构 AIS 信号,BP 算法相比 OMP 算法具有更好的稀疏表示精度和误码率,但是信号处理时间会更长.由于 AIS 对实时性的高要求,OMP 算法更适合用于 AIS 自主定位系统.

关键词: 船舶自动识别系统(AIS);稀疏表示;实时信号;基追踪(BP);正交匹配追踪(OMP)

中图分类号: U675.7 **文献标识码:** A **doi:** 10.7511/dllgxb201805014

0 引言

船舶自动识别系统(automatic identification system, AIS)由岸基设施和船载设备共同组成,是一种数字助航系统,可以实现识别船只、简化信息交流以及提供其他辅助信息以避免碰撞的发生等功能^[1].为了实现国际海事组织关于将来船舶应装备天基和陆基双备份定位系统以及为了沿海船舶航行安全将要使用陆基定位系统的迫切要求,Hu 等利用已有 AIS 岸站进行定位,使其成为沿海船舶陆基无线定位系统,从而使船载 AIS 设备能够发挥通信和定位两种功能,称其为 AIS 自主定位系统^[2].然而现有的 AIS 本质上是一个通信系统,在最初的系统设计和建设中并没有考虑实现定位功能的要求,因此利用 AIS 岸站进行定位需要解决许多技术问题,载波的相位测量技术就是其中之一^[3-6].目前应用的定位系统的载波都是双相调制的同频载波,能够利用载波相位测量技术得到高精度的定位.但是 AIS 的载波根据通信的需求是双频点的高斯滤波最小移频键控

(GMSK)调制方式,比同频双相调制要复杂,其非同频和非周期特性使得传统定位系统的载波跟踪测量方法不能应用于 AIS 信号.因此,需要一种新的针对 AIS 信号的载波测量技术,称之为 AIS 信号全息相关检测技术.由于 AIS 是一个实时系统,要测量其载波,首要的问题是要获得 AIS 信号在一定时间间隔内的全息数据.为了解决这个问题,本文将信号稀疏表示理论首次引入到实时信号处理领域.

Mallat 和 Zhang 在 1993 年提出了信号在冗余字典上进行稀疏表示的思想^[7].所谓冗余字典是一个超完备的冗余函数库,信号在冗余字典上进行稀疏分解的结果是该信号的展开中只有少数的基函数具有比较大的非零系数,其他大部分基函数的系数等于零.这里的基函数就是字典中的元素,也被称为原子,因此该信号的主要特征就可以由少量原子来表示.

当前稀疏表示更多的是应用在图像处理领域,其处理时间相对于信号的实时处理而言,可以

收稿日期: 2018-03-09; 修回日期: 2018-07-28.

基金项目: 国家自然科学基金资助项目(61231006,61501078,61501079).

作者简介: 怀率恒(1989-),男,博士生,E-mail: huaishuaiheng@dmlu.edu.cn; 张淑芳*(1955-),女,博士,教授,博士生导师,E-mail: sfzhang@dmlu.edu.cn.

认为是后处理. 本文的研究内容是将图像处理领域的方法引入到信号实时处理领域, 主要的难点在于算法在不降低性能和精度的基础上, 实现快速处理, 满足信号实时处理要求. 本文使用基追踪 (basis pursuit, BP) 和正交匹配追踪 (orthogonal matching pursuit, OMP) 算法, 对基于 K 项奇异值分解 (K -SVD) 算法构造的冗余字典, 获得 AIS 信号的稀疏表示, 并且对比解调之后获得的 AIS 二进制序列, 从信号处理时间、稀疏表示精度、解调误码率和受噪声影响等方面对比两种算法, 为 AIS 信号的全息相关检测技术提供参考.

1 AIS 信号稀疏表示原理

1.1 AIS 信号模型

AIS 自主定位系统由一个主 AIS 基站、多个从 AIS 基站和车载 AIS 设备组成, 系统配置如图 1 所示. 所有的 AIS 基站都需要时钟同步, 主基站和从基站轮流发送信号, 车载 AIS 设备接收来自基站的信号.

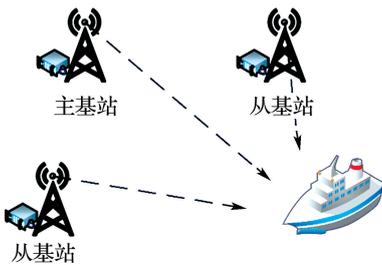


图 1 AIS 自主定位系统配置

Fig. 1 AIS autonomous positioning system configuration

由于 AIS 信号的调制方式是 GMSK, 其信号模型可以等价于 GMSK 信号模型, 表示为

$$s(t) = A \cos[\omega_c t + \theta(t)] = A [\cos \theta(t) \cos \omega_c t - \sin \theta(t) \sin \omega_c t] \quad (1)$$

其中相位 $\theta(t)$ 如下式所示:

$$\theta(t) = \int_{-\infty}^t \omega d\tau = 2\pi h \int_{-\infty}^t \sum_{n=-\infty}^{\infty} a_n g(\tau - nT_b) d\tau = 2\pi h \sum_{n=-\infty}^{\infty} a_n \int_{-\infty}^t g(\tau - nT_b) d\tau \quad (2)$$

其中 $g(t)$ 为高斯滤波器的矩形脉冲响应.

GMSK 基带波形的信号特征完全由高斯滤波器确定, 该高斯滤波器是一个单位脉冲响应服从高斯分布并且均值为 0 的低通滤波器, 其单位

脉冲响应为

$$h(t) = \frac{1}{\sqrt{2\pi}\delta} \exp\left(-\frac{t^2}{2\delta^2}\right) \quad (3)$$

其中 δ^2 是高斯分布的方差. 在信号处理领域, 时间带宽积 BT 通常用来表征高斯滤波器, 其中 B 为高斯滤波器 3 dB 带宽, T 是比特周期. BT 和 δ 满足以下关系:

$$\delta = \sqrt{\ln 2} / 2\pi BT \quad (4)$$

将式(4)代入式(3), 可以得到单位脉冲响应的另一个表达式:

$$h(t) = \sqrt{\frac{2\pi}{\ln 2}} BT \exp\left[-\frac{2\pi^2 t^2}{\ln 2} (BT)^2\right] \quad (5)$$

因此, 从式(5)可以看出, GMSK 通信系统中基带波形的信号特征仅与参数 BT 有关. 国际电信联盟给出的建议书^[8]规定, 在 AIS 中, 用于传输数据的 GMSK 调制器 BT 最大应为 0.4, 用于数据接收的 GMSK 解调器的 BT 最大应为 0.5. 为了统一计算, 本文中的 BT 都设为 0.3. 建议书^[8]还规定, AIS 数据传输用 24 bits 的训练序列开始, 训练序列中包括一段同步比特, 由交替的 0 和 1 组成, 使用非归零反向编码, 可以用 1 或 0 开始, 如图 2 所示, 图中纵坐标是归一化幅度.

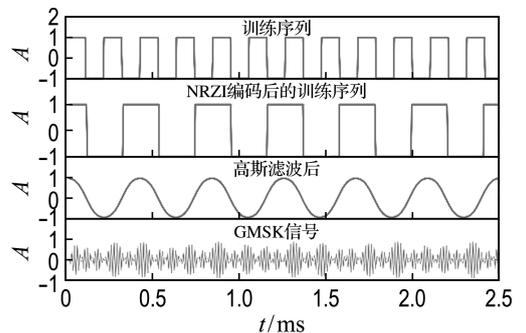


图 2 AIS 信号训练序列

Fig. 2 Training sequence of AIS signal

1.2 稀疏表示理论

根据文献^[9]所述, 信号的稀疏表示源于非线性逼近理论, 可以总结为给定一个字典 $\mathbf{D} = \{d_i, i=1, 2, \dots, M\}$, 它的原子是张成整个希尔伯特空间的单位矢量, 即 $\mathbf{R}^N = \text{span}(\mathbf{D}), M \gg N$. 对于任意的信号 $x \in \mathbf{R}^N$, 在 \mathbf{D} 中自适应地选取 K ($K \ll N$) 个原子对信号 x 作 K 项逼近:

$$\mathbf{x}_K = \sum_{i \in I_k} \boldsymbol{\alpha}(i) d_i \quad (6)$$

其中 I_k 是 d_i 的下标集合, $\boldsymbol{\alpha} = (\boldsymbol{\alpha}(1) \cdots$

$\boldsymbol{\alpha}(M))^T$ 是分解系数向量, 逼近绝对误差定义为 $\epsilon_m = \|\mathbf{x} - \mathbf{x}_K\|_2$. 当 \mathbf{D} 是正交基时, 保留 $|\boldsymbol{\alpha}(i)|$ 最大的 K 个原子, 就得到了信号 \mathbf{x} 的最佳 K 项逼近. 当 \mathbf{D} 是冗余字典时, 式(6)有多个解, 信号的稀疏表示就是从中选出 K 取值最小的即分解系数最稀疏的一组原子, 这等同于解决 0-范数问题, 对于一个随机的冗余字典来说, 这是一个 NP (non-deterministic polynomial) 问题^[9].

1.3 基于 K-SVD 的冗余字典

矩阵的奇异值通常对应于矩阵中隐藏的信息, 重要性与奇异值的大小正相关, K -SVD 算法正是基于这个概念. 在 AIS 全息相关检测中, 需要获得一段时间间隔内的 AIS 信号, 并且不能丢失信号的主要特征. K -SVD 算法对于给定的一组训练信号, 能够自适应地按照稀疏约束条件, 根据训练信号提取其特征, 训练出稀疏表示的冗余字典. 其算法模型可以描述为

$$\begin{aligned} \min_{\mathbf{D}, \mathbf{X}} \{ \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 \} \\ \text{s. t. } \forall i, \|\mathbf{x}_i\|_0 \leq T_0 \end{aligned} \quad (7)$$

其中 \mathbf{D} 为待训练的字典, \mathbf{Y} 为训练信号, \mathbf{X} 是稀疏系数, T_0 是稀疏度.

K -SVD 能在不影响稀疏性限制的情况下保证均方误差减小或者不变, 迭代之后的均方误差是单调递减的, 这也保证了 K -SVD 收敛于一个局部最小值, 但是这并不代表 K -SVD 算法的收敛性是一定成立的, 而是依赖于追踪算法的收敛性. 然而在 T_0 足够小的时候, BP 和 OMP 算法都有非常好的表现, 其收敛性是能够保证的^[10].

2 追踪算法

如 1.2 节所述, 求解式(6)是一个 NP 问题. 表面上看似似乎是无望的, 但是由于信号是稀疏的这个前提, 这个问题是可以解决的. 研究者提出了许多获取信号稀疏逼近的追踪算法, 通常分为贪婪追踪算法和凸松弛法^[11]. 贪婪追踪算法的基本思想是每次迭代时求取一个局部最优解, 逐步逼近原始信号, 主要代表有 OMP 算法. 凸松弛法则

是将非凸问题转化为凸问题, 并对其求解找到原信号的逼近, 代表性的是 BP 算法.

2.1 BP 算法

BP 算法的基本思想是找到系数具有最小 0-范数的信号表示, 其求解模型可以描述为

$$\begin{aligned} \min \|\mathbf{X}\|_0 \\ \text{s. t. } \mathbf{Y} = \mathbf{D}\mathbf{X} \end{aligned} \quad (8)$$

由于对于式(8)的求解是一个 NP 问题, 根据 1-范数在一定条件下和 0-范数具有等价性的条件, 求解一个更加简单的 1-范数问题会得到同样的解^[12].

$$\begin{aligned} \min \|\mathbf{X}\|_1 \\ \text{s. t. } \mathbf{Y} = \mathbf{D}\mathbf{X} \end{aligned} \quad (9)$$

这使得问题变成了一个凸优化问题. 由于式(9)中的 \mathbf{X} 没有非负约束, 要将 \mathbf{X} 变为两个非负变量 \mathbf{u} 和 \mathbf{v} 的差, 即 $\mathbf{X} = \mathbf{u} - \mathbf{v}$, 这样 \mathbf{X} 既可以是正也可以是负, 因此式(9)中的约束条件可以写为 $\mathbf{D}(\mathbf{u} - \mathbf{v}) = \mathbf{Y}$, 等同于 $(\mathbf{D} \quad -\mathbf{D}) \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \mathbf{Y}$. 根据范数的定义, 式(9)中目标函数可以写为 $\|\mathbf{X}\|_1 = \sum_i |X_i| = \sum_i |u_i - v_i|$.

因此, 这个 1-范数问题可以化简为线性规划问题:

$$\begin{aligned} \min \mathbf{c}^T \mathbf{x} \\ \text{s. t. } \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \end{aligned} \quad (10)$$

其中 $\mathbf{c}^T = (1 \quad 1 \quad \cdots \quad 1)_{1 \times 2p}$, $\mathbf{x} \in \mathbf{R}^{2p}$, $\mathbf{A} = (\mathbf{D} \quad -\mathbf{D})$, $\mathbf{b} = \mathbf{Y}$.

1-范数的性质决定了其无法分辨稀疏系数的位置, 所以尽管整体上重构信号在欧氏距离上是逼近原信号的, 但是存在位置错位的现象, 从而出现一些不期望的人工效应. 并且在 1-范数的框架下, BP 算法计算复杂度的量级在 $O(N^3)$, 这对于 AIS 实时信号处理, 其运算时间需要进行实验验证.

2.2 OMP 算法

匹配追踪算法的本质是以贪婪迭代算法的思想来选择测量字典矩阵中的列向量. 首先从字典矩阵 \mathbf{D} 中选择一个与原始信号 \mathbf{Y} 最匹配的原子, 构建稀疏逼近, 求出信号残差, 然后继续选择一个

与信号残差最匹配的原子,反复迭代;最后原始信号 \mathbf{Y} 则可以由这些原子的线性组合和最后的残差值来表示. 如果最后的残差值满足预先设置的阈值,在可忽略的范围内,原始信号 \mathbf{Y} 就可使用这些原子的线性组合来表示.

OMP 算法继承了匹配追踪算法原子选择的思想,但是为了克服匹配追踪算法的非最优性问题,其每次迭代过程中均对选择的原子进行施密特正交化,从而保证了每次迭代之后的残差与选中的原子都是正交的. 这样就避免了重复选择,保证了迭代的最优性,同时减少了迭代次数,在精度要求相同的情况下,加快了收敛速度. OMP 算法计算复杂度的量级是 $O(NK^2)$.

OMP 算法的求解模型跟式(8)描述的 BP 算法一样. 算法中当前残差 \mathbf{r} 的定义如下:

$$\mathbf{r}_k = \mathbf{Y} - \mathbf{D}\hat{\mathbf{X}} \quad (11)$$

其中 $\hat{\mathbf{X}}$ 代表由当前字典中的原子得到的原信号的近似解. 当残差或稀疏度达到要求的时候,迭代停止.

3 实验结果及分析

通过仿真 AIS 信号来对比两种追踪算法的性能. 如图 3 所示,在 AIS 默认传输分组中,数据部分的长度为 168 bits,以此为例产生 64 组长度为 168 bits 的 AIS 信号作为训练数据. AIS 是一个实时系统,信号处理的时间必须满足要求,AIS 中使用帧的概念,一帧的时长等于 1 min,分成 2 250 个时隙,一个时隙的时长约为 26.7 ms. 因此要获取 AIS 信号的数据,对其进行稀疏表示必须在一个时隙之内完成^[8]. 本文中实验平台的硬件条件是 Intel Core i7-6700 CPU@3.40 GHz、16 GB 内存.

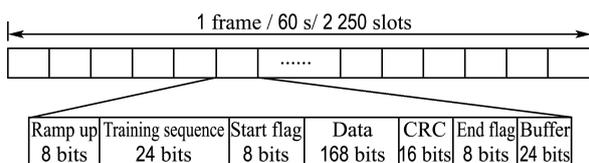


图 3 AIS 信号结构

Fig. 3 Structure of AIS signal

图 4 和 5 分别给出了原始 AIS 信号与使用

BP 和 OMP 算法获得的稀疏信号的对比图,截取了前 2 ms 的波形. 考虑到噪声的影响,信噪比设置为 10 dB. 图 6 和 7 则是 AIS 二进制消息与从两种稀疏信号解调获得的二进制消息的对比图.

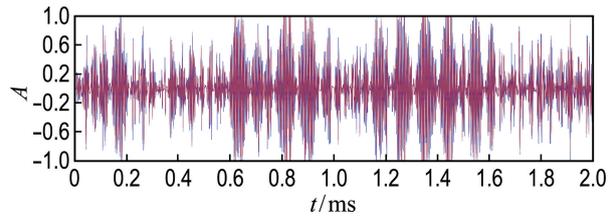


图 4 原始 AIS 信号与 BP 稀疏信号对比

Fig. 4 Comparison of original AIS signal and BP sparse signal

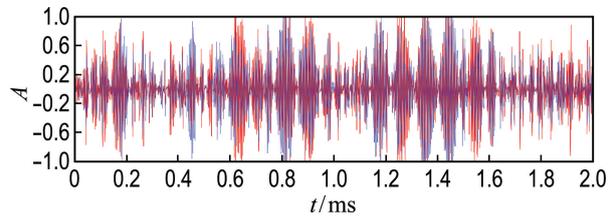


图 5 原始 AIS 信号与 OMP 稀疏信号对比

Fig. 5 Comparison of original AIS signal and OMP sparse signal

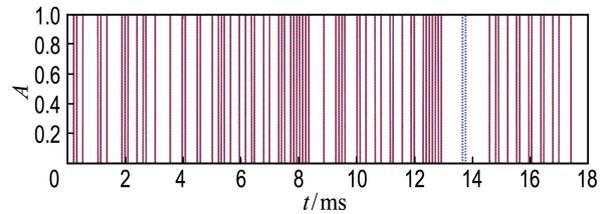


图 6 AIS 二进制消息与解调 BP 稀疏信号获得的二进制消息对比

Fig. 6 Comparison of AIS binary message and binary message from demodulated BP sparse signal

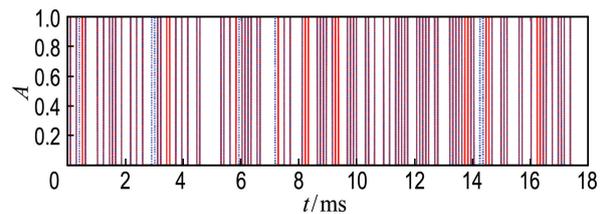


图 7 AIS 二进制消息与解调 OMP 稀疏信号获得的二进制消息对比

Fig. 7 Comparison of AIS binary message and binary message from demodulated OMP sparse signal

4幅图中的纵坐标是归一化幅度,红色实线代表的是原始信号,蓝色虚线代表的是得到的信号。

可以看出,图4中蓝色虚线代表的信号与红色实线代表的信号的重合程度是高于图5中的。为了使观察结果趋于稳定,重复实验500次,比较两种算法的运算时间和稀疏表示精度,得到BP算法的用时为24.4 ms,精度用均方根误差来表示是0.016;OMP算法的用时为13.9 ms,精度是0.25。从实验结果可以看出,由于BP算法的运算复杂度比OMP算法的运算复杂度高,相应的其运算时间也比较长,两种算法的运算时间都满足AIS实时信号处理的时间要求。BP算法的稀疏表示精度高于OMP算法的稀疏表示精度。从图6和7可看出,图6中解调之后的二进制码与原二进制码只有一个不同,而图7中有11个不同。同样的,重复实验500次,比较两种算法的解调之后的误码率,得到BP算法的误码率是0.6%,OMP算法的误码率是7.7%。AIS应用的是循环冗余校验,可以接受的误码率是10%左右^[13]。

表1给出了在信噪比不同的情况下,两种算法的误码率。

表1 不同信噪比下误码率

Tab.1 Bit error rate under different SNRs

信噪比/dB	误码率/%	
	BP	OMP
0	4.1	11.5
5	1.1	9.2
10	0.6	7.7
15	0.12	7.2
20	0.006	6.9

从表1可以看出,在不同信噪比的条件下,BP算法的误码率都低于OMP算法的误码率,但是BP算法的误码率随着信噪比的降低而变化的程度是要高于OMP算法的,这说明BP算法对噪声的变化更敏感。

4 结论

(1)对基于K-SVD算法构造的冗余字典,相比使用OMP算法获得的AIS信号的稀疏表示,使用BP算法获得的AIS信号的稀疏表示具有更

好的精度。

(2)BP算法的信号处理时间比OMP算法的信号处理时间长。

(3)在相同的噪音条件下,BP算法的误码率低于OMP算法的误码率,但是其对噪声的变化更敏感。

(4)AIS是一个实时系统,算法计算的简单快速显得尤为重要,运算时间的多少也代表着算法的可操作性和实现的复杂程度。因此重构算法在保证算法性能的基础上,计算复杂度应尽可能的低。从这一点上来说,OMP算法比BP算法更适合于AIS自主定位系统。

参考文献:

- [1] HU Bin, LIU R W, WANG Kai, *et al.* Statistical analysis of massive AIS trajectories using Gaussian mixture models [C] // **Second International Conference on Multimedia and Image Process**. New York: IEEE, 2017.
- [2] HU Qing, JIANG Yi, ZHANG Jingbo, *et al.* Development of an automatic identification system autonomous positioning system [J]. **Sensors**, 2015, **15**(11):28574-28591.
- [3] JIANG Yi, ZHANG Shufang, YANG Dongkai. A novel position estimation method based on displacement correction in AIS [J]. **Sensors**, 2014, **14**(9):17376-17389.
- [4] ZHANG Jingbo, ZHANG Shufang, WANG Jinpeng. Pseudorange measurement method based on AIS signals [J]. **Sensors**, 2017, **17**:1183.
- [5] WANG Xiaoye, ZHANG Shufang, SUN Xiaowen. The additional secondary phase correction system for AIS signals [J]. **Sensors**, 2017, **17**:736.
- [6] WANG Xiaoye, ZHANG Shufang. Evaluation of multipath signal loss for AIS signals transmitted on the sea surface [J]. **Ocean Engineering**, 2017, **146**: 9-20.
- [7] MALLAT S G, ZHANG Zhifeng. Matching pursuits with time - frequency dictionaries [J]. **IEEE Transactions on Signal Processing**, 1993, **41**(12):3397-3415.
- [8] International Telecommunication Union. Technical Characteristics for An Automatic Identification

- System Using Time-Division Multiple Access in the VHF Maritime Mobile Band; M. 1371-4 [S]. Geneva: Radio Communication Sector of ITU, 2010.
- [9] COHEN A, DALES J P. Nonlinear approximation of random functions [J]. **SIAM Journal on Applied Mathematics**, 1997, **57**(2):518-540.
- [10] RUBINSTEIN R, PELEG T, ELAD M. Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model [J]. **IEEE Transactions on Signal Processing**, 2013, **61**(3):661-677.
- [11] 石光明, 刘丹华, 高大化, 等. 压缩感知理论及其研究进展[J]. 电子学报, 2009, **37**(5):1070-1081.
SHI Guangming, LIU Danhua, GAO Dahua, *et al.* Advances in theory and application of compressed sensing [J]. **Acta Electronica Sinica**, 2009, **37**(5): 1070-1081. (in Chinese)
- [12] CANDES 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.
- [13] 孙 乾, 张陶盛, 李洪星, 等. 基于 CRC 纠错与迭代干扰消除的星载 AIS 接收机[J]. 电讯技术, 2016, **56**(10):1140-1146.
SUN Qian, ZHANG Taosheng, LI Hongxing, *et al.* A satellite-borne AIS receiver based on CRC error correction and iterative interference cancellation [J]. **Telecommunication Engineering**, 2016, **56**(10):1140-1146. (in Chinese)

Research on pursuit algorithms for sparse representation of AIS real-time signal

HUAI Shuaiheng, ZHANG Shufang*

(Information Science and Technology College, Dalian Maritime University, Dalian 116026, China)

Abstract: For the carrier tracking problem of non-periodic automatic identification system (AIS) signal, the signal sparse representation is used to reconstruct the AIS real-time signal to obtain the channel information of AIS signal, thereby lay the foundation for realizing the measurement of positioning information. Based on the characteristics of AIS signals and the requirements of AIS real-time signal processing, for the adaptive redundant dictionary based on the K-SVD algorithm, two kinds of pursuit algorithms, basis pursuit (BP) and orthogonal matching pursuit (OMP), are compared in terms of the signal processing time, sparse representation accuracy, error rate and noise impact. The experimental results indicate that both of the algorithms can effectively and accurately reconstruct AIS signal, the BP algorithm has better sparse representation accuracy and error rate than the OMP algorithm, but the signal processing time is longer. Due to the high requirements of real-time performance of the AIS, the OMP algorithm is more suitable for the AIS autonomous positioning system.

Key words: automatic identification system (AIS); sparse representation; real-time signal; basis pursuit (BP); orthogonal matching pursuit (OMP)