

整体最小二乘法在精同步中的应用

罗海军^{1,2}, 彭卫东¹

(1. 空军工程大学 综合电子信息系统与电子对抗技术研究中心, 西安 710051;

2. 空军工程大学 装备管理与安全工程学院, 西安 710051)

摘要: 为了验证整体最小二乘法 (TLS, Total Least Squares) 比最小二乘法 (LS, Least Squares) 更适用于精同步, 提出了基于整体最小二乘的精同步方法; 该方法以 PN 码基于 BPSK 调制的同步模型为基础, 在接收端得到鉴相曲线后, 对鉴相曲线零值点附近的几个固定位置的数据点利用 QR 分解和奇异值分解的整体最小二乘法拟合出鉴相曲线, 求出同步误差; 最后通过在 Matlab 软件上的仿真, 得出结论: 在噪声环境下, 整体最小二乘法在同步中的应用可以得到优于基于普通最小二乘法的同步法的测量精度和稳定性; 特别是在信噪比较低的条件下, 基于整体最小二乘的精同步测量精度提升了 3 倍多, 同步稳定性提升 5 倍多。

关键词: 整体最小二乘法; BPSK; 精同步; 测量精度; 抗噪声

Application of Total Least Squares in Precise Synchronization

Luo Haijun^{1,2}, Peng Weidong¹

(1. Research Center for Integrated Electronic & Information System and Electronic Countermeasure Technology,

Air Force Engineering University, Xi'an 710051, China; 2. Equipment Management and Safety Engineering

Collge, Air Force Engineering University, Xi'an 710051, China)

Abstract: To verify that the total least-squares method is a better suitable for fine synchronization than the least-squares method, propose the precise synchronization method based on the total least squares. This method is on the base of PN code's BPSK modulation. After the receiver get the phase curve, to get the data of curve's several fixed position near the zero point, then using the total least squares based on QR decomposition and singular value decomposition fitting out the phase line, and estimate the synchronization error. According to the result of the simulation in Matlab software, draw the conclusion; under the noise environment, the Total least squares method in the application of synchronous can be better than that of synchronization based on ordinary least squares method on measuring precision and stability. Especially under the condition of low signal noise ratio (SNR), based on the least squares of synchronous measurement precision is improved more than 3 times, and the stability performance is improved more than 5 times.

Keywords: total least squares (TLS); BPSK; precise synchronization; measuring precision; anti noise

0 引言

PN 码相位的测量精度精同步的关键在于, 测量精度直接影响到同步精度, 测量方法的好坏直接影响通信系统同步性能。目前的 PN 码相位测量方法有很多: 如基于 FFT 的 DSSS 信号 PN 码相位测量法^[1]、两点线性内插法和三点线性内插法^[2], 这些方法由于测量精度都不够高。龚国辉等^[3]提出的三点二次插值法利用 PN 码相关谱谱峰及邻近的 2 个数据点求得二次插值多项式, 有着很高的测量精度。但是由于要专门搜索谱峰所以计算量较大, 而且所用的数据点就是谱峰和附近的两个点, 数据不具有代表性, 抗干扰能力不强。胡修林等^[4]提出来的利用鉴相曲线和最小二乘法直线拟合来估计相位差的方法, 计算量小, 实现简单, 测量精度高, 且抗干扰能力较强, 所以该方法在很多数字通信领域得到运用。近年来整体最小二乘法提出并被证实比最小二乘法有着更好的性能, 但是该方法在精同步中的性能并没有得到验证。

本文介绍了整体最小二乘法在精同步中的应用, 并建立了

基于 BPSK 传输的 PN 码的测量的仿真验证了整体最小二乘法的优势。

1 整体最小二乘法直线拟合

用整体最小二乘法^[5-6]做直线拟合, 既考虑了直线因变量 y 的测量误差, 也考虑自变量 x 的误差。整体最小二乘法的直线方程为:

$$y_i + v_{yi} = \hat{a}(x_i + v_{xi}) + \hat{b} \quad (1)$$

普通最小二乘只考虑的拟合直线与测量值 y 的误差平方和最小, 而整体最小二乘考虑的是各数据点到拟合直线的距离的平方和最小, 得出其拟合准则如下:

$$R(\tau - T_c/2) \quad (2)$$

拟合准则的复杂性决定了 TLS 的参数估计不能如 LS 那样通过极值原理求导出。丁克良^[7]等人提出的用 EIV 模型描述直线方程, 构造增广矩阵 C 并通过矩阵分解, 求出直线的估计参数。

以一组 n 维的数据 $(x_i, y_i), i = 1, 2, 3, \dots, n$ 为例:

首先, 构造增广矩阵:

$$C = \begin{bmatrix} 1 & x_1 & y_1 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix} \quad (3)$$

然后, 对 C 进行 QR 分解:

收稿日期: 2014-02-22; 修回日期: 2014-03-29。

基金项目: 国家部委基金资助课题 (51310020401)。

作者简介: 罗海军 (1990-), 男, 硕士研究生, 主要从事跳频同步技术方向的研究。

$$C = QR \quad (4)$$

Q 是正交阵, 可以通过施密特正交化得出;

$$R = Q^T C = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \end{bmatrix} \quad (5)$$

矩阵满足:

$$R_{11}\hat{b} + R_{12}\hat{a} = R_{13} \quad (6)$$

$$R_{22}\hat{a} = R_{23} \quad (7)$$

再对 R 矩阵的部分做奇异值分解:

令: $C_R = [R_{22} \ R_{23}]$

则奇异值分解结果为:

$$C_R = U \cdot \Sigma \cdot V^T \quad (8)$$

其中: $\Sigma = \text{diag}(\sigma_1, \sigma_2)$, $\sigma_1 > \sigma_2$

参数 \hat{a} 的估计值为:

$$\hat{a} = (R_{22}^T R_{22} - \sigma_2^2)^{-1} R_{22}^T R_{23} \quad (9)$$

回代到方程 (6) 得到参数 \hat{b} 的值:

$$\hat{b} = R_{11}^{-1} (R_{13} - R_{12}\hat{a}) \quad (10)$$

文献[7-8]中都验证了 TLS 方法拟合结果保持在一个方向, 相比 LS 法更加稳定, 但是没有对在噪声条件下的情况进行仿真, 同时由于给出的是拟合的点, 没有一个准确的参考值, 所以对最小二乘法的测量精度比较上, 没有足够的说服力, 本文通过对 TLS 方法在 PN 码相位测量中的应用, 仿真实验是预先设定好的偏差值, 可以准确的比较两种方法的测量精度, 同时也能验证 TLS 更适合 PN 码的相位测量。

2 整体最小二乘在同步中应用

基于最小二乘法的同步方法中, 最小二乘直线拟合主要用于拟合鉴相直线。理想条件下, 本地 PN 码和接收 PN 码之间可以通过相关求得相关峰, 同时通过移位相减可以得到鉴相曲线如图 1 所示。

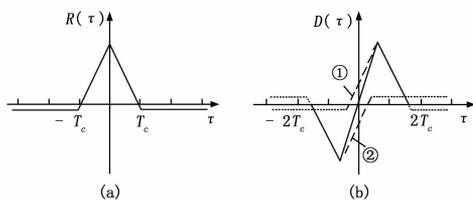


图 1 伪码序列的相关曲线和鉴相曲线

如图 1 所示, 图 1 (a) 是 PN 码的自相关函数 $R(\tau)$, 图 1 (b) 的虚线①是自相关函数 $R(\tau)$ 右移 $T_c/2$ 后的相关函数 $R(\tau - T_c/2)$, 虚线②是自相关函数 $R(\tau)$ 左移 $T_c/2$ 后的自相关函数 $R(\tau + T_c/2)$, 实线 $D(\tau)$ 为 $R(\tau - T_c/2)$ 和 $R(\tau + T_c/2)$ 相减后得到的 PN 序列的鉴相曲线。分析可知, 原点附近的这条鉴相曲线是 (和 (相减得到的, 这条直线的零点横坐标是 PN 码的相位偏差。基于最小二乘同步方法的是用最小二乘直线拟合的方法拟合出这条直线, 从而估算出其零值点的横坐标。在实际的无线通信系统中, 带宽是有限的, PN 序列不可能直接传输, 这样接收端也很难完全还原出 PN 码的原始信号, 相关峰和鉴相曲线也就没有图 1 中那么理想。文献 [9] 中提到用 BPSK 来传输 PN 序列, 但是并没有对该方法做详细说明。本文建立基于 BPSK 的传输模型通过仿真实验验证了 PN 序列通过 BPSK 传输的可行性。其流程图如图 2 所示。

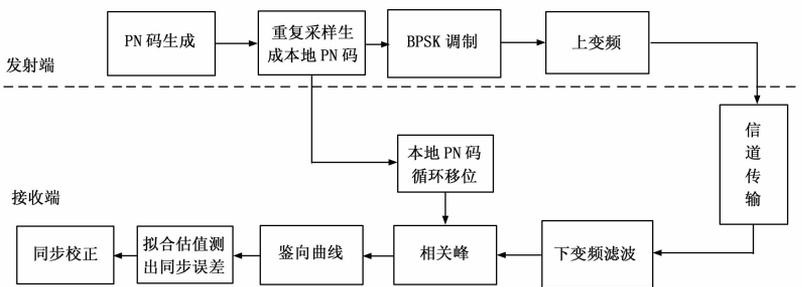


图 2 基于 BPSK 的伪码同步流程

仿真实验使用的到的是 63 bit 的双极性伪码, 采样率 $f_s = 11$, PN 码做 11 次的过采样, 得到的信号如图 3 (a) 所示, 为了便于观察文中只给出了前 20 bit 的示意图。如果直接把如图 3 (a) 中的信号调制成方波进行传输, 带宽有限接收端很难恢复出方波信号。所以发射端对 PN 序列调制成三角函数波效率更高, 本文选择 BPSK 方式来传输序列。为了使每一位伪码的调制波形是同极性的, 同时保证整个序列调制完形成一个连续的三角函数波形, 本文设定每一位伪码根据自己的极性调制成正半周期的正弦波:

其中 PN 序列调制成方波的表达式为:

$$C(t) = \sum_{i=0}^{\infty} a_i \{u[t - (i-1)T_c] - u(t - iT_c)\} \quad (11)$$

其中: a_i 是双极性的伪码序列, 取值为 ± 1 。

这样得到基带的调制信号为:

$$S(t) = C(t) \cdot \left| \sin\left(\frac{\pi}{T_c} t\right) \right| \quad (12)$$

PN 码在发射端调制完成后经过上变频后发送出去, 在接收端经过下变频滤波恢复出如图 3 (b) 的信号, 由于正弦波良好的传输特性接收端能较好的恢复出基带信号。接收端的本地 PN 码和发送端未调制的 PN 码序列是一致的如图 3 (a) 所示, 通过循环移位与接收信号相关, 可以得到相关峰, 进而得出鉴相曲线, 如图 4 所示。经过调制后, 相关峰和鉴相曲线和图 1 有所差别, 但是鉴相曲线在零值点附近依然是一条直线。所以在做直线拟合的时候, 对零值点附近的少数几个采样点做直线拟合可以求出鉴相曲线的相位估值, 这里取鉴相曲线上的第 11 到 16 这 6 个点, 分析可知: 在无相偏的情况下, 第 13 个点是零值点, 所以构造出如下的拟合数据: $\mathbf{X} = [-2, -1, 0, 1, 2, 3], \mathbf{Y} = [pe_{11}, pe_{12}, pe_{13}, pe_{14}, pe_{15}, pe_{16}]$ 。

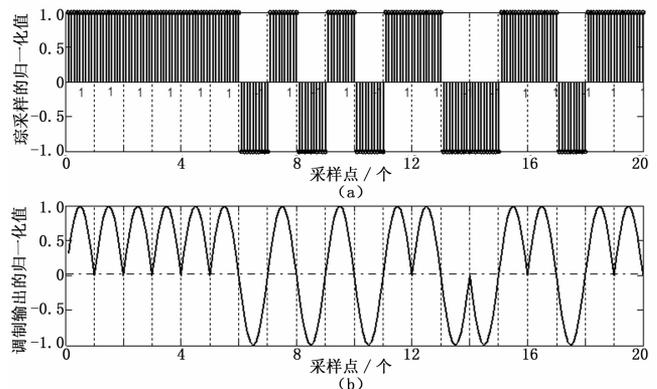


图 3 过采样的 PN 码和 BPSK 调制后的 PN 码

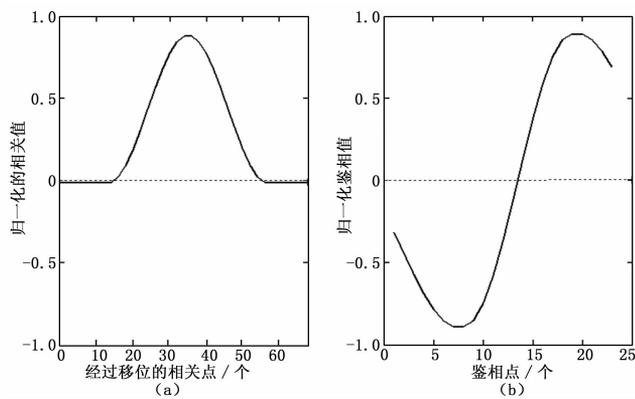


图 4 经过 BPSK 调制后的相关曲线和鉴相曲线

X 是固定值, Y 是对应得鉴相曲线上的测量值。根据式 (3) 构造出增广矩阵如下:

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 & 3 \\ pe_{11} & pe_{12} & pe_{13} & pe_{14} & pe_{15} & pe_{16} \end{bmatrix}^T$$

通过式 (4) ~ (10) 的运算方法运用整体最小二乘法计算出拟合直线的估计参数, 得到直线方程:

$$y = \hat{a}x + \hat{b} \quad (13)$$

直线的零点横坐标, 也就是相位测量值为:

$$P_c = -\frac{\hat{b}}{\hat{a}} \quad (14)$$

整体二乘法直线拟合在精同步中的应用, 源于整体最小二乘法相对于最小二乘同步方法的改进, 由于同时考虑了测量值的误差和采样时间点的误差, 所以可以得到更好的结果, 以上的分析也可以看出, 整体最小二乘法和最小二乘法一样适用于精同步的伪码相位测量。

3 仿真实验

文献 [5-6] 虽介绍了整体最小二乘法的直线拟合方法, 通过分析和数据证明了整体最小二乘法的优势。但是只给出了拟合数据本没有一个正确的参考值所以无法正真体现出整体最小二乘法优势。通过预设相偏再通过同步模型测量基于整体最小二乘法的同步精度, 同时与最小二乘同步做比较就可验证整体最小二乘法的优势。

本文用 63bitPN 序列, 通过 BPSK 对 PN 码进行调制, 再在调制信号加上预定的相偏 $\frac{T_c}{10}$ 。通过给信号加入高斯白噪声验证两种方法的抗噪性能。

3.1 测量精度的比较

为了比较两种方法在噪声条件下的测量精度, 预设相位偏移量是 $\frac{T_c}{10}$, 采样率 $f_s = 11$, 比较信噪比在 (-20 dB, 0 dB) 两种方法的测量精度, 求得两种方法的测量误差, 如图 5 所示。

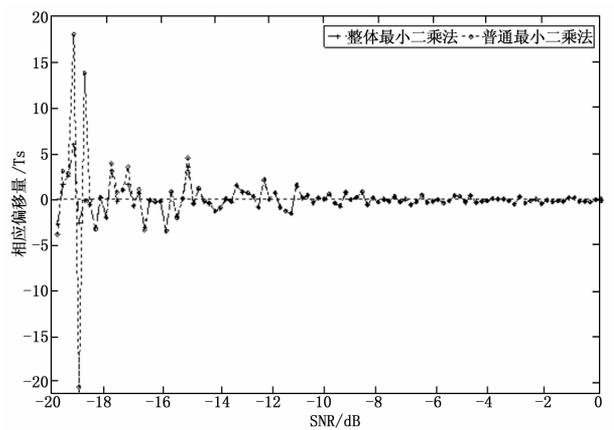


图 5 测量误差受噪声影响的比较

从图 5 可以看出, 整体最小二乘法抗噪声干扰的能力明显比普通最小二乘法好, 特别是信噪比很低的情况下, 信噪比到 -20dB 整体最小二乘法测量误差在 $\frac{T_c}{2}$ 以内 (T_s 是采样周期且满足 $T_c = f_s \times T_s$), 普通最小二乘法的误差达到 $1.5T_c$, 从可以看出在信噪比小于 -18 dB 时, TLS 法的测量精度是 LS 法的 3 倍以上。

3.2 测量稳定性的比较

为了分析两种方法在不同信噪比下的测量稳定性, 在信噪比范围是 (-20 dB, 0 dB) 高斯环境下进行仿真实验, 每个选取的信噪比条件下进行 200 次仿真并获取其统计数据。表 1 给出了两种算法在不同的信噪比条件下测量值的均方误差。

表 1 显示, 随着信噪比降低, TLS 稳定性比 LS 更加好, 在信噪比降到 -20 dB 时 TLS 的误差精度是 LS 的 5 倍。整体而言, 在不同的信噪比条件下整体最小二乘法都显示出了优于普通最小二乘同步法的稳定性。仿真实验显示: 相比于普通最小二乘同步方法, 整体最小二乘法无论是在高信噪比还是低信噪比其测量稳定性都明显优于前者, 同样该方法的测量精度受噪声干扰较小, 用整体最小二法做精同步在低信噪比环境下相比普通最小二乘法测量精度也有很大的提高。所以在实际环境下, 整体最小二乘法比普通最小二乘法更适用于精同步。

4 结束语

整体最小二乘法是近年来出现的同时顾及系数矩阵和观测值误差的数据处理方法, 本文把整体最小二乘法应用到精同步的 PN 码测相中来。建立基于 BPSK 调制方式的同步模型, 通过分析, 证明了 TLS 在精同步中的可行性。通过仿真实验, 验证了整体最小二乘法相对于普通最小二乘法的优势。仿真结果可以看出, 其测量结果受噪声影响相对较小。特别是噪声环境恶劣, 信噪比 (SNR) 较低的环境里, 信噪比在 -20dB 条件下, 其测量精度可以达到 LS 方法的 3 倍以上, 而均方差更

表 1 两种方法的测量均方差比较/ T_c

算法	信噪比/dB										
	-20	-18	-16	-14	-12	-10	-8	-6	-4	-2	0
普通最小二乘法	0.867 2	0.487 2	0.444 2	0.167 7	0.125 8	0.076 4	0.073 2	0.045 6	0.038 7	0.028 6	0.023 0
整体最小二乘法	0.166 1	0.126 9	0.103 2	0.097 7	0.072 9	0.067 7	0.044 7	0.037 0	0.038 1	0.028 3	0.022 7

是能达到 LS 同步方法的 5 倍以上。得出结论，在低信噪比环境下整体最小二乘法比普通最小二乘法更合适于做同步误差估计。

参考文献:

[1] Vanneer D J R, Coenen A J R M. New fast GPS code-acquisition technique using FFT [J]. Electronics Letters, 1991, 27: 158-160.

[2] Kaplan E D. Understanding GPS: Principles and Application [M]. Boston: Artech House Publishers, 1996.

[3] 龚国辉, 李思昆. 提高 DSSS 信号 PN 码相位测量精度的三点二次插值法 [J]. 通信学报, 2007, 28 (2): 130-133.

[4] 胡修林, 曾臻, 张俊, 等. 直扩系统伪码精确同步及 FPGA

实现 [J]. 华中科技大学学报 (自然科学版), 2005, 33 (6): 44-46.

[5] Golub G H, Van Loan C F. An analysis of the Total Least Squares problem [J]. SIAM J Numer Anal, 1980, 17 (6): 883-893.

[6] Van Huffe S I, Vandewalle J. The Total Least Squares Problem Computational Aspects and Analysis [M]. SIAM, Philadelphia, 1991.

[7] 丁克良, 沈云中, 欧吉坤. 整体最小二乘法直线拟合 [J]. 辽宁工程技术大学学报 (自然科学版), 2010, 29 (1): 44-47.

[8] 李红伟, 魏少春, 陈安平, 等. 总体最小二乘法在直线拟合中的应用 [J]. 地矿测绘, 2010, 26 (2): 4-5.

[9] 王徐华, 柏鹏, 李寰宇, 等. 基于多速率内插和最小二乘法的精确同步方法 [J]. 计算机应用研究, 2012, 29 (10): 3922-3925.



(上接第 2290 页)

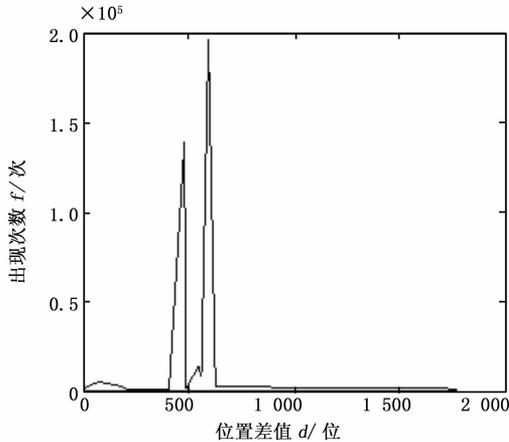


图 7 报文长度推断

分析图 7 可知，位置差值 592 与位置差值 480 出现次数远多于其他差值，且前者多于后者。分析 Wireshark 对两种报文的解析结果可知，报文长度推断正确，进一步验证了算法的有效性；其中位置差值 592 出现次数更多的原因是 ICMP 报文的特征序列位长大于 ARP 报文的特征序列，而提取的频繁序列位长较短，使得同一段特征序列重复计数。

4 小结

本文以数据挖掘中频繁集和关联规则相关理论为基础，针对长度可变的报文集合设计了基于自适应权值的数据报指纹特征提取方案。真实数据测试表明，该方案对于数据报的指纹特征提取是有效的，自适应权值的引入使得筛选率和命中率都有提高且更为稳定，有效控制了冗余数据段对于提取结果的干扰。该方案在实际应用中还应该根据实际情况合理设置参数以有效可靠地提取指纹特征^[8-14]。

参考文献:

[1] Kim M S, Won Y J, Hong J W K. Application-level traffic monitoring and an analysis on IP networks [J]. ETRI Journal, 2005, 27 (1): 22-42.

[2] Callado A, Kamienski C, Szabo G, et al. A survey on internet traffic identification [J]. Communications Surveys & Tutorials, IEEE, 2009, 11 (3): 37-52.

[3] Trifilo A, Burschka S, Biersack E. Traffic to protocol reverse engineering [A]. Computational Intelligence for Security and Defense Applications, CISDA 2009, IEEE Symposium on. IEEE [C], 2009: 1-8.

[4] 张一嘉. 局域网链路层数据帧识别算法的设计与实现 [J]. 通信对抗, 2007, (4): 41-44.

[5] 白彧, 杨晓静, 张玉. 基于高阶统计处理技术的 m-序列帧同步码识别 [J]. 电子与信息学报, 2012, 34 (1): 33-37.

[6] 金凌. 面向比特流的未知帧头识别技术研究 [D]. 上海: 上海交通大学, 2011.

[7] 马占欣, 王新社, 黄维通, 等. 对最小置信度门限的置疑 [J]. 计算机科学, 2007, 34 (6): 216-218.

[8] 王祥斌. 数据挖掘技术在入侵检测系统中的应用研究 [J]. 计算机测量与控制, 2012, 20 (2): 321-323.

[9] 张文博, 姬红兵, 王磊. 一种自适应权值的多特征融合分类方法 [J]. 系统工程与电子技术, 2013, 35 (6): 1133-1137.

[10] 裴颂文, 吴百锋. 动态自适应特征权重的多类文本分类算法研究 [J]. 计算机应用研究, 2011, 28 (11): 4092-4096.

[11] Jiang H C, Xu J Z, Shi W S, et al. Stellar spectra association rule mining method based on the weighted frequent pattern tree [J]. Research in Astron. Astrophys, 2013, 13 (3): 3434-3442.

[12] 张昆明, 甘文丽, 李元臣. Master-Worker 模式的并行关联规则挖掘算法 [J]. 计算机测量与控制, 2013, 21 (4): 1008-1010.

[13] Tian X G, Duan M Y, Sun C L, et al. Detecting network intrusions by data mining and variable-length sequence pattern matching [J]. Journal of Systems Engineering and Electronics, 2009, 20 (2): 405-411.

[14] 刘兴彬, 杨建华, 谢高岗, 等. 基于 Apriori 算法的流量识别特征自动提取方法 [J]. 通信学报, 2008, 29 (12): 51-59.