

文章编号 1004-924X(2010)10-2280-08

面向 IP 模式测控系统的 PLSR-SBR 双层压缩

洪晓斌, 刘桂雄

(华南理工大学 机械与汽车工程学院, 广东 广州 510640)

摘要:针对以太网测控网络存在数据冲突导致系统实时性、可靠性降低问题,提出了基于偏最小二乘回归(PLSR)SBR 的双层压缩方法。第一层建立主参量与所有辅助参量的确定模型,利用压缩有效性指标确定主成分,完成主参量的信息压缩。第二层基于改进的 SBR,通过选取辅助参量中的基础序列,建立基础信号;在满足拟合误差条件下,逐步将每一个辅助参量序列映射到基础信号上,完成对辅助参量的数据压缩。该方法重点解决辅助参量和主参量中的解释潜变量和反映潜变量相关程度最大、基础信号由最少基础序列组成、辅助参量实现最小变长分解个数及基础信号独立更新原则等关键问题。最后将该方法应用于 IP 模式乙醇浓度测控系统。实验结果表明,在 IP 模式测控系统同时具有主参量和辅助参量,且不同参量间存在相关性时,该方法可在允许拟合相对误差为 5% 的情况下,使压缩率达到 68% 以上,从而有效地降低以太网测控网络数据冲突程度。

关键词:IP 模式测控系统;数据压缩;偏最小二乘回归;改进的 SBR

中图分类号:TN911.7;TP273 **文献标识码:**A **doi:**10.3788/OPE.20101810.2280

Double-layer compression method based on PLSR-SBR for IP mode measurement and control system

HONG Xiao-bin, LIU Gui-xiong

(School of Mechanical and Automotive Engineering,
South China University of Technology, Guangzhou 510641, China)

Abstract: To address the existing problems in Ethernet measurement and control networks, such as real-time and reliability caused by the collision of measurement and control data, a double-layer compression method based on Partial Least Square Regression-self Based Regression(PLSR-SBR) was induced to the IP Mode Measurement and Control System(IMMCS). In the first layer, the model of main parameters and all auxiliary parameters was built up based on PLSR. The principal score of main parameter was determined by the efficient compression index to finish the information compression of main parameter. In the second layer, the auxiliary parameters were compressed based on a modified SBR to set up a basic signal. To ensure the fitting error to be smaller than the setting value, every auxiliary parameter sequence in the auxiliary parameters or the decomposed sequences were mapped in turn to a definitive basic signal to complete the data compression of the auxiliary parameters. Innovative technologies for some key problems were addressed in the method, including the confirmation of maximum correlation degree of interpreting latent variables and reflecting latent variables,

收稿日期:2009-12-10;**修订日期:**2010-03-10.

基金项目:教育部新世纪优秀人才支持计划资助项目(No. NCET-08-0211);广东省自然科学基金资助项目(No. 9151007010000001);广东省科技计划资助项目(No. 2009B010900045);中央高校基本科研业务费专项资金资助项目(No. 2009ZM0042)

basic signal with least basic sequences, the minimum number of decomposed sequences from auxiliary parameters and the independent update rule of basic signal. Finally, experiments using IMMCS for alcohol concentration were performed. The results show that under the conditions that IMMCS possesses main parameters and auxiliary parameters, and simultaneously data variety of parameters is wide, the compression ratio based on the method reaches 68%, which is higher than that based on the Swinging Door. The data collision of Ethernet measurements and control networks are decreased effectively.

Key words: IP Mode Measurement & Control System (IMMCS); data compression; Partial Least Squar Regression(PLSR); modified Self Based Regression(SBR)

1 引言

随着现代工业生产的发展, IP 模式测控系统(IP Mode Measurement and Control System, IMMCS)得到了进一步发展,促使系统内的测控节点大幅度增加,测控数据量急速增加,而数据压缩技术的引入可以大大降低以太网测控网络的数据冲突问题^[1]。目前工业数据压缩算法主要有两种:一种以 ENDA 数据库为代表,采用 Huffman 压缩算法;另一种以 PI 实时数据库为代表,采用旋转门压缩算法,其中旋转门算法比 Huffman 算法具有更高的压缩率^[2],但应用于 IP 模式时实时数据库系统存在两点不足:(1)“过渡数据集”容量大小存在随机性。如果“过渡数据集”太大,会造成内存的极大浪费;随着“过渡数据集”的增大,判断过渡数据集内数据是否不在“平行四边形”内的时间将以几何级数增长;(2)该压缩算法属于单节点压缩技术,没考虑到分布式系统中多节点数据

中存在的大量冗余信息。Deligiannakisd. A 等提出的 SBR (Self-Based Regression)方法是一种基于数据线性相关性的数据压缩方法,目前主要应用于无线传感器网络中^[3-4]。即使某最小二乘回归的总误差符合要求,该算法也只能记录当前回归拟合过程中最小误差,导致结果有时不符合用户的误差要求。面向 IMMCS 的实时压缩方法必须与 IMMCS 研究相结合,因此本文提出了一种新型的面向 IMMCS 的双层压缩方法。

2 IMMCS 以及双层压缩一些定义

文献[5]已经成功地实现了一种 IMMCS(见图 1)。从图 1 可以看出,整个网络测控系统分为三层次,即第一层为现场测控层,具体包括分布式智能传感器、IP 智能测控装置、IP 模糊控制装置和现场受控设备组;第二层为企业级监控层,具体包括以太网交换机或集线器、现场工作站和企业服务器等;第三层为基于 Internet 异地远程网络

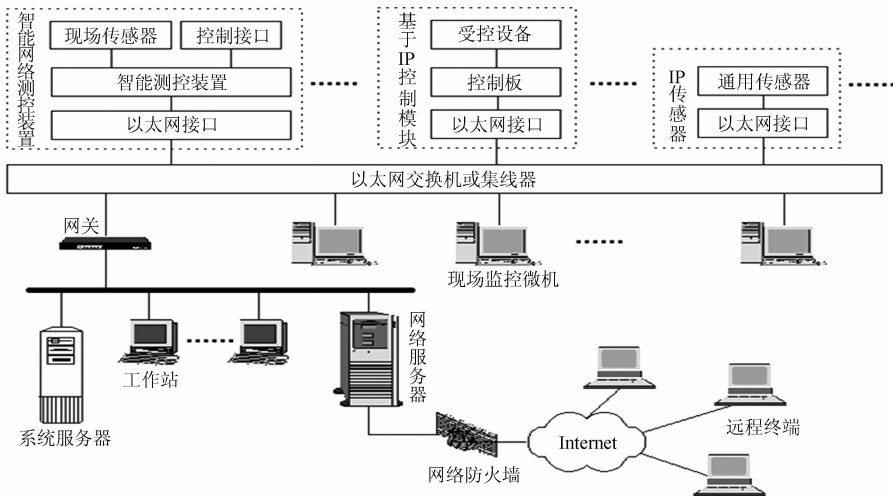


图 1 系统总体结构

Fig. 1 General structure of system

层主要包括远程工作站和网络服务器等。系统采用一网到底模式,通过现场测控网络、内部网和 Internet 把分布于各局部现场、独立完成特定功能的测控单元互连起来,以达到资源共享、协同工作的目的。

IMMCS 的实时测控数据具有如下几个主要特点:(1)采集速度快;(2)有效性受时间限制;(3)分布式数据存在一定的冗余度;(4)传感参量间存在多重相关性^[6]。一方面,由于偏最小二乘回归(Partial Least Square Regression, PLSR)理论能在样本容量小于变量个数的情况下,容易消除变量间多重相关性^[7],因此可借助 PLSR 消除多传感参量间的耦合关系。另一方面,IP 模式实时数据库除了规定压缩率和数据序列总体误差外,不用考虑通信带宽及最大的节点数据存储空间等约束^[8],因此可进一步改进 SBR,消除冗余度进行数据压缩,从而形成基于 PLSR-SBR 的双层压缩方法。这里首先对本方法涉及一些概念进行定义。

定义 1 假定 IMMCS 中存在 m 个传感参量,把其中受到若干个参量影响的参量称为主参量。相应地,影响主参量的其它参量称为辅助参量。其中有 p 个辅助参量 $\{x_1, x_2, \dots, x_p\}$,在观测 n 个样本点后,构成辅助参量数据表 $X = [x_1, x_2, \dots, x_p]_{n \times p}$;其中有 q 个主参量 $\{y_1, y_2, \dots, y_q\}$,同样构成主参量数据表 $Y = [y_1, y_2, \dots, y_q]_{n \times q}$, $p+q=m$ 。

定义 2 基于 PLSR 的第 h 个成分压缩有效性为

$$Q_h^2 = 1 - \frac{PRESS_h}{SS_{(h-1)}}, \quad (1)$$

其中: $PRESS_h = \sum_{i=1}^n (y_i - \hat{y}_{h(-i)})^2$, $SS_h = \sum_{i=1}^n (y_i - \hat{y}_{hi})^2$ 。 $\hat{y}_{h(-i)}$ 为除去某样本点 i 的剩余所有样本点集合后,采用 h 个成分拟合一个回归方程,再把被排除的样本点 i 代入拟合的方程所求出样本点 i 上的拟合值。 \hat{y}_{hi} 为采用所有样本点拟合含 h 个成分的回归方程下,第 i 个样本点的拟合值。

定义 3 基于 $\bar{x}_i \{x_{i1}, \dots, x_{im}\}$ 对 $\bar{x}_j \{x_{j1}, \dots, x_{jm}\}$ 进行最小二乘线性拟合,拟合结果为 $\hat{\bar{x}}_j \{x_{j1}, \dots, x_{jm}\}$,基于欧几里德距离,其拟合误差可表示为 $d(\bar{x}_j, \hat{\bar{x}}_j) = \left(\sum_{k=1}^m (\bar{x}_{jk} - \hat{x}_{jk})^2 \right)^{1/2}$ 。

定义 4 基于 $\bar{x}_i \{x_{i1}, \dots, x_{ik}\}$ 对 $\bar{x}_j \{x_{j1}, \dots, x_{jm}\}$ 进行最小二乘线性拟合,其中 $k < n$,则称 $\bar{x}_i \{x_{i1}, \dots, x_{ik}\}$ 的使用量为 k 。

定义 5 基于 $\bar{x}_i \{x_{i1}, \dots, x_{ik}\}$ 对 $\bar{x}_j \{x_{j1}, \dots, x_{jm}\}$ 进行最小二乘线性拟合,其中 $k < n$,共存在 $n-k+1$ 次完全最小二乘线性拟合计算, $n-k+1$ 次拟合误差之和称为 $\bar{x}_i \{x_{i1}, \dots, x_{ik}\}$ 对 $\bar{x}_j \{x_{j1}, \dots, x_{jm}\}$ 的累计误差 δ_{add} 。

3 基于 PLSR 的主参量压缩机制

3.1 数据标准化处理

分别对主参量和辅助参量进行标准化。

$$F_0 = (F_{01}, F_{01}, \dots, F_{0q})_{n \times q}, F_{0j} = y_j^* = \frac{y_j - E(y_j)}{S_{y_j}}, \quad (j=1, 2, \dots, q) \quad (2)$$

$$E_0 = (E_{01}, E_{01}, \dots, E_{0p})_{n \times p}, E_{0i} = x_i^* = \frac{x_i - E(x_i)}{S_{x_i}}, \quad (i=1, \dots, p) \quad (3)$$

式中: F_0, E_0 分别为 Y, X 的标准化矩阵; $E(y_j), E(x_i)$ 分别为 Y, X 的均值; S_{y_j}, S_{x_i} 分别为 Y, X 的均方差; n 为样本容量。数据标准化处理后,数据的重心与原点重合。

3.2 h 成分提取

已知 F_0, E_0 。从 E_0 中提取第一个成分 $t_1, t_1 = E_0 W_1, W_1$ 是 E_0 的第一个轴,且 $\|W_1\| = 1$; t_1 是标准化变量 $x_1^*, x_2^*, \dots, x_p^*$ 的线性组合,为原信息的重新整合。从 F_0 中提取第一个成分 $u_1, u_1 = F_0 C_1, C_1$ 是 F_0 的第一个轴, $\|C_1\| = 1$ 。

要求 t_1, u_1 能分别很好地代表 X 与 Y 中的数据变异信息,同时满足 t_1 对 u_1 有最大的解释能力。在 $\|W_1\| = 1$ 和 $\|C_1\| = 1$ 的约束条件下,去求 $(W_1^T E_0^T F_0 C_1)$ 的最大值,采用拉格朗日算法,经推导得出^[9]:

$$E_0^T F_0 F_0^T E_0 W_1 = \theta_1^2 W_1, \quad (4)$$

$$F_0^T E_0 E_0^T F_0 C_1 = \theta_1^2 C_1, \quad (5)$$

其中: W_1 为 $E_0^T F_0 F_0^T E_0$ 的特征向量, θ_1^2 是对应的特征值。要 θ_1 取最大值,则 W_1 为 $E_0^T F_0 F_0^T E_0$ 矩阵最大特征值的单位特征向量, C_1 是对应于矩阵 $F_0^T E_0 E_0^T F_0$ 最大特征值 θ_1^2 的单位特征向量。

由于在 IMMCS 中 F_0 每次只有一个变量,因此 C_1 是一个常数,可得

$$\theta_1^2 = \|E_0^T F_0\|^2, \quad (6)$$

$$W_1 = \frac{E_0^T u_1}{\theta_1} = \frac{E_0^T F_0}{\|E_0^T F\|}, \quad (7)$$

因为 E_0, F_0 均是单位向量,所以有

$$W_1 = \frac{1}{\sqrt{\sum_{j=1}^p r^2(x_j, y)}} \begin{bmatrix} r(x_1, y) \\ \dots \\ r(x_p, y) \end{bmatrix}, \quad (8)$$

其中: $r(x_j, y)$ 为辅助参量 x_j 和主参量 y 的相关系数。求得轴 W_1 后,即可得分

$$t_1 = E_0 W_1 = \frac{1}{\sqrt{\sum_{j=1}^p r^2(x_j, y)}} [r(x_1, y) E_{01} + \dots + r(x_p, y) E_{0p}] \quad (9)$$

分别求 F_0, E_0 对 t_1 的回归方程:

$$E_0 = t_1 p_1^T + E_1, \quad (10)$$

$$F_0 = t_1 r_1 + F_1, \quad (11)$$

式中: $p_1 = \frac{E_0^T t_1}{\|t_1\|^2}, r_1 = \frac{F_0^T t_1}{\|t_1\|^2}$ 为相应的回归系数向量(标量);记残差矩阵

$$E_1 = E_0 - t_1 p_1^T = [E_{11}, \dots, E_{1p}], \quad (12)$$

$$F_1 = F_0 - t_1 r_1. \quad (13)$$

以 E_1 取代 E_0, F_1 取代 F_0 , 用上面的方法求第二个轴 W_2 和第二个成分 t_2 , 有

$$W_2 = \frac{E_1^T F_1}{\|E_1^T F_1\|} = \frac{1}{\sqrt{\sum_{j=1}^p Cov(E_{1j}, F_1)}} \begin{bmatrix} Cov(E_{11}, F_1) \\ \dots \\ Cov(E_{1p}, F_1) \end{bmatrix}, \quad (14)$$

$$t_2 = E_1 W_2. \quad (15)$$

W_2 为矩阵 $E_1^T F_1 F_1^T E_1$ 最大特征值 θ_2^2 的单位特征向量,

施行 E_1, F_1 对 t_2 的回归,有

$$E_1 = t_2 p_2^T + E_2, \quad (16)$$

$$F_1 = t_2 r_2 + F_2, \quad (17)$$

式中: $p_2 = \frac{E_1^T t_2}{\|t_2\|^2}, r_2 = \frac{F_1^T t_2}{\|t_2\|^2}$ 。

同理,可推求第 h 成分 t_h 。最终 h 交叉有效性可由定义 2 确定。

3.3 主参量压缩方程

根据上述 h 成分计算完成,可得到 E_0, F_0 关于 t_1, t_2, \dots, t_h 的最小二乘回归方程:

$$E_0 = t_1 p_1^T + t_2 p_2^T + \dots + t_h p_h^T, \quad (18)$$

$$F_0 = t_1 r_1 + t_2 r_2 + \dots + t_h r_h. \quad (19)$$

由于 t_1, t_2, \dots, t_h 均是 $E_{01}, E_{02}, \dots, E_{0p}$ 的线性组合,依据 PLSR 性质,知

$$t_h = E_{h-1} W_h = E_0 \prod_{j=1}^{h-1} (I - W_j p_j^T) W_h = E_0 W_h^*, \quad (20)$$

$$\text{记 } W_h^* = \prod_{j=1}^{h-1} (I - W_j p_j^T) W_h.$$

所以,

$$F_0 = E_0 (r_1 W_1^* + \dots + r_p W_p^*), \quad (21)$$

$$\text{若记 } x_j^* = E_{0j}, y^* = F_0, a_j = \sum_{h=1}^m r_h W_{hj}^*,$$

可得主参量压缩方程为:

$$\hat{y}_j^* = \alpha_{1j} x_1^* + \alpha_{2j} x_2^* + \dots + \alpha_{pj} x_p^*. \quad (22)$$

依上述的推导可以看出,第一层压缩方法在辅助参量 $\{x_1, x_2, \dots, x_p\}$ 中逐次提取解释潜变量 $t_1, t_2, \dots, t_h (h < p)$, 这相当于对 x_1, x_2, \dots, x_p 中的信息进行重新组合及提取,从而得到对 Y 的解释能力最强,同时又最能概括自变量集合 X 中信息的综合变量,很好地进行了信息压缩。如果存在多个主参量,那么继续运行第一层压缩过程。经上述分析可知,在基于 PLSR-SBR 的双层压缩方法中,第一层压缩的结果 I 可将每个主参量数列表示为如下 $(\text{seat}, a_{1j}, a_{2j}, \dots, a_{pj})$ 多元组的形式存储,即有 $(1+p)$ 个元素。其中 seat 为多维参量中主参量的序号; $a_{1j}, a_{2j}, \dots, a_{pj}$ 为回归模型参数, $1, 2, \dots, p$ 为多维参量中辅助参量的序号。

4 基于改进 SBR 的辅助参量压缩机制

基于 PLSR-SBR 的双层压缩方法经过上述第一层压缩后,剩下了相关性较强的其他多个自变量(如温度、湿度、压力等),可紧接进行基于改进 SBR 方法的第二层数据相关压缩计算。首先定义 5 个重要变量,包括:(1) row: 序列排号;(2) shift: 数据序列在基础信号中的偏移量;(3) a: 一元线性回归方程的参数;(4) b: 一元线性回归方程的参数;(5) error: 定义为一元回归的近似误差。

4.1 基础信号建立

基础信号的建立是第二层压缩中最重要的一环,它是辅助参量序列进行拟合压缩的基准序列,决定了压缩率和压缩误差大小。由于基础信号是由若干个基础序列组成的,并体现所有的数据特征,因此选择最少基础序列是研究的关键。

将 p 个辅助参量序列(每个序列具有个 n 数

据点)串接组成一个大序列,并重新分成 k 个包含 w 个数据点的基础序列,其中 $W = \sqrt{p \times n}$,每 w 个数据点组成的序列称为一个基础序列^[3]。然后依次以其中一基础序列 i 分别使用最小二乘法去拟合每个基础序列 j (包括本身),可得出每次拟合的误差值 $Error_j$,最终可得出一个当前序列 i 的 benefit 值。

$$\text{benefit} = \sum_{j=1}^k [\text{linearErr}(\text{Cand}_j) - \text{Error}_j]. \quad (23)$$

其中: $\text{linearErr}(\text{Cand}_j)$ 为被拟合序列的自身最小二乘拟合误差。

依 benefit 值的大小按降序排列所有基础序列,并存到 List 变量中。选择前 S 个基础序列组成基础信号, S 值采用折半查找法确定(见图 2)^[10]。其确定要求是用前 S 个基础序列被每个辅助参量数据序列拟合的最小误差之和 $Error[S]$, 小于前 $S-1$ 或 $S+1$ 个基础序列所拟

合得到的最小误差之和 $Error[S-1]$ 或 $Error[S+1]$ 。图中 Start, End, Middle 值代表被拟合序列由 List 中第一个、全部及其一半基础序列的个数。

$Error[S]$ 函数的计算方法实现如下。以所选择的 S 个基础序列组成基准序列 X , 每个辅助参量序列 I 从基础序列 X 的最左端开始最小二乘法拟合, 拟合一次可得到一个 5 元素数组 (row, shift, a_1 , b_1 , $error_1$)。然后辅助参量序列 X 向基础序列右边移一位, 即 $\text{shift} = \text{shift} + 1$, 又可得到另外一个 5 元素数组 (row, shift, a_2 , b_2 , $error_2$)。每个辅助参量序列依此移位, 直到 $\text{shift} = X \text{ length} - I \text{ length}$ 为止, 得到总共 $X \text{ length} - I \text{ length} + 1$ 个 5 元素数组, 比较每个 5 元素数组中的 error 值, error 值最小(表示为 $Error_{\min}$) 的那个 5 元素数组即代表该辅助参量数据序列。总而言之, $Error[S]$ 结果即为用前 S 个基础序列作为被拟合序列, 所有辅助参量序列各自所对应的 $Error_{\min}$ 之和。

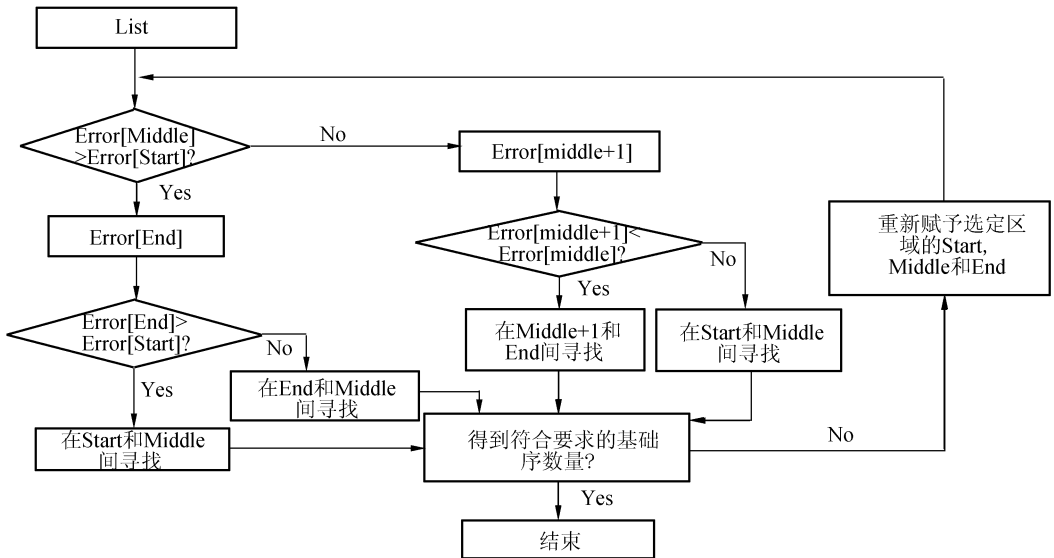


图 2 折半查找法

Fig. 2 Binary search method

4.2 辅助参量数据序列压缩计算

确定了基础信号, 分别将每个辅助参量序列建立到基础信号的映射。当辅助参量序列所对应 $Error_{\min}$ 小于系统设定的 δ_0 时, 即可用辅助参量对应的 $Error_{\min}$ 5 元素向量中的 (row, shift, a , b) 4 个元素来最终表示。

当辅助参量序列所对应 $Error_{\min}$ 大于系统设

定的 δ_0 时, 需对该辅助参量序列进行变长分解计算。将该辅助参量序列平分为 2 个子序列, 可在该序列的 row 加上 1 或 2 来区分分解后的左子序列和右子序列(如排号 row 为 1, 即变为 11 和 12, row 为 11 代表分解后左序列的新排号 row, row 为 12 代表分解后右序列的新号 row)。如果分解后的子序列对应的 $Error_{\min}$ 仍然大于系统设定的

δ_0 时,按同样的方法,变成更多位,以便在解压缩时可以识别。然后再分别建立到基础信号的映射。为了最大程度地减少最终子序列的数目,提高压缩率,当分解后序列所对应 $Error_{min}$ 小于系统设定的 δ_0 时即可结束,分解后序列用对应的 $Error_{min}$ 5 元素向量中另外(row, shift, a, b) 4 个元素来最终表示。一旦出现分解后子序列的数据点小于或等于 4,而其对应 $Error_{min}$ 还是大于系统设定的 δ_0 时,则这样的分解结果是不能满足要求的,因此需要寻找减少它的 $Error_{min}$ 的方法。本文在采取辅助参量序列进行变长分解计算方法不变的基础上,对基础信号进行更新。采用最低使用量和最大累计误差综合评价原则进行补偿,即在基础信号的各数据序列中,应用定义 4 中的使用量 k ,可选出被所有辅助参量序列最小二乘拟合使用次数最少的基础序列。

当出现最少使用次数相同的多个基础序列时,应用定义 5 中累计误差计算方法,可得出所有辅助参量在基础信号中各个基础序列的累计误差,通过比较它们各自的累计误差,可以找到最大累积误差的一个基础序列。再由之前 List 中未选出组成基础信号中最大 benefit 值的基础序列在位置上完全替代被删除的基础序列,形成对被分解的辅助参量独用的基础信号,其 row 排号设为 $s+1$ 。从而完成相关性较强的多个辅助参量的压缩计算。

5 实验与结果

液态乙醇浓度检测中存在明显的多传感耦合信息问题。在进料浓度和压力恒定情况下,影响乙醇浓度的因素主要有塔底温度、塔顶温度、湿度等 3 个影响因素^[11],因此本实验将本文提出的双层压缩方法应用于乙醇浓度测控环境中,搭建的现场精馏实验系统组成如图 3 所示。

所采用的 CS-00A 型蒸馏装置的精馏塔是塔身内径为 50 mm 的玻璃钢填料塔,保温良好;塔釜装有 6 根功率为 2 kW 的电加热棒;塔顶、塔身、塔釜有取样口,可分别取液相物料进行分析;冷凝器采用列管式换热器,同时塔体上有热敏电阻接口、湿度传感器接口以及乙醇浓度传感器接口;试剂组成为乙醇—水体系,水溶剂选用实验用蒸馏水;用 NTC 型热敏电阻测温, WYG1 型湿度

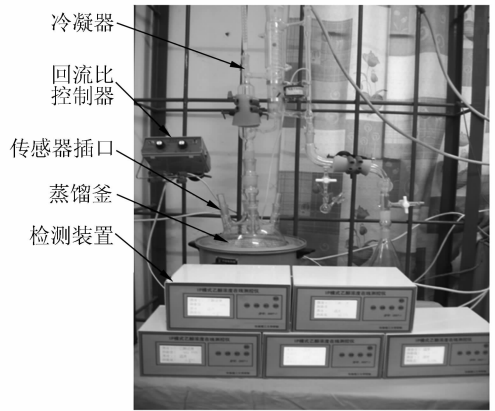


图 3 实验平台

Fig. 3 Experimental platform

传感器测湿度, TGS2620 型乙醇气敏传感器测乙醇浓度, 5 台 SCUT-1 型 IP 模式测控装置可计算不同位置传感信息并进行结果显示。基于 PLSR-SBR 的双层实时压缩方法软件模块,启动 IP 模式实时数据库。本实验系统共有 5 台 IP 模式智能测控装置,每台装置具有 4 个通道,分别进行 4 个参数采集(塔底温度、塔顶温度、湿度和乙醇浓度),共有 20 组序列。根据不同的数据量来验证该方法的运行情况,分为两种情况进行对比,来分析数据量与压缩率之间存在的可能关系:

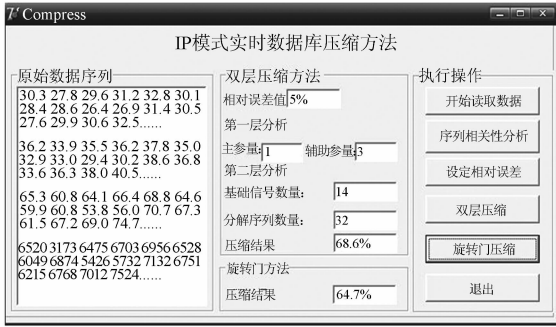
(1) 每个参量采集 54 个数据,共 1 080 个数据。先单击“开始读取数据”,即立刻输入原始数据序列,接着单击序列相关性分析,可得表 1 所示结果,可见系统 4 个参量相关性强。最后单击压缩结果和旋转门压缩,结果如图 4(a)所示。系统包括 1 个主参量和 3 个辅助参量,在设定允许相对误差为 5%时,基础信号由 14 个基础序列组成,辅助参量原始序列最终分解成为 32 个变长序列,压缩结果为 68.6%,旋转门方法压缩结果为 64.7%。

表 1 序列相关性分析

Tab. 1 Analysis of sequence correlation

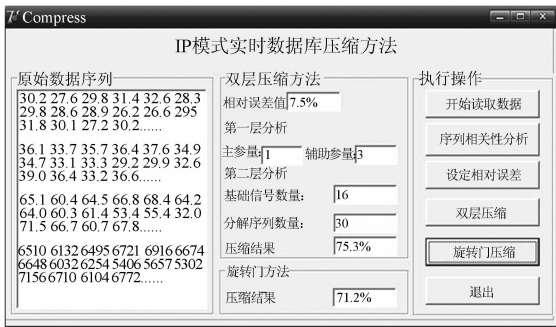
x_1	x_2	x_3	y_1	
x_1	1.000 0	-0.842 2	-0.694 3	-0.766 4
x_2	0	1.000 0	0.605 6	0.819 7
x_3	0	0	1.000 0	0.740 7
y_1	0	0	0	1.000 0

(2) 每个参量采集 115 个数据,共 2 300 个数据。与第一种情况的操作过程相同,结果如图 4



(a) 第一种情况

(a) Running result for the first condition



(b) 第二种情况

(b) Running result for the second condition

图 4 实验结果

Fig. 4 Results for experiments

(b)所示。在设定允许相对误差 7.5% 时,基础信号由 16 个基础序列组成,辅助参量原始序列最终分解成为 30 个变长序列,压缩结果为 75.3%,旋转门方法压缩结果为 71.2%。

6 结 论

本文提出的面向 IMMCS 压缩方法必须与 IMMCS 研究相结合。基于 PLSR-SBR 的双层压缩方法,通过设定符合实际要求的相对误差值,能够满足 IMMCS 测控点不断扩展的发展需求。系统中参量之间相关性越强,其压缩效果越好,当允许拟合相对误差为 5% 时,该方法压缩率可达到 68% 以上;在有效时间的限制下,分布式多参量数据量越大,其压缩率越高。在辅助参量序列进行变长分解计算方法不变的基础上,为满足数据压缩率,提出最低使用量和最大累计误差综合评价原则对基础信号进行更新,最大程度地保证压缩数据质量。如果在测控数据量非常大的情况下,可考虑针对具体同一参量不同时间段进行压缩,以进一步提高压缩率。

参考文献:

- [1] PETER N. Communication in industrial automation—What is going on? [J]. *Control Engineering Practice*, 2007, 15(11):1332-1347.
- [2] MAKAROV Y V, LOUTAN C, MA J, et al.. Operational impacts of wind generation on california power systems[J]. *IEEE Transactions on Power Systems*, 2009, 24(2):1039-1050.
- [3] DELIGIANNAKIS A, KOTIDIS Y, ROUSSOPOULOS N. Compressing historical information in sensor networks[C]. *Proceeding of International Conference on Management of Data*, 2004: 527-538.
- [4] 任倩倩. 传感器网络中感知数据压缩技术研究[D]. 哈尔滨:黑龙江大学,2006.
REN Q Q. *Study on compression method for sensing data in sensor networks* [D]. Harbin: Hei-

longjiang University, 2006. (in Chinese)

- [5] 洪晓斌,刘桂雄,周钦河. 网络化测控系统选播运行模式策略[J]. *光学 精密工程*, 2008, 16(5):957-964.
HONG X B, LIU G X, ZHOU Q H. Strategy of anycast mode in networked measuring & controlling platform[J]. *Opt. Precision Eng.*, 2008, 16(5): 957-964. (in Chinese)
- [6] 张雷,黄廉卿,赵唯佳. 一种超光谱图像分层压缩方法[J]. *光学 精密工程*, 2006, 14(3):478-484.
ZHANG L, HUANG L Q, ZHAO W J. A multi-layered decomposition of hyper-spectral image compression[J]. *Opt. Precision Eng.*, 2006, 14(3): 478-484. (in Chinese)
- [7] 刘桂雄,林绪虹. 鱼类超微弱发光的偏最小二乘回归分析与建模[J]. *华南理工大学学报(自然科学版)*, 2006, 34(11):29-32.
LIU G X, LIN X H. Partial least-square regressive analysis and modeling of ultraweak photon emission

of fish[J]. *Journal of South China University of Technology: Natural Science Edition*, 2006, 34(11):29-32. (in Chinese)

- [8] 邓家先. 基于重要系数提升的遥感图像压缩[J]. *光学精密工程*, 2006, 14(5):910-916.

DENG J X. A remote-sensing image coding algorithm based on significant coefficients lifting[J]. *Opt. Precision Eng.*, 2006, 14(5): 910-916. (in Chinese)

- [9] 王惠文. 偏最小二乘回归方法及其应用[M]. 北京: 国防工业出版社, 2000:150-171.

WANG H W. *Partial Least-Squares-Regression*

Method and Applications[M]. Beijing: Defense Industry Publishing Company, 2000:150-171. (in Chinese)

- [10] DELIGIANNAKIS A, KOTIDIS Y, ROUSSOPOULOS N. Dissemination of compressed historical information in sensor networks [J]. *The VLDB Journal*, 2007, 16:439-461.

- [11] HONG X B, LIU G X, DU X P. Implementation method for region vector of IP mode multi-sensing cooperated measurement based on synergetics[J]. *IEEE*, 2008, 7: 644-647.

作者简介:



洪晓斌(1979—),男,广东揭阳人,博士后,副研究员,2007年于华南理工大学获得工学博士学位,主要从事网络化测控技术的研究。E-mail: mexbhong@scut.edu.cn



刘桂雄(1968—),男,广东揭阳人,教授,博士生导师,1990年、1995年于重庆大学分别获得学士和博士学位,主要从事智能传感技术、现代检测技术与网络化控制方面的研究。E-mail: megxliu@scut.edu.cn

● 下期预告

采用负刚度调谐的硅调谐式陀螺仪

杨 波,王寿荣,李坤宇,朱 熙,曹慧亮

(东南大学 仪器科学与工程学院 微惯性仪表与先进导航技术教育部重点实验室,江苏 南京 210096)

为了验证硅调谐式陀螺仪原理的可行性,对硅调谐式陀螺仪的调谐机理,硅转子运动模态和加工工艺,信号检测与再平衡控制回路进行了研究。提出了一种利用力矩器的负刚度效应来实现硅调谐式陀螺仪调谐的新方法,推导了静电负刚度调谐的理论公式;设计、仿真、加工了硅薄片式转子平衡环部件和电容器极板,设计了电容信号敏感接口电路、升压放大电路、反馈校正电路和再平衡控制回路;在以上基础上,实现了硅调谐式陀螺仪原理样机。试验结果证实利用负刚度调谐是可行的,接口电路和再平衡控制回路设计是正确的。初步性能测试表明硅调谐式陀螺仪实现标度因数为 $1.42 \text{ mV}/((^\circ)/\text{s})$, 标度因数非线性为 2.47% , 量程为 $\pm 200 (^\circ)/\text{s}$ 。试验结果验证该陀螺仪基本原理是可行的。