

基于联合卡尔曼滤波的数字化文件融合方法研究

肖建毅[†], 蔡海滨, 李宏亮

(南方电网数字企业科技(广东)有限公司, 广东 广州 510000)

摘要:针对电力相关数字文件数据融合方法的数据纠偏机制与数据的实时状况不相符,导致数据融合效率低、数据融合效果不佳的问题,提出了基于联合卡尔曼滤波方法的电力数字文件数据融合方法。所述方法从数字文件中获取数据源,按照数据的实时情况构建对应的数据纠偏机制,并对数据降噪处理实现数据的预处理,然后,对预处理后的数据进行特征提取并进行归一化处理,利用联合卡尔曼滤波算法对电力数字文件中的数据进行快速融合。仿真实验表明,相比于改进的BP神经网络的数字文件数据融合方法和基于离散小波分解与重构的数字文件数据融合方法,所述方法的平均融合效率达到402.3条/min,具有更好的计算和融合效率。

关键词:联合卡尔曼滤波;数字文件;数据融合;方法设计

中图分类号:G642

文献标识码:A

Research on Data Fusion Method of Digital Files Based on Joint Kalman Filter Method

XIAO Jianyi[†], CAI Haibin, LI Hongliang

(CSG Digital Enterprise Technology (Guangdong) Co., Ltd., Guangzhou, Guangdong 510000, China)

Abstract: Aiming at the problem of low data fusion efficiency and poor data fusion performance due to the inconsistency between the constructed data correction mechanism and the real-time status of the data, this paper proposed a data fusion method for digital files of power system based on joint Kalman filter method. First, the data from the digital file and the real-time data are used to build data correction mechanism. Then a noise reduction calculation is used to pre-process the data. Based on normalization of the extracted data, the joint Kalman filter method is used to calculate the correlation, and realize the quick fusion of data. Simulation results show that compared to the improved BP neural network data fusion method and the discretized wavelet decomposition and reconstruction data fusion method, average fusion efficiency of the proposed method reaches 402.3 items/min, which shows that the proposed method has a higher data fusion speed and a better data fusion efficiency.

Key words: joint Kalman filter; digital files; data fusion; method design

随着电力系统信息化的不断发展,电力系统运行过程中产生的数字文件种类越来越多,数据量也越来越大。由于数字文件中的数据来源复杂、种类繁多,导致在进行数字文件信息处理时,难以从海量多类别信息里快速提取有用信息,所以有必要对

电力系统的多源、多类别数据进行数据融合。但是由于数据种类的不同,导致在对该类电力数据进行处理时难以进行快速且有效的融合,数据融合效果较差。因此,研究对电力系统多源、多种类数据的快速融合对于提高电力信息的利用效率,辅助相关

电力系统的运营做出高效的决策具有重要意义。

现有针对该类多源数据进行融合的问题已有较多研究,由于联合卡尔曼滤波算法能够对多传感器的输出结果进行融合^[1-4],因此,在电力行业相关的多个领域中应用广泛^[5-6],且有着较好的应用效果。如文献[7]提出利用扩展的卡尔曼滤波算法,对电池能量和功率状态进行数据采集和数据填充,计算采集的数据之间的关联性,以此为基础,实现数据的联合估计,实验结果表明取得了较好估计效果;文献[8]通过对磷酸铁锂电池的荷电状态进行实时监测和数据采集,将采集到的电池荷电状态数据进行聚类处理,提出利用扩展卡尔曼滤波算法,将聚类后的数据进行二次处理,完成对电池荷电状态的估算,实验结果表明,该方法估算结果较为精确;文献[9]利用配电网多源异构数据的特性,对数据进行标准化处理,提出采用联合卡尔曼滤波算法,根据数据的标准化处理结果,计算多源异构数据之间的关联性,实现多源异构数据的有效融合。实验结果表明,该方法融合效果较好;文献[10]针对电动汽车电池荷电状态的估计问题,提出一种联合扩展卡尔曼滤波的滑模观测器算法,通过在扩展卡尔曼滤波算法中加入防抖函数,在滤除噪声的同时实现在复杂车载环境下对电池SOC的高效精度;文献[11]针对电力线通信(PLC)中谐波引起的窄带干扰问题,提出了一种合并未知参数的联合卡尔曼滤波方法,用于提高增益性能、降低计算量;文献[12]针对电动汽车的锂电池荷电状态(SOC)与健康状态(SOH)问题,提出了双自适应无迹卡尔曼滤波算法(DAUKF),用于SOC与SOH的联合估算,实验结果表明,所述方法具有较高的估算精度。

基于现有卡尔曼滤波算法的应用,为了提高电力系统相关数字文件数据融合的效率,本文提出了基于联合卡尔曼滤波方法的电力数字文件数据融合方法,通过对电力数字文件数据进行预处理,利用联合卡尔曼滤波算法对预处理后的数据特征进行提取,从而实现电力数字文件数据的高效融合,为电力企业基于数据支持的决策和发展提供了可靠支撑。

1 电力企业数字文件数据融合方法

1.1 数字文件数据预处理

在现有电力企业数据库中,由于电力系统运行过程中存在多种不同类型的信息测量设备,因此所产生的数字文件中含有大量不同来源、不同类型的

数据,此外,所收集到的电力系统数据中还存在噪声,在进行数据融合时,这些噪声数据的存在会降低数据融合方法的性能,影响数据融合的效果^[13-15]。因此,在进行数据融合前,需要对数据进行预处理。本文针对数字文件的数据预处理过程主要包含三个部分,分别为数据临时存储、纠偏机制构建和数据处理。文中所述数字文件的数据预处理流程如图1所示。

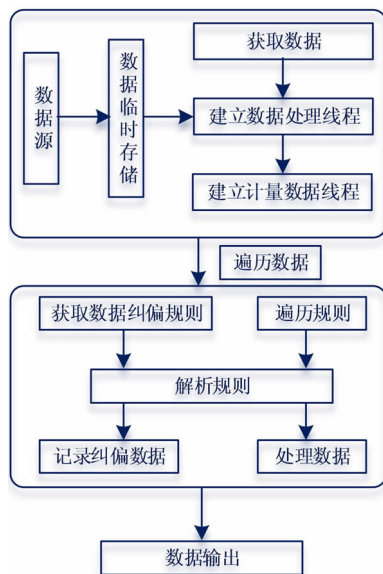


图1 数字文件预处理流程

如图1所示,进行预处理时,首先从数字文件中提取数据源,然后再根据实时数据将提取到的数据源存储至临时数据存储模块,并根据处理的数据类型,建立相对应的数据计量线程。之后,对临时数据存储模块中所存储的提取数据进行遍历,根据遍历的结果,构建相对应的数据纠偏机制。数据纠偏机制的具体构建规则如表1所示,其中 N 表述字段长度。

表1 数据纠偏规则

序号	纠偏规则名称	规则内容
1	字段长度书写规则	$= N$
		$> N$
		$< N$
		$\geq N$
		$\leq N$
2	字段范围书写规则	最小值 $<$ 且 $<$ 最大值
		最小值 \leq 且 \leq 最大值
		最小值 $<$ 且 \leq 最大值
		最小值 \leq 且 $<$ 最大值
		无穷大
3	字体要求	无穷小
		统一为英文字体

根据表 1 所构建的纠偏规则,对所提取数据的遍历结果进行判断:如果所遍历的数据中不存在异常情况,则直接将遍历数据输出;如果遍历的数据中存在异常情况,则需要对其数据进行处理,并生成相应的数据纠偏记录。之后,在对数据进行纠偏处理后,采用降噪方法对纠偏后所得到的数据进行进一步处理。具体的降噪处理计算过程如下所示:

$$\begin{cases} \text{pmu}_i = \alpha_1 y_i + (1 - \alpha_1)(\text{pmu}_{i-1} + \text{trend}_{i-1}) \\ \text{trend}_i = \alpha_2 (\text{pmu}_i - \text{pmu}_{i-1}) + (1 - \alpha_2)t_{i-1} \\ y_{i+h} = \text{pmu}_i + h \times \text{trend}_i \end{cases} \quad (1)$$

式中, pmu_i 表示时刻 i 的真实数据数值, y_i 表示降噪后的数据取值, trend_{i-1} 表示对数据进行二次降噪后所得到的数值, α_1 和 α_2 表示对数据进行降噪处理的参数, h 表示进行降噪处理的数据长度, y_{i+h} 表示经过降噪处理后的数据, t_{i-1} 表示进行降噪处理的数据。通过上述降噪处理计算,获得降噪处理后的数据,为后续的数据融合提供了较为精准的数据。

1.2 数字文件特征提取及数据归一化

在完成数据的预处理后,为了更好地确定数字文件中数据的聚类效果和分布范围,提高数据的分析和处理效率,需要对数据的特征进行提取,特征提取的具体过程如下所示:

$$\begin{cases} f = \prod_{n>1}^I Q(f_n | g_n) \\ A(g) = \frac{1}{X} \times \exp(-B \times O(g)) \end{cases} \quad (2)$$

式中, f 表示在数字文件中提取到特征数据的概率, n 表示数字文件中数据的数量, I 表示数字文件中数据的扰动系数, $A(g)$ 表示针对数字文件中数据的特征提取结果, X 表示需要进行特征提取的数据, B 表示数据特征的验证次数, Q 表示特征数据集, $O(g)$ 表示特征数据的信息量, g 表示数字文件中的数据。利用式(2)即可对数字文件中的数据进行特征提取。为了便于后续进行数据融合,文中进一步将特征数据进行归一化处理,以保证数据维度的统一^[9]。数据进行归一化的具体计算过程如下所示:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (3)$$

式中, X' 表示对特征数据进行归一化处理的结果, X_{\min} 表示特征数据的最小值, X_{\max} 表示特征数据的最大值。通过式(3)对特征数据进行处理,即可保证数据的有效值在某一个固定的区间之内。在对特征数据进行归一化处理前,需要确保数据没有异常情况,以免数据的归一化处理无法达到预期

的效果,对后续的数据融合造成影响^[16-18]。

1.3 基于联合卡尔曼滤波的融合方法

对数字文件中的数据进行预处理和特征提取后,即可对数据进行融合处理。文中采用基于联合卡尔曼滤波的融合方法对数据进行融合处理,主要处理过程如下:首先,为了将数字文件中相同的数据融合到同一个数据类型中,需要对数字文件中的数据进行关联度的计算^[11],具体计算过程如下所示:

$$\begin{cases} A_i = \frac{|z_i|}{\sum_{i=1}^n |z_i|} \\ m(\mu_i) = \frac{\sum_A m_1(A_i) \cdot m_2(B_j)}{1 - K} \end{cases} \quad (4)$$

式中, A_i 表示输出结果, z_i 表示在 i 时刻测量数据, $m(\mu_i)$ 表示数据的相关度, $\sum_A m_1(A_i)$ $m_2(B_j)$ 表示两个数据之间的冲突, K 表示数据出现冲突的时间段。由上述公式可知,数据的冲突性越强,说明两个数据的相关性越差,反之,相关性则越强。由此可知, $m(\mu_i)$ 的数值越大,说明当前两个数据的相关性越小,反之则越强^[12]。在上述计算的基础上,利用联合卡尔曼滤波算法,实现数字化文件的数据融合^[13]。其中,联合卡尔曼滤波的计算过程如图 2 所示。

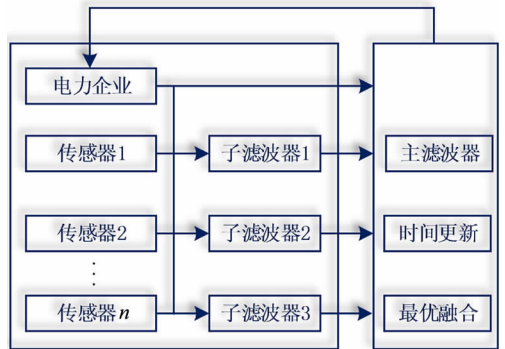


图 2 联合卡尔曼滤波计算过程

如图 2 所示,利用联合卡尔曼滤波方法对数据进行处理时,将数据的处理过程划分为两部分,分别是主滤波数据处理和局部滤波数据处理,在进行数据处理时,这两个数据处理过程是互相独立的。因此,在进行数据融合时,采用联合卡尔曼滤波算法,能够对数据进行二次融合,基于联合卡尔曼滤波方法进行数据融合的具体过程如下:

$$\begin{cases} \mathbf{X}_{k+1} = \mathbf{A}_i \mathbf{x}_k + \mathbf{B} \omega_k \\ Z'_k = \mathbf{A}_i P_k + \mathbf{Q} \\ q = \Omega_k \mathbf{A}_i (\Omega'_k)^{-1} \mathbf{Q}^{-1} \end{cases} \quad (5)$$

式中, \mathbf{X}_{k+1} 表示 k 时刻数据的状态估计值, \mathbf{x}_k 表示 k 时刻目标数据的状态向量, \mathbf{B} 表示目标数据

的状态转移矩阵, ω_k 表示数据噪声, Z_k 表示数据的实际状态值, P_k 表示数据的观测噪声, q 表示数据的融合结果, Ω_k 表示数据状态信息。利用式(5), 将数字文件中的数据以状态矩阵的形式描述, 进而进行数据的融合, 随后, 重复上述过程, 直到所有数据完成融合^[15]。

2 仿真实验

2.1 参数设置

为验证本文所述基于联合卡尔曼滤波方法的电力企业数字文件数据融合方法的效果, 文中进行了仿真实验。仿真实验所采用的实验平台参数和数据相关参数的设置如表2所示。

表2 实验参数的设置

序号	实验参数	参数设置
1	操作系统	Windows 10
2	内存	32 GB
3	操作平台	Intel Xeon CPU E5-2603v4@2.20GHz
4	数据长度	100 bit
5	数字文件大小	500 bit
6	数据传输能耗	45 nJ/bit
7	数据传输距离	50 m
8	数据点数量	100 个
9	数据接收能耗	30 nJ/bit

基于表2的仿真实验环境对所述方法进行验证时, 以某电力企业的数据库为实验过程的数据来源, 实验数据具体情况如表3所示, 本次实验所用的3个数据集均包含来源多样的电力系统运行相关数据。

表3 实验数据

数据集	数据来源	数据格式
数据集1	生产管理数据	结构化数据
	电力营销数据	非结构化数据
	地理位置数据	非结构化数据
数据集2	电力企业自动化数据	结构化数据
	电力调度数据	结构化数据
数据集3	用电信息采集数据	非结构化数据
	天气数据	非结构化数据
	舆情监督数据	结构化数据

首先对上述实验数据进行预处理。限于本文篇幅, 仅展示对数据集1进行预处理后的结果, 如图3所示。由图可知, 预处理后的数据量纲进行了统一表示, 波动幅度更小, 便于后续进行数据融合处理。

2.2 结果分析

将上述经过预处理后的数据作为本次实验数

据, 进行相关实验测试。为保证本次实验的可靠性以及展示所述方法的效果, 文中实验过程中设置了对照实验, 进行数据融合效果的对比。其中, 文中所述基于联合卡尔曼滤波方法的数字文件数据融合方法记为方法1, 基于改进的BP神经网络的数字文件数据融合方法记为方法2, 基于离散小波分解与重构的数字文件数据融合方法记为方法3。

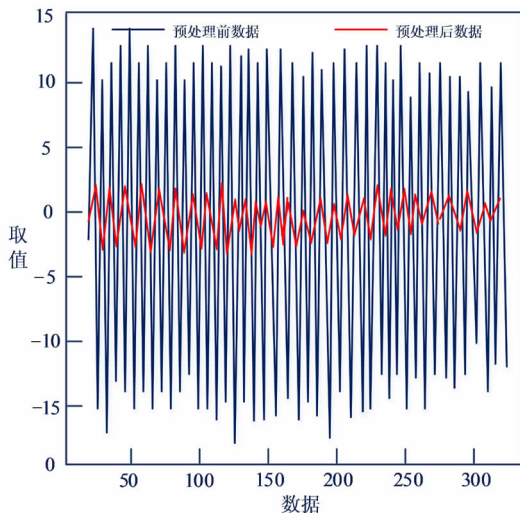


图3 数据集1的预处理结果

为对比三种方法进行数据融合的效果, 本次实验以数据的融合效率, 即进行数据融合所花费的时间为评价指标, 对比三种方法在电力数字化文件数据中的融合效果。在实验中, 利用三种方法分别对表3中的三个实验数据集进行数据融合, 对比三个数据集的融合效率。其中, 数据集1的融合效率如图4所示, 数据集2的融合效率如图5所示, 数据集3的融合效率如图6所示。

由图4、图5和图6可知, 当测试所用数据的数据规模较小时, 三种方法的融合效率接近, 而随着测试所用数据数量的增多和数据集规模的增大, 三种方法融合效率的差距逐渐增加, 其中, 方法1随着数据规模的加大, 计算时间变化幅度较小, 表明方法1的融合效率逐渐提高, 在数据集1中, 当数据规模从100条增加到700条时, 平均融合效率达到约402.3条/min, 方法2和方法3随着数据规模的加大, 计算时间增加幅度较大, 融合效率逐渐降低, 其中, 当数据规模从100条增加到700条时, 方法2的平均融合效率约为100.4条/min, 方法3的平均融合效率约为213.2条/min。由上述仿真实验可知, 相比于方法2和方法3, 使用方法1进行数据融合的计算效率更高, 这表明本文所述基于联合卡尔曼滤波算法的电力企业数字文件数据融合方法具有较好的融合速度和效率, 具有较大的潜在应用价值。

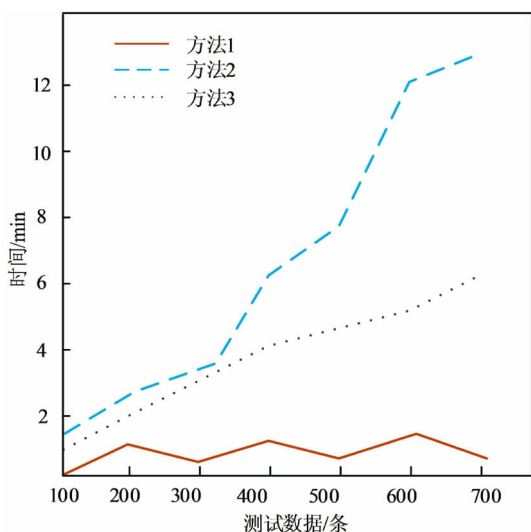


图 4 数据集 1 的融合效率对比

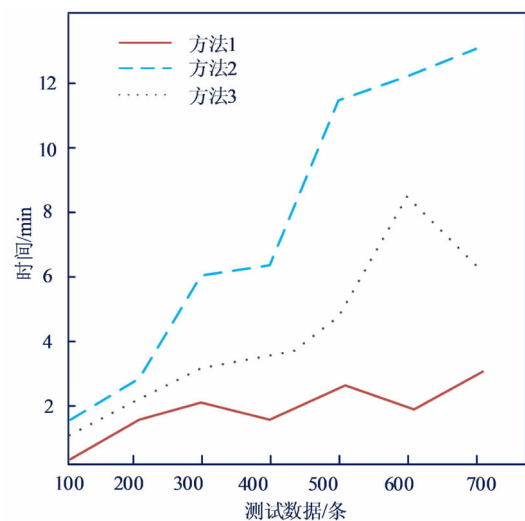


图 5 数据集 2 的融合效率对比

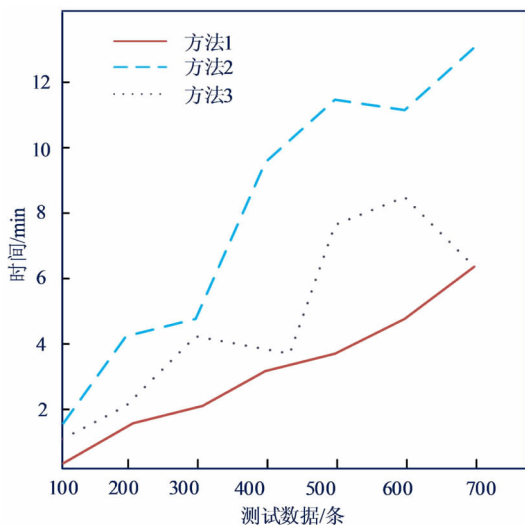


图 6 数据集 3 的融合效率对比

3 结 论

针对电力系统相关的数字档案数据融合问题,提出了基于联合卡尔曼滤波算法的数据融合方法,所述方法通过对数字文件中的数据进行预处理,并基于联合卡尔曼滤波以提高数据融合的效率。仿真实验表明,当数据规模从 100 条增加到 700 条时,所述方法的平均融合效率分别约对比方法的 4 倍和 2 倍,因此具有更好的潜在应用价值,为电力企业运营提供更高效的数据支持。

参考文献

- [1] 姜鑫蕾,杨文武,韦焱文,等. 基于卡尔曼滤波的 GNSS 系统联合定位技术研究[J]. 电子测量技术, 2021, 44(3): 109-113.
- [2] 崔深远,黄晓瑞. 基于联合卡尔曼滤波的多传感器信息融合算法及其应用[J]. 电机与控制学报, 2001, 5(3): 204-207.
- [3] 陈玉坤,司锡才,李志刚. 联合卡尔曼滤波在多传感器信息融合技术中的应用[J]. 弹箭与制导学报, 2005, 25(SD): 946-948+952.
- [4] 胡敬,田俊曦,邹世明,等. 基于自适应卡尔曼滤波的联合 RSS/TOA/INS 无人机定位算法[J]. 无人系统技术, 2022, 5(2): 62-70.
- [5] 储建新,潘炫霖,雷健新,等. 基于联合集合卡尔曼滤波的锂电池 SOC 估计[J]. 浙江电力, 2021, 40(1): 123-130.
- [6] 李晓宇,朱春波,魏国,等. 基于分数阶联合卡尔曼滤波的磷酸铁锂电池简化阻抗谱模型参数在线估计[J]. 电工技术学报, 2016, 31(24): 141-149.
- [7] 刘子豪,张雪松,孙立清,等. 基于扩展卡尔曼滤波的储能电池能量和功率状态联合估计方法[J]. 储能科学与技术, 2023, 12(3): 913-922.
- [8] 林加顺,吴乃豪,杨晓全,等. 一种联合 PID 控制与扩展卡尔曼滤波的磷酸铁锂电池荷电状态估算方法[J]. 电网技术, 2023, 47(4): 1623-1632.
- [9] 夏伟,蔡文婷,刘阳,等. 基于联合卡尔曼滤波的配电网多源异构数据融合[J]. 电力系统保护与控制, 2022, 50(10): 180-187.
- [10] 周娟,王梅鑫,杨晓全,等. 联合扩展卡尔曼滤波的滑模观测器 SOC 估算算法研究[J]. 中国电机工程学报, 2021, 41(2): 692-703.
- [11] 范永全,张家树. 基于联合卡尔曼滤波的电力线通信谐波干扰抑制[J]. 计算机应用研究, 2009, 26(10): 3706-3708.
- [12] 王若琦,王晓佳,杨淇,等. 基于双自适应无迹卡尔曼滤波算法的锂电 SOC/SOH 联合估计[J]. 机械设计与制造, 2023(1): 1-48.
- [13] 蒋清乐,吴庆先,黄永安,等. 基于多源数据融合的可视化电网发展诊断系统研究[J]. 机械与电子, 2023, 41(6): 51-54+59.
- [14] 李岩. 基于多源数据融合感知的电网设备供应链管控与预警算法设计[J]. 电子设计工程, 2023, 31(12): 169-173.
- [15] 吕鹏飞,薛国强,武欣. 利用多源地球物理图像融合方法识别物性异常——以航空电磁和航磁数据为例[J]. 地球物理学报, 2023, 66(6): 2658-2669.
- [16] 张维星,党超群,张锁平,等. 基于多传感器融合的波浪谱数据获取与传输方法研究[J]. 海洋与湖沼, 2023, 54(3): 645-652.
- [17] 封永铭,吴文宝,张书启,等. 融合振动数据和 SCADA 数据的风电机齿轮箱故障预警方法[J]. 工业控制计算机, 2023, 36(5): 1-3.
- [18] 祁炜雯,郭勤慧,朱传古,等. 基于数据融合并行特征提取的调峰电源设备状态评估方法[J]. 水电能源科学, 2023, 41(5): 203-206+215.