Lithium Battery SOC Estimation Method Study Based on Principal Component Analysis

Haiying Wang, Chao Xue, Qi Fan and Peng Liu

School of Automation, Harbin University of Science and Technology, Harbin, China
wanghy@hrbust.edu.cn

Abstract

Battery SOC is acted as an essential parameter for EV and HEV, its accurate estimate relates to whether the vehicle control strategy is implemented correctly. There has been no other prediction method, which can obtain a more accurate prediction result. To overcome this deficiency, this paper presents a method to build a battery SOC estimation models using principal component analysis (PCA). PCA cannot manage to extract non-linear factors in the parameters, which may induce prediction error based on the analysis result, in view of this kernel principle component analysis (KPCA) is applied to establish the prediction model. Corresponding simulation and experiments are conducted to verify the proposed model. The simulation results have shown that the revised model can be applied to various working condition with superior real-time ability, reliability and improved precision. The average error of the prediction results can be reduced to 1.46% in comparison with Ah measurement.

Keywords: state of charge; principle component analysis; KPCA

1. Introduction

Estimating the SOC of battery is one of the most basic and important steps in the battery management, reliable and accurate estimates remaining battery capacity is not only able to provide accurate battery for electric vehicle driver of the remaining power status, but also could provide valuable reference for the scientific management of the battery [1]. At present, common estimation methods of SOC at home and abroad are: Kalman filtering method, Ampere hour measurement, resistance method, and neural network method and extended Kalman filtering method, and so on [2-3].

Kalman filtering method can be adapted to all kinds of batteries, more accurate estimate, but the equivalent circuit model of the battery requires higher, the accuracy of prediction is often largely dependent on the equivalent circuit model of a battery, it would cause larger error because of the uncertainty of the model; resistance method is more difficult to measure the internal resistance of the battery cell, and the initial stage of discharge resistance is relatively large [4]; Ah measurement method has larger error under the condition of high temperature and current volatile situation; the artificial neural network method requires large amounts of training data for reference, it is affected by the training data and the training method; promotion of Kalman filter algorithm leads to filter divergence because of limited computer word, it would lead to inaccurate model eventually [5-6].

Using the above methods and models to estimate the state of charge of battery, consideration is unitary, and the efficiency of the battery charge and discharge, battery hysteresis, the cycle life of the batteries are not taken into account, it causes the error estimation of state of charge of battery, but it can't meet the need of real-time monitoring, environment variable and state variable, so it cannot adapt to the rapid detection of battery accurately in EV/HEV system requirements [7]. The goal of this research is to establish a

effective lithium power battery SOC estimation model of electric vehicles, aiming at the battery multivariable, nonlinear electrochemical characteristics, using mathematical statistical methods to reflect the status of lithium battery SOC parameters for analysis, on the basis of the features that reflecting power battery in real time, establishing the state of charge lithium power battery estimation model through PCA, KPCA methods, it provides theoretical and technical support about reliability and security to the conventional performance simulation testing, battery-related key performance in the course of the actual vehicle operation [8-9].

2. Introduction of Principal Component Analysis and Kernel Principal Component Analysis Principle

2.1 Principle of Principal Component Analysis

We assume that there are p random variables in the problem to be solved, we put them separately expressed as: $X_1, X_2, ..., X_p$, the main idea of principal component analysis is to transform the original variables linearly, to obtain the linear combination on these variables finally, $F_1, F_2, ..., F_k (k \leq p)$ as a new variable after the transformation, they are independent of each other, and to retain the most useful information in the original variables, and to exclude some redundancy and interference between each variable. So, the main purpose of PCA is to try to find a group to represent a linear combination of the original variables F_i, F_i is shown as formula (1).

$$\begin{cases} F_1 = u_{11} X_1 + u_{12} X_2 + ... + u_{1p} X_p, \\ F_2 = u_{21} X_1 + u_{22} X_2 + ... + u_{2p} X_p, \\ ...... \\ F_p = u_{p1} X_1 + u_{p2} X_2 + ... + u_{pp} X_p \end{cases} \tag{1}$$

2.2. Kernel Principal Component Analysis Principle

We take $x_1, x_2, ..., x_M$ as the required training sample set, the input space is expressed as $\{x_i\}$, take $\{x_i\}$ of their own dimension from space map to a high dimensional space (in feature space), then the correlation operation methods of principal component analysis in high dimensional space is the kernel principal component analysis method. Assuming the corresponding mapping is Φ, its definition as follows:

$$\begin{aligned} \Phi : R^d &\to F \\ x &\mapsto \xi = \Phi(x) \end{aligned} \tag{2}$$

In the formula, R^d is d dimension Euclidean space, it is the dimensional space that initial data $\{x_i\}$ are in, it is often called the input space, Hilbert space F is the space after the original data mapping to, it is called feature space.

After the kernel function transform, taking initial data point x mapping to F on map Φ, after the transformation of data can meet the requirements of centralized in the feature space.

$$\sum_{\mu=1}^{M} \Phi(x_\mu) = 0 \tag{3}$$

In the feature space, the covariance matrix is:

$$C = \frac{1}{M} \sum_{\mu=1}^{M} \Phi(x_\mu) \Phi(x_\mu)^T \tag{4}$$

Among it, λ is the characteristic value of C, v is the corresponding feature vector, then we could get:

$$V \in F \setminus \{0\}, C v = \lambda v \tag{5}$$

$$(\Phi(x_v) \cdot C v) = \lambda(\Phi(x_v) \cdot v) \tag{6}$$

We could prove that $\Phi(x_1), \Phi(x_2), ..., \Phi(x_M)$ can be linearly expressed all eigenvectors.

$$v = \sum_{i=1}^{M} \alpha_i \Phi(x_i) \tag{7}$$

$$\frac{1}{M} \sum_{\mu=1}^{M} \alpha_\mu (\sum_{w=1}^{M} (\Phi(x_v) \cdot \Phi(x_w) \Phi(x_w) \Phi(x_\mu))) = \lambda \sum_{\mu=1}^{M} (\Phi(x_v) \cdot \Phi(x_\mu)) \tag{8}$$

$$v = 1, 2, ..., M$$

Defining matrix K as follows:

$$K_{\mu v} := (\Phi(x_\mu) \cdot \Phi(x_v)) \tag{9}$$

$$M \lambda K \alpha = K^2 \alpha \tag{10}$$

$$M \lambda \alpha = K \alpha \tag{11}$$

Calculating the characteristic values and characteristic vectors, the projection of the test sample in the feature vector space v^k is:

$$(v^k \cdot \Phi(x)) = \sum_{i=1}^{M} (\alpha_i)^k (\Phi(x_i), \Phi(x)) \tag{12}$$

Inner product will be replaced with the kernel function:

$$(v^k \cdot \Phi(x)) = \sum_{i=1}^{M} (\alpha_i)^k K(x_i, x) \tag{13}$$

Kernel matrix can be modified as.

$$K_{\mu v} \to K_{\mu v} - \frac{1}{M} (\sum_{w=1}^{M} K_{\mu w} + \sum_{w=1}^{M} K_{wv}) + \frac{1}{M^2} \sum_{w,\tau=1}^{M} K_{w\tau} \tag{14}$$

3. Battery SOC Estimation Model using Kernel Principal Component Analysis

Step 1, taking the data that to be analyzed (including n variables, each variable has m samples) as a ($m \times n$) dimensional data matrix form.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Step 2, calculating kernel matrix K. Specific approach is to first select the Gauss radial kernel function, obtaining kernel matrix according to formula $K_{\mu v} := (\Phi(x_\mu) \cdot \Phi(x_v))$.

Step 3, through revising kernel matrix using formula (14) to obtain KL.

Step 4, calculating characteristic value $\lambda_1, ..., \lambda_n$ and characteristic vector $v_1, ..., v_n$ in kernel matrix KL through Jacobi iterative method, and take the characteristic values

sorted in descending order as $\lambda_1' > ... > \lambda_n'$, and adjusting the characteristic vectors as $v_1', ..., v_n'$.

Step 5, taking the unit orthogonal processing to characteristic vectors through Schmidt orthogonal method, obtaining $\alpha_1, ..., \alpha_n$.

Step 6, calculating the cumulative contribution rate $B_1, ..., B_n$ according to characteristic values, then give the number of principal components appropriately, marking as $\alpha_1, ..., \alpha_t (t \leq n)$.

Step 7, calculating the characteristic vector projection $Y = KL \cdot \alpha$ according to the revised kernel matrix KL and the principal components obtained in above step $\alpha = (\alpha_1, ..., \alpha_t)$, so Y is the final data through kernel principal component analysis.

Step 8, taking the least squares regression modeling according to the final data Y, to get the final battery SOC estimation model.

Establishing the battery SOC estimation model according to the estimation procedure kernel principal component analysis:

Using step 1 to step 5, to calculate the characteristic values and the corresponding units of orthogonal characteristic vectors according to voltage V_o, DC resistance R and current I, data are shown in table 1.

Table 1. Table of Eigenvalues and Eigenvectors value

Characteristic values	Characteristic vectors
1.9466	0.6998 0.7069 0.1028
0.9988	-0.1412 0.0042 -0.9900
0.0545	-0.7002 0.7073 -0.0969

Calculating the various principal component contribution rate and cumulative contribution rate using step 6, characteristic values and contribution rate values are shown in Table 2.

Table 2. Table of Eigenvalues Contribution Rate and Their Cumulative Contribution Rate

Characteristic values	Contribution rate (%)	Cumulative contribution rate (%)
1.9466	0.6489	0.6489
0.9988	0.3329	0.9818
0.0545	0.0182	1

From Table 2, the first two standard samples of principal component contribution rate has reached 98.18%, the cumulative contribution rate of the first two principal components meet the requirements according to the calculation of the principal components, so it only needs to take the first two principal components.

We could obtain the final desired data according to step 7, taking the least squares regression modeling using these data according to step 8, we could obtain SOC estimation model as follows when $T = 15 \sim 55 ℃$.

$$SOC_{KPCA} = 1.9871 + \begin{bmatrix} 0.0086 & -0.1683 & -23.3239 \end{bmatrix}$$

$$.* \begin{bmatrix} I & V_o & R \end{bmatrix}^T \tag{15}$$

Among them: $\begin{bmatrix} I & V_o & R \end{bmatrix}^T$ is the vector composed by voltage V_o, current I and DC resistance R .

4. Model Simulation and Experimental Validation

Battery operating temperature in normal temperature condition, in order to estimate the performance of the battery model SOC principal component analysis algorithm and kernel principal component analysis algorithm verification under ambient conditions, using a thermostat to control the experimental temperature, all temperatures are controlled deviation $\pm 2 ℃$. Designing the pulse discharge experiments, there are 7800 points to be sampled, sampling time and sampling interval and application of ampere hour measurement is 1s, the total amount of data are 7800. Current sampling, sampling voltage and sampling temperature change curves are shown from Figure 1 to Figure 3. Figure 4 is about contrast analysis diagram of battery SOC estimation. Table 3 is about PCA estimation results error comparison under normal temperature.
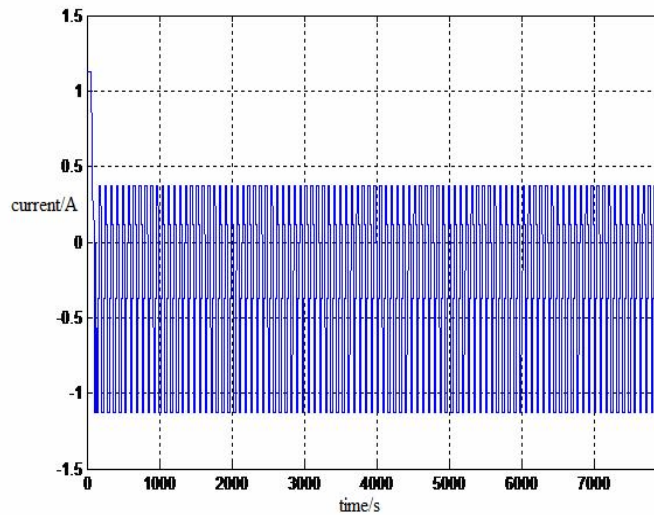

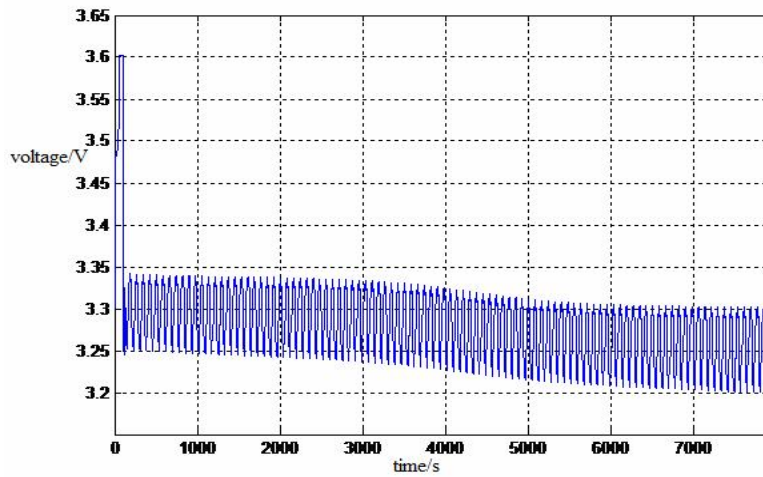**Figure 1. Curves of Sampling Current Changing**

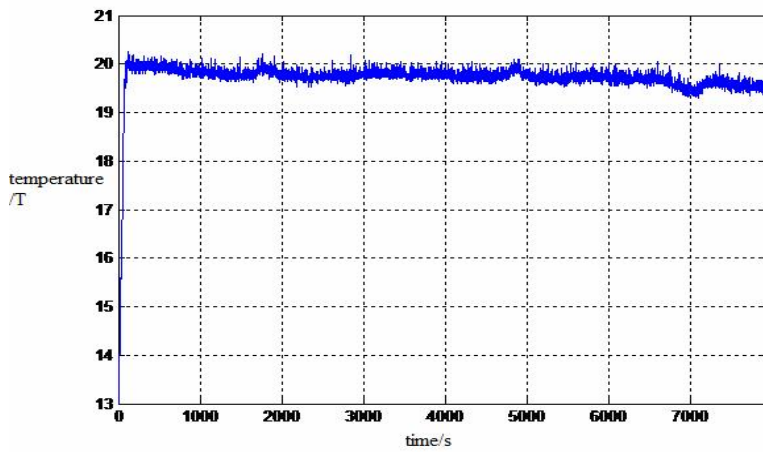**Figure 2. Curves of Sampling Voltage Changing**

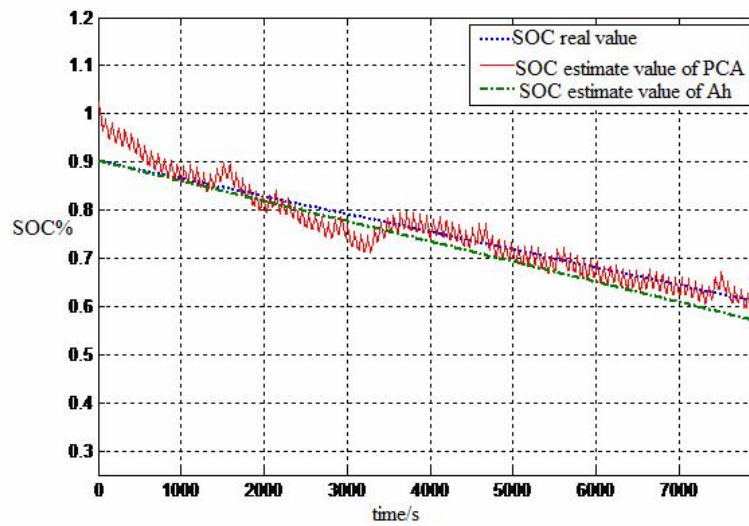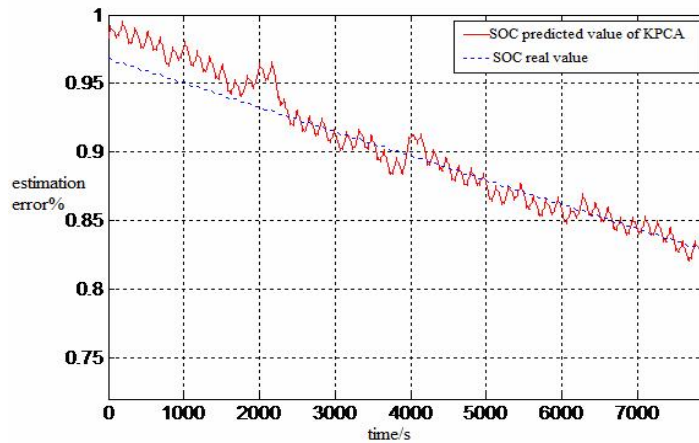

Figure 3. Curves of Temperature Changing

Figure 4. Comparative Analysis Curve of Estimated the SOC of Battery

Table 3. PCA Estimation Error Comparison Results Under Room Temperature

Algorithm	Maximum error	Root mean square error	Average estimated error
Ann measurement	5%	0.0263	2.37%
Principal component analysis method	8%	0.0230	2.29%
Error of ascension	-3%	0.0033	0.08%

From the above analysis, principal component analysis of the battery SOC estimation model is superior to the Ann measurement in a certain extent, but the advantages are very small, for the estimation of SOC battery, the influence of the parameters of the battery SOC data analysis in consideration of the linear relationship are more, and without considering the nonlinear relationship between variables, the expected effect is not ideal. So we use the kernel principal component analysis algorithm to establish the estimate model, figure 5 is the battery SOC estimation curves through kernel principal component analysis method under the room temperature. Table 4 is about PCA and KPCA estimation error comparison results under room temperature.

Figure 5. Curves of Estimated the SOC of Battery Using KPCA under Room Temperature

Table 4. PCA and KPCA Estimation Error Comparison Results under Room Temperature

Algorithm	Maximum error	Root mean square error	Average estimated error
KPCA	3.54%	0.019	1.46%
PCA	8%	0.023	2.29%
Error reduced	4.46%	0.004	0.83%

5. Conclusion

This article from the multi variable forecasting battery SOC perspective, analyzing the state parameters that reflect lithium-ion power battery SOC, based on principal component analysis and its improved algorithm to complete the battery SOC estimation model, the model on the basis of reflecting power battery real-time characteristics, it can accurately estimate the battery SOC, it could provide a certain theoretical and technical support for the actual production. We obtain the following conclusions:

1. Establishing the battery SOC estimation model through the principal component analysis algorithm combined with the least square regression method, experimental verification of the model has strong practicality, the estimation accuracy is 2.29%, and it is superior to estimation model for ampere hour measurement.

2. In view of the fact that the PCA algorithm cannot extract nonlinear factor in the parameters, we put forward to build battery SOC estimation model using kernel principal component analysis algorithm, experimental validation of the predictive accuracy of the model is 1.46%, the estimation accuracy is higher than PCA estimation model, it can give reference and guidance for the actual battery SOC prediction.

Acknowledgements

References

[1]   H. Liu, "Study on SOC estimation method of lithium-ion battery based on KEF for electric vehicle[D]", Master degree thesis of Beijing Jiaotong University, vol.8, no.5, (2010), pp.195-199.

[2]   J. Chatzakis, K. Kalaitzakis, N. C. Voulgaris and S. N. Manias, "Designing a new generalized battarry management system", IEEE Trnations on industrial electronics, vol.50, no.5, pp. 990-999.

[3]   B. Liu, Y. Wang and T. Yin, "Battery management system for electric vehicle", Electrical automation, vol.32, no.1, (2010), 60-62.

[4]   K. W. E. Cheng, B.P. Divakar, H. J. Wu, "Battery-Management System(BMS) and SOC development for electrical vehicles", IEEE Transactions on vehicular technology, vol.60, no.1, (2011), 76-88.

[5]   X. Gong, B. Qi and Y. Liu, "Electric vehicle power battery model and estimation strategy of SOC", Power supply technology, no. 10, (2004), pp.633-636.

[6]   J. Kim, G.-S. Seo, C. Chun, "OCV hysteresis effect-based SOC estimation in extended Kalman filter algorithm for a LiFePO4/C Cell", Electric Vehicle Conference (IEVC), vol.4-8, (2012), pp.1-5.

[7]   G. Wang and H. Ye, "Principal component analysis and partial least square method", Beijing: Tsinghua University press, (2012), pp.7-9.

[8]   W. Ding and Y. Fu, "Automatic target recognition of KPCA feature extraction algorithm", Application of science and technology, no.38, (2011), pp.1-2.

[9]   Z. Du, H. Tu and G. Geng, "Research and application of KPCA method", Computer engineering and Application, vol.46, no.7, (2011).