Fault Detection Based on Incremental Locally Linear Embedding for Satellite TX-I

Cheng Yuehua(程月华)[1*], *Hu Guofei*(胡国飞)[2], *Lu Ningyun*(陆宁云)[2], *Jiang Bin*(姜斌)[2], *Xing Yan*(邢琰)[3]

1. College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P. R. China
2. College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P. R. China
3. National Laboratory of Space Intelligent Control, Beijing Control Engineering Institute, Beijing 100000, P. R. China

(Received 4 November 2014; revised 24 March 2015; accepted 25 April 2015)

Abstract: A fault detection method based on incremental locally linear embedding (LLE) is presented to improve fault detecting accuracy for satellites with telemetry data. Since conventional LLE algorithm cannot handle incremental learning, an incremental LLE method is proposed to acquire low-dimensional feature embedded in high-dimensional space. Then, telemetry data of Satellite TX-I are analyzed. Therefore, fault detection are performed by analyzing feature information extracted from the telemetry data with the statistical indexes T^2 and squared prediction error(SPE) and SPE. Simulation results verify the fault detection scheme.

Key words: incremental locally linear embedding (LLE); telemetry data; fault detection; dimensionality reduction; statistical indexes

CLC number: V467 **Document code**: A **Article ID**: 1005-1120(2015)06-0600-10

0 Introduction

Satellite engineering has been increasingly critical in spacecraft research[1]. Satellites in orbit transmit telemetry data to identify satellite running status and implement tasks[2]. Significant improvements in space technology inevitably lead to more complex behavior of satellites, which in turn brings about higher demand for reliability. Meanwhile, large-volume telemetry data make data analysis difficult. Conventional data-driven fault diagnosis encounters significant challenges when applied to satellites.

Since fault information is contained in satellite telemetry data, quite a number of fault diagnosis methods have been proposed based on the features extracted from the data. Yu et al. developed a fault detection and diagnosis (FDD) scheme by combining principal component analysis (PCA) with support vector machines (SVRs)[3]. However, feature extraction based on PCA do not work well with nonlinear data. Fujimaki et al. proposed a novel "knowledge-free" anomaly detection method for spacecrafts based on kernel feature space (KFS) and directional distribution, which employed normal historical telemetry data to construct a system behavior model and then monitored current system running by comparing incoming data with the model[4]. However, the selection of the consistency of appropriate kernels for heterogeneous multidimensional data is inconvenient for its applications. Yang et al. presented several data-mining methods based on data-driven FDD for satellite telemetry data, which showed great potential[5]. Recently, several novel feature extraction methods based on manifold learning have been reported for handling nonlinear high-dimensional data. A new approach was adapted for intelligent fault diagnosis based on locally linear embedding (LLE)[6], a kind of

* **Corresponding author**: Cheng Yuehua, Associate Professor, E-mail: chengyuehua@nuaa.edu.cn.

How to cite this article: Cheng Yuehua, Wang Tao, Lu Ningyun, et al. Fault detection based on incremental locally linear embedding for satellite TX-I[J]. Trans. Nanjing U. Aero. Astro., 2015,32(6):600-609.
http://dx.doi.org/10.16356/j.1005-1120.2015.06.600

manifold. An extended LLE algorithm was also introduced in Ref.[7]. Manifold learning has been demonstrated in various applications such as face recognition[6] and machinery fault diagnosis[8].

Most of these algorithms are operated in a batch mode, i. e., all data are available for training. Nevertheless, they fail to provide a model or formula to extract features for sequentially incoming data. When new data arrive, one needs to execute the entire algorithm again with original data augmented by new data. This restricts these algorithms' applications, especially in a changing and dynamic environment.

Information can be ascertained with new incoming incremental data. Various incremental manifold learning methods have been proposed[9-11]. We introduce Satellite TX-I, as the research object, with telemetry data. It is unrealistic to expect to directly grasp the features of these data because of complex working conditions. To extract features of parameters in telemetry data, an incremental LLE algorithm is employed herein. Then, the anomalies hidden in the data can be detected by using statistical indexes T^2 and squared prediction error (SPE) and SPE acting on the extracted features.

1 Preliminary Overview

How to extract useful information from massive data in this age is a big challenge for data analysis. In the past decades, manifold learning, a new type of feature extraction method, was developed. Compared with other linear methods, manifold learning provides an effective way to handle nonlinear data. The LLE algorithm is an unsupervised manifold learning method that extract features based on geometric intuition.

1.1 Conventional LLE

Given a data set $\boldsymbol{X}=\{\boldsymbol{x}_i\in\mathbf{R}^D, i=1,2,\cdots,N\}$ in a high-dimensional input space, the LLE algorithm is used to regain the low-dimensional representation of $\boldsymbol{X}$, which is presented as $\boldsymbol{Y}=\{\boldsymbol{y}_i\in\mathbf{R}^d, i=1,2,\cdots,N\}\ (d<D)$. Based on geometric intuition, the LLE algorithm aims to preserve the local structure of the high-dimensional data while mapping the original data into low-dimensional space. In summary, the LLE algorithm mainly contains three steps as follows:

Step 1 Find the k nearest neighbors $\{\boldsymbol{x}_j^i, j=1,2,\cdots,k\}$ for each $\boldsymbol{x}_i\in\boldsymbol{X}$;

Step 2 Reconstruct weights w_{ij} of $\boldsymbol{x}_j^i$ to minimize the reconstruction error of $\boldsymbol{x}_j^i$ and $\boldsymbol{x}_i$;

Step 3 Compute the low-dimensional embeded $\boldsymbol{Y}$ of $\boldsymbol{X}$ without changing the reconstruction weights obtained in Step 2.

In Step 1, the Euclidean distance is always taken as the criteria to select k nearest neighbors $\boldsymbol{x}_j^i$ for $\boldsymbol{x}_i$. Then, the local geometry of the data can be characterized by linear coefficients w_{ij}, which are acquired by reconstructing the data points from their own k neighbors.

In Step 2, the optimal weights can be obtained by solving the constrained least-squares problem as

$$\boldsymbol{W}=\arg\min_{\boldsymbol{W}}\left(\sum_{i=1}^{N}\|\boldsymbol{x}_i-\sum_{j=1}^{k}w_{ij}\boldsymbol{x}_j^i\|^2\right) \tag{1}$$

where $\sum_{j=1}^{k}w_{ij}=1$. In addition, $\boldsymbol{W}_i=[w_{i1},w_{i2},\cdots,w_{ik}]$ stands for the ith row vector of the reconstruction weight matrix $\boldsymbol{W}$. Moreover, $w_{ij}=0$ considering that $\boldsymbol{x}_j^i$ does not belong to the neighbor of $\boldsymbol{x}_i$. The whole $N\times N$ weight matrix $\boldsymbol{W}$ for the input data $\boldsymbol{X}$ will be obtained based on steps 1—2.

In step 3, the dimensionality reduction is utilized by retaining the same reconstruction weights between the points and the corresponding neighbors. Each high-dimensional input $\boldsymbol{x}_i$ is mapped into a low-dimensional output $\boldsymbol{y}_i$ representing global internal coordinates in the manifold, which can be accomplished by minimizing the embedding cost function

$$\boldsymbol{Y}=\min_{\boldsymbol{Y}}\left(\sum_{i=1}^{N}\|\boldsymbol{y}_i-\sum_{j=1}^{k}w_j^i\boldsymbol{y}_j^i\|^2\right) \tag{2}$$

where $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}_i\boldsymbol{y}_i^{\mathrm{T}}=\boldsymbol{I}$, $\sum_{i=1}^{N}\boldsymbol{y}_i=0$, and $\boldsymbol{I}$ stands for an identity matrix. Note that the embedding is calculated by the reconstruction weights without involving original data $\boldsymbol{x}_i$ in the constrained equation. In other words, the embedding is determined entirely by the geometric information encoded by the weights. To solve the optimization

problem, a new $N\times N$ sparse and symmetric matrix $\boldsymbol{M}=(\boldsymbol{I}-\boldsymbol{W})^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{W})$ is defined. Furthermore, the d-th coordinate output by LLE corresponds to the $(d+1)$th smallest eigenvector of $\boldsymbol{M}$. Thus, for efficiency and convenience, the low-dimensional embedding calculation is finally translated into a matrix eigenvector computation.

However, the LLE algorithm performs dimensionality reduction in a batch mode, which means that the LLE fails to provide an explicit way of mapping the high-dimensional space into the embedded feature space. When new data points arrive, one has to rerun the LLE algorithm to achieve the low-dimensional features with an augmented data set. Considerating some special cases, like large data sets with high dimensionality in a dynamic environment, the LLE algorithm seems to be less attractive because it fails to accommodate new data. Thus, the incremental LLE (I-LLE) algorithm gradually becomes interesting and appealing.

1.2 Incremental LLE

Many attempts have tried to tackle the problems mentioned above, including local linear regression (LLR)[12] and local linear projection (LLP)[13]. Assuming that the embedded manifold is locally linear, LLR and LLP are conducted to reveal implicit projection between high-dimensional space and embedded feature space. The implementation of LLP algorithm is presented in Table 1.

Table 1 Implementation of LLP

Interface	Input	$\boldsymbol{x}_{\mathrm{new}}$ New high-dimensional sample
	Output	$\boldsymbol{y}_{\mathrm{new}}$ d dimensional embedded coordinate
Procedure		**Step 1** Based on the Euclidean distance metrices, one finds the k nearest neighbors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_k$ for new sample $\boldsymbol{x}_{\mathrm{new}}$ in high-dimensional space; **Step 2** By solving a linear least-squares regression problem, a mapping matrix $\boldsymbol{A}$ satisfying $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$ can be obtained, where $\boldsymbol{A}=[a_1, a_2, \cdots, a_d]^{\mathrm{T}}\in\mathbf{R}^{d\times D}$, $\boldsymbol{X}=[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k]$, and $\boldsymbol{Y}=[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k]$ represent the low-dimensional coordinate corresponding to the k nearest neighbors of $\boldsymbol{x}_{\mathrm{new}}$; **Step 3** Compute d dimensional embedding by means of the following equation $\boldsymbol{y}_{\mathrm{new}}=\boldsymbol{A}\boldsymbol{x}_{\mathrm{new}}$

In Step 2, the linear least-squares regression problem can be presented as

$$\boldsymbol{A}=\arg\min_{\boldsymbol{A}}\left(\sum_{i=1}^{k}\|\boldsymbol{y}_i-\boldsymbol{A}\boldsymbol{x}_i\|^2\right) \tag{3}$$

With the aid of LLP, one can directly calculate the low-dimensional representation without repeatedly carrying out the complex LLE algorithm. Eq. (3) shows that different mapping matrices are developed for each new sample, which can ensure the accuracy of the calculated embedding to a certain extent. Furthermore, a fixed mapping matrix can be used because the new samples are geometrically similar. Then, the LLP algorithm can be seen as LLR, which only uses the whole structure of the data set without finding k nearest neighbors for a new incoming sample. Namely, the solid mapping matrix $\boldsymbol{A}$ is calculated with the original training set $\boldsymbol{X}$ and the associated low-dimensional coordinates $\boldsymbol{Y}$ and $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$.

However, LLP is to find low-dimensional representations of the new data without a feedback loop. By extending the neighbors directly, the dimensionality reduction algorithm cannot guarantee optimal results. In other words, not only the low-dimensional embedding of the new incoming data but also the effects of these data on the original data set should be taken into account. In this paper, novel I-LLE is presented to map new samples into the low-dimensional feature space with a feedback loop.

Given the training set $\boldsymbol{X}=\{\boldsymbol{x}_i\in\mathbf{R}^D, i=1, 2, \cdots, N\}$ and the weight matrix $\boldsymbol{W}$, the k nearest neighbors of $\boldsymbol{x}_i$ can be presented as $\{\boldsymbol{x}_{i(1)}, \boldsymbol{x}_{i(2)}, \cdots, \boldsymbol{x}_{i(k)}\}$, denoting Δ_i^j as the distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. Assume that the nearest neighbors are listed in order, then $\Delta_i^{i(1)}\leqslant\Delta_i^{i(2)}\leqslant\cdots\leqslant\Delta_i^{i(k)}$. The task of I-LLE is to update the weight matrix $\boldsymbol{W}$ and the cost function $\boldsymbol{M}$ mentioned in Section 1.1 without rerunning the entire LLE algorithm. The procedure of I-LLE can be accomplished through two sequential steps as follows:

Step 1 Update the weight matrix $\boldsymbol{W}$ with the help of the Euclidean distance criteria;

Step 2 Recompute the cost function $\boldsymbol{M}$ corresponding to the change of $\boldsymbol{W}$.

Given a new incoming sample $\boldsymbol{x}_{\mathrm{new}}$, find its k nearest neighbors and compute the corresponding

weight vector, denoted as $\boldsymbol{W}_{\text{new}}$. Step 1 is described in Table 2.

Table 2 Weight matrix update

Function	Weight matrix update
Procedure	(1) Compute the distances between the new sample and each data point in the original set, denoting Δ_i^{new} ($i=1,2,\cdots,N$), as the distances; (2) For each point $\boldsymbol{x}_i$, if $\Delta_i^{\text{new}}<\Delta_i^{i(k)}$, then the neighbors of $\boldsymbol{x}_i$ should be updated as $\{\boldsymbol{x}_{i(1)},\boldsymbol{x}_{i(2)},\cdots,\boldsymbol{x}_{i(k)},\boldsymbol{x}_{\text{new}}\}$. Put the new neighbors in an ascending order $\{\boldsymbol{x}'_{i(1)},\boldsymbol{x}'_{i(2)},\cdots,\boldsymbol{x}'_{i(k+1)}\}$. Simultaneously, the update of the corresponding weight vector for $\boldsymbol{x}_i$ should be executed, following procedure (1).

Only the changed neighbors have been considered when updating the weight matrix, which greatly improves the efficiency of the I-LLE algorithm. From a mathematics perspective, procedure (1) focuses on $\boldsymbol{x}_i$, whose neighbors have been updated. Finally, a new $(N+1)\times(N+1)$ weight matrix $\boldsymbol{W}'$ is obtained from updating the weight matrix $\boldsymbol{W}$ augmented by the weight vector $\boldsymbol{W}_{\text{new}}$.

Determining How to facilitate the new weight matrix to find low-dimensional representation naturally promotes research on cost matrix $\boldsymbol{M}'$. As mentioned in Section 1.1, the calculation of the cost function can be "translated" as

$$\boldsymbol{M}'=(\boldsymbol{I}-\boldsymbol{W}')^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{W}') \tag{4}$$

Since only a local part of the weight matrix is updated, one can calculate the cost function in terms of the mutative points. Compared to the conventional LLE, I-LLE greatly improves the efficiency of calculating a new embedding when an existing data set is being slightly modified. Furthermore, it considers not only the explicit mapping from samples into embedded space but also the effects of the new samples on the original data set.

I-LLE provides a way to acquire low-dimensional embedding of the samples without entirely rerunning the LLE algorithm.

2 Feature Extraction for Telemetry Data

Since Satellite TX-I transmit large amounts of telemetry data to ground stations every day, it is an realistic to grasp the embedded features through visualization. However, the I-LLE algorithm can exposed hidden features of the data only if specific cases are taken into account.

2.1 Basic features of telemetry

2.1.1 Observable area

A satellite travels in its orbit and transmits data to ground stations in a specified area. Telemetry data of Satellite TX-I are from an arc area in orbit, e. i., the so-called observable area. The observable area is so small that the observation time is very limited. Therefore, only finite telemetry data are sent to ground stations. Furthermore, the cycle period of the observable area can be known only if the satellite is operated in a normal model.

2.1.2 Rising orbit and descending orbit

The wholel orbit around the Earth can be divided into two parts: Rising orbit and descending orbit, which correspond to the local nighttime and daytime, respectively. In terms of latitude, a satellite in rising orbit moves from south to north; while the opposite moving tendency means the satellite in descending orbit. Moreover, these two tendencies have different effects on system parameters. Even the same parameters may present different features in different orbit regions. In this study, parameters related to the satellite attitude control system are taken into account to demonstrate the effects of the orbit tendencies.

For Satellite TX-I, the telemetry data of three parameters, pitch angle, pitch velocity, and x axis of magnetometer A, are presented in Fig. 1. Fig. 1 implies the effects of the couple orbits on the parameters. Furthermore, the single orbit (rising or descending) remains for a short period for each parameter as the small observable area mentioned in subsection 2.1.1. Meanwhile, the curves in Fig. 1 are unsmooth because of the noise in telemetry data. For Satellite TX-I, the realtime telemetry data points are transmitted to a ground station every 2 s. In Fig. 1, the descending orbit of the pitch angle covers almost 286 points. In other words, the descending orbit re-

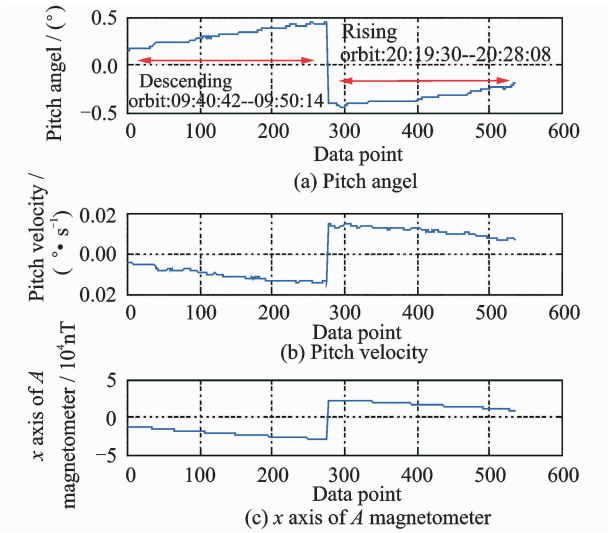

Fig. 1 Telemetry data of different parameters on 1 January 2012

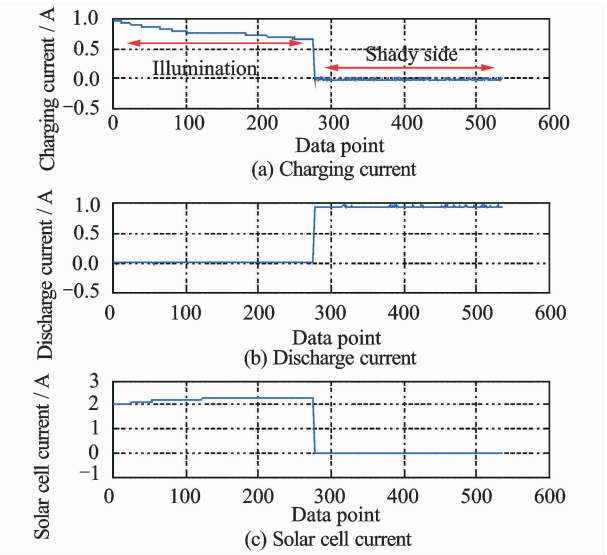

Fig. 2 Telemetry data of current on 1 January 2012

mains for almost 10 min, which can be calculated from the detailed time shown in Fig. 1.

The features of the couple orbits perfectly demonstrate that parameters hold various features in different orbits, which should be considered for future feature extraction.

2.1.3 Illumination

As they revolve around the Earth, satellites are not always exposed to sunlight, which can affect the operation mode. Specifically, the energy working condition on the shady side significantly differs from that in the illumination region. As an example, the charging current and the discharge current of TX-I on 1 January 2012 are shown in Fig. 2. Charging current runs beyond zero under the illumination condition, while the discharge current stays around zero, which illustrates that illumination 'perturbs' the working condition of the energy system in satellites. To comprehend the energy system of Satellite TX-I, the solar cell current on the same day is also presented in Fig. 2.

Therefore, features of the parameters are greatly affected by working conditions and environment. It is possible to capture the feature of a single parameter from corresponding telemetry data but the fact that hundreds of satellite parameters remain to be analyzed makes this single-parameter-based analysis impractical. Furthermore, several parameters are needed to study a subsystem, including a satellite attitude control system, which greatly increases the difficulty. Thus, the I-LLE algorithm is employed to perform feature extraction.

2.2 Application of I-LLE to feature extraction

Feature extraction mainly in clude two aspects: (1) Dimensionality reduction, which provides the low-dimensional presentation for the high-dimensional space; and (2) the preservation of the main features. In other words, the main features of the data are retained, while the redundant information should be ignored or discarded. Therefore, feature extraction plays an important role in data analysis. Furthermore, feature extraction should be performed by considering that the real-time telemetry data are affected by working conditions.

In this section, the telemetry data of some parameters are used as "guinea pigs" to validate the I-LLE algorithm. To avoid losing generality, five parameters related to the satellite attitude from 1—5 January 2012, are employed. Additionally, only one kind of the orbits (the rising or the descending) are considered due to the feature differences between the two orbits. However, the I-LLE algorithm can not be operated until it complete some preliminary treatments on the telemetry data. First of all, it must remove the outliers falling short of the overall trend. Then, filtering method is highly recommended as it can consequently improves the feature extraction efficiency.

In addition, normalization for different units of different parameters is in great demand.

The pretreated 276 data points on 1 January 2012 are used as the training data, while the data points from 2—5 January 2012, are used as the testing data, which are shown in Fig. 3(a).

The ILLE-based feature extraction should set two parameters in advance: Reduced dimensionality (intrinsic dimensionality) d and the number of neighbors k. In addition, the effects of these two parameters should not be ignored. Take d as an example. If d is set too large, the mapping of I-LLE will contain a great deal of redundant information and enhance the noise; on the contrary, a small d may lead to embedded data overlapping. The selection of these parameters is beyond the scope of this paper. Several methods of determining the parameters have been provided[6,14,15]. We set $d=3$ and $k=12$. To directly compare the I-LLE algorithm with the conventional LLE algorithm, the dimensionality reduction results for the telemetry data from 2—5 January 2012 are shown in Fig. 3.

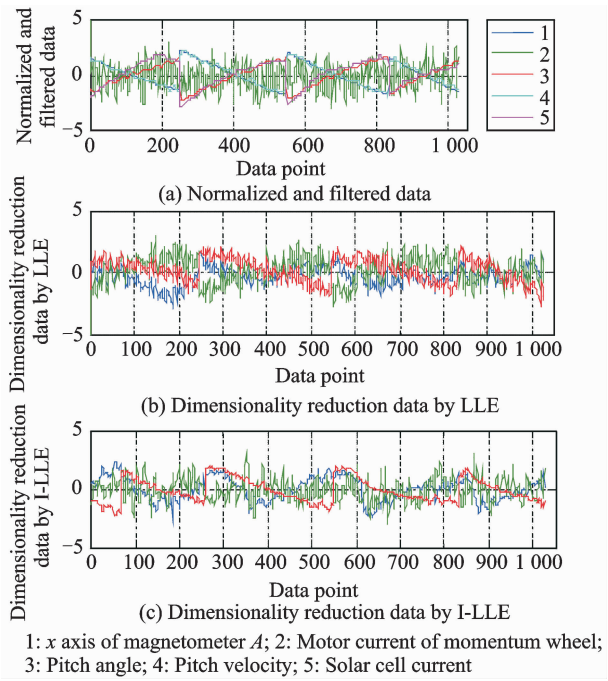

Fig. 3 Pretreated data for 2—5 January 2012, and dimensionality reduction results

Fig. 3 shows that the I-LLE algorithm can also reduce dimensionality like the conventional one. As discussed in Section 2, I-LLE performs the feature extraction with a feedback loop, which greatly improves the accuracy of the extraction. The accuracy of the conventional LLE algorithm can be somewhat improved by changing its parameters, like k without a feedback loop while neglecting performance efficiency. The difference is that the I-LLE algorithm only takes the changed part into account without running the entire algorithm, which improves execution efficiency. And using feedback loop directly affects the accuracy of feature extraction. Meanwhile, the I-LLE algorithm has the potential to map the high-dimensional fault samples into embedded low-dimensional space with obtaining the fault information, which finally can be used as a tool to perform fault detection.

3 Fault Detection

Unlike conventional model-based fault detection methods, the I-LLE algorithm is good at solving data-based problems. In many cases, the reality model of the samples was located in a low-dimensional manifold embedded in the high-dimensional sample space[16,17]. From a geometry perspective, the data belonging to the same fault class lie in or near a manifold embedded in the high-dimensional space [18]. In other words, with the help of the feature extraction of I-LLE, the fault information can finally be enveloped in the embedded space. Thus, a new fault detection based on an I-LLE algorithm is presented. By capturing the embedded manifold feature, the algorithm maps the high-dimensional space into a low-dimensional embedded space to detect faults. To visualize the fault information, the statistical indexes T^2 and SPE and SPE are employed herein.

3.1 Fault detection scheme

As mentioned previously, fault detection is performed in the embedded space by using the statistical indexes T^2 and SPE and SPE. The main problem is how to use these indexes to monitor the embedded space. At the macroscopic level, the I-LLE algorithm is a kind of modeling problem. Thanks to the new incoming data, I-LLE continually updates its original database (seen as the model), starting with the training data. Meanwhile, the statistical indexes T^2 and SPE and

SPE should be introduced to sense the variation of the new incoming samples and use the normal operating condition model.

SPE is adopted to measure the noisy part of the process information and the variability deviating from the normal operating condition model, which is also called the residual space. The calculation of SPE for a new incoming sample is presented as

$$\mathrm{SPE}=\| \boldsymbol{x}_{\mathrm{new}}-\hat{\boldsymbol{x}}_{\mathrm{new}} \|^2 \tag{5}$$

where $\hat{\boldsymbol{x}}_{\mathrm{new}}$ represents the reconstruction of the sample $\boldsymbol{x}_{\mathrm{new}}$. Moreover, $\hat{\boldsymbol{x}}_{\mathrm{new}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\ \boldsymbol{x}_{\mathrm{new}}$, where $\boldsymbol{A}=\boldsymbol{Y}\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}})^{-1}$, which can be calculated from Eq. (3)[19,20]. The popular Hotelling's T^2 statistic is used to measure the variation within the model space. Given the embedded data $\boldsymbol{y}_{\mathrm{new}}$ of the input sample $\boldsymbol{x}_{\mathrm{new}}$ in the I-LLE model space, the statistic associated with x_{new} is measured by

$$T^2=\boldsymbol{y}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{y}_{\mathrm{new}} \tag{6}$$

where $\boldsymbol{\Lambda}$ is the sample covariance matrix of $\boldsymbol{Y}$ in the normal operating condition that is

$$\boldsymbol{\Lambda}=\frac{\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}}{N-1} \tag{7}$$

The statistical indexes T^2 and SPE and SPE not only ensure the use of the normal model for I-LLE but also promote the fault detection scheme that is presented as follows:

Step 1 Collect the normal operating condition data $\boldsymbol{X}_0=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N\}$ as the training data and filter and normalize $\boldsymbol{X}_0$;

Step 2 Perform the conventional LLE algorithm on $\boldsymbol{X}_0$ to obtain the embedded low-dimensional coordinates $\boldsymbol{Y}_0$;

Step 3 Approximate the mapping matrix $\boldsymbol{A}_0$ that satisfies $\boldsymbol{Y}_0=\boldsymbol{A}_0\boldsymbol{X}_0$ by using the LLR method;

Step 4 Obtain the new incoming data; filtering and normalization should be conducted to obtain $\boldsymbol{X}_{\mathrm{new}}$;

Step 5 Conduct dimensionality reduction with the help of the mapping matrix, $\boldsymbol{A}_0$ i. e., $\boldsymbol{Y}_{\mathrm{new}}=\boldsymbol{A}_0\boldsymbol{X}_{\mathrm{new}}$;

Step 6 Calculate the statistical indexes T^2 and SPE and SPE for the new embedding and present the fault detection result;

Step 7 Update the mapping matrix $\boldsymbol{A}_0$ by performing the I-LLE algorithm with the normal new data $\boldsymbol{X}_{\mathrm{new}}$;

Step 8 Repeat Steps 4—7.

It can be seen that the I-LLE algorithm is performed on the new incoming data that are taken to update the I-LLE model. The database is refreshed, and then the embedding space is updated. Undoubtedly, the mapping matrix will be revised, which can ensure the accuracy of the dimensionality reduction. Meanwhile, the LLR method is used herein because of the geometrical similarity of the telemetry data between different dates.

3.2 Case Study

In this study, the telemetry data of Satellite TX-I on 1 January 2012 are employed as the training set, which triggers the execution of the I-LLE algorithm on the telemetry data from 2—5 January 2012. Moreover, only the data of the descending orbit are adopted. The intrinsic dimensionality $d=3$ and the number of the neighbors $k=12$. To verify the performance of the fault detection scheme based on the I-LLE algorithm, numerical simulations of different fault types are presented in the following sections.

3.2.1 Case 1: Constant fault for pitch velocity

The constant bias occurs from 120 to 170 points of pitch velocity. To illustrate the I-LLE algorithm, the statistical detection results are shown in Fig. 4. Furthermore, Fig. 5 presents the detection results based on the conventional LLE algorithm.

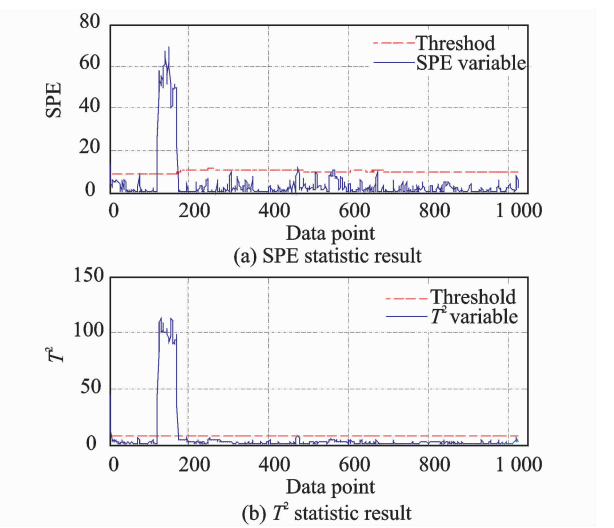

Fig. 4 Statistic detection results for constant fault based on I-LLE algorithm

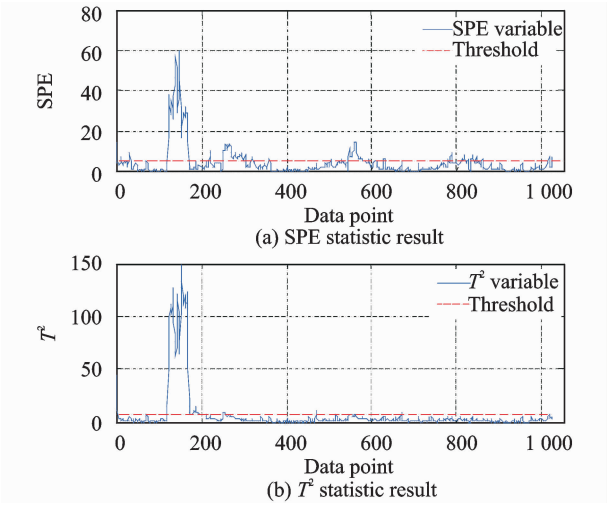

Fig. 5 Statistic detection results for constant fault based on LLE algorithm

Fig. 4 shows that both T^2 and SPE and SPE change over the threshold when fault information is introduced. In addition, the false alarm rate (FAR) and the missing alarm rate (MAR) for these statistical indexes are given in Tables 3, 4. The results based on LLE, by contrast, do not show a promising perspective. The corresponding FAR and MAR are also given in Tables 3, 4. Although both algorithms can correctly detect the fault information, the results based on LLE sometimes provide misleading data. Owing to the effect of the feedback loop, the thresholds of the statistical indexes in I-LLE are metabolic. In the conventional LLE algorithm, there is no feedback loop, which means that the database will not be refreshed. Additionally, the LLR method, which is used to find the mapping matrix, determines the static thresholds for the statistical indexes because the training data are set.

Table 3 FAR and MAR of SPE statistic with constant fault

Algorithm type	FAR (SPE statistic)/%	MAR (SPE statistic)/%
I-LLE	2.148 4	0
LLE	23.471 4	0

Table 4 FAR and MAR of T^2 statistic with constant fault

Algorithm type	FAR (T^2 statistic)/%	MAR (T^2 statistic)/%
I-LLE	1.171 8	0
LLE	11.045 4	0

3.2.2 Case 2: Time-varying fault for pitch velocity

Time-varying fault occurs from 120 to 170 points. The fault detection scheme based on the I-LLE algorithm is applied to obtain the statistical results presented in Fig. 6. Fault detection based on the conventional LLE algorithm is shown in Fig. 7. Intuitively, FAR and MAR of the statistical indexes for the algorithms are first provided in Tables 5, 6.

Table 5 FAR and MAR of SPE statistic with time-varying fault

Algorithm type	FAR (SPE statistic)/%	MAR (SPE statistic)/%
I-LLE	5.664 1	2
LLE	23.511 3	18

Table 6 FAR and MAR of T^2 statistic with time-varying fault

Algorithm type	FAR (T^2 statistic)/%	MAR (T^2 statistic)/%
I-LLE	5.078 1	4
LLE	10.266 9	20

Compared with Case 1, the time-varying fault detection does not perform as well. However, the statistical indexes based on the I-LLE algorithm show great adaptability, while those based on the conventional LLE algorithm remain in a conservation position, which can also be seen in Figs. 6, 7.

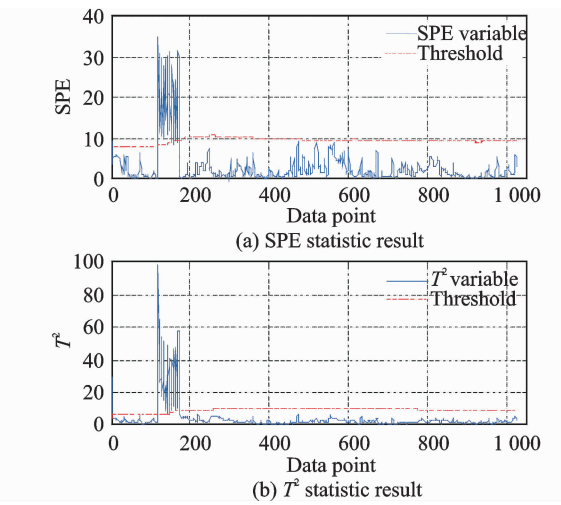

Fig. 6 Statistic detection results for time-varying fault based on I-LLE algorithm

Fig. 6 shows that the indexes jump over their associated threshold when the fault occurs and return to normal when the fault disappears. However, the performance of the fault detection based

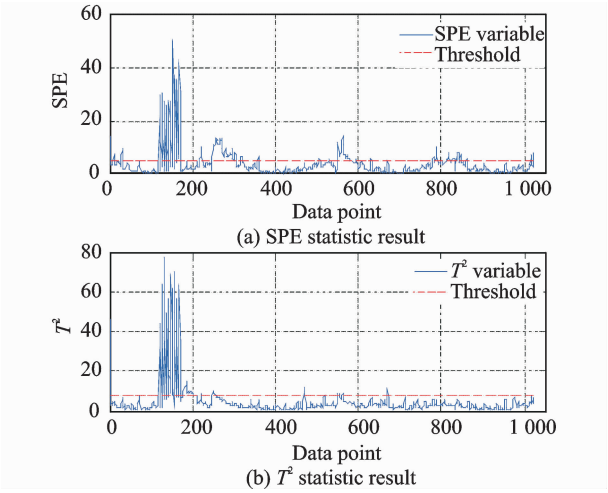

Fig. 7 Statistic detection results for time-varying fault based on LLE algorithm

on LLE does not meet expectations, which can determined from Tables 5, 6. Overall, both I-LLE and LLE can complete fault detection tasks, but I-LLE greatly improves detection accuracy.

4 Conclusions

A fault detection approach based on I-LLE is presented. After a brief introduction of the telemetry data of Satellite TX-I, I-LLE is employed to extract data features and capture the low-dimensional embedding, which leads to the development of fault detection scheme. Finally, fault detection for the TX-I telemetry data based on the I-LLE algorithm is carried out on the embedded low-dimensional space. To illustrate the potential of the proposed scheme, simulations are provided.

However, the parameters of I-LLE, like intrinsic dimensionality, should be set in advance, which will influence the fault detection performance. Therefore, the detailed selection of these parameters is worthy of further study.

Acknowledgement

This work was supported by the Fundamental Research Funds for the Central Universities(No. 2016083).

References:

[1] Hou Q, Cheng Y, Lu N, et al. Study on FDD and FTC of satellite attitude control system based on the effectiveness factor[C]// Systems and Control in Aerospace and Astronautics. China:IEEE, 2008:1-6.

[2] Crowley N L, Apodaca V. Analysis of satellite telemetry data [C] // Aerospace Conference, 1997. USA:IEEE, 1997: 57-67.

[3] Gao Y, Yang T, Xing N, et al. Fault detection and diagnosis for spacecraft using principal component analysis and support vector machines [C] // 2012 7th IEEE Conference on Industrial Electronics and Applications. Singapore:IEEE, 2012:1984-1988.

[4] Fujimaki R, Yairi T, Machida K . An approach to spacecraft anomaly detection problem using Kernel Feature Space [C] // IEEE Aerospace Conference. USA:IEEE, 1997: 57-67.

[5] Yang T, Chen B, Gao Y, et al. Data mining-based fault detection and prediction methods for in-orbit satellite[C] // Measurement, Information and Control (ICMIC), 2013 International Conference on. Egypt: IEEE, 2013:805-808.

[6] Jiang Q, Jia M, Lv J. New approach of intelligent fault diagnosis based on LLE algorithm[C]// Control and Decision Conference, 2008. China:IEEE, 2008: 522-526.

[7] Miao A, Song Z, Zhiqiang G E, et al. Nonlinear fault detection based on locally linear embedding[J]. Journal of Control Theory & Applications, 2013, 11(4): 615-622.

[8] Zhang W, Yang X. Fault detection method based on extended locally linear embedding[C]// Industrial Electronics and Applications, 2006 1st IEEE Conference on. Singapore: IEEE, 2006:1-4.

[9] Kouropteva O, Okun O, Pietikäinen M. Incremental locally linear embedding algorithm[M]// Image Analysis. Germany: Springer Berlin Heidelberg, 2005: 521-530.

[10] Law M H C, Jain A K. Incremental nonlinear dimensionality reduction by manifold learning[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2006, 28(3):377-391.

[11] Shi Lukui, Yang Qingxin, Liu Enhai, et al. An incremental manifold learning algorithm based on the small world model[C]// Final Program and Book of Abstracts of the 2010 International Conference on Life System Modeling and Simulation & 2010 International Conference on Intelligent Computing for Sustainable Energy and Environment. China: LNCS, 2010: 324-332.

[12] Ma Y, Wang M, Shi H. Fault detection for chemical process based on locally linear embedding[J]. Ciesc Journal, 2012,63(7):2121-2127.

[13] Zhang J, Li S Z, Wang J. Nearest manifold approach for face recognition[C]// 2013 10th IEEE Interna-

tional Conference and Workshops on Automatic Face and Gesture Recognition. Korea:[s. n.], 2004: 223-228.

[14] Li B, Zhang Y. Supervised locally linear embedding projection (SLLEP) for machinery fault diagnosis [J]. Mechanical Systems & Signal Processing, 2011, 25(8):3125-3134.

[15] Saul L K, Roweis S T. Think globally, fit locally: Unsupervised learning of nonlinear Manifolds [J]. Journal of Machine Learning Research, 2003, 4(2): 119-155.

[16] Seung H S, Lee D D. The manifold ways of perception [J]. Science, 2000, 290(5500):2268-2269.

[17] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 290(5):2323-2326.

[18] Jiang Q, Jia M, Hu J, et al. Machinery fault diagnosis using supervised manifold learning[J]. Mechanical Systems & Signal Processing, 2009, 23(7): 2301-2311.

[19] Cai L, Tian X, Zhang Y. Dynamic process monitoring based on orthogonal locality preserving projections and exponentially weighted moving average[C]// The 25th Chinese Control and Decision Conference. China: Northeastern University Press, 2013: 4337-4342.

[20] Zhang Ni, Tian Xuemin. Nonlinear dynamic fault detection method based on isometric mapping[J]. Journal of Shanghai Jiaotong University, 2011, 45(8): 1202-1206. (in Chinese)

(Executive Editor: Zhang Bei)