

# Fault Detection Based on Modified t-SNE

Decheng Liu<sup>1</sup>, Tianxu Guo<sup>1</sup>, Maoyin Chen<sup>\*1</sup>

1. Department of Automation, Tsinghua University, Beijing 100084, P. R. China  
 E-mail: mychen@mail.tsinghua.edu.cn

**Abstract:** Dimension reduction is a general step to process high dimensional data for fault detection. Principal component analysis (PCA) divides data space into principal component space and residual space. But it is a global method without considering local geometric properties between data points. Concentrating on local structure of data, manifold learning can be introduced in dynamic and continuous process for fault detection. It can extract latent features of data, and also be viewed as nonlinear dimension reduction. In this paper, we propose a modified t-SNE algorithm for fault detection, simultaneously considering local structure and different scales of variables. Modified t-SNE converts Mahalanobis distance to the conditional probability for representing pairwise similarities instead of Euclidean distance, which satisfies the characteristics of industrial process data. A subspace can be obtained from high-dimension to low-dimension by applying modified t-SNE, which effectively preserves local structure. Simulation on Tennessee Eastman process (TEP) demonstrates the effectiveness of our proposed method.

**Key Words:** dimension reduction, fault detection, local structure, modified t-SNE, mahalanobis distance.

## 1 Introduction

With the great enhancement of intelligent instruments and plant intelligence, modern industrial processes become more complex, generating a large amount of process data [1]. In real-world industrial cases, process data usually contains noise in high-dimensional space [2]. For multivariate statistical process monitoring (MSPM), one of classic methods is principal component analysis (PCA) [3], and dimension reduction can be used to reduce the noise and extract hidden features in process data [4] [5]. However, PCA can only find the global structure in process data, without considering the local structure. PCA divides data space into residual space and principal component space using the correlation of process data [6] [7]. Hotelling's  $T^2$  and squared prediction error (SPE) statistic can be used for fault detection. Based on PCA, there are many improvements for fault detection, such as DPCA [8], KPCA [9], RDTCSA [10] and RCDR [1]. Dynamic PCA (DPCA) uses the time lag shift technique to extend PCA models into dynamic process [8]. Kernel PCA (KPCA) was then proposed to deal with the nonlinearity of data [9]. KPCA can transform the original variables into a high-dimensional feature space, by utilizing a kernel function [9]. Considering dynamic process monitoring, RDTCSA is more sensitive to dynamic changes [10]. Robust characteristic dimensionality reduction (RCDR) was developed and a time-constrained sparse representation for fault detection under strong disturbances was introduced [1]

From information redundancy and geometric viewpoint, manifold learning algorithms can find latent low-dimensional structures embedded in high-dimensional observations, such as MDS [11], Isomap [12], LLE [13], LE [14], and t-SNE [15]. t-SNE is a manifold learning algorithm and can be used for nonlinear dimension reduction. It assumes that data in high-dimension and low-dimension have the same probability distributions. The difference between two probability distributions is represented by information entropy, which Kullback-Leibler divergence can indicate. The metric representing pairwise similarities in t-

SNE is based on Euclidean distance between points. Note that PCA is a linear algorithm and can not explain the complex polynomial relationship of features. In contrast, t-SNE is based on the probability distribution of random walks on the neighborhood graph, and can find the structural relationship. Therefore, t-SNE can preserve both local and global structure of data in process control monitoring.

It is known that Mahalanobis distance can eliminate different dimensions among variables. In this paper, we propose a modified t-SNE algorithm, simultaneously considering local structure and different scales of sensors. Modified t-SNE embeds Mahalanobis distance into conditional probability for representing pairwise similarities, which can satisfy the characteristics of industrial process data. It shows that modified t-SNE obtains good performance of fault detection in Tennessee Eastman process (TEP) [16]. Take fault 5 as an example, the fault detection rate reaches 100% by our proposed method.

The paper is organized as follows. t-SNE is reviewed in Section 2. Then we propose modified t-SNE algorithm based on fault detection in Section 3. Simulation using modified t-SNE is provided in Section 4. And Section 5 is the conclusion of the paper.

## 2 Review of t-distribution stochastic neighbor embedding

t-SNE is able to reveal global and local structures of high-dimensional data [15]. The goal of t-SNE is to minimize the difference between two distributions of all points using the Kullback-Leibler divergence. Data in high-dimension and low-dimension should have the same probability distribution. The pairwise similarities metric is based on Euclidean distance.

For a data set  $X = \{x_1, x_2, \dots, x_n\}$  ( $x_i \in \mathbb{R}^m$ ), t-SNE computes the probability  $p_{ij}$ , which is proportional to pairwise similarity of objects  $x_i$  and  $x_j$ .

In low-dimensional space, t-SNE aims to learn a  $d$ -dimensional map  $Y = \{y_1, y_2, \dots, y_n\}$  ( $y_i \in \mathbb{R}^d$ ,  $d < m$ ). It measures the pairwise similarity  $q_{ij}$  between  $y_i$  and  $y_j$ , using t-distribution.

If  $p_{ij}$  is the same as  $q_{ij}$ , local structures of data can be preserved completely. Motivated by information entropy,

This work was supported by NSFC under Grants No.61873143.

\*Corresponding author.

Kullback-Leibler divergence can be chosen to measure the difference between the probabilities  $p_{ij}$  and  $q_{ij}$  [15].

Then, gradient descent method can be used to optimize the cost function.

### 3 Modified t-SNE

#### 3.1 Mahalanobis distance

In modified t-SNE, Mahalanobis distance is an effective metric for calculating the pairwise similarity. It is suitable to measure the distance between pairs of industrial process data. Firstly, the variance of the measured values from sensors (e.g., speed sensor, pressure sensor) is greatly different. Secondly, variables resulting from different sensors are typically on different scales. For instance, reactor pressure in TEP [16] varies from 2689.9 (kscmh) to 2724.7 (kscmh) in normal condition. Although there is only a slight change in reactor temperature, from 120.34°C to 120.46°C. Euclidean distance and the widely-used normalization by many kinds of fault detection methods neglect the difference among various scales.

In modified t-SNE, we use Mahalanobis distance to calculate the pairwise similarity by considering different scales of sensors. For  $x_k$  and  $x_l$ , Mahalanobis distance can be calculated by:

$$d(x_k, x_l) = \sqrt{(x_k - x_l)^T S^{-1} (x_k - x_l)} \quad (1)$$

where  $S$  is the covariance matrix of data set  $X$ .

In modified t-SNE, conditional probability of pairwise similarity between  $x_k$  and  $x_l$  in high-dimension can be calculated as follows:

$$p_{kl} = \frac{\exp(-d(x_k, x_l)^2)/(2\sigma^2)}{\sum_{g \neq h} \exp(-d(x_h, x_g)^2)/(2\sigma^2)} \quad (2)$$

In low-dimensional space, the t-distribution is used to calculate the pairwise similarity [15]. For low-dimensional data points  $y_k$  and  $y_l$ , conditional probability can be calculated as follows:

$$q_{kl} = \frac{(1 + d(y_k, y_l)^2)^{-1}}{\sum_{g \neq h} (1 + d(y_g, y_h)^2)^{-1}} \quad (3)$$

#### 3.2 Modified t-SNE for fault detection

##### 3.2.1 Compute projection space

Inspired by the projection space [17], low-dimensional data from high dimensional via manifold learning can be considered as a result of linear projection from high-dimensional data. The explanation of projection is illustrated in Fig. 1, where  $X$  is the high-dimensional data and  $Y$  is the low-dimensional data.  $Y$  can be calculated by using manifold learning from high-dimensional data  $X$ . Hence, there exists a linear projection matrix  $A$  from high-dimension to low-dimension.

Consider  $n$  observations under normal conditions in industrial process, each observation has  $m$  measured variables. Let  $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$  be the original training data with  $n$  samples, where  $x_i \in \mathbb{R}^m$  ( $i = 1, 2, \dots, n$ ) denote  $m$  variables measured from sensors at the  $i$ th sampling time. Namely, each sample  $x_i$  is a  $m$ -dimensional vector. Let  $Y = [y_1, y_2, \dots, y_n] \in \mathbb{R}^{l \times n}$  be the data embedded in

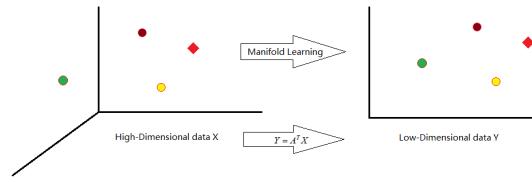


Fig. 1: Illustration of projection based on manifold learning

$n$ -dimensional space, where  $y_i \in \mathbb{R}^l$  ( $i = 1, 2, \dots, n$ ) is the projected point from  $x_i$  in  $l$ -dimensional space.

$$x_i \rightarrow y_i = A^T x_i \quad (4)$$

Based on the hypothesis of potential manifold locality, a linear projection is established to approximate the mapping relationship between high-dimensional data space and low-dimensional embedded space. Note that  $A$  is a  $m \times l$  matrix, namely  $A \in \mathbb{R}^{m \times l}$ . It indicates the projection and can be obtained via linear least squares regression:

$$A = (X X^T)^{-1} X Y^T \quad (5)$$

##### 3.2.2 Calculate $T^2$ statistic

$T^2$  statistic in fault detection is usually calculated using Mahalanobis distance. In modified t-SNE, process data can be mapped from high-dimension to low-dimension, preserving the local structure at the meaning of Mahalanobis distance. For computing  $T^2$  statistic, the projection matrix  $A$  has been solved in (5), and then

$$Y = A^T X \quad (6)$$

where  $X$  is the original training data and  $Y$  is the projected data. From the projection matrix  $A$ ,  $T^2$  statistic can be defined and evaluated for process data  $x_i$  in modified t-SNE as follows:

$$T_i^2 = y_i^T \left( \frac{Y Y^T}{n-1} \right)^{-1} y_i \quad (7)$$

where  $y_i$  is the projection of  $x_i$ .

##### 3.2.3 Calculate control statistic limit

In modified t-SNE algorithm, control statistic limit can be calculated using KDE method [18]. For  $T^2$  statistic ( $T_1^2, T_2^2, \dots, T_n^2$ ) in normal data, the probability density of  $T^2$  statistic is  $f(T^2)$ :

$$\hat{f}_\sigma(T^2) = \frac{1}{n\sigma} \sum_{i=1}^n K\left(\frac{T^2 - T_i^2}{\sigma}\right) \quad (8)$$

where  $K(\cdot)$  is a kernel function and  $\sigma$  is bandwidth. Given a significance level  $\alpha$  (usually  $\alpha = 0.01$  or  $\alpha = 0.05$ ),  $T_{limit}^2$  can be calculated.

##### 3.2.4 Calculate $T^2$ statistic for new data

As stated above, we propose a subspace projection matrix  $A$  in (5). When the new process data  $x_{new}$  in high-dimensional space is measured, it can be projected in the low-dimensional space:

$$y_{new} = A^T x_{new} \quad (9)$$

Then we can calculate  $T^2$  statistic for new data  $x_{new}$  as follows:

$$T_{new}^2 = y_{new}^T \left( \frac{YY^T}{n-1} \right)^{-1} y_{new} \quad (10)$$

If  $T_{new}^2 \leq T_{limit}^2$ ,  $x_{new}$  is a normal data point. Otherwise,  $x_{new}$  is a fault data point.

From the above analysis, the algorithm of modified t-SNE can be briefly formulated by an algorithm.

---

**Algorithm :** Modified t-SNE based on fault detection

---

**Offline training :**

*Step 1* : Collect normal process measurements as a training set.  $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$ .

*Step 2* : Compute Mahalanobis distance  $d(i, j)$  among training data as (1).

*Step 3* : Set low-dimension  $l$ , and minimize the KL divergence between two distributions as (2), (3).

*Step 4* : Compute projection matrix  $A$  in modified t-SNE via (5).

*Step 5* : Calculate control statistic limit  $T_{limit}^2$  via (8).

**Online detection :**

*Step 1* : For test data  $x_{new}$ , calculate  $y_i = A^T x_{new}$  via (9) and statistic  $T_{new}^2$  via (10).

*Step 2* : If  $T_{new}^2 \leq T_{limit}^2$ ,  $x_{new}$  is a normal data point; otherwise,  $x_{new}$  is a fault data point.

---

## 4 Simulation

### 4.1 Tennessee Eastman Process

In this section, we use Tennessee Eastman process (TEP) [16]. TEP is a famous benchmark for fault detection and widely used for process monitoring. In TEP, we choose 33 variables for fault detection, including 22 continuous process measurements and 11 manipulated variables [10].

In TEP, the training data set is measured under normal condition, and contains 500 samples. It is worth noting that only normal data can be known. The simulation generates 21 types of fault data sets. Each type of fault has 960 samples and the fault is introduced at the 161st sampling instant.

### 4.2 Fault detection

We use the same parameters in PCA, t-SNE and modified t-SNE. The dimension of principal component space in PCA is 17 when 0.90 cumulative percent variance (CPV) is used. The significance level is 0.01. Statistic control limit is solved by using KDE method ( $\sigma = 0.3$ ) and other non-parametric methods. Note that fault 1 can be easily detected using PCA, but fault 5 is hard to detect. However, the fault detection rate (FDR) on fault 5 reaches 100% by our proposed method.

Fig. 2 is the detection results of faults 5, 10, 19. Modified t-SNE has higher FDRs. For 21 faults in TEP, the results of FDRs and fault alarm rates (FARs) from PCA, t-SNE and modified t-SNE are listed in Table 1. It shows that modified t-SNE can obtain a better performance of FDRs for all faults.

Table 1: FDRs and FARs using PCA, t-SNE and modified t-SNE (%).

Fault	PCA		t-SNE	modified t-SNE
	$T^2$	SPE		
01	99.25	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>
02	98.38	99.12	98.75	<b>99.62</b>
03	8.12	6.12	10.00	<b>12.50</b>
04	41.62	<b>100.00</b>	67.63	<b>100.00</b>
05	28.75	31.50	<b>100.00</b>	<b>100.00</b>
06	99.50	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>
07	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>
08	97.50	95.62	98.25	<b>98.75</b>
09	6.50	5.12	9.00	<b>10.00</b>
10	48.00	57.75	85.25	<b>89.62</b>
11	52.88	81.50	66.88	<b>84.75</b>
12	98.62	95.12	99.88	<b>100.00</b>
13	94.38	95.50	95.25	<b>95.75</b>
14	99.62	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>
15	9.88	10.50	18.28	<b>19.25</b>
16	33.50	53.00	86.62	<b>91.00</b>
17	81.38	<b>97.00</b>	91.12	96.12
18	90.00	90.50	90.00	<b>90.75</b>
19	17.00	37.25	70.75	<b>85.88</b>
20	45.00	63.75	84.75	<b>89.50</b>
21	43.88	58.25	61.50	<b>67.75</b>
FAR	2.59	3.78	4.23	4.22

### 4.3 Visualization

t-SNE and modified t-SNE can also be used to visualize the process of detection shown in Fig. 3. Take faults 1, 2, 5 as examples, modified t-SNE has a better visualization than t-SNE. This shows Mahalanobis distance could preserve local structure better in low-dimensional space.

## 5 Conclusion

In this paper, we propose a fault detection method based on modified t-SNE algorithm. Concentrating on local structure of data, modified t-SNE can extract latent features of normal data. Thus, a subspace can be obtained from high-dimension to low-dimension. Simulation on TEP demonstrates the effectiveness of our proposed method.

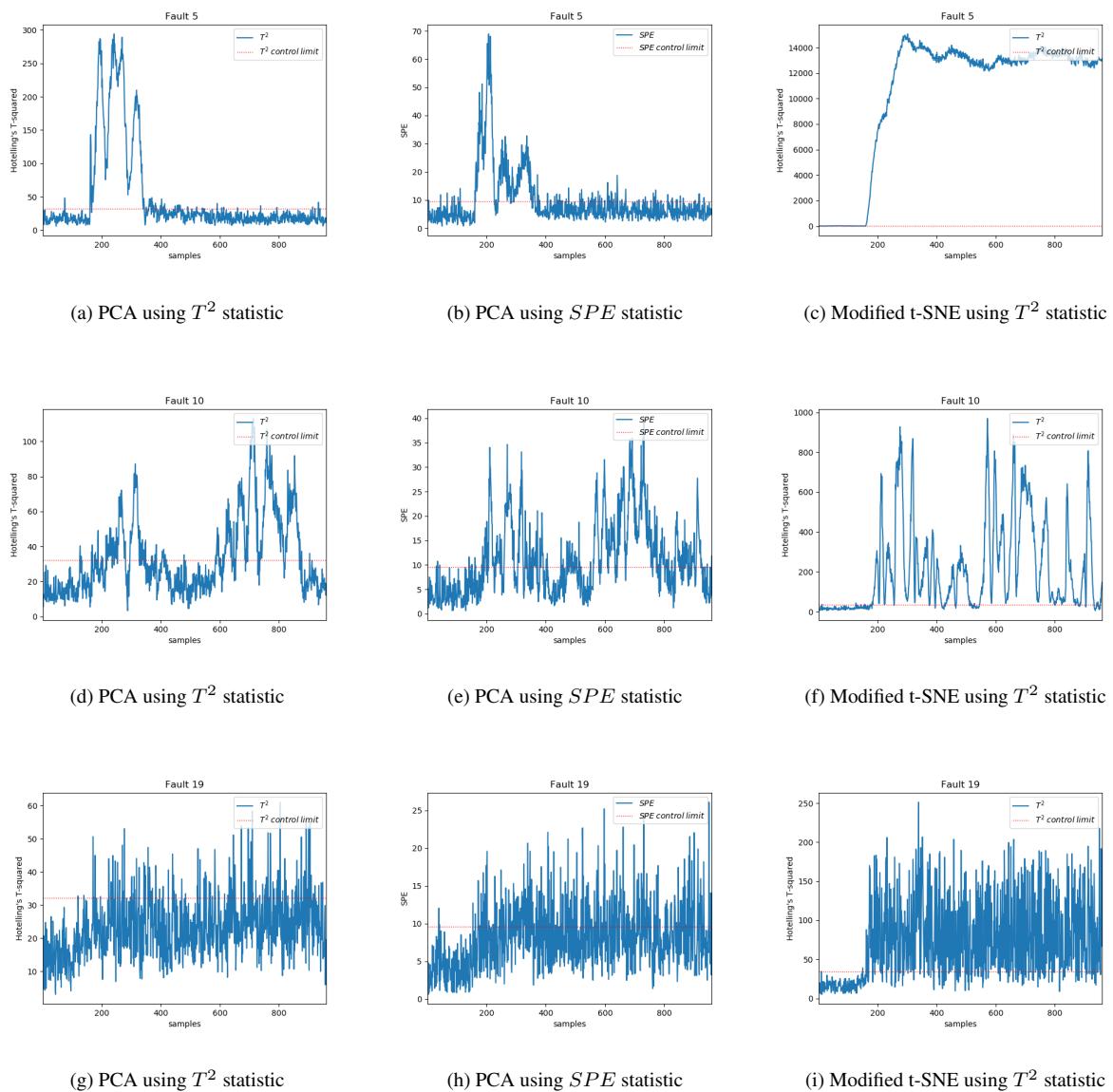


Fig. 2: Fault detection on faults 5, 10, 19 using PCA and modified t-SNE, respectively.

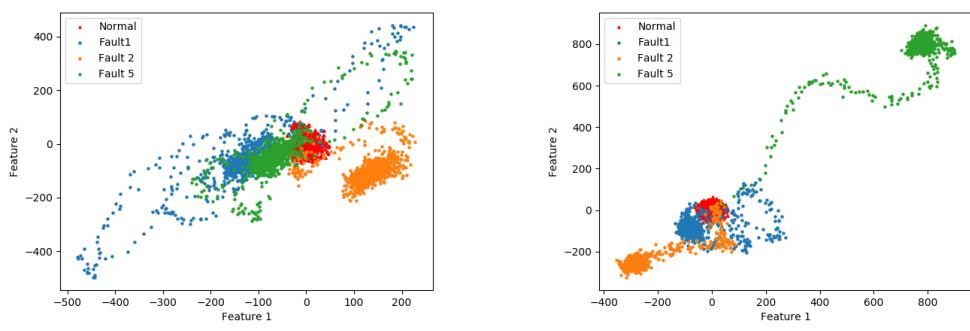


Fig. 3: Feature visualization using t-SNE and modified t-SNE.

## References

- [1] T. Guo, D. Zhou, J. Zhang, M. Chen, and X. Tai, Fault detection based on robust characteristic dimensionality reduction, *Control Engineering Practice*, 84: 125–138, 2019.
- [2] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, A review of process fault detection and diagnosis: Part iii: Process history based methods, *Computers & chemical engineering*, 27(3): 327–346, 2003.
- [3] S. J. Qin, Statistical process monitoring: basics and beyond, *Journal of Chemometrics: A Journal of the Chemometrics Society*, 17(8-9): 480–502, 2003.
- [4] D. Zhou, G. Li, and S. J. Qin, Total projection to latent structures for process monitoring, *AICHE Journal*, 56(1): 168–178, 2010.
- [5] M. Chen and J. Shang, Recursive spectral meta-learner for online combining different fault classifiers, *IEEE Transactions on Automatic Control*, 63(2): 586–593, 2017.
- [6] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, A survey of fault detection, isolation, and reconfiguration methods, *IEEE transactions on control systems technology*, 18(3): 636–653, 2009.
- [7] S. J. Qin, Survey on data-driven industrial process monitoring and diagnosis, *Annual reviews in control*, 36(2): 220–234, 2012.
- [8] W. Ku, R. H. Storer, and C. Georgakis, Disturbance detection and isolation by dynamic principal component analysis, *Chemometrics and intelligent laboratory systems*, 30(1): 179–196, 1995.
- [9] J.M. Lee, C. Yoo, S. W. Choi, P. A. Vanrolleghem, and I.B. Lee, Nonlinear process monitoring using kernel principal component analysis, *Chemical engineering science*, 59(1): 223–234, 2004.
- [10] J. Shang and M. Chen, Recursive dynamic transformed component statistical analysis for fault detection in dynamic processes, *IEEE Transactions on Industrial Electronics*, 65(1): 578–588, 2017.
- [11] I. Borg and P. Groenen, Modern multidimensional scaling: Theory and applications, *Journal of Educational Measurement*, 40(3): 277–280, 2003.
- [12] J. B. Tenenbaum, V. De Silva, and J. C. Langford, A global geometric framework for nonlinear dimensionality reduction, *science*, 290(5500): 2319–2323, 2000.
- [13] S. T. Roweis and L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, *science*, 290(5500): 2323–2326, 2000.
- [14] M. Belkin and P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, *Neural computation*, 15(6): 1373–1396, 2003.
- [15] L. v. d. Maaten and G. Hinton, Visualizing data using t-sne, *Journal of machine learning research*, 9(11): 2579–2605, 2008.
- [16] J. J. Downs and E. F. Vogel, A plant-wide industrial process control problem, *Computers & chemical engineering*, 17(3): 245–255, 1993.
- [17] X. He, D. Cai, S. Yan, and H.-J. Zhang, Neighborhood preserving embedding, in *Tenth IEEE International Conference on Computer Vision (ICCV'05)*, 2005: 1208–1213.
- [18] P. Phaladiganon, S. B. Kim, V. C. Chen, and W. Jiang, Principal component analysis-based control charts for multivariate nonnormal distributions, *Expert Systems with Applications*, 40(8): 3044–3054, 2013.