

Article

Sensors Information Fusion System with Fault Detection Based on Multi-Manifold Regularization Neighborhood Preserving Embedding

Jianping Wu ^{1,2}, Bin Jiang ^{1,2,*} , Hongtian Chen ^{1,2} and Jianwei Liu ^{1,2} 

¹ College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; wujianping@nuaa.edu.cn (J.W.); chtbaylor@163.com (H.C.); ljw301@nuaa.edu.cn (J.L.)

² Jiangsu Key Laboratory of Internet of Things and Control Technologies, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

* Correspondence: binjiang@nuaa.edu.cn; Tel.: +86-025-8489-2305

Received: 28 January 2019; Accepted: 21 March 2019; Published: 23 March 2019



Abstract: Electrical drive systems play an increasingly important role in high-speed trains. The whole system is equipped with sensors that support complicated information fusion, which means the performance around this system ought to be monitored especially during incipient changes. In such situation, it is crucial to distinguish faulty state from observed normal state because of the dire consequences closed-loop faults might bring. In this research, an optimal neighborhood preserving embedding (NPE) method called multi-manifold regularization NPE (MMRNPE) is proposed to detect various faults in an electrical drive sensor information fusion system. By taking locality preserving embedding into account, the proposed methodology extends the united application of Euclidean distance of both designated points and paired points, which guarantees the access to both local and global sensor information. Meanwhile, this structure fuses several manifolds to extract their own features. In addition, parameters are allocated in diverse manifolds to seek an optimal combination of manifolds while entropy of information with parameters is also selected to avoid the overweight of single manifold. Moreover, an experimental test based on the platform was built to validate the MMRNPE approach and demonstrate the effectiveness of the fault detection. Results and observations show that the proposed MMRNPE offers a better fault detection representation in comparison with NPE.

Keywords: sensor information fusion; locality preserving embedding (LPP); multi-manifold regularization neighborhood preserving embedding (MMRNPE); fault detection

1. Introduction

As a foundational component of industrial development, sensors of various types are applied and equipped in diverse systems [1,2], which perfectly meet the demand of data gathering [3,4] and fault detection [5–7]. In addition, the attractions of the information fusion process lie in its ability to eliminate the redundancy and contraction among sensor data sets, as well as its decision-making capacity with uncertainty information. However, once potential information gathered by sensors is not dealt with promptly, some tiny faults may extend unrestrictedly and result in breakdown of systems [7]. Therefore, for the sake of responsiveness, sensors information fusion is an essential technology for improving security and reliability of the system.

Several issues arise when fusing information is introduced to the system, including information uncertainty and data management. An efficient method to deal with the massive data information is feature extraction [8–11]. There are various techniques reported in the literature for fault diagnosis

with the combination of algorithms' fusion. Jafarian et al. [12] used fast Fourier transform as a feature extraction methodology, after which the artificial neural networks, support vector machines and k nearest neighbor classification algorithms are employed to verify the multiple performance metrics and realize signal monitoring. In addition, Saimurugan and Ramprasad [13] fused diverse algorithms for separate purposes to realize fault diagnosis. Wavelet transform and the decision tree were employed for feature extraction, in addition to the artificial neural network to classify the faulty situation. More recently, Liu et al. [14] proposed an intelligent multi-sensor data fusion method with the help of a relevance vector machine for gearboxes' fault detection, and an ant colony optimization algorithm is involved.

However, the proposed compound methods are composed of diverse algorithms, which may result in the complexity of computational system for fault detection. For the purpose of detecting consistently and integrally, there should be some concise algorithms in data preprocessing, or one algorithm to realize the monitoring and detection of the whole system [15–18]. For example, Yunusa-Kaltungo et al. [19] proposed an improved composite spectrum data fusion technique to retain amplitude and phase information by applying cross power spectrum density to fault diagnosis in rotating machines. Moreover, Yunusa-Kaltungo et al. [20] used the data combination method as a preprocessing way for obtaining composite higher order spectra, after which principal component analysis was employed for fault detection. Jing et al. [21] employed deep convolutional neural networks to address an adaptive multi-sensor data fusion problem, which was capable of detecting the conditions of the planetary gearbox effectively with the best diagnosis accuracy. Most of the fusion strategies show their efficiency in small systems, whereas, in practical real-world scenarios, where data generated by sensors might be tremendous, these approaches might lose their advantages and lead to erroneous results.

In many critical fields, the systems are generally confronted with large scale and complicated logic, contributing to a high dimension of data collected from sensors. Under such circumstances, a manifold learning algorithm aiming at dimensionality reduction shows its advantage in data mining [22,23], such as neighborhood preserving embedding (NPE) [24], locality preserving projection (LPP) [25], Laplacian eigenmap (LE) [26], locally linear embedding (LLE) [27] and more [28–31]. It has been proved that the discriminative ability will be enhanced tremendously once the intrinsic manifold structure is considered. To be specific, NPE is a linear technology for combining neighbored data points together to seek an optimal local distribution. Nevertheless, NPE merely concerns the designated points, which means a lack of concern regarding the paired data.

At the same time, motivated by manifold learning algorithms, manifold regularized techniques, which also take local geometric structure into account, are proposed to learn a low-rank approximation [32,33]. For example, in [34], a structure cluster ensemble method is proposed to capture the structure information of the original data set with a manifold regularized objective function. In [35], Chuang et al. employ a manifold regularized distribution adaptation algorithm to classify both multi-spectral and hyper-spectral remote sensing data as well.

Inspired by the aforementioned research, this paper proposes a new algorithm named multi-manifold regularization neighborhood preserving embedding (MMRNPE). Different from the previous local manifold methods based on merely minimizing Euclidean distance between designated point and its neighbors, our framework pays extra attention to paired points in low dimension manifolds, along with the proportion adjustment between designated points and paired ones, which comprises the global information. It is also attractive because the selection of multi-manifold feature will avoid the disturbance and uncertainty of noise. Furthermore, multiple parameters are included for regularization or optimization purposes, which are capable of judging the membership between local and partial global message. In addition, as an iterable algorithm, MMRNPE is capable of choosing the iteration time by required accuracy but with consistent convergence.

The remainder of this paper is organized as follows. In Section 2, a brief review of some preliminaries and related works is given, including our small sensor information fusion system,

as well as the NPE and LPP algorithms. In Section 3, the MMRNPE algorithm is proposed to extract both designated point information and paired points information, together with some discussions on algorithms and parameters. In Section 4, complete experiments based on the fusion system and MMRNPE are presented and verified. Conclusions are drawn in Section 5.

2. Background and Related Theoretical Reviews

2.1. One Small Sensor Information Fusion System

The small sensor information fusion system is actually an electrical drive system in high speed trains. Figure 1 presents the schematic diagram components of this platform [36]. It is an experimental platform of a high-speed train from the China Railway Rolling Stock Corporation in Zhuzhou, China. Several sensors are installed in different parts of traction components. By fusing the sensors and computational modules together with the aid of MATLAB R2014a (Mathworks, Natick, MA, USA) and dSPACE models (2014-A, dSPACE, Paderborn, Germany), the features of unexpected faults can be distinguished and the faults will be detected finally.

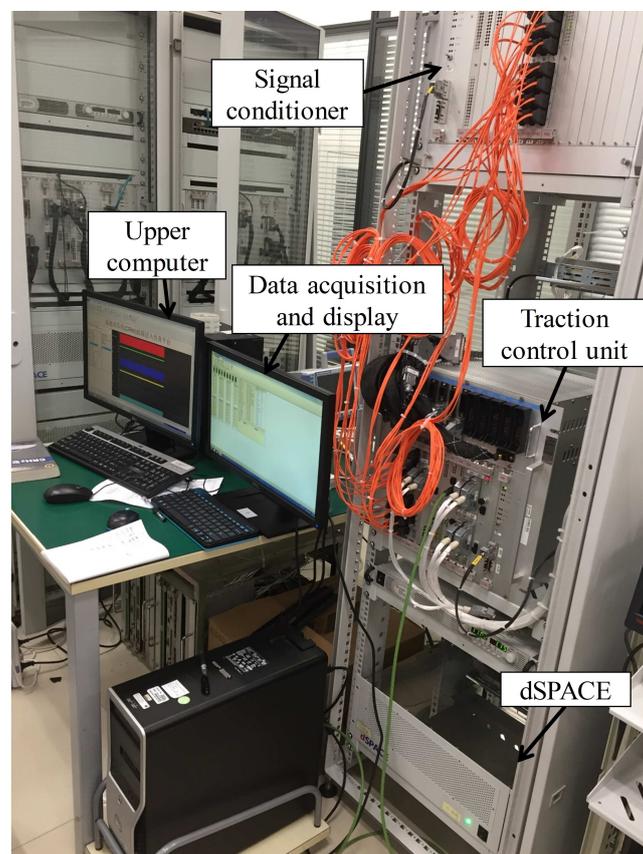


Figure 1. Small sensor information fusion system.

Effective sensor data is crucial to increase reliable detection capacity of this system. The more effective and complete the data set compiled from various sensors, the greater the system's ability to extract features. In our electrical data-driven platform, three-phase output current signals collected from sensors equipped in traction motors are indicated as i_a, i_b, i_c . At the same time, the voltage from the line-side of a transformer is represented as u_{net} while the two-phase input voltage signals of inverters are labeled as u_{d1} and u_{d2} . The rotation signal of traction motor is s . In addition, sensors in traction inverters can also acquire boolean values as the judgement of switch state.

Moreover, information fusion is a fundamental and essential part of sensor management. Since information from a single sensor may be inaccurate and uncertain, not only data collected from

multiple sensors ought to be fused in time, but also the acquired data in the inference and calculation process needs preprocessing. In addition, the following information fusion steps for analysis are based on the newly algorithm MMRNPE, which will be elaborated in Section 3. With the aid of the inference and calculation process, the control and management of the whole system will be realized once the monitoring of fault detection is enabled. The information fusion is, in fact, a data-based method. The process of information fusion structuring is shown in Figure 2.

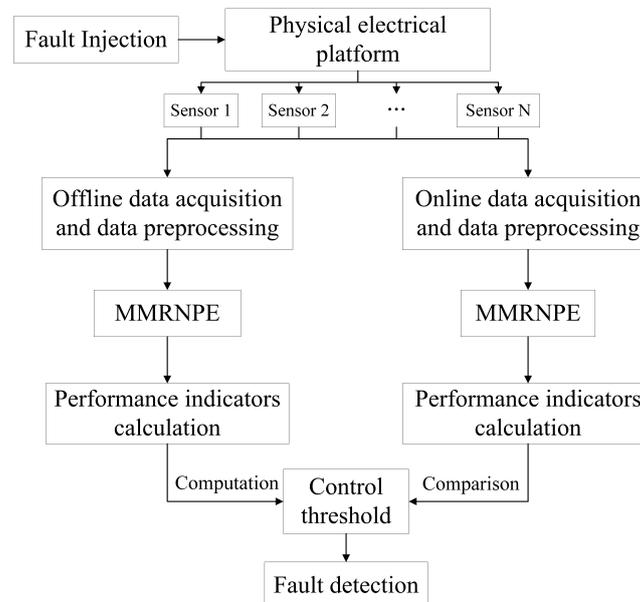


Figure 2. Construction of data-based information fusion structure.

To be specific, the offline data mainly consists of normal data sets, which means the calculated performance indicators represent the normal part. Once faults are injected into this physical electrical platform, the online data will change rapidly with the occurrence of jumping transition in performance indicators. With the aid of one accurate control threshold, the accuracy of faults detection might be guaranteed effectively. This sensor information fusion system is constructed for the purpose of fault detection. It fuses the information from multiple sensors to recognize abnormal signals through the proposed algorithm and make the right decisions accordingly.

2.2. Neighborhood Preserving Embedding

Neighborhood preserving embedding (NPE) is a recently proposed feature extraction method. The basic idea of NPE is to seek a lower dimensional projection of input sensor data set $X = [x_1, x_2, \dots, x_N] \in R^{D \times N}$.

To obtain the optimal projection matrix $A \in R^{N \times D}$, the NPE algorithm first constructs a neighborhood graph and then finds the weight matrix of each edge by minimizing the following reconstruction error [24,27]:

$$\phi(W) = \min \sum_{i=1}^N \|x_i - \sum_{j=1}^N w_{ij}x_j\|^2, \quad (1)$$

where w_{ij} is the neighbored weight from x_i to x_j , with constraints $\sum_{j=1}^N w_{ij} = 1$ and $0 \leq w_{ij} < 1$. After acquiring the basic weight matrix, the minimized cost function with regard to the output matrix $Y \in d \times N (d < N)$ is then chosen as follows:

$$\begin{aligned}\Phi(Y) &= \min_{yy^T=1} \sum_{i=1}^N \left\| y_i - \sum_{j=1}^N w_{ij} y_j \right\|_2^2 \\ &= \min_{yy^T=1} a^T X(I - W)^T (I - W) X^T a \\ &= \min_{yy^T=1} a^T R_{NPE} a,\end{aligned}\quad (2)$$

where $R_{NPE} = X(I - W)^T (I - W) X^T$ and a is the column vector of matrix A . Now that $Y = A^T X$ is a linear projection, NPE is a linear approximate method that contributes to the fast computation ability among massive manifold learning algorithms. In addition, its focus on the relationship between the designated point and its neighbors is widely applied.

2.3. Locality Preserving Projection

The process of calculating the optimal weight matrix and reconstruction error of locality preserving projection (LPP) is the same as NPE. The only difference between these two algorithms is in their eigenmaps, i.e., the objective function of LPP is totally different from that of NPE.

$$\begin{aligned}\Phi(Y) &= \frac{1}{2} \min_{yy^T=1} \sum_{i=1}^N \sum_{j=1}^N (y_i - y_j)^2 w_{ij} \\ &= \frac{1}{2} \min_{yy^T=1} \sum_{i=1}^N \sum_{j=1}^N (a^T x_i - a^T x_j)^2 w_{ij} \\ &= \min_{yy^T=1} \left(\sum_{i=1}^N a^T x_i D_{ii} x_i^T a - \sum_{i=1}^N \sum_{j=1}^N a^T x_i W_{ij} x_j^T a \right) \\ &= \min_{yy^T=1} a^T X(D - W) X^T a \\ &= \min_{yy^T=1} a^T X L_{LPP} X^T a,\end{aligned}\quad (3)$$

where D is a diagonal matrix with entries calculated from the column sum of W , i.e., $D_{ii} = \sum_j W_{ji}$. In addition, $L_{LPP} = D - W$ is a Laplacian matrix.

What makes LPP so attractive is its exploration towards finding another relationship between all paired points, that is, the variance based on the projected data with Euclidean distance. The detailed explanation is given in Section 3.

3. Multi-Manifold Regularization Neighborhood Preserving Embedding

As is mentioned in previous research [24,25], both NPE and LPP are linear manifold learning algorithms concentrating on extracting neighbored connections. In NPE, it reveals that NPE and LPP provide two different ways to linearly approximate the eigenfunction of Laplace Beltrami operator. However, in fact, they have totally different concerns that NPE focuses on the variance about the projected designated point and its reconstructed point while LPP concentrates on the variance about the projected paired points. In other words, NPE is more concerned about single data while LPP cares more about paired data. When it is projected into all data sets, NPE is a local algorithm and LPP is to some extent a global one.

Hence, this proposed multi-manifold regularization neighborhood preserving embedding is developed by combining them together to obtain an overall optimal algorithm. The detailed derivation of this MMRNPE algorithm is presented as below.

In LPP, $L_{LPP} = D - W$, where W is the weight matrix whose element w_{ij} represents the neighborhood relationship between x_i and x_j . However, in this MMRNPE, we replace W by H , where H is the matrix packaging the neighbored information:

$$H_{ij} = \begin{cases} 0, & \text{if } x_i \text{ and } x_j \text{ are neighbors,} \\ 1, & \text{if } x_i \text{ and } x_j \text{ are not neighbors.} \end{cases} \quad (4)$$

Actually, once the data set is given, two kinds of graph Laplacian can be established, including an unsupervised graph in which L is constructed by unlabeled data and a supervised graph in which L is constructed by labeled data. Unsupervised graph Laplacian is defined as $\tilde{L}^{(m)} = D^{(m)} - H^{(m)}$, where m is the number of manifolds. If labeled information is achievable, discriminative information is obtainable for separate samples of different labels. Therefore, a supervised graph Laplacian is constructed as $\tilde{L}^{(m)} = \tilde{L}_{pos}^{(m)} - \beta \tilde{L}_{neg}^{(m)}$, where $\tilde{L}_{pos}^{(m)}$ and $\tilde{L}_{neg}^{(m)}$ represent normal data and faulty data, respectively. In addition, β is a regulating parameter. In this paper, a normalized graph Laplacian is proposed and applied as follows:

$$L_m^{(i)} = D_m^{(i)-\frac{1}{2}} (D_m^{(i)} - H) D_m^{(i)-\frac{1}{2}}, \quad (5)$$

where $D_m^{(i)}$ is the diagonal matrix similar to that in LPP and i here means the i -th manifold with particular setting. Thus, $i = 1, 2, \dots, m$.

However, there exists noise corruption in every manifold of $L_m^{(i)}$, which may result in the failure of exploring intrinsic distribution of samples and then the inaccuracy of fault detection. To avoid the accidental errors caused by a single manifold, one conjoint multi-manifold algorithm is proposed. The core of this multiple method is as follows:

$$\begin{aligned} L &= \sum_{i=1}^m \alpha^{(i)} L_m^{(i)}, \\ \text{s.t. } \sum_{i=1}^m \alpha^{(i)} &= 1, \end{aligned} \quad (6)$$

where $L_m^{(i)}$ are various manifolds that stem from different settings of neighborhood and $\alpha^{(i)}$ are the parameters to match the optimal multi-manifold combination. By expanding the choices of manifolds, the terminal selected L is in $L_{ter} = \text{span}\{L_m^{(1)}, L_m^{(2)}, \dots, L_m^{(m)}\}$.

Actually, the idea of multi-manifold is confined to the introduction of $\alpha^{(i)}$, which is quite remarkable in our algorithm. Obviously, this approach is based on an assumption that the intrinsic manifold exactly lies in the convex hull of all the pre-given manifold candidates and these manifolds are the same as the concept of graph Laplacians. Thus, several manifolds corresponding to diverse neighborhood settings are gathered in the spanning set L_{ter} , which will ensure different features being collected and filtered. In addition, at the same time, the disturbance of noise and uncertainty from a single manifold are diminished.

Taking NPE and the above multi-manifold ideology from LPP into consideration, the objective function is changed:

$$\begin{aligned} \min \quad & \mathbf{a}^T R_{NPE} \mathbf{a} - k \cdot \mathbf{a}^T X L X^T \mathbf{a} \\ = \min \quad & \mathbf{a}^T R_{NPE} \mathbf{a} - k \cdot \sum_{i=1}^m \alpha^{(i)} \mathbf{a}^T X L_m^{(i)} X^T \mathbf{a} \\ \text{s.t. } \quad & \sum_{i=1}^m \alpha^{(i)} = 1, \quad \mathbf{a}^T X X^T \mathbf{a} = 1, \end{aligned} \quad (7)$$

where $L = \sum_{i=1}^m \alpha^{(i)} (D_m^{(i)-\frac{1}{2}} (D_m^{(i)} - H) D_m^{(i)-\frac{1}{2}})$ and k is a regularization parameter.

The parameter k here takes a role to scale the contribution of NPE and LPP. That is to say, although MMRNPE takes both local structure from NPE and the variance structure from LPP into consideration, there is no exact method to measure their own membership. Affiliated with the objected function, k adapts its role of adjusting the proportion of locally neighbored information and partial global variance information, which are critical to the distributive balance.

Use the Lagrange multiplier to solve the minimized problem, and the question is transformed into one generalized eigenvalue problem:

$$[X^T M X - k \cdot X L X^T] \mathbf{a} = \lambda X X^T \mathbf{a}. \quad (8)$$

However, experiments indicate that the above objective function may concentrate on a series of problems with regard to manifold selection. To be specific, $\alpha = [\alpha_1 \alpha_2 \cdots \alpha_m]$ may meet with the following situation:

$$\alpha_i = \begin{cases} 0, & i \neq k, \\ 1, & i = k, \end{cases} \quad (9)$$

where k is the k -th manifold. This means that only the k -th parameter is efficient. In other words, although multiple manifolds are collected, only one particular manifold is chosen. Now that only one single manifold information is preserved and revealed, L_{ter} is not used well. Hence, the objective function needs further improvement and constraint, which are about $\alpha^{(i)}$ to check and balance the decisive leadership of k -th manifold. Therefore, the task is to find a function $f(\alpha^{(i)})$ with negative correlation of $\mathbf{a}^T X L X^T \mathbf{a}$. Entropy of information [37–39] is an efficient method to make adjustment to the contribution on $\alpha^{(i)}$. Thus, another objective function is constructed to obtain $\alpha^{(i)}$:

$$\begin{aligned} \min \quad & \mathbf{a}^T R_{NPE} \mathbf{a} - k \cdot [\mathbf{a}^T X L X^T \mathbf{a} + \gamma \sum_{i=1}^m \alpha_i \log \alpha_i], \\ \text{s.t.} \quad & \sum_{i=1}^m \alpha^{(i)} = 1, \quad \mathbf{a}^T X X^T \mathbf{a} = 1, \end{aligned} \quad (10)$$

where γ is a parameter to adjust the proportion of entropy.

With the aid of the entropy of information about $\alpha^{(i)}$, the condition that one single manifold is emphasized obsessively will be avoided reasonably. With the introduction of γ , the cost function of $\alpha^{(i)}$ is modified. The existence of parameter γ along with the entropy of information is a penalty term of a generalized regularization, which does well in monitoring and regulating the validity of the objective function.

Therefore, the optimized objective function not only considers the local and partly global information, but also fuses multiple manifolds with sensor information. In addition, the added regularized parameters are also significant for selection of an optimal solution.

In a gesture to solve the minimization problem in Equation (10), Lagrange multipliers are introduced for constructing Lagrange function:

$$\begin{aligned} Q = & \text{tr}(\mathbf{a}^T R_{NPE} \mathbf{a}) - k \cdot \sum_{i=1}^m \alpha^{(i)} \text{tr}(\mathbf{a}^T X L^{(i)} X^T \mathbf{a}) \\ & - k \cdot \gamma \cdot \sum_{i=1}^m \alpha_i \log \alpha_i - \lambda_1 (\sum_{i=1}^m \alpha^{(i)} - 1) \\ & - \lambda_2 \text{tr}(\mathbf{a}^T X X^T \mathbf{a} - I), \end{aligned} \quad (11)$$

where λ_1 and λ_2 are Lagrange multipliers.

By setting the derivative of Q w.r.t. $\alpha^{(i)}$ along with λ_1 and λ_2 to zero, we have

$$\begin{aligned} \frac{\partial Q}{\partial \alpha^{(i)}} &= -k \cdot \text{tr}(\mathbf{a}^T X L^{(i)} X^T \mathbf{a}) - k \cdot \gamma \cdot \log \alpha_i - k \cdot \gamma - \lambda_1, \\ \frac{\partial Q}{\partial \lambda_1} &= -\sum_{i=1}^m \alpha^{(i)} + 1, \\ \frac{\partial Q}{\partial \lambda_2} &= -\text{tr}(\mathbf{a}^T X X^T \mathbf{a} - I), \end{aligned} \tag{12}$$

so that we obtain $\alpha^{(i)}$ as follows:

$$\alpha^{(i)} = \frac{\exp\left(\frac{-\mathbf{a}^T X L^{(i)} X \mathbf{a} - \gamma}{\gamma}\right)}{\sum_{i=1}^m \exp\left(\frac{-\mathbf{a}^T X L^{(i)} X \mathbf{a} - \gamma}{\gamma}\right)}. \tag{13}$$

Now that $\alpha^{(i)}$ is deeply relevant to \mathbf{a} , the variable $\alpha_1^{(i)}$ is initialized with a constant value, such as $\alpha_1^{(i)} = (1/m)$. Given $\alpha_1^{(i)}$, \mathbf{a}_1 can be calculated from Equation (8). Then, a reliable $\alpha_2^{(i)}$ will be obtained from Equation (13). With the updated $\alpha_2^{(i)}$, L is completely new, and so are the manifolds. Thus, the eigenvectors of Equation (8) constitute \mathbf{a}_2 , which are the column vectors of projection matrix A . It is remarkable that the manifolds constructed by $\alpha_2^{(i)}$ are the expected results which are sensitive to suppressing the noise disturbance.

In accordance with the above analysis, the algorithmic procedure of the proposed MMRNPE can be formally summarized as below:

1. Compute the normalized graph Laplacians $L_m^{(i)}$ of different manifolds with Equation (5).
2. Compute the initial L_1 with pre-given manifold candidates:

$$L_1 = \sum_{i=1}^m \alpha_1^{(i)} L_m^{(i)}, \quad \alpha_1^{(i)} = \frac{1}{m}, \tag{14}$$

where $i = 1, 2, \dots, m$.

3. Solve the generalized eigenvectors of the following equation as \mathbf{a}_1 :

$$[X M X^T - k(X L_1 X^T)] \mathbf{a}_1 = \lambda X X^T \mathbf{a}_1, \tag{15}$$

where $M = (I - W)^T(I - W)$.

4. Compute L_2 with a series of optimized $\alpha_2^{(i)}$.

$$\begin{aligned} \alpha_2^{(i)} &= \frac{\exp\left(\frac{-\mathbf{a}_1^T X L^{(i)} X^T \mathbf{a}_1 - \gamma}{\gamma}\right)}{\sum_{i=1}^m \exp\left(\frac{-\mathbf{a}_1^T X L^{(i)} X^T \mathbf{a}_1 - \gamma}{\gamma}\right)}, \\ L_2 &= \sum_{i=1}^m \alpha_2^{(i)} L_m^i. \end{aligned} \tag{16}$$

5. Solve the generalized eigenvectors of the following equation as \mathbf{a}_2 :

$$[X M X^T - k(X L_2 X^T)] \mathbf{a}_2 = \lambda X X^T \mathbf{a}_2. \tag{17}$$

6. Obtain the embedding as follows:

$$x_i \rightarrow y_i = A^T x_i, \quad A = [\mathbf{a}_2^{(1)}, \mathbf{a}_2^{(2)}, \dots, \mathbf{a}_2^{(d)}]. \tag{18}$$

Furthermore, the calculation process of our framework is an iterable one with the verification of regularization consistency shown in Figure 3.

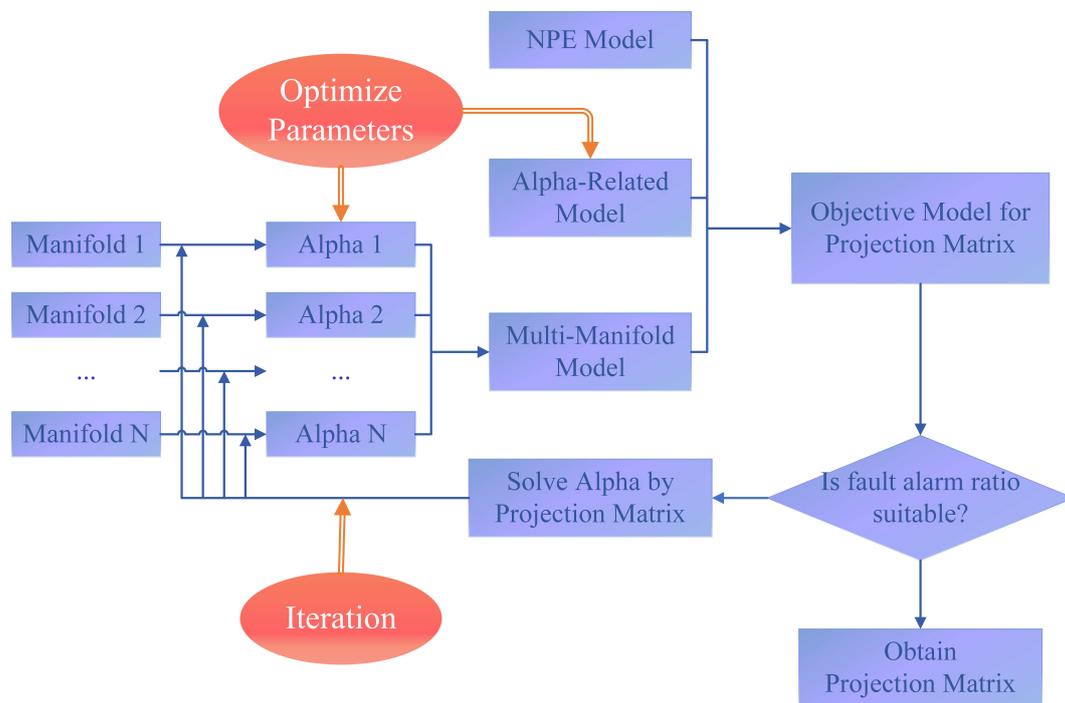


Figure 3. Iterable structure for MMRNPE.

The core of this process lies in the update of $\alpha^{(i)}$. More specifically, the iteration round will continue when a_2 in Equation (18) is substituted to Equation (13) and then $\alpha_3^{(i)}$ is obtained. After substituting $\alpha_3^{(i)}$ to Equation (8), it is easier to observe the projection matrix A. Such alternating iteration round is continuous and the convergency of this learning algorithm can always be guaranteed. Moreover, along with the increasing number of iterations, the algorithm is theoretically deeper, i.e., researchers may choose the iteration times freely and obtain the designing curve based on the requirement of experiment error plenarily.

The iteration procedure is an optimal process of parameter $\alpha^{(i)}$ as well.

In general, MMRNPE is proved to be a successful method in exploiting the underlying geometry structure of selected data sets with the aid of NPE and LPP. Actually, by incorporating the ideology of LPP or graph Laplacians, MMRNPE takes advantage of NPE's local structure and LPP's partial global variance structure successfully. It is worthwhile to highlight the marvelous properties of this proposed approach:

- In our MMRNPE, it takes the Euclidean distance of both designated point and paired points into adequate consideration, which guarantees the balance between local and global information from sensor data.
- Multiple parameters are included in this algorithm, some of which are of regularization purpose and the others are of limitation consideration. Now that various choices of parameters will result in distinction of performance, some optimization algorithms can be chosen to promise the fault detection rate.
- Some of the regularized parameters are able to judge the membership relationship of elements, i.e., the membership between local information with NPE algorithm and variance information with LPP algorithm is displayed intuitively, which realizes the sensor information fusion.

4. Experiments

4.1. Fault Detection Strategy

The fault detection is based on the small sensor information fusion system referred to in Section 2.1. The detailed parameters information is given in Table 1 [7], which includes parameters both in physical space and in computational space.

Table 1. The parameters for the platform.

Symbol	Quantity	Value (Unit)
T_s	sampling time	4×10^{-4} (s)
p	pole pairs	2 (1)
U_s	voltage of charge source	400 (V)
R_s	resistance in stator side	0.228 (Ω)
R_r	resistance in rotor side	0.1267 (Ω)
L_s	inductance in stator side	0.0281 (H)
L_r	inductance in rotor side	0.0280 (H)
L_{msr}	mutual inductance of motor	0.0268 (H)
L_{ls}	leakage inductance in stator side	0.0013 (H)
L_{lr}	leakage inductance in rotor side	0.0013 (H)
U_d	intermediate voltage	3300 (V)
C_d	capacitor of direct current link	8×10^{-3} (F)
J	rotary inertia	100 ($\text{kg}\cdot\text{m}^2$)
L	filter inductance	0.42×10^{-3} (H)
C	filter capacitor	6×10^{-3} (F)

Actually, data sets collected from sensors in physical space are sent to the computational space promptly, after which the calculation and monitoring process begins. The calculation and monitoring process includes two steps—offline calculation and online monitoring.

Now that MMRNPE seeks a latent variable space that represents the high dimensional space relatively, the monitoring statistics Hotelling's T^2 statistic is constructed as a measurement of the performance of fault detection [40,41]. Here is the definition of T^2 :

$$\begin{aligned}
 T^2 &= X^T A \Lambda^{-1} A^T X \\
 &= X^T A (\text{cov}(Y_{\text{offline}}))^{-1} A^T X \\
 &= Y^T (\text{cov}(Y_{\text{offline}}))^{-1} Y \\
 &= \frac{1}{n-1} Y^T Y_{\text{offline}}^T Y_{\text{offline}} Y,
 \end{aligned} \tag{19}$$

where $\Lambda = \text{cov}(Y_{\text{offline}})$ is the covariance matrix of the offline data set and Y is the sample after dimensional reduction process. It is obvious that with sampling n , T^2 is the sum of the normalized squared scores, making it possible to measure the performance of the chosen projection matrix A [42,43]. Another statistic SPE plays this role as well [40,41].

With the performance statistics obtained, the offline modeling procedure is as follows:

1. Collect original data set X and normalize it with zero mean and unit variance.
2. Compute the projection matrix with the proposed MMRNPE algorithm.
3. Calculate the dimensional-reduction data set Y with the linear mapping.
4. Compute the performance statistics T^2 and SPE of offline data set.
5. Construct the upper control limits of T^2 and SPE as the standard of online data.

After the offline modeling process, the upper control limits are obtained, and then we can implement the online monitoring procedure:

1. Collect online data set X_{online} and normalize it with zero mean and unit variance.
2. Calculate the dimensional-reduction data set Y with the projection matrix obtained in offline procedure.
3. Compute the performance statistics T^2 and SPE of the online data set and compare them with the upper control limits of T^2 and SPE from the offline process.
4. Compute the fault alarm ratio (FAR), non-detection ratio (NDR) and total detection rate (TDR) to evaluate the fault detection ability of this MMRNPE algorithm.

4.2. Experiments Verification with the Proposed MMRNPE

The platform with multiple sensors to fuse information derived in Section 2.1 is selected to verify this multi-manifold algorithm.

Several experiments are carried out in the normal state and various bias faulty states on the test bench. Both the normal and faulty expressions are listed in Table 2.

Table 2. The parameters for concerned states.

Notations	Fault Description	Expression	Sample Number
N	normal	normal	250
F_1	current sensor fault	$f_1 = 0.05 \text{ A}$	8695
F_2	voltage sensor fault	$f_2 = 2 \text{ V}$	8695
F_3	speed sensor fault	$f_3 = 0.5 \text{ rad/s}$	8695

Numerous sensors are equipped and distributed across all aspects and locations of the system. Several typical signals attract much attention to evaluate the performances of motors, including current signals, voltage signals and speed signals. Hence, we locate the sensor faults at current path and voltage path, as well as the speed sensor itself separately. As is obviously shown in Table 2, three types of faults are manually induced in various sensors under different operation conditions, i.e., current sensor fault, voltage sensor fault and speed sensor fault.

Once sensors located in other locations break down and affect the security of system operation, the typical signals mentioned above also changes abnormally. In a gesture to evaluate the severities of faults, three different ranks of signal amplitude are set. The degree of voltage sensor fault is 0.01% of the running voltage amplitude while that of current sensor fault is 0.05%. In addition, the speed sensor fault is 0.5% of the normal condition. It should also be noted that the training data of three different faults share the same normal data sets. The only difference lies in the test data, where samplings collected during the faulty operation periods have the same length. Figures 4–6 give the evolution processes of fault injections with different sensor faults.

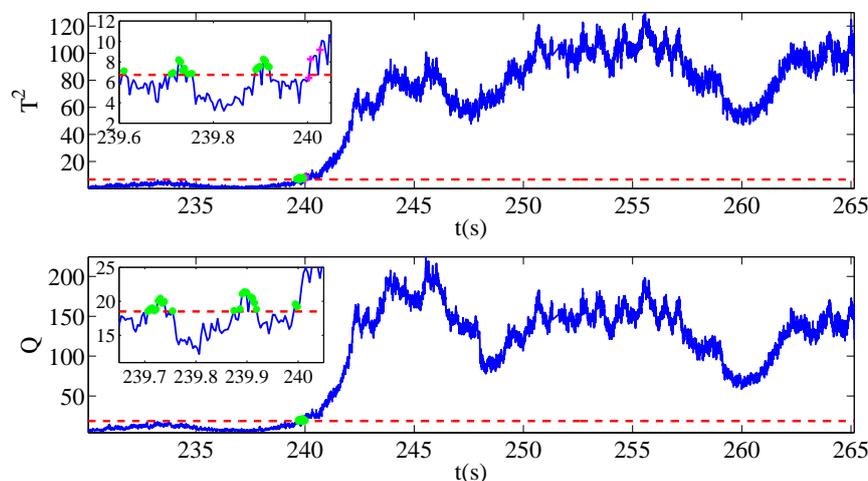


Figure 4. Fault detection results for f_1 .

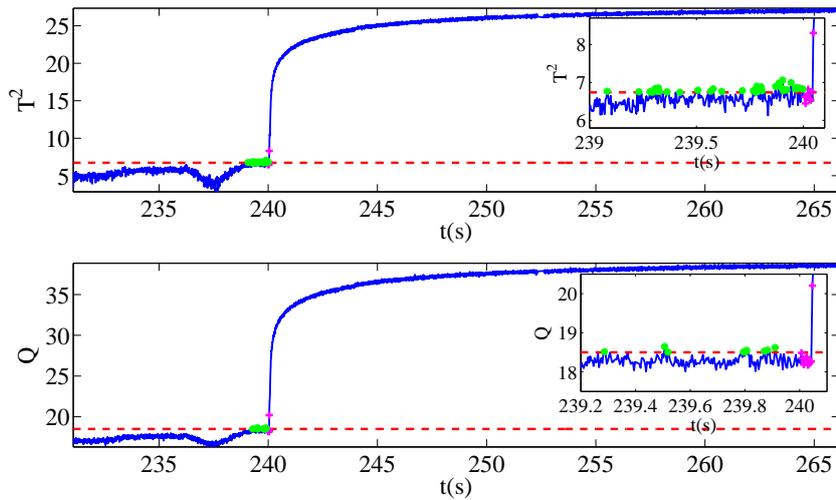


Figure 5. Fault detection results for f_2 .

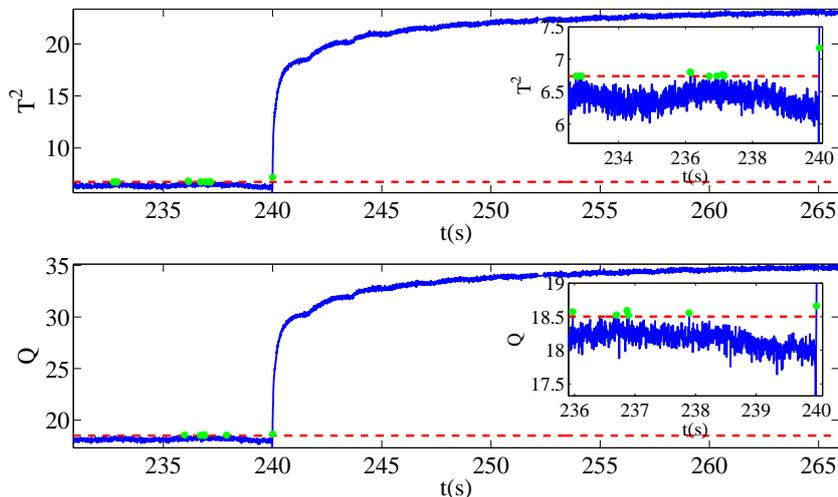


Figure 6. Fault detection results for f_3 .

The sensor faults are injected into the platform at 240 s. With the aid of this MMRNPE structure proposed for fault detection, three faults are detected successfully in our experiments.

The parameters set in this experiment are chosen as below. Three manifolds are constructed with $\alpha_1^{(1)} = \alpha_1^{(2)} = \alpha_1^{(3)} = \frac{1}{2}$. By choosing $k = 8$ and $\gamma = 5689.9$, $\alpha_2^{(i)}$ can be calculated that $\alpha_2^{(1)} = 0.3826$, $\alpha_2^{(2)} = 0.3375$, and $\alpha_2^{(3)} = 0.2799$.

As is shown in Figures 4–6, there are some fault alarming points which are below control limits after faults occur and non-detection points that are above control limits before faults happen. Both fault alarming points and non-detection points are marked in the figures with striking colors. At the same time, fault alarming ratio (FAR), non-detection ratio (NDR) and total detection ratio (TDR) with performance statistics T^2 and SPE are calculated under both MMRNPE and NPE algorithms, which are shown in Table 3.

Table 3. The detection indexes of various sensor faults with MMRNPE and NPE.

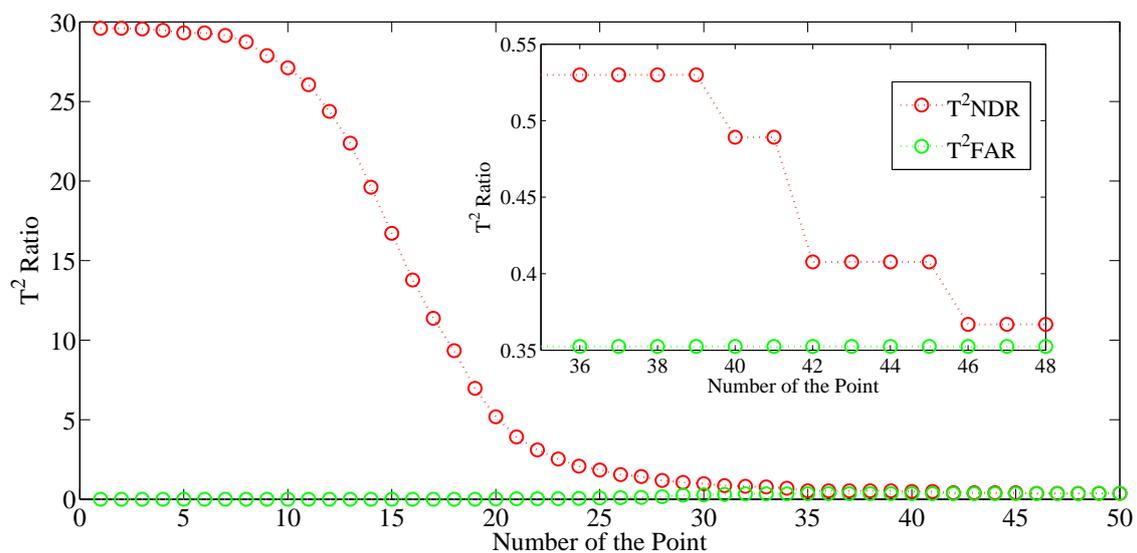
Algorithm		MMRNPE			NPE		
Sensor Fault	Statistics	FAR(%)	NDR(%)	TDR(%)	FAR(%)	NDR(%)	TDR(%)
f_1	SPE	0.8464	0	0.0024	0.8464	0	0.0024
	T^2	0.6449	0.0322	0.0021	0	12.3270	0.0881
f_2	SPE	0.3136	0.1393	0.0021	0.4032	0.1547	0.0022
	T^2	1.4337	0.1238	0.0048	11.0663	0.0464	0.0288
f_3	SPE	0.1754	0	0.000345	0.2630	0	0.0069
	T^2	0.2630	0	0.0010	0.3725	0.0156	0.0015

The detection indexes of various sensor faults with MMRNPE and NPE shown in Table 3 can obviously verify the superiority of our MMRNPE, especially when it comes to statistics T^2 . The results of FAR, NDR and TDR with performance statistics T^2 perform better accuracy of fault detection. Furthermore, with a careful observation of the magnified figures, it is apparent that several misclassifications occur. To be more specific, the non-detection point may be marked in the fault alarming notation after the fault is injected while the fault alarming point may be marked in a non-detection notation before 240 s. Such mistakes take place due to the setting of sampling time. The sampling time T_s is 4×10^{-4} s, which will make it difficult to inject the sensor faults exactly at 240 s, i.e., the injected time 240 s is between sampling points m and $m + 1$.

Several parameters are induced into this MMRNPE algorithm and each of them plays its own role. Here, we will discuss the role of γ to select an optimal one. By using different γ , the monitoring and detection results of f_1 with MMRNPE are totally different, which are shown in Figure 7.

As is shown in Figure 7, the FAR of T^2 increases very slightly while NDR decreases like an inverse sigmoid curve with the gradual increase of γ .

The x-axis of Figure 7 is the number of γ , where the series of γ is actually a geometrical sequence with the initial value is 1 and the terminal value is 10^4 .

**Figure 7.** Fault detection results for f_3 .

5. Conclusions

Multiple sensors located at various positions of this electrical drive system are comprised of numerous amounts of characteristic information. In this research, we discussed the excavation of sensor data sets obtained from sensors information fusion systems to detect faults via an adapted

MMRNPE algorithm. There are three key components of this improved algorithm. Firstly, as a combination of NPE and LPP, the objective function of MMRNPE inspired by manifold learning algorithms considers both the designated points and paired points to find the intrinsic incorporation. Secondly, multiple manifolds with various neighborhood points are merged together closely while keeping their distinct characteristics. Thirdly, diverse parameters are introduced and presented into this methodology to play their own role, some of which bear the responsibility of adjusting the proportion of locally neighbored information and partial global variance information. In addition, some of the parameters are used for weighting adjustments with different manifolds. The experimental results demonstrate that MMRNPE realizes data processing and information fusion successfully in terms of its sufficient fault detection effects. With three different sensor faults injected and detected promptly and efficiently, this approach is verified and confirmed adequately.

In this study, parameters are selected in an enumeration or experimental way, which means future work will focus on the optimization algorithms for our strategy. By taking the optimization algorithms as well as iteration into account, the detection efficiency will be promoted to a higher level.

Author Contributions: Data curation, J.W. and H.C.; Funding acquisition, B.J.; Methodology, J.W.; Project administration, B.J.; Resources, B.J. and J.L.; Software, J.W. and H.C.; Validation, J.W.; Writing—original draft, J.W.; Writing—review and editing, B.J., H.C. and J.L.

Funding: This research was supported in part by the National Natural Science Foundation of China under Grant 61490703 and Grant 61533008, in part by the Open Fund for Postgraduate Innovation Laboratory of Nanjing University of Aeronautics and Astronautics under Grant kfj20180320, and in part by A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors have no role in the design of the study; in the collection, analysis or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

NPE	Neighborhood Preserving Embedding
MMRNPE	Multi-Manifold Regularization Neighborhood Preserving Embedding
LPP	Locality Preserving Embedding
LE	Laplacian Eigenmap
LLE	Locally Linear Embedding
SPE	Square Prediction Error
FAR	Fault Alarm Ratio
NDR	Non-Detection Ratio
TDR	Total Detection Ratio

References

1. Vitola, J.; Pozo, F.; Tibaduiza, D.A.; Anaya, M. A sensor data fusion system based on k-nearest neighbor pattern classification for structural health monitoring applications. *Sensors* **2017**, *17*, 417. [[CrossRef](#)] [[PubMed](#)]
2. Garramiola, F.; del Olmo, J.; Poza, J.; Madina, P.; Almandoz, G. Integral sensor fault detection and isolation for railway traction drive. *Sensors* **2018**, *18*, 1543. [[CrossRef](#)] [[PubMed](#)]
3. Zhang, Y.; He, S.; Chen, J. Data gathering optimization by dynamic sensing and routing in rechargeable sensor networks. *IEEE/ACM Trans. Netw.* **2016**, *24*, 1632–1646. [[CrossRef](#)]
4. Chen, H.; Jiang, B.; Chen, W.; Yi, H. Data-driven detection and diagnosis of incipient faults in electrical drives of high-speed trains. *IEEE Trans. Ind. Electron.* **2019**, *66*, 4716–4725. [[CrossRef](#)]
5. Jlassi, I.; Estima, J.O.; El Khil, S.K.; Bellaaj, N.M.; Cardoso, A.J.M. A robust observer-based method for IGBTs and current sensors fault diagnosis in voltage-source inverters of PMSM drives. *IEEE Trans. Ind. Appl.* **2017**, *53*, 2894–2905. [[CrossRef](#)]
6. Chen, H.; Jiang, B.; Lu, N. A newly robust fault detection and diagnosis method for high-speed trains. *IEEE Trans. Intell. Transp. Syst.* **2018**. [[CrossRef](#)]

7. Chen, H.; Jiang, B. A review of fault detection and diagnosis for the traction system in high-speed trains. *IEEE Trans. Intell. Transp. Syst.* **2019**. [[CrossRef](#)]
8. Yunusa-Kaltungo, A.; Sinha, J.K. Faults diagnosis in rotating machines using higher order spectra. In Proceedings of the ASME Turbo Expo 2014: Turbine Technical Conference and Exposition, Düsseldorf, Germany, 16–20 June 2014; p. V07AT31A002.
9. Ehlenbröker, J.F.; Mönks, U.; Lohweg, V. Sensor defect detection in multisensor information fusion. *J. Sens. Sens. Syst.* **2016**, *5*, 337–353. [[CrossRef](#)]
10. Najjar, N.; Gupta, S.; Hare, J.; Kandil, S.; Walthall, R. Optimal sensor selection and fusion for heat exchanger fouling diagnosis in aerospace systems. *IEEE Sens. J.* **2016**, *16*, 4866–4881. [[CrossRef](#)]
11. Yunusa-Kaltungo, A.; Sinha, J.K. Generic vibration-based faults identification approach for identical rotating machines installed on different foundations, VIRM 11-Vibrations in Rotating. *Machinery* **2016**, *11*, 499–510.
12. Jafarian, K.; Mobin, M.; Jafari-Marandi, R.; Rabiei, E. Misfire and valve clearance faults detection in the combustion engines based on a multi-sensor vibration signal monitoring. *Measurement* **2018**, *128*, 527–536. [[CrossRef](#)]
13. Saimurugan, M.; Ramprasad, R. A dual sensor signal fusion approach for detection of faults in rotating machines. *J. Vib. Control* **2018**, *24*, 2621–2630. [[CrossRef](#)]
14. Liu, Z.; Guo, W.; Tang, Z.; Chen, Y. Multi-sensor data fusion using a relevance vector machine based on an ant colony for gearbox fault detection. *Sensors* **2015**, *15*, 21857–21875. [[CrossRef](#)]
15. Irhoumah, M.; Pusca, R.; Lefevre, E.; Mercier, D.; Romary, R.; Demian, C. Information fusion with belief functions for detection of interturn short-circuit faults in electrical machines using external flux sensors. *IEEE Trans. Ind. Electron.* **2018**, *65*, 2642–2652. [[CrossRef](#)]
16. Luwei, K.C.; Yunusa-Kaltungo, A.; Sha’aban, Y.A. Integrated Fault Detection Framework for Classifying Rotating Machine Faults Using Frequency Domain Data Fusion and Artificial Neural Networks. *Machines* **2018**, *6*, 59. [[CrossRef](#)]
17. Luwei, K.C.; Sinha, J.K.; Yunusa-Kaltungo, A.; Elbhah, K. Data fusion of acceleration and velocity features (dFAVF) approach for fault diagnosis in rotating machines. *MATEC Web Conf.* **2018**, *211*, 21005. [[CrossRef](#)]
18. Rizal, M.; Ghani, J.A.; Nuawi, M.Z.; Haron, C.H. Cutting tool wear classification and detection using multi-sensor signals and Mahalanobis-Taguchi System. *Wear* **2017**, *15*, 1759–1765. [[CrossRef](#)]
19. Yunusa-Kaltungo, A.; Sinha, J.K.; Elbhah, K. An improved data fusion technique for faults diagnosis in rotating machines. *Measurement* **2014**, *58*, 27–32. [[CrossRef](#)]
20. Yunusa-Kaltungo, A.; Sinha, J.K.; Nembhard, A.D. A novel fault diagnosis technique for enhancing maintenance and reliability of rotating machines. *Struct. Health Monit.* **2015**, *14*, 604–621. [[CrossRef](#)]
21. Jing, L.; Wang, T.; Zhao, M.; Wang, P. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors* **2017**, *17*, 414. [[CrossRef](#)] [[PubMed](#)]
22. Ge, Z.; Song, Z.; Ding, S.X.; Huang, B. Data mining and analytics in the process industry: The role of machine learning. *IEEE Access* **2017**, *5*, 20590–20616. [[CrossRef](#)]
23. Harandi, M.; Salzmänn, M.; Hartley, R. Dimensionality reduction on SPD manifolds: The emergence of geometry-aware methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 48–62. [[CrossRef](#)]
24. He, X.; Cai, D.; Yan, S.; Zhang, H. Neighborhood preserving embedding. In Proceedings of the 10th IEEE International Conference on Computer Vision, Beijing, China, 17–21 October 2005; pp. 1208–1213.
25. He, X.; Niyogi, P. Locality preserving projections. *Adv. Neural Inf. Process. Syst.* **2004**, *16*, 153–160.
26. Belkin, M.; Niyogi, P. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Adv. Neural Inf. Process. Syst.* **2002**, *10*, 585–591.
27. Roweis, S.T.; Saul, L.K. Nonlinear dimensionality reduction by locally linear embedding. *Science* **2000**, *290*, 2323–2326. [[CrossRef](#)] [[PubMed](#)]
28. He, X.; Yan, S.; Hu, Y.; Niyogi, P.; Zhang, H.J. Face recognition using laplacianfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 328–340. [[PubMed](#)]
29. Belhumeur, P.N.; Hespanha, J.P.; Kriegman, D.J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1997**, *19*, 711–720. [[CrossRef](#)]

30. Jolliffe, I. Principal Component Analysis. In *International Encyclopedia of Statistical Science*; Miodrag Lovric; Springer: Berlin/Heidelberg, Germany, 2011; pp. 1094–1096.
31. Mika, S.; Ratsch, G.; Weston, J.; Scholkopf, B.; Mullers, K.R. Fisher discriminant analysis with kernels. In Proceedings of the 1999 IEEE Signal Processing Society Workshop, Madison, WI, USA, 25 August 1999; pp. 41–48.
32. Belkin, M.; Niyogi, P.; Sindhvani, V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *J. Mach. Learn. Res.* **2006**, *10*, 2399–2434.
33. Li, X.; Ng, M.K.; Cong, G.; Ye, Y.; Wu, Q. MR-NTD: Manifold regularization nonnegative tucker decomposition for tensor data dimension reduction and representation. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 1787–1800. [[CrossRef](#)] [[PubMed](#)]
34. Li, X.; Lu, Q.; Dong, Y.; Tao, D. SCE: A manifold regularized set-covering method for data partitioning. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 1760–1773. [[CrossRef](#)] [[PubMed](#)]
35. Luo, C.; Ma, L. Manifold regularized distribution adaptation for classification of remote sensing images. *IEEE Access* **2018**, *6*, 4697–4708. [[CrossRef](#)]
36. Chen, H.; Jiang, B.; Ding, S.X.; Lu, N.; Chen, W. Probability-relevant incipient fault detection and diagnosis methodology with applications to electric drive systems. *IEEE Trans. Control Syst. Technol.* **2018**. [[CrossRef](#)]
37. Jiang, Q.; Shen, Y.; Li, H.; Xu, F. New fault recognition method for rotary machinery based on information entropy and a probabilistic neural network. *Sensors* **2018**, *18*, 337. [[CrossRef](#)] [[PubMed](#)]
38. Gao, Y.; Vilecco, F.; Li, M.; Song, W. Multi-Scale permutation entropy based on improved LMD and HMM for rolling bearing diagnosis. *Entropy* **2017**, *19*, 176. [[CrossRef](#)]
39. Sawalhi, N.; Randall, R.B.; Endo, H. The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. *Mech. Syst. Signal Process.* **2007**, *21*, 2616–2633. [[CrossRef](#)]
40. Luo, H.; Yang, X.; Krueger, M.; Ding, S.X.; Peng, K. A plug-and-play monitoring and control architecture for disturbance compensation in rolling mills. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 200–210. [[CrossRef](#)]
41. He, W.; He, Y.; Zhang, C. A new fault diagnosis approach for analog circuits based on spectrum image and feature weighted kernel Fisher discriminant analysis. *Rev. Sci. Instrum.* **2018**, *89*, 074702. [[CrossRef](#)] [[PubMed](#)]
42. Jiang, Q.; Yan, X. Parallel PCA–KPCA for nonlinear process monitoring. *Control Eng. Pract.* **2018**, *80*, 17–25. [[CrossRef](#)]
43. [[CrossRef](#)] Tian, E.; Wang, Z.; Zou, L.; Yue, D. Probabilistic-constrained filtering for a class of nonlinear systems with improved static event-triggered communication. *Int. J. Robust Nonlinear Control.* **2018**, *29*, 1484–1498, doi:10.1002/rnc.4447.

