electronics

MDPI

Article

Nonintrusive Load Monitoring Based on Complementary Features of Spurious Emissions

Donglin Su *, Qian Shi, Hui Xu and Wang Wang

School of Electronic and Information Engineering, Beihang University, Beijing 100191, China
* Correspondence: sdl@buaa.edu.cn; Tel.: +86-010-8231-4905

Received: 10 August 2019; Accepted: 4 September 2019; Published: 7 September 2019

Abstract: In this paper, a novel method that utilizes the fractional correlation-based algorithm and the B-spline curve fitting-based algorithm is proposed to extract the complementary features for detecting the operating states of appliances. The identification of appliance operating states is one of the key parts for nonintrusive load monitoring (NILM). Considering the individual spurious emissions generated because of nonlinear components in each electronic device, the spurious emissions from the power cord can be picked up to solve the problem of data storage. Five types of common household appliances are considered in this study. The fractional correlation-based algorithm and B-spline curve fitting-based algorithm are used to extract two groups of complementary features from the spurious emissions of those five types of appliances. The experimental results show that the feature vectors extracted using the proposed method are obviously distinguishable. In addition, the features extracted show a good long-time stability, which is verified through a five-day experiment. Finally, based on support vector machine (SVM) and Dempster–Shafer (D-S) evidence theory, the identification accuracy reaches 85.5% using a combining classifier incorporated with the features extracted from the proposed methods.

Keywords: feature extraction; fractional correlation; B-spline curve fitting; combining classifier; support vector machine (SVM); Dempster–Shafer (D-S) evidence theory; nonintrusive load monitoring (NILM)

1. Introduction

Buildings account for a major portion of both primary energy and electricity consumption, and previous investigations indicated that building electricity consumption can be reduced by up to 10%–15% using energy management [1,2]. The smart grid and its supporting home automation network can promote residents' understanding of household appliance consumption and have the potential to be an effective means of energy management. To make household appliance consumption clear and understandable, each appliance requires a sensor, which can be achieved by modifying the existing appliances. This is associated with a low customer acceptance and thus is an obstacle in the promotion of the smart grid. Nonintrusive load monitoring (NILM), instead of placing a sensor on each appliance, reveals appliance usage at home from a single sensing point, which has been shown to increase customer acceptance of the home automation network [3–5].

NILM was first proposed by Hart [6] in the 1980s, in which the steady-state active power (P) and reactive power (Q) consumptions were used as load signatures. However, in certain situations, two appliances may consume the same amount of P and Q, which makes the two appliances indistinguishable [7]. Therefore, researchers extract appliance signatures from electrical parameters in switch states [8–10]. Nevertheless, transient signals require non-stop monitoring, which imposes data transmission and storage challenges. To observe the phase noise near 60 Hz, the current was measured 170 times per s in [8]. Furthermore, some of the transient signal can be missed,

while some of the non-transient signal can be mistaken as transient [11]. In this paper, to tackle this problem, spurious emissions in the spectrum of each appliance are employed to identify different appliances. This technique is based on the fact that spurious emissions coupled into the power cord from various appliances are different due to their unique nonlinear components, such as the diode. The employment of spurious emissions can obviously reduce the amount of data storage and still guarantee identification accuracy.

To recognize the operating states of household appliances, the features of historical data are firstly extracted to establish a classifier model, and then, the operating states of home appliances are identified through real-time data. The issue of how to extract individual features from measured data is critical. In our study, the features of spurious emissions are extracted from the time–frequency distribution and their physical properties, respectively. The peak intensity of the fractional correlation is higher than the traditional correlation, and the correlation operation can be extended to any angle of the time–frequency plane by the fractional Fourier transform (FrFT) [12]. Furthermore, electromagnetic emissions can be divided into different components based on physical characteristics, presented in the "basic emission waveform theory" [13]. The signal can be accurately separated through B-spline curve fitting on the physical layer to obtain the statistical features of different components. Therefore, in this paper, the fractional correlation-based algorithm and B-spline curve fitting-based algorithm are proposed to extract two groups of complementary features, which show long-time stability and then are used in a classifier. There are various classification algorithms, such as an improved fuzzy clustering algorithm [14], neural networks [15,16], and so on. However, considering the amount of training sample, a combining classifier is used. The combining classifier is based on support vector machine (SVM) and Dempster–Shafer (D-S) evidence theory, which shows the final identification results.

The main contributions in this paper are summarized as follows:

- Different from previous studies, spurious emissions in the spectrum of each appliance are used to identify the operating states based on the nonlinear components, which can obviously reduce the amount of data storage and still guarantee identification accuracy;
- Two groups of complementary features are extracted from the spurious emissions of household appliances. The fractional correlation-based algorithm exploits the correlation coefficient to evaluate the similarity between transformed sequences in the time–frequency plane, and the B-spline curve fitting-based algorithm obtains statistical features of different components on the physical layer.

The paper is organized as follows. In Section 2, the feature extraction method is proposed based on the fractional correlation-based algorithm and the B-spline curve fitting-based algorithm. Section 3 briefly introduces the combining classifier based on SVM and D-S evidence theory. The measurement system and experimental verification are presented in Section 4. The conclusion comes in Section 5.

2. Feature Extraction

To reduce redundant information and improve the accuracy, features are extracted from the original sequence. In this section, two feature extraction algorithms are proposed as follows: the fractional correlation-based algorithm is introduced in Section 2.1, which extracts features by calculating the fractional correlation coefficient between transformed sequences; the B-spline curve fitting-based algorithm is proposed in Section 2.2, which measures the uniformity of different components.

2.1. Fractional Correlation-Based Algorithm

The time–frequency distributions of signals from the same device are similar, while those from different devices are significantly different. Therefore, the peak intensity of the fractional correlation

coefficient between transformed sequences is used as the identification feature herein. The fractional correlation coefficient between sequences $y^i(x)$ and $y^j(x)$ is defined as:

$$CORR_p^{i,j}(\rho) = F^{-\pi/2}\left[Y_{p+1}^i(u)\left(Y_{p+1}^j(u)\right)^*\right](\rho) \tag{1}$$

$$Y_p(u) = \sum_x \widetilde{K_p}(u,x)\, y(x), \tag{2}$$

where $*$ denotes the conjugate; $Y_{p+1}^i(u)$ and $Y_{p+1}^j(u)$ are $p+1$ order FrFT of $y^i(x)$ and $y^j(x)$, respectively; $\widetilde{K_p}(u,x)$ is the kernel function, defined as $\widetilde{K_p}(u,x) = A_\alpha \exp\left[j\pi\left(u^2\cot\alpha - 2ux\csc\alpha + x^2\cot\alpha\right)\right]$ with $p \neq 2a$ (a is an integer) and $A_\alpha = \sqrt{1-j\cot\alpha}$, $\alpha \neq p\pi/2$. FrFT is essentially a traditional time-frequency transform. Compared to the common time-frequency distribution, in the fractional Fourier domain, there is no cross-term problem. Besides, FrFT shows all the characteristics of a signal in the time–frequency plane when the order p increases from 0–4 [17,18]. To explore the features of the spectral data in the time-frequency plane, only the idea of FrFT is used in this paper, and different from the traditional FrFT, spectral data are treated as a time domain signal with the same phase. To save computational cost, the multi-path parallel fast algorithm of FrFT is utilized, and the flow chart is shown in Figure 1. For discrete signals, its idea is to split the samples into the even ones and the odd ones and do all the operations on these two parts separately [19]. In our deployments, $p = 0.1$ is an optimal choice by 10-fold cross-validation [20], and fractional correlation peak intensity is extracted as the feature. For each sequence, the dimension of the feature vector depends on the number of training sequences. The extracting procedures for the training and the test vectors are summarized in Algorithm 1.

Algorithm 1 The extracting procedures for the training and the test vectors of the fractional correlation-based algorithm.

Training vectors:

1: Let $Y_{0.1}^i(u)$ and $Y_{0.1}^j(u)$ be 0.1 order FrFT of the training sequence $y^j(x)$ and $y^i(x)$, respectively, where $i, j = 1, 2, \cdots, N$, where N is the total number of training sequences over all classes;

2: Calculate fractional correlation coefficient $CORR_{0.1}^{i,j}$ between $Y_{0.1}^i(u)$ and $Y_{0.1}^j(u)$ using Equation (1);

3: Mark the maximum amplitude of $CORR_{0.1}^{i,j}$ as $MC_{0.1}^{i,j}$;

4: For the training sequence $y^i(x)$, the training vector can be denoted as $MC_{0.1}^i = \left(MC_{0.1}^{i,1}, MC_{0.1}^{i,2}, \cdots, MC_{0.1}^{i,N}\right)$, and used to train the SVM classifier.

Test vectors:

5: Let $Y_{0.1}^k(u)$, $k = 1, 2, \cdots, M$, be 0.1 order FrFT of test sequence $y^k(x)$, where M is the total number of test sequences;

6: For test sequence $y^k(x)$, calculate correlation coefficient $CORR_{0.1}^{k,i}$ between $Y_{0.1}^k(u)$ and $Y_{0.1}^i(u)$ using Equation (1), and the test vector $MC_{0.1}^k = \left(MC_{0.1}^{k,1}, MC_{0.1}^{k,2}, \cdots, MC_{0.1}^{k,N}\right)$ is obtained. The identification is implemented by employing the SVM classifier defined by the above training vectors.

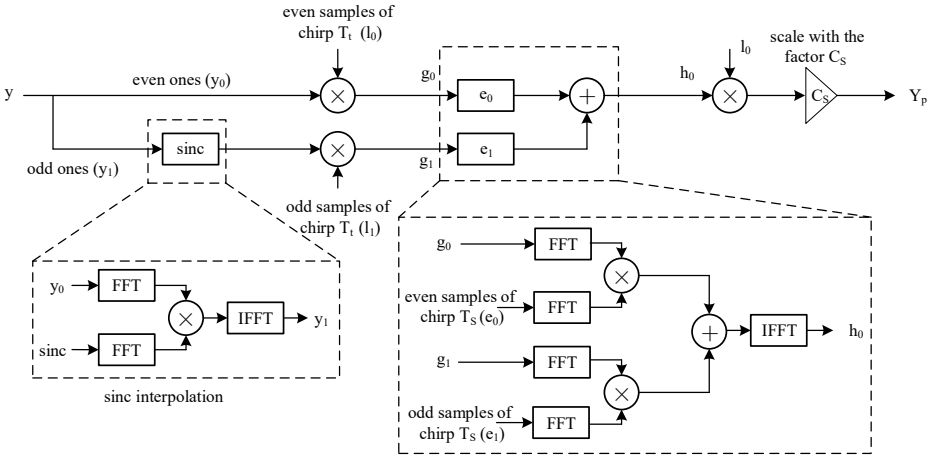


Figure 1. Multi-path parallel fast algorithm of fractional Fourier transform (FrFT).

2.2. B-Spline Curve Fitting-Based Algorithm

Electromagnetic emissions can be characterized by four basic waveforms: the square wave, sine wave, damped oscillation, and spike wave. The sine wave and square wave appear as a single-frequency point or clustered single-frequency points, referred to as narrowband components in the spectrum; while the damped oscillation and spike wave occupy a certain bandwidth and are defined as wideband components [21]. Roughly speaking, the emission spectrum can be decomposed into narrowband components and broadband components based on their physical properties. Since the narrowband components are sensitive to the environment, the broadband components are mainly considered in our study. The uniformity of the narrowband components and broadband components is exploited for identification.

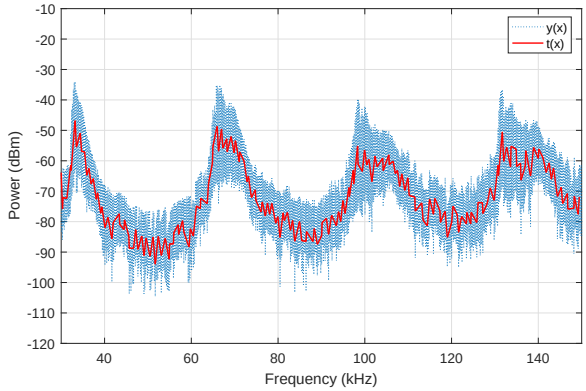
In the proposed method, the broadband component is firstly fitted using the B-spline curve, and then, the sequences are disaggregated. The B-spline fitting function can be expressed as [22]:

$$B(x) = \sum_{i=1}^{n} P_i N_{i,k}(x), 0 < x < 1, \tag{3}$$

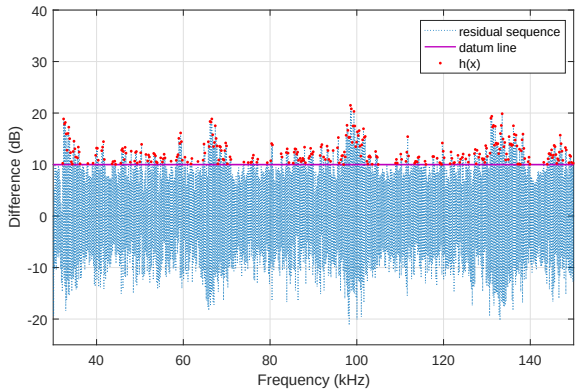
$$\begin{cases} N_{i,0}(x) = \begin{cases} 1, & x_i \le x \le x_{i+1} \\ 0, & x \notin [x_i, x_{i+1}) \end{cases} \\ N_{i,k}(x) = \frac{x - x_i}{x_{i+k} - x_i} N_{i,k-1}(x) + \frac{x_{i+k+1} - x}{x_{i+k+1} - x_{i+1}} N_{i+1,k-1}(x), \quad k > 0 \\ 0/0 = 0 \end{cases}. \tag{4}$$

where $N_{i,k}(x)$ denotes the basis function, which is obtained recursively and shown in Equation (4), and P_i is the position vector of polygon control points. Note that the choice of control points directly determines the trend of fitting curves. Here, some extreme points are selected and combined as double control points. Then, the fitting curve $t(x)$ of the broadband component is obtained, including 1000 sampling points, as shown in Figure 2a, taking a fluorescent table lamp as an example. After removing the broadband components from the original signal, we only keep the remaining signal, denoted as $h(x)$ in Figure 2b, which mainly consists of the obvious narrowband components. Considering the narrowband components that are an order of magnitude larger than the amplitude of the wideband components, the datum line was set to 10 dB. Both $t(x)$ and $h(x)$ were used to construct the other group of features for appliance identification. Spectral similarity was estimated using the mean and standard deviation of the two overall components and the segmented broadband sequence. The segmentation interval was set to 10 through cross-validation, which can guarantee the integrity of the information

and the burden of the classifier. The extracting procedures for the training and the test vectors are summarized in Algorithm 2.



(**a**)



(**b**)

Figure 2. B-spline curve fitting-based algorithm for the spectrum of a fluorescent table lamp: (**a**) original data $y\left(x\right)$ and fitting curve $t\left(x\right)$; (**b**) residual sequence after removing the broadband component and obvious narrowband sequence $h\left(x\right)$.

Algorithm 2 The extracting procedures for the training and the test vectors of the B-spline curve fitting-based algorithm.

Training vectors:

1: Fit the trend of the broadband component using the B-spline curve and separate the spectrum into two components $t\left(x\right)$ and $h\left(x\right)$;

2: Calculate the mean and standard deviation of $t\left(x\right)$ and $h\left(x\right)$, denoted as $(\mu_t, \sigma_t, \mu_h, \sigma_h)$;

3: Taking 10 as the interval, the segmented mean and standard deviation of $t\left(x\right)$ is computed, denoted as $(\mu_1, \sigma_1, \mu_2, \sigma_2, \cdots, \mu_{100}, \sigma_{100})$;

4: Combined with the length l of $h\left(x\right)$, the training vector of training sequence $y^i\left(x\right)$ is formed as $v^i = \left(\mu_t^i, \sigma_t^i, \mu_h^i, \sigma_h^i, l^i, \mu_1^i, \sigma_1^i, \cdots, \mu_{100}^i, \sigma_{100}^i\right)$, which is used to train the SVM classifier.

Test vectors:

5: As the training vector v^i of training sequence $y^i\left(x\right)$, the test vector of test sequence $y^k\left(x\right)$ is obtained as $v^k = \left(\mu_t^k, \sigma_t^k, \mu_h^k, \sigma_h^k, l^k, \mu_1^k, \sigma_1^k, \cdots, \mu_{100}^k, \sigma_{100}^k\right)$. The identification is implemented by employing the SVM classifier defined by the above training vectors.

3. Combining Classifier

The features extracted by different algorithms compose different feature spaces [23]; some samples are in an overlapping region of one feature space, while they may be absent from the overlapping region of another feature space. Therefore, high-dimensional feature spaces, superimposed directly on low-dimensional feature spaces, may deteriorate the aggregation of samples belonging to the same class. To solve this problem, a combining classifier was introduced, and the combination of several complementary classifiers improved the performance of individual classifiers.

In this paper, considering the high cost of sample collection, the SVM was employed as the individual classifiers. Its core idea is that of using a kernel function to map a non-linearly-separable space to a linear separable high-dimensional space and then searching for the optimal separating hyperplane [24]. For linearly-separable data, the decision function of SVM is defined as:

$$f(x) = sign\left(w \cdot \Phi(x) + b\right), \tag{5}$$

where $\Phi(x)$ is the kernel function and w and b determine the hyperplane of the mapped space. For linear inseparable data, the decision function of SVM is defined as:

$$f(x) = sign\left[\left(\sum_{i=1}^{l} \alpha_i y_i K\left(x_i, x_j\right)\right) + b\right], \tag{6}$$

where $K\left(x_i, x_j\right)$ is the kernel function, l is the number of training data, α_i is the i^{th} mode embedded dimension, and y_i is the i^{th} label. The classification problem is converted into a constrained optimization function, given by:

$$\begin{array}{ll} \min\limits_{\alpha} & \frac{1}{2}\sum\limits_{i=1}^{l}\sum\limits_{j=1}^{l} y_i y_j \alpha_i \alpha_j K\left(x_i, x_j\right) - \sum\limits_{j=1}^{l} \alpha_j \\ s.t. & \sum\limits_{i=1}^{l} y_i \alpha_i = 0, \\ & 0 \le \alpha_i \le C,\ i = 1, 2, \cdots, l \end{array}, \tag{7}$$

where C is a penalty factor. By solving the optimization problem in Equation (7), α and b are sequentially obtained. The classifier model is established.

Furthermore, the combination of individual classifiers was based on D-S evidence theory [25] in this study. By only using the recognition, substitution, and rejection rates of each individual classifier as the prior knowledge, the performance indexes of a classifier were well represented by testing the classifier with a test sample set. Firstly, the basic probability assignment (BPA) function of each classifier was determined. Suppose that the classification result of classifier C_k is j_k for an input feature vector X, $j_k = 1, 2, \cdots, M, M+1$; $j_k = M+1$ means the classifier C_k does not recognize the input X. M is the total number of identified classes. The error of classifier C_k is described by its confusion matrix as follows:

$$PT_k = \begin{bmatrix} n_{11}^k & n_{12}^k & \cdots & n_{1(M+1)}^k \\ n_{21}^k & n_{22}^k & \cdots & n_{2(M+1)}^k \\ \vdots & \vdots & \vdots & \vdots \\ n_{M1}^k & n_{M2}^k & \cdots & n_{M(M+1)}^k \end{bmatrix}, \tag{8}$$

$$n_{ij}^k = sum(C_k(X) = j | X \in i), \tag{9}$$

where n_{ij}^k denotes the number that the sample of class i has been assigned a label j by classifier C_k. The uncertainty of event $C_k(X) = j$ can be given by conditional probabilities expressed as:

$$\begin{aligned} P(X \in i | C_k(X) = i) &= \frac{n_{ii}^k}{\sum_{j=1}^{M} n_{ji}^k} \\ P(X \notin i | C_k(X) = i) &= 1 - \frac{n_{ii}^k}{\sum_{j=1}^{M} n_{ji}^k} \\ P(X \in C_{M+1} | C_k(X) = i) &= \frac{n_{i(M+1)}^k}{\sum_{i=1}^{M+1} \sum_{j=1}^{M} n_{ji}^k} \end{aligned} . \tag{10}$$

The BPA function of classifier C_k is $m_k^0 = 1$, if $j = M+1$. While $j = 1, 2, \cdots, M$, the focal elements are C_i, $\overline{C_i}$, Θ, and the corresponding BPA functions can be calculated by the following formula:

$$\begin{aligned} m_k^0(C_i) &= P(X \in i | C_k(X) = i) \\ m_k^0(\overline{C_i}) &= 1 - m_k^0(C_i) \\ m_k^0(\Theta) &= P(X \in C_{M+1} | C_k(X) = i) \end{aligned} . \tag{11}$$

$m_k^0(C_i)$, $m_k^0(\overline{C_i})$, and $m_k^0(\Theta)$ can be normalized to obtain $m_k^1(C_i)$, $m_k^1(\overline{C_i})$, and $m_k^1(\Theta)$. Then, considering the recognition performance of the classifier for different samples of the same class, the BPA functions are modified by the defined posterior probability function, $p_i^k(X) = \frac{1}{1+\exp(\|g^k(X) - E_i^k(X)\| - 1)}$, where $g^k(X)$ is the output vector of the classifier C_k for the input X and $E_i^k(X)$ is the average output vector of the classifier C_k for samples of class i. The modified BPA functions are:

$$\begin{aligned} m_k(C_i) &= m_k^l(C_i)\, p_i^k(X) \\ m_k(\overline{C_i}) &= m_k^l(\overline{C_i}) - \left[m_k^l(C_i) - m_k(C_i) \right] \\ m_k(\Theta) &= 1 - m_k(C_i) - m_k(\overline{C_i}) \end{aligned} . \tag{12}$$

The different identification results of the individual classifiers can be combined by D-S evidence theory using Equation (13), and the belief value $bel(j)$ expresses the final beliefs with uncertainty on each M mutually-exclusive propositions $X \in j$; the higher the $bel(j)$, the more likely $X \in j$ is to be true. The identification result is determined by a threshold $bel(j) > \alpha$, $\alpha = 0.5$ in this deployment.

$$\begin{aligned} m(C_i) &= m_i(C_i) \prod_{j=1, j \neq i}^{n} \left[m_j(\overline{C_j}) + m_j(\Theta) \right] + m_i(\Theta) \prod_{j=1}^{n} m_j(\overline{C_j}) \\ m(\Theta) &= 0 \\ m(\Theta) &= 1 - \sum_{i=1}^{n} m(C_i) \end{aligned} . \tag{13}$$

4. Experiment and Verification

In this section, we evaluate the identification performance of the proposed method through various numerical experiments. The measurement system is described in Section 4.1. The distinguishability and time stability of the proposed features are discussed in Sections 4.2 and 4.3, respectively. The identification performance by the combining classifier is verified in Section 4.4.

4.1. Measurement System

To validate the proposed method, some experiments were carried out in a shielded chamber, and the power of spurious emission was measured, which was different from the measurement of active power in [4]. The measurement system, mimicking the usage of household appliances, is shown in Figure 3, where a line impedance stabilization network (LISN, Schwarzbeck NNBL 8225) accessed the power cord, external noise was filtered out, and a spectral analyzer (CETC-41 AV4037MB) was used to provide the real-time spurious emission spectrum on the neutral or live line while the appliances were operating. Since the difference between a neutral and live line signal is not critical for identification, the access point of the spectral analyzer could be selected at either. Spectrum data were recorded with 1200 sample points per 15 s, to reduce data storage effectively [8], in which an WCS2210 Hall sensor and an UNO R3 Arduino were used, as well as the current with a measurement interval of 1/170 s obtained. However, the phase information was missing during the spectrum measurement, leading to an increase in the number of training sample, which was discussed in Section 4.2.

The actual access locations in the power cord of some household appliances were variable, that is the distance between Access Point A, marked as A1, A2, and A3 for Appliance-1, Appliance-2, and Appliance-3 as shown in Figure 3, and Monitoring Point B, the location of LISN, was not fixed in the measurement system. Therefore, low-frequency spectra were recorded to ensure that the measurement results were not affected by the access location of the appliance. In addition, the frequency of the tested grid was 50 Hz; hence, the measuring spectral range was set from 30 kHz–150 kHz with a resolution of 200 Hz to obtain the spurious emissions from the power cord. To gather as much useful information as possible, both the maximum and minimum were recorded.

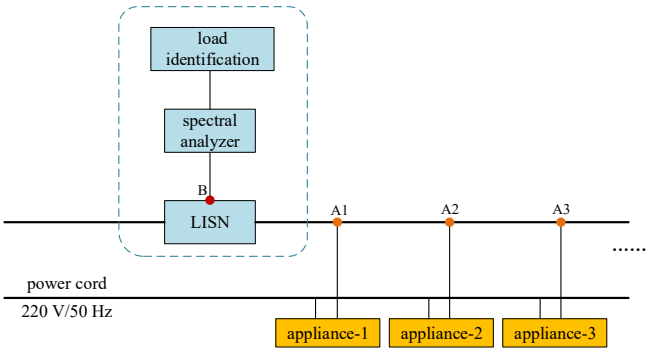


Figure 3. Schematic diagram of the measurement system. LISN, line impedance stabilization network.

4.2. Different Appliances' Measurement

In this subsection, the proposed method was verified using household appliances including a laptop, fluorescent table lamp, hair dryer, far infrared heater, and ultrasonic humidifier. The brand and model are shown in Table 1. In the following, only 50 groups of measured data, $N = 50$, 10 for each appliance, were used to demonstrate the results of our proposed methods for clear visualization. However, in our training procedure (discussed in Section 4.4), 50 groups of measured data for each appliance were used to train the combined classifier.

Table 1. The brand and model of the tested appliances.

Item	Brand and Model
Laptop charger, connected to a laptop	Lenovo, ADLX65YCC3A, 65 W(Lenovo, ThinkPad E480)
Fluorescent table lamp	Guandi, GD-3021, 11 W
Hair dryer	Panasonic, EH-WNA3B, 1800 W
Far infrared heater	Qianyou, NX, 1000 W
Ultrasonic humidifier	DEERMA, DEM-F470, 25 W

In the fractional correlation-based algorithm, 50 measured sequences were used as training sequences, thereby the dimension of feature vector $MC_{0.1}^{i} = \left(MC_{0.1}^{i,1}, MC_{0.1}^{i,2}, \cdots, MC_{0.1}^{i,N}\right)$ was 50 for each sequence, as shown in Figure 4. It can be seen that the correlation coefficients of 10 sequences were almost identical for the same appliance, while distinguishable from those of other appliances. In the B-spline curve fitting-based algorithm, the fitted broadband component of each appliance is shown in Figure 5, which obviously showed differences in the peaks, the amplitude, and other features, which led to obvious differences in the statistical parameters used.

Measurements covering different operating states were taken. Since an additional circuit was developed during the "on" state for the laptop, some narrowband components were introduced, which could be removed using curve fitting, because in our study, broadband components were mainly considered. In addition, the oscillation frequency of the broadband component for the ultrasonic humidifier varied with the states due to different circuit parameters, which could be used for the identification of operating states.

The results in Figures 4 and 5 show that most devices in isolation were identifiable by the extracted features. However, spurious emissions from the power cord involved, in practicality, multiple appliances in parallel and not a simple amplitude addition due to the phase [26]. Therefore, further measurements were required for appliance combinations.

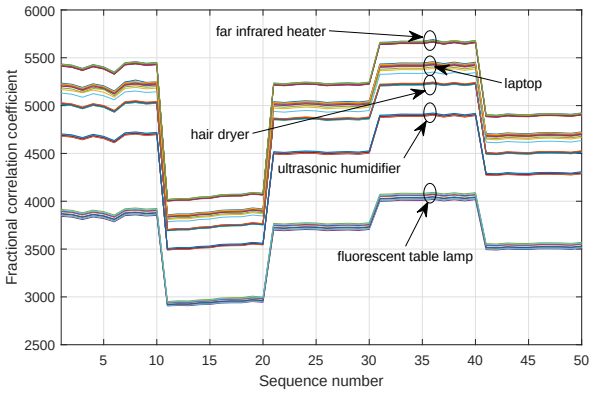


Figure 4. Fractional correlation coefficients for 50 training sequences.

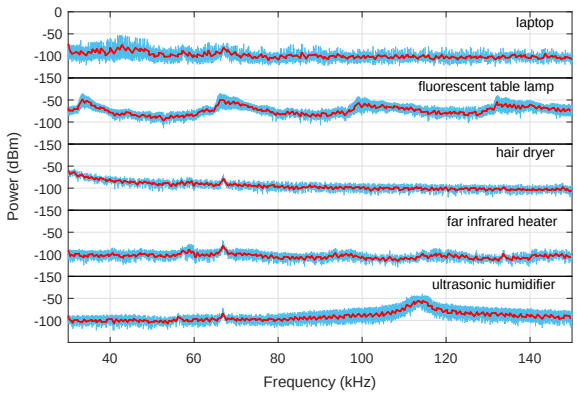


Figure 5. Fitted broadband components for five household appliances.

4.3. Long-Time Stability Measurement

The power grid was not stable due to the access of high-power appliances and the power quality, which indicated that the measured spectra for each appliance may be different over time. Although the appliances were in parallel, the current on the power cord would increase when a high-power appliance was operating, which would increase the voltage V_L across the loss resistance R_L of the

power cord and decrease the voltage V_A across the appliances, as shown in Figure 6. Therefore, the stability of the features extracted for identifying various loads was required. In our study, the table lamp was taken as an example to validate the long-time stability of the extracted features.

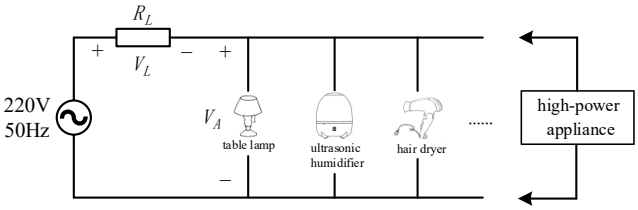
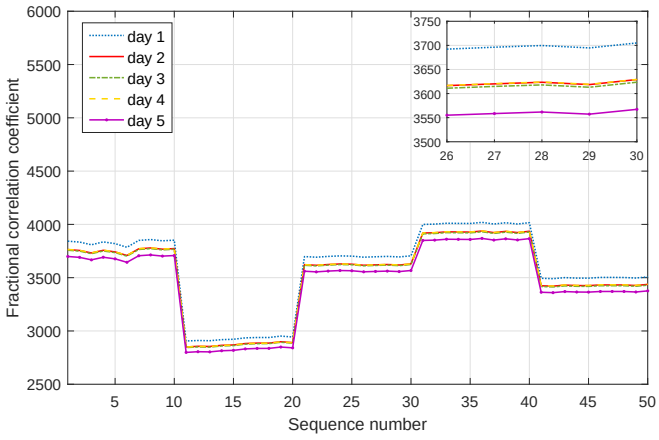
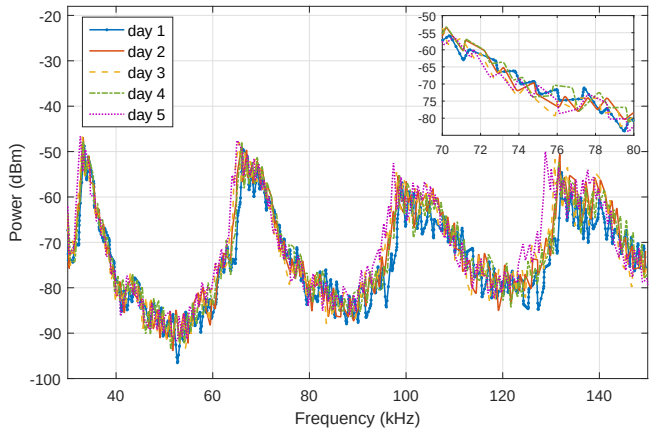


Figure 6. Schematic diagram of the power grid.

The spurious emissions of the table lamp were measured over five consecutive days, and the features were extracted using the proposed methods, as shown in Figure 7. Figure 7a demonstrates that the difference in correlation coefficients was limited to 150, and the coefficients over three days almost overlapped. In addition, Figure 7b indicates that the broadband components of the spectra were consistent over time, although there was some variation around 130 kHz. However, this small variation was acceptable for a classifier.



(**a**)



(**b**)

Figure 7. Correlation coefficients and fitted broadband components for the fluorescent table lamp over five consecutive days: (**a**) Fractional correlation coefficient; (**b**) fitted broadband component by B-spline curve fitting.

4.4. Identification Accuracy

To evaluate the identification accuracy of the features extracted from the proposed method, the probability of correct identification was used as a performance measure. Three types of appliances were considered: the fluorescent table lamp, hair dryer, and ultrasonic humidifier. Fifty measured data groups for each appliance were used, and the total number of training sequences was $N = 50 * 2^K$, where K is the number of appliances, $K = 3$ and $N = 400$ in this subsection.

Different feature spaces were constructed by the fractional correlation-based algorithm and B-spline curve fitting-based algorithm. Two SVM classifiers were trained by those two feature spaces, respectively, and the results from the two classifiers were combined for high identification accuracy. Here, two SVM classifiers were first implemented where the Gaussian radial basis function (RBF) was selected as the kernel function [27], then the combined results using D-S evidence theory were calculated. The accuracy was 85.5% for 800 test sequences, which were collected in different time periods over one day and covered all appliance combinations, showing better performance than individual classifiers.

Note that the household meter included a main switch, toilet switch, and kitchen switch, amongst others, which controlled the power cord of the apartment and the power cords of different rooms covering specific appliances. In our study, the power cord of the room was mainly considered, that is the number of measured appliances K and the total number of training sequences N were acceptable and would not cause an index explosion.

5. Conclusions

In this paper, a method to extract complementary spectrum features from spurious emissions was proposed for detecting the operating states of appliances, using the fractional correlation-based algorithm and B-spline curve fitting-based algorithm, which had the capacity to support NILM and home automation networks. Firstly, the spurious emissions in the spectrum coupled to the power cord were measured based on the existence of nonlinear components inside appliances. Then, two groups of complementary features were extracted using the similarity between transformed sequences in the time–frequency plane and the statistical features of different components on the physical layer. Finally, the feasibility and long-time stability of the proposed method were validated using several measurements, and the identification accuracy reached 85.5% using the combining classifier based on SVM and D-S evidence theory. It was found that the proposed feature extraction methods could be used to identify the appliances and showed good performance.

In future studies, further measurements should be performed for more household appliances in an actual home and the extracted features incorporated with other features to improve the identification accuracy. In addition, the identification results of appliance operating states would be applied to recover the energy consumption per appliance.

Author Contributions: Conceptualization, D.S.; methodology and software, Q.S.; validation, Q.S. and W.W.; formal analysis, Q.S.; investigation, Q.S. and W.W.; data curation, Q.S.; writing, original draft preparation, Q.S.; writing, review and editing, W.W. and H.X.; supervision, D.S.; project administration, H.X.; funding acquisition, D.S.

Funding: This research was funded by the National Natural Science Foundation of China Grant Number 61427803.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zeifman, M.; Roth, K. Nonintrusive appliance load monitoring: Review and outlook. *IEEE Trans. Consum. Electron.* **2011**, *57*, 76–84. [CrossRef]
2. Salerno, V.; Rabbeni, G. An extreme learning machine approach to effective energy disaggregation. *Electronics* **2018**, *7*, 235. [CrossRef]

3. Suzuki, K.; Inagaki, S.; Suzuki, T.; Nakamura, H.; Ito, K. Nonintrusive appliance load monitoring based on integer programming. In Proceedings of the 2008 SICE Annual Conference, Tokyo, Japan, 20–22 August 2008; pp. 2742–2747.
4. Biansoongnern, S.; Plangklang, B. Nonintrusive load monitoring (NILM) using an Artificial Neural Network in embedded system with low sampling rate. In Proceedings of the 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand, 28 June–1 July 2016; pp. 1–4.
5. Zhong, M.; Goddard, N.; Sutton, C. Signal aggregate constraints in additive factorial HMMs, with application to energy disaggregation. In Proceedings of the Advances in Neural Information Processing Systems, Montréal, QC, Canada, 8–13 December 2014; pp. 3590–3598.
6. Hart, G.W. Nonintrusive appliance load monitoring. *Proc. IEEE* **1992**, *80*, 1870–1891. [CrossRef]
7. Laughman, C.; Lee, K.; Cox, R.; Shaw, S.; Leeb, S.; Norford, L.; Armstrong, P. Power signature analysis. *IEEE Power Energy Mag.* **2003**, *1*, 56–63. [CrossRef]
8. Lee, D. Phase noise as power characteristic of individual appliance for non-intrusive load monitoring. *Electron. Lett.* **2018**, *54*, 993–995. [CrossRef]
9. Ducange, P.; Marcelloni, F.; Antonelli, M. A novel approach based on finite-state machines with fuzzy transitions for nonintrusive home appliance monitoring. *IEEE Trans. Ind. Inform.* **2014**, *10*, 1185–1197. [CrossRef]
10. Lin, Y.H.; Tsai, M.S. Development of an improved time–frequency analysis-based nonintrusive load monitor for load demand identification. *IEEE Trans. Instrum. Meas.* **2013**, *63*, 1470–1483. [CrossRef]
11. Chen, K.Y.; Gupta, S.; Larson, E.C.; Patel, S. DOSE: Detecting user-driven operating states of electronic devices from a single sensing point. In Proceedings of the 2015 IEEE International Conference on Pervasive Computing and Communications (PerCom), St. Louis, MO, USA, 23–27 March 2015; pp. 46–54.
12. Zalevsky, Z.; Mendlovic, D.; Caulfield, J.H. Localized, partially space-invariant filtering. *Appl. Opt.* **1997**, *36*, 1086–1092. [CrossRef] [PubMed]
13. Su, D.; Xie, S.; Chen, A.; Shang, X.; Zhu, K.; Xu, H. Basic Emission Waveform Theory: A Novel Interpretation and Source Identification Method for Electromagnetic Emission of Complex Systems. *IEEE Trans. Electromagn. Compat.* **2017**, *60*, 1330–1339. [CrossRef]
14. Liu, Q.; Kamoto, K.M.; Liu, X.; Sun, M.; Linge, N. Low-Complexity Non-Intrusive Load Monitoring Using Unsupervised Learning and Generalized Appliance Models. *IEEE Trans. Consum. Electron.* **2019**, *65*, 28–37. [CrossRef]
15. Zhang, C.; Zhong, M.; Wang, Z.; Goddard, N.; Sutton, C. Sequence-to-point learning with neural networks for non-intrusive load monitoring. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018.
16. Bonfigli, R.; Felicetti, A.; Principi, E.; Fagiani, M.; Squartini, S.; Piazza, F. Denoising autoencoders for non-intrusive load monitoring: improvements and comparative evaluation. *Energy Build.* **2018**, *158*, 1461–1474. [CrossRef]
17. Ozaktas, H.M.; Arikan, O.; Kutay, M.A.; Bozdagt, G. Digital computation of the fractional Fourier transform. *IEEE Trans. Signal Process.* **1996**, *44*, 2141–2150. [CrossRef]
18. Almeida, L.B. The fractional Fourier transform and time-frequency representations. *IEEE Trans. Signal Process.* **1994**, *42*, 3084–3091. [CrossRef]
19. Bultheel, A. *A two-phase implementation of the fractional Fourier transform*; TW Reports; Department of Computer Science, KU Leuven: Leuven, Belgium, 2011.
20. Yadav, S.; Shukla, S. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In Proceedings of the 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 27–28 Feburary 2016; pp. 78–83.
21. Shang, X.; Su, D. Use modified Lomb-Scargle method to analyze electromagnetic emission spectrum. In Proceedings of the 2015 IEEE 6th International Symposium on Microwave, Antenna, Propagation, and EMC Technologies (MAPE), Shanghai, China, 28–30 October 2015; pp. 415–420.
22. Su, D.; Wu, J.; Shang, X.; Zhu, K.; Peng, Z. A novel method of detecting and analyzing electromagnetic emission. In Proceedings of the 2017 IEEE Asia Pacific Microwave Conference (APMC), Kuala Lumpar, Malaysia, 13–16 November 2017; pp. 1156–1158.

23. Ho, T.K.; Hull, J.J.; Srihari, S.N. Decision combination in multiple classifier systems. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16*, 66–75.
24. Lopes, F.F.; Ferreira, J.C.; Fernandes, M.A. Parallel Implementation on FPGA of Support Vector Machines Using Stochastic Gradient Descent. *Electronics* **2019**, *8*, 631. [CrossRef]
25. Xu, L.; Krzyzak, A.; Suen, C.Y. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Trans. Syst. Man Cybern.* **1992**, *22*, 418–435. [CrossRef]
26. Lane, R.; Badran, T.; Marshall, G. Measurements for remote identification of electrical equipment. *Electron. Lett.* **2017**, *53*, 1001–1005. [CrossRef]
27. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 27. [CrossRef]