J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 7645--7658

Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Polynomial matrix sparse coding algorithm for extracting non-electrical signal characteristic parameters of key components of high-voltage direct current converter valves

Qinze Yang[1,✉], Yi Jiang[1], Hai Jiang[1], Haiyang Chu[1], Hongnie Cai[1], Ao Feng[2]

[1] *Tianshengqiao Bureau, Ultra-High Voltage Transmission Company of China Southern Power Grid Co., Ltd., Xingyi, Guizhou, 562400, China*
[2] *Wuhan University, Wuhan, Hubei, 430072, China*

ABSTRACT
Sparse decomposition has been generally emphasized in signal processing theory. In this paper, a non-electrical signal feature dataset of key components of high-voltage DC converter valve is established by using principal component analysis to streamline the data volume. The compression-aware feature extraction algorithm based on polynomial matrix sparse coding is used to extract and collect the non-electrical signal parametric data. Through the performance over the experimental signal analysis, it can be known that the eigenvalues of a total of 10 parameters, including the infrared temperature measurement results, the appearance, the presence of corrosion or dirt, and the presence of abnormal vibration and sound, are all greater than 1. Therefore, these 10 parameters are identified as the key parameters. When the number of measurement points is between 64 and 200, the algorithm in this paper can satisfy the need of feature extraction when the signal length is insufficient, compared with the traditional approach. In the empirical analysis of the vibration signal as an example, the method of this paper can effectively extract the frequency and time domain of the vibration signal.

Keywords: sparse coding, polynomial matrix, non-electrical signals, key components, feature extraction

1. Introduction

In the process of global energy transition and power infrastructure modernization, high-voltage direct current (HVDC) transmission technology has become an important technical means for building smart grids due to its excellent transmission efficiency, long-distance transmission capability, and powerful grid interconnection [1-3]. With the increasing global demand for clean energy, the advantages of high-voltage direct current transmission (HVDC) technology in long-distance transmission and cross-

✉ Corresponding author.
E-mail address: xbkgyyds0102@163.com (Q. Yang).
Received 20 February 2024; Revised 15 April 2024; Accepted 20 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-423

regional grid interconnection are particularly significant. Compared with traditional AC transmission, HVDC technology can not only effectively reduce energy loss, but also improve the stability and reliability of the power system, especially in large-scale renewable energy grid integration and cross-country power transmission plays an irreplaceable role [4-7].

As the core facility to realize the mutual conversion of DC and AC in the power system, the converter station reduces the energy loss in the high-voltage transmission process, and at the same time, ensures that the AC generated by the power plant can be transmitted over long distances in the form of high-efficiency DC, and finally converted back to the AC form required by the users [8-10]. In the technical composition of the converter station, the valve control system is one of the core control units of the whole converter station, which is responsible for accurately managing the working status of the converter valves in order to realize the efficient conversion between DC and AC power, and its stability is crucial for guaranteeing the safe operation of the power system [11-13].

Thyristor converter valve is the key equipment of DC power transmission system, and its reliability is of great significance to the performance of the whole power system. Under normal operating conditions, the stress environment of the thyristor in the converter valve is uniformly distributed and the degree of change is not large, and the reliability of the high-voltage DC transmission converter valve system can be directly assessed using data from engineering statistics [14-16]. However, in special environments, due to the existence of parasitic capacitance parameters in the valve body, the electrical and thermal stresses inside the converter valve will change, and the intensity of the stress change will affect the failure rate of the key components, which in turn affects the reliability of the entire high-voltage direct current transmission system [17-19]. Thyristor failure rate is the core parameter for evaluating the reliability of the converter valve system and even the whole HVDC transmission system [20]. Therefore, it is necessary to extract and analyze the non-electrical signal characteristic parameters of these key components.

In order to realize the extraction of non-electrical signal feature parameters of key components of high-voltage DC converter valve, this paper combines coefficient coding with polynomial observation matrix to construct a compressed perception algorithm. Based on this, a feature extraction algorithm for compressed sensing is proposed. Aiming at the main components of the converter valve, principal component analysis is applied to select 10 principal component feature parameters, including infrared temperature measurement results, appearance, the presence of corrosion or dirt, etc., to construct a feature data set. The polynomial matrix sparse coding algorithm is compared with the traditional method to judge the feasibility of the method in this paper. The N205M converter valve is taken as the measured object, and the vibration signal is taken as an example to explore the effectiveness of signal feature extraction.

2. Selection and dimensionality reduction of non-electrical signal characteristic data of converter valve

2.1. Selection of non-electrical signal characterization parameters

The main components of the thyristor converter valve include the thyristor assembly, the valve cooling assembly and the valve arrester in 3 parts, and each part is independent of each other. Each part contains original data, operation data, maintenance test data and so on.

1) Thyristor assembly: The thyristor assembly consists of thyristor elements and their required triggering, protection and monitoring circuits, damping circuits, reactors, etc. The thyristor assembly is the main component of the converter valve. The thyristor assembly is the most basic functional unit of the converter valve, which determines the converter valve's through-flow capability.

2) Valve cooling components: the main task of valve cooling components is to dissipate the heat loss generated by the thyristor converter valve in all kinds of operating conditions, so that the temperature

of the thyristor is maintained at a relatively low level, to ensure the reliability of the converter valve operation, and at the same time, to improve the power transfer capability.

3) Valve arrester: The valve arrester is suspended from the outside of the valve tower by insulators. Each single valve is equipped with a valve arrester to prevent excessive voltage on the converter valve. The valve arrester is connected in parallel with each single valve through a flexible connection bus, forming a flexible connection system. In this paper, on the basis of the State Grid enterprise standard "Guidelines for evaluating the state of key elements of HVDC converter valves", the key state quantities of three components are selected, and the influence of environmental factors on the operation of converter valves is also taken into account, so that a total of 19 non-electrical signal characteristics are selected, i.e., 19 characteristic indexes, which are represented by $X = (i = 1,2,\ldots,19)$. Among the above 19 state quantities, some of them can be obtained by monitoring, while others need to be scored and quantified by testing or visual inspection.

2.2. Non-electrical Signal Characteristic Parameter Matrix Establishment

The parameters in the basic index system for condition evaluation of key components of high-voltage DC converter valves are very complex, and there are many judgment conditions that can be selected, so how to select judgment conditions is very important for the reliability of the condition evaluation results. Generally speaking, the principle of comprehensiveness, importance and practicability should be considered when selecting the key index criteria. The principle of comprehensiveness means that the selected criteria can reflect the basic attributes of each basic state parameter in a more comprehensive way, so as to achieve the goal of using the key parameters to characterize all the basic parameters. The principle of importance means that the more critical covariates should be ranked according to their importance. Practicality means that it is easier to obtain the basic parameters in engineering practice for evaluation. Aiming at the shortcomings of the large storage capacity of feature data in the key components of the high-voltage DC converter valve, the principal component analysis method is used to reduce the dimensionality of the features, and the non-electrical signal feature data set of the key components of the high-voltage DC converter valve is constructed.

In this paper, five judgments are selected for key indicator parameter analysis. Different types of converter valves have different state parameters to produce different changes, the data as the first criterion. In the high-voltage direct current converter valve key components factory test experiments and quota data can also be used as a criterion. As the important reference documents in operation and maintenance of HVDC transmission projects, the HVDC converter valve condition maintenance guideline and HVDC converter valve condition evaluation guideline are also of guiding value for the selection of key condition parameters. In addition, the real-time monitoring data of the converter valve during the operation of the key components of the converter valve as an important part of the condition evaluation is also one of the criteria for the key parameters. Summarize the types of criteria as shown in Table 1.

Table 1. The selection criterion of key parameter index system of thyristor

Array element	Meaning
P_1	The percentage of the parameters in the data statistics of the past years
P_2	The existence of the parameter in the transmission of the pressure dc transmission flow valve is indicated that the existence of the parameter is not present
P_3	The existence of the parameter in the direction of the pressure dc transmission flow valve is indicated that the existence of the parameter is not present
P_4	The existence of the factory test and the quota information is said to be that the existence of 0 is not present
P_5	The existence of this parameter in real-time monitoring data indicates that the existence of 0 is not present

2.3. *Non-electrical signal base indicator system*

According to the above established high-voltage DC converter valve key component characteristic parameters of the basic index system and the listed criterion matrix to establish the basic index criterion matrix, the basic index system is shown in Table 2.

Table 2. The matrix of the base index of the aging state of the thyristor

	Classification	State quantity	P_1	P_2	P_3	P_4	P_5
X1	Appearance	Appearance	66.2%	1	1	0	1
X2	Thermal stress and mechanical stress test	Thermal resistance and impedance	90.2%	1	1	0	1
X3		Temperature change	21.2%	1	0	0	0
X4		Shock test	0.2%	1	1	0	1
X5		Salt fog test	36.2%	0	0	0	1
X6		High temperature storage test	21.2%	0	1	0	0
X7	Inspection information	There is rust or dirt	36.7%	1	0	1	1
X8		There is an abnormal vibration and harmony	30.4%	1	1	1	1
X9		Infrared temperature measurement	91.3%	0	1	1	1
X10	Failure check information	Control unit function	40.1%	1	0	0	1
X11		Reverse recovery phase protection unit function	33.2%	1	1	1	0
X12		Control component appearance	10.2%	0	1	1	1
X13		Mechanical pressure of thyristor	4.6%	0	1	0	1
X14	Overhaul record	Overhaul	99.5%	1	1	0	1
X15		Defective condition	51.2%	1	1	0	0
X16	Operating environment	Valve room temperature	99.8%	0	0	1	1
X17		Valve room humidity	41.3%	0	0	0	1
X18		Valve hall electromagnetic interference	0.2%	0	0	1	0
X19		Operating electrical conditions	89.3%	1	1	1	0

The correlation coefficient matrix R of the criterion matrix was calculated using SAS software as shown in Figure 1. The table contains the eigenvalue, variance contribution rate and cumulative variance contribution rate of each base state parameter. According to the two principles that the eigenvalue of the base state parameter is greater than 1, and the cumulative variance contribution rate of each principal factor is more than 0.85, it can be seen that the eigenvalue of these four principal factors is greater than 1 when the first four factors are used as the principal factors and their cumulative variance contribution rate reaches 0.95, which is greater than the 0.85 requirement. That is to say, it is considered that these four main factors have the ability to characterize all the basic parameters and their ability to explain the original basic index system is strong enough.

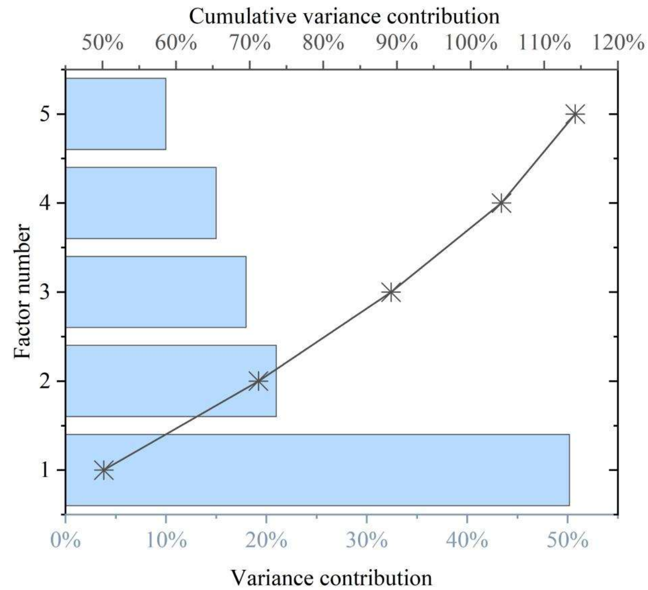


Fig. 1. The contribution rate of the main cause

The eigenvalues and contributions of the main factors are shown in Table 3.

Table 3. The main cause of the child eigenvalue and the contribution rate

Factor	Eigenvalue	Variance contribution(%)	Cumulative variance contribution(%)
1	3.62	50.2%	50.2%
2	2.12	20.3%	70.5%
3	1.74	13.0%	83.5%
4	1.06	12.0%	95.5%
5	0.17	4.5%	100.0%
6	0	0%	100.0%
...	...	...	...

The first four factors as the main factor, the original base index criterion to take the variance maximization method to rotate the matrix, so as to have the factor score matrix as shown in Table 4.

Table 4. The main cause of the child score

	Factor			
	1	2	3	4
D_1	0.6672	-0.2179	0.6748	0.2763
D_2	0.8715	0.15095	-0.3847	0.1225
D_3	0.9124	0.0162	-0.2163	0.2578
D_4	-0.2879	0.9224	0.1524	0.2618
D_5	0.7223	0.3748	0.1742	-0.5776

A scatter plot can be made for each base indicator covariate expressed using three principal factors as shown in Figure 2, where X1-X19 denote the serial numbers of each base covariate.

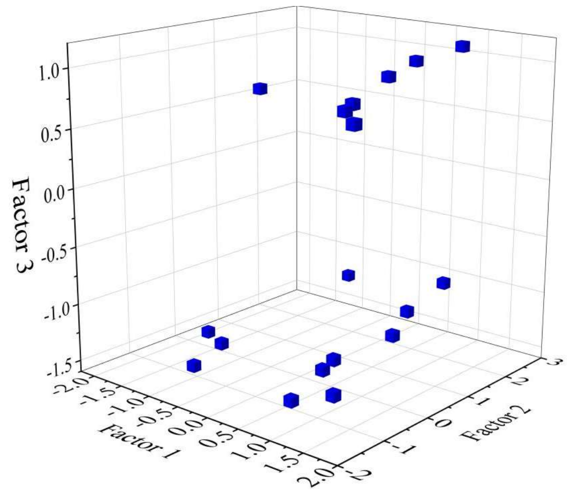


Fig. 2. Each basic parametric scatter diagram

The composite score of each basic indicator according to the above four main factors can be calculated, as shown in Table 5. In the table of factor scores and composite scores, there is a part of the factor in a judgment under the score or the composite score is negative, here the positive and negative does not have practical significance, but indicates the relative degree with the average level, the positive number indicates that in the main factor of the interpretation of all the basic parameters of the system in the average level, while the negative number indicates that in the average level below.

Table 5. The basic parameters score for the main cause

	Factor1	Factor2	Factor3	Factor4	Comprehensive score
X9	0.5170	1.944	0.9725	0.7573	1.7122
X1	0.6624	1.0425	0.9245	0.8712	1.6241
X19	0.6422	1.6427	-1.1748	0.9528	1.4857
X7	0.8412	-0.0457	0.8122	0.9975	1.3826
X16	-0.9823	2.5842	-1.2123	1.0426	1.3744
X8	0.8677	-0.2536	0.7771	1.0203	1.3628
X10	0.6312	-0.1189	-1.3224	1.5147	1.2010
X11	0.6693	-0.4225	-1.3526	1.2234	1.1647
X17	-1.8475	0.2745	-1.5426	1.8975	1.0853
X12	0.8092	-1.3026	-1.4328	1.3028	1.0126
X2	1.0526	2.2236	1.115	-1.6254	-0.3526
X14	0.8002	2.3316	-0.9925	-1.4628	-0.5262
X5	-1.1032	0.5823	0.7132	-0.6125	-0.7285
X15	1.0845	0.5863	-1.1526	-1.238	-0.7885
X3	1.5723	-1.0945	0.8224	-1.235	-0.9526
X13	1.3625	-1.1722	-1.3122	-1.0516	-1.0526
X2	-1.2024	-0.2124	-1.4677	-0.3824	-1.0845
X5	-1.2024	-0.2124	-1.4677	-0.3824	-1.0845
X18	-1.0845	-0.9638	-1.5248	-0.2824	-1.1849

Analyzing the data, we can find that not all data elements are necessary, they are often full of redundancy and noise. We can only find the most important elements in the original data, we can use these components to approximate the original data to achieve the purpose of data dimensionality reduction.The mathematical principle of PCA is to use a set of orthogonal bases to describe the original data space, which is required to be as small as possible and can maximize the approximation of all the data in the original space. In other words, the most important principal elements are found and data redundancy is removed to the greatest extent possible. The combined scores of each base parameter are shown in Figure 3. A total of 10 basic parameters, including infrared temperature measurement results, appearance, the presence of corrosion or dirt, and the presence of abnormal vibration and sound, which have positive scores, are selected to be constructed as the key index system for evaluating the condition performance of key components of the high-voltage direct current (HVDC) converter valve.

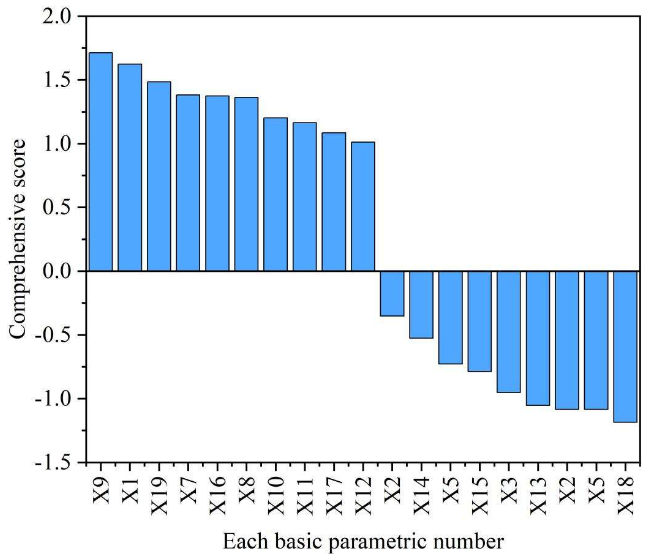


Fig. 3. Basic index parameter composite score

3. Feature extraction for polynomial matrix sparse coding algorithms

3.1. Mathematical description of sparse coding

Sparse representation, proposed in the 1990s, laid the foundation for the study of sparse coding theory, which has been developing rapidly since then, and has been widely used in signal processing fields such as image compression, denoising, and feature extraction. In addition, sparse representation theory is particularly important in the field of artificial intelligence. The theory gives an effective way from machine intelligence to artificial intelligence.

The concept of sparse coding originates from the study of optic neural networks, which is a representation of neural networks with multidimensional data in which only a small number of neurons are active at the same time. Biological experiments have shown that the processing of external stimuli in the optic cortex adopts the principle of neural sparse representation, which both provides a simple representation of complex and redundant information and facilitates the extraction of the most essential features of the stimulus by the upper sensory nerves.

The basic principle of sparse representation: based on the overcomplete dictionary, seeking to express the original information as much as possible with as few atoms as possible in the transform domain, in short, representing complex signals in the simplest way possible, this is sparse representation. For a signal, viewed as a vector in a finite-dimensional vector space, a signal is strictly sparse if only N sample points have non-zero values and all other sample points are zero, and the signal can be called a sparse signal. Even though some signals are not sparse in the current domain of expression, it is possible to sparsely represent them by transforming them to the appropriate domain of expression, e.g., a sinusoidal signal, which is not sparse in the time domain, is very sparse in the frequency domain after Fourier transform.

Mathematically speaking, sparse coding is currently assumed to be a representation of a linear decomposition of multidimensional signal data, and its mathematical description is very simple. Suppose the input data $X=\left(x_1,x_2,\cdots,x_n\right)^T$ is a n-dimensional random vector. Denote the linearly transformed m-dimensional vector by $S=\left(s_1,s_2,\cdots,s_m\right)^T$, then the linear transformation matrix is $m\times n$-dimensional, denoted as M. The linear transformation representation is given as follows:

$$S = MX \tag{1}$$

M in the above equation is also known as the sparse transformation matrix, and each of its row vectors is analogous to a wavelet basis in a wavelet transform; the linearly transformed sparse component S satisfies the requirements of a sparse (super-Gaussian) distribution, and the vectors $\left(s_l,s_2,\cdots,s_k\right)$ are as independent of each other as possible.

The sparse coding model is able to represent the processes and characteristics of the simple cells of the primary visual cortex that encode the images of external visual stimuli, and this dynamic mapping and the issuing properties of the cells of the expressive layer lead to a sparse distribution of cellular issuance. Since mammalian visual physiological processes are complex, when simulating the receptive field properties of neurons in the primary visual system with sparse coding models. It is usually necessary to set up the following assumptions:

1) The input data has a sparse (or super-Gaussian) structure.
2) The individual sparse coefficients are statistically independent of each other.
3) The sparse transform is a linear transform.
4) The basis functions are non-singular.
5) The sparse components and the noise signal are independent of each other.

3.2. Principles of Polynomial Observation Matrix Construction

Polynomial deterministic matrices are a method of constructing deterministic observation matrices. It has been used by many researchers because of its advantages of both determinism and sparsity. The polynomial observation matrix is constructed as follows:

Step 1: r th order polynomial construction. Given a prime number p, a finite field of order p is $F=\{0,1,2\ldots p-1\}$, if any given natural number r, $0<r<p$, p_r denote the set of polynomials $Q(x)$ whose highest number is not greater than r:

$$Q(x)=a_0+a_1x+\ldots+a_rx^r \tag{2}$$

where $Q(x)$ has coefficient $a_i\in F$, and as coefficient a_i varies there are p^{r+1} polynomials of order less than or equal to r. Each polynomial can construct a column of the observation matrix, so the observation matrix can contain up to p^{r+1} columns.

Step 2: Observe the column construction method of the matrix. For each polynomial $Q(x)$, first define a $p\times p$ zero matrix E that maps the value of 0 at the $Q(x)$ th position $(x,Q(x))$ of the x th column in the matrix E to 1. The mapped matrix is then stacked column-wise into a $M\times 1$ column vector where $M=p^2$.

Following step 1 polynomial coefficient variation produces p^{r+1} different column vectors, and following step 2 each column vector has dimension $p^2\times 1$, so the maximum capacity of the composed observation matrix is $p^2\times p^{r+1}$.

Polynomial Matrix Generation and Observation Matrix Construction Methods, given the prime number p and the highest number r, the polynomial matrix can be uniquely determined, and since the polynomial matrix formula of $p\times p$ contains only 1 non-zero element per column and p non-zero elements per matrix, the observation array of $p^2\times N$ which it is constructed also has only p non-zero values of 1 out of the p^2 elements of each column, and all the other matrix elements have values of 0, i.e., the observation matrix is sparse. Moreover, the sparse value positions are determined after the construction columns of the observation array are selected. Compared with random matrices, such matrices not only have low computational complexity for observation and reconstruction, but also the certainty of their elements is more favorable for hardware implementation.

3.3. Compression perception based on polynomial matrix sparse coding algorithm

If the N long signal s is time domain k sparse, i.e., s contains only k non-zero values $\left(N \gg k\right)$, then the signal s can be compression-aware to realize both sampling and compression. On the other hand, if N the long signal f is sparse under some transform domain Ψ, the signal f satisfies the requirement of signal compressibility. The process of compression perception is divided into the following three steps:

1) Sparse Transform. By means of a suitable sparse transform, a compressible signal f is converted into a sparse representation s in the transform domain Ψ. Let the signal $f \in R^N$, with basis vector $\varphi_i (i = 1, 2, \ldots, N)$, be the decomposition transform of f:

$$f = \sum_{i=1}^{N} s_i \varphi_i \; Or \; f = \Psi s \tag{3}$$

If the transform coefficients s decay exponentially and converge to zero after sorting from large to small under the transform domain Ψ, the signal f can be considered as compressible, and s is a sparse representation of f over the sparse domain Ψ. The sparse representation of the signal is a necessary condition for compressed perception.

2) Signal observation. Design a $M \times N\left(M \ll N\right)$ -dimensional observation matrix Φ that is uncorrelated with the sparse transformation basis Ψ to observe the signal f and obtain an $M \times 1$ -dimensional observation vector y. The compression observation projects the compressible signal f onto the $M \times N\left(M \ll N\right)$ observation matrix Φ that is uncorrelated with the sparse transformation matrix:

$$y = \Phi f \tag{4}$$

The $M \times 1$ observation vector y containing the necessary information of f reduces the representation of the observation object from N to M dimensions, a process that simultaneously realizes the perception and compression of the signal. More particularly, since the observation is non-adaptive, the choice of the observation matrix Φ does not depend on the signal f, thus the system is able to realize the sampling of arbitrarily sparse signals with the same observation matrix. Considering further the sparse transform Ψ performed on the compressible signal f, Eq. (4) can be expressed as:

$$y = \Phi f = \Phi \Psi s = \Theta s \tag{5}$$

$$\Theta = \Phi \Psi \tag{6}$$

where s is a vector of $N \times 1$ and Θ is a $M \times N$ composite matrix. For a compressible signal f, if the observation matrix Φ is uncorrelated with the sparse basis Ψ, then $\Theta = \Phi \Psi$ the RIP property is satisfied to a large extent.

The colored squares of s denote non-zero value locations and the blank cells denote signal values 0. The observation vector y is the projection of the sparse signal s onto the matrix Θ so that the less correlation there is between the column vectors of Θ the more information about the signal s is retained by y the signal.

3) Estimation reconstruction. Reconstruction of the original signal from an $M \times 1$ -dimensional observation vector. The reconstruction of a signal requires solving the original signal s of length N from an M long observation vector y and the reconstruction can be described as a problem of minimizing the number of l_0 paradigms:

$$\left(P_0\right)\min \Psi\left|s_0\right| \; s.t. \; y=\Phi\Psi s \tag{7}$$

The l_0 paradigm of a signal denotes the number of nonzero elements of the signal.

The solution of the P_0 problem is an NP problem, and in order to avoid traversing all C_N^K possible sparse combinations of positions of the original signal s when reconstructing it with y , the solution is based on a fast optimization strategy. Its typical fast approximation solution is the matched tracking (MP) family of algorithms.

It can be shown that, under the RIP condition constraints, the optimal l_1 paradigm instead of the l_0 optimal paradigm can accurately reconstruct the k sparse vectors with high probability when obeying the independently and identically distributed Gaussian observables $M \geq c \cdot k \cdot \log N / k$:

$$\left(P_1\right)\min \Psi\left|s_1\right| \; s.t. \; y=\Phi\Psi s \tag{8}$$

3.4. Compression-aware feature extraction

Among the feature extraction algorithms based on compressed sensing, the OMP algorithm is widely used because of its fast convergence speed, simple principle and easy implementation, and it is also one of the important feature extraction algorithms in the wireless channel feature extraction technology. The OMP algorithm is based on the improvement of the matched tracking algorithm, and the two adopt the same method of selecting atoms, which is to select the atom that matches with the observed signal or iterative residuals the best atom at the time. The innovation of the OMP algorithm is that it adds the step of Schmidt orthogonalization of the selected atoms, which projects the signal onto the space formed by the orthogonalized atoms in order to obtain the projected and residual components, which ensures that the selection of the atoms is not repeated, and thus guarantees the validity of the iterations. As the number of iterations increases, the residual component decreases and the original signal can be approximated with high accuracy using only a limited number of atoms. This orthogonalization speeds up the convergence of the algorithm and improves the reconstruction quality of the signal, which can be accurately represented by a sufficiently small number of atoms.

The main steps of the OMP algorithm are described below:

1) Initialization. Approximation coefficient $\hat{x}_0=0$, residuals $r_0=y$, candidate atom index set $I_0=\phi$, number of iterations $k=0$.

2) Search for matching atoms. $k=k+1$, search for the atom most relevant to the residual and add the corresponding index to the candidate atom index set I_k.

3) Update decomposition coefficients and residuals. For the new candidate atom index set I_k, calculate the new decomposition coefficients and new residuals.

4) Judge whether the stopping condition $(k>K)$ is satisfied. If the stopping condition is not satisfied, return to step (2), if the stopping condition is satisfied, output the sparse approximation signal $\hat{x}$ of x, i.e., the sparse feature vector.

The core idea of the OMP algorithm is that in the case of the predicted sparsity K, in the iterative process, the atom in the measurement matrix that is most relevant to the observed signal or iterative residuals, i.e., the one with the largest absolute value of the inner product, is added to the atom set, and its corresponding index value, i.e., the column number, is added to the support set, and the atom set and the support set are continuously expanded in many iterations to estimate the reconstructed value of the signal until the stop condition is satisfied at the end of the algorithm.

4. Analysis of feature extraction results

4.1. Comparative analysis of model performance

In the data acquisition process of 10 non-electrical signal parameters constructed for the key components of the high-voltage DC converter valve, the N value represents the length of each segment of information and the M value represents the number of measurement points under sampling. When the N is constant and the M value decreases, the data represent the same length of information, the number of measurement points decreases, and the average sampling frequency decreases with M decreasing. When the value of M is constant and the value of N increases, the number of measurement points remains constant, the data represents an increase in the length of the message, and the average sampling frequency decreases as N increases.

The feature extraction and pattern recognition are carried out respectively on the data of different lengths obtained by the two sampling methods, the data length ranges from 32 to 800 points, and the recognition rate is calculated at an interval of 4 data points each, and the results obtained are shown in Fig. 4. According to the figure, it can be seen that as the data length decreases to a certain extent, after a certain degree, the statistical feature quantity as a feature parameter will not be able to completely express the characteristics of a signal, the feature extraction effect will be greatly reduced, and the recognition rate begins to decline significantly. When the number of measurement points is between 200 and 770, the compressed perceptual feature extraction algorithm is basically the same as the traditional method, at about 95%. When the number of measurement points is between 64 and 200, the length of the signal obtained by the traditional sampling method cannot meet the needs of feature extraction, and the length of the data collected by applying the polynomial matrix sparse coding based compressed perceptual feature extraction method carries more information, and its feature recognition rate is higher than that of the traditional sampling method, which is basically the same as that of the previous recognition rate. When the data length is reduced to 0-32 points, the recognition rate obtained by the polynomial matrix sparse coding algorithm also gradually starts to decrease due to the insufficient amount of information, and the recognition rate obtained by the traditional method still continues to decrease. It shows that the application of compressed perceptual feature extraction algorithm can realize the recognition task with fewer sampling points.

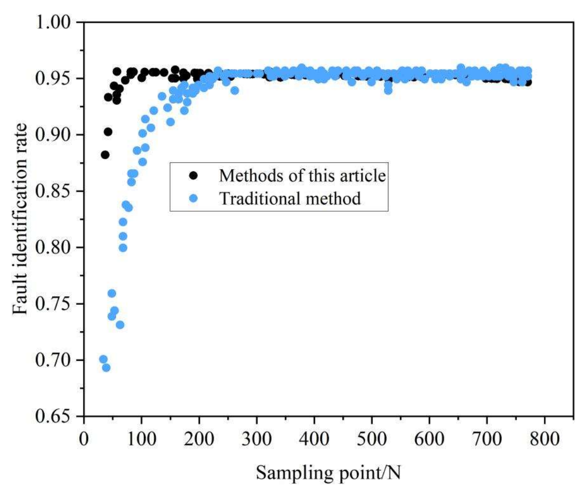


Fig. 4. Fault identification rate of two methods

In order to further verify the advantages of the compressed perceptual feature extraction algorithm, a longitudinal comparison is made, and the effects of different information lengths N on the recognition rate are shown in Table 6 when the number of measurement points M is determined. When the traditional sampling method is used to collect data and the number of measurement points is 64, the recognition rate is only about 80% because the signal length is not enough to express the features accurately. When the data is collected using the compressed perceptual feature extraction algorithm, when the number of measurement points is also 64, the recognition rate is approximately the same as that obtained using the ordinary collection method, indicating that no matter which sampling method is used, the information can not be extracted sufficiently when the length of the information is insufficient. As the information length grows, within a certain compression ratio, the number of measurement points remains constant, and the recognition rate gradually increases as the value of N gradually increases, and the final recognition rate increases to about 93%. with the value of M unchanged, the sampling time of the same information length gradually extends as the value of N increases, and the average sampling rate over a period of time gradually decreases, indicating that the application of the compression-aware feature extraction method based on multi-polynomial matrix sparse coding can effectively reduce the number of features to achieve the task of feature extraction, and the average sampling rate over a period of time decreases. realizing the feature extraction task can effectively reduce the average sampling rate.

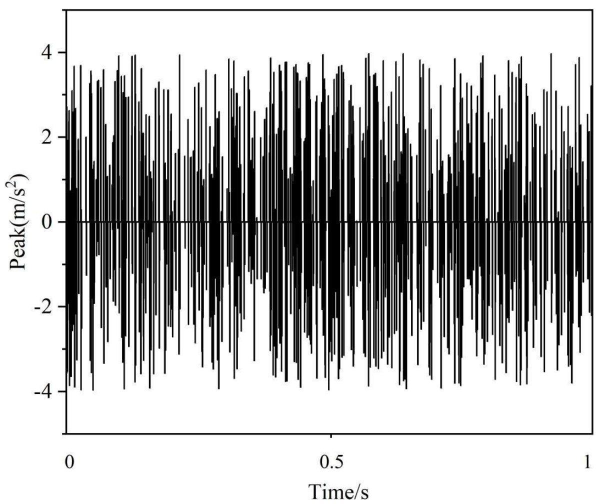
Table 6. Comparison of fault identification rate distribution

Measurement point		Sample	Training sample	Test sample	Identification sample	Recognition rate
Traditional method	64	500	250	250	198	79.2%
	64	500	250	250	204	81.6%
	80	500	250	250	206	82.4%
This method	96	500	250	250	213	85.2%
	128	500	250	250	231	92.4%
	256	500	250	250	231	92.4%
	320	500	250	250	235	94.0%
	384	500	250	250	234	93.6%
	512	500	250	250	232	92.8%
	768	500	250	250	232	92.8%

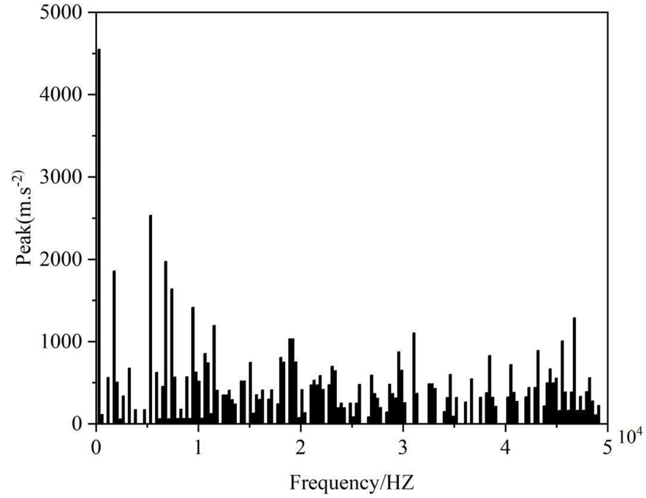
4.2. Experimental signal analysis of the converter valve

In order to further verify the applicability of the proposed dataset construction method and the effectiveness of the compression domain feature extraction method, the measured polynomial matrix sparse coding algorithm commutator valve signals adopted are compressed and feature extracted. The N205M converter valve is taken as the test object and experiments are carried out on the experimental bench.

Taking the vibration signal as an example, the signal with a time length of 1s is intercepted from the acquired vibration signal for analysis, and the time domain waveform and frequency spectrum of the obtained signal are shown in Figure 5, with Figure (a) as the time domain and Figure (b) as the frequency. As can be seen from the figure, the time-domain signal has no obvious shock component due to strong noise interference, and there is no obvious resonance peak in the frequency spectrum. Therefore, the compressed perception algorithm based on polynomial matrix sparse coding can effectively extract signal features.



(a)Signal time domain



(b)Signal spectrum

Fig. 5. The time domain and spectrum of vibrating signals

5. Conclusion

In order to extract the feature data of non-electrical signal key parameters of the key elements of the high-voltage DC converter valve, this thesis proposes a compressed feature perception algorithm based on polynomial matrix sparse coding. As shown by experimental data, a data set containing 10 state parameters is established by comprehensively considering various types of parameters affecting the aging state of the key element of the converter valve. When the number of measurement points is between 64 and 200, the polynomial matrix sparse coding compressed feature perception algorithm constructed by this method has a higher recognition rate than the traditional method, and completes the feature extraction task with fewer sampling points. With the number of measurement points determined, the final recognition rate of this paper's algorithm reaches about 93% as the value of information length gradually increases. Empirical analysis with the collected vibration signals reveals that the time-domain signal of the N205M converter valve has no obvious shock components, and there are no obvious resonance peaks in the frequency spectrum.

Funding

This work was supported by the Science and Technology Project of China Southern Power Grid Company Limited under (CGYKJXM20220335).

References

[1] Adam, G. P., Gowaid, I. A., Finney, S. J., Holliday, D., & Williams, B. W. (2016). Review of dc–dc converters for multi-terminal HVDC transmission networks. IET Power Electronics, 9(2), 281-296.

[2] May, T. W., Yeap, Y. M., & Ukil, A. (2016, November). Comparative evaluation of power loss in HVAC and HVDC transmission systems. In 2016 IEEE Region 10 Conference (TENCON) (pp. 637-641). IEEE.

[3] Muniappan, M. (2021). A comprehensive review of DC fault protection methods in HVDC transmission systems. Protection and Control of Modern Power Systems, 6(1), 1-20.

[4] Gao, S. P., Liu, Q., & Song, G. B. (2017). Current differential protection principle of HVDC transmission system. IET Generation, Transmission & Distribution, 11(5), 1286-1292.

[5] Kalair, A., Abas, N., & Khan, N. (2016). Comparative study of HVAC and HVDC transmission systems. Renewable and Sustainable Energy Reviews, 59, 1653-1675.

[6] Sun, J., Li, M., Zhang, Z., Xu, T., He, J., Wang, H., & Li, G. (2017). Renewable energy transmission by HVDC across the continent: system challenges and opportunities. CSEE Journal of Power and Energy Systems, 3(4), 353-364.

[7] Alassi, A., Bañales, S., Ellabban, O., Adam, G., & MacIver, C. (2019). HVDC transmission: Technology review, market trends and future outlook. Renewable and Sustainable Energy Reviews, 112, 530-554.

[8] Bang, S., Kim, H. S., Koo, J. H., & Lee, B. W. (2021). Consideration of the insulation design method on a±200 kV converter valve unit in an HVDC converter hall. Energies, 14(8), 2296.

[9] Kumar, A., Jhampati, S., & Suri, R. (2017, December). HVDC converter stations design for LCC based HVDC transmission system-key consideration. In 2017 14th IEEE India Council International Conference (INDICON) (pp. 1-6). IEEE.

[10] Lee, C. H., Kwak, J. S., & Kim, S. W. (2020, October). Verification on Valve Losses of LCC HVDC Converter Station in Korea. In 2020 3rd International Conference on Power and Energy Applications (ICPEA) (pp. 46-51). IEEE.

[11] Judge, P. D., Chaffey, G., Wang, M., Zerihun Dejene, F., Beerten, J., Green, T. C., ... & Leterme, W. (2019). Power-system level classification of voltage-source HVDC converter stations based upon DC fault handling capabilities. IET Renewable Power Generation, 13(15), 2899-2912.

[12] Jung, J. H., Nho, E. C., Chung, Y. H., Baek, S. T., & Lee, J. H. (2017). Test circuit for MMC-based VSC valves in HVDC power station. Electronics Letters, 53(4), 272-273.

[13] Anuoluwapo, G. P. (2018). ‘Power losses in HVDC converter stations. University of Kwazulu-Natal, Dissertation.

[14] Zhu, R., Lin, N., Dinavahi, V., & Liang, G. (2019). An accurate and fast method for conducted EMI modeling and simulation of MMC-based HVdc converter station. IEEE Transactions on Power Electronics, 35(5), 4689-4702.

[15] Bisewski, B., & Hu, J. (2024). Auxiliary Power Systems for HVDC Converter Stations. In High Voltage DC Transmission Systems: HVDC (pp. 1-29). Cham: Springer International Publishing.

[16] Yin, S., Li, X., Li, J., Liu, D., & Cai, Z. (2021). Real-time virtual measurement technique for the thyristor valve current based on the converter terminal current. IEEE Transactions on Power Delivery, 37(5), 3447-3457.

[17] Gou, R. (2019). Research on 1100-kV/5500-A ultra-high voltage thyristor valve key technology and its application. IEEE Transactions on Power Electronics, 34(11), 10524-10533.

[18] Yue, K., Pang, L., You, H., Li, S., Kong, D., Li, Y., ... & Liu, L. (2017). Reverse recovery characteristics of high power thyristors in HVDC converter valve. IEEE Transactions on Dielectrics and Electrical Insulation, 24(4), 2132-2140.

[19] Caiwang, S. H. E. N. G., Juanjuan, Z. H. A. N. G., Jie, L. I. U., Tingting, L. I., Jun, Y. A. N. G., & Libo, S. U. N. (2022). Thyristor Turn-off Time Under Actual Operational Condition in UHVDC Project Converter Valve. Power System Technology, 46(8), 3076-3083.

[20] Yao, W., Liu, Y., Guo, L., Ding, G., Zhang, C., & Wang, S. (2024). Research on equivalent circuit model of HVDC valve and calculation of thyristor junction temperature. Active and Passive Electronic Components, 2024(1), 6671153.