Article

A Feature Extraction Method Using Auditory Nerve Response for Collapsing Coal-Gangue Recognition

Huadong Pang^{1,2}, Shibo Wang¹, Xijie Dou¹, Houguang Liu¹ , Xu Chen¹, Shanguo Yang^{1,*} , Teng Wang² and Siyang Wang¹

¹ School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou 221116, China; panghuadong@cumt.edu.cn (H.P.); wangshb@cumt.edu.cn (S.W.); douxijie@cumt.edu.cn (X.D.); liuhg@cumt.edu.cn (H.L.); chenx@cumt.edu.cn (X.C.); wangsiyang@cumt.edu.cn (S.W.)

² Sun Yueqi Honors College, China University of Mining and Technology, Xuzhou 221116, China; wangtengttt@cumt.edu.cn

* Correspondence: ysgcumt@cumt.edu.cn

Received: 29 September 2020; Accepted: 22 October 2020; Published: 23 October 2020

Abstract: To intelligentize the top-coal caving's process, many data-driven coal-gangue recognition techniques have been proposed recently. However, practical applications of these techniques are hindered by coal mine underground's high background noise and complex environment. Considering that workers distinguish coal and gangue by hearing the impact sounds on the hydraulic support, we proposed a novel feature extraction method based on an auditory nerve (AN) response model simulating the human auditory system. Firstly, vibration signals were measured by an acceleration sensor mounted on the back of the hydraulic support's tail beam, and then they were converted into acoustic pressure signals. Secondly, an AN response model of different characteristic frequencies was applied to process these signals, whose output constituted the auditory spectrum for feature extraction. Meanwhile, a feature selection method integrated with variance was used to reduce redundant information of the original features. Finally, a support vector machine was employed as the classifier model in this work. The proposed method was tested and evaluated on experimental datasets collected from the Tashan Coal Mine in China. In addition, its recognition accuracy was compared with other coal-gangue recognition methods based on commonly used features. The results show that our proposed method can reach a superior recognition accuracy of 99.23% and presents better generalization ability.

Keywords: top-coal caving; coal-gangue recognition; auditory nerve response; auditory spectrum; support vector machine

1. Introduction

There are about 1341 billion tons of proven coal resources in China, of which thick coal seam accounts for 44 percent [1]. The longwall top-coal caving (LTCC) mining method, which makes the upper part of the coal seam collapse under the action of gravity, is an efficient and productive method for mining thick coal seams [2]. During the top-coal caving mining process, the workers control the hydraulic support's window to cave the upper part coal based on the recognition of coal and gangue. Therefore, coal-gangue recognition (CGR) is one of key techniques to intelligentize the process of top-coal caving [3].

Since the 1960s, over 20 types of coal-gangue interface identification methods have been developed. Representative methods for coal and gangue interface identification contain artificial γ -ray, nature γ -ray [4,5], and radar [6]. The artificial γ -ray method measures the top coal's thickness and recognizes the interface between coal and gangue by back scattering. However, researchers have progressively

abandoned the artificial γ -ray method because it is harmful and has limited penetration [5]. By detecting the gamma ray energy through the remaining coal seam in the roof, a natural γ -ray detection method [7] can measure the thickness of the coal seam, according to the law of its intensity attenuation. Adopting the natural γ -ray detection method is relatively mature for coal-gangue interface recognition. However, the drawback of such a method is the high detection cost [5]. In addition, when the top coal contains few or no radioactive elements or consists of exorbitant gangue, natural γ -ray detection is not practical. In order to measure the top coal's thickness, radar detection [6] uses electromagnetic waves' reflection at the coal-gangue interface. Although radar detection has the advantage of a wide application range, such a method is not useful when the top coal is excessively thick.

The problems of high cost of equipment and complex equipment operation impede the practical application of the above methods, and hence some scholars have explored recognition methods based on acoustic and vibration signals to identify coal and gangue. Primarily, three methods of coal-gangue identification have been investigated for top-coal caving mining. One method is based on acoustic signals acquisition, the second method is based on vibration signals technique, the last method is based on a combination of sound and vibration signals. In the first method, a coal-gangue interface identification method based on the mel-frequency cepstrum coefficient (MFCC) and back propagation (BP) neural network is introduced in top-coal caving using acoustic sensors mounted on the tail beam of a hydraulic support [8]. Furthermore, Li et al. [9] analyzed the permutation entropy obtained by the local mean decomposition of coal and gangue acoustic signals. In the second method, Wang et al. proposed a vibration model of the tail beam using time series analysis models and employed the power spectrum as a feature to recognize coal and gangue [10]. Moreover, on the basis of vibration signals, Liu et al. [11] and Zhang et al. [5] recognized coal and gangue by applying information entropy of Hilbert spectrum and stacked sparse autoencoders, respectively. Hua et al. [12] extracted dimensionless parameters, i.e., peak factor, margin factor, and another seven features as feature vector to classify the coal and gangue. In the last method, Wang et al. extracted vibration and acoustic signal features by employing multiple signal processing methods and adopted the multiclass F-score (MF-score) feature dimension reduction method to eliminate the redundant information of those features in top-coal caving [13].

Although the above studies, based on acoustic and vibration signals, have achieved good results under certain conditions, practical applications in the top-coal caving process are hindered by the high background noise and the complex environment of an underground coal mine. Up to now, it still depends on workers to distinguish coal and gangue by hearing impact sounds of the hydraulic support. Considering that the recognition ability of the human auditory system shows superior robustness as compared with conventional automatic recognition technology [14], we propose a new feature extraction technology integrated with an auditory nerve (AN) response model, which simulates the human auditory system.

The primary contributions of the work can be summarized as follows:

(1) Inspired by the important roles of auditory in coal-gangue recognition during an actual top-coal caving process, we proposed an original feature extraction method integrated with an auditory nerve (AN) response model and auditory spectrum. In this method, the AN response model was used to process coal and gangue's acoustic pressure signals, and each characteristic frequency of such model's output constituted the auditory spectrum. The distribution properties of the auditory spectrum were analyzed, and the effective statistical features were extracted.

(2) Because each type of feature behaves differently at various frequencies, we proposed a feature selection method integrated with variance to reduce noise and redundant information of original feature set and verified the effectiveness of this dimensionality-reduction method.

(3) Through the analysis of the AN-based support vector machine (SVM) recognition results and comparison experiments, we demonstrated that more accuracy and generality than traditional extracting feature methods on the basis of signal processing have been achieved in this proposed method.

The rest of this paper is organized as follows: In Section 2, we illustrate vibrational signals' acquisition and preprocessing; in Section 3, we introduce an AN response model and feature extraction and selection based on auditory spectrum; in Section 4, we describe the experimental results of the proposed method and its prediction results are compared with that of the recognition methods based on other commonly used features; finally, the conclusions we reached are described in Section 5. Figure 1 shows the proposed coal-gangue recognition framework on the basis of AN response.

Figure 1. The proposed coal-gangue recognition framework based on the auditory nerve (AN) response.

2. Signal Acquisition and Preprocessing

2.1. Signal Acquisition

Underground noise, such as the noise generated by the coal shearer cutting coal, pollute the vibration signal of the coal and gangue and increase the difficulty of coal-gangue identification. It is important to fix the sensor on a position that can avoid the shearer's noise in the longwall working face. Hence, we installed the vibration sensor on the back of hydraulic support's tail beam, which was also sensitive to the impact of the collapsed coal and gangue on the hydraulic support.

The experiment was conducted on the No. 8222 working platform in the Tashan Coal Mine, as shown in Figure 2. The data collection procedure was as follows: First, an integrated electronics piezoelectric (IEPE) acceleration sensor detected the vibrational signal of the tail beam; then, the vibrational signal transmitted by the data acquisition cable was collected by a signal acquisition system (DH5925N, Donghua Testing Technology Co., Shanghai, China). The signals were obtained

with sampling frequency of 12.8 kHz. The collected signals were segmented into 0.25 s for analysis. Figure 3 shows part of the coal and gangue vibrational signals.

Figure 2. The coal-gangue vibration data acquisition experiment conducted in the Tashan Coal Mine. (a) Schematic view of the experiment; (b) Components of the experimental platform; (c,d) Signal acquisition process (signals are transmitted from the acceleration sensor to signal acquisition system via data acquisition cable).

Figure 3. Coal and gangue vibrational signal. (a) Coal vibrational signal; (b) Gangue vibrational signal.

2.2. Signal Preprocessing

Because coal falling is a continuous process, the last part of the collected vibration signal is gangue by default, and the front part is coal. The vibration signal of coal and gangue is decomposed into a series of subsignals of the same length. In order to meet the requirements of the AN response model for processing vibration signals, the trapezoidal integral and sound pressure calculation formula were used to convert the collected vibrational signals into sound pressure signal, as given in Figure 4. The trapezoidal integral and sound pressure calculation formulas are shown as follows:

$$\begin{cases} a^* = \frac{a-\mu}{\sigma} \\ v_i = \frac{(a_i^* + a_{i+1}^*)\Delta t}{2} \\ P = \rho cv \end{cases} \quad (1)$$

where $P_{ref} = 2 \times 10^{-5}$ Pa and a denotes the sensor signal; μ and σ denote mean and variance of sensor signals, respectively; v_i denotes the result of trapezoidal integration of sensor signal; Δt denotes the sampling interval during signal acquisition; ρ and c are constants, representing the density of the medium and the speed of sound, respectively; and P denotes the sound pressure to be calculated.

Figure 4. Acoustic pressure signal. (a) Coal acoustic pressure signal; (b) Gangue acoustic pressure signal.

3. Feature Extraction and Selection

3.1. AN Response Model

The AN response model is a phenomenological model that can simulate the human auditory system. Each section provides a phenomenological description of the primary functional parts of the auditory periphery, from the middle ear to the AN. In this paper, we used Bruce et al.'s model of the auditory periphery [15], which is one of the most comprehensive auditory models, to calculate the AN response. Briefly, four sections of this auditory model are mainly used, i.e., the middle ear filter, C1 and C2 filter, outer hair cell section, and inner hair cell section. The schematic diagram of the model is shown in Figure 5.

Figure 5. Schematic diagram of the AN response model [15]. Abbreviations: inner hair cell (IHC), outer hair cell (OHC), low-pass (LP) filter, static nonlinearity (NL), characteristic frequency (CF), and inverting nonlinearity (INV).

The stimulus' instantaneous pressure signals (in Pa) are firstly input into the middle ear (ME) filter. The model's ME section is implemented by a series of three digital filters [16]:

$$\begin{cases} ME_1(z) = 0.0127 \left(\frac{1.0000 + 1.0000z^{-1}}{1.0000 - 0.9986z^{-1}} \right) \\ ME_2(z) = 0.0127 \left(\frac{1.0000 + 1.0000z^{-1}}{1.0000 - 0.9986z^{-1}} \right) \\ ME_3(z) = \frac{1.0000 - 1.9943z^{-1} + 0.9973z^{-2}}{1.0000 - 1.9856z^{-1} + 0.9892z^{-2}} \end{cases} \quad (2)$$

Three correlated filter paths, the C1, C2 filters in the signal path simulating frequency selection characteristics of the basilar membrane, and the broadband filter in the feedforward control path, follow the ME filter. There are two transduction functions following the C1 and C2 filters [17,18].

The transduction function of C1 filter:

$$\begin{aligned} V_{ihc,C1} &= A_{ihc}[P_{C1}] \log_{10}(1 + B_{ihc}|P_{C1}|) \\ A_{ihc}[P_{C1}] &= \begin{cases} A_{ihc0} & \text{for } P_{C1} > 0 \\ -\frac{|P_{C1}|^{C_{ihc}+D_{ihc}}}{3|P_{C1}(t)|^{C_{ihc}+D_{ihc}}} A_{ihc0} & \text{for } P_{C1} < 0 \end{cases} \end{aligned} \quad (3)$$

where P_{C1} denotes the output of the signal path C1 filter and A_{ihc0} , B_{ihc} , C_{ihc} , and D_{ihc} represent parameters which are equal to 0.1, 2000, 1.74, 6.87×10^{-9} , respectively.

The transduction function of C2 filter:

$$\begin{aligned} V_{ihc,C2} &= \text{signum}(-P_{C2}) \times 2 \times A_{ihc0} \times \log_{10}(1.0 + 5 \times 10^{-5} B_{ihc} \times (P_{C2} \cdot CF)^2) \\ \text{signum}(-P_{C2}) &= \begin{cases} -1 & \text{for } P_{C2} \geq 0 \\ 1 & \text{for } P_{C2} < 0 \end{cases} \end{aligned} \quad (4)$$

where P_{C2} represents the signal path C2 filter's output; CF represents the characteristic frequency; and A_{ihc0} and B_{ihc} denote the parameters which are equal to 0.1 and 2000, respectively.

The feedforward control path is composed of broadband filter and outer hair cell section including low-pass filter, static nonlinearity, and scale zooming, which can adjust the gain and bandwidth of the C1 filter to reflect several level-dependent characteristics in the cochlea [16,19].

The two transduction functions' joint responses after the C1 and C2 filters are input into a seventh-order IHC low-pass filter to obtain the result of the inner hair cell section with a single characteristic frequency [17,18]. Finally, a series of the inner hair cell model's outputs corresponding to all characteristic frequencies can be combined to form the auditory spectrum.

Detailed information of the model segments can be found in Zilany and Bruce et al. [15,17,20].

3.2. Feature Extraction

3.2.1. Auditory Spectrum

In this paper, the signals are obtained with a sampling frequency of 12.8 kHz. According to the sampling theory, the available signal frequency ranges from 0 to 6.4 kHz. Hence, the range of characteristic frequencies of the AN response model can be set as 0.2–6 kHz, which are divided into 20 characteristic frequency by logarithmic spacing. The corresponding characteristic frequencies are shown in Table 1.

Table 1. Characteristic frequencies and their corresponding serial number.

Serial Number	Characteristic Frequency/Hz	Serial Number	Characteristic Frequency/Hz
1	200	11	1198.01
2	239.21	12	1432.87
3	286.10	13	1713.75
4	342.18	14	2049.71
5	409.26	15	2451.52
6	489.49	16	2932.10
7	585.44	17	3506.89
8	700.22	18	4194.35
9	837.48	19	5016.58
10	1001.66	20	6000

We apply the AN response model to process the sound pressure signals of coal and gangue and the auditory spectrum, consisting of each characteristic frequency of the AN response model's outputs, is obtained. Figures 6 and 7 represent the auditory spectrum of coal and gangue.

From Figures 6 and 7, we can see that the auditory signals at each characteristic frequency are significantly different. The auditory spectrum of coal reaches maximum at 200 Hz and as the characteristic frequency increases, the signal amplitude decreases progressively. The auditory spectrum of gangue reaches its greatest value at 286.1 Hz. Although both spectrums reach their maximum at low frequencies, the output voltage of gangue signal's spectrum is more than that of coal signal's spectrum. In addition, within 409.26–2932.1 Hz, the auditory spectrum of the coal sample is similar to that of the gangue sample but the whole trend of coal has increasing overall fluctuation as compared with that of gangue.

Figure 6. The coal signal's auditory spectrum.

Figure 7. The gangue signal's auditory spectrum.

3.2.2. Feature Extraction Based on Auditory Spectrum

In pattern recognition, it is significant to extract effective features which are significant to boost the recognition accuracy. According to the analysis of time-frequency distribution characteristics of the auditory spectrum, statistical features including auditory spectrum energy (ASE), auditory spectrum energy moment (ASEM), auditory spectrum energy entropy (ASEE), skewness, kurtosis, and variance are extracted on the basis of auditory spectrum. These statistical characteristics are as follows:

(1) Auditory spectrum energy (ASE) can intuitively reflect the distribution characteristics of energy under the characteristic frequency. The energy of auditory signals at each characteristic frequency E_i is:

$$E_i = \sum_{t=1}^N |x_i(t)|^2 \quad (5)$$

where $x_i(t)$ represents the auditory signal at the i -th characteristic frequency, N represents the time span of such auditory signal. Moreover, ASE is normalized to better highlight the difference in auditory spectrum distribution property as:

$$ASE = [E_1, E_2, \dots, E_M] / \sum_{i=1}^M E_i \quad (6)$$

(2) Auditory spectrum energy moment (ASEM) demonstrates both energy and distribution characteristics of auditory signals. The energy moment of signals at each characteristic frequency $ASEM_i$ is:

$$ASEM_i = \sum_{t=1}^N \frac{t}{f_s} \cdot |x_i(t)|^2 \quad (7)$$

where f_s represents the sampling frequency of vibration signal, which is equal to 12.8 kHz.

(3) Auditory spectrum energy entropy (ASEE) reflects the chaotic degree of energy distribution in the characteristic frequency. The energy distribution gradually tends to be uniform with an increase in energy entropy and a decrease in energy entropy will lead to the disordered energy distribution in this characteristic frequency. The energy entropy of signals at each characteristic frequency $ASEE_i$ is:

$$ASEE_i = - \sum_{t=1}^N \frac{|x_i(t)|^2}{E_i} \cdot \ln\left(\frac{|x_i(t)|^2}{E_i}\right) \quad (8)$$

where E_i represents the energy of i -th auditory signal.

(4) Skewness represents the degree of skewness in auditory signal distribution within each characteristic frequency, which can be calculated as follows:

$$Skewness_i = \frac{E(x_i(t) - \mu)^3}{\sigma^3} \quad (9)$$

where μ and σ represent mean value and standard deviation of the auditory signal.

(5) Kurtosis demonstrates the gentleness of the waveform, which can be calculated as follows:

$$Kurtosis_i = \frac{E(x_i(t) - \mu)^4}{\sigma^4} \quad (10)$$

(6) Variance reflects the degree of auditory signal fluctuation, which can be calculated as follows:

$$Variance_i = \sum_{t=1}^N \frac{(x_i(t) - \mu)^2}{N - 1} \quad (11)$$

According to Functions (5) to (11), the features of the auditory spectrum are extracted. Figure 6 shows the normalized comparison of various features extracted from coal and gangue. Each row represents the distribution of a type of feature within different characteristic frequencies, and each column demonstrates the distribution of a different type of feature at a certain characteristic frequency. The features of the auditory spectrum extracted from the acoustic signals which are processed by the auditory nerve response model are quite distinct, as represented in Figures 8 and 9.

Figure 8. Features of coal signals' auditory spectrum.

Figure 9. Features of gangue signals' auditory spectrum.

It shows that the difference between features extracted from coal and gangue auditory signals is obvious within low frequency and high frequency. Within the range of 200–409.26 Hz, ASE, ASEE, and variance of falling gangue samples are greater than those of falling coal samples. When the characteristic frequency ranges from 2932.1 to 6000 Hz, skewness and kurtosis of falling coal samples are greater than that of falling gangue samples, but ASEM is less than that of falling coal samples. When the characteristic frequency is 409.26–2932.1 Hz, the features of coal and gangue are not distinguishable, which is difficult to be used in coal and gangue identification. The above conclusions are consistent with the distribution properties of the auditory spectrum under different characteristic frequencies.

As mentioned above, in terms of a single feature, each type of coal-gangue feature is significantly distinct. Take ASEE as an example, the comparison of such features is shown in Figure 10. As shown in Figure 10, at a low frequency, ASEE of falling gangue samples is significantly greater than that of falling coal samples, which can represent more disorder of gangue energy distribution. Hence, the characteristics of coal-gangue feature are distinguishable and can be applied as the features of coal-gangue recognition.

Figure 10. Comparison of auditory spectrum's auditory spectrum energy entropy (ASEE) at different frequencies.

3.3. Feature Selection

Feature selection can be summarized as a process of selecting several of the most effective feature subsets from the original features set in order to reduce the dimension of dataset. In this section, variance is chosen as the evaluation parameter of features effectiveness because it can reflect the difference among various indicators. The variance of each type of feature within various characteristic frequencies is ranked and the recognition result of SVM model is used as the selection standard.

The detailed feature selection procedures are described as follows:

Step 1. Calculate the variance vector $V = [v(1,1), v(1,2), \dots, v(i,f)]_{1 \times 20}$ where $v(i,f)$, ($i = 1, 2, \dots, 6; f = 1, 2, \dots, 20$), is the variance between the i -th feature's values corresponding to signal of the f characteristic frequency and calculate a variance matrix as:

$$v(i, f) = \sum_{n=1}^N \frac{(F(n, i, f) - \mu)^2}{N - 1} \quad (12)$$

where $F(n, i, f)$ represents the i -th feature value of n -th sample corresponding to f characteristic signal, N represents the number of samples, μ and σ represents mean value and standard deviation of $F(n, i, f)$.

$$V = \begin{bmatrix} v(1,1) & \cdots & v(6,1) \\ \vdots & \ddots & \vdots \\ v(1,20) & \cdots & v(6,20) \end{bmatrix} \quad (13)$$

Step 2. Rank the variance of each type of feature at different characteristic frequencies ($v(1,1)$, $v(1,2)$, $\dots$, $v(1,20)$).

Step 3. The first m ($m = 1, 2, \dots, 20$) optimal characteristic frequencies of each feature are selected each time and processed by the SVM model to obtain the recognition result as the selection criterion. The result is shown in Figure 11.

Figure 11. The change curve of coal-gangue recognition accuracy with an increase in selected features.

In Figure 11, the AN-based SVM recognition accuracy generally increases with the increasing number of features. The maximum recognition rate reached 99.23% when the first 13 optimal characteristic frequencies of each feature are selected. As the number of optimal characteristic frequencies of each feature exceeds 13, with an increase in dimension, some secondary relations, noise, and redundant information of the original auditory signal are introduced, which decline the identification accuracy. When the optimal characteristic frequency of each feature reaches 19, the recognition accuracy achieves the optimal value again, but the increase in feature dimension leads to a decrease in model recognition efficiency.

4. Experimental Results

4.1. Result of AN-Based SVM

The SVM is a machine learning classification method developed by Cortes and Vapnik which is widely adopted for both classification and regression [21]. The binary SVM linearly classifies the sample data in input space by finding an optimal hyperplane to separate the two classes while maximizing generalization. Specifically, given the training data $x_1, \dots, x_n \in R^d$ with labels $y_i = \pm 1$, the purpose of SVM is to detect the best hyperplane that maximizes the class margin among support vectors. Therefore, we used the SVM model to recognize coal and gangue on the basis of the features extracted from the auditory spectrum. In addition, such an SVM model adopts radial basis kernel function (RBF) [22]:

$$\kappa(\mathbf{x}_i \cdot \mathbf{x}_j) = e^{-\|\mathbf{x}_i - \mathbf{x}_j\|_2^2 / (2\sigma^2)} \quad (14)$$

where σ is the width parameter.

The AN-based SVM takes features extracted from the auditory spectrum as the input and takes the labels of coal-gangue samples as output. Both penalty term and the kernel parameter (σ) influencing the performance of SVM are optimized through cross-validation of the grid search method, and then

parameters with the highest cross-validation accuracy are selected. The experiment results of the AN-based SVM are shown in Figure 12 and Table 2, which indicate that the SVM model reaches a testing accuracy of 99.23%.

Figure 12. Receiver operating characteristics (ROC) graphs showing classification performances. (a) Song et al.'s method [13]; (b) Our proposed method.

Table 2. Different feature extraction methods' recognition results.

	Kurtosis	SEI	EEWPC	SC	MMFCC	Song et al. [13]	Proposed Method
Test accuracy	63.79%	46.55%	93.10%	74.14%	81.03%	91.38%	99.23%
AUC	0.7667	0.4560	0.9286	0.7600	0.8131	0.9798	1.0000

4.2. The Result of Analysis

To further assess the proposed method's performance, receiver operating characteristics (ROC) analysis is applied to demonstrate its superiority. ROC analysis is a classification performance evaluation method based on statistical decision theory. The SVM classification method's performance of coal-gangue signals can be assessed according to receiver operating characteristics (ROC) graphs [23]. In addition, such performance is decided by using 10-fold cross-validation, avoiding the bias produced by selecting a given training or test set. The relative balance between true positive rates (benefits) and false positive rates (costs) are shown in the ROC diagrams represented by two-dimensional diagrams [24]. We obtain the ROC diagram by calculating the parameters of sensitivity (Sen) and specificity (Spe). In this article, sensitivity measures the fraction of real positive samples which have been accurately recognized, such as the proportion of coal samples which have been correctly identified as coal. Specificity is a parameter of real negative samples' proportions which have been accurately identified such as gangue samples' percentages which have been accurately identified as gangue.

$$Sen = \frac{TP}{TP + FN} \quad (15)$$

$$Spe = \frac{TN}{TN + FP} \quad (16)$$

where, TP and TN indicate the total number of real positive samples and real negative samples identified correctly, respectively. FN and FP indicate the total number of samples mistakenly treated as negative and positive samples, respectively.

We plot sensitivity and 1-specificity to obtain the ROC diagram, as represented in Figure 12. The advantage of the classifier for recognizing a specific class is demonstrated by its sensitivity to a class of data, and the accuracy of the classifier in not identifying other classes as a particular class can be represented by its 1-specificity to a class of data. Meanwhile, AUC which is the area under the ROC curve is employed to evaluate the detection performance [25], that is, $AUC \in (0,1)$, which can be used to measure pattern recognition methods' performance. The corresponding method is better with an increasing value of AUC.

The test accuracy and AUC of the proposed method are 99.23% and 1.00, respectively. As such, the above evaluation parameters of recognition result confirm that the proposed method performs well at identifying coal and gangue.

4.3. Comparison with Recognition Methods Based on Commonly Used Features

In this section, the performances of commonly used features such as kurtosis, sum energy of IMFs (SEI), energy entropy of wavelet packet coefficient (EEWPC), spectral centroid (SC), and mean of MFCC (MMFCC) are compared with the proposed method to illustrate its superiority in this paper. The relevant outcomes are shown in Figure 13 and Table 2.

Figure 13. Recognition results of different feature extraction methods.

For traditional methods, the training accuracy and test accuracy of energy entropy of wavelet packet coefficient is best, reaching 93.10% and 0.9286, respectively. It means that the recognition accuracy of the time-frequency domain features is significantly higher than that of time domain features and frequency domain features, which is because coal and gangue's vibration signals are non-stationary, and neither time domain features nor frequency domain features can fully demonstrate the subtle difference of collected signals. However, the time-frequency domain features can solve such a problem, thus, achieving higher recognition accuracy. In addition, the mean of MFCC shows good performance in identifying coal and gangue, reaching a recognition accuracy of 81.03%, which means advantages of auditory in recognizing coal and gangue signals. It further verifies the effectiveness of applying the

AN response model to recognize coal and gangue. However, as compared with the mean of MFCC, a higher recognition accuracy is achieved based on auditory spectrum features, which can be explained by the consideration of more auditory characteristics in the AN response model. Although the above features show excellent performance for identifying coal and gangue, their recognition accuracy is still less than that of the auditory spectrum-based features, which means such features extracted by the AN response model is more advantageous.

4.4. Comparison with the Recognition Method Based on MF-Score Feature Reduction

We further compare the proposed method's performance with that of Song et al. [13] who, recently, proposed a method, which was based on feature selection. The method was conducted on the same datasets and classes as in the proposed method. The primary steps of the comparison include preprocessing, which is followed by traditional feature extraction, and automatic selection of features based on the MF-score feature reduction method and coal-gangue recognition based on the SVM classifier. Such extracted features are generated by analyzing signals in domains of time, frequency, or time frequency. Moreover, a feature reduction method on the basis of MF-score, employed to optimize the feature attributes, is achieved based on coal-gangue samples. Then, the information of feature subsets after feature reduction is inputted as SVM classifier's training samples and test samples.

Therefore, residual variance and fractal dimensions are selected to obtain the optimum of feature subset. As shown in Figures 12b and 13 and Table 2, the test accuracy and AUC of the compared methods are 91.38% and 0.98, respectively. The proposed method's test accuracy, test accuracy and AUC are 99.23% and 1.00, respectively. The result shows that the features extracted by the AN response model are more effective and advantageous than those obtained by the MF-score feature reduction method.

On the basis of the above comparison, the superiority of our proposed method can be summarized as follows: The advantageous classification performance has been achieved by using a single type of feature. By imitating the process of classifying coal and gangue in the human ear, the features extracted from the AN response model contained more effective information than those extracted from time, frequency, and time-frequency domains. The limitations of our work include the following: In fully mechanized caving faces, coal mixed with gangue falls down. With the continuous discharge of coal, the proportion of gangue continues to increase up to a certain percentage of gangue. In the experiment, the sample data collected is only divided into coal and gangue, and therefore we do not study the impact properties of different mixing ratios of coal and gangue. Hence, the proposed method is not applicable to all fully mechanized caving faces.

5. Conclusions

In this work, a feature extracted method to identify coal and gangue based on the AN response model and SVM was proposed. By using the vibration signal of hydraulic support's tail beam, the features were extracted using auditory spectrum and consisted of each characteristic frequency of the AN response model's output, which fully reflected the difference between coal and gangue. The feature selection method based on variance was employed to choose an effective feature subset to reduce redundant information. Finally, the selected feature subset was inputted into SVM for coal-gangue identification, and recognition results were further compared with traditional methods. The main contributions of our proposed method are stated as follows:

(1) We proposed an original feature extraction method that integrated the AN response model with auditory spectrum. In this system, the AN response model was applied to process the coal and gangue's acoustic pressure signal to simulate the decision-making judgment of workers' ears during the actual coal mining process. The auditory spectrum, constituted of each characteristic frequency of the AN response model, was used to extract features.

(2) To verify validity of features extracted by the auditory spectrum, each type of feature's behavior at various frequencies was analyzed, and a feature selection method integrated with variance

was developed to decrease the noise and redundant information of the initial feature set based on distribution properties of each feature at different frequencies.

(3) Several traditional methods and a recent coal-gangue recognition method on the basis of feature selection, described in Song et al. [13], were utilized to demonstrate our proposed method's effectiveness. Through a comparison of the recognition results among the features of kurtosis, sum energy of IMFs, energy entropy of wavelet packet coefficient, spectral centroid, mean of MFCC, and MF-score for feature selection, it indicated that the proposed method had outstanding performance. The positive result of this research clearly demonstrated the potential of this proposed approach in coal-gangue identification.

Future work will focus on the impact characteristics of different coal-gangue mixing ratios. Moreover, we also look forward to conducting more studies on improving the AN response model for practical applications.

Author Contributions: Conceptualization, H.L., H.P., and S.Y.; methodology, H.L., X.C., and S.Y.; software, T.W. and S.W. (Siyang Wang); validation, H.P., X.D., and S.W. (Shibo Wang); writing, H.P. All authors have read and agreed to the published version of the manuscript.

Funding: The National Key Research and Development Program of China (grant number 2018YFC0604503) and the Priority Academic Program Development of Jiangsu Higher Education Institutions funded this work.

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

References

1. Wang, J.; Yu, B.; Kang, H.; Wang, G.; Mao, D.; Liang, Y.; Jiang, P. Key technologies and equipment for a fully mechanized top-coal caving operation with a large mining height at ultra-thick coal seams. *Int. J. Coal Sci. Technol.* **2015**, *2*, 97–161. [\[CrossRef\]](#)
2. Alehossein, H.; Poulsen, B.A. Stress analysis of longwall top coal caving. *Int. J. Rock Mech. Min. Sci.* **2010**, *47*, 30–41. [\[CrossRef\]](#)
3. Sun, J.; Su, B. Coal-rock interface detection on the basis of image texture features. *Int. J. Min. Sci. Technol.* **2013**, *23*, 681–687. [\[CrossRef\]](#)
4. Wang, Z.C.; Wang, R.L.; Xu, J.H. Research on coal seam thickness detection by natural Gamma ray in shearer horizon control. *J. China Coal Soc.* **2002**, *27*, 425–429.
5. Zhang, G.; Wang, Z.; Zhao, L. Recognition of rock-coal interface in top coal caving through tail beam vibrations by using stacked sparse autoencoders. *J. Vibroeng.* **2016**, *18*, 4261–4275.
6. Strange, A.D. Robust Thin Layer Coal Thickness Estimation Using Ground Penetrating Radar. Ph.D. Thesis, Queensland University of Technology, Brisbane City, Australia, 2007.
7. Wang, Z.; Fu, Q.; Xie, Z. Research on the development of rock-coal interface monitor in top-coal caving mining. In Proceedings of the ICMIT 2005: Information Systems and Signal Processing, Chongqing, China, 20 February 2006; p. 60411W.
8. Junkai, X.; Zengcai, W.; Wanzhi, Z.; Yanpeng, H. Coal-rock interface recognition based on MFCC and neural network. *Int. J. Signal Process. Image Process. Pattern Recognit.* **2013**, *6*, 191–200.
9. Li, Y.; Cheng, G.; Chen, X.; Liu, C. Coal-rock interface recognition based on permutation entropy of LMD and supervised Kohonen neural network. *Curr. Sci.* **2019**, *116*, 96. [\[CrossRef\]](#)
10. Wang, B.; Wang, Z.; Zhu, S. Coal-rock interface recognition based on time series analysis. In Proceedings of the 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), Taiyuan, China, 22–24 October 2010; pp. V8-356–V8-359.
11. Liu, W.; He, K.; Liu, C.-Y.; Gao, Q.; Yan, Y. Coal-gangue interface detection based on Hilbert spectral analysis of vibrations due to rock impacts on a longwall mining machine. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2015**, *229*, 1523–1531. [\[CrossRef\]](#)
12. Hua, W.; Zhao, X.; Luo, C.; Xue, G.; Wu, M. Coal-rock interface recognition method based on dimensionless parameters and support vector machine. *Electron. J. Geotech. Eng.* **2016**, *21*, 5477–5486.
13. Song, Q.; Jiang, H.; Zhao, X.; Li, D. An automatic decision approach to coal-rock recognition in top coal caving based on MF-Score. *Pattern Anal. Appl.* **2017**, *20*, 1307–1315. [\[CrossRef\]](#)

14. Li, Z.; Gao, Y. Acoustic feature extraction method for robust speaker identification. *Multimed. Tools Appl.* **2016**, *75*, 7391–7406. [[CrossRef](#)]
15. Zilany, M.S.; Bruce, I.C.; Nelson, P.C.; Carney, L.H. A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-Term adaptation with power-law dynamics. *J. Acoust. Soc. Am.* **2009**, *126*, 2390–2412. [[CrossRef](#)] [[PubMed](#)]
16. Bruce, I.C.; Sachs, M.B.; Young, E.D. An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. *J. Acoust. Soc. Am.* **2003**, *113*, 369–388. [[CrossRef](#)] [[PubMed](#)]
17. Zilany, M.S.; Bruce, I.C. Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery. *J. Acoust. Soc. Am.* **2006**, *120*, 1446–1466. [[CrossRef](#)]
18. Zilany, M.S.; Bruce, I.C. Representation of the vowel/ε/in normal and impaired auditory nerve fibers: Model predictions of responses in cats. *J. Acoust. Soc. Am.* **2007**, *122*, 402–417. [[CrossRef](#)]
19. Zhang, X.; Heinz, M.G.; Bruce, I.C.; Carney, L.H. A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J. Acoust. Soc. Am.* **2001**, *109*, 648–670. [[CrossRef](#)]
20. Zilany, M.S.; Bruce, I.C.; Carney, L.H. Updated parameters and expanded simulation options for a model of the auditory periphery. *J. Acoust. Soc. Am.* **2014**, *135*, 283–286. [[CrossRef](#)]
21. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
22. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
23. Çinarer, G.; Emiroğlu, B.G.; Yurttakal, A.H. Prediction of Glioma Grades Using Deep Learning with Wavelet Radiomic Features. *Appl. Sci.* **2020**, *10*, 6296. [[CrossRef](#)]
24. Adil, M.; Javaid, N.; Qasim, U.; Ullah, I.; Shafiq, M.; Choi, J.-G. LSTM and Bat-Based RUSBoost Approach for Electricity Theft Detection. *Appl. Sci.* **2020**, *10*, 4378. [[CrossRef](#)]
25. Jadin, M.S.; Taib, S.; Ghazali, K.H. Feature extraction and classification for detecting the thermal faults in electrical installations. *Measurement* **2014**, *57*, 15–24. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).