Research Article

Denoising Method Based on Spectral Subtraction in Time-Frequency Domain

Lei Hao,1 Shuai Cao,1 Pengfei Zhou,2 Lei Chen3,4, Yi Zhang3,4, Kai Li,1 Dongdong Xie,2 and Yijun Geng5

1School of Qilu Transportation, Shandong University, Jinan 250002, China
2Shandong Hi-speed Group Co., Ltd., Jinan 250098, China
3Geotechnical and Structural Engineering Research Center, Shandong University, Jinan 250061, China
4School of Civil Engineering, Shandong University, Jinan 250061, China
5Yellow River Survey, Planning, Design and Research Institute Co., Ltd., Zhengzhou 450003, China

Correspondence should be addressed to Lei Chen; clei667@163.com and Yi Zhang; zy15522208076@163.com

Received 2 December 2020; Revised 9 March 2021; Accepted 29 April 2021; Published 8 July 2021

Academic Editor: Liborio Cavaleri

Copyright © 2021 Lei Hao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In view of the key problem that a large amount of noise in seismic data can easily induce false anomalies and interpretation errors in seismic exploration, the time-frequency spectrum subtraction (TF-SS) method is adopted into data processing to reduce random noise in seismic data. On this basis, the main frequency information of seismic data is calculated and used to optimize the filtering coefficients. According to the characteristics of effective signal duration between seismic data and voice data, the time-frequency spectrum selection method and filtering coefficient are modified. In addition, simulation tests were conducted by using different S/R, which indicates the effectiveness of the TF-SS in removing the random noise.

1. Introduction

The data acquisition with good signal-to-noise ratio (S/R) plays an important role in the exploration of oil and gas reservoirs, providing good basis for imaging underground geology [1]. Similarly, seismic data with high S/R are also the premise of seismic forward-prospecting in the tunnel. However, because of the limited space and construction noise in the tunnel, the seismic data obtained in tunnel environment often contain stronger noise than data collected in ground surface, which affects the extraction of effective signals and leads to anomalies in imaging results. To improve the imaging accuracy, the denoising method is developed. The denoising methods can be divided into time-domain method and transform-domain method. The typical time-domain method consists of convolution filtering, singular value decomposition (SVD), etc., and the transform-domain method consists of F-K filter, Radon transform filter, and time-frequency (TF) filter. For example, Sun [2] proposed an adaptive method for seismic data denoising; Li [3] proposed an adaptive denoising method based on ensemble empirical mode decomposition (EEMD); and Wang [4] used wavelet transform to eliminate and suppress the random noise.

In seismic forward-prospecting in the tunnel, the S/R of data is seriously affected by the environmental noise. It is necessary to study denoising methods for improving imaging results. In addition to band-pass filtering [5, 6], τ-p transformation is widely used to separate effectively reflected waves from the waves with different apparent directions and then achieve data denoising. However, traditional denoising methods are based on the assumption of stable noise, which cannot be suitable for the nonstationary noise in the tunnel well. Evaluation of noise using frequency domain information cannot meet the requirement of denoising in tunnel environment. To solve this problem, time-frequency (TF) information was analysed and introduced in data processing in tunnel environment. For example, band-pass filtering, τ-p filtering, and F-K filtering are used and/or combined to remove noise in seismic ahead-prospecting (SAP), tunnel
reflection tomography (TRT), and so on, further improving the S/R of data.

On this basis, the time-space filtering method was developed and widely used. The typical method is spectral subtraction (SS), which is a transform domain filtering method. The SS has been widely used in enhancing speech signal due to its advantages of good denoising function and fast calculation speed [7, 8]. For the music noise, Berouti et al. [9], Mcaulay and Malpass [10], and Lockwood et al. [11] reduced the music noise in the case of low and medium S/R. Singh and Sridharan [12], Kamath Sunil and Loizou [13], Sim [14], Oh and Lee [15] further reduced the music noise in high S/R by improving the SS. To solve the music noise caused by the location estimation error of effective data and noise distribution, Pascal and Filho [16] and Plapous et al. [17] estimated the data position by prior estimation and voice detection and Zhang et al. [18], Cao [19], Ahmed [20], and Ye [21] used neural network and other methods to detect the effective signal boundary and improve data quality. Because of the similarity between voice signal and seismic detection signal in trigger time, the SS has application potential in seismic data denoising. The seismic data in TF involve more information than those in frequency domain. The TF characteristics are beneficial to the denoising of seismic data. This method also can be useful in global seismology. Chen et al. [22] proposed a framework to preserve amplitude, which is useful for denoising/reconstructing weak seismic arrivals of the USArray data. Thus, a good denoising method could greatly advance the understanding of geological structures in detection.

In actual, important information should be observed both in the time domain and TF transformation [23]. In this paper, a time-frequency spectral subtraction (TF-SS) method for improving seismic data by removing random noise was studied. At first, the TF-SS is introduced into the data processing to improve the S/R by suppressing the noise. Furthermore, the filter window function of TF-SS was optimized by analyzing the dominate frequency of seismic data. According to the characteristics of effective signal duration in seismic data and voice noise, the selection of TF noise spectrum and adaptive filtering coefficient is improved. Finally, the numerical simulations verified the effectiveness of this method.

2. Method

Spectral subtraction (SS) is a denoising method for removing random noise [7]. The key processes are as follows: (1) transform the original data from time-domain to frequency domain; (2) estimate the power spectrum of noise and then subtract it from the power spectrum of original data to obtain the power spectrum of the required signal; and (3) reconstruct the power spectrum of denoised data and the phase spectrum together to obtain time-domain data.

Because of the difference between the basic assumption and the actual situation, “music noise” and “phase noise” will be produced in the process of SS denoising and “music noise” has a bad impact on the processing results. Because of the complexity of the noise, it is difficult to keep the stability in the time domain, which will lead to the error of noise spectrum estimation. To solve this problem, the SS is converted to the time-frequency (TF) domain for processing and the noise spectrum error of SS is corrected better.

The window length, filtering window function, and S/R of the signal have significant impact on the denoising result. Therefore, the TF spectral subtraction was optimized in the following: (1) according to the difference in trigger and duration between seismic data and voice data, the time window selection position of seismic data noise spectrum estimation is adjusted adaptively; (2) based on the dominate frequency of detection data, the filtering window function of seismic forward-prospecting is improved by adopting a weighting coefficient in the frequency domain; and (3) according to the empirical formula of optimal window function size of Gaussian filter, the optimal window length of TF-SS is selected.

2.1. Window Length for Filtering. In order to keep the signal approximately stable within the filtering window range, Lin analysed various parameters, such as the arrival time, waveform, wave crest, trough, and phase of wavelet, to select the filter window with least influence on Ricker wavelet [24]:

$$1 \leq WL \leq \frac{0.384 f_s}{f_d}$$

where WL is the selected window length, $f_s$ is the sampling frequency, and $f_d$ is the main frequency of the effective signal. This equation is used as the basis for the calculation of the filter window length.

2.2. Filtering Based on Dominate Frequency. The Gaussian window function is a kind of traditional filtering window function, which has an excellent side lobe suppression effect and has been widely used in different fields. This section will propose adaptive improvement methods on the basis of the Gaussian window function. The traditional two-dimensional Gaussian filter window function is constructed as follows [25]:

$$G(t, \omega) = Ae^{\left[-\left((t-t_0)^2/2\sigma_t^2+(\omega-\omega_0)^2/2\sigma_\omega^2\right)\right]}, -k \leq t, \omega \leq k,$$

where $\sigma_t$ and $\sigma_\omega$ are the standard deviations of the window function and frequency directions, respectively; $t_0$ and $\omega_0$ are the coordinates of the peak points in the time and frequency directions in each window; and $A$ is the amplitude of the window function. According to equation (3), the amplitude of the traditional two-dimensional Gaussian filtering window function does not change with the change of time or frequency. On this basis, a frequency-related weighting coefficient $\lambda_\omega$ is used to enhance the weight of the data near the main frequency of the signal.

$$G(t, \omega) = A\lambda_\omega e^{\left[-\left((t-t_0)^2/2\sigma_t^2+(\omega-\omega_0)^2/2\sigma_\omega^2\right)\right]}, -k \leq t, \omega \leq k.$$
In equation (4), $\sigma$ is the variance of the weighting coefficient and $\mu$ is the position of the peak point of the weighting coefficient. Among them, $\mu$ and $\sigma$ are the coefficients to be determined additionally. It is planned to establish the relationship between the peak value of the weighted coefficient $\mu$, variance $\sigma$, and wavelet coefficient matrix to achieve adaptive selection.

The numerical value of the wavelet coefficient matrix after TF transformation corresponds to the original signal strength. When the effective signal is stronger than the noise, the peak point $\mu$ of the weighted coefficient can be selected by the maximum value in the wavelet coefficient matrix. After synchronous extrusion wavelet transforms, the original time-domain data can be transformed into the wavelet coefficient matrix in the TF domain, which is divided into frequency segments and N time periods. By calculating the values of each element in the matrix, the element with the largest wavelet coefficient can be selected.

2.3. Flow of Time-Frequency Spectrum Subtraction. The processing flow of TF-SS for seismic data is introduced as follows (as is shown in Figure 1):

1. Input the time-domain seismic records to obtain the sampling frequency information
2. Transform the data into frequency domain and get the dominate frequency
3. Transform the data in the time-domain data into the time-frequency (TF) domain
4. Establish the relationship between the filter window length and the dominate frequency of the signal and then calculate the optimal length of filtering window
5. Estimate the position of the dominate frequency and select the peak position of the Gaussian filter adaptively
6. Based on the Gaussian probability distribution density function and the analysis of seismic data spectrum, the variance of the Gaussian filtering coefficient is adaptively selected
7. Calculate the average value of each frequency band
8. The noise spectrum estimation matrix is subtracted
9. Transforming the wavelet coefficient matrix into the time-domain

3. Numerical Simulations

3.1. Denoising Using Different Lengths of Filtering Window. To study the characteristics of the empirical formula of filtering window length and reveal its influence on the denoising, different window lengths of filtering is selected to denoise the data and summarize the variation law. Figure 2 shows seismic record with Gaussian white noise. The dominate frequencies of these two Rickers are 50 Hz and 150 Hz, respectively. The sampling rate is 1 ms. The optimal window lengths of these two seismic data are 7 and 3, respectively. Figure 3 shows the seismic record comparison between the original data and the denoised data with the window length of 3, 7, and 9, respectively.

According to the denoising results, it is easy to find the following characteristics: (1) when the filtering window is 3, two effective signals have the most accurate amplitude of all; however, the noise still exists in record misleading the effective waves’ recognition and extraction; (2) when the length of the window increases to 7, the amplitude and waveform of the signal with the frequency of 150 Hz change (disagree with the actual waveform), but the noise is suppressed better than using the length of 3; (3) with the length increasing continuously (such as 9), the amplitude and waveform of the two signals (with the frequency of 50 Hz and 150 Hz) decrease together and the denoised signal has the slight more noise than that using the window length of 7. Table 1 shows the S/R of the denoised data using TF-SS. When the filter window length is 3, the denoised data still contain more noise, but the effective signal has the best amplitude of all. When the filter window length is 7, the S/R is the highest of all.

The relationship between the TF window length and the denoising characteristic can be found as follows: (1) when the window length is short, the signal has the relative higher linear degree, which is conducive to reduce the error caused by the nonlinear characteristics and maintain the amplitude, but at the same time, the noise suppression ability is relatively poor; (2) when increasing the window length, the denoising ability is improved, but the amplitude and waveform of effective signal will also be changed; and (3) when the window length is too long, the signal and noise cannot be distinguished.

3.2. Denoising Using Different Frequency Domain Weighted Coefficient. In this section, the linear model and sigmoid model are used to construct a series of datasets with different frequency band by changing the data sampling rate, respectively. Then, the weighted coefficient variance with the best denoising result is selected, and the relationship between the weighted coefficient variance and the number of frequency band is summarized. Typical numerical examples of the linear model and sigmoid model are used to illustrate the process flow as follows.

The seismic record of the linear model is shown in Figure 4. The main parameters used in the linear model are as follows: the sampling rate of 500 Hz, sampling length of 1 s, and dominate frequency of 20 Hz. There are 120 traces in total, and Gaussian white noise with the S/R of 0 dB is added.

Four line events with two intersections are in the linear model (Figure 4(a)), which are parallel but have different first arrival times. The effective wave events are concentrated in the period of 0–1.6 s. The model with Gaussian white noise (S/R of the data is 0) is shown in Figure 4(b), which has large noise interfering signal recognition.

Figure 5 is the TF spectrum of the linear model. In the frequency domain, the energy in the middle part is relatively
concentrated and the main energy intensity is mainly distributed in the range of 20–30 Hz. In the time domain, the signal is relatively concentrated in the range of 0.2–0.5 s. The data in the range of 0.8–1 s is dominated by noise, leading to a relative chaotic time spectrum.

According to the result of denoising by SS (Figure 6(a)), the noise can be suppressed effectively. However, because of the noise in the last 20 ms, many errors will be produced by the nonstationary characteristics of noise, as well as some signals will be wrongly suppressed as noise (Figure 6(b)). The denoising results are shown in Figures 6(c) and 6(d), which indicate that the signal is preserved and the noise is suppressed.

According to the local similarity [26] comparison (Figures 6(e) and 6(f)), the effective signal can be seen vaguely in the result of denoising with SS. However, after denoising by TF-SS, the effective signal is preserved and recognized. Table 2 shows that when the selected weighted variance coefficient is 19, the S/R of the denoised data is relatively high, and the increase or decrease in the variance will lead to the decrease in the S/R. Meanwhile, the root-mean-square error (RMSE) is low when the selected weighted variance coefficient is 19, so the weighted coefficient variance of 19 is initially used as the variance of the empirical formula summarized by the model.

The relationship between the number of data frequency bands and the optimal filtering variance is summarized (Figure 7). In the figure, the blue dot is the calculated value of the variance of the optimal weighting coefficient of different frequency bands’ number and the red line is the function obtained by linear fitting the calculated value.

It can be seen from Figure 8 that the variance of the optimal weighting coefficient is approximately linear with...
the number of frequency bands. The relationship between the variance \( \sigma \) and the number of data frequency bands \( f \) is assumed to be

\[
\sigma = a \times f + b,
\]

(5)

where \( a \) and \( b \) are fixed constant real numbers.

Substituting the straight line fitting results in Figure 7 into equation (5), then the empirical formula can be obtained as follows:

\[
\sigma = 0.063 \times f - 0.169.
\]

(6)

3.3. Comparative Analysis of Denoising Effects of Different S/R

3.3.1. Denoising of Data with S/R = 0 dB. In this section, Gaussian white noise is added to the sigmoid model to form a seismic record with the S/R of 0 dB and then the denoising results (using SS and TF-SS, respectively) were analysed. The comparison of seismic records and local similarities by using the sigmoid model is shown in Figure 7.

As is shown in Figure 7, when the S/R of the data after denoising is 0 dB, the SS method has poor effect of denoising, some of the signals are still treated as noise and removed (Figure 7(d)). The comparison of denoising results shows that the SS method has poorer denoising ability than the TS-SS method for complex seismic records. In a word, TF-SS has great potential in dealing with complex data.

3.3.2. Denoising of Data with S/R = −6 dB. In this section, Gaussian white noise is added to the sigmoid model to form the data with S/R of −6 dB, supporting to simulate the strong noise in the detection environment. The seismic records and spectrum after denoising by using the SS method and TF-SS method are analysed, respectively. The comparison of seismic records and spectrum of the sigmoid model for denoising comparison is shown in Figure 9.

According to the comparison of seismic records before and after denoising (Figure 9), the characteristics are summarized. For the sigmoid model, the events in data after denoising can be indistinctly distinguished, but the fault
interface in the data cannot be distinguished. As is shown in Figures 9(c) and 9(d), most of the effective signals are regarded as noise and removed by SS denoising and the denoising results are hard to be processed further. However, the result of TF-SS indicates most of the effective signals are preserved and the reflections by interface can be recognized after denoising.

3.3.3. Summary. Through the comparative analysis of seismic records before and after denoising, the denoising ability can be evaluated. However, many factors designed by human experience often make it is difficult to evaluate this processing. Therefore, the S/R and RMSE of the denoised data by the above two typical models with different S/R are studied and summarized, then a objective evaluation indicators for supplementary evaluation based on statistical characteristics was developed.

It can be seen that the S/R trend of denoised data is similar to the original noisy data (Figure 10), which can be regarded as a straight line without obvious salient point. For the data with the same noise intensity, the S/R of TF-SS processing result is higher than that of SS processing result. When the S/R of the seismic data is larger than 5 dB, the S/R after denoising by TF-SS or traditional way changes slightly.

In the window weighting process, some effective signals with high frequency will be removed together with the noise by TF filtering, resulting in the loss of high-frequency signals. However, for the TF-SS, the denoising ability is better than the traditional method. When there are noisy data with

Figure 4: Seismic record of the linear model. (a) Original signal. (b) Data with Gaussian white noise (S/R = 0 dB).

Figure 5: TF spectrum of the linear model.
Figure 6: Continued.
Figure 6: Comparison of denoising results. (a) Denoising result by SS. (b) The difference between noisy data and noise elimination by SS. (c) Denoising result using TF-SS. (d) The difference between noisy data and noise elimination by TF-SS. (e) The local similarity between (a) and (b). (f) The local similarity between (c) and (d).

| Data with noise | Weighted variance coefficient | S/R (dB) Before denoising | S/R (dB) After denoising | RMSE |
|-----------------|------------------------------|--------------------------|-------------------------|------|
|                 | 11                           | 0                        | 3.401                   | 0.626|
|                 | 15                           | 0                        | 5.639                   | 0.484|
| Linear model    | 19                           | 0                        | 5.859                   | 0.472|
|                 | 23                           | 0                        | 5.168                   | 0.511|
|                 | 27                           | 0                        | 5.117                   | 0.522|

Figure 7: Continued.
Figure 7: Denoising results of the sigmoid model with Gaussian white noise (S/R = 0 dB). (a) Original seismic record. (b) Noisy seismic record. (c) Denoising by SS. (d) Noise removed by SS. (e) Denoised by TF-SS. (f) Noise removed by TF-SS.

Figure 8: Relationship between the number of frequency bands and the optimal filtering variance.
the S/R of less than 5 dB, with the increasing of noise intensity, the denoising effect of the TF-SS is improved. The reason is that the SS produces a large error in the noise spectrum estimation and the error will be more serious with the increasing of noise intensity, resulting in the music noise.

Differently, the TF-SS can effectively suppress the emergence of music noise and improve the denoising effect by dividing the data into different frequency bands.

To make the result display more intuitive, the semilogarithmic coordinate system is used for the comparison of mean

\[ \text{Figure 9: Denoising results of the sigmoid model with Gaussian white noise (S/R = −6 dB).} \]

(a) Original seismic record. (b) Noisy seismic record. (c) Denoising by SS. (d) Noise removed by SS. (e) Denoised by TF-SS. (f) Noise removed by TF-SS. (g) Local similarity between (c) and (d). (h) Local similarity between (e) and (f).
For the sigmoid model with the S/R of 0 dB∼15 dB, the root-mean-square error of data denoised by TF-SS is higher than that processed by SS. If the root-mean-square error of the data before denoising is small, then the TF-SS has a weaker effect compared to the traditional SS, and if the root-mean-square error of the data before denoising is large, then the TF-SS can play a more significant role in reducing the root-mean-square error compared to the SS.

When the S/R of the noisy data is less than 5 dB, with the increase of noise intensity, the denoising effect of the TF-SS is improved compared with the traditional method. The reason is that the SS will produce a large error in the noise spectrum estimation. With the increase of noise intensity, the influence of the error is more and more serious, resulting in the music noise. Differently, the TF-SS can effectively suppress the emergence of music noise and improve the denoising effect by dividing the data into different frequency bands.

### 5. Conclusion

In order to solve the problem of large error in noise spectrum estimation due to the unified processing of all band data by SS, a TF-SS, which transforms SS into the time-frequency (TF) domain for processing, is proposed to optimize the discrimination ability of SS in the frequency direction.

In view of the key problem that a large amount of noise affects the data quality and easily induces interpretation deviation in tunnel seismic forward-prospecting, the TF spectral subtraction method is introduced into the data processing, and the noise reduction and removal in tunnel environment are realized.

Based on the characteristics of the main frequency concentration of tunnel seismic data, the filtering window function is improved. Through a large number of simulation examples, the relationship between the improved window function coefficient and the original data is summarized and the adaptive selection of the window function coefficient is realized.

### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This research was supported by the Science and Technology Program of Department of Transport of Shandong Province (2019B47_2).

### References

[1] K. Zhang, “Reservoir characterisation method with multicomponent seismic data by unsupervised learning and colour feature blending,” *Exploration Geophysics*, vol. 50, no. 3, pp. 269–280, 2019.

[2] L. Sun, “An alternative adaptive method for seismic data denoising and interpolation,” *Mathematical Problems in Engineering*, vol. 2020, Article ID 6295902, 16 pages, 2020.

[3] Bo Li, “An EEMD-based denoising method for seismic signal of sigh arch dam combining wavelet with singular spectrum analysis,” *Shock and Vibration*, vol. 2019, Article ID 4937595, 9 pages, 2019.
[4] G. Wang, "Application of a new wavelet threshold method in unconventional oil and gas reservoir seismic data denoising," Mathematical Problems in Engineering, vol. 16, pp. 1–7, 2015.

[5] A. Andisheh, M. Ali, N. Reza, S. M. Zab, and E. Afshin, "Prediction of geological hazardous zones in front of a tunnel face using TSP–203 and artificial neural networks," Tunnelling and Underground Space Technology, vol. 23, no. 6, pp. 711–717, 2008.

[6] M. Deng, "Challenges and thoughts on risk management and control for the group construction of a super-long tunnel by TBM," Engineering, vol. 4, no. 1, pp. 112–122, 2018.

[7] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Transactions on Acoustics Speech & Signal Processing, vol. 27, no. 2, pp. 113–120, 1979.

[8] B. Liu, "Structural complexity-guided predictive filtering," Geophysical Prospecting, vol. 68, no. 5, pp. 1509–1522, 2020.

[9] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise Acoustics, Speech, and Signal Processing," in Proceedings of the IEEE International Conference on ICASSP’79 IEEE, Washington, DC, USA, April 1979.

[10] R. J. McAulay and M. L. Malpass, "Speech enhancement using a soft-decision noise suppression filter," IEEE Transactions on Acoustics Speech & Signal Processing, vol. 28, no. 2, pp. 137–145, 1980.

[11] P. Lockwood, J. Boudy, and M. Blanchet, "Experiments with a nonlinear spectral subtractor (NSS), hidden markov models and the projection, for robust speech recognition in cars," Speech Communication, vol. 11, pp. 215–228, 1992.

[12] L. Singh and S. Sridharan, "Speech enhancement using critical band spectral subtraction," in Proceedings of the International Conference on Spoken Language Processing DBLP, Denver, CO, USA, September 1998.

[13] D Kamath Sunil and P. C. Loizou, "A multi-band spectral subtraction method for enhancing speech corrupted by colored noise," in Proceedings of the icassp.International Conference, Orlando, FL, USA, May 2002.

[14] B. L. Sim, "A parametric formulation of the generalized spectral subtraction method," IEEE Transactions on Acoustics Speech & Signal Processing, vol. 6, pp. 328–337, 1998.

[15] I. Oh and I. Lee, "Speech enhancement system using the multi-band coherence function and spectral subtraction method," The Journal of the Acoustical Society of Korea, vol. 38, no. 4, pp. 406–413, 2019.

[16] Scalart Pascal and J. V. Filho, "Speech enhancement based on a priori signal to noise estimation," in Proceedings of the IEEE International Conference on Acoustics IEEE, Toronto, ON, Canada, June 2002.

[17] C. Plapous, C. Marro, and P. Scalart, "Improved Signal-to-noise ratio estimation for speech enhancement," IEEE Transactions on Audio Speech and Language Processing, vol. 14, no. 6, pp. 2098–2108, 2006.

[18] Y. Zhang, K. Wang, and B. Yan, "Speech endpoint detection algorithm with low signal-to-noise based on improved conventional spectral entropy," in Proceedings of the 2016 12th World Congress on Intelligent Control and Automation (WCICA) IEEE, Guilin, China, June 2016.

[19] Y. Cao, "Voice activity detection algorithm based on entropy in noisy environment," in Proceedings of the 2016 Chinese Control and Decision Conference (CCDC), pp. 3799–3803, Yinchuan, China, May 2016.

[20] H. M. Ahmed, "Active voice detection using ridgelet transform," AL-Mansour Journal, vol. 26, pp. 47–56, 2016.

[21] B. Ye, "Voice activity detection based on time-delay neural networks," in Proceedings of the 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), pp. 1173–1178, Lanzhou, China, November 2019.