Research on Single Channel Speech Noise Reduction Algorithm Based on Signal Processing

Chen Chen

Applied Physics, Faculty of Science, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, China

Keywords: Speech noise reduction, Single channel microphone, Wiener filtering, Spectral subtraction, Mmse, Subspace

Abstract: As an analog signal carrying specific information, speech has become an important means of obtaining and disseminating information in people's lives. However, in real life, speech signals are polluted by various noises during the encoding and transmission processes. At this time, you can use the voice noise reduction technology to suppress, reduce noise interference, and improve the quality of the voice. The voice noise reduction technology can be divided into single channel and multi-channel according to the number of channels of the microphone. Because of the simple model and low cost, single-channel voice noise reduction Algorithms have been the focus of research. Based on a single-channel speech algorithm for signal processing, this paper presents several basic algorithms such as Wiener filter algorithm, spectral subtraction, signal subspace algorithm, MMSE method, LMS and RLS algorithm, and digital filter, as well as their basic principles, algorithm implementation steps and advantages and disadvantages analysis.

1. Introduction

Since the invention of the telephone by the British Bell in 1876, there has been speech signal processing. With the rapid development of science and technology and the Internet, speech processing technologies such as speech communication and speech recognition have been widely used in all aspects of people's lives. In the actual environment, speech signals are affected by background noise to a certain extent during the acquisition, transmission, encoding, and output processes. For example, when people use smart devices to make two-party calls, as long as one party has background noise interference, it will cause speech quality. The decline in speech causes obstacles to voice communication. It can be seen that how to reduce or eliminate various background noises and improve the clarity and signal-to-noise ratio of speech signals is the significance of speech noise reduction.

Since 1950, speech noise reduction has become a hot area for speech signal processing, and scholars at home and abroad have conducted extensive and in-depth research on it. After the 1960s, the development of electronic computers caused significant changes in speech noise reduction technology. With the maturity of digital processing technology, in the 1970s, speech noise reduction algorithms have achieved certain results. To date, speech noise reduction algorithms have been developed and can be divided into single-channel systems and multi-channel systems based on the number of channels.

The speech noise reduction algorithm history [1] of single-channel speech is as follows: In 1974, Weis, Paron, and Aschkenasy proposed spectral subtraction. In 1979, Lin and Oppeheim proposed a Wiener filtered speech enhancement algorithm based on the minimum mean square error criterion. 1983 In 1984, Yarv Ephraim and David Malah proposed the Minimum Mean Square Estimation Algorithm (MMSE), and in 1984 they proposed the minimum mean square error criterion short-term amplitude spectrum(MMSE-STSA). In 1995, Ephaim proposed a signal noise reduction algorithm based on signal subspace. In 1996, Scalart P proposed a Wiener filtering algorithm and spectral subtraction based on prior signal-to-noise estimation. Since the beginning of the 21st century As the technology of integrated circuits and signal processing chips continues to develop and mature, a large number of new speech noise reduction algorithms have emerged, such as

  DOI: 10.25236/ictmic.2020.018

methods based on neural networks and wavelet transforms that are getting better and better.

This article studies several speech noise reduction algorithms for single-channel microphones, including Wiener filtering algorithms, spectral subtraction, MMSE algorithms, subspace algorithms, LMS and RLS algorithms, and FIR and IIR digital filters. The concepts of these algorithms are mainly introduced, Basic principles, algorithm implementation steps and advantages and disadvantages comparison.

2. Single Channel Speech Noise Reduction Algorithm

2.1 Wiener Filter Algorithm

The American scientist n. Wiener originally proposed Wiener filtering to solve the problem of aerial shooting, which is an algorithm that uses the minimum mean square error criterion to make the speech signal achieve the best linear filtering under short-term stationary conditions.

2.1.1 Basic Principles of Wiener Filtering Algorithm

The Wiener filtering algorithm first uses Fourier transform to estimate the power spectrum of the noisy speech signal and the noise signal in the frequency domain, and constructs a gain function based on it. Then calculates the noisy speech power and gain function to obtain the filtered pure speech signal Power, and then inverse Fourier transform to get the enhanced speech signal in the time domain.

2.1.2 Wiener Filtering Algorithm Based on Prior Snr Estimation

The Wiener filtering algorithm [2] based on prior signal-to-noise ratio estimation is a speech noise reduction algorithm proposed by Scalart P and others in 1996, and it is also the most frequently used speech enhancement algorithm in recent years.

Assume that the impulse response of a linear system is h (n), and put a random signal with an observation value of x (n), which contains the signal s (n) and the noise signal n (n), and its formula is as follows:

$$x(n)=s(n)+n(n) \tag{1}$$

The Fourier transform of the above formula (1-1) can be obtained:

$$X(k, j)=S(k, j)+N(k, j) \tag{2}$$

X(k, j), S(k, j) and N(k, j) are the time-domain signals in equation (2-1) at the j-th frame and the k-th spectral component, respectively.

Corresponding amplitude spectrum representation.

The gain function of the Wiener filter algorithm based on the prior signal-to-noise ratio can be expressed as:

$$H(k, j)=\frac{\xi(k, j)}{1+\xi(k, j)} \tag{3}$$

$\xi(k, j)$ represents the prior signal-to-noise ratio, which can be obtained by the direct judgment method [3][4]:

$$\xi(k, j)=a \cdot H^2(k, j-1) \cdot \gamma(k, j-1)+(1-a) \cdot \max[\gamma(k, j)-1, 0] \tag{4}$$

$\gamma(k, j)$ is the posterior signal-to-noise ratio, a is the smoothing factor, and $\gamma(k, j)$ can be expressed as:

$$\gamma(k, j)=\frac{|X(k, j)|^2}{P_{dd}(k, j)} \tag{5}$$

$P_{dd}(k, j)$ is the noise power spectrum, which can be evaluated when the speech is at rest:

$$P_{dd}(k,j)=\mu\cdot P_{dd}(k,j)+(1-\mu)\cdot|X(k,j)|^2 \tag{6}$$

μ is the noise smoothing factor.

The enhanced speech can be expressed as:

$$S(k,j)=H(k,j)\cdot X(k,j) \tag{7}$$

Algorithm implementation steps:

(1) Fourier transform the noisy speech to obtain the frequency domain amplitude spectrum of the speech.

(2) Generally, the first 120ms of noisy speech is judged as noise, and the average power of these 6 frames (20ms per frame) is taken as the initial noise power.

The newly obtained noise power spectrum value $P_{dd}(k,j)$.

(3) Calculate the posterior signal-to-noise ratio $\gamma(k,j)$ by formula (5), and then obtain the prior signal-to-noise ratio $\xi(k,j)$ by the direct decision method in formula (4).

(4) Obtain the gain function H(k, j) of the Wiener filter algorithm based on the prior signal-to-noise ratio according to formula (3).

(5) Substituting the obtained gain function into equation (7) to obtain the enhanced speech amplitude spectrum $S(k,j)$. The inverse Fourier transform can be used to obtain the final enhanced speech signal.

2.1.3 Analysis of the Advantages and Disadvantages of Wiener Filtering Algorithm

The advantages of the Wiener filtering algorithm [5] are as follows:

(1) In terms of speech quality, the Wiener filter algorithm pays attention to the reduction of noise during filtering, so the speech enhanced by the Wiener filter algorithm has a higher signal-to-noise ratio and improved quality compared with noisy speech.

(2) The residual noise after speech enhancement is similar to white noise, which makes the human ear feel better.

The disadvantage of the Wiener filtering algorithm is that while reducing the noise to improve the quality of speech, it will filter out useful speech, cause speech distortion and do not achieve good results in terms of enhanced intelligibility.

2.2 Spectral Subtraction

The Wiener filtering algorithm is to enhance speech by estimating the signal spectrum. Later, Boll also proposed spectral subtraction [6] using spectral estimation. Spectral subtraction is one of the most classic speech noise reduction algorithms, which uses statistical stability of noise. The characteristics of additive and additive noise are not related to pronunciation for speech noise reduction.

2.2.1 Basic Principles of Spectral Subtraction

Spectral subtraction assumes that the noise is statistically stable. The estimated value of the noise spectrum calculated using the non-speech gap measurement replaces the spectrum with the speech interval noise and is subtracted from the noisy speech spectrum to obtain the estimated speech spectrum. When the difference is set to zero when negative.

The conditions for using spectral subtraction must satisfy the following assumptions:

(1) All noise in the speech signal is additive and is a normally distributed white noise with a mean value of 0.

(2) Noise and speech are locally transient and stable, that is, the statistical characteristics of noise in a noisy speech are the same as the statistical characteristics of the noise before the beginning of the speech segment, and remain unchanged throughout the speech segment.

(3) Noise is independent or uncorrelated with speech statistics.

(4) The energy of the speech signal is concentrated in certain frequency bands, and the amplitude is high.

(5) Because the human ear is not sensitive to signal phase information, the phase of the noisy speech signal is used instead of the phase of the pure speech.

2.2.2 Spectral Subtraction Algorithm [7]

Assume that the original noiseless speech signal is s (n), the noise signal is x (n), and the noise-containing speech signal is y (n).

$$y(n) = s(n) + x(n) \tag{8}$$

Fourier transforming the above formula (8), we get:

$$Y(\omega) = S(\omega) + x(\omega) \tag{9}$$

Perform square operation on both sides of the above formula (9), and get:

$$|Y(\omega)|^2 = |S(\omega)|^2 + |x(\omega)|^2 + 2\,\mathrm{Re}[S(\omega) * X(\omega)] \tag{10}$$

Performing mathematical expectation calculations on both sides of the above formula (10), we get:

$$E(|Y(\omega)|^2) = E(|S(\omega)|^2) + E(|x(\omega)|^2) + 2E\{\mathrm{Re}[S(\omega) * X(\omega)]\} \tag{11}$$

From hypothesis (1)(3), $\mathrm{Re}[S(\omega) * X(\omega)] = 0$, so:

$$E(|Y(\omega)|^2) = E(|S(\omega)|^2) + E(|x(\omega)|^2) \tag{12}$$

For a transient stationary process, there are:

$$|Y(\omega)|^2 = |S(\omega)|^2 + |x(\omega)|^2 \tag{13}$$

Use the noise power spectrum calculated without speech gap measurement to estimate the frequency spectrum of speech interval noise $|X(\omega)|^2$ available:

$$|S(\omega)|^2 = |Y(\omega)|^2 - |X(\omega)|^2 \tag{14}$$

The original noiseless speech signal is:

$$|S(\omega)| = \sqrt{|Y(\omega|^2 - |X(\omega)|^2} \tag{15}$$

2.2.3 Analysis of Advantages and Disadvantages of Spectral Subtraction

The advantages of spectral subtraction are as follows:

(1) The algorithm is as simple as the Wiener filtering algorithm, easy to implement, has a small amount of calculation, and takes up little computer memory.

(2) The effect of smooth noise is better, and the signal-to-noise ratio can be greatly improved.

The disadvantages of spectral subtraction are as follows:

(1) Spectral subtraction ignores the random characteristics of noise and speech, which will cause speech distortion when the signal-to-noise ratio is increased. If too much spectral component is subtracted, some speech information may be lost; on the contrary, too much speech will remain Noise, especially at low signal-to-noise ratios, spectral subtraction is difficult to improve speech quality, and it is even more difficult to improve speech intelligibility.

(2) The noise estimation of spectral subtraction is to average the noise. After the spectrum is converted to the time domain by inverse Fourier transform, multi-frequency sound quality vibrato is produced, which is called music noise [8]. The ear's hearing fatigue cannot be removed by spectral subtraction, which greatly affects and limits the performance of the noise reduction algorithm.

2.3 Minimum Mean Square Error (Mmse)

Because spectral subtraction generates music noise, Yariv Ephraim and David Malah proposed the minimum mean square estimation algorithm (MMSE) [9] in 1983, which can solve part of the problem of music noise. It is a distortion criterion to be determined and a posteriori probability The insensitive estimation method uses the noise power spectrum information to estimate the pure speech spectral components from the noisy speech spectrum, and then uses the noisy speech phase to obtain an enhanced speech signal [10].

Assuming that the short-term spectrum of the speech signal s (n) is Gaussian, s (n) is a pure speech signal, and d (n) is a noise signal, then the speech signal with noise $y(n) = s(n) + d(n)$. $Y_k = R_k \exp(j\theta_k)$, N_k , $S_k = A_k \exp(ja_k)$ repesent the k-th spectral component of a noisy speech signal, a noisy signal, and a pure speech signal. Under the MMSE criterion, the estimation formula for A_k is obtained from the Bayesian formula[11]: Yariv Ephraim and David Malah proposed the short-term magnitude of the minimum mean square error criterion Spectrum (MMSE-STSA) in 1984, which is based on the theory that the human ear's perception of sound intensity is proportional to the logarithm of the spectral amplitude. Assuming that the noise signal n(i) is stationary Gaussian noise, we have:

$$\hat{A}_k = E(A_k \mid Y_k) = \frac{\int_0^\infty \int_0^{2\pi} a_k p(a_k, a_k) p(Y_k \mid a_k, a_k) da_k da_k}{\int_0^\infty \int_0^{2\pi} p(a_k, a_k) p(Y_k \mid a_k, a_k) da_k da_k} \tag{16}$$

$$p(Y_k \mid a_k, a_k) = \frac{1}{\pi \lambda_d(k)} \exp\{-\frac{1}{\lambda_d(k)} |Y_k - a_k e^{ja_k}|^2\} \tag{17}$$

$$p(a_k, a_k) = \frac{a_k}{\pi \lambda_s(k)} \exp\{-\frac{a_k^2}{\lambda_s(k)}\} \tag{18}$$

Substituting (17) (18) into (16) gives the estimated speech spectrum:

$$\hat{A}_k = \Gamma(1.5) \exp(-\frac{v_k}{2})[(1+v_k) I_0(\frac{v_k}{2}) + v_k I_1(\frac{v_k}{2})] R_k = G_{MMSE} R_k \tag{19}$$

Where $\Gamma()$ is a gamma function. $I_0()$ and $I_1()$ represent the zero-order and first-order modified Bessel functions, respectively.

The advantages of the MMSE algorithm are as follows:

(1) The MMSE method can effectively suppress noise interference and enhance speech signals, achieving a good balance between speech intelligibility and noise reduction ratio.

(2) Compared with the spectral subtraction method, the MMSE method is more effective in suppressing music noise. In the case of low signal-to-noise ratio, it is more effective and low-recovery signal waveform than other methods, and the applicable signal-to-noise ratio range is wider.

The disadvantage of the MMSE algorithm is that the performance of the MMSE algorithm in processing Gaussian white noise is low [12].

2.4 Algorithms Based on Signal Subspace

The Wiener filter, spectral subtraction, and MMSE algorithm introduced above are all speech based on signal spectrum estimation Noise reduction method, but in 1995, two scholars, Ephaim and Van Trees, proposed a signal subspace-based speech noise reduction method [13] different from the above three methods.

2.4.1 Basic Principles of the Subspace Algorithm

In the subspace-based speech noise reduction technology, the noisy signal subspace is divided into a signal subspace and a noise subspace. The noisy speech signal is projected onto the signal subspace and the noise subspace, and then the noise subspace is filtered out as much as possible. And preserve the signal part of the signal subspace, thereby recovering an approximately pure speech signal.

Decomposing the signal space into two subspaces mainly includes two decomposition methods: singular value decomposition (SVD) and eigenvalue decomposition (EVD) [14]. Singular value decomposition is mainly in the space of noisy speech, using singular values and their corresponding feature vector is used to decompose the speech subspace, and then reconstructed to obtain the speech signal. The eigenvalue decomposition is to use the eigenvalue decomposition of the noisy subspace to process the eigenvectors of the eigenvalues obtained by the decomposition separately, and then only retain the feature vector (corresponding to the signal subspace) corresponding to zero eigenvalues, thereby reconstructing a pure speech signal.

The subspace speech enhancement algorithm estimates pure speech signals from noisy speech. There are two types of linear estimators: time domain constraint estimator (TDC) and frequency domain constraint estimator (SDC).

Because the eigenvalue-based singular value-based decomposition method has a better noise reduction effect, the time-domain constraint estimator can more effectively suppress music noise. This article chooses a subspace method based on eigenvalue decomposition and time-domain constraint estimation as an example. Introduction.

2.4.2 Subspace Algorithm Based on Eigenvalue Decomposition with Time-Domain Constraint Estimation

Experimental steps [15]:

(1)Define a model of signal x. Let x be a K-dimensional zero-mean random vector, covariance matrix R_x is a positive definite matrix and rank is M. There are M positive definite eigenvalues and K-M zero eigenvalues.

$$R_x = E\{xx''\} \tag{20}$$

(2)Let d be the K-dimensional zero-mean additive white noise that is uncorrelated with the speech signal, and the covariance matrix R_x has been known and positive, there is

$$R_d = E\{dd''\} \tag{21}$$

(3)Let y(n) be a noisy signal, and y(n) = x(n) + d(n), then the covariance matrix of the noisy speech can be calculated:

$$R_y = E\{yy''\} = E(x+d)(x''+d'') = E[xx''] + E[dd''] = R_x + R_d \tag{22}$$

(4) Estimation matrix $\Sigma = R_d{}^{-1} R_y - I$, and calculate the noise covariance matrix R_d .

(5) Eigenvalue decomposition of $\Sigma : \Sigma V = V \Lambda$.

(6) Let eigenvalues of Σ be arranged as $\lambda_\Sigma^1 \geq \lambda_\Sigma^2 \geq \cdots\cdots \geq \lambda_\Sigma^K$, and estimate the subspace dimension:

$$M = \arg \max_{0 \leq k \leq K} \{\lambda_\Sigma^K > 0\} \tag{23}$$

(7) Calculate the SNR value:

$$SNR = \frac{tr(V^T R_x V)}{tr(V^T R_d V)} = \frac{\sum_{k=1}^{M} \lambda_\Sigma^K}{K} \tag{24}$$

(8) $SNR_{dB} = 10 \log SNR$, calculate the value of μ :

$$\mu = \begin{cases} \mu_0 - \dfrac{SNR_{dB}}{s}, -5 < SNR_{dB} < 20 \\ 1, SNR_{dB} \geq 20 \\ 5, SNR_{dB} \leq -5 \end{cases} \tag{25}$$

(9) Calculate the optimal linear estimator H_{opt} :

$$H_{opt} = R_d V \Lambda(\Lambda + \mu I)^{-1} V^T = V^{-T} \Lambda(\Lambda + \mu I)^{-1} V^T \tag{26}$$

(10) Use H_{opt} to calculate the enhanced pure speech signal: $x = H_{opt} \cdot y$.

2.4.3 Analysis of Advantages and Disadvantages of Algorithms Based on Signal Subspace

The advantage of the subspace method is that because the subspace method can adjust the quality of the output speech by controlling both the degree of noise elimination and the degree of speech distortion, and it has the best decorrelation to the speech signal. So contrast spectrum subtraction, subspace algorithm, the output speech has a higher signal-to-noise ratio, that is, the quality of the processed speech is higher.

The disadvantages of the subspace method are as follows:

(1) The subspace algorithm has higher computational complexity and time consuming than the Wiener filtering method, spectral subtraction method and MMSE method, and cannot be well applied to real-time systems.

(2) When the signal-to-noise ratio is very low, there will be residual noise, and the noise reduction effect is worse than when the signal-to-noise ratio is high.

2.5 Lms Algorithm

Adaptive filtering originates from linear filtering methods such as Wiener filtering. After improvement, adaptive filtering currently has the best noise reduction effect on noisy speech. Among the adaptive filtering algorithms, the most commonly used is the minimum mean square error algorithm (LMS) and recursive least squares algorithm (RLS) .The LMS algorithm is based on Wiener filtering, with the fastest descent algorithm, taking the minimum mean square value of the error between the known expected response and the filter output signal as the standard, based on the input signal During the iteration process, the gradient vector is estimated, and the weight coefficients are updated to achieve the optimal solution recursively [16].

LMS algorithm steps:

(1) Let the input vector be $X(n) = [x(n), x(n-1), \cdots, x(n-N+1)]^T$, the time series is n, and the weight vector $W(n) = [w_0(n), w_1(n-1), \cdots, w_{N-1}(n)]^T$, where N is the filter order.

(2) Filter output y(n):

$$y(n) = W^T(n) X(n) \tag{27}$$

(3)Let d(n) be the signal that the filter wants to approximate, and the error estimate is:

$$e(n) = d(n) - y(n) \tag{28}$$

(4)Update weight coefficient:

$$W(n) = W(n-1) + 2\mu e(n-1) x(n-1) \tag{29}$$

Substituting (28) into (29) gives the recursive formula:

$$w(n) = w(n-1) + 2\mu x(n-1)[d(n-1) - x^H(n-1) w(n-1)] \tag{30}$$

Where μ is the step factor, $0 < \mu < (MP_{in})^{-1}, P_{in} = E\{|x_1(n)|^2\}$, and M is the number of filter taps.

The advantages of the LMS algorithm are as follows:

(1) Easy to implement and simple algorithm.

(2) It does not depend on the model and its performance is robust.

The disadvantages of the LMS algorithm are as follows:

(1) The convergence speed of the LMS algorithm is controlled by the distribution range of the eigenvalues. The larger the distribution range, the slower the convergence, and the speech signal just has a larger eigenvalue distribution range, so the LMS algorithm has a slower convergence speed.

(2) When processing non-stationary signals, the adaptive performance of the LMS algorithm is poor.

2.6 Rls Algorithm

The recursive least squares algorithm RLS is different from LMS. The RLS algorithm is to examine the average power of an error signal output by a system that inputs a stable signal over a period of time, and make it to a minimum as a criterion for adaptive system performance. Its working process the new parameters are compared with the previous parameters, and the previous data is corrected according to the recursive algorithm to reduce the estimation error, thereby updating the parameter estimates [17] over and over until the parameter estimates meet the experimental requirements.

RLS algorithm steps:

(1) Same as LMS algorithm step (1).

(2) Filter output y(n):

$$y(n) = W^T(n-1)X(n) \tag{31}$$

(3) Error estimation:

$$e(n) = d(n) - y(n) \tag{32}$$

(4) Update weight vector:

$$W(n) = W(n-1) + g(n)e(n) \tag{33}$$

Substituting (32) into (33) gives the recursive formula:

$$w(n) = w(n-1) + g(n)[d(n) - x^T(n-1)w(n-1)] \tag{34}$$

Where the gain coefficient $g(n) = P(n-1)X(n)/[\lambda + X^T(n)P(n-1)X(n)]$, where λ is the forgetting factor and $0 < \lambda < 1$, P(n) is the autocorrelation matrix Inverse matrix of $P_{xx}(n)$.

The advantages of the RLS algorithm are as follows:

(1) The convergence speed of the RLS algorithm is many times faster than LMS.

(2) No matter what kind of changes occur to the input signal, the adaptability of RLS to the signal is better.

The disadvantages of the RLS algorithm are as follows:

(1) The computational complexity of the RLS algorithm is very high, and the required amount of storage is huge, which is not conducive to real-time implementation.

(2) If the inverse of the estimated autocorrelation matrix loses its positive definite characteristics, it will cause the algorithm to diverge.

2.7 Fir and Iir Filters

One of the main purposes of voice noise reduction is to eliminate environmental noise, and the digital filter is a method to filter out noise outside the frequency band of the voice signal under high noise. According to the characteristics of the digital filter on the impulse response, it can be classified into limited Impulse response (FIR) digital filters and infinite impulse response (IIR) digital filters.

The transfer function of the FIR filter has only zero points, and its unit impulse response h(k) contains only a finite number of non-zero values, that is, the impulse response of this digital filter is limited in time and the output is zero after a certain time. The IIR filter has both zero and poles, and its impulse response h(k) contains an infinite number of non-zero values, that is, the impulse response of this filter is an infinitely long sequence, which may change after a certain time. Small but not zero [18].

FIR and IIR filter design includes three steps [19];

(1) Give the technical specifications of the filter.

(2) Design a technical index required by H(z) to approximate it.

(3) Implement the designed H(z).

Advantages and disadvantages of FIR filters [20]:

(1) The non-recursive operation of the fir filter makes the operation error caused by the finite word length effect not cause the system to be unstable.

(2) The fir filter can obtain strict linear phase characteristics to ensure that the signal is not distorted during transmission.

(3) Because the unit pulse response of the fir filter is finite, a fast Fourier transform can be used to implement filtering processing and improve the operation rate.

(4) The fir filter does not have a ready-made design formula, and the calculation workload is large, and it is generally completed by means of a computer.

The advantages and disadvantages of IIR filters:

(1) The IIR filter must adopt a recursive structure. The rounding process in the operation sometimes causes parasitic oscillation.

(2) The better the selectivity of the IIR filter, the more serious the non-linearity of the phase. This point the fir filter is better than the IIR filter.

(3) The IIR filter is easier to implement than the FIR filter. Under the same conditions, the IIR filter is designed. The filter requires fewer parameters, fewer operations, and is more economical.

(4) The unit pulse of the IIR filter is infinite, and the fast Fourier transform algorithm cannot be used, and the operation speed is slow.

(5) The calculation workload of the IIR filter is small, and the calculation tools are not high.

3. Summary and Prospect

Speech noise reduction technology is a hot research area of speech signal processing technology. The main significance is to suppress background noise and improve speech quality. It has applications in many fields. This paper introduces several speech noise reduction algorithms for single channel microphones. Works as follows:

(1) The basic principle of the Wiener filtering algorithm is briefly explained, the steps of the Wiener filtering algorithm based on the prior signal-to-noise ratio estimation are introduced, and its advantages and disadvantages are summarized.

(2) Then introduced the basic principles of spectral subtraction, listed the assumptions made by spectral subtraction, summarized several main formulas in spectral subtraction, and compared the advantages and disadvantages of Wiener filtering.

(3) The principle of the minimum mean square error method (MMSE) and the minimum mean square error criterion (MMSE-STSA) and the speech spectrum estimation formula are introduced, and their advantages and disadvantages are summarized.

(4) There are many algorithms for subspace-based speech noise reduction techniques. This article chooses a subspace method based on eigenvalue decomposition and time-domain constraint estimation as an example to introduce the specific experimental steps and feature analysis.

(5) The two most common algorithms in the adaptive filtering algorithm-the minimum mean square error algorithm (LMS) and the recursive least squares method (RLS) are introduced and their recursive formulas are derived and listed the advantages and disadvantages of both.

(6) According to the characteristics of the impulse response, the digital filter is divided into FIR and IIR, and the characteristics of the two filters are compared in detail.

In addition to the work done above, there are still some areas for improvement in this article:

(1) This article only introduces several basic speech noise reduction algorithms. In fact, there are many emerging effective algorithms, such as neural network-based speech noise reduction algorithms, etc., which have not been summarized.

(2) The algorithm steps described in this article are relatively simple, and only a few main formulas are selected. The complete inference calculation process needs to be further improved.

(3) This article is only about the noise reduction algorithm for single-channel systems. In fact, there are many studies on multi-channel microphone arrays, which need to be further understood.

(4) This article does not mention specific applications that are generalized to life, but only explains some theoretical algorithms, which need to be further improved.

References

[1] Loizou P. Speech Enhancement: Theory and Practice[M]. Boca Raton: Florida: CRC Press LLC, 2007.

[2] Scalart P, Vieira-Filho V. Speech enhancement based on a priori signal to noise estimation[C]. Proc. 21st IEEE Int. Conf. Acoustics. Speech Signal Processing, Atlanta, 1996:629-632.

[3] Yang L, Loizou P C. Speech enhancement by combining statistical estimators of speech and noise[C]. Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. IEEE, 2010:4754-4757.

[4] Meng Xin, Ma Jianfen, Zhang Xueying, a highly intelligible speech enhancement algorithm with low signal-to-noise ratio [j].Computer Applications and Software, 2016, 33 (10): 145-147.

[5] Bai Wenya, Huang Jianqun, Chen Zhiling, Improved implementation of speech enhancement algorithm based on Wiener filter, Electroacoustic Technology, 2007, 31 (1): 44-50.

[6] Boll S. Suppression of acoustic noise in speech using spectral subtraction. IEEE Transaction on Acoust, Speech, Signal Process. 1979, 27(2): 113-120.

[7] Wang Tao, Research on speech noise reduction processing technology, Lanzhou Jiaotong University, 2018: 14-15.

[8] M. Berouti, R., Schwartz, J. Makhoul. Enhancement of speech corrupted by acoustic noise [J]. Proc. IEEE ICASSP. Washington, 1979, 208-211.

[9] Yang Xingjun, Chi Huisheng, et al. Digital Processing of Speech Signals, Beijing, Electronic Industry Press, 1995.

[10] Zhi-Heng Lu, Huai-Zong Shao, and Tai-Liang Ju. Speech Enhancement Algorithm Based on MMSE Short Time Spectral Amplitude in Whispered Speech. Journal of Electronic Science and Technology of China, 2009, 7(2): 115-119.

[11] Zhiheng Lu, Research on Speech Enhancement Technology and Real-time Implementation of dsp Multi-channel, University of Electronic Science and Technology of China, 2009:25(2-19).

[12] Meng Jing, Research on Evaluation of Speech Enhancement Algorithm Performance, Lanzhou Jiaotong University, 2010.

[13] Ephaim Y and Van Trees H L. A signal Subspace Approach for Speech Enhancement [C]. IEEE Trans. Speech and Audio Processing, 1995, 3: 251-266. [14] Wang Wenjie, Research on microphone array speech enhancement technology, Hebei University of Technology, 2010: 14-15.

[14] Jin Qian, Research on Speech Enhancement Algorithm Based on Signal Subspace, Jilin University, 2011: 19.

[15] Wang Haifeng, Research on real-time processing algorithms for speech noise reduction, Wuhan University of Technology, 2009.

[16] Weidrow, Wang Yongde, Adaptive Signal Processing [m], Beijing: Mechanical Industry Press, 2008: 54-69.

[17] Zhang Linlin, Research on joint noise reduction method based on dsp, Henan University of Science and Technology, 2010. [19] Leng Bangping, Implementation of high-speed fir digital filter on fpga [d], Chengdu: University of Electronic Science and Technology of China Electronics and Solid State Electronics, 2008.

[18] Zhang Lifei, Research on Application of BP Neural Network Based on Genetic Algorithm in Digital Filter, Xihua University, 2007.