Integrated speech signal processor based on spectrum analysis and LMS algorithm

Fei Shao*

School of Electronic Information, Wuhan University, Wuhan, China

Abstract. Spectrum analysis is often used for sound signal analysis. LMS algorithm is an improved algorithm of steepest descent algorithm, which is an optimized extension of wiener filter theory and is often used in audio noise reduction. In this paper, we display the time-frequency domain image of audio, use spectrum analysis for audio gender recognition, and use LMS algorithm for noise reduction. Experimental results show that the method adopted in this paper has high accuracy for sound discrimination, and can significantly eliminate the environmental noise, which can not only be heard but seen in the frequency domain figure. After our audio enhancement, the speaker's voice can be very clear.

1 Introduction

As the leading direction of intelligent computer research and the key technology of human-machine voice communication, speech recognition technology has always received extensive attention from the scientific community of various countries. The range of human pitch is about 70～350Hz. Due to the different physiological structure, male and female voices show different auditory characteristics. The pitch frequency of male voice is mostly between 100-200HZ, while female voice is between 200-350HZ. Compared with male voices, the average and standard deviation of the former are about twice that of the latter. The pitch frequencies of different speakers are normally distributed on the logarithmic frequency axis for male and female voices. The average and standard deviation of the pitch frequencies of male voices are 125Hz and 20Hz, respectively. The female voice is about twice that of the male voice. In view of the obvious difference in the pitch frequency of male and female voices, the pitch frequency can be used as the basis for the recognition of male and female voices. We designed a player based on this feature using Matlab. The player can record audio, play audio, display the time domain and frequency domain waveforms of the audio, double the speed of the audio, adjust the volume, and perform male and female voice recognition on the audio. In order to improve the accuracy of the male and female voice recognition technology in noisy environments, we have added a noise reduction function based on the improved LMS algorithm to process the voice signal and display the time domain and frequency domain of the processed audio.

The least mean square algorithm, or LMS algorithm, is an improved algorithm of the steepest descent algorithm. It is an optimized extension of wiener filtering theory using the

* Corresponding author: 1392737826@qq.com

fast descent method. The algorithm does not need to know the statistical characteristics of the input signal and the expected signal, and the weight coefficient of the "current moment" is obtained by the weight coefficient of the "previous moment" plus a proportional term of a negative mean square error gradient.

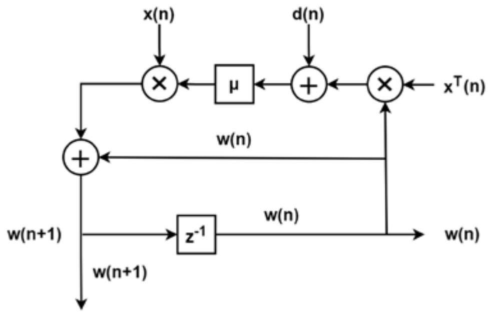

Fig. 1. Vector flow chart of LMS algorithm.

2. The proposed models

2.1 Spectrum analysis

Time domain describes the relationship of mathematical functions or physical signals to time. The frequency domain is a coordinate system used to describe the frequency characteristics of a signal. In electronics, control system engineering, and statistics, frequency domain diagrams show the amount of signal in each given frequency band within a frequency range. Time domain analysis and frequency domain analysis are two observation planes of analog signals. Time domain analysis uses the time axis as the coordinate to express the relationship of the dynamic signal; frequency domain analysis changes the signal to the frequency axis as the coordinate to express. Generally speaking, the time domain representation is more vivid and intuitive, and the frequency domain analysis is more concise, and the analysis of the problem is more profound and convenient. In order to display the sound processing effect more conveniently, we use the Fourier transform function of Matlab to process the signal and display the frequency domain and time domain waveforms of the signal.

$$s(f) = \int_{-\infty}^{+\infty}(s(t) \cdot e)dt \tag{1}$$

$$s_{D(t)} = \frac{dS(t)}{dt} \tag{2}$$

$$s_D(f) = \int_{-\infty}^{+\infty}(s_D(t) \cdot e^{-j2\pi ft})dt = j \cdot 2\pi f \cdot s(f) \tag{3}$$

2.2 LMS algorithm

LMS algorithm is updated point by point. Every time an input signal and error signal are sampled, the filter coefficient will be updated once, so that the error signal E (n) becomes smaller and smaller until it is close enough to 0 to judge that the algorithm has converged.

Initialization must be set up before performing the LMS algorithm. Determine the update step μ and filter order N, and then initialize the filter coefficient, generally initialized to 0, for specific noise, you can first let the algorithm convergence once, the convergence of the filter coefficient set as the initial value, so that you can get better noise reduction effect, finally start the LMS algorithm operation.

Assuming that the filter is of order N, the coefficient is W(N), and the input signal is x(N), the output signal y(N) of the filter should be:

$$y(n) = \sum_{i=0}^{N-1} w_i(n)x(n-1) = W^T(n)X(n) = X^T(n)W(n) \tag{4}$$

Its Wiener solution should be:

$$W_{opt} = R^{-1}r \tag{5}$$

LMS filter coefficient update formula should be:

$$\mathrm{W(n+1) = W(n) + 2\mu X(n)e(n)} \tag{6}$$

LMS algorithm is used to iterate, and the appropriate filter weight matrix is obtained. Finally, the original signal can be successfully recovered.

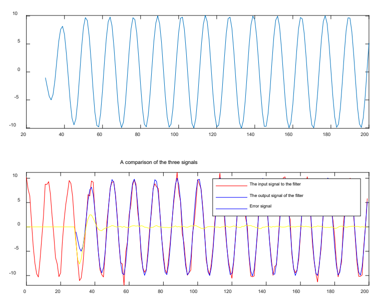

Fig. 2. Waveform recovered by LMS filter.

3 Results

This project adopts MATLAB software for graphic interface design, and finally obtains an App with comprehensive functions.

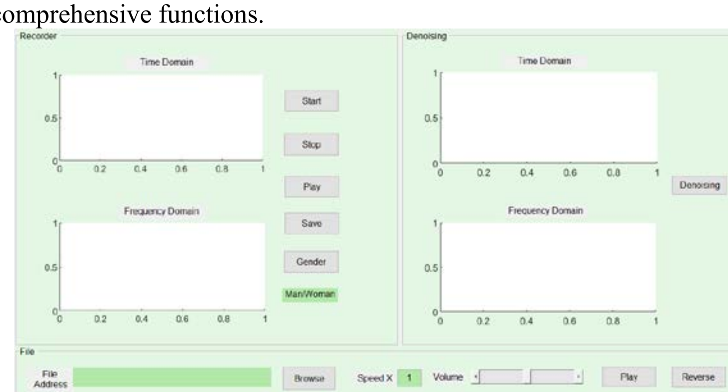

Fig. 3. Integrated audio processor interface.

Now the specific functions are described as follows:

First press the Start button to Start recording sound, such as recording a human voice in a noisy environment, and real-time signal time domain waveform will be displayed on the interface.

After recording, press the Stop button to Stop recording. Click the Play button to Play back, and automatically draw the spectrum of the original signal. Click the Save button to select a local path to Save the recorded audio.

Click the Gender button to identify the Gender of the speaker and display it at the bottom of the screen.

Click the Denoise button on the right for Denoise processing, and the Denoise signal waveform and spectrum will be displayed on the interface. Click the Play button at the bottom right of the interface to play the noise-reduced audio.

To demonstrate, we will test two example audio clips and check the effect. The first audio clip is a three-second speech voice.

The time and frequency domain waveform of the original signal is:

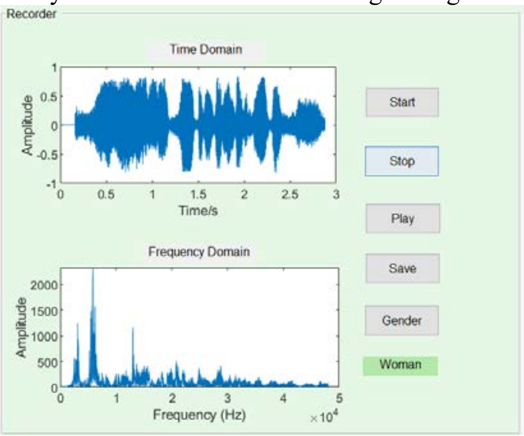

Fig. 4. First audio clip original signal.

The time and frequency domain waveform of the denoised signal is:

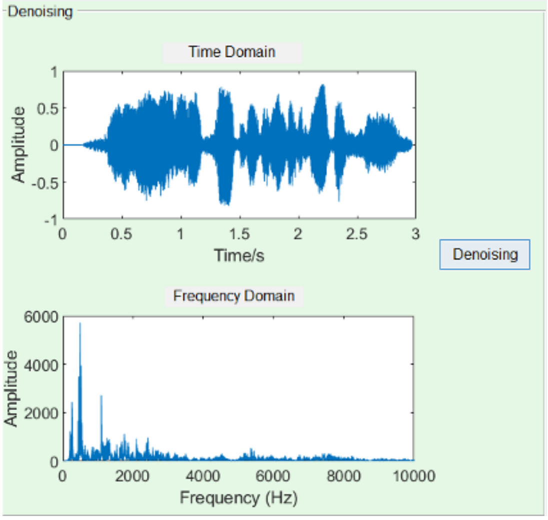

Fig. 5. First audio clip denoised signal.

The results were fantastic. According to common sense, the frequency of human voice is concentrated in 300-3400Hz. After noise reduction, the spectrum value of the waveform is concentrated between 300-3400Hz, which perfectly conforms to our expectations. In fact, after playing the audio noise reduction, you can hear the obvious noise reduction effect.

4 Conclusion

This project realizes the recording, playback, and processing of sound signals (LMS algorithm noise reduction processing, volume adjustment, reverse playback, double speed, etc.), and can display waveforms in the time domain and frequency domain, display the fundamental frequency of the sound, and perform men and women. Acoustic recognition. Through the LMS algorithm filtering, the audio tested in the experiment achieves a good noise reduction effect, the environmental noise is greatly reduced, and the human voice becomes very clear. The volume of the human voice can be artificially amplified by changing the amplitude coefficient, although there is also a non-negligible energy loss. Therefore, this paper realizes the extraction and enhancement of human voice, which also makes the judgment result more accurate and the sound processing effect more obvious.

References

1. Feriadi, Y. Simulation of active noise reduction using LMS algorithm: Synthetic and field data.
2. Do Prado, Robson A. On the Analysis of the Incremental ℓ -LMS Algorithm for Distributed Systems.
3. Staniec, Kamil. Analysis of the single frequency network gain in digital audio broadcasting networks.
4. Mendiratta, Arnav. Adaptive Noise Cancelling for audio signals using Least Mean Square algorithm.
5. https://blog.csdn.net/TIQCmatlab/article/details/114410085
6. https://ww2.mathworks.cn/help/matlab/ref/audiorecorder.html
7. https://baike.baidu.com/item/%E6%97%B6%E5%9F%9F%E9%A2%91%E5%9F%9F/9399325?fr=aladdin
8. https://www.cnblogs.com/sowhat1412/p/12734361.html
9. https://www.bilibili.com/video/BV1YW411K7bf?from=search&seid=18025662790709903802&spm_id_from=333.337.0.0
10. https://github.com/CodeSama346623/Bilibili346623