

Research on Acquisition and Process of Speech Signals

Xinxin Wang, Jianlin Zhao, Jinhui Guo, Lijun Wang*

School of Mechanical Engineering, North China University of Water Resources and Electric Power, Zhengzhou, Henan, 450045, China

Abstract — This system collected and processed speech signals through computer sound card, LabVIEW and MATLAB. Include two parts: acquisition playback and analysis treatment. Create interface through LabVIEW to control the sound card and realize the real-time recording and playback of speech signal. Low efficiency on numerical analysis and algorithm tool of LabVIEW does not meet the requirements. MATLAB has the function of mathematical calculation and powerful drawing, which can be well make up for lack of LabVIEW. This system combines LabVIEW and MATLAB to complete the basic research and operates simply. The MATLAB script node realizes communication between them. Connect data flow as the script node from LabVIEW to the input script node, open m file in the call node to realize signal processing. This system is a part of "human-computer interaction" to achieve the recording, playback, inflection play, time domain and frequency domain analysis of speech signal.

Keywords - speech signal, signal acquisition, signal processing, human computer interaction

I. INTRODUCTION

Speech signal is the most common mode of information communication in people's daily life and it is typical non-stationary random signal. Play an extremely important role in every aspect of people's lives. Along with the great-leap-forward development and the popularization of computer, interactive voice technology is used more and more widely. Speech recognition, synthesis, coding and recognition technology is more widely used in industry and daily life of the people. Through this study, can be collected to the characteristic parameters of speech signal, corresponding to one to achieve a specific speech and specific instructions, so as to achieve machine "understand" function, it will bring revolutionary changes to the whole world. Speech signal acquisition is the premise and foundation of the human-computer interaction. The sound card is an important part of a computer to process speech signal with amplifying filtering sampling, A/D and D/A module conversion. The computer comes with a tape recorder, and the audio can be saved as a document. Rely on external software to further processing and analysis, but the recording function is very limited and not extended, which gives the usual study of inconvenience [1, 2].

LabVIEW is a graphical compilation platform developed by national instrument company, which contains a complete set of function library to complete any programming tasks, to realize data acquisition, GPIB, serial control, data analysis, data display and data storage, etc. But for signal analysis and processing function, LabVIEW is a bit short to call the Script MATLAB Script node. The software MATLAB has powerful data analysis and processing functions, can be perfect to achieve the analysis and

processing of signal. Combine the advantage of LabVIEW and MATLAB, collect information effectively, analyze perfectly, provide effective parameters for the speech signal processing conveniently and quickly [3, 4].

The main work of this study mainly includes: ① A virtual experiment system for signal acquisition using LabVIEW. Achieve the acquisition of the voice signal stored in wav format, that is, record and storage through the sound card. ② Construct a music player using LabVIEW to playback of the acquisition of the speech signal and realize the pause, play, variable sound play, stop functions. ③ Analysis and processing by Fourier transform, noise, and filtering of speech signal by MATLAB to get the time domain waveform. Call MATLAB script node in the LabVIEW program block diagram to realize the software communication to analyze the speech signal quickly and accurately.

II. PREPARATION OF SIGNAL ACQUISITION AND PROCESSING

Sound card is also called the audio card, there are three basic functions: the first one is the music synthesis pronunciation function; the second one is the mixer processor and digital sound effects function; the third one is the simulation input and output functions of audio signal. In this paper adapt integrated sound card which is the sound card supports the integration of the chip sets. This sound card is integrated on the motherboard that not only does not occupy the PCI interface but also the cost is much lower. It is a very good choice that the use of the computer with a sound card to data acquisition. It is not only convenient and low cost of research and development, but also has the

function of A/D and D/A, good compatibility and high stability [5].

Laboratory Virtual Instrumentation Engineering Workbench (LabVIEW) is a graphical programming platform developed by American National Instrument Company. LabVIEW is also a general programming system the same as C language and assembly language. LabVIEW is a general programming system which contains a complete function library. It achieves data acquisition, GPIB, data analysis, data storage and display, and other functions. It is a graphical programming language that uses icons instead of text to create an application program. It is different from the traditional programming language that executes in the order of the sequence of statements and instructions. The flow direction of data flow in LabVIEW determines the execution order of VI and function. LabVIEW in the front panel control options designed a lot of controls similar to the actual instrument appearance, it can be easily created user interface in the front panel. The function control board provides a large number of function control icons in the rear panel (the program block diagram). Control the front panel objects by the icon link which is called G code. The functions used in this paper are mainly located in programming - graphics and sound - sound options, including the input, output, and the various functions of the file [6].

MATLAB is a commercial mathematics software produced by Math Works which is mainly used in the development of algorithms, data visualization, data analysis and numerical calculation. It mainly includes MATLAB and Simulink, which is integrated in an window environment easy to use with many powerful functions such as numerical analysis, matrix computation, scientific data visualization, and nonlinear dynamic system modeling and simulation.^[3] MATLAB has a series of functions and powerful signal processing ability. To realize the time domain transform of the speech signal by writing program and add noise or noise reduction processing.

LabVIEW combines the function of powerful Numerical calculation and drawing which greatly improves the development of Virtual Instruments. LabVIEW uses the script node to execute the MATLAB script.

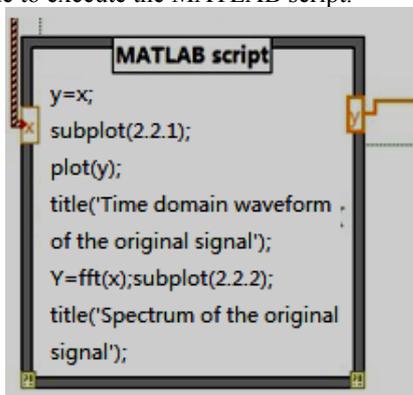


Fig.1 MATAB script

The MATLAB script node (Fig. 1) which is located in the function selection board - Mathematics - script and formula. Open the script node to operate, adjust the size and placement location. Add the input and output and edit the corresponding names for them by clicking the node box. The import of M file can be realized through two ways, directly paste the m file in advance to the node or import the program by clicking on the import function in the option. When using MATLAB script it should be noted that the user needs to develop the type of each data because MATLAB script cannot judge the data type. Select the appropriate type of data type by click on the option-data type. LabVIEW and MATLAB must match the data type. There are eight types of MATLAB data supported by LabVIEW such as real, complex, 1-D Array of real, 1-D Array of complex, 2-D Array of real, 2-D Array of complex, string path [7].

There are two aspects to be noted. Firstly LabVIEW uses MATLAB script node on ActiveX technology, so MATLAB script node can only be used on the windows platform. Secondly the computer must be installed with MATLAB 6.5 or more versions to use MATLAB script node.

III. SPEECH SIGNAL ACQUISITION AND PROCESSING

Speech signal is a typical non-stationary sound signal. The breathing movement and vocal cord vibration produces sound. Voice signal has "short term". In a short period of speech, the speech signal is kept in a relatively stable state. It is mainly concentrated in 300Hz to 3400Hz, when the sampling frequency is two times higher than the highest frequency of the speech signal, the information can be preserved. Therefore the sampling rate of 8000Hz can be used to obtain the discrete speech signal.

A. Collect Speech Signal

From Nyquist sampling theorem, in the process of analog-to-digital conversion, if the sampling frequency of f_s is greater or equal to two times of the maximum frequency, that is $f_s \geq 2f_{max}$, the information of the original speech signal can be preserved completely after sampling. In general practical application, the sampling frequency is five times to ten times of the signal frequency [8].

Sampling frequency, that is, the number of sound signals per second, the greater the sampling frequency, the more close to the real sound. It is mainly used to measure the quality of the sound card and voice file. The sampling frequency is generally divided into 8 KHz, 22.05 KHz, 11.025 KHz, and 44.1 KHz, a few can reach 48 KHz in today's mainstream civil sound card. The sampling number can be understood as the resolution of sound card to the processing of sound information. The bigger the sampling frequency is, the more close to the real record and playback sounds are, and now the popular is 16 bits sound card [9].

Use microphone built-in the computer, say a short speech to the microphone, the voice signal will be transmitted through the microphone to the sound card to get digital signal. There are many functions to support the sound card located in the VI Library of programming - graphics and

sound - the sound to control the sound card. Collect sound signal to create the front panel and the program block diagram (Fig. 2) which is mainly to achieve the recording and storage function. All the collected audio information are automatically saved as WAV sound format.

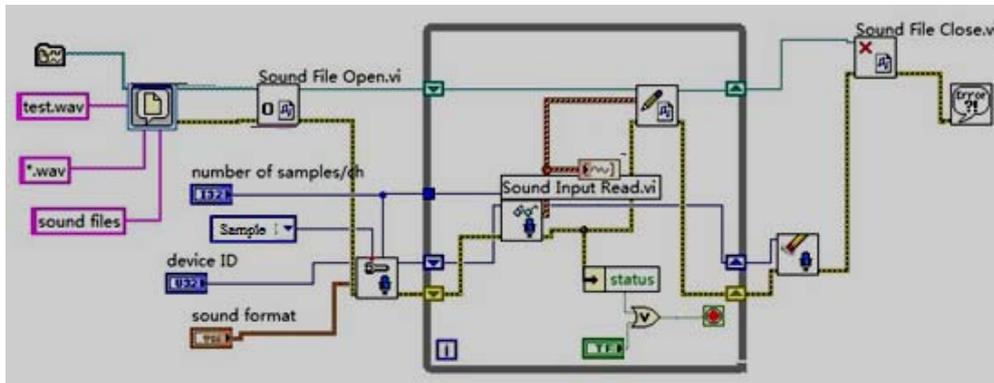


Fig.2 Program Block Diagram of Sound Signal Acquisition

The program block diagram takes the while loop structure. It is located in the function - structure -while loop, repeat the internal subroutine block diagram, until the conditional terminal (input) receives a specific Boolean value. Connect Boolean value to the terminal of while loop. Stop or continue while the true (T) is selected in the shortcut menu. While cycle operates at least one time and stops when the condition terminal is true. Error output and stop switch carries out “or” operation, stop recording as long as there is a real. That is, program runs stops when the program detects a mistake or users click on the stop button. The letter i lies on the left lower corner of while loop structure block diagram is used to record the number of cycles. File dialog box lies in the function - programming - file I/O- advanced file function, start path is default data directory, default name is test.wav, file types are wav format, type label is for sound file. When performing sound

collection procedures, a dialog box appears to tip for selecting the storage path of the sound file.

The configuration of sound output (Fig. 3) can be achieved on the acquisition of sound settings lies in function - Programming - graphics and sound - input. For example, the sampling number per channel is set to 10000, the sampling mode is continuous, voice format is the setting of the collection of sound, sampling rate is 22050, channel number is 2, and sampling bit number is 16.

The sound input is cleared, the sound file is closed, and the simple error processor (Fig. 4). The first two are in programming - graphics and sound. The last one lies in programming - dialog with the user interface - simple error handler. A dialog box appears and tips error when the program executes an error.

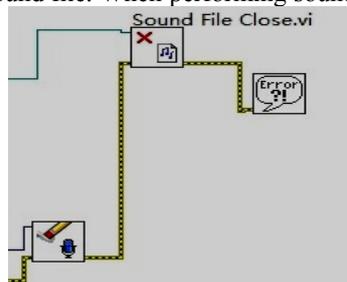


Fig.3 the Configuration of Sound Output

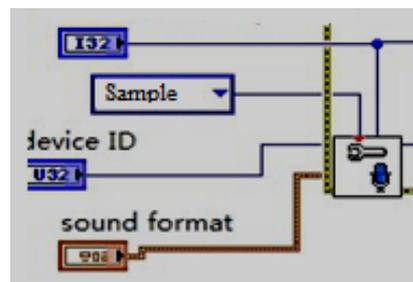


Fig.4 the Cleared Off

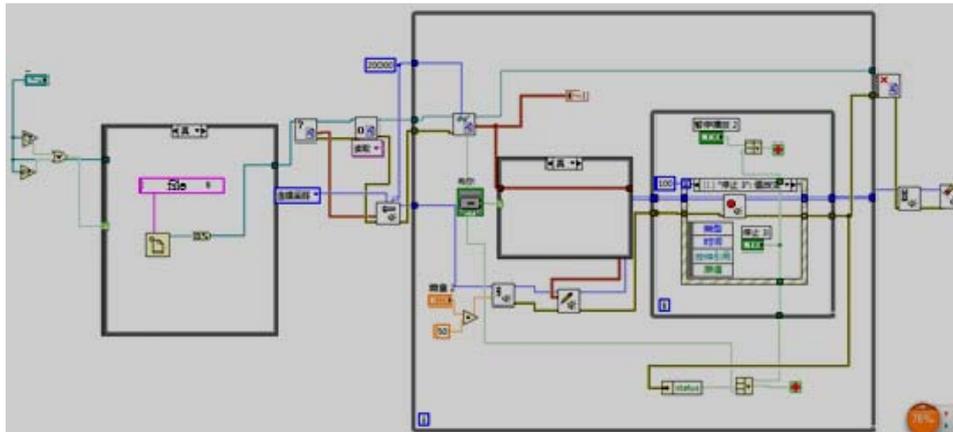


Fig.5 Program Block Diagram of Voice Signal

B. Playback Speech Signal

The program block diagram of voice signal (Fig. 5) used three different structures, respectively, the conditional structure, the while loop, the event structure. The conditional structure performs different paths according to different conditions of the value in the selector terminal, includes one or more sub block diagram or branch. When the structure is executed, it can only execute a subroutine block diagram or branch. The value of the selector terminal can be Boolean, string, integer, and enumerated type, used to determine the implementation of the branch. Add or remove a branch by right clicking on the structure border and enter the value of the conditional selector tag by the label tool and configure the value of each branch [10].

In the process of the playback of sound signals, there are two applications for the condition frame. The first one is to judge whether the path is empty strings, empty paths or illegal paths. Execute true branch if it conforms to one of them, system interface will jump out of a dialog box with “open a recording file”, that is, open a recorded wav format file. If judge a false branch, the path is valid (a wav format used to play the recording file) and execute it.

The second one is to judge whether to change sound playback. Connect a Boolean switch to the terminal of the condition selector. Produce true/false signal corresponding with true or false of the branch when the switch is pressed to perform acoustic playback or change playback. Resample based on the original speech waveform to get different play forms. Change the play effect through changing the

sampling point interval to reach the purpose of change sound. But the resampling leads to the change of sampling frequency which results in the decrease of the sound quality, there will be problems such as speed change.

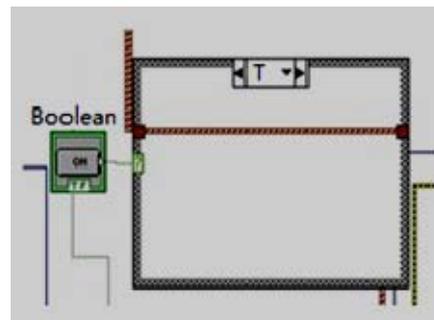


Fig.6 the True Branch

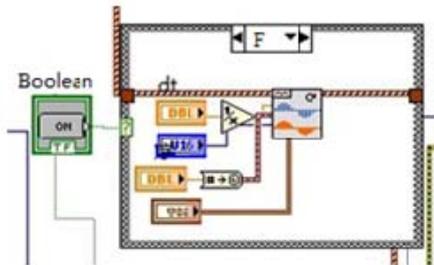


Fig.7 the False Branch

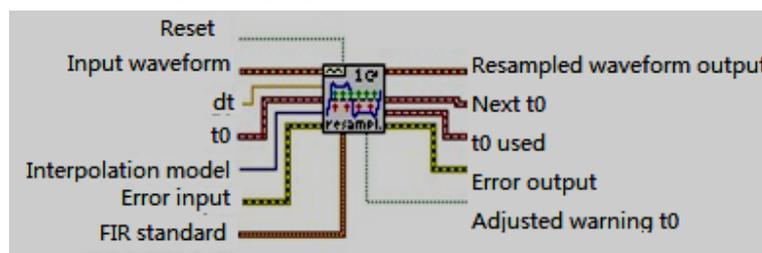


Fig.8 Waveform Resample

The true (Fig. 6) and false branch (Fig. 7), when select audio playback, perform true branch (Fig. 8). Data flow is directly through the branch and does not make any changes. Resample, adjust dt and interpolation mode to change the waveform so as to realize the different forms of sound. Wave resample is shown as Figure 8. Use continuous sampling, reset the value 0 of the phase input control, and default FALSE. dt is the sampling interval defined by the resampled waveform, obtained from reciprocal of the input control. Interpolation model is algorithm for specifying resampling, there are mandatory, linear, spline, FIR filter. Here adapt force, set each the value of output sample to be the closest to the input sample [11].

Event structure is located in function - programming - structure as Figure 13. Include one or more sub block diagrams or time branches. It waits until the event occurs and executes the corresponding branch to handle the event. Add a new branch and configure the event to be executed by right click the structure box. Line the "timeout" terminal in the left upper corner of the connection time frame, specify time in milliseconds waiting for the event to occur. Default value is -1 indicating that never timeout. In this paper set timeout waiting time is 100 milliseconds, that is, execute another branch the after the time and play sound from the cache. Also stop by the "stop" button, at the same time the outmost while loop will stop. Here use "or" operation of a composite operation. The three input is the error output, the end of the file, stop play button. Any one of the three is true, while loop stops so as to realize the stop of playing.

C. Time Display and Playback

Time display module is located in programming - time - acquisition time. Adopt the sequential structure in order to achieve the serial execution of the acquisition and playback. Appear a box to prompt the selection or input file path after pressing the button. Then input audio, appear a tip box "Open a recording file", select the recording file for music playback.

D. Analysis and Processing of Speech Signals

Analysis work mainly includes time domain and frequency domain analysis of speech signal noise reduction and filtering. Sample the speech signal by the wavread function and get the sampling frequency and points in MATLAB software programming platform. Call format of wavread function

`x=wavread(file)`, read the wav file provided by file, return the sample values in the vector x

`[x,fs,nbits]=wavread(file)`, the sampling value returns to the vector x , fs indicates the sampling frequency (Hz), $nbits$ represents the sampling number

`x=wavread(file, N)`, place the sampling value of the first N points in the vector y

`x=wavread(file, [N1,N2])`, the sample values of $N1$ to $N2$ points are placed in the vector x .

Firstly, draw the time domain waveform of the speech signal, secondly make fast Fourier transform using Fourier transform function `fft` to get the spectrum characteristics.

Procedures are as follows

```
[y,fs,nbits]=wavread('test.wav'); %Load the speech
signal is into MATLAB
sound(y,fs,nbits);%Playback speech signal
n=length(y); % the length of speech signal
Y=fft(y,n); %fast Fourier transform
subplot(2,1,1);plot(y);title('Time domain waveform of
the original signal');
subplot(2,1,2);plot(abs(Y));title('Original signal
spectrum')
```

Run result is as follows (Fig. 9):

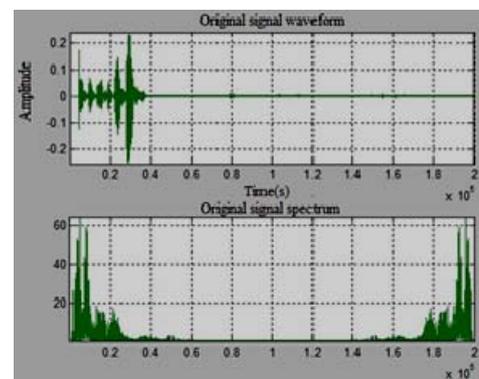


Fig.9 Signal Waveform and Spectrum

Procedure for adding noise signals are follows:

```
y=wavread('test.wav');
N=length(y);
noise=rand(N,2)/20;
k=y+noise; %Superposition of noise signals
K=fft(k);
subplot(2,1,1);plot(k);grid on;
title('Time domain waveform after adding noise');
subplot(2,1,2);plot(abs(K));title('Frequency domain
waveform after adding noise');
grid on; %Add grid
axis([0,45000,0,200]);
Waveform after adding noise (Fig. 10).
```

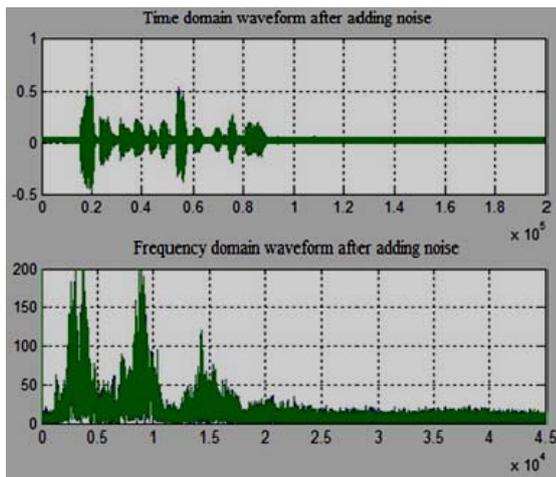


Fig.10 Waveform after Adding Noise

E. Digital Filter

The task of designing a digital filter is to find a linear time invariant system with specified frequency characteristics. The digital filter can be divided into two kinds of infinite impulse response (IIR) filter and finite impulse response (FIR) filter according to the time domain. IIR is commonly used to implement with a recursive model sometimes called recursive filter. The impulse response of FIR filter can only continue for a certain time [12].

IIR filter can only guarantee its amplitude frequency response to meet the performance index, phase characteristics cannot be considered and often nonlinear. The prominent advantage of the FIR filter is that it not only can satisfy the filter amplitude frequency response request but also obtain the strict linear phase characteristic. It is very important for high fidelity signal processing such as speech processing, data processing and testing. At the time of collecting sound signal, inevitably there will be some noise signal input. In this paper, FIR filter is adopted to filter the sound signal. Specifically the Kaiser window is selected, its characteristic (Fig. 11). Simulation results (Fig. 12) can be found that the effect of the noise signal can be effectively filtered through the Kaiser Window filter, and the processing effect of the speech signal is improved.

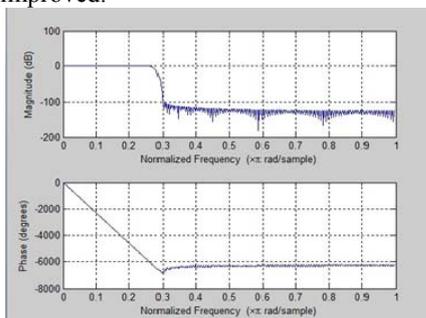


Fig.11 Characteristic Chart of Kaiser Window

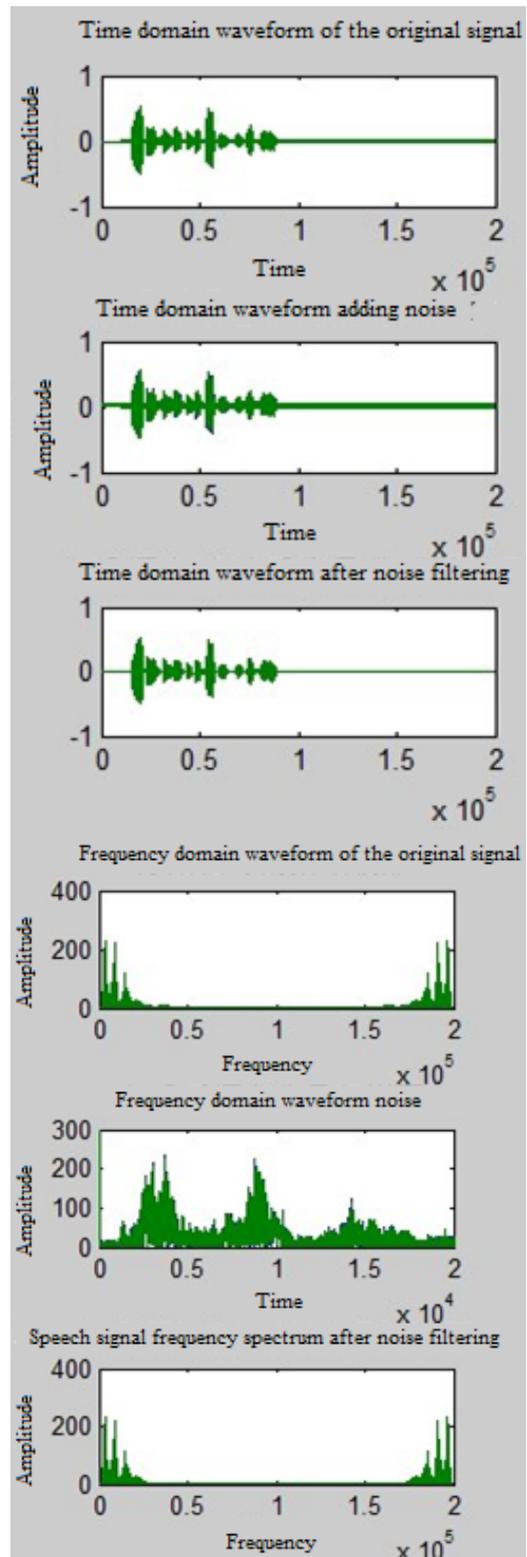


Fig.12 Comparison Chart of the Original Signal and Filtered Signal

IV. CONCLUSION

Collect and analyze the speech signal using computer sound card, LabVIEW and MATLAB in this paper. Structure interface to realize the control of sound card so as to input and playback real-time speech signal by LabVIEW. Powerful mathematical calculation and drawing functions of MATLAB meet the requirements of processing speech signal. This study realized successfully acquisition, playback, audio playback, change sound playback and performed time domain analysis, at last filter processing of white noise signal which is added to the sound signal and reached good effect.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

REFERENCES

- [1] Maas, Roland, On Bayesian networks in speech signal processing, 11th ITG Symposium on Speech Communication, September 24, 2014 - September 26, 2014.
- [2] LIU Qing-hua, CHEN Zi-qiang, The Construction of the Speech Signal Processing Course Based on MATLAB and DSP, *JOURNAL OF EEE*. Vol. 28 No. 4 Aug. 2006.
- [3] Chen Weiguo, Real Time Digital Speech Signal Processing System, *Theory and Applications*, Zhejiang University, 2004.
- [4] Zhong Linpeng, Studies on the Speech Signal Processing of the Speaker Recognition System, *University of Electronic Science and Technology of China*, 2013.
- [5] Despotovic, Vladimir, Design of nonlinear predictors for adaptive predictive coding of speech signals, 2013 21st Telecommunications Forum Telfor, November 26, 2013 - November 28, 2013.
- [6] Kolokolov, A.S. A method for speech signal processing based on band filtering of the logarithmic spectrum, *Automation and Remote Control*, v75, n3, pp.496-502, 2014.
- [7] Petrovsky, Alexander, Instantaneous harmonic analysis: Techniques and applications to speech signal processing, *Lecture Notes in Computer Science*, v8773, pp.24-33, 2014.
- [8] Tu, Lingying, Spectrum analysis of speech signal based on SPTool, *Advanced Research and Technology in Industry Applications*, WARTIA 2014, December 4, pp.899-902, 2014.
- [9] Wang Xiaofei, A reverberation robust target speech detection method using dual-microphone in distant-talking scene, *Speech Communication*, v72, pp.47-58, 2015.
- [10] Chen Hai-Hua, Speech signal dynamic time programming based on ant colony algorithm, *Harbin Gongye Daxue Xuebao*, v38, n10, pp.1758-1780, 2006
- [11] Ali Syed Imran, Speech and audio processing laboratory: Speech coding related signal processing modules, 2nd World Symposium on Web Applications and Networking, March 21, 2015 - March 23, 2015.
- [12] Ogunfunmi, Tokunbo, Speech over VoIP networks: Advanced signal processing and system implementation, *IEEE Circuits and Systems Magazine*, v12, pp.35-55, 2012.