Analysis of Speech Signals based on Short-time Fourier Method

Yu Jie

Shanghai University of Electric Power, Shanghai, China

Keywords: Non-equal interval sampling; fast Fourier transform; signal processing; spectrum analysis; speech signal; road target perception

Abstract: Aiming at the problem that the heterodyne signals inherent in pseudo-random (PN) code modulated speech signals cannot be equally spaced, a new short-time Fourier transform (DFT) method for non-uniform sampling signals is proposed. Firstly, the speech signal system model of distance velocity measurement is given, and the reason why the heterodyne signal can't be equally spaced is pointed out. Then, a new spectrum analysis method of non-uniform sampling signal is proposed by theoretical derivation. Finally, the simulation is verified by simulation. The method can be used to analyze the spectrum of non-equally spaced sampled signals. The results show that the method can still effectively analyze the frequency of unequal interval sampling Doppler signals in the Doppler frequency range generated by road moving targets with the received signal signal-to-noise ratio (SNR) of 0 dB.

1. Introduction

Fast Fourier transform algorithm plays an important role in the field of digital signal processing. It is an important means of signal spectrum analysis so far. It can compress the calculation of short-time Fourier transform of N-point data from O(N2). To O(NlogN), it is possible to analyze the signal spectrum of a large amount of data on an electronic computer. However, the execution of the FFT algorithm requires that its data distribution be limited to a uniformly distributed standard grid, which is often difficult to meet in practical applications. Non-standard sampling fast Fourier transform algorithm has a wide range of applications in practice. For example, non-uniform sampling caused by instability of electronic clock, reconstruction of computed tomography image, radar signal processing, etc. The problem of equally spaced sampling[1].

Currently, voice receivers only measure the target distance without measuring the target speed. It is proposed to use the continuous wave method to realize the synchronous measurement of the target distance and speed of the voice radar, but the peak power of the speech machine operating in the continuous wave mode cannot be too high, especially in the case of using the semiconductor speech device, the peak transmission power is only a few Milliwatts. Due to the diversity of road targets, it is highly probable that the target of low-reflectivity will reach the theoretical detection limit of voice radar when the peak power is low. Therefore, practical speech receivers use a pulsed method to measure the distance of the target. In order to ensure the correlation of the transmitted pulse sequence on the speech and audio, the random code is used to amplitude modulate the speech machine operating in the continuous mode, and then the modulated pulse signal is amplified by the voice fiber amplifier, so that the peak power of the transmission reaches about 100 W. Pulse detection method for voice radar. The reflected echo of the speech radar is mixed with the local continuous reference speech signal to obtain a discontinuous heterodyne signal, so that the equal interval sampling of the heterodyne signal cannot be realized[2].

The existing non-equal interval sampling Fourier spectrum analysis method includes: using the signal reconstruction method such as interpolation method, linear regular transformation, polynomial estimation and least squares method to realize the spectrum analysis method of the non-equal interval sampling signal. The characteristics of these methods are that the signal-to-noise ratio of the signal is high. If the signal is submerged in noise, the algorithm is not reliable. Radar echoes are usually very weak and are generally submerged in the noise of the receiver, so spectrum analysis cannot be performed by interpolation or zero-crossing detection. More difficult to use in

radar signal processing. The statistical average method requires a large amount of data, which inevitably reduces the real-time performance of the voice radar. Aiming at the specific application of speech receiver, this paper deduces a method to analyze the spectrum of non-equal interval sampling signals, and applies this method to the spectrum analysis of speech radar heterodyne signals, which can achieve better results.

2. Problem description

The speech device emits continuous speech, which is modulated by the external speech modulator and then projected by the collimating lens to the target. The target reflected signal is collected by the receiving lens and coupled into the voice fiber, and then the reflected echo is split into two paths by the voice separator, one of which is irradiated onto the voice electrical converter PD2, and the other is mixed with the reference signal in the voice coupler. The frequency is radiated onto the voice electrical converter PD1. Since the voice electrical converter is a device sensitive to speech intensity, its output is a heterodyne signal that reflects the echo and the local reference signal[3].

To detect the target distance with high precision, it is necessary to increase the bandwidth of the transmitted signal. In this paper, the pseudo-random code is used to modulate the intensity of the voice radar transmit signal to achieve spread spectrum and improve the range resolution. At the receiving end, the electrical signal output by the PD2 is correlated with the modulation code, and the time point corresponding to the correlation peak is found to measure the flight time of the speech from the transmission to the reception, thereby measuring the distance of the target.

3. Signal modulation

If the transmitted speech is unmodulated continuous speech, the signal output by PD1 will be a continuous sine wave whose frequency corresponds to the frequency of the Doppler signal produced by the target motion velocity. However, when the transmitted signal is a signal modulated by a pseudo random code, the signal output by the PD1 will become a discontinuous sine wave, wherein the shape of the sine wave is still maintained in the "1" time slot of the modulation code, and the modulation code is "0". The signal within the time slot will tend to zero level. Due to the randomness of the pseudo-random code, that is, "0" can exist at any time, regardless of the sampling rate used, equal interval sampling of the Doppler signal cannot be obtained. The FFT algorithm that uses the equally spaced sampling of the sampled data directly cannot stably obtain the peak of the signal spectrum, that is, the speed of the target cannot be reliably measured[4].

In the communication system, the signal-to-noise ratio of the received signal is generally above 20 dB, and the time position level of the code "0" can be recovered by the envelope detection method. However, in the radar system, since the received echo signal passes through object reflection, spatial scattering, atmospheric absorption, etc., the echo signal arriving at the receiver is very weak, usually submerged in the thermal noise of the receiver, and the signal-to-noise ratio of the received signal. Generally below 0dB. The signal provided by the position of the code "1" is not reliable in itself, so the data obtained by the envelope detector does not represent the envelope of the signal itself, that is, the original continuous sine wave can't be obtained by envelope detection. Le signal. For the same reason, the interpolation method cannot be used, whether it is sin function interpolation or B-spline interpolation, and the frequency analysis work is performed on the signal submerged in the noise. In order to analyze the frequency of the heterodyne signal submerged in the noise, it is necessary to design a new method that does not use the amplitude at the code "1" for evaluation.

In order to perform spectrum analysis on a pseudo-random code modulated speech radar receiving heterodyne signal, it is necessary to know the interval between the transmitted pulses. These intervals can be counted by the counter and stored in memory during the phase of pseudo-random code generation. However, due to the existence of the random delay of the received signal, the position of the first pulse in the heterodyne signal cannot be predicted. Therefore, the

position of the Doppler signal must be found by means of the distance information calculated by the output signal of the PD2, and then pulse interval determines the data points of the individual Doppler signals, ensuring that the sampled data points are Doppler signals that are subject to noise interference, rather than pure noise. Since the delay between PD1 and PD2 is fixed, depending on the length of the speech fiber between the two, it is obviously feasible to determine the head position of the Doppler signal by the distance measured by PD2. The calculation matrix is shown in Equation 1.

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{cases} c_1 + c_2 + c_3 + c_4 = 0 \\ c_1 + c_5 + c_6 + c_7 = 0 \\ c_2 + c_5 + c_8 + c_9 = 0 \\ c_3 + c_6 + c_8 + c_{10} = 0 \\ c_4 + c_7 + c_9 + c_{10} = 0 \end{cases}$$

(1)

4. Simulation results and analysis

Since the distance measurement part only calculates the correlation function between the PD2 output signal and the modulation code, and determines the speech time of flight from the peak position of the correlation function, there is no problem in principle, so only the speed measurement part is simulated, and the non-equal interval sampling signal spectrum is verified. The feasibility of the analysis determines the feasibility of measuring the velocity while measuring the distance. In the case where the speech radar transmits a signal with a wavelength of 1550 nm, a target speed of 1 m/s (corresponding to the walking speed of the pedestrian) produces a Doppler signal with a frequency of 1.29 MHz. Assuming that the moving target speed is in the range of 1 m/s (walking speed) to 100 m/s (relative speed of the highway vehicle in the direction of relative motion), the frequency of the Doppler signal varies over a wide frequency range of 1.29 to 129 MHz. In addition, the voice radar receives the reflected signal of the target, and the received signal is generally weak, limited by the current electronic level, and the received signal of the PD2 output and the heterodyne signal of the PD1 output are drowned in the noise. In the simulation, the echo power is calculated according to the radar equation on the condition that the peak power of the transmission is 100W and the target reflectivity is 10%. According to the noise level of the existing transimpedance amplifier, the signal-to-noise ratio of the PD1 output Doppler signal is set to 0 dB. The schematic diagram of depth simulation is shown in Figure 1.

The width of the modulation code is set to 2 ns, and the signal length of 770 ns is used to complete a single measurement of the target. The setting of these two parameters ensures that the voice radar has a high range resolution (tens of centimeters) and a high scanning speed. (to reach or exceed 200,000 times per second) to meet the requirements of the road environment in the smart driving field. To demonstrate how the algorithm works, noise is ignored in the graph. The dotted line represents the pseudo-random code modulated Doppler signal output by PD1. The dotted waveform is sampled, and the FFT is used and frequency-converted. The resulting signal spectrum does not contain a peak corresponding to the signal frequency, so the Doppler signal cannot be resolved. Frequency of. The amplitude correction is performed on the code "1" waveform before the code "0" interval, and one data is obtained by sampling each modified code "1", and then the FFT of the sampled data can be used to correctly parse the pseudo-output of the PD1 output. The frequency of the Doppler signal modulated by the random code[5].

The sine wave of the six frequencies with Doppler frequencies of 1.29, 5, 10, 20, 50, and 129 MHz is multiplied by the PN code to obtain the heterodyne signal of the PD1 output, and the Doppler signal modulated by the PN code is superimposed on Gaussian. Noise, the output signal-to-noise ratio is 0dB; then the code "1" time slot signal amplitude correction, the results show that the proposed method can correctly resolve the frequency of the Doppler signal generated by the

road moving target, so as to measure the target distance at the same time, the speed of the target can be measured.
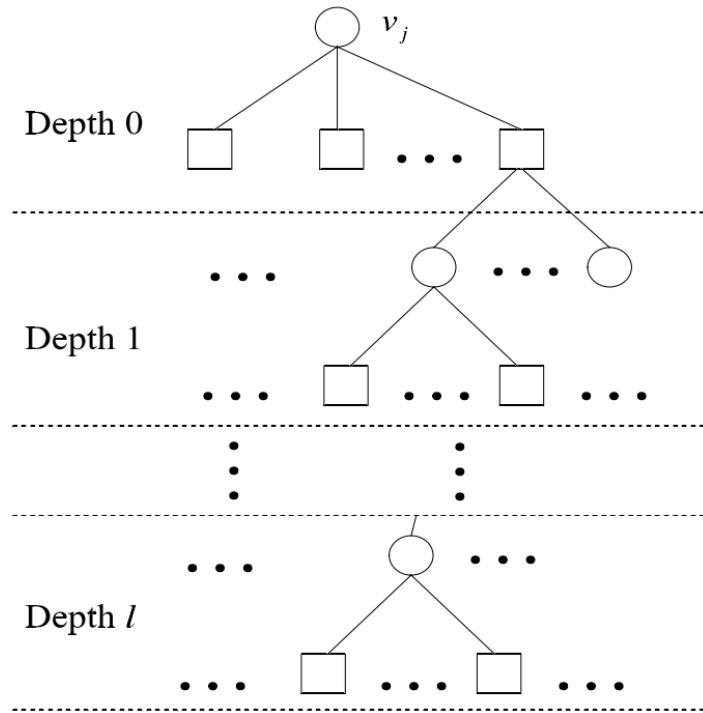

Figure 1 Schematic diagram of depth simulation

5. Conclusion

This paper first introduces the speech receiver model, working principle and its inherent non-equal interval sampling heterodyne signals for simultaneous measurement of target distance and velocity. It is pointed out that the frequency of Doppler signals cannot be directly analyzed by FFT algorithm; The method for calculating the signal spectrum under the condition of knowing the interval of each sampling point is obtained, that is, the method of correcting the signal waveform of the voice radar by using the sampling point interval, and directly using the FFT method for the corrected waveform; finally, receiving the signal The Doppler frequency is in the range of 1.29～129MHz. The effectiveness of the method is verified by computer simulation by selecting six frequency points, and the method is independent of the initial phase of the received signal. In particular, in the case where the received signal signal-to-noise ratio is 0 dB, the method of the present invention can effectively calculate the spectrum of the signal.

References

[1] Xiurong Ma,Dan Liu,Yunlong Shan. Intra-pulse modulation recognition using short-time ramanujan Fourier transform spectrogram[J]. EURASIP Journal on Advances in Signal Processing,2017,2017(1).

[2] Pravin Vasantrao Gadkari,M.G. Shashidhar,Manohar Balaraman. Delivery of green tea catechins through Oil-in-Water (O/W) nanoemulsion and assessment of storage stability[J]. Journal of Food Engineering,2017,199.

[3] Berrin Tansel,Nadir Dizge,Ibrahim N. Tansel. Analysis of high resolution flux data to characterize fouling profiles of membranes with different MWCO under different filtration modes[J]. Separation and Purification Technology,2017,173.

[4] Nokhaeng Lee,Youngjin Park,Gun Woo Lee. Frequency-domain active noise control for magnetic resonance imaging acoustic noise[J]. Applied Acoustics,2017,118.

[5] C. Montalvo,C.J. Gavilán-Moreno,A. García-Berrocal. Cofrentes nuclear power plant instability analysis using ensemble empirical mode decomposition (EEMD)[J]. Annals of Nuclear Energy,2017,101.