UDC 004.048

SENSOR DATA FREQUENCY FEATURE EXTRACTION

ZHAO DI

Belarusian State University of Informatics and Radioelectronics, Republic of Belarus

Received March 8, 2022

Abstract. Frequency domain features are important for human activity recognition. However, the raw signal needs to be converted to frequency domain features using a fast Fourier transform. In the frequency domain, the time series data of each component is converted by using the Fast Fourier Transform (FFT) algorithm.

Keywords: Fast Fourier algorithm, Feature extraction.

Introduction

In a human activity recognition (HAR) system, to build a good performance recognition system, feature extraction is an important aspect. Some studies suggest that the use of efficient feature extraction methods may improve the final performance better than using an effective classifier [1]. Frequency domain analysis can observe the characteristics of the signal based on the frequency. Generally speaking, the frequency domain analysis is more concise, and observing the signal on the frequency domain makes the analysis of the problem more profound and convenient. The frequency domain feature extraction algorithm adopted in this article is a fast Fourier algorithm, which can transform the sample signal to the frequency domain and extract the frequency domain features of the sample signal to facilitate the frequency domain analysis.

The frequency domain characteristic parameters

Spectrum analysis includes amplitude spectrum and phase spectrum, and amplitude spectrum is the most commonly used. When the health state of the key components of the reducer changes, the frequency component in the signal spectrum of the sample data changes accordingly [2]. Therefore, we can first accurately characterize the signal spectrum information by analyzing the frequency domain characteristics of the signal, and then obtain the human behavior state under different situations. The frequency domain characteristic parameters used in this thesis include center of gravity frequency, mean frequency, root mean square frequency and frequency standard deviation.

1. Center of gravity frequency.

The center of gravity frequency (FC) can describe the frequency of large signal components in the spectrum, reflecting the distribution of the signal power spectrum. In other words, for a given band range, the energy contained below the center of gravity frequency is half the total energy of the signal. In other words, the center of gravity frequency is a weighted average of the amplitude of the power spectrum as the weights, so the center of gravity frequency moves to locations where the power spectrum amplitude is larger. The parameters defined by using the expression:

$$\mathrm{FC} = \frac{\int_0^{+\infty} f\,\overline{p}(f)\,df}{\int_0^{+\infty} p(f)\,df}, \tag{1}$$

where f is frequency of the power spectrum $p(f)$ of the signal, $p(f)$ is the power spectrum of the signal.

2. Mean square frequency.

The mean square frequency is the square of the RMS frequency. Unlike the center of gravity frequency, the mean square frequency is the weighted average of the square of the signal frequency, weighted by the amplitude of the power spectrum. The mean square frequency (MSF) is describing the change in the position of the main band of the power spectrum. The parameters defined by using the expression:

$$\mathrm{MSF} = \frac{\int_0^{+\infty} f^2 p(f)\,df}{\int_0^{+\infty} p(f)\,df}, \tag{2}$$

where f is frequency of the power spectrum $p(f)$ of the signal, $p(f)$ is the power spectrum of the signal.

3. Center of gravity frequency.

The root mean square frequency is the calculated square root of the mean square frequency. It also describes the position changes of the main frequency band of the power spectrum. The parameters defined by using the expression:

$$\mathrm{RMSF} = \sqrt{\mathrm{MSF}}. \tag{3}$$

4. Frequency variance.

The frequency variance is the square of the frequency standard deviation, another metric measure of the energy dispersion of the power spectrum. Frequency standard deviations Frequency standard deviation is the radius of inertia centered on the center of gravity frequency. If the amplitude of the spectrum near the center of gravity is large, the frequency standard deviation is small; if the spectrum near the center of gravity is small. The frequency standard deviation describes the degree of dispersion of the power spectrum energy distribution. The parameters defined by using the expression:

$$\mathrm{VF} = \frac{\int_0^{+\infty} (f - \mathrm{FC})^2 p(f)\,df}{\int_0^{+\infty} p(f)\,df}, \tag{4}$$

where f is frequency of the power spectrum $p(f)$ of the signal, $p(f)$ is the power spectrum of the signal.

In order to quickly calculate the above parameters quickly the following formula based on the discrete Fourier transform properties is used:

$$F_{\mathrm{FC}} = \frac{\sum_{n=1}^{N} p(f) p(f)}{2\pi \sum_{n=1}^{N} p(f)^2}, \tag{5}$$

$$F_{\mathrm{FC}} = \frac{1}{N} \sum_{n=1}^{N} p(f), \tag{6}$$

$$F_{\mathrm{RMSF}} = \sqrt{\frac{\sum_{n=1}^{N} p(f)^2}{4\pi^2 p(f)^2}}, \tag{7}$$

$$F_{\mathrm{RVF}} = \sqrt{\frac{\sum_{n=1}^{N} (-F_{\mathrm{FC}})^2 p(f)}{\sum_{n=1}^{N} p(f)}}, \tag{8}$$

where $\overline{p}(f) = \frac{p(f) - p(f-1)}{\Delta}$, Δ is the frequency resolution interval, and N is the number of frequency sample points.

Pros and cons of frequency feature extraction

1. Pros of frequency feature extraction:

– frequency domain analysis can observe the signal features according to the frequency, and the frequency domain representation is more concise;

– the physical significance of frequency domain analysis is clear and it can provide more intuitive feature information than the time domain.

2. Cons of frequency feature extraction:

– frequency feature extraction can only be conducted in the effective interval;

– because the time domain truncation causes energy leakage, the amplitude of the discrete spectrum, phase and frequency can produce large errors.

Program code implementation

In Matlab, firstly we need to fast Fourier transform the signal, and then use the four features of the center of gravity frequency, the mean square frequency, root mean square frequency and the Frequency variance to calculate the eigenvalue of the frequency domain. The program code implementation of the article in Figure 1.

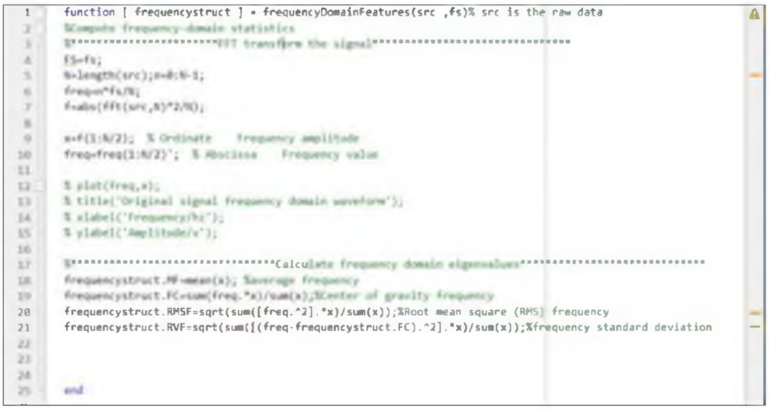

Figure 1. Program code implementation

Conclusion

Frequency domain analysis can observe the signal characteristics according to the frequency. Generally, the time domain analysis is more intuitive, but the frequency domain analysis is more concise, and observing the signals in the frequency domain makes the analysis of the problem more profound and convenient. At present, frequency domain analysis has become a trend of signal analysis. Through feature extraction, we can extract the center of gravity frequency, mean frequency, root mean square frequency and frequency standard deviation in three directions of x, y, and z, a total of twelve feature vectors, we can carry out the next step of data analysis work through these twelve feature vectors, that will become convenient and concise.

References

1. Wu X. [et al.] // IEEE Trans. On KDE. 2014. Vol. 26. P. 97.
2. Manual J. [et al.] // IJCV. 2022. Vol. 130. P. 199–200.