Discourse Feature Recognition for Text Dynamic Translation

Meng Li
Guangzhou Huali College, Guangzhou, 511325, China

Keywords: Text, Dynamic Translation, Discourse Feature Recognition, Parameter Modulation, Spectrum Feature Extraction.

Abstract: In order to improve the ability of dynamic translation of foreign media discourse, this paper puts forward a method of discourse feature recognition and recognition based on spectral density feature decomposition. Using time-frequency dynamic feature analysis, the dynamic compensation model of foreign media text translation discourse signal is established, and the noise parameters of foreign media text translation discourse signal are analyzed and recognized by combining the frequency spectrum density feature estimation method. The feature decomposition model of foreign media text translation discourse signal is constructed by multi-frame speech compensation and parameter modulation method, and the wavelet multi-level feature detection method is adopted. The anti-interference filtering analysis of foreign media text translation discourse is realized. The conditional coding of text translation discourse signal is carried out by the method of spectral density dynamic feature decomposition, and the statistical information analysis model of foreign media text translation discourse signal is established. The signal detection and feature recognition of foreign media text translation discourse signal are realized by text parameter clustering and envelope amplitude-frequency feature detection. The test results show that the accuracy of foreign media text translation discourse feature recognition by this method is high, and the output signal-to-noise ratio of the signal is improved by discourse feature recognition.

1 INTRODUCTION

With the development of foreign media text and discourse recognition technology, the accuracy of foreign media text translation and discourse recognition is required higher and higher. A discourse feature recognition model oriented to dynamic text translation is established, and a foreign media text translation discourse signal collection and feature recognition model is built by combining signal processing and speech signal analysis methods (John, 2019). The feature recognition and detection of foreign media text translation discourse signals are carried out by combining anti-interference design methods. Improve the ability of signal detection and recognition, so as to improve the output clarity of foreign media text translation discourse signals and reduce the interference of signals. Studying the feature recognition technology of foreign media text translation discourse signals plays an important role in the further development of foreign media texts (JIA, LAI, YU et al., 2021).

The feature recognition of foreign media text translation discourse signal is based on the automatic test technology of the signal, and adopts the embedded signal processing method combined with signal feature analysis to carry out anti-interference processing and feature analysis of foreign media text translation discourse signal. Among the current methods, the feature recognition methods of foreign media text translation discourse signal mainly include time-frequency analysis method, statistical feature analysis method, differential control method and spectral feature recognition method, etc (AI, YANG, XIONG, 2018). The error gain control model of foreign media text translation discourse signal is constructed, and combined with signal gain control, the foreign media text translation discourse feature recognition processing is realized, and the signal interference is reduced. However, the traditional method of foreign media text translation discourse feature recognition has poor output signal-to-noise ratio and weak adaptive control ability (Wang, Li, Meng, 2022).

Aiming at the disadvantages of traditional methods, this paper puts forward a method of foreign media text translation discourse feature recognition and

Li, M.
Discourse Feature Recognition for Text Dynamic Translation.
DOI: 10.5220/0011752300003607
In *Proceedings of the 1st International Conference on Public Management, Digital Economy and Internet Technology (ICPDI 2022)*, pages 578-583
ISBN: 978-989-758-620-0

recognition based on spectral density feature decomposition. Firstly, the collected foreign media text translation discourse signals are preprocessed, including noise reduction filtering, anti-interference design, feature extraction, etc., and then the signals are decomposed by time-frequency decomposition and empirical mode decomposition to realize foreign media text translation discourse feature recognition. Finally, the simulation test shows the superior performance of this method in improving foreign media text translation discourse feature recognition and recognition ability.

2 SIGNAL ACQUISITION DETECTION AND PREPROCESSING

2.1 Signal Sampling Analysis

In order to realize foreign media text translation discourse feature recognition based on spectral density feature decomposition, firstly, the foreign media text translation discourse signal is sampled, the time-frequency dynamic feature analysis of foreign media text translation discourse is adopted, and a dynamic compensation model of foreign media text translation discourse signal is established. M omni-directional signal sources are set to collect foreign media text translation discourse signal, an expected foreign media text translation discourse signal and p interference feature quantitie (Zheng, 2022) s. The collected foreign media text translation discourse signals are input into the TV signal detection system in array, and there is a foreign media text translation discourse signal with a length of N, which can be represented as array, and it is represented as $x(n)\in R^N$. Through the encoding and decoding of foreign media text translation discourse signals, the characteristic marker model of foreign media text translation discourse signals is obtained, and the acquisition output of foreign media text translation discourse signals is expressed as follows:

$$x=\sum_{i=1}^{N} s_i\Psi_i=\Psi s \tag{1}$$

Wherein, s_i is the time domain component of foreign media text translation discourse signal, Ψ_i is the phase of foreign media text translation discourse signal detection, and Ψs is the frequency domain component of foreign media text translation discourse signal. The acquisition model of foreign media text translation discourse signal is constructed, and the best sampling starting point is obtained:

$$J^*(m)=\max_{\tau}\{J^*(\tau)+D_m(\tau)+C\},\quad J^*(0)=0 \tag{2}$$

Wherein, $J^*(\tau)$ is the time delay of foreign media text translation discourse signal detection, $D_m(\tau)$ is the characteristic decomposition level of foreign media text translation discourse signal, and C is the bandwidth of foreign media text translation discourse signal. According to the constraint characteristics of the modulation mode of foreign media text translation discourse signals, the discrete components of broadband multi-frame foreign media text translation discourse signals after basic decoding are obtained, and the discrete characteristic equation is expressed as follows:

$$x(n)=s(n)+v(n)=\omega_{k-1}^{(i)}\frac{p\left(y_k\mid X_k^{(i)},Y_{k-1}\right)p\left(x_k^{(i)}\mid X_{k-1}^{(i)},Y_{k-1}\right)}{q(x_k^{(i)}\mid .)} \tag{3}$$

In the above formula, $s(n)$ represents the discrete sequence of foreign media text translation discourse signal, $v(n)$ represents the color noise component, $\omega_{k-1}^{(i)}$ represents the scale of foreign media text translation discourse signal, $q(x_k^{(i)}\mid .)$ is the prior probability density of foreign media text translation discourse signal detection, $p\left(x_k^{(i)}\mid X_{k-1}^{(i)},Y_{k-1}\right)$ is the conditional probability density of foreign media text translation discourse signal reliability collection, and SD is the random probability density component. The domain feature expression is carried out on the broadband multi-frame foreign media text translation discourse signal, and the joint parameter identification is carried out on the foreign media text translation discourse signal components in Y to construct the conditional coding model of foreign media text translation discourse signal. The coding output is as follows:

$$p(\boldsymbol{x}_l^l\mid\alpha)=\prod_{i=1}^{L}p(\boldsymbol{y}_i\mid\alpha,r_i,l) \tag{4}$$

Wherein, $p(y_i\mid\alpha,r_i,l)$ is the encoded output of foreign media text translation discourse frame format, and L is the sampling sequence length of foreign media text translation discourse signal. It is assumed that the foreign media text translation discourse signal is decomposed by the adaptive spectrum separation method, and the sampling model of the foreign media text translation discourse signal is obtained (Wang, 2021).

2.2 Signal Filtering Pretreatment

By using time-frequency dynamic feature analysis, a dynamic compensation model of foreign media text translation discourse signal is established (Lin, 2021), and the signal anti-interference design is carried out by combining the estimation method of frequency spectrum density feature when foreign media text translation discourse signal is used, and the modulation components of foreign media text translation discourse signal are obtained:

$$E_j = \sum_{k=1}^{n} E_{j,k} \tag{5}$$

$$P_{j,k} = E_{j,k}/E_j \tag{6}$$

Wherein, $E_{j,k}$ is the order of foreign media text translation discourse signal, and E_j is the Doppler frequency offset of foreign media text translation discourse. Time-frequency analysis method is used to filter the foreign media text translation discourse signal in the interval, and wavelet multi-level feature detector is used to obtain the modulation frequency spectrum WE_k of foreign media text translation discourse signal as follows:

$$WE_k = -\sum_j P_{j,k} \, ln(P_{j,k}) \tag{7}$$

Wherein, $P_{j,k}$ is the carrier frequency of foreign media text translation discourse signal, and $\ln(P_{j,k})$ is the amplitude of filtering detection. The interference suppression component is determined by the order determination algorithm, and the modulation processing of the broadband multi-frame foreign media text translation discourse signal is carried out. The input characteristic sequence of the broadband multi-frame foreign media text translation discourse signal is $\mathbf{x} = [x(0), ..., x(N-1)]$, which is a discrete sequence of the finite-length broadband multi-frame foreign media text translation discourse signal. $0 \le n \le N-1$ D, the empirical mode decomposition method is adopted to filter and detect the broadband multi-frame foreign media text translation discourse signal, and the average sum of the filtered output of foreign media text translation discourse is obtained.

$$S_x = E[x^3(t)] + \sqrt{s}b \tag{8}$$

$$K_x = E\left[x^4(t)\right] - 3E^2\left[x^2(t)\right]b \tag{9}$$

Wherein, $E[x^3(t)]$ is the passband of foreign media text translation discourse signal transmission, b represents multi-frame information gain, $E\left[x^4(t)\right]$ is the modulation frequency offset of foreign media text translation discourse, and s is the joint probability density distribution. According to the filtering results, the multi-frame speech compensation and parameter modulation method are used to construct the feature decomposition model of foreign media text translation discourse signal, which can improve the ability of discourse feature recognition and processing (ABDUL RAUF, SCHWENK, 2011).

3 DISCOURSE FEATURE RECOGNITION AND PROCESSING

3.1 Empirical Mode Decomposition of Signal

The method of multi-frame speech compensation and parameter modulation is used to construct the feature decomposition model of foreign media text translation discourse signal, and the method of spatial parameter recognition is used to construct the ambiguity detection model of foreign media text translation discourse signal (MARIE, FUJITA, 2017). After the interference position is determined and zeroed, the dynamic amplitude feature recognition model of broadband multi-frame foreign media text translation discourse signal is obtained. The joint decision function is:

$$\begin{cases} H_0: x'(t) = w(t) \\ H_1: \sqrt{E}s'(t) + w(t) \end{cases} \quad 0 \le t \le T \tag{10}$$

In the formula, x'(t) and s'(t) are:

$$x'(t) = x(t)*h_w(t) \tag{11}$$

$$s'(t) = s(t)*h_w(t) \tag{12}$$

Wherein, $w(t)$ is the single-line spectrum of foreign media text translation discourse signal, T is the time sampling interval, E is the distributed energy spectrum of foreign media text translation discourse, $x'(t)$ is the convolution of foreign media text translation discourse signal, $h_w(t)$ is the filter transfer function of foreign media text translation discourse signal, the M-order phase polynomial is determined, and the improved empirical mode decomposition method is adopted. The scale decomposition coefficient of the broadband multi-frame foreign media text translation discourse signal is $N^{(1)} = N, \quad N^{(j)} = N_0^{(j-1)}, \quad 2 \le j \le J$, and each sub-signal is described by $\sum_j P_j = 1$, so the output probability density characteristic quantity of the broadband multi-frame foreign media text translation discourse signal satisfies. The energy set of the broadband multi-frame foreign media text translation discourse signal is described by $\{P_1, P_2, \dots, P_j\}$, and the single-line spread characteristic decomposition method is adopted. The residual frequency offset of the broadband multi-frame foreign media text translation discourse signal after discrete orthogonal wavelet transform is obtained, and $\{P_1, P_2, \dots, P_j\}$ is used to represent the finite time sequence of the broadband multi-frame foreign media text translation discourse signal, that is, the discrete sequence output by empirical mode decomposition of the foreign media text translation discourse signal is:

$$\mathbf{X} = [X(0), \dots, X(N-1)] \tag{13}$$

Wherein, N is the discrete sequence length of foreign media text translation discourse signal. According to the empirical mode decomposition of foreign media text translation discourse signal, the frequency parameters of foreign media text translation discourse signal are estimated as follows:

$$\hat{f}_i(n) = \frac{1}{2\pi} \sum_{i=0}^{p} i a_i n^{i-1} \tag{14}$$

Wherein, i is the sampling sequence point of foreign media text translation discourse signal, and a_i is the sampling track of foreign media text translation discourse signal, n^{i-1} is Z transform. At the receiving end, through information output identification, the time domain discrete components of foreign media text translation discourse signals are expressed as follows:

$$F = \{x_1, x_2, \cdots, x_n\} \tag{15}$$

Thus, the discrete components of the signal are obtained by empirical mode decomposition. According to the constraint feature quantity of foreign media text translation discourse signal modulation mode, the signal discrete sequence of foreign media text translation discourse signal modulation is obtained, and the feature decomposition model of foreign media text translation discourse signal is constructed by multi-frame speech compensation and parameter modulation method, so as to improve the feature recognition ability of the signal (ZHANG, LI, YANG et al., 2020).

3.2 Realization Of Discourse Feature Recognition

The amplitude of modulation mode of foreign media text translation discourse signal obtained by frequency domain parameter identification is:

$$y(t) = \iint_{a,b} \rho(a,b) - \frac{1}{2\pi} \frac{d}{df} \arg[Z(f)] \frac{dadb}{a^2} \tag{16}$$

In the above formula, $f(t)$ is the estimated value of modulation parameters of foreign media text translation discourse signal, $\rho(a, b)$ is the ambiguity of foreign media text translation discourse signal, and a is the broadband multi-frame length of foreign media text translation discourse signal. The modulation model of foreign media text translation discourse signal is constructed, and the modulation output expression is as follows:

$$K(u) = Z(f) + A(f) e^{j\theta(f)} \tag{17}$$

Wherein, $Z(f)$ represents the real part of foreign media text translation discourse signal, $A(f)$ represents the i333maginary part, and $e^{j\theta(f)}$represents the frequency offset. According to the foreign media text translation discourse signal modulation processing, the output expression of discourse feature recognition can be expressed as:

$$Y(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(f) e^{j2\pi ft} df \tag{18}$$

Wherein, $S(f)$ represents the energy spectrum component of foreign media text translation discourse, and f represents the modulation frequency of foreign media text translation discourse. The steep gradient of modulation of foreign media text translation discourse signal is calculated. Based on the intelligent recognition method, the conditional coding of foreign media text translation discourse signal is carried out, and the statistical information analysis model of foreign media text translation discourse signal is established. Through text parameter clustering and envelope amplitude-frequency feature detection, the signal detection and feature recognition of foreign media text translation discourse signal are realized.

4 SIMULATION TEST

In order to verify the application performance of this method in foreign media text translation discourse feature recognition, Matlab is used for experimental analysis, and the sample number of foreign media text translation discourse signal is set to 20, the signal frame length is 2000~4000, and the chip length is 4016. See Table 1 for the initial parameter settings of the signal.

Table 1: Initial parameter setting of signal.

Signal sequence	Bandwidth /Kbps	Interference intensity /dB	Sampling frequency /KHz
Sample 1	12.034	3.772	21.000
Sample 2	12.592	3.476	21.516
Sample 3	12.695	3.592	21.370
Sample 4	12.016	3.194	21.886
Sample 5	12.439	3.310	21.262
Sample 6	12.548	3.468	21.982
Sample 7	12.316	3.014	21.372
Sample 8	12.537	3.462	21.695
Sample 9	12.184	3.249	21.661
Sample 10	12.525	3.649	21.775
Sample 11	12.168	3.052	21.585
Sample 12	12.761	3.486	21.181
Sample 13	12.339	3.479	21.558
Sample 14	12.538	3.696	21.358
Sample 15	12.209	3.502	21.312
Sample 16	12.088	2.981	21.577
Sample 17	12.674	3.066	14.176
Sample 18	12.034	3.772	21.000
Sample 19	12.592	3.476	21.516
Sample 20	12.695	3.592	21.370

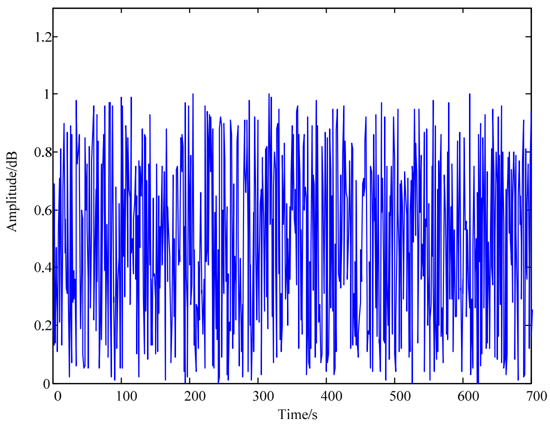

Figure 1: Original signal.

According to the parameter settings in Table 1, foreign media text translation discourse feature recognition processing is performed, and the original signal is shown in Fig 1, and the signal after feature recognition is shown in Fig 2.

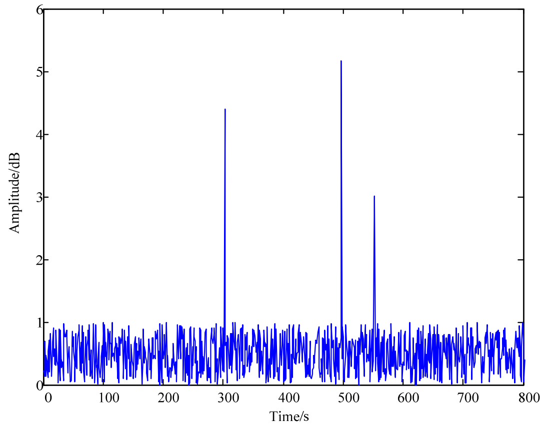

Figure 2: Discourse feature recognition.

By comparing the results of Fig 1 and Fig 2, it can be seen that the peak gain of the foreign media text translation speech signal is obviously expressed by this method, which indicates that the foreign media text translation speech signal has strong anti-interference ability, and the output signal-to-noise ratio (SNR) is tested, and the comparison result is shown in Fig 3. By analyzing the results of Fig 3, it can be seen that the output signal-to-noise ratio (SNR) of the foreign media text translation speech signal is higher than that of the traditional method.

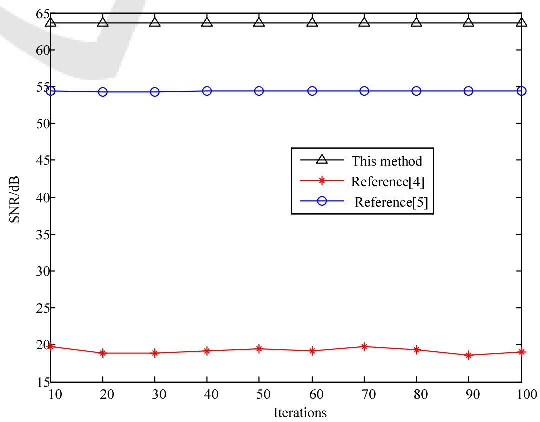

Figure 3: Contrast test of output signal-to-noise ratio.

5 CONCLUSIONS

In this paper, a model of foreign media text translation discourse signal collection and feature recognition is constructed, and the feature recognition and detection of foreign media text translation discourse signal are carried out by combining anti-interference design methods, so as to improve the ability of signal detection and recognition, and thus improve the ability of foreign media text translation discourse feature recognition. In this paper, a method of speech feature recognition and recognition of foreign media text translation based on spectral density feature decomposition is proposed. Based on the intelligent recognition method, the conditional coding of text translation speech signal is carried out, the statistical information analysis model of foreign media text translation speech signal is established, and the speech feature recognition is realized by combining signal filtering and detection recognition. The test shows that the output signal-to-noise ratio (SNR) of the feature recognition processing of multi-frame foreign media texts is high and the feature recognition effect is good.

REFERENCES

ABDUL RAUF S,SCHWENK H. Parallel sentence generation from comparable corpora for improved SMT[J]. Machine Translation, 2011,25(4):341-375.

AI Z L,YANG M M, XIONG D Y. Data augmentation for neural machine translation[J]. Journal of Chinese Information Processing,2018,32(7):30-36.

Haixing Wang. Exploration on the Innovative Path of College Students' Mental Health Education under the Background of Epidemic Situation--Take the Psychological Salon as an Example[J]. Advances in Educational Technology and Psychology, 2021, 5(8):66-70.

John K. Aderibigbe and Sihle Mahola and Tendia Chimucheka. Influence of entrepreneurship risk perceptions and aversion on entrepreneurial intention among university students in South Africa[J]. AFFRIKA Journal of Politics, Economics and Society, 2019, 9(1): 239-252.

JIA C X, LAI H, YU Z T, et al. Pseudo-parallel corpus generation for Chinese-Vietnamese neural machine translation based on pivot language[J]. Computer Engineering and Science,2021, 43(3):543-550.

MARIE B, FUJITA A. Efficient extraction of pseudo-parallel sentences from raw monolingual data using word embeddings[C]//Proceedings of the 2017 55th Annual Meeting of the Association for Computational Linguistics. Stroudsburg:ACL,2017:392-398.

Peng Wang, Dan Li, Xiangyin Meng. The Management Mode of College Students' Mental Health Based on Wireless Sensor Network[J]. International Journal of Frontiers in Sociology, 2022, 40(2):45-49.

Yi Lin. Analysis on the Innovative Strategies of College Students' Mental Health Education under the Background of We-Media[J]. International Journal of Education and Teaching Research, 2021, 2(3):32-38.

ZHANG X L,LI X,YANG Y T,et al. Analysis of bi-directional reranking model for Uyghur-Chinese neural machine translation[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2020,56(1):31-38.

Zheng Wanting. Cluster Analysis Algorithm in the Analysis of College Students' Mental Health Education.[J]. Applied bionics and biomechanics, 2022, 63(6):19-21.