

Retraction

Retracted: Intelligent English Translation Based on Intelligent Speech Waveform Analysis

Mobile Information Systems

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

Copyright © 2023 Mobile Information Systems. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] J. Meng, "Intelligent English Translation Based on Intelligent Speech Waveform Analysis," *Mobile Information Systems*, vol. 2022, Article ID 9408361, 13 pages, 2022.

Research Article

Intelligent English Translation Based on Intelligent Speech Waveform Analysis

Jing Meng 

Jilin Agricultural University, Changchun 130000, China

Correspondence should be addressed to Jing Meng; mengjing@jlau.edu.cn

Received 20 May 2022; Revised 6 June 2022; Accepted 27 July 2022; Published 21 August 2022

Academic Editor: Yajuan Tang

Copyright © 2022 Jing Meng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the effectiveness of an intelligent English translation, this paper combines intelligent speech waveform technology to construct an intelligent English translation system. When the modulation method of each signal component is known, this paper takes the overlapping point as the fitting point set and performs fitting with each data point. Moreover, this paper takes the obtained modulation method parameters as eigenvectors and finally performs clustering to obtain the parameter estimation of each single-component signal. In addition, in the case of overlapping points between each linear English speech waveform dataset, this paper finds the overlapping points and processes the overlapping point sets. The simulation test and data statistical analysis results show that the intelligent English translation method based on intelligent speech waveform analysis proposed in this paper can effectively improve the effect of English translation.

1. Introduction

The main functions of English translation interaction are defined as auxiliary tools, such as English translation input and search or simple function repetition set by the user. At present, the products of English translation mobile terminals in the domestic market are mainly positioned as auxiliary products such as English translation assistants or driving assistants. The intelligent English translation terminal equipment is mainly positioned as a multidevice linkage with practical controllable equipment under the control of English translation. However, their similarities are gradually opening up third-party applications to establish a highly collaborative ecological platform, integrating more Internet application functions online and offline new service experiences.

As the technological progress of AI is gradually applied to English translation technology, deep learning based on the analysis of a large amount of pure data has promoted the rapid development of English translation interaction technology. This progress is mainly reflected in three dimensions: natural cognitive ability, sensory feedback ability, and English translation output ability. These three dimensions constitute the core competency of AI.

The English translation interaction design framework requires the ability of language perception understanding and perception feedback in the process of language communication behavior, and the process of communication interaction needs to have the characteristics of multiple rounds of interaction, and human-computer English translation interaction generally follows this basic feature. According to the particularity of English translation interaction as a medium, this paper analyzes the products of English translation interaction in the market at this stage and summarizes the basic design principles [1].

“Lightweight” operation reduces the user’s burden to improve user efficiency. Therefore, reducing the exclusive use of the interface and not increasing the user’s unnecessary burden is one of the design principles that need to be considered [2]. In the design process of English translation interaction, the complexity of English translation interaction behavior in different scenarios should be considered. In a specific functional scenario, such as when the user confirms the operation or the user is not advised to make other decisions at this time, the interface can be exclusively used, and the visual focus is not easily distracted and does not affect the user’s judgment and decision-making. In order to

improve the use efficiency, the proportion of the interface can be increased to leave a visual space for the user to click to select and confirm. The learnability of deeply customized smart products that users are used to is also a basic requirement of people for smart products, and users have high expectations for smart products, and products that are not intelligent enough will easily affect the user experience of products [3]. The technical learnability of the product is mainly reflected in the adaptability to the voiceprint, which improves the accuracy of information recognition. In terms of design, it is necessary to set up a database about the habits of users in different scenarios, that is, multiple or repeated operations, such as checking the weather, setting an alarm clock, checking in, booking a flight ticket, and dialing a contact number. When the user triggers it again, it should jump directly without repeating the operation, shortening the user's operation path. Not only the design of English translation interaction but the English translation assistant also needs to obtain reasonable judgment through the data generated by various modules of the intelligent terminal and improve the human-machine intelligence and ease of use of the product through the data [4]. Intelligent English translation products have the ability to be emotional. Intelligent English translation products have the same emotion as human beings and can even perceive each other and generate "sympathy." These have always been the good wishes of human beings. However, the technical bottleneck cannot be achieved in the short term. It is truly anthropomorphic and emotional, but it can increase the degree of emotionality of the product through various ways of expression, the change of intonation, the corpus with anthropomorphism, etc. [5].

With the popularization of the concept of artificial intelligence, some products dedicated to intelligent education have also appeared on the market. Intelligent learning products can be roughly divided into intelligent learning software and intelligent learning terminals [6]. Today, when smartphones are changing people's lives, intelligent learning software is popular with parents and students because of its low price and convenient use. At present, various intelligent learning software covers functions such as online classrooms, topic photo search, and online problem solving [7]. Online classrooms record the teaching process of outstanding teachers in key schools into videos for students to watch and learn. Although this allows students to preview before class or review after class, it takes up students' precious after-class time. In fact, students are in the opportunity to communicate and learn synchronously with teachers in the classroom which is more valuable, and watching other teachers' videos after class is more effective [8]. Question photo search is that students can turn to the Internet when they encounter a question they cannot understand and get the correct answer quickly. However, due to various reasons such as the OCR technology of each company, the quality of the question bank, and the user's photo scene, many times it is not possible to search for the correct topic [9]. The online question-making function has a built-in massive question bank for students to practice, but the fly in the ointment is that most of the products provide students with questions

that are not targeted, and students cannot avoid the pain of the sea of questions. The intelligent learning terminal also occupies part of the market because of its professionalism and the characteristics of integrating multiple functions, such as learning machines, reading machines, and learning robots produced by various manufacturers. Learning machines mainly include functions such as online classrooms, online question-making, electronic dictionaries, and gamified learning [10]; point-reading machines are specially designed for foreign language subjects and teach students oral language by clicking to read aloud; learning robots are aimed at young students and stimulate children's interest in learning by reading poems, singing nursery rhymes, and reciting math problems. The functions of intelligent learning terminals and intelligent learning software are similar, and they also have the disadvantage that it is difficult to improve the efficiency of classroom and after-school learning [11].

Intelligent virtual assistants can be categorized from different perspectives, and some common perspectives include generating models from problem-solving knowledge domains and techniques. From the analysis of the knowledge domain of solving the problem, it can be divided into a closed domain and an open domain. A closed domain is based on a specific topic, focusing on answering questions in a specific domain, and professional knowledge such as a domain knowledge base can be imported to effectively improve the performance of the system [12]. On the other hand, based on the open domain, there are no limited topics, and users can talk freely, so the difficulty is relatively increased, and the knowledge base and models to be prepared are much more complicated. From the perspective of the technical generation model of the intelligent virtual assistant, it can be divided into retrieval type and generative type. Retrieval-based intelligent virtual assistants use the defined question base and answer knowledge base and return the user answers by matching user questions with the knowledge base [13]. This method requires the knowledge base to be as large as possible, and its advantage is that the answer quality is high. Generative intelligent virtual assistants do not rely on a specific answer library but use certain technical means to automatically generate answers. The advantage of this type is that user questions can cover arbitrary topics, and the disadvantage is that the quality of the generated responses may be affected [14].

At the system support level, intelligent virtual assistants generally exist in learners' personal devices and are often associated with accounts, in real-time communication, and learners' conversational context information, such as learners' geographic locations, learning trajectories, duration, schedule, personalized preferences, styles, and more. This contextual and personalized information provides the data foundation for learning analytics. At the same time, the intelligent assistant provides the support of the domain knowledge base at the system level, and a large-scale, high-quality domain knowledge base is the key to realize the effective learning support function [15].

Through the data measurement and collection of learners and their learning situations, the obtained personalized data of learners can be used for effective learning

analysis through data mining, atmosphere computing, affective computing, and other methods to understand and optimize learning and learning situations and facilitate the construction of more complete learner models. This will provide support and feedback for learning applications so that the learning system can truly focus on the individual development and needs of each learner [16].

On the basis of system support and learning analysis, intelligent assistants can provide learners with a variety of learning support applications, such as personalized learning guidance, interactive Q&A and emotional communication, learning planning and scheduling, and situational learning. As a personal teaching assistant in the learning process, the intelligent assistant can assist in the management of the learning process. When the learner asks questions about the learning content, it can analyze the learning situation of the learner, provide one-to-one individual guidance, and achieve personalized and emotional communication in the learning process [17].

This paper combines the intelligent speech waveform technology to construct an English intelligent translation system to improve the effect of intelligent English translation and improve the interactivity of English translation.

2. Intelligent Speech Waveform Signal Analysis

2.1. English Speech Waveform Clustering and Its Application in Multicomponent Signal Parameter Estimation. Nonlinear English speech waveform clustering is a kind of data clustering aiming at the distribution of sample points in a nonlinear English speech waveform distribution. For multi-English speech waveform data with a certain mathematical model, a simple and effective method is to map the data to a high-dimensional space, and the data can be linearly separable in the high-dimensional space. This requires mapping the original one-dimensional space x to the three-dimensional space (x, x^2, x^3) and then looking for a model between the characteristics and the results in the high-dimensional space. This feature transformation is called feature mapping. The mapping function is called ϕ , and in this example,

$$\phi(x) = \begin{bmatrix} x \\ x^2 \\ x^3 \end{bmatrix}. \quad (1)$$

In a model of a linear English speech waveform clustering algorithm, when the formula can be represented by the inner product between data points, the kernel trick can be used to avoid explicitly mapping the data into a high-dimensional space. For the sample point x and the sample point y , the kernel function $k(x, y)$ represents a function of the inner product of x and y . For N sample points, x_1, x_2, \dots, x_n , the kernel matrix is shown in the formula:

$$K := \{k(x_i, x_j)\}_{1 \leq i, j \leq N}. \quad (2)$$

The kernel matrix needs to be a positive semidefinite symmetric matrix, and any kernel function can be expressed in the form of an inner product, as shown in the formula:

$$k(x, y) = \langle \phi(x), \phi(y) \rangle. \quad (3)$$

Common kernel functions include polynomial, Gaussian, and linear functions, each of which has advantages and disadvantages. The linear kernel does not perform feature transformation and can be used directly, that is, without kernel skills. The advantages are high computational efficiency and good interpretation of the results, but the disadvantage is that the data need to be linearly separable. The general form of the polynomial kernel is shown in the following formula:

$$k(x, y) = (a + bx \cdot y)^c. \quad (4)$$

The advantage is that compared with the linear kernel, the data requirements are relatively loose, and the disadvantage is that there are many coefficients to be selected, which makes the model more complicated. The Gaussian kernel is also known as the radial basis kernel, and the general form is shown in the following formula:

$$k(x, y) = \exp(-a\|x - y\|). \quad (5)$$

Among them, a is a constant, generally taken as the reciprocal of the total number of data categories. Its advantage is that it has fewer coefficients to choose from. It is more widely used than linear kernel and polynomial kernel, and it can process almost all data. Moreover, the kernel function can map the sample features to an infinite-dimensional space and is not prone to computational accuracy problems. The disadvantage is that the interpretability of infinite dimension is poor, and it is too powerful, easy to overfit, sensitive to the choice of coefficients, and computationally expensive. The biggest advantage of the Gaussian kernel is that it maps the original data into an infinite dimensional space. When a is 1,

$$\begin{aligned} k(x, y) \exp(-\|x - y\|^2) &= \exp(-x^2) \exp(-y^2) \exp(2xy), \\ &= \exp(-x^2) \exp(-y^2) \left(\sum_{i=0}^{\infty} \frac{[2xy]^i}{i!} \right), \\ &= \sum_{i=0}^{\infty} \left(\exp(-x^2) \exp(-y^2) \sqrt{\frac{2^i}{i!}} \sqrt{\frac{2^i}{i!}} x^i y^i \right), \\ &= \sum_{i=0}^{\infty} \left(\left(\exp(-x^2) \sqrt{\frac{2^i}{i!}} x^i \right) \left(\exp(-y^2) \sqrt{\frac{2^i}{i!}} y^i \right) \right) \\ &= \phi(x) \cdot \phi(y). \end{aligned} \quad (6)$$

Among them, the features are transformed into

$$\phi(x) = \exp(-x^2) \left(1, \sqrt{\frac{2^1}{1!}} x^1, \sqrt{\frac{2^2}{1!}} x^2, \dots, \sqrt{\frac{2^i}{1!}} x^i, \dots \right). \quad (7)$$

In the transformed high-dimensional space, the dimension is infinite, so almost all data can be classified by using the Gaussian kernel function.

KSCC is the kernel function method of SCC, but the calculation is not calculated in a high-dimensional space but

is calculated in a low-dimensional space, which is also an advantage of the kernel technique. It is assumed that the dataset $X = \{x_1, \dots, x_n\}$ has a total of N sample points and K types of data. It is assumed that the distribution of each type of data conforms to a certain mathematical model, the original data are in the d -dimensional space, and the mapped data are in the l -dimensional space.

For a set of $l+2$ data points $I = \{i_1, \dots, i_{l+2}\}$ in any mapping dataset, the corresponding kernel matrix is denoted by $K_{I,l}$:

$$K_{I,l} := (K_{ij})_{i,j \in I} \quad (8)$$

The formula for calculating the polar curvature of the eigenvector is

$$c_p^2(I) = \frac{1}{l+2} \cdot \max_{i,j \in I} (K_{ii} + K_{ij} - 2K_{ij}), \quad (9)$$

$$\sum_{i \in I} \frac{\det(K_{I,l} + 1)}{\prod_{j \in I, j \neq i} K_{ii} + K_{ij} - 2K_{ij}}.$$

In the abovementioned formula, if the denominator is 0, the polar curvature is directly assigned to 0, and the KSCC algorithm calculates attractiveness for any $l+2$ data point set:

$$A_p(I) := e^{-c_p^2(I)(2\sigma^2)}. \quad (10)$$

Among them, σ is a tuning parameter. Its attractiveness is closer to 1 for the set of data points from the same latent subspace. For the set of data points from different latent subspaces, its attractiveness is closer to 0, and its computation is still performed in a low-dimensional space due to the use of the kernel trick. Then, the multiway similarity matrix W is calculated:

$$W_{ij} = \sum_J A_p(i, J) \cdot A_p(j, J). \quad (11)$$

Among them, $J = (i_1, \dots, i_{l+1}) | 1 \leq i_1, \dots, i_{l+1} \leq N$. Finally, the dataset is clustered by spectral clustering. The KSCC algorithm is used to estimate the parameters of the actual multicomponent signal. The parameters of the multicomponent signal are shown in Table 1, and the expression is as follows:

$$s(t) = \exp[j2\pi(0.2t)] + \exp\left[j2\pi\left(0.175 \sin\left(\frac{\pi}{32}t\right) + 0.225\right)\right]. \quad (12)$$

Its time-frequency distribution after rearrangement spectrum transformation is shown in Figure 1.

First, select a polynomial kernel function, the kernel function is $(1 + x' \cdot y)^2$, and the high-dimensional space dimension l is 4. The clustering result is shown in Figure 2. It can be seen from the figure that most of the sample points can achieve component signal clustering. However, the clustering of sample points where the two signal components overlap will be less accurate.

If the Gaussian kernel function is selected, the kernel function is $\exp(-a\|x - y\|^2)$, a , the value is 1, and the dimension l of the high-dimensional space is 5, and the clustering result is shown in Figure 3. As can be seen from the figure, the two signal components can be separated. Then, according to the modulation mode of the signal, the parameters of each signal component can be fitted according to the corresponding time-frequency sample points of the single-component signal, and the parameters of each single-signal component can be obtained.

The original dataset is $X = \{x_1, x_2, \dots, x_n\}$, $x_i \in R^d$ is the dimension of the original dataset, and the dataset is mapped to $Y = \{y_1, y_2, \dots, y_n\}$, $y_i \in R^m$ by the LLE method, where m is the dimension of the dataset after dimensionality reduction, $m < d$. The main steps of the method are as follows:

- (1) The algorithm calculates the Euclidean distance between sample points in a high-dimensional space, and for each sample point, it selects the K nearest sample points as the nearest neighbors.
- (2) For each sample point x_i , the algorithm calculates the weight W_{ij} between the point and each of its neighbors and constructs the loss function as follows:

$$\varepsilon(W) = \sum_{i=1}^N \left| x_i - \sum_{j=1}^K W_{ij} x_j \right|^2. \quad (13)$$

Among them, $\sum_{j=1}^K W_{ij} = 0$. The algorithm obtains W_{ij} by minimizing the loss function.

- (3) The algorithm calculates y_i and y_j in the low-dimensional embedding space according to the W_{ij} obtained in the second step and constructs the loss function as follows:

$$\Phi(Y) = \sum_{i=1}^N \left| y_i - \sum_{j=1}^K W_{ij} y_j \right|^2. \quad (14)$$

The algorithm obtains the dataset Y in the low-dimensional space by minimizing the loss function.

For the multicomponent signal composed of two non-linear FM signals and a constant frequency signal in Table 2, its expression is

$$s(t) = \exp[j2\pi(0.45t)] + \exp\left[j2\pi\left(0.1 \sin\left(\frac{\pi}{64}t\right) + 0.3\right)\right] + \exp\left[j2\pi\left(0.05 \sin\left(\frac{\pi}{16}t\right) + 0.1\right)\right]. \quad (15)$$

The time-frequency distribution spectrogram of the rearranged spectrum is obtained as shown in Figure 4. When the sampling frequency is too large, there are too many sample points in the time-frequency distribution, which will slow down the calculation of the clustering algorithm. Therefore, for each time point, only the n points with the highest energy peak are extracted as sample points, where n

TABLE 1: Multicomponent signal (12) parameters.

	Modulation method	Minimum frequency	Maximum frequency
Signal component 1	Sine	0.050	0.400
Signal component 2	Constant frequency	0.200	0.200

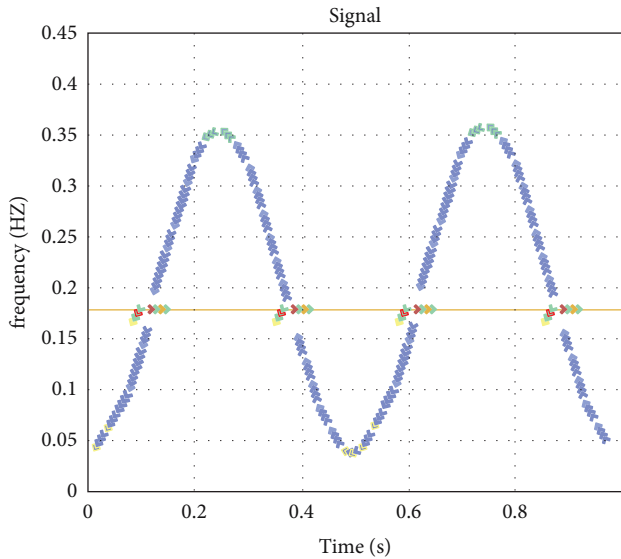


FIGURE 1: Rearranged spectral distribution of a multicomponent signal (12).

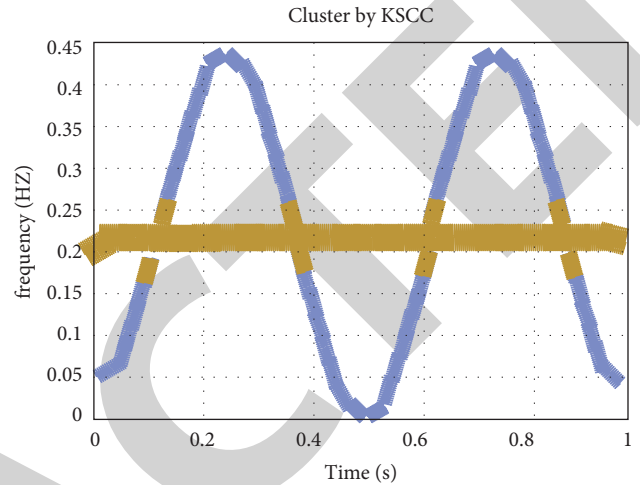


FIGURE 3: Clustering results of the Gaussian kernel function.

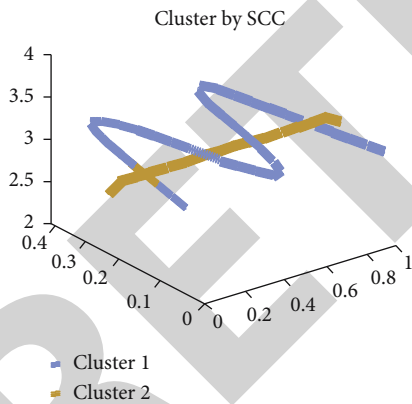


FIGURE 2: Clustering results of a polynomial kernel function.

is the number of signal components, then the extracted main sample points are clustered, and then the signals are sorted.

The obtained samples are dimensionally reduced by the LLE algorithm, and then the clustering results are obtained by the K -means clustering algorithm. As can be seen from Figure 5, in the case that the multicomponent signal includes a pulse signal with a fixed frequency and two sinusoidal FM signals, the time-frequency distribution sample points of each component signal can be separated. After that, parameter estimation is performed on the time-frequency distribution of each component signal, and the accuracy of parameter estimation is high.

2.2. Clustering of Nonlinear English Speech Waveforms Based on Overlap Points. The two signal component parameters are shown in Table 3, and the expressions is

$$s(t) = \exp[j2\pi(0.2t)] + \exp[j2\pi(0.175 \sin(4\pi t) + 0.225)]. \quad (16)$$

The time-frequency distribution of the rearranged spectrum is obtained as shown in Figure 6. It can be seen from the figure that there are some overlapping points in the time-frequency distribution of the two signals. That is, if the frequencies of the two signals at the same time are equal, the overlapping points between the time-frequency linear English speech waveform distributions of each single-component signal can be obtained through an algorithm. Then, curve fitting is performed on each data point and the overlapping point set to obtain the corresponding curve parameters, and the curve parameters of the data points are used as feature vectors, and then the sorting results are obtained by clustering.

As shown in Figure 6, there are 4 overlapping points $\{p1, p2, p3, \text{ and } p4\}$ for the two signal components. Since there are few modulation modes of the signal, it is assumed that we know the modulation modes of the two signal components through their time-frequency distribution. As shown in the figure, the modulation modes of the two signal components are constant frequency modulation and sinusoidal frequency modulation, respectively. In MATLAB, use the `lsqcurvefit` function to perform polynomial fitting on the data points x_1 and $\{p1, p2, p3, \text{ and } p4\}$ with $y = a \cdot x + b$ as the parameter. The principle of the `lsqcurvefit` function is that the sum of squares of the error is the smallest, and the obtained curve parameters are about $[0 \text{ and } 0.2]$. If the sum of squared

TABLE 2: Multicomponent signal (15) parameters.

	Modulation method	Minimum frequency	Maximum frequency
Signal component 1	Sine	0.200	0.400
Signal component 2	Sine	0.050	0.150
Signal component 3	Constant frequency	0.450	0.450

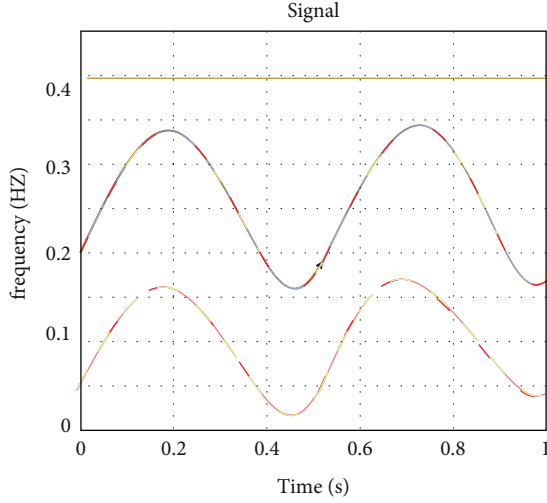


FIGURE 4: Rearranged spectral time-frequency distribution of the multicomponent signal (15).

errors between the real data point and the fitted data point is less than a certain threshold ε , its parameters are used as the feature vector of the data point. If the sum of the squared errors between the real data points and the fitted data points is greater than a certain threshold ε , the modulation formula $y = a \sin(bx + c) + d$ is used as a parameter, and the `lsqcurvefit` function is used to perform curve fitting on the data points x_2 and $\{p1, p2, p3, p4\}$. The corresponding sinusoidal parameters are obtained to be about $[0.175, 12, -1.3, 0.225]$. After clustering each sample point, it is added to the corresponding fitting point set to improve the accuracy of the fitting parameters.

Then, the curve parameters of each data point are used as eigenvectors, and the *K*-means clustering method is used to cluster the data points, and the clustering result is shown in Figure 7, where, $y_1 = a \cdot x_1 + b$, $y_2 = a \sin(bx_2 + c) + d$, and $\varepsilon = 0.00005$.

First, as shown in Figure 8 and Table 4, it finds the overlapping point set between the time-frequency distribution sample points according to the method of density peaks, and in the case of many overlapping points, curve fitting is performed on each sample point and the overlapping point set together. Since it already has certain prior knowledge and knows the modulation mode of the component signal, the frequency modulation parameter formula

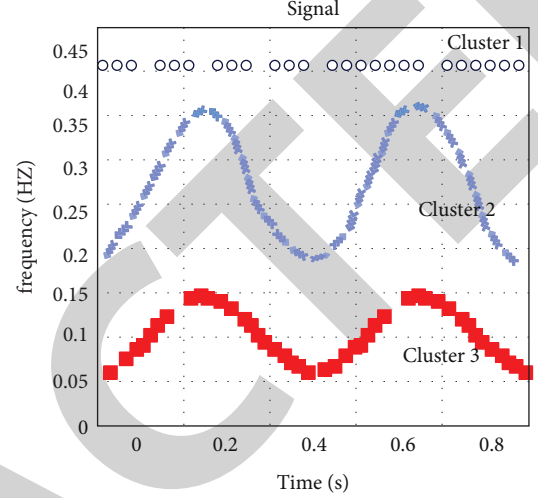


FIGURE 5: Clustering results of the LLE algorithm.

$y = a \sin(bx + c) + d$ can be directly fitted to the curve. At this time, it is only necessary to obtain the modulation formula parameters, and the curve parameters of each data point are obtained by the above method as the feature vector, and the clustering result is shown in Figure 9. It can be seen from the figure that the clustering effect is very ideal, the two sinusoidal frequency-modulated component signals can be separated, and the estimated accuracy of the component signal parameters is also ideal.

2.3. Improved Spectral Curvature Clustering Algorithm for Initial Sampling. Spectral curvature clustering linear English speech waveform clustering algorithm is based on multi-channel clustering technology. Usually, at least $d + 1$ data points are required to define a d -dimensional affine subspace. Spectral curvature clustering uses a multichannel similarity measure to describe the possibility that $d + 2$ data points come from the same potential linear English speech waveform, so as to construct a similarity measure matrix W between data points.

D represents the original dimension of the dataset, and for any $d + 2$ data points z_1, \dots, z_{d+2} , where $0 \leq d < D$, $V_{d+1}(z_1, \dots, z_{d+2})$ represents the volume of the $d + 1$ simplex formed by these points, and the polar sine of each data point is

$$p \sin_{z_i}(z_1, \dots, z_{d+2}) = \frac{(d+1)! \cdot V_{d+1}(z_1, \dots, z_{d+2})}{\prod_{1 \leq j \leq d+2, j \neq i} \|z_j - z_i\|}, \quad 1 \leq i \leq d+2. \quad (17)$$

TABLE 3: Multicomponent signal (16) parameters.

	Modulation method	Minimum frequency	Maximum frequency
Signal component 1	Sine	0.050	0.400
Signal component 2	Constant frequency	0.200	0.200

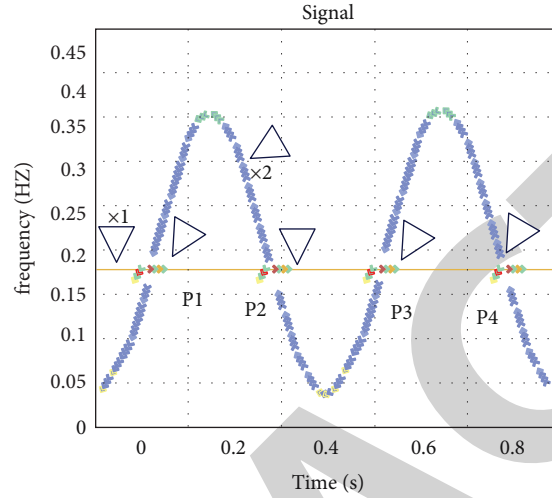


FIGURE 6: Schematic diagram of the overlapping point of the multicomponent signal (16).

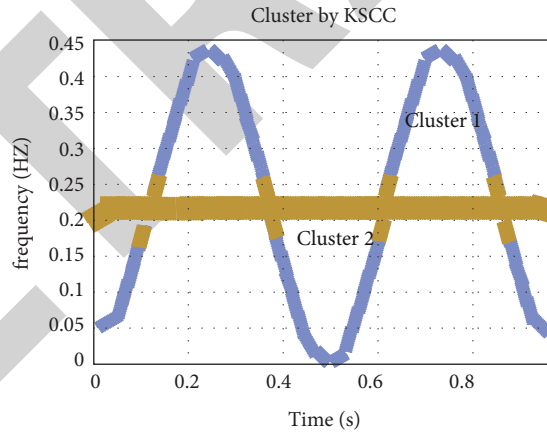


FIGURE 7: Clustering results of multicomponent signals (16).

The polar curvature of $d + 2$ data points is?

$$c_p(z_1, \dots, z_{d+2}) = \text{diam}(\{z_1, \dots, z_{d+2}\}) \cdot \sqrt{\sum_{i=1}^{d+2} (p \sin_{z_i}(z_1, \dots, z_{d+2}))^2}. \quad (18)$$

Among them, $\text{diam}(S)$ is the diameter of the set S . When d is 0, the definition of polar curvature is equivalent to Euclidean distance. $\Omega_{d+2} = \{x_{il}\}_{l=1}^{d+2}$ is $d + 2$ randomly

selected data points, and spectral curvature clustering is based on the probability of polar curvature to define a multiway similarity $S_{i_1, i_2, \dots, i_{d+2}}$ as follows:

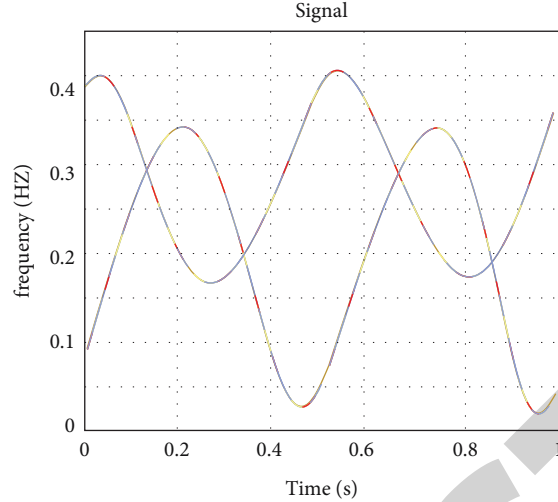


FIGURE 8: Rearranged spectral distribution of two sinusoidally modulated signals.

TABLE 4: Single-component signal parameters of two sinusoidal modulation methods.

	Modulation method	Minimum frequency	Maximum frequency
Signal component 1	Sine	0.050	0.350
Signal component 2	Sine	0.200	0.450

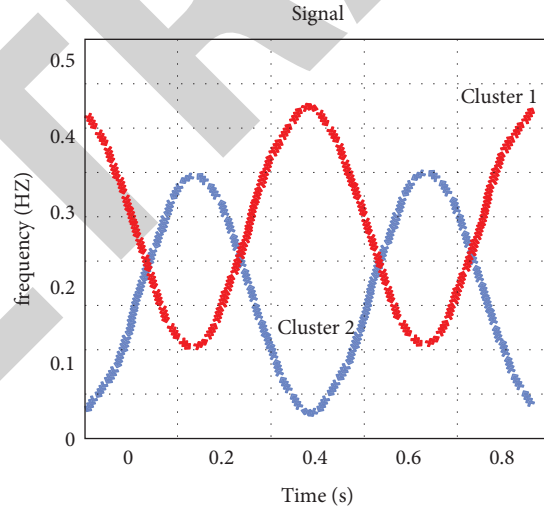


FIGURE 9: Clustering results of multicomponent signals for two sinusoidal modulation methods.

$$S_{i_1, i_2, \dots, i_{d+2}} = \exp\left(-\frac{1}{2\sigma^2} \text{diam}^2(\Omega_{d+2}) \sum_{l=1}^{d+2} \frac{(d+1)!^2 \text{vol}^2(\Omega_{d+2})}{\prod_{1 \leq m \leq d+2, m \neq l} \|x_{im} - x_{il}\|^2}\right). \quad (19)$$

The above formula holds if and only if i_1, i_2, \dots, i_{d+2} comes from different linear English speech waveforms, otherwise $S_{i_1, i_2, \dots, i_{d+2}} = 0$. Among them, $\text{diam}(\Omega)$ represents the radius of the set Ω , $\text{vol}(\Omega)$ represents the volume of the

$d+1$ simplex formed by the midpoint of the set Ω , and σ is a constant. The similarity matrix W is defined as

$$w_{ij} = \sum_{k_2, \dots, k_{d+2}} S_{i, k_2, \dots, k_{d+2}} S_{j, k_2, \dots, k_{d+2}}. \quad (20)$$

For the K -type data C_1, \dots, C_K obtained by clustering, a measure for evaluating the spectral curvature clustering

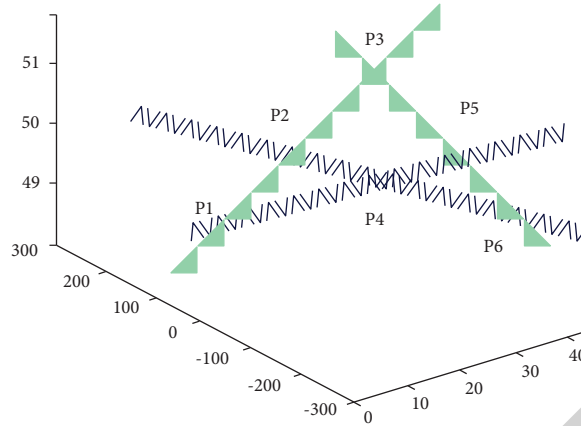


FIGURE 10: Linear English speech waveform distribution data.

Input: a sample set with N types of data.

Output: a set of initial sample points.

Step 1:

For the sample set with N types of data, the overlapping point set $P = \{p_1, p_2, \dots, p_k\}$ is obtained through the algorithm.

(2) It is assumed that the number of overlapping points is K , a vector can be formed between any two points, and K overlapping points can form $M = K \cdot (K - 1)/2$ vector groups $V = \{V_1, V_2, \dots, V_m\}$.

(3) For M vectors, the absolute value of the cosine value of the angle between the vectors is calculated as the similarity between the vectors. The formula is as follows:

$$s_{ij} = \text{abs}(V_i \cdot V_j / \|V_i\| \cdot \|V_j\|).$$

(4) The similarity between vectors in the same linear English speech waveform distribution will be close to 1, and the similarity between vectors in different linear English speech waveform distributions will be close to 0. Therefore, N groups of vectors with higher similarity can be obtained according to the similarity between the vectors, and at the same time, N groups of data point sets on the same linear English speech waveform distribution can be obtained.

ALGORITHM 1: Determines the initial sampling point.

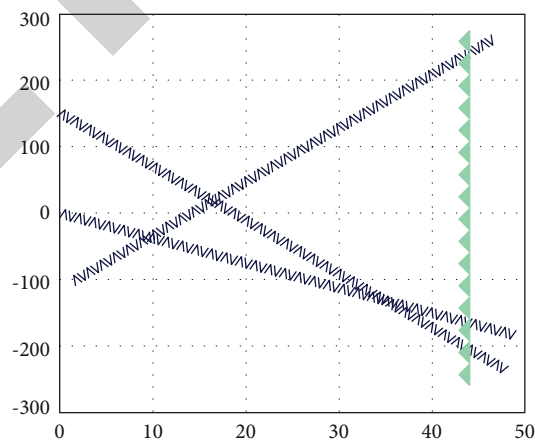


FIGURE 11: Original dataset.

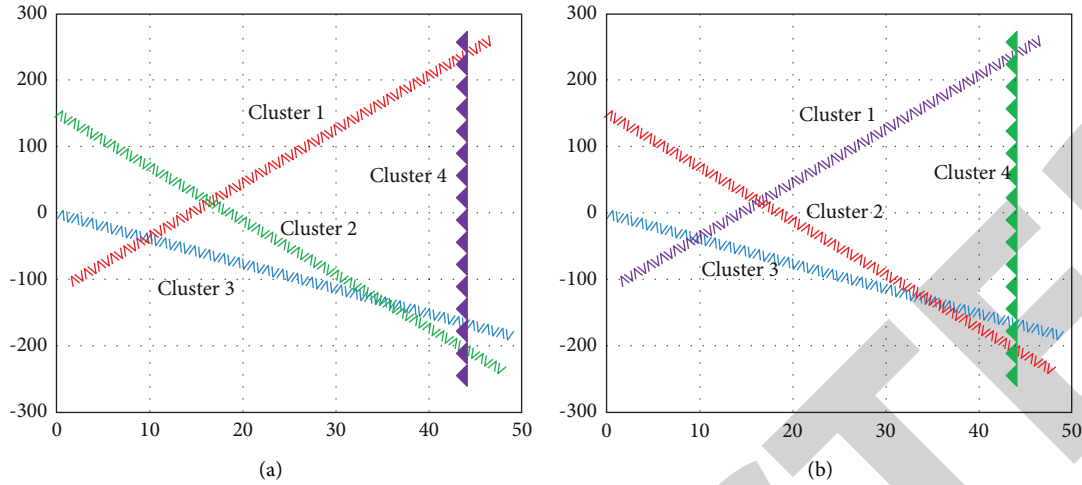


FIGURE 12: Clustering results of the SCC algorithm and the SCC algorithm with improved initial sampling. (a) Clustering results of the SCC algorithm and (b) clustering results of the SCC algorithm with improved initial sampling.

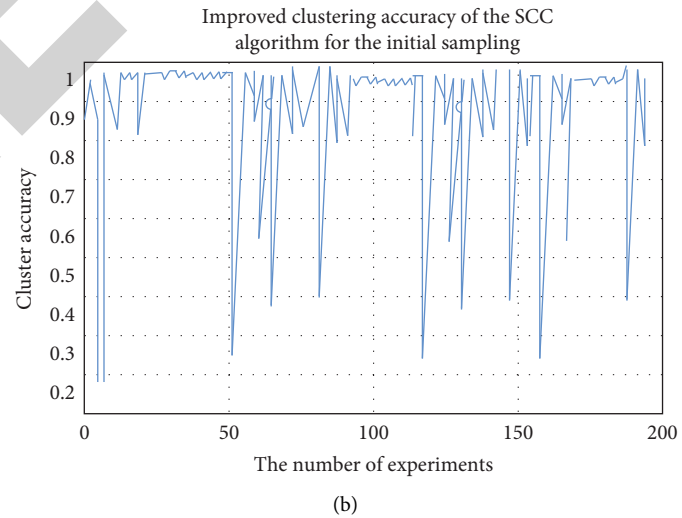
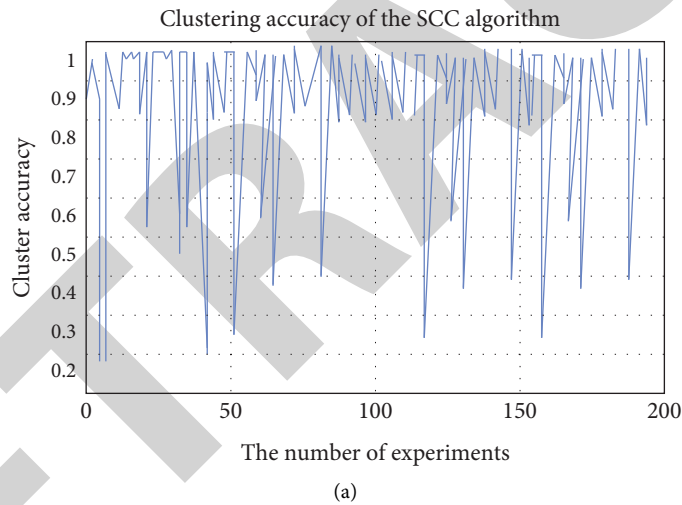


FIGURE 13: Comparison of clustering accuracy between the SCC algorithm and the SCC with improved initial sampling.

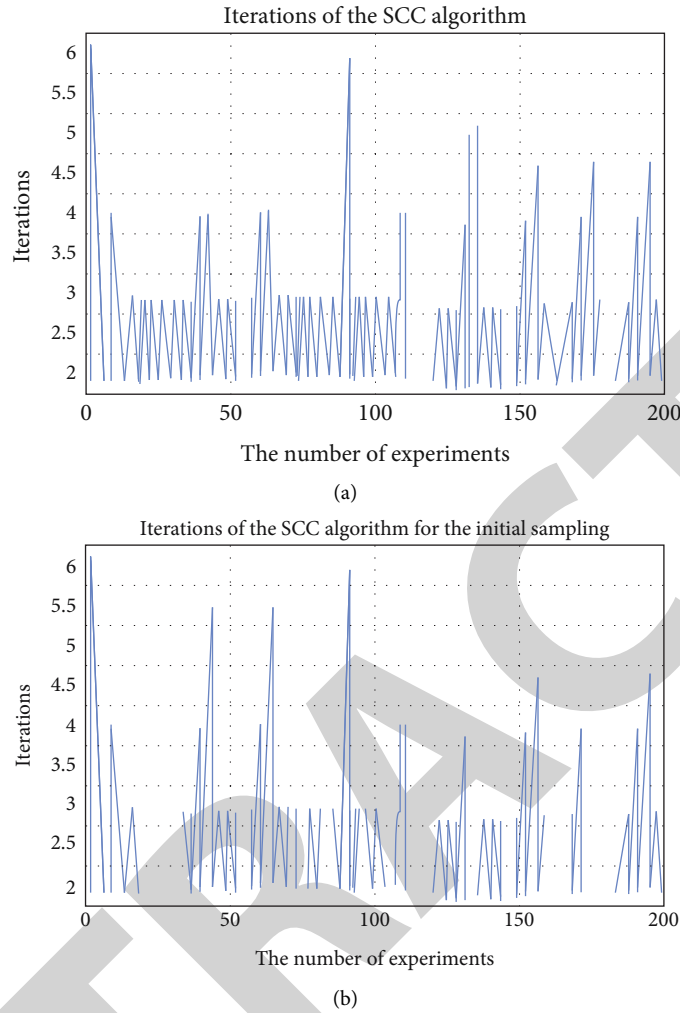


FIGURE 14: Comparison of iteration times between the SCC algorithm and the SCC with improved initial sampling.

TABLE 5: Comparison of the average clustering accuracy and the average number of iterations between the SCC algorithm and the SCC with improved initial sampling.

	SCC	SCC with improved initial sampling
Average clustering accuracy	87.61%	93.21%
Average number of iterations	2.4661	2.392

algorithm is the average orthogonal least squares error (OLS), and the calculation formula is as follows:

$$e_{OLS} = \sqrt{\frac{1}{N} \sum_{k=1}^K \sum_{x \in C_k} \text{dist}^2(x, F_k)}. \quad (21)$$

Among them, F_k represents the d-dimensional approximation value of the OLS of the dataset C_k , and $\text{dist}(x, F_k)$ represents the orthogonal distance from x to F_k .

The algorithm for determining the initial sampling point is described in detail as follows.

As shown in Figure 10, four groups of data point sets $\{p_1, p_2, p_3\}$, $\{p_1, p_4, p_5\}$, $\{p_2, p_4, p_6\}$, and $\{p_3, p_5, p_6\}$ can be obtained, and these four groups of data point sets can be determined to be on the same linear English speech waveform distribution.

Therefore, when using the spectral curvature clustering algorithm to cluster multiple sets of sample point sets with overlaps, Algorithm 1 can first be used to obtain N sets of data point that are known to be on the same linear English speech waveform distribution. Then, the algorithm selects the initial sampling points from these sets of data points, which can avoid falling into the local optimal solution,

TABLE 6: English intelligent translation effect based on intelligent speech waveform analysis.

Number	Translation effect	Number	Translation effect	Number	Translation effect
1	91.95	10	88.14	19	88.67
2	87.21	11	88.43	20	88.41
3	89.54	12	90.89	21	87.46
4	87.17	13	87.89	22	91.65
5	92.80	14	88.79	23	88.56
6	92.58	15	89.13	24	88.91
7	89.16	16	91.91	25	91.94
8	90.17	17	88.94	26	90.22
9	90.60	18	88.07	27	92.25

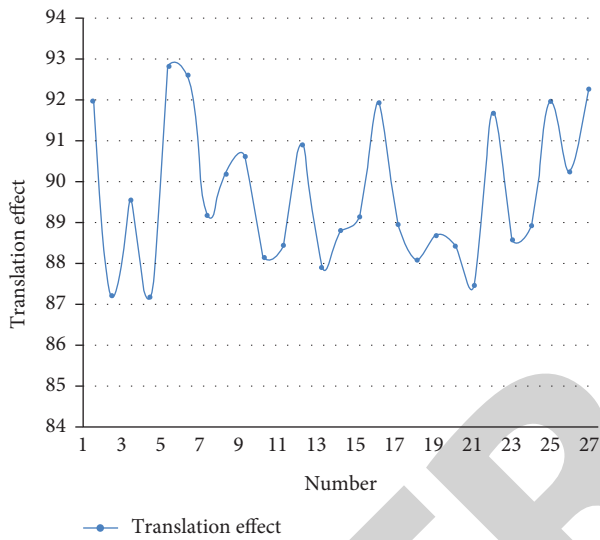


FIGURE 15: Statistical diagram of translation effects.

reduce the instability of the algorithm, and reduce the number of iterations of the algorithm.

3. Experimental Results and Analysis

Four types of data conforming to the distribution of linear English speech waveforms are generated through simulation; each type of data has 150 sample points and a total of 600 sample points. Moreover, the 4 sets of data overlap each other, and there are 6 overlapping points. The original dataset is shown in Figure 11.

Figure 12 shows the clustering effect of SCC and the SCC algorithm with improved initial sampling. It can be seen from Figure 12 that the clustering effects of the two algorithms are very satisfactory, and in most cases, good clustering results can be obtained.

Then, the SCC algorithm and the SCC with improved initial sampling are tested 200 times, and the 200-time experimental accuracy and the number of experimental iterations of the two algorithms are shown in Figures 13 and 14. As can be seen from Figure 13, the SCC algorithm itself is very unstable. Although 113 of the 200 experiments have a clustering accuracy higher than 95%, there are many times

where the clustering accuracy is less than 60%. Compared with the SCC algorithm, the SCC with improved initial sampling is much more stable, and the clustering accuracy of 177 out of 200 experiments is higher than 95%. As can be seen from Figure 14, the average number of iterations of the 200 experiments of the SCC algorithm is 2.47 times, and the average number of iterations of the SCC algorithm with improved initial sampling is 2.39 times, and the number of iterations is also slightly reduced.

The average clustering accuracy and the average number of iterations of the SCC algorithm and the SCC with improved initial sampling for 200 experiments are shown in the table. It can be seen from Table 5 that the average accuracy and the average number of iterations of the SCC algorithm with improved initial sampling are higher than those of the SCC algorithm.

Based on the abovementioned research, the analysis effect of the intelligent speech waveform is verified. Next, the application effect of the method proposed in this paper in English translation is evaluated, and the test results are obtained, which are shown in Table 6 and Figure 15.

From the above analysis, we can see that the intelligent English translation method based on intelligent speech waveform analysis proposed in this paper can effectively improve the effect of English translation.

4. Conclusion

English translation has become a new innovation in human-computer interaction. In the development of human-computer interaction, "technological progress" and "carrier innovation" alternately promote the continuous improvement of the efficiency of people's access to information and the reduction of the cost of use. Moreover, it will also speed up the commercialization of technology applications, and technological innovation will not only bring users a new way of life and a friendly human-computer interaction mode but also subvert the traditional interface-based operation and interaction method. This paper combines the intelligent speech waveform technology to construct an English intelligent translation system to improve the effect of English intelligent translation. The simulation test and data statistical analysis results show that the intelligent English translation method based on intelligent speech waveform analysis proposed in this paper can effectively improve the effect of English translation.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was sponsored by Jilin Agricultural University.

References

- [1] E. Brynjolfsson, X. Hui, and M. Liu, "Does machine translation affect international trade? Evidence from a large digital platform," *Management Science*, vol. 65, no. 12, pp. 5449–5460, 2019.
- [2] R. Schulte, "Expanding the horizon of translation studies," *Translation Review*, vol. 106, no. 1, pp. 1–3, 2020.
- [3] K. Gurumurthy, D. Anand, and K. Karthik, "Democratizing translation using digital methods," *Language and Language Teaching*, vol. 10, no. 1, pp. 30–36, 2021.
- [4] K. Karlgren, M. Lakkala, A. Toom, L. Ilomäki, P. Lahti-Nuutila, and H. Muukkonen, "Assessing the learning of knowledge work competence in higher education—cross-cultural translation and adaptation of the Collaborative Knowledge Practices Questionnaire," *Research Papers in Education*, vol. 35, no. 1, pp. 8–22, 2020.
- [5] C. Dinneen, "Students' use of digital translation and paraphrasing tools in written assignments on Direct Entry English Programs," *English Australia Journal*, vol. 37, no. 1, pp. 40–51, 2021.
- [6] Y. Sun and D. Li, "Digital humanities approaches to literary translation," *Comparative Literature Studies*, vol. 57, no. 4, pp. 640–654, 2020.
- [7] M. M. J. F. Farooque and B. Jadad, "A comparative analysis of translation software using data science approach to Arabic statements," *International Journal of Advanced Computer Systems and Software Engineering*, vol. 2, no. 3, pp. 13–19, 2021.
- [8] J. C. S. Hvidt, L. F. Christensen, C. Sibbersen, S. Helweg-Jørgensen, J. P. Hansen, and M. B. Lichtenstein, "Translation and validation of the system usability scale in a Danish mental health setting using digital technologies in treatment interventions," *International Journal of Human-Computer Interaction*, vol. 36, no. 8, pp. 709–716, 2020.
- [9] B. Türkmen, "Utilising digital media as a second language (L2) support: a case study on Netflix with translation applications," *Interdisciplinary Description of Complex Systems*, vol. 18, no. 4, pp. 459–470, 2020.
- [10] H. Minors, "Desblache, lucile," *Music and Translation: New Mediations in the Digital Age*, pp. 230–232, Macmillan, London, 2021.
- [11] Y. A. Rankin, J. O. Thomas, and N. M. Joseph, "Intersectionality in HCI: lost in translation," *Interactions*, vol. 27, no. 5, pp. 68–71, 2020.
- [12] A. A. Fadele, A. Kamsin, K. Ahmad, and R. A. Rasheed, "A novel Hadith authentication mobile system in Arabic to Malay language translation for android and iOS Phones," *International Journal of Information Technology*, vol. 13, no. 4, pp. 1683–1692, 2021.
- [13] P. Tracey, E. Dalpiaz, and N. Phillips, "Fish out of water: translation, legitimation, and new venture creation," *Academy of Management Journal*, vol. 61, no. 5, pp. 1627–1666, 2018.
- [14] R. N. Parsa, "Trends in e-tools and resources for translators and interpreters," *Translation & Interpreting*, vol. 13, no. 2, pp. 183–186, 2021.
- [15] C. Salter, "Disturbance, translation, enculturation: necessary research in new media, technology, and the senses," *Visual Anthropology Review*, vol. 34, no. 1, pp. 87–97, 2018.
- [16] L. Lin, P. Hu, X. Tong et al., "High-speed three-dimensional photoacoustic computed tomography for preclinical research and clinical translation," *Nature Communications*, vol. 12, no. 1, pp. 882–903, 2021.
- [17] L. Bowker, "Machine translation literacy: academic libraries' role," *Proceedings of the Association for Information Science and Technology*, vol. 56, no. 1, pp. 618–619, 2019.