WHAT IS CLAIMED IS:

1. A speech recognition method, comprising:

obtaining a to-be-recognized speech signal, obtaining a decoding result by decoding the to-be-recognized speech signal, and extracting a first feature from the decoding result, wherein the first feature contains prepositive word information of the to-be-recognized speech signal;

extracting a target speech segment from the to-be-recognized speech signal, obtaining a logarithmic amplitude spectrum of the target speech segment, and determining a second feature according to the logarithmic amplitude spectrum, wherein the second feature is a two-dimensional time-frequency domain feature of the to-be-recognized speech signal;

obtaining a third feature by combining the first feature with the second feature; and

classifying the third feature using a trained classifier, to determine whether the third feature contains a prepositive word.

2. The method according to claim 1, wherein obtaining a decoding result by decoding the to-be-recognized speech signal comprises:

decoding the to-be-recognized speech signal using a trained acoustic model and a trained language model, and extracting top three candidate decoding results among candidate decoding results as the decoding result.

3. The method according to claim 2, wherein extracting a first feature from the decoding result comprises:

obtaining an acoustic-model score and a language-model score of the decoding result, and obtaining a normalized acoustic-model score and a normalized language-model score as the first feature by performing normalization on the acoustic-model score and the language-model score.

4. The method according to claim 3, wherein obtaining a normalized acoustic-model score and a normalized language-model score by performing normalization on the acoustic-model

score and the language-model score comprises:

obtaining the normalized acoustic-model score by dividing the acoustic-model score by an acoustic-model score of an optimal decoding result, wherein the optimal decoding result is a candidate decoding result ranked first among the candidate decoding results; and

obtaining the normalized language-model score by dividing the language-model score by a language-model score of the optimal decoding result.

5. The method according to claim 1, wherein

determining the prepositive word information as 1 in response to containing of the prepositive word in the decoding result; and

determining the prepositive word information as 0 in response to not containing of the prepositive word in the decoding result.

6. The method according to claim 5, wherein

the prepositive word is at least one of “no” and “not”.

7. The method according to claim 1, further comprising:

determining a sensitive word set containing the prepositive word, and extracting the first feature from the decoding result in response to containing of any element in the sensitive word set in the decoding result.

8. The method according to claim 7, wherein extracting a target speech segment from the to-be-recognized speech signal comprises:

determining a time starting point and a time ending point corresponding to the element according to time information of the decoding result; and

extracting a speech segment between the time starting point and the time ending point from the to-be-recognized speech signal as the target speech segment.

9. The method according to claim 1, wherein obtaining a logarithmic amplitude spectrum of the target speech segment comprises:

dividing the target speech segment into a preset number of sub segments;

obtaining a speech spectrogram by performing a short-time Fourier transform on each of the sub segments regarding a preset number of points; and

obtaining the logarithmic amplitude spectrum according to the speech spectrogram.

10. The method according to claim 9, wherein determining a second feature according to the logarithmic amplitude spectrum comprises:

obtaining a normalized logarithmic amplitude spectrum by normalizing the logarithmic amplitude spectrum to a range of 0 to 1; and

extracting the second feature from the normalized logarithmic amplitude spectrum.

11. The method according to claim 10, wherein extracting the second feature from the normalized logarithmic amplitude spectrum comprises:

dividing the normalized logarithmic amplitude spectrum into multiple sub bands;

obtaining values of time-smoothed sub band energy by smoothing sub band energy of the sub bands in a time direction;

calculating time jump ratios of the sub band energy according to the values of time-smoothed sub band energy;

obtaining a full-band time jump ratio corresponding to a time instant by averaging the time jump ratios of the sub band energy of the multiple sub bands corresponding to the time instant;

obtaining a maximum of the time jump ratio, an average of the time jump ratio, and a standard deviation of the time jump ratio from multiple full-band time jump ratios corresponding to multiple time instants as the second feature;

obtaining values of frequency-smoothed sub band energy by smoothing sub band energy of the sub bands in a frequency direction;

calculating frequency jump ratios of the sub band energy according to the values of

frequency-smoothed sub band energy;

obtaining a full-band frequency jump ratio corresponding to a time instant by averaging the frequency jump ratios of the sub band energy of the multiple sub bands corresponding to the time instant; and

obtaining a minimum of the frequency jump ratio, an average of the frequency jump ratio, and a standard deviation of the frequency jump ratio from multiple full-band frequency jump ratios corresponding to multiple time instants as the second feature.

12. The method according to claim 11,

wherein obtaining values of time-smoothed sub band energy by smoothing sub band energy of the sub bands in a time direction comprises:

obtaining the values of time-smoothed sub band energy by averaging the sub band energy at a current time instant and the sub band energy at an adjacent time instant; and

wherein obtaining values of frequency-smoothed sub band energy by smoothing sub band energy of the sub bands in a frequency direction comprises:

obtaining the values of frequency-smoothed sub band energy by averaging the sub band energy at a current frequency and the sub band energy at an adjacent frequency.

13. The method according to claim 11,

wherein calculating time jump ratios of the sub band energy according to the values of time-smoothed sub band energy comprises:

obtaining quotients of a value of time-smoothed sub band energy corresponding to a preset time instant and a value of time-smoothed sub band energy corresponding to a current time instant as the time jump ratios of the sub band energy; and

wherein calculating frequency jump ratios of the sub band energy according to the values of frequency-smoothed sub band energy comprises:

obtaining quotients of a value of frequency-smoothed sub band energy corresponding to

a preset frequency and a value of frequency-smoothed sub band energy corresponding to a current frequency as the frequency jump ratios of the sub band energy.

14. The method according to claim 13, wherein the preset time instant is a 5^{th} time instant after the current time instant, and the preset frequency is a 5^{th} frequency after the current frequency.

15. An electronic device, comprising:

a processor; and

a memory for storing one or more programs, wherein when the one or more programs are executed by the processor, the speech recognition method according to any of claims 1 to 14 is caused to be implemented by the processor.

16. A method for training a classifier, comprising:

obtaining a sample speech signal, obtaining a decoding result by decoding the sample speech signal, and extracting a first feature from the decoding result, wherein the first feature contains prepositive word information of the sample speech signal;

extracting a target speech segment from the sample speech signal, obtaining a logarithmic amplitude spectrum of the target speech segment, and determining a second feature according to the logarithmic amplitude spectrum, wherein the second feature is a two-dimensional time-frequency domain feature of the sample speech signal;

obtaining a third feature by combining the first feature with the second feature; and

training an untrained classifier using the third feature, to obtain a trained classifier.

17. The method according to claim 16, wherein extracting a first feature from the decoding result comprises:

obtaining an acoustic-model score and a language-model score of the decoding result, and obtaining a normalized acoustic-model score and a normalized language-model score as the first feature by performing normalization on the acoustic-model score and the language-model

score.

18. The method according to claim 16, wherein determining a second feature according to the logarithmic amplitude spectrum comprises:

dividing the target speech segment into a preset number of sub segments;

obtaining a speech spectrogram by performing a short-time Fourier transform on each of the sub segments regarding a preset number of points;

obtaining the logarithmic amplitude spectrum according to the speech spectrogram.

obtaining a normalized logarithmic amplitude spectrum by normalizing the logarithmic amplitude spectrum to a range of 0 to 1; and

extracting the second feature from the normalized logarithmic amplitude spectrum.

19. The method according to claim 18, wherein extracting the second feature from the normalized logarithmic amplitude spectrum comprises:

dividing the normalized logarithmic amplitude spectrum into multiple sub bands;

obtaining values of time-smoothed sub band energy by smoothing sub band energy of the sub bands in a time direction;

calculating time jump ratios of the sub band energy according to the values of time-smoothed sub band energy;

obtaining a full-band time jump ratio corresponding to a time instant by averaging the time jump ratios of the sub band energy of the multiple sub bands corresponding to the time instant;

obtaining a maximum of the time jump ratio, an average of the time jump ratio, and a standard deviation of the time jump ratio from multiple full-band time jump ratios corresponding to multiple time instants as the second feature;

obtaining values of frequency-smoothed sub band energy by smoothing sub band energy of the sub bands in a frequency direction;

calculating frequency jump ratios of the sub band energy according to the values of frequency-smoothed sub band energy;

obtaining a full-band frequency jump ratio corresponding to a time instant by averaging the frequency jump ratios of the sub band energy of the multiple sub bands corresponding to the time instant; and

obtaining a minimum of the frequency jump ratio, an average of the frequency jump ratio, and a standard deviation of the frequency jump ratio from multiple full-band frequency jump ratios corresponding to multiple time instants as the second feature.

20. The method according to claim 16, wherein training an untrained classifier using the third feature, to obtain a trained classifier comprises:

training an untrained naive Bayes classifier using the third feature, to obtain a trained naive Bayes classifier.