Claims

1. An audio recognition method, comprising:

performing diffusion processing on a plurality of first feature points in a spectrogram of a to-be-recognized audio file to obtain a feature point map;

searching in a spectrogram of a target audio file to determine whether second feature points that respectively correspond to the diffused first feature points in the feature point map exist;

if yes, determining that the to-be-recognized audio file is a part of the target audio file.

2. The method according to claim 1, wherein, before the step of performing diffusion processing on first feature points in a spectrogram of a to-be-recognized audio file, the method further comprises:

normalizing energy values of the first feature points in the spectrogram of the to-be-recognized audio file within a range of grayscale values.

3. The method according to claim 1 or 2, wherein the diffusing processing comprises at least one of Gauss function diffusion processing or amplification processing.

4. The method according to claim 2, wherein the step of normalizing energy values of the first feature points in the spectrogram of the to-be-recognized audio file to grayscale values of the first feature points comprises:

traversing the spectrogram frame by frame with a first preset length as a window;

acquiring a local maximum value and a local minimum value among the energy values of the first feature points within the window;

normalizing, according to the local maximum value and the local minimum value, the energy values of the first feature points to grayscale values of the first feature points.

5. The method according to claim 1 or 2, wherein the step of searching in a spectrogram of a target audio file to determine whether second feature points that respectively correspond to the diffused first feature points in the feature point map exist comprises:

traversing the spectrogram of the target audio file frame by frame with the feature point map as a window;

determining, in the each traversing process, feature points in the spectrogram of the target audio file inside the window that have coordinates within a range of coordinates of the

first feature points in the window after the diffusion processing as the second feature points;

searching in the spectrogram of the target audio file inside the window to determine whether the second feature points that respectively correspond to the diffused first feature points exist.

6. The method according to claim 5, wherein the step of determining feature points in the spectrogram of the target audio file inside the window that have coordinates within a range of coordinates of the first feature points in the window after the diffusion processing as the second feature points comprises:

determining a degree of matching between the first feature points and the feature points in the spectrogram of the target audio file inside the window that have coordinates within the range of coordinates of the first feature points in the window after the diffusion processing;

determining feature points with the degree of matching greater than a first threshold value as the second feature points.

7. The method according to claim 6, wherein the degree of matching comprises a ratio of a number of feature points in the spectrogram inside the window that are within a range of coordinates of the diffused first feature points to a number of the first feature points, or a sum of energy values or grayscale values of the first feature points corresponding to the feature points in the spectrogram inside the window that are within a range of coordinates of the diffused first feature points.

8. The method according to claim 1 or 2, wherein, before the step of diffusing first feature points in a spectrogram of a to-be-recognized audio file, the method further comprises:

using feature points comprised in the spectrogram of a to-be-recognized audio file that have energy values or grayscale values greater than a second threshold value as key points;

if the energy value or grayscale value of the key point is a maximum value in a preset region, determining the key point as the first feature point.

9. The method according to claim 1, wherein the target audio file carries audio information, and the audio information includes song names.

10. An audio recognition system, comprising:

a diffusing unit configured to perform diffusion processing on a plurality of feature points in a spectrogram of a to-be-recognized audio file to obtain a feature point map;

a searching unit configured to search in a spectrogram of a target audio file to determine whether second feature points that respectively correspond to the diffused first feature points in the feature point map exist;

a determining unit configured to determine that the to-be-recognized audio file is a part of the target audio file when the second feature points that respectively correspond to the diffused first feature points in the feature point map are found in the spectrogram of the target audio file.

11. The system according to claim 10, wherein, before the diffusing unit, the system further comprises:

a normalizing unit configured to normalize the energy values of first feature points in the spectrogram of the to-be-recognized audio file within a range of grayscale values.

12. The system according to claim 10 or 11, wherein the diffusing processing comprises at least one of Gauss function diffusion processing or amplification processing.

13. The system according to claim 11, wherein the normalizing unit comprises:

a first normalizing sub-unit configured to traverse a spectrogram frame by frame with a first preset length as a window;

a second normalizing sub-unit configured to acquire a local maximum value and a local minimum value among the energy values of the first feature points within the window;

a third normalizing sub-unit configured to normalize, according to the local maximum value and the local minimum value, the energy values of the first feature points to grayscale values of the first feature points.

14. The system according to claim 10 or 11, wherein the searching unit comprises:

a first searching sub-unit configured to traverse the spectrogram of the target audio file frame by frame with the feature point map as a window;

a second searching sub-unit configured to determine, in the each traversing process, feature points in the spectrogram of the target audio file inside the window that have coordinates within a range of coordinates of the first feature points in the window after the diffusion processing as the second feature points;

a third searching sub-unit configured to search in the spectrogram of the target audio file inside the window to determine whether the second feature points that respectively correspond to the diffused first feature points exist.

15. The system according to claim 14, wherein the second searching sub-unit comprises:

a fourth searching sub-unit configured to determine a degree of matching between the first feature points and the feature points in the spectrogram of the target audio file inside the window that have coordinates within the range of coordinates of the first feature points in the window after the diffusion processing;

a fifth searching sub-unit configured to determine feature points with the degree of matching greater than a first threshold value as the second feature points.

16. The system according to claim 15, wherein the degree of matching comprises a ratio of a number of feature points in the spectrogram inside the window that are within a range of coordinates of the diffused first feature points to a number of the first feature points, or a sum of energy values or grayscale values of the first feature points corresponding to the feature points in the spectrogram inside the window that are within a range of coordinates of the diffused first feature points.

17. The system according to claim 10 or 11, wherein, before the diffusion processing, the system further comprises:

a first processing unit configured to use feature points comprised in the spectrogram of a to-be-recognized audio file that have energy values or grayscale values greater than a second threshold value as key points;

a second processing unit configured to determine, if the energy value or grayscale value of the key point is a maximum value in a preset region, the key point as the first feature point.

18. The system according to claim 10, wherein the target audio file carries audio information, and the audio information includes song names.