CLAIMS

What is claimed is:

1. A method for detecting a melody of an audio signal, comprising:

performing Short-Time Fourier Transform (STFT) on the audio signal, wherein the audio signal is a humming or cappella audio signal;

acquiring a pitch frequency by pitch frequency detection on a result of the STFT, wherein the pitch frequency is configured to detect a pitch value;

inputting an interpolation frequency at a signal position corresponding to each frame of audio sub-signal in response to detecting no pitch frequency;

determining the interpolation frequency corresponding to the frame as the pitch frequency of the audio signal;

dividing the audio signal into a plurality of audio segments based on a beat;

equally dividing each of the audio segments into several audio sub-segments, separately detecting the pitch frequency of each frame of audio sub-signal in each of the audio sub-segments, and determining a mean value of the pitch frequencies of a plurality of continuously stable frames of audio sub-signals in the audio sub-segment as a pitch value of each of the audio segments;

determining a pitch name corresponding to each of the audio segments based on a frequency range of the pitch value;

acquiring a musical scale of the audio signal by estimating a tonality of the audio signal based on the pitch name of each of the audio segments; and

determining a melody of the audio signal based on a frequency interval of the pitch value of each of the audio segments in the musical scale.

2. The method for detecting the melody of the audio signal according to claim 1, wherein the step of dividing the audio signal into the plurality of audio segments based on the beat comprises:

determining a duration of each of the audio segments based on a specified beat type; and

dividing the audio signal into several audio segments based on the duration, wherein the audio segments are bars determined based on the beat.

3. The method for detecting the melody of the audio signal according to claim 1 or 2, wherein upon the step of determining the mean value of the pitch frequencies of the plurality of

continuously stable frames of the audio sub-signals in the audio sub-segment as the pitch value, the method further comprises:

calculating a stable duration of the pitch value in each of the audio sub-segments; and

setting the pitch value of the audio sub-segment to zero in response to the stable duration being less than a specified threshold.

4. The method for detecting the melody of the audio signal according to claim 1, wherein the step of determining the pitch name corresponding to each of the audio segments based on the frequency range of the pitch value comprises:

acquiring a pitch name number by inputting the pitch value into a pitch name number generation model; and

searching, based on the pitch name number, a pitch name sequence table for the frequency range of the pitch value of each of the audio segment, and determining the pitch name corresponding to the pitch value.

5. The method for detecting the melody of the audio signal according to claim 4, wherein in the step of acquiring the pitch name number by inputting the pitch value into the pitch name number generation model, the pitch name number generation model is expressed as:

$$\mathrm{K} = \left(12 \times log_2\left(\frac{f_{m-n}}{a}\right)\right) mod\ 12 + 1,$$

wherein K represents the pitch name number, f_{m-n} represents a frequency of the pitch value of an nth note in an mth audio segment of the audio segments, a represents a frequency of a pitch name for positioning, and *mod* represents a mod function.

6. The method for detecting the melody of the audio signal according to claim 1, wherein the step of acquiring the musical scale of the audio signal by estimating the tonality of the audio signal based on the pitch name of each of the audio segments comprises:

acquiring the pitch name corresponding to each of the audio segments in the audio signal;

estimating the tonality of the audio signal by processing the pitch name using a toning algorithm; and

determining a number of semitone intervals of a positioning note based on the tonality, and acquiring the musical scale corresponding to the audio signal by calculation based on the number of semitone intervals.

7. The method for detecting the melody of the audio signal according to claim 1, wherein the step of determining the melody of the audio signal based on the frequency interval of the pitch value of each of the audio segments in the musical scale comprises:

acquiring a pitch list of the musical scale of the audio signal, wherein the pitch list records a correspondence between the pitch value and the musical scale;

searching the pitch list for a note corresponding to the pitch value based on the pitch value of each of the audio segments in the audio signal; and

arranging the notes in time sequences based on the time sequences corresponding to the pitch values in the audio segments, and converting the notes into the melody corresponding to the audio signal based on the arrangement.

8. The method for detecting the melody of the audio signal according to claim 1, wherein prior to the step of dividing the audio signal into the plurality of audio segments based on the beat, the method further comprises:

generating a music rhythm of the audio signal based on specified rhythm information; and

generating reminding information of beat and time based on the music rhythm.

9. An apparatus for detecting a melody of an audio signal, comprising:

a pitch detection unit, configured to: perform Short-Time Fourier Transform (STFT) on the audio signal, wherein the audio signal is a humming or cappella audio signal; acquire a pitch frequency by pitch frequency detection on a result of the STFT, wherein the pitch frequency is configured to detect a pitch value; input an interpolation frequency at a signal position corresponding to each frame of audio sub-signal in response to detecting no pitch frequency; determine the interpolation frequency corresponding to the frame as the pitch frequency of the audio signal; divide an audio signal into a plurality of audio segments based on a beat, equally divide each of the audio segments into several audio sub-segments, separately detect the pitch frequency of each frame of audio sub-signal in each of the audio sub-segments, and determine a mean value of the pitch frequencies of a plurality of continuously stable frames of audio sub-signals in the audio sub-segment as a pitch value of each of the audio segments;

a pitch name detection unit, configured to determine a pitch name corresponding to each of the audio segments based on a frequency range of the pitch value;

a tonality detection unit, configured to acquire a musical scale of the audio signal by estimating a tonality of the audio signal based on the pitch name of each of the audio segments; and

a melody detection unit, configured to determine a melody of the audio signal based on a frequency interval of the pitch value of each of the audio segments in the musical scale.

10. An electronic device, comprising:

a processor; and

a memory configured to store one or more instructions executable by the processor,

wherein the processor is configured to perform the method for detecting the melody of the audio signal as defined in any one of claims 1 to 8.

11. A non-transitory computer-readable storage medium storing one or more instructions wherein the one or more instructions, when executed by a processor of an electronic device, cause the electronic device to perform the method for detecting the melody of the audio signal as defined in any one of claims 1 to 8.